ABSTRACT

Title of Document: MULTIPLE INTERROGATIVES: SYNTAX, SEMANTICS, AND LEARNABILITY

Lydia Grebenyova, Ph.D., 2006

Directed By: Distinguished University Professor, Howard Lasnik, Department of Linguistics

The dissertation consists of theoretical and experimental studies of multiple interrogatives (i.e., sentences containing more than one wh-phrase, like *Who bought what?*). First, I examine the status of Superiority effects in contexts with and without subject-aux(iliary) inversion cross-linguistically. The relevant contrast from English is between *Who bought what?*, $^{??}$*What did who buy?*, and ***I wonder what who bought*.**, where (*) indicates a greater degree of unacceptability by native speakers than ($^{??}$). I argue that the presence of subject-aux inversion in main clauses in English is responsible for the given asymmetry, and I attribute the degraded status of $^{??}$*What did who buy?* to the independent semantic properties of questions.

Next, I explore the semantic properties of multiple interrogatives in detail. I develop an analysis that does not rely on covert wh-movement, relying instead on the syntactic position of the Question morpheme. I also explore the nature of complex

wh-phrases (e.g., *what boy, which book*). I propose that choice functions are part of complex wh-phrases but not bare wh-phrases.

I then explore the behavior of multiple interrogatives under Sluicing (i.e., clausal ellipsis). I observe that, in Slavic, it is possible to have multiple wh-phrases as well as focused referential expressions as remnants of sluicing. Based on this data, I argue that clausal ellipsis is licensed by focus in general. I also explore the apparent Superiority effects under sluicing in Russian and Polish and conclude that those are, in fact, parallelism effects, and not minimality effects.

Finally, I present the results of several language acquisition studies on at what age and how English-, Russian-, and Malayalam- speaking children acquire the language-specific syntactic and semantic properties of multiple interrogatives, given the limited evidence in the input. I report the results of the corpus studies of parental speech with respect to the frequency of occurrence of multiple interrogatives, as well as the results of the studies, where multiple interrogatives were elicited from children and adults in specific contexts. I conclude that young children acquire syntax and semantics of multiple interrogatives quite successfully. I then discuss what evidence in the input they might be using.

MULTIPLE INTERROGATIVES: SYNTAX, SEMANTICS, AND LEARNABILITY

By

Lydia Grebenyova

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2006

Advisory Committee:
Distinguished University Professor Howard Lasnik, Chair
Professor Željko Bošković
Professor Norbert Hornstein
Associate Professor Jeffrey Lidz
Professor Richard Brecht

Acknowledgements

A lot of people made this dissertation and my overall transition through the graduate school possible and immensely enjoyable. First of all, I would like to thank Daniel Seely, my advisor in the M.A. program in linguistics at Eastern Michigan University, for getting me interested in generative linguistics back in 1999 and for his support and encouragement ever since.

I am also deeply indebted to Howard Lasnik, my Ph.D. advisor at the University of Maryland, for being such a great teacher and mentor to me, for giving me confidence in my ideas, and for always pointing out issues I had never thought of before. I also thank my dissertation committee members Željko Bošković, Norbert Hornstein, Jeff Lidz, and Richard Brecht for their invaluable comments on the drafts of the chapters, and especially to Željko Bošković, who did it all long-distance from the University of Connecticut.

I also would like to thank all the members of the Department of Linguistics at the University of Maryland for making the place so much fun to be around. My special regards go to my officemates, Tomo Fujii and Scott Fults, and to my long time friend Heather Taylor.

Many thanks go to all my many informants and participants of the acquisition experiments. Without their help, this work would not be possible.

I am also grateful to Tatiana Grebenyova, Rosmin Mathew, Utako Minai, and Lisa Pearl, for assisting me in running the acquisition experiments. Many thanks also go to the Kerala School in New Delhi, India, and the Day Care Center #6 in Voronezh, Russia, for their hospitality during the experiments.

And most of all, I would like to express my deepest thanks to my family and friends in Russia, and especially to my mother, who has always been the source of inspiration and encouragement for me. Ogromnoje spasibo!

Table of Contents

Chapter 1: Introduction

The goal of this work is to address certain issues in syntax, semantics, and acquisition of multiple interrogatives. A multiple interrogative is an interrogative clause with more than one wh-phrase in it, as in an English example in (1). I will use the terms *multiple interrogative* and *multiple question* interchangeably to refer to such expressions.

(1) Who bought what?

The thesis is organized as follows. In Chapter 2, I examine how T-to-C movement and semantics of multiple questions affect Superiority effects cross-linguistically. The novel generalization is that an asymmetry in Superiority effects between matrix and embedded multiple questions arises in languages that show an asymmetry in the availability of T-to-C movement in main vs. embedded clauses.

Adopting the minimality account of Superiority of Chomsky (1995), I argue that the presence of T-to-C movement relaxes Superiority effects in certain contexts. I present evidence for this hypothesis from a number of languages such as English, Icelandic, Brazilian Portuguese, and Bulgarian. I also argue that the semantics of multiple questions, particularly the availability of Single-pair readings in bare multiple questions in a given language, crucially affects Superiority effects.

I conclude that what we know as Superiority effects reveal a complex interplay between syntactic and semantic factors such as minimality, T-to-C movement and the interpretation of multiple questions. The analysis has consequences for clausal structure, the locality of Attract, and the status of T-to-C movement in embedded clauses in English and in the grammar in general (i.e., whether T-to-C movement takes place in overt syntax or whether it is a PF phenomenon).

In Chapter 3, I explore the semantics of multiple interrogatives, focusing on the Pair-list (PL) and Single Pair (SP) readings in these structures. First, I examine the distribution of these readings in a variety of contexts, such as Superiority-obeying vs. Superiority-violating contexts, questions with bare vs. complex wh-phrases, local vs. long-distance multiple questions, and questions where the wh-phrases are separated by an island boundary.

Since choosing the right tools for the semantics of multiple interrogatives crucially depends on whether covert wh-movement exists, I explore the potential semantic and syntactic evidence for it and conclude that, if covert wh-movement exists, it must be motivated by a purely formal requirement. That is, there seems no semantic evidence for covert wh-movement. This conclusion leads me to using choice functions in developing an account of the PL/SP readings distribution.

My account is based on the compositional semantics of the PL/SP readings of Hagstrom (1998). With that as a starting point, I explore what is responsible for the cross-linguistic parameterization with respect to the availability of the SP readings in multiple questions with bare wh-phrases. I relate this parameterization to the

selectional restrictions of the interrogative morpheme (Q-morpheme), which, as I argue, vary across languages. I then account for the distribution of the PL/SP reading in the contexts mentioned above.

In Chapter 4, I explore how multiple interrogatives behave under clausal ellipsis (Sluicing). I focus on the phenomenon known as multiple sluicing: sluicing with multiple remnants. Multiple sluicing is very productive in Slavic; therefore most of this chapter deals with the data from Slavic languages, particularly Bulgarian, Russian, Polish, and Serbo-Croatian. Contrary to the previous proposals that an interrogative +*wh* complementizer licenses TP-ellipsis, as in Lobeck (1995) and Merchant (2001), I argue that it is in fact the +*focus* feature that is responsible for licensing this ellipsis operation in all languages. I assume the relevant operation to be deletion, following Ross (1969), Lasnik (1999) and Merchant (2001). The evidence I present for the focus licensing of sluicing comes from the data from Slavic languages like Russian and Polish, where it is possible to have not only wh-phrases but also focused R-expressions as remnants of sluicing. I also demonstrate how the unavailability of SP readings in multiple interrogatives in a given language is found even under sluicing. This presents a new argument for the full clausal structure of the sluice, as opposed to the structure of the sluice consisting of just the remnant material. Finally, I explore Superiority effects under sluicing in languages that do not show Superiority effects in non-elliptical structures. I derive those effects from an independent property of ellipsis, namely, scope parallelism.

In Chapter 5, I report the results of the acquisition studies on how English-, Russian-, and Malayalam- speaking children produce and interpret multiple

interrogatives. Because the majority of the studies on acquisition of questions focus only on single interrogatives, my goal here is to make the initial steps in approaching the learnability issues in multiple interrogatives.

First, I explore how much evidence for the syntactic and semantic properties of multiple interrogatives children get in the linguistic input. To do that, I conducted a corpus analysis of parental speech in the CHILDES database. The results show a great asymmetry between the frequencies of occurrence of single vs. multiple interrogatives in the parental speech. Multiple interrogatives occur much more rarely than single interrogatives. This suggests that children acquire the language-specific facts about multiple interrogatives at a later age than they do single interrogatives. Therefore, I investigate next at what age children exhibit the knowledge of the syntax and semantics of multiple interrogatives by eliciting those structures from children and adults in specific contexts.

The overall conclusion is that, with the apparently limited direct evidence in the input, children are still able to acquire the language-specific facts about multiple interrogatives at quite an early age.

The specific results show that Russian-, English-, and Malayalam-speaking children by 4;9 exhibit adult-like semantic knowledge of the restriction on the SP readings in these languages. I develop a learning algorithm for acquisition of these semantic properties based on children deducing the relevant properties of multiple interrogatives from an independent property of language, namely, the presence of an independent Focus projection above TP in a given language. The evidence children use in acquiring this property consists of the easily observable focus and

complementizer morphology in some languages and the distribution of the focus-fronted expressions in other.

As for the acquisition of syntax of multiple interrogatives, I report that unlike English- and Malayalam-speaking children, Russian-speaking children produce some non-adult-like structures, where only one wh-phrase is fronted in a multiple question, when all wh-phrases are fronted in these contexts in adult Russian. I relate this behavior to acquisition multiple wh-fronting via learning the crucial properties of contrastive focus, which has been argued in the literature to be the underlying trigger of multiple wh-fronting. I also discuss how Russian-speaking acquire the asymmetry between bare and complex wh-phrases in this language and identify this as an additional factor contributing to the acquisition of multiple wh-fronting.

Chapter 2: Superiority – Syntactic and Interpretive

1. Introduction

In this chapter, I develop an analysis of Superiority that is based on both the syntactic and semantic properties of multiple interrogatives. I begin with the observation that Superiority violations in embedded clauses are more severe than those in main clauses in English. Given our current understanding of Superiority, this contrast is not expected. For instance, an Economy approach to Superiority, based on Minimal Link Condition of Chomsky (1995), rules out such violations in both main and embedded clauses in exactly the same way. Therefore, the matrix-embedded asymmetry in Superiority presents a challenging puzzle for the current theorizing. Resolving this puzzle will be one of the main goals of this chapter.

I develop an account of Superiority which largely maintains the spirit of Attract Closest of Chomsky (1995), where the strong uninterpretable +*wh* feature of C^0 attracts the *closest* wh-phrase to SpecCP for feature checking.[1] The novel part of the analysis is that, in accounting for Superiority, it considers the effect of other syntactic and semantic processes occurring in the derivation of multiple interrogatives which turn out to have an effect on Superiority. Particularly, I investigate the effect of head-movement on the locality of Attract (i.e., how T-to-C movement affects the

[1] Whether feature-checking takes place in a Spec-head configuration or under an operation Agree will become relevant later on in Section 3 of this chapter. For consistency, I will frame the discussion in terms of the Spec-head relation.

locality of C^0 attracting a wh-phrase), and how the distribution of pair-list and single-pair readings in multiple interrogatives contributes to Superiority effects. The resulting account makes correct predictions about Superiority effects in main and embedded clauses cross-linguistically, in local and long-distance questions, and in questions with bare and complex wh-phrases.

The chapter is organized as follows. In Section 2, I demonstrate the central contrast between Superiority effects in main and embedded questions. In Section 3, I examine how the presence of T-to-C movement in main clauses and its absence in embedded clauses in English affects locality conditions associated with Attract F. I conclude that T-to-C movement is a crucial factor contributing to the Superiority asymmetries in question. However, T-to-C movement alone is not sufficient to account for all of the data. In Section 4, I explore independent semantic properties of multiple wh-questions and their contribution to the rise of the observed Superiority contrasts. Section 5 examines the predictions of the analysis for Superiority effects in languages with and without T-to-C movement and with the varying interpretive possibilities in multiple interrogatives. The specific languages I examine include Icelandic, Bulgarian, Serbo-Croatian, and Brazilian Portuguese. Finally, Section 6 discusses the implications of this overall account for the theory of movement, considering Attract vs. Move based theories, as well as examining several accounts of successive cyclicity. Section 7 is the summary of conclusions from this chapter.

2. Matrix-embedded Asymmetry in Superiority Effects

The phenomenon of Superiority has been explored since Chomsky (1973). The empirical generalization is that in a question involving more than one wh-phrase, it is

the *superior* wh-phrase (i.e., the one that asymmetrically c-commands other wh-phrases) that is moved to the clause-initial position. For example, consider the acceptability of (2a) and the degraded status of (2b). In (2b), the lower wh-phrase is fronted over the superior one, unlike in (2a).

(2) a. Who$_1$ did John persuade t_1 to buy what?
    b. *What$_1$ did John persuade who buy t_1?

Chomsky (1973) postulates the Superiority Condition, given in (3).

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

The Superiority condition correctly rules out (2b), where wh-movement applies to *what* even though *who* is superior to *what*.

In the Minimalist approach to Superiority in Chomsky (1995), the main generalization is captured through the economy condition *Attract Closest* as part of the definition of *Attract* (the operation responsible for feature checking). The basic idea is that it is most economical for a head K with an uninterpretable feature to attract the *closest* element with a matching feature. Chomsky (1995:311) formulates Attract Closest as in (4).[2]

[2] Chomsky (1995:311) actually refers to the condition in (4) as Minimal Link Condition (MLC). However, MLC was originally an output condition on chains, as the Minimize Chain Link Principle in Chomsky and Lasnik (1993:90). The condition in (4), however, has nothing to do with chains. Therefore, I refer to it as *Attract Closest*, which is often done in the literature.

(4) K attracts α only if there is no β, β closer to K than α, such that K attracts *β*.

β is closer to K than α if β asymmetrically c-commands α. And K c-commands α and β. With respect to Superiority, the interrogative complementizer with an uninterpretable [+wh] feature must attract the closest wh-phrase for feature checking. Attract Closest rules out (2b) by virtue of the fact that the object wh-phrase *what* is not the closest wh-element to C^0 and therefore cannot be attracted by C^0.

However, there are data that Attract Closest alone cannot capture. Consider the following asymmetry found in main and embedded clauses in English. Superiority violations in embedded questions, as in (5d), are judged by English native-speakers as stronger than Superiority violations in matrix questions, as shown in (5b).

(5) a. Who bought what?
b. ??What$_1$ did who buy t_1?
c. *John wonders what$_1$ who bought t_1.

A similar contrast can be found in Serbo-Croatian, a Slavic language with multiple wh-fronting. Bošković (1997a, 1998, 2002a) reports that, while Serbo-Croatian main clauses with null C^0 do not exhibit Superiority effects, those effects emerge in embedded clauses. While both (6a) and (6b) are acceptable, demonstrating the absence of Superiority effects in the main clauses, the embedded question in (7b) with the object wh-phrase fronted over the subject wh-phrase is degraded, unlike (7a).

(6) a. ***Ko*** ***šta*** o njemu govori? *Serbo-Croatian*
who what about him says
'Who says what about him?'

b. ***Šta ko*** o njemu govori t_1?
what who about him says

(7) a. Pavle je pitao ***ko šta***$_1$ o njemu govori.
Pavle Aux asked who what about him says
'Pavle asked who says what about him'

b. ??Pavle je pitao ***šta***$_1$ ***ko*** o njemu govori t_1.
Pavle Aux asked what who about him says

Note that, when it comes to Superiority in multiple wh-fronting languages, it is not sufficient for C^0 to attract the closest wh-phrase to its Spec. Under any analysis, it is important to ensure that the wh-phrase that is first in linear order is the one that moves first (and stays first). This can be implemented by assuming that, when the next wh-phrase is attracted, it either tucks-in underneath the first one, as in Richards (1997a), or right-adjoins to the first wh-phrase, as in Rudin (1998) and Bošković (1998, 2002a).[3]

The generalization we can draw so far is that the degree of Superiority effects increases in embedded clauses in English and Serbo-Croatian. Crucially, Attract

[3] Determining which of these options is the right one seems to depend on whether multiple specifiers, which are part of the tucking-in approach, should be allowed in our theory. Prohibiting multiple specifiers would require an extra stipulation and therefore is undesirable, which makes tucking-in technically plausible. Preferring the lower Spec over the higher one, however, remains stipulative. Richards (1997) argues that it follows from Shortest Move. However, Shortest Move itself does not follow from anything, especially with Attract Closest in the system. Tucking-in may seem to violate the Extension Condition of Chomsky (1993), but not if the Extension Condition is derived from feature strength, as in Bošković and Lasnik (1999). Note, however, that deriving the Extension Condition from feature strength is only possible if the strong feature is on an attractor and not on the moving item, which is currently under debate. See Section 6 of this chapter for more on the nature of this debate.

Closest alone cannot distinguish between Superiority violations in matrix and embedded clauses.

There is an alternative explanation for the Serbo-Croatian facts. Bošković (1998, 2002a) proposes that the absence of Superiority effects in matrix clauses in Serbo-Croatian can be explained if the matrix phonetically null C^0 with the strong *+wh* feature can be inserted in covert syntax. In this case, although there is overt multiple wh-fronting in Serbo-Croatian, it is claimed to have nothing to do with a *+wh* feature. Instead, as Bošković argues, overt wh-fronting in Slavic is driven by contrastive focus. That is, the wh-phrases move overtly to a focus projection, which is lower than CP in Serbo-Croatian.[4]

On this analysis, wh-movement to SpecCP in clauses with covert C^0 merger, takes place covertly, preventing the possibility of the Superiority effects being observed in overt syntax. Such covert merger of C^0 is impossible in embedded clauses because it would violate strict cyclicity (i.e., the Extension Condition or any of its substitutes) because Merge would not apply at the root of the tree, hence not extending the tree. That is why the embedded C^0, even if phonetically null, must be merged with TP overtly, producing Superiority effects in embedded clauses in Serbo-Croatian.

[4] The question arizes as to why focus-movement does not exhibit Superiority effects. Bošković suggests that the uninterpretable *+focus* feature on the functional projection bearing contrastive focus is hypothesized to be an Attract-all feature, such that it is not checked until all the elements with interpretable *+focus* features are attracted for feature-checking. In the end result, it does not matter in which order the wh-phrases are attracted if all of them are attracted to check the very same feature. This leaves it as puzzle though, why the *+wh* feature in principle could not be an Attract-all feature in Slavic.

However, this analysis does not extend to English since, unlike Serbo-Croatian, English exhibits Superiority effects in matrix clauses with null C^0. Bošković analyses the English matrix null interrogative C^0 as a PF verbal affix, which can be seen from the obligatory status of inversion in English main clauses. That is, the phonological information in the English C^0, by hypothesis, prevents it from merging into the structure covertly. Hence, an explanation is needed at least for the English matrix-embedded asymmetry in Superiority effects, with a potential extension to Serbo-Croatian as well as other languages.

There is an independent matrix-embedded clause asymmetry which seems relevant here. While subject-auxiliary inversion, standardly analyzed as T-to-C movement, occurs in main clauses, it does not take place in embedded clauses in English, as shown in (8).[5]

(8) a. What can John buy?
b. *What John can buy?
c. John wonders what Mary can buy.
d. *John wonders what can Mary buy.

We can now formulate a tentative generalization about the matrix-embedded asymmetry in Superiority in English in terms of T-to-C movement, as in (9).

(9) Superiority effects are stronger in contexts without T-to-C movement.

[5] See Pesetsky and Torrego (2001) for an analysis challenging the claim that T-to-C movement does not take place in embedded clauses in English.

Why would the absence of T-to-C movement cause a higher degree of unacceptability of Superiority violations? In order to answer this question, in the next section, I will explore how exactly T-to-C movement affects the derivation of a multiple wh-question.

3. T-to-C movement and locality of Attract

3.1. Equidistance via head-movement

It has been previously proposed that head-movement has an effect on the locality of XP-movement. For instance, Chomsky (1993) argues that head-movement licenses extraction of elements from otherwise non-local positions. He formulates the notions of domain and minimal domain of α as in (10), where α is a head or a feature, and CH is the chain (α, *t*) or a trivial chain α.

(10) a. Max(α) is the smallest maximal projection including α.
b. The *domain* δ(CH) of CH is the set of categories included in Max(α) that are distinct from and do not contain α or *t*.
c. The *minimal domain* Min (δ(CH)) of CH is the smallest subset K of δ(CH) such that for any $\gamma \in \delta$(CH), some $\beta \in$ K reflexively dominates γ.

Consider the derivation in (11).

(11) $[_{TP}$ T $[_{AGRoP}$ NP_2 $[Agr_O$ - $V_1]$ $[_{VP}$ NP t_1 $t_2]]]$

On Chomsky's (1993) analysis, the chain [V_1, t_1], with the head V^0 in Arg_O, extends the minimal domain of V^0 to include Arg_OP, making SpecAgr_OP and SpecVP

equidistant from the canonical object position. Thus, V-to-Agr_O movement in (11) allows for the object to move to $SpecAgr_OP$ over the subject in SpecVP without violating minimality.

Bobaljik and Jonas (1996) further explicate this idea in their study of how SpecTP positions are used by subjects in Icelandic. Consider (12), where the movement of Agr_O to T^0 makes SpecTP and $SpecAgr_O$ part of the same minimal domain and therefore equidistant from the subject in SpecVP, allowing for the subject to move over the object in $SpecAgr_O$ without violating minimality.[6]

(12) $[_{TP}\ NP_2\ [T\ [_{AGRo}\ Agr_O\text{-}V]_1]\ [_{AGRoP}\ NP\ \ t_1\ \ [_{VP}\ t_2\ \ldots\ ]]]$

Notice that the analyses of (11) and (12), described above, are developed from the perspective of Move, where movement is viewed as triggered by the moving item and not by the target of movement. This certainly works for A-movement, where the two potential landing sites are "competing" on the basis of how close they are to the moving item. Wh-movement, however, is crucially different. When a wh-phrase moves to SpecCP, SpecCP and the position occupied by another (potentially intervening) wh-phrase do not compete with respect to the moving wh-phrase. It is the positions of the two wh-phrases that are in competition with respect to the target position of movement. This property of wh-movement has been reflected in another view of movement, the one involving an operation Attract, where the trigger of

[6] The Split VP Hypothesis and the overt object shift analysis of Koizumi (1995) (see also Bobaljik 1995 and Lasnik 1995, 1999) avoid the problem of object and subject raising over each other due to the subject originating higher than the target position of movement of the object. This avoids the situation where objects and subjects ever cross each other. However, see McCloskey (2000) and Bošković (1997b) for some arguments against this account.

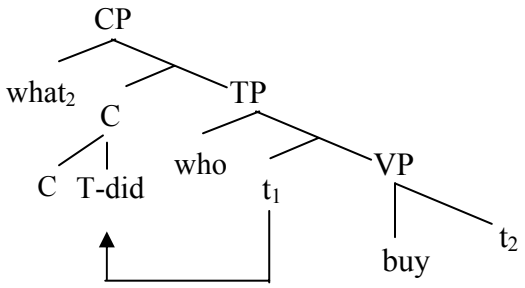
movement is an uninterpretable feature of a functional head. This feature attracts an interpretable matching feature from its c-command domain for feature checking. And by Attract Closest (i.e., by the definition of Attract), it attracts the *closest* matching feature. This puts the positions of the wh-phrases into the locality competition that we observe in Superiority. An idea very similar to Attract closest was first introduced by Oka (1993), who formalizes it in terms of his Shallowness, which is parallel to Closeness of Chomsky (1995).

Can the gist of the head-movement analysis of Chomsky (1993) be captured in a system with Attract? Chomsky (1995:299) proposes a way to do this by suggesting that when X is in the minimal domain of a chain CH with a head Y adjoined to an attracting head Z, X does not have to be preferred for the purposes of Z attracting elements into its minimal domain. The idea is that head-movement to a head Z extends the minimal domain of Z, and Z no longer prefers elements inside its minimal domain and can attract something from outside this domain. (We will later explore the possibility of strengthening this by requiring that Z must not attract elements from inside its minimal domain.) Let us apply the analysis of Chomsky (1995), described above, to T-to-C movement and Superiority. Consider the derivation of an example with a Superiority violation in the main clause, given in (13), and its graph-theoretic representation in (14).[7]

[7] The structure of the verbal domain is somewhat simplified here for the purposes of exposition.

(13) ??[$_{CP}$ What$_2$ [C-[T-did]$_1$] [$_{TP}$ who t_1 buy t_2]]

(14)



Spec,TP and Spec,CP are in the same relation with respect to the chain created by T-to-C movement: they are both within the minimal domain of this chain. Recall that the definition of minimal domain in (10) is formulated with respect to chains, where heads are viewed as trivial chains. To ensure that a functional head does not prefer elements within its minimal domain to elements outside this domain for the purposes of Attract, Chomsky (1995:299) defines *closer to,* which is crucial for Attract Closest, as follows.

(15) β is *closer to* HP (headed by H) than α if β c-commands α and is not in the minimal domain of CH (CH = (γ, t) and γ is adjoined to H).

The formulation of Attract Closest in (4), combined with the notion of closeness as in (15), ensures that a feature within the minimal domain of a chain whose head is adjoined to the attracting head does not count as an intervener for the purposes of Attract. Thus in (13), a *+wh* feature is within the minimal domain of the chain (T, *t*) created by T-to-C movement, and T^0 is adjoined to C^0. By the definition in (15), *who*

in Spec,TP is **not** *closer to* C^0 than *what* is. Therefore, C^0 is free to attract *what* without violating Attract Closest. On this account so far, C^0 can attract either *who* or *what*.

There is one aspect of this analysis that needs further explanation. Notice that SpecTP is not located in the minimal domain of C^0, by definitions of Chomsky (1993) given in (10), but rather it is in the minimal domain of the chain whose head is adjoined to C^0. However, it is C^0 that checks its uninterpretable +*wh* feature. This raises a question as to why C^0 is sensitive to the elements in the minimal domain of the chain (T, *t*) when attracting a certain feature (beyond the stipulation in (15)).

To understand this, let us examine the precise effect of head-adjunction on the nature of the resulting complex head. What does the feature composition of a given head before and after adjunction look like? It is plausible that head-adjunction destroys the autonomy of both heads with respect to their features, producing one complex head. This, in turn, makes it impossible to determine which head exactly attracts the +*wh* feature for feature-checking. It is possible that C^0 and T^0 do it together as one unit (i.e., as one bundle of features). It is only natural then for this single bundle of features to share the minimal domain, which is the union of their former individual minimal domains.[8]

[8] Note that we must be careful not allow head-movement in V2 languages to create one minimal domain as big as a clause. Even if there is a prior step of head-movement to T^0 before T^0 moves to C^0, it does not affect the possibility of extracting the lowest wh-phrase since the calculation of the minimal domain of a chain is not transitive: each new chain link does not extend the minimal domain of the previous link. Such lack of transitivity is assumed in Chomsky (1993) and Bobaljik and Jonas (1996), and can be extended to the analysis in Chomsky (1995). It is less doable in the revised version of Chomsky (1995) that I am entertaining here. I will return to this issue in Subsection 3.3.

3.2. Equidistance or a stronger condition?

On Chomsky's definition of *closer to* in (15), the attracting head (or Probe) still has an option of attracting an element from its minimal domain; it simply does not have to do that since there is no reason for it to prefer such an element to an element with the matching feature outside its minimal domain. By (15), the two potential attractees (or Goals) are made equidistant with respect to the target position.

Now, consider (16).

(16) *Who did leave?

The degraded status of this example has been often approached by trying to prohibit T-to-C movement in the context of subject wh-movement.[9] However, if the account of Chomsky (1995) is strengthened by requiring that an attracting head can never attract elements from its minimal domain, a new analysis of (16) emerges. On this analysis, it is the subject wh-movement that is not permitted when T-to-C movement takes place. T-to-C movement forms a complex head $[C\text{-}T\text{-}did]^0$, as demonstrated in (17), which makes SpecTP part of the minimal domain of this complex head, in the way developed in the previous section. This prevents $[C\text{-}T\text{-}did]^0$ from attracting the subject *who*, correctly ruling out this derivation.

[9] The first account of this phenomenon is in Chomsky (1955, 1957), where adjacency between the tense affix in Aux and the verb is required for the Auxiliary Transformation (Affix Hopping) to apply. When Subject-Aux Inversion applies, it separates the affix from the verb. However, then wh-movement places the wh-subject to a position higher than the target position of the auxiliary. This brings back the needed adjacency of the affix and the verb. Finally, Affix Hopping applies, bleeding Do-support.

(17) *Who$_2$ [C-T$_1$-did]^0 [$_{TP}$ t$_2$ t$_1$ leave]?

The question arises whether the trace (or the lower copy) of T^0 also participates in attracting a *+wh* feature for feature-checking. There is an additional question, answering which seems to provide an answer to this question as well. Is being in the minimal domain of a head sufficient for the feature-checking in general? If it were, a wh-phrase could check the *+wh* feature of the attractor from SpecTP, as in (18), incorrectly letting in the unacceptable sentence.

(18) *[C-T$_1$-did]^0 [$_{TP}$ who t$_1$ leave]?

Thus, we must ensure that *who* cannot check the uninterpretable *+wh* feature of the attractor from SpecTP, even though it is in its minimal domain. One way to do this is by saying that Spec-head configuration is required for feature-checking. This brings us to our original question of whether the trace of T^0 can be considered part of the attractor (or Probe) and whether it can be used for checking the *+wh* feature. For the Spec-head requirement to hold, the answer would have to be negative. That is, the trace (or the lower copy) of T^0 is distinct from the higher copy enough that it cannot be involved in feature-checking. Chomsky (1995) also assumes that traces cannot participate in feature-checking. In addition, under the copy-theory, we can allude to the differences in the feature matrices of these two copies: the higher copy now contains not only the features of T^0 but also the features of C^0, which makes it distinct from the lower copy. Since transferring the features of C^0 to the lower copy of T^0 would require an extra (possibly, ad-hoc) operation, this copy distinction seems a simple and desirable property of grammar. Copy distinction has also been used for the

purposes of linearization, particularly for determining which copy must be deleted at PF, in Nunes (2004). Thus, the obligatoriness of Spec-head configuration, combined with the copy distinction, prohibits the wh-phrase to check the *+wh* feature of [C-T] from SpecTP.

Another way to get this result is with an operation Agree, a long-distance feature-checking operation of Chomsky (1999). If Agree and not Spec-head configuration is required for feature-checking, it is plausible that Agree is subject to the same conditions as Attract. That is, a head cannot establish an Agree relation with an element in its minimal domain. This ensures that the subject wh-phrase cannot check the *+wh* feature of the attractor even without the distinction between the higher and the lower copies of T^0. The attractor then could be as complex as [C-T-t_T].

Thus, our attempt to strengthen Chomsky's condition on Attract seems successful so far. However, there is a potential problem for this new system, containing a stronger condition on Attract. Consider an instance of the object DP moving to Spec of *v*P for Case, as in (19). I am suppressing the fact that the subject *John* moved from Spec*v*P to SpecTP.

(19) [$_{TP}$ John [$_{vP}$ Mary$_2$ [*v*-likes$_1$] [$_{VP}$ t_1 t_2]]]

In (19), the main verb likes moves to *v*, presumably for the assignment of the external theta-role to the subject originating in Spec*v*P. According to our analysis above, the complex head [*v*-likes$_1$]^0 should not be able to attract the object DP since it is in the minimal domain of the chain {likes$_1$, t_1}. The same problem arises if *v* assigns

Accusative Case to the object DP via Agree, assuming that the main verb raises to *v* in this case as well.

The proposed analysis would be compatible with this particular clause structure if the main verb did not actually move to *v*. Let us explore if this is feasible. The motivation for V-to-*v* movement is the assignment of the external theta-role to the subject DP in Spec*v*P. How will the subject DP get its theta-role if such movement does not take place? What if *v* is itself an external theta-role assigner? In this case, no V-to-*v* movement is needed. The head attracting (or Agreeing with) the object DP for Case is then just *v* and not the complex head [*v*-V]^0. The object DP is not in the minimal domain of *v*, hence can be attracted (or Agreed with). Additional arguments for the separation of the external theta-role from the main verb can be found in Kratzer (forthcoming), who argues on semantic grounds that the external theta-role is actually assigned by Voice^0. This is consistent with our analysis, since no V-movement is needed there either.

Another way to approach the problem of the direct object in (19) is by examining whether object shift in English is overt or covert. In the structure in (19), if the main verb does not move any higher than *v* and assuming that the higher copy of the object DP is pronounced, object shift must be covert, otherwise we will get the wrong word order. However, there is evidence that object shift in English is overt, based on the properties of ECM, Pseudogapping, among other phenomena, as argued by Postal (1974), Bošković (1997c, 2004), Lasnik (1999b), and McCloskey (2000).[10] Overt object shift analyses require the verb to move higher than the shifted object, so that the verb still precedes the object in overt syntax. This has been captured by

[10] Lasnik (1999b) also argues that object shift in English is optional.

hypothesizing an Agr_O projection splitting the two verbal projections, as in Koizumi (1995), Bobaljik (1995), and Lasnik (1999a, 1999b). The clause structure on these analyses is as shown in (20). I refer to the higher verbal projection as *v*P for a closer comparison with the alternative structure in (19).

(20) $[_{AgrSP}$ John $[_{TP}$ $[_{vP}$ [*v*-likes$_2$] $[_{AgrOP}$ Mary$_1$ Agr_O $[_{VP}$ t_2 t_1]]]]]

On this analysis, the verb does not move to Agr_O, which attracts the object DP, but rather moves directly to *v*. Therefore, Agr_O can attract the object DP without any problem since the object DP is not part of its minimal domain.

Thus, strengthening Chomsky's condition on Attract seems to work in both A and A'-domains. However, we still need to instantiate this formally. Chomsky's definition of *closer to* in (15) is too weak for our purposes. Hence, I propose to keep the original simpler definition of closeness, the one which is merely based on c-command, as in (21), where K is the attractor; and redefine the definition of Attract, as in (22).

(21) β is closer to K than α if β asymmetrically c-commands α.

(22) K attracts α only if α is outside the minimal domain of K and there is no β, β closer to K than α, such that K attracts *β*.

The definition in (22) essentially keeps the minimality part of the standard definition of Attract and specifies the domain of Attract, which was standardly assumed to be the c-command domain of the attracting head, as made explicit in Kitahara (1997).

On the new definition, the domain of Attract is explicitly identified as the c-command domain outside the minimal domain of the attracting head.

3.3. No-turning-back Principle

Recall from the Footnote 8 that we must not allow head-movement in V2 languages to create one minimal domain as big as a clause. Even if there is a prior step of head-movement to T^0 before T^0 moves to C^0, it should not affect the possibility of extracting the lowest wh-phrase. That is, the calculation of the minimal domain of a chain should not be transitive: each new chain link should not extend the minimal domain of the previous link. Such lack of transitivity is assumed in Chomsky (1993) and Bobaljik and Jonas (1996), and can be extended to the analysis in Chomsky (1995). It is less doable in the revised version of Chomsky (1995) that I have been pursuing here.

Another undesirable property of the system is the stipulative nature of the definitions of domain, maximal domain and minimal domain, adopted from Chomsky (1993, 1995). To avoid both of these problems, it is worth exploring an entirely different approach to the relation between T-to-C movement and Attract. I sketch such an approach below.

The idea is based on the fact that T^0 at some point establishes a checking relation with the material in SpecTP, namely, in checking the Φ-features and Case. When T^0 further moves to C^0, forms a feature bundle with it and becomes part of the attractor that attracts SpecTP, there is an effect of T^0 coming into a checking relation with SpecTP for the second time. Hence, what seems to be active here is a condition against a given head coming into a checking relation with the same element more

than once. In other words, if an element X has already been in a checking relation with an element Y and then moved on to a different position in the structure, it cannot establish a checking relation with that element again. That is, elements have only one shot at checking all the relevant features against a given item. We can refer to it as a No-turning-back (NTB) Principle, which can be formulated as in (23).

(23) No-turning-back Principle:

A feature-checking relation between X and Y cannot be established more than once at different points in the derivation.

It should be clarified that checking multiple features of a single head at a single point in a derivation is not considered a multiple checking relation because both elements remain in the same positions throughout feature-checking.

On this analysis, we still maintain the special status of a wh-phrase in SpecTP with respect to T-to-C movement and Superiority by virtue of SpecTP being the only position that establishes a checking relation with T^0 before T-to-C movement takes place. This prevents C^0 from attracting the elements from the closest SpecTP, allowing lower wh-elements to be attracted instead. This analysis does not have the transitivity problem in V2 languages and the stipulative notions of domain, maximal domain, and minimal domain are not needed here.

3.4. Back to main-embedded clause asymmetry

Now that we have examined the effect of T-to-C movement on the locality of Attract, recall that this produces the weakening effect on Superiority only in matrix questions in English, as shown in the paradigm in (24), repeated from (5).

(24) a. Who bought what?

b. ??What$_1$ did who buy t_1?

c. *John wonders what$_1$ who bought t_1.

Let us now consider the situation in embedded clauses, as in (24c). Since the embedded clauses in English do not involve T-to-C movement, T^0 will establish a checking relation with SpecTP only once and C^0 can freely establish its own checking relation with SpecTP. Hence, the object wh-phrase cannot be attracted by C^0 in this case over the subject wh-phrase (by Attract Closest).

Now a question arises as to why matrix questions like (24b), although better than their embedded counterparts, are still degraded to some extent. According to the analysis so far, nothing prevents the interrogative C^0 in (24b) from attracting the object wh-phrase, because of T-to-C movement. Therefore, the sentence should be fine. In the next section, I address this remaining degraded status of (24b).

4. Interpretive Superiority

Considering the effect of T-to-C movement on the derivation, as discussed in Section 3, the degraded status of (24b) cannot be a result of a minimality violation. Hence, it must be caused by some independent factor. I suggest that the badness of (24b) results from the independently present semantic properties of multiple interrogatives. Particularly, my account is concerned with the licensing conditions on Single-Pair (SP) and Pair-List (PL) readings in multiple interrogatives.

Multiple interrogatives can potentially have a PL or a SP reading. The question in (26) with the PL reading is felicitous in a scenario as in (25). An expected

response to such a question constitutes a list of propositions involving ordered pairs, as in (27).

(25) *PL Scenario:* John is at a formal dinner where there are diplomats and journalists. Each journalist was invited by a different diplomat. John wants to find out all the details, so he asks the host:

(26) Who invited who to the dinner?

(27) Mr. Smith invited Mr. Jones, Ms. Black invited Mr. Green…

A scenario corresponding to the SP reading is given in (28). English lacks the SP reading in questions with bare wh-phrases as in (26), as first observed by Wachowicz (1974). However, we can use a question with discourse-linked (D-linked) wh-phrases, where the SP reading is available in English, as shown in (29).[11] A felicitous response to a single-pair question is given in (30).

(28) *SP Scenario:* John knows that a very important diplomat invited a very important journalist to a private dinner. John wants to find out all the details, so he asks the caterer:

(29) Which diplomat invited which journalist to the dinner?

(30) Ms. Black invited Mr. Smith.

[11] I use the notion of D-linking as in Pesetsky (1987), referring to a wh-phrase whose meaning involves a presupposition that the speaker and the addressee share the knowledge of the exact members of the set over which such wh-phrase ranges.

The distribution of PL/SP readings is subject to cross-linguistic variation, as observed by Hagstrom (1998), Bošković (2003), and Grebenyova (2004). As mentioned before, the SP reading is unavailable in the English bare multiple wh-questions, (31a). The same is true of Bulgarian and Russian, as demonstrated in (31b) and (31c). However, the SP reading is freely available in Serbo-Croatian and Japanese, as can be seen in (32a) and (32b) respectively. That is, unlike the questions in (31a) – (31c), the questions in (32a) – (32b) are felicitous in both PL and SP scenarios.

(31) a. PL/*SP

Who invited who to the dinner?

b. *PL/*SP*

Koj kogo e pokanil na večerjata? *Bulgarian*
who whom Aux invited to dinner
'Who invited who to the dinner?'

c. *PL/*SP*

Kto kogo priglasil na užin? *Russian*
who whom invited to dinner
'Who invited who to the dinner?'

(32) a. *PL/SP*

Ko je koga pozvao na večeru? *Serbo-Croatian*
who aux whom invited to dinner
'Who invited who to the dinner?

b. *PL/SP*

Dare-ga dare-o syokuzi-ni manekimasita-ka? *Japanese*
who-Nom who-Acc dinner-Dat invited-Q
'Who invited who to the dinner?

In languages that allow SP readings in multiple interrogatives, fronting the lower wh-phrase over the higher wh-phrase forces the SP reading. Hagstrom (1998) observes this phenomenon with respect to Japanese (33a) and Bošković (2003) reports the same for Serbo-Croatian (33b). Bošković (2003) refers to this phenomenon as *Interpretive Superiority*, meaning that movement of the lower wh-phrase over the higher takes away only one of the two potential readings, instead of producing complete unacceptability.

(33) a. **PL/SP*

Nanio$_1$ darega t_1 katta no? *Japanese*
what$_{ACC}$ who$_{NOM}$ bought Q
'Who bought what?'

b. **PL/SP*

Šta$_1$ je ko kupio t_1? *Serbo-Croatian*
what is who bought
'Who bought what?'

What would happen if a similar fronting of the lower wh-phrase over the higher one took place in a language where SP readings are unavailable in multiple interrogatives, like in English? It is plausible that, if a SP reading is forced in such a

language, the complete unacceptability should be expected.[12] Let us consider our crucial example in (24b), repeated below as (34).

(34) ??What$_1$ did who buy t_1?

In this example, the object wh-phrase *what* is fronted over the subject wh-phrase *who*. Such fronting forces the SP reading, as we observed in Japanese and Serbo-Croatian, languages that actually allow such a reading in questions with bare wh-phrases. But as was demonstrated in (31), the SP reading is unavailable in English bare multiple questions. I suggest that this is precisely what causes the degraded status of (34).

Let us now consider the embedded clauses in English. There is no asymmetry between main and embedded clauses with respect to the PL/SP readings distribution: the embedded interrogative in (35) has only a PL reading.[13]

(35) John wonders who bought what. *PL/*SP*

This means that fronting the object wh-phrase over the subject wh-phrase in an embedded clause in English, as in (36), should invoke the same effect of Interpretive Superiority as in the main clause.

(36) *John wonders what$_1$ who bought t_1.

[12] I abstract away from the interrogatives with D-linked wh-phrases for now.

[13] The infelicitous scenario is where John wonders about the identity of exactly one individual and of exactly one item which that individual bought. Instead, the English speakers understand (35) as describing a situation where John wonders about the list of pairs of individuals and items they bought.

Thus, Interpretive Superiority is one of the sources of the badness of (36). In addition, the absence of T-to-C movement in these contexts invokes a violation of Attract Closest (as discussed in section 2). These two factors combined make the Superiority effects in embedded clauses worse than those in main clauses, where only one factor (Interpretive Superiority) is involved.

I will examine the formal nature of the PL and SP readings and what underlies the phenomenon of Interpretive Superiority in Chapter 3, which is devoted to the semantics of multiple interrogatives. For the remainder of this chapter, let us explore the predictions and consequences of the analysis above for the Superiority effects cross-linguistically, as well as for the theory of syntactic movement.

5. Implications and Consequences

5.1. Cross-linguistic Predictions

One straightforward prediction of the present analysis is that in a language where T-to-C movement takes place in both main and embedded clauses and SP readings are available in bare multiple questions, we should not expect to find any Superiority effects in either main or embedded clauses. Such a language is Icelandic, where V2 (i.e., verb movement to C^0 via T^0) occurs in both main and embedded clauses. Icelandic also allows SP readings in wh-questions: (37a) and (38a) are perfectly acceptable on the SP reading. As expected, there are no Superiority effects in either main or embedded clauses, as demonstrated by the lack of contrast between (37a) and (37b), and between (38a) and (38b).

(37) a. *PL/SP*

Hver bauð hverjum í veisluna? *Icelandic*
who invited whom in the-dinner
'Who invited who to the dinner?'

b. *?PL/SP*

Hverjum bauð hver í veisluna?
whom invited who in the-dinner
'Who invited who to the dinner?'

(38) a. *PL/SP*

Jón veit ekki hver bauð hverjum í veisluna.
John knows not who invited whom in the-dinner
'John does not know who invited who to the dinner.'

b. ?PL/SP

Jón veit ekki hverjum bauð hver í veisluna.
John knows not whom invited who in the-dinner
'John does not know who invited who to the dinner.'

The PL reading is harder to get in (37b) and (38b). The SP reading is preferred in these contexts. This seems to be another instance of Interpretive Superiority, similar to the facts from Japanese and Serbo-Croatian, discussed in section 3. The effect is, however, weaker in Icelandic since, unlike in Japanese and Serbo-Croatian, the PL reading is still available in the context of object fronting in Icelandic, just not as easily available as without the object fronting.

Our overall analysis has certain predictions about the structure of (37a) and (38a), where the subject wh-phrase seems to be in SpecCP. On our analysis the

clauses in these examples cannot be CPs. Otherwise, the complex head $[C\text{-}T]^0$ would not be able to attract the subject wh-phrase. The problem does not arise if subject V2 clauses are TPs and the subject wh-phrase in these examples is in SpecTP. This view is supported by the work of Travis (1991) and Zwart (1991, 1993), who argue that subject, unlike non-subject V2 clauses, in Germanic are TPs. Later in this section, we will extend this analysis to English subject vs. non-subject wh-questions.

The mirror image of Icelandic is Brazilian Portuguese, where T-to-C movement does not take place in either main or embedded clauses and the language does not allow SP readings in bare multiple questions. The lack of T-to-C movement is shown in (39b) in the context with a main verb and (39d) and (39e) demonstrate the same with an auxiliary.

(39) a. O quê (que) o Diogo comprou? *Brazilian Portuguese*
the what that the Diogo bought
'What did Diogo buy?'

b. *O quê comprou o Diogo?
the what bought the Diogo

c. O quê (que) o Diogo vai comprar?
the what that the Diogo will buy
'What will Diogo buy?'

d. *O quê (que) vai o Diogo comprar?
the what that will the Diogo buy

e. *O quê (que) vai comprar o Diogo?
the-what that will buy the Diogo

More extensive arguments for the absence of T-to-C movement in Brazilian Portuguese can be found in Silva (2001). There is also an explanation of the historical loss of T-to-C movement in this language in Pires (2004), based on clitic placement.

As for the interpretation of multiple questions in Brazilian Portuguese, consider the examples in (40a) and (40c) below, which only allow PL readings, and are unacceptable on the SP readings. Given these facts, the analysis developed in this chapter predicts Superiority effects to be equally strong in Brazilian Portuguese in both main and embedded clauses. The prediction is borne out: (40b) and (40d) are equally unacceptable.

(40) a. PL/*SP
Quem (que) comprou o quê. *Brazilian Portuguese*
who that bought the what
'Who bought what?'

b. *O quê (que) quem comprou?
the what (that) who bought
'What did who buy?'

c. PL/*SP
Max (me) perguntou quem (que) comprou o quê?
Max to-me asked who (that) bought the-what
'Max asked me who bought what?'

d. *Max (me) perguntou o quê (que) quem comprou?
Max to-me asked the what (that) who bought
'Max asked me who bought what?'

There is another language that behaves similarly to Brazilian Portuguese with respect to Superiority effects, namely, Bulgarian. It exhibits equally strong Superiority effects in matrix and embedded clauses. However, it has an interfering factor. That is, subject-aux(iliary) inversion is required in Bulgarian wh-questions.[14] The obligatoriness of inversion in main clauses is demonstrated in (41), as argued for in Rivero (1994), among others.

(41) a. Koe pismo napisa deteto? *Bulgarian*
which letter wrote the-child
'Which letter did the child write?'

b. *Koe pismo deteto napisa?
which letter the-child wrote

Izvorski (1993) reports that the obligatoriness of inversion also holds in the embedded questions in Bulgarian, as shown in (42).[15]

(42) a. Tja me popita kâde živee Ivan. *Bulgarian*
she me asked where lives Ivan
'She asked me where Ivan lives.'

b. *Tja me popita kâde Ivan živee.

[14] The inversion in Bulgarian actually applies to main verbs as well as auxiliaries.

[15] The examples in (41) and (42) are from Izvorski (1993).

The question arises whether the inversion in Bulgarian is an instance of T-to-C movement. First, it is already different from English, because, unlike in English, it takes place in both main and embedded clauses and can apply to main verbs. But that by itself is not enough to conclude that we are dealing with a different kind of movement in Bulgarian. Recall, for instance, the presence of T-to-C movement with main verbs in the embedded clauses in Icelandic. However, there are still reasons to believe that the inversion in Bulgarian is not a result of T-to-C movement. Izvorski (1993) provides several arguments to this effect, one of which is based on the fact that adverbs can precede the verb in wh-questions in Bulgarian. This is demonstrated with an IP-adverb in (43a), and with a VP-adverb in (43b). These data indicate that the verb remains in the TP domain after inversion.

(43) a. Za kakvo včera spomena Ivan pred Maria? *Bulgarian*
about what yesterday mentioned Ivan to Maria
'What did Ivan mention to Maria yesterday?'

b. Kakvo veče kupi Ivan?
what already bought Ivan
'What did Ivan buy already?'

Izvorski (1993) argues that the inversion phenomenon is not an instance of the rightward movement of the subject, based on the data in (44) with an extra argument in the VP and the subject preceding that argument instead of occurring sentence finally. The same point can be made with respect to (43a).

(44) Otkâde znae Paulina vsičko tova? *Bulgarian*
from-where knows Paulina all this
'Where does Paulina know all this from?'

Based on these and a few other arguments, Izvorski reaches the conclusion that the verb moves to T^0 in Bulgarian, while the subject remains in situ. Thus, we can conclude that the source of inversion in Bulgarian is not T-to-C movement. Given that Bulgarian lacks T-to-C movement and given that it lacks SP readings in multiple interrogatives, as was demonstrated in (31b), we predict there to be no contrast between the matrix and embedded Superiority violations in this language. The prediction is borne out, as shown below.

(45) a. *Kogo koj e pokanil na večerjata? *Bulgarian*
whom who Aux invited to dinner
'Who invited who to the dinner?'

b. *Tja me popita kogo koj e pokanil na večerjata?
she me asked whom who Aux invited to dinner
'She asked me who invited who to the dinner.'

Thus, the cross-linguistic data from English, Icelandic, Brazilian Portuguese and Bulgarian support the proposed analysis of Superiority.

5.2. Superiority in Non-subject Questions

One of the implications of the present analysis is that, in English, T-to-C movement should not affect the locality of Attract in multiple questions that do not involve a wh-phrase in matrix SpecTP. Thus, we expect to find an asymmetry between subject and

non-subject wh-questions with respect to Superiority. This is precisely what we find in English (46a) – (46d).

(46) a. ??What did who buy?
b. *What did Mary tell who to buy?
c. *Bill wonders what Mary told who to buy.
d. *Bill wonders what who bought.

The sentence in (46a), questioning the matrix subject, is less degraded than all the other members of this paradigm: (46b), questioning the object of a control clause, (46c), where the embedded clause is questioning the direct and the indirect objects, and the familiar (46c) with the embedded clause questioning embedded subject.

To control for the degree of clausal complexity in (46a) and (46b), I have tested the paradigm in (47), where both examples are mono-clausal.[16]

(47) a. ??What did who buy?
b. ???Who did John give what to?

The contrast goes in the direction that is predicted, although is not as clear because speakers seem to slightly prefer (47) to the bi-clausal examples in (46b) and (46c). That is why I marked them with ??? to express this contrast.[17]

[16] Interestingly, for some speakers, the example in (47b) improves with preposition pied-piping: To whom did John give what? I do not have an explanation of this at this point.

[17] All three of my informants prefer the form *who* to *whom* and freely allow preposition stranding.

5.3. Subject Extraction and T-to-C Movement

We have discussed subject wh-questions in English in Section 3 briefly. Let us now examine it in more detail. Recall that my analysis prohibits attracting elements from SpecTP if T-to-C movement has taken place (by NTB Principle). This offers a potential answer to the long-standing question of why sentences like (48) are unacceptable in English.

(48) *Who did leave?

The problem is often approached with an attempt to prohibit T-to-C movement in the context of subject wh-movement.[18] However, on the analysis developed here, it is the subject wh-movement that is not permitted when T-to-C movement has applied. This correctly rules out (48).

Now consider the paradigm in (49).

(49) a. *Did who leave?
b. Who left?
c. Who bought what?

The contrast between (49a) and (49b) can be captured if the complementizer C^0 is not present in the structure in these particular cases and therefore T-to-C movement cannot take place. The absence of CP would then also apply to (49c). Thus, the subject wh-phrase may not be raising higher than TP in these configurations. The interrogative force must be then located in T^0 in these constructions. For a similar proposal, see George (1980), Chomsky (1986) and Pesetsky (1989).

[18] But see Footnote 9 for an alternative analysis of this from Chomsky (1955, 1957).

And note that T^0 cannot attract the object wh-phrase in a single interrogative like in (50a).

(50) a. What Mary bought?

b. *[$_{TP}$ What$_2$ Mary$_1$ T^0 t_1 bought t_2]?

As demonstrated in the derivation in (50b), there is no room in SpecTP for *what*, if *Mary* is already in SpecTP to satisfy the EPP requirement. Recall that English does not allow multiple specifiers, as indicated by the lack of multiple wh-fronting in this language.

The conclusion that subject wh-phrases do not move to SpecCP in main clauses in English has a particular implication for the analysis of main clause sluicing with a subject wh-remnant, as in (51).

(51) *Speaker A:* Someone left.

Speaker B: Who ~~[left]~~?

If *who* in (51) is not in SpecCP, how can it survive sluicing, under the standard assumption that sluicing is an instance of TP-ellipsis? I suggest that the surface position of *who* is actually SpecAgr$_S$P, as in (52) or, possibly, a focus phrase, as will be discussed in Chapter 4).

(52) *Speaker A:* Someone left.

Speaker B: [$_{AgrSP}$ Who [$_{TP}$ ~~left~~]]?

This analysis of the optional presence of CP projection can be extended to other languages where wh-phrases have been argued to not move overtly all the way to SpecCP, as, for example, in Bošković's (1997a, 1998, 2002a) treatment of Serbo-Croatian main clauses with the null complementizer and Stepanov's (1998) treatment of Russian wh-questions in both main and embedded contexts.

There are, however, arguments in the literature for the existence of the vacuous movement of the subject wh-phrase to SpecCP. Those can be found in Cheng (1991), Rizzi (1990, 1996), Boeckx (2003), and An (To appear). Most of the arguments motivate subject wh-movement through clausal typing or feature checking. That is, on these accounts, if a subject wh-phrase does not move to SpecCP, the clause will not be typed as interrogative or the *+wh* feature of C^0 would not be licensed. This, however, is not a problem for my particular account, on which the *+wh* feature may reside on T^0 and hence can be checked without movement to SpecCP. The clausal typing requirement can also be dealt with if we assume that a clause can be typed at any projection that happens to be the highest phonologically realized projection in a clause (which can be lower than CP). See Bošković's (2002a) for making the same assumption about clausal typing with respect to many other phenomena.

An (To appear) presents a new kind of argument for subject wh-movement to SpecCP, based on his Intonational Phrase Edge Generalization (IPEG) that prohibits the specifier and the head of a clause to be empty at the same time. This crucially presupposes that T-s cannot be heads of clauses for this purpose. Once again, if we allow an interrogative feature to sometimes reside on T^0, there seems to be no reason

for T^0 not to be able to function as a head of the clause. The IPEG seems quite similar to clausal typing of Cheng (1991), as a more rationalized and formalized version of it. If this correlation is on the right track and the precise interaction of IPEG and clausal typing can be established, the status of subject wh-movement to SpecCP can be made more precise.

Let us explore what the system developed in this chapter predicts if it turns out that subject wh-phrases do actually move to SpecCP overtly. In the current system, this would be possible only if this movement takes place before T-to-C movement takes place. This would bring back the original optionality of the account of Chomsky (1993, 1995), described in Section 3. The difference is that, by assuming the NTB Principle, we can account for the relevant empirical facts without appealing to the notions of minimal domain and 'neighborhood'.

6. Attract vs. Move

The analysis developed in this chapter crucially relies on the view of movement where the trigger for movement is not on the moving element but rather on the target of movement (e.g., the uninterpretable *+wh* feature on an interrogative C^0). The movement operation on this view is Attract (i.e., C^0 with the *+wh* feature attracts a wh-phrase for feature checking). Attract can be further decomposed into Agree and Move and Move can be further decomposed into Copy and Merge, as in Chomsky (2000). In what follows, I will refer to the overall concept of placing the trigger for movement on the target of movement as the Attract-analysis.

The alternative view of movement is from the perspective of the moving item where some feature of the phrase that undergoes movement triggers the movement

(i.e., an uninterpretable *+wh* feature on a wh-phrase triggers wh-movement in order to check that feature against the interpretable *+wh* feature of C^0). I will refer to this as the Move-analysis. As discussed in section 2 of this chapter, the approach based on Move is incompatible with capturing the interaction between head-movement and Superiority. However, the Attract-approach has a potential problem of explaining successive cyclicity. The goal of this section is to explore the two approaches to movement in some detail, focusing on how each approach handles successive cyclicity.

6.1. Look-ahead

One of the main motivations for introducing Attract in Chomsky (1995) was eliminating the look-ahead problem with respect to the moving element. With the Bare Phrase Structure (BPS) of Chomsky (1994), the phrase structure is no longer introduced into the derivation all at once but rather is built in a piece-meal fashion, through the successive application of Merge (an operation putting two elements into a set). This creates a look-ahead problem for the approach to movement based on Move: the (strong) uninterpretable feature on a given phrase remains unchecked until the target of movement is introduced, which could be arbitrarily many clauses away, given the possibility of long-distance movement. The problem is that the *+wh* feature of the wh-phrase cannot be checked for an arbitrary long period of time.

Attract reduces the 'waiting' problem because the uninterpretable feature is introduced into the derivation at the time when the target head is Merged into the structure, which triggers immediate Attract (for feature-checking). Note, however, that this only partially eliminates look-ahead because it is still there when it comes to

phases and the Activation Condition of Chomsky (1999, 2000). We will turn to that shortly.

6.2. Successive Cyclicity

Many cross-linguistic facts indicate that phrasal movement proceeds successive cyclically, as explored in Chomsky (1973, 1986, 2000, 2001), McCloskey (1990), Chung (1982), Torrego (1984), among others. For instance, McCloskey (1990) demonstrates that in long-distance extraction in Irish, the complementizer which is morphologically specified for wh-movement appears in every clause, including the intermediate CPs. The data in (53) shows that this particular complementizer is *a* and (54) shows that it can be realized as the intermediate C^0s, which themselves do not carry a +*wh* feature.

(53) a. Dúirt sé [$_{CP}$ **gur** bhuail tú é] *Irish*
said he COMP struck you him
'He said that you struck him'

b. an fear [$_{CP}$ **a** bhuail tú *t*]
the man COMP struck you
'the man that you struck'

(54) an rud [$_{CP}$ **a** shíl mé [$_{CP}$ **a** dúirt tú [$_{CP}$ **a** dhéanfá]]]
the thing COMP thought I COMP said you COMP do-COND-2SNG
'the thing that I thought you said you would do'

The situation is schematized below:

(55) $[_{NP}$ NP $[_{CP}\ a \ldots [_{CP}\ a \ldots\ [_{CP}\ a \ldots]]]]$

The effects of successive cyclicity can be found in English as well in a form of a familiar wh-island effect, represented by the contrast in (56). Although wh-movement is ultimately unbounded, as indicated by (56a), it must proceed in a step-by-step fashion. The presence of *how* in the embedded SpecCP blocks the movement of *what* into that position, preventing the movement from proceeding successive cyclically.

(56) a. What_1 does Mary think $[_{CP}\ t_1$ that John fixed $t_1]$?
b. ??What_1 does Mary wonder $[_{CP}$ how John fixed $t_1]$?

Chomsky's (1973) and (1986) analyses of successive-cyclicity captured this phenomenon from the perspective of Move. In a framework with feature-checking, as I mentioned earlier, Move has the look-ahead problem with respect to the checking of the strong uninterpretable feature. Attract-approach makes successive cyclicity quite mysterious as well, for it is not clear why the moving item needs to stop at any intermediate position on the way to its ultimate landing site. Given this, there are several ways to address successive cyclicity.

One possibility is to adopt Chomsky and Lasnik's (1993) Minimize Chain Link Principle, which requires chain links to be as short as possible (i.e., subject to Subjacency, as in Chomsky (1986)). Thus, in (56a), the *+wh* feature of the matrix C^0 attracts the wh-phrase and the movement goes through the intermediate SpecCP by MCLP. This is the direction Takahashi (1994a) pursues. Such an approach, however, requires an additional operation Form Chain, since the analysis crucially relies on chains, and is Move-based.

An alternative approach to successive cyclicity that does not make use of Form Chain is that of Chomsky (1999, 2000, 2001). It is based on Phase-Impenetrability Condition (PIC) and Generalized EPP. PIC is a condition on extraction out of certain categories that are considered phases and it states that only the head and the Spec of the phase, constituting the edge of the phase, are available to the operations outside that phase.[19] The movement of a phrase to the edge of the phase is driven by the EPP property optionally assigned to the head of the phase. The optionality of the EPP assignment is needed in order to solve the look-ahead problem. This kind of EPP is known as Generalized EPP (the original EPP is a constant property of T^0 in a language like English, requiring for SpecTP to be occupied at some point in the derivation). Bošković (2005) points out that Generalized EPP creates a problem for (57) since nothing prevents the declarative complementizer *that* from having an EPP property, triggering the movement of what to the intermediate SpecCP and staying there. Note that the +wh feature in the main clause is checked by *who*.

(57) *Who thinks [$_{CP}$ what$_1$ that Mary bought t_1]?

On the MCLP analysis, where there is no relation of any sort between the complementizer *that* and a wh-phrase, and no PIC either, (57) is ruled out by Last Resort: once *who* is attracted to check the +*wh* feature of C^0 (this feature can also be

[19] The intuition behind PIC is that Spell-out takes place cyclically and it applies to the complement of the head of each phase. The actual deduction of PIC from the properties of Spell-out can be found in Fox and Pesetsky (2005), and Bošković (2005). Uriagereka (1999) reaches a similar result without dealing with PIC proper.

located on Agr_S^0 or Foc^0, as suggested in the previous section), nothing motivates the overt movement of *what*.

Generalized EPP and the general approach of attributing the driving force of the intermediate steps of movement to the intermediate projections face another problem. It has to do with Agree being a prerequisite for movement, as proposed in Chomsky (2000). This requires the intermediate steps of successive cyclic movement to involve feature-checking. However, Bošković (2002b) and Boeckx (2003) present a number of arguments against feature-checking in the intermediate positions.

6.3. Bošković (2005)

Bošković (2005) presents an alternative account of successive-cyclicity, which, as its goal, does not make use of either Form Chain or Generalized EPP. The proposal is that the uninterpretable feature driving movement is always on the moving item. That is, in wh-movement in English, it is always on one of the wh-phrases and not on C^0. Bošković also adopts PIC, although deriving it from cyclic linearization, as in Fox and Pesetsky (2005). Given PIC, when the intermediate CP is created in a derivation, the uninterpretable feature of a wh-phrase inside the domain of the phase motivates the movement of the wh-phrase to the intermediate SpecCP. This movement step does not involve feature checking but is rather 'agnostic', in the sense that it happens just to ensure that feature-checking at a later point can take place.[20] The analysis of Bošković (2005) crucially relies on the view of movement from the perspective of the moving item, so I will refer to this analysis as Move-analysis.

[20] A similar idea can be found in Chomsky (1995, 1999).

There are several potential difficulties with this analysis. First, in a deterministic system, it is difficult to determine when a wh-phrase is assigned an uninterpretable feature without look-ahead.

(58) a. What did John buy?
b. Who bought what?

What moves in (58a), but it does not in (58b). The only difference between the two examples is the presence of another wh-phrase, *who*, in (58b). On the Attract-analysis, the picture is rather clear: there is only one interrogative complementizer and it is the complementizer that has the uninterpretable *+wh* feature. Since *who* checks that feature in (58b), *what* can remain in situ. However, on the Move-analysis of Bošković (2005), it needs to be determined when *what* can appear in a derivation with the *+wh* feature and when it cannot bear this feature. The distribution of two different lexical items for *what* depends on the presence of another wh-phrase in the structure, which can be introduced at a much later point in the derivation. In a deterministic system, this creates a look-ahead problem, similar to the one associated with Generalized EPP, where the assignment of the EPP requirement to a given head depends on whether there is another wh-phrase higher in the structure. The same factor seems to govern the assignment of the uninterpretable feature to a wh-phrase.

The problem goes away if the wh-phrases can optionally occur in the derivation with or without the *+wh* feature, producing among the successful derivations many crashing derivations. This is the strategy that is used by Bošković (2005). The same logic seems to be used in the theory of Generalized EPP by

Chomsky: allowing phase-heads to appear optionally with or without the EPP-feature. The question is whether this is a satisfactory solution. The answer seems to depend on how much optionality (if any) an optimal system should have. For the moment, we can at least conclude that both the Move-analysis and the Generalized EPP analysis face similar challenges with respect to look-ahead or optionality.

There also seems to be a conceptual problem in allowing 'agnostic' initiation of movement. It undermines the motivation behind feature-checking, since its basic purpose is to drive movement. Besides, movement in order to do something at a later point in the derivation, which could be indefinitely many steps away, seems again to involve look-ahead.

Move-analysis of Bošković (2005) also loses the account of Superiority that the Attract-analysis made possible and quite insightful. Park (2005) proposes an alternative analysis of Superiority that is compatible with the Move-analysis. It is based on the notion of Chain Uniformity. The idea is that, when movement violates locality, the violation is encoded by placing a * on the trace, the moving element, and a barrier that is crossed. The uniform chains, on Park's analysis, are the ones that either have * on all the members of the chain or on no members of the chain.

In case of a Superiority violation, Park proposes that the intervening wh-phrase is also marked with a *, as demonstrated schematically in (59a).

(59) a. *what$_1$ *who *t_1

b. *what$_1$ who$_2$ *t_2 *t_1

It is assumed that the actual principle that is violated here is Relativized Minimality of Rizzi (1990). The two chains (the chain of movement of *what* and the trivial chain of *who*) are uniform so far since all their members are marked with *. However, the further movement of *who*, is local and creates a non-uniform chain since only the trace of *who* remains marked with a *. The movement of the intervening wh-phrase is most clearly seen in Bulgarian, where all wh-phrases eventually move to SpecCP overtly. As for English, the crossing could happen inside *v*P, and *who* would then move to SpecTP.[21]

This novel approach to Superiority, with a few additional assumptions about focus and reconstruction, captures all of the Superiority facts that the Attract-analysis does.[22] The question is how compatible it is with the Move-analysis of Bošković (2005). Recall that one of the main motivations behind Bošković's analysis is to avoid Form Chain, which is the ingredient in MCLP of Chomsky and Lasnik (1993) and the analysis of Takahashi (1994), who adopts MCLP. However, the analysis of Park (2005) crucially relies on the existence of chains in the system, bringing back Form Chain. Given that Generalized EPP, the other undesirable ingredient that Bošković (2005) is trying to eliminate, is not part of Takahashi's Attract-based analysis, the analyses of Bošković (2005) and of Takahashi (1994) again become equally plausible. The only remaining difference is that, on Takahashi's account, movement to the intermediate projections in successive-cyclic movement takes place after the actual target of movement is introduced, while, on Bošković's account, the

[21] It is unclear what happens in questions with two object wh-phrases, or with an object and an adjunct wh-phrases, where the intervening wh-phrase may not necessarily move. Perhaps, covert wh-movement would solve this potential problem.

[22] For more details, please see Park (2005).

intermediate steps happen early, making this account more compatible with general idea of multiple spell-out. We will discuss multiple spell-out in the next section.

The fact that the Attract-analysis inherently avoids the specific look-ahead problem of the Move-analysis described earlier and the fact that it allows us to capture the cross-linguistic effects of the interaction of T-to-C movement and Superiority, which we observed in the previous sections, lead me to the conclusion that the Attract-analysis is still a strong alternative to the Move-analysis.

6.4. Another look at successive cyclicity

The conclusion reached in the discussion above is that it seems beneficial to keep the Attract-approach to movement in the system. However, where does this leave us with respect to successive cyclicity? As was discussed above, Generalized EPP triggering movement to intermediate positions is not likely to be on the right track. The alternative analysis, based on MCLP, avoids the problems of Generalized EPP but uses Form Chain. Another approach to successive cyclicity that keeps Attract in the system, yet avoids Generalized EPP, is that of Lasnik and Uriagereka (2005). It relies on Chain Uniformity of Chomsky (1991) and lexical selection. Without going into the details of this account, it is clear that, like the MCLP account, it relies on Form Chain. To be precise, both of these accounts rely on the existence of chains. Whether chains are created by a separate operation Form Chain or are merely a by-product of the Copy Theory of movement is a separate question. Even if the operation Form Chain is not necessary for the existence of chains, the question still remains whether chains as a theoretical construct should be part of our theory; that is, whether conditions on chains, such as MCLP and Uniformity, should exist.

At this point it is not clear if we can dispense with chains as theoretical constructs, given that a number of principles rely on them. Besides, not only principles, but other modules might operate on syntactic chains as well, for instance, linearization at PF, as in Nunes (2004). The account of the interaction of T-to-C movement and Attract, developed in this chapter, also relies on chains, providing another argument for the existence of chains. Thus, the accounts of successive cyclicity that involve MCLP or Chain Uniformity might still be on the right track.[23]

The discussion above should not, however, stop us from searching for an alternative account of successive cyclicity which would allow us to maintain Attract in the system, yet without relying on chains or Generalized EPP. I suggest that one such analysis is that of Ochi (1999), which builds on the theory of feature-movement developed in Chomsky (1995). On this view, only formal features are attracted by the attracting head X since that is all that X needs for feature checking. Nothing else happens if Attract F (i.e., the operation that only applies to formal features) takes place in LF. If, however, it takes place in overt syntax, the category left in-situ is uninterpretable at PF without the missing feature. Hence, the category must pied-pipe to the minimal domain of the head that attracted its feature and reunite with that feature.[24]

Ochi (1999) develops this analysis further by proposing that category pied-piping is subject to Subjacency of Chomsky (1986). In other words, category pied-

[23] Choosing between these two accounts is beyond the scope of this particular work.

[24] For several insightful analyses based on this theory of feature movement, see Lasnik (1995, 1999, 2001) and Lasnik and Uriagereka (2005).

piping proceeds successive cyclically.[25] This analysis allows us to have both Attract and Move in the system in a rational way: Attract is associated with feature-movement and Move is associated with category pied-piping. The two happen for two different reasons; hence, it is plausible that they are different in nature.

Notice that on this account, the successive-cyclic wh-movement cannot begin until the formal feature is attracted, which means that it cannot begin until the target of movement is introduced. (The same is true in Takahashi's (1994) analysis). But this does not create any look-ahead because the wh-phrase has no inadequacy of any sort (i.e., there is no uninterpretable feature on the wh-phrase). This analysis seems to have an important consequence for Attract and PIC: the matrix C^0 must have the access to the *+wh* feature of the wh-phrase no matter how far down it might be in the structure. Bošković (2005) reaches the same conclusion about Agree, namely, that it is not subject to PIC (or whatever underlies it).[26]

The question still remains why category pied-piping is successive cyclic. Let us consider a few possibilities. First, could some version of multiple spell-out motivate the intermediate movement steps of the pied-piped category? Given the timing of this movement (i.e., it begins only after the target of movement is introduced and the relevant feature is attracted), it would be difficult to achieve, requiring a separate cycle for spell-out and linearization.

---

[25] Ochi (1999) actually treats the wh-island effects as due to Relativized Minimality, and the other islands, as due to the category pied-piping being subject to Subjacency. This is still consistent with this overall account of successive cyclicity.

[26] With Agree, however, no feature-movement takes place since it is a long-distance checking operation. Hence, no category pied-piping will be necessary. To maintain a feature-movement analysis, Attract seems necessary.

Another option is to derive successive-cyclic category pied-piping from computational necessity. Consider what happens immediately after a given head X attracts a feature F of a category Y. How does Y know where exactly to find the missing feature? Attract F does not leave an "address" where a given feature is taken. If the whole purpose for the category pied-piping is to find the missing feature F, how would this movement proceed? Between any two categories A and B that might host the missing feature F of Y, it is most reasonable for Y to move first to the category which is the closest and check if the feature is there. And only then, if the feature is not there, Y would move to the next category, and so on until it finds its feature. This is demonstrated below with the trajectory of pied-piping of *what*:

(60) [$_{CP}$ **What** does John$_2$ [$_{vP}$ ***t'''*** [t_2 believe [$_{CP}$ ***t''*** that Mary$_1$ [$_{vP}$ ***t'*** [t_1 likes ***t***]]]]?

In addition, if pied-piping of the category is sensitive to c-command (i.e., movement is to a c-commanding position, prohibiting lowering), successive cyclic pied-piping is *required* because, if any projection is skipped, any further movement will only be up and not down. Thus, if the missing feature happens to be in some position that was skipped, as in (61), it will never be found, incorrectly ruling out (61).

(61) John wonders [$_{CP}$ ***what*** Mary likes ***t***]?

This can be what motivates the intermediate steps in successively cyclic movement.

In addition to capturing successive cyclicity without appealing to Generalized EPP or Form Chain, this approach has important consequences for Spell-out. As was

mentioned above, the timing of the category pied-piping is such that it begins only after the target of movement is introduced and the relevant feature is attracted. Does that lead to abandoning multiple Spell-out? For the main 'stem' of the phrase structure tree, that is indeed the case. This is not, however, a bad result. The correlation between Spell-out and phases as being the relevant domains for this operation was not motivated to start with. Phases were introduced in Chomsky (2000) as subnumerations, for Economy considerations having to do with the familiar paradigm in (62).

(62) a. There was believed [$_{IP}$ *t* to be [a unicorn in the garden]]].
b. *There was believed [$_{IP}$ a unicorn to be [*t* in the garden]]].
c. A rumor emerged [$_{CP}$ that there was a unicorn in the garden].
d. There emerged a rumor [$_{CP}$ that a unicorn was in the garden].

The analysis went as follows. An economy principle Merge-over-Move is responsible for the badness of (62b): at the point of creating the structure *to be [a unicorn in the garden]*, instead of merging *there* to SpecTP (to satisfy EPP), *a unicorn* is moved to that position, which is more costly since movement consists of two operations Copy and Merge. The acceptability of (62d) is then a problem since the movement of *a unicorn* takes place here over merging an expletive. Considering that Merge-over-Move is an economy strategy and not an absolute condition on convergence and under the assumption that only derivations based on the same numerations can be compared, Chomsky proposes that the numerations are constructed cyclically (i.e., sub-numeration by sub-numeration), and sub-numerations are considered phases. It is

then necessary that CPs are phases, as in (62c) and (62d), while IPs are not, as in (62a) and (62b).[27]

Thus, if phases are sub-numerations, as in Chomsky (2000), the requirement that Spell-out must be sensitive to sub-numerations is not straightforwardly motivated. Besides, there are other contexts where Spell-out seems to have nothing to do with phases. For instance, complex specifiers are associated with linearization as being the factor that triggers Spell-out in these contexts. Specifically, there is a linearization conflict posed by such specifiers under the view of linearization as the base part of LCA of Kayne (1994), where asymmetric c-command maps to precedence. As is known from the work of Uriagereka (1999), the elements inside a complex specifier do not c-command the elements inside a complement, so the linearization cannot obtain. Spelling-out the specifier turns that category into a single word thus allowing it to be linearized as a non-complex specifier. This way Uriagereka also captures the ban on extraction out of specifiers.

However, CP phases cannot be treated the same way because, crucially, they are complements and the same linearization problem does not apply to complements. Besides, extraction out of complements is allowed.

There is an undesirable disjunction in defining the operation Spell-out. It applies to a category *X* if there is a linearization problem or if it is a phase. Furthermore, in case of phases, Spell-out applies to only the complement of the

[27] For Chomsky, what distinguishes CP from IP is the propositional force of CP. Lasnik and Uriagereka (2005) propose another idea that a phase is a cyclic domain where all relations of a certain type are satisfied. However, it is not easy to determine what relations those are, since, for instance, the establishment of Case/agreement relations seems to make IP a phase. To my understanding, the issue of what categories constitute phases is still not fully worked out.

phase-head, in accordance with PIC, which does not happen in Spelling-out specifiers.

Thus, the consequence of the pied-piping analysis of successive cyclicity is that Spell-out applies early only if triggered by interface conditions, such as linearization or other morpho-phonological requirements, and then it applies once more at the end of the whole derivation.

The potential problem of the feature-movement analysis is in the counter-cyclicity associated with category pied-piping. The intermediate steps of pied-piping do not extend the tree, potentially violating the Extension Condition of Chomsky (1993), which requires that every movement step extends the tree. This potential problem can be resolved if the Extension Condition is deduced from feature strength, as in Bošković and Lasnik (1999).[28] The deduction is achieved by appealing to the 'virus' approach to feature strength of Chomsky (1995), where a strong feature must be checked as soon as it is introduced into the derivation (i.e., a node containing a strong feature cannot be embedded). In this case, category pied-piping is not technically counter-cyclic because it takes place immediately after the feature of the root node is checked and the intermediate steps of pied-piping do not involve feature checking. Thus, the apparent counter-cyclicity of pied-piping is a good result, for it allows us to understand cyclicity better. What looks like counter-cyclic movement is actually allowed by the grammar in this one instance, where the movement takes place for a purpose other than feature checking.

[28] It is possible to use feature uninterpretability instead of feature strength, depending on whether covert movement for feature-checking purposes exists.

7. Summary

Let me summarize the main ideas of this chapter. I have presented an analysis of the contrasts in Superiority effects in main vs. embedded clauses in a number of languages. As a result, we have a refined account of Superiority, which considers both syntactic and semantic properties of multiple interrogatives. On this analysis, one factor (Interpretive Superiority) contributes to the Superiority effects in both main and embedded clauses in English, with an additional factor (the absence of T-to-C movement, allowing for the Attract Closest to be operative) present in the embedded clauses. The evidence based on the Superiority effects (or lack there of) in Icelandic, Bulgarian, and Brazilian Portuguese further support this analysis.

If this analysis is correct, it makes head-movement quite relevant for syntactic processes, which means that it should not be treated as a PF operation, as in Chomsky (2000). This analysis also puts into question the analyses suggesting that there is, in fact, T-to-C movement in embedded clauses in English (e.g., Pesetsky and Torrego (2001)). The proposed analysis can also be viewed as a new argument for the Attract-based approach to movement, as was discussed in Section 6.

In the next chapter, I explore further what underlies the phenomenon of Interpretive Superiority, the distribution of Pair-list and Single-pair readings, and the semantics of multiple interrogatives in general.

Chapter 3: Semantics of Multiple Interrogatives

1. The Phenomena

The analysis in the previous chapter relies on the syntactic effect of T-to-C movement on the derivation of multiple interrogatives. In addition, it relies on the semantics properties of multiple interrogatives cross-linguistically. In this chapter, I explore these semantic properties more closely.

Our goal will be to account for the following phenomena. First, recall from the previous chapter the cross-linguistics distribution of the Pair-list (PL) and Single-pair (SP) readings in multiple interrogatives.[29] I summarize these facts below. The data in (63) and (64) suggest that the PL reading is freely available in mono-clausal multiple questions across languages (except in the context of Interpretive Superiority, which we will turn to shortly). On the other hand, the SP reading is more limited in its distribution. The examples in (63) are from languages that disallow the SP reading in this context: English, Bulgarian, Russian, and Brazilian Portuguese. However, this reading is freely available in languages like Serbo-Croatian, Japanese, and Icelandic, as demonstrated in (64).[30]

(63) a. *PL/*SP*

Who invited who to the dinner? *English*

[29] I will use the term 'interrogative' and 'question' interchangeably to refer to syntactic objects. When referring to the corresponding semantic representations, I will use the term 'denotation of'.

[30] See Section 4 of Chapter 2 for the example scenarios for each reading.

b. *PL/*SP*

Koj kogo e pokanil na večerjata? *Bulgarian*
who whom Aux invited to dinner
'Who invited who to the dinner?'

c. *PL/*SP*

Kto kogo priglasil na užin? *Russian*
who whom invited to dinner
'Who invited who to the dinner?'

d. *PL/*SP*

Quem convidou quem para (o) jantar? *Brazilian Portuguese*
who invited whom to (the) dinner
'Who invited who to (the) dinner?'

(64) a. *PL/SP*

Ko je koga pozvao na večeru? *Serbo-Croatian*
who aux whom invited to dinner
'Who invited who to the dinner?

b. *PL/SP*

Dare-ga dare-o syokuzi-ni manekimasita-ka? *Japanese*
who-Nom who-Acc dinner-Dat invited-Q
'Who invited who to the dinner?

c. *PL/SP*

Hver bauð hverjum í veisluna? *Icelandic*
who invited whom in the-dinner
'Who invited who to the dinner?'

To capture this distribution will be one of the goals of this chapter.

Another phenomenon in need of explanation is Interpretive Superiority. Recall again from the previous chapter that, in languages that allow SP readings, fronting the lower wh-phrase over the higher one forces the SP reading, eliminating the PL reading, as demonstrated below.

(65) a. **PL/SP*

Nanio$_1$ darega t_1 katta no? *Japanese*

what$_{ACC}$ who$_{NOM}$ bought Q

'Who bought what?'

b. **PL/SP*

Šta$_1$ je ko kupio t_1? *Serbo-Croatian*

what is who bought

'Who bought what?'

Questions with complex wh-phrases present another puzzle. In languages that lack SP readings in questions with bare wh-phrases, both SP and PL readings are available in questions with complex wh-phrases, as in (66) from English and Russian.

(66) a. *PL/SP*

Which diplomat invited which journalist to the dinner?

b. Kakoj diplomat kakogo žurnalista priglasil na užin? *Russian*

which diplomat which journalist invited to dinner

'Which diplomat invited which journalist to the dinner?'

Russian allows for the second complex wh-phrase to optionally remain in situ. The judgments with respect to PL/SP readings remain the same in this case.[31]

When the lower complex wh-phrase is fronted over the higher one, the Interpretive Superiority effect does not arise. That is, both PL and SP readings remain available, as shown below.[32]

(67) a. *PL/SP*

Which journalist did which diplomat invite to the dinner?

b. Kakogo žurnalista kakoj diplomat priglasil na užin? *Russian*
which journalist which diplomat invited to dinner
'Which journalist did which diplomat invite to the dinner?'

Another context in which SP readings show up (and even seem to be forced) is when the wh-phrases are separated from their scope position by an island boundary, as in (68) with the *if*-clause, and in (69) with a relative clause. This holds for other islands as well. This locality phenomenon with respect to interpretation was observed in Mahajan (1990), Dayal (1996, 2002), Hagstrom (1998), and Aoun and Li (2003). Given the asymmetry between the bare and complex wh-phrases, discussed above, I provide the examples with both types of wh-phrases. In this case, there seems to be no asymmetry.

[31] The reason I refer to the wh-phrases in these examples as complex rather than D-linked is that the same facts hold of questions with wh-phrases of the type 'whose NP', 'what NP', 'what kind of NP', which are not D-linked. I assume Pesetsky (1987) notion of D-linking, where the exact individuals in the set denoted by a D-linked wh-phrase are known to both the speaker and the addressee.

[32] See Barss (2000) for a judgment different from that of my informants. He reports the PL reading to be unavailable in (67a). Even if the speakers might vary on this, we need to explain the judgment of the speakers for whom the PL reading survives such contexts (contrary to Interpretive Superiority).

(68) a. Which linguist will be offended if we invite which philosopher? **PL/SP*
b. Who will be offended if we invite who? **PL/SP*
c. Who will be offended if we break what? **PL/SP*

(69) a. Which student read the book that which professor wrote? **PL/SP*
b. Who read the book that who wrote? **PL/SP*
c. Who read the book that describes what? **PL/SP*

This is the opposite of what we saw in the mono-clausal contexts. Recall that English interrogatives of the type *Who bought what?* have PL and no SP reading. The island boundary seems to switch the two readings. Howard Lasnik (p.c.) points out that the definiteness effect of *the book* might be a factor here in the examples in (69). Without the definiteness effect, as in (70), my informants' judgments diverged: some could get the PL reading and some could not.

(70) a. Which student read a book that which professor wrote? **PL/SP*
b. Who read a book that who wrote? **PL/SP*
c. Who read a book that describes what? **PL/SP*

The facts in (68) and (69) are even less clear if we consider a scenario in (71) for the question in (68b), provided by Norbert Hornstein (p.c.). All my informants are capable of obtaining the PL reading in this case.

(71) We know that some Americans dislike some Russians and vice versa, but we need to invite the representatives of both embassies to certain event. So who will be offended if we invite who?

See Hagstrom (1998) for another scenario that can also bring out the PL reading across strong islands, provided by Noam Chomsky in p.c. with Paul Hagstrom.

Thus, in this chapter, I will discuss how my system would handle the facts as they are reported in (68) and (69), although it is important to keep in mind the controversial status of these facts.

A more established fact is that PL readings are available across a wh-island, the fact known since Baker (1970). Although not all English speakers get the matrix reading of *what* in (72), those who do get it, prefer the PL reading for the resulting multiple question, with the expected answer listing the pairs of wonderers and the things John bought, as in (73). Whether, the SP reading is available in this context is unclear, at least to my informants.

(72) Who wonders where John bought what?
(73) Mary wonders where John bought a car, Sue wonders where John bought a cat…

To summarize, the following phenomena are in need of explanation: (i) cross-linguistic variation with respect to the PL/SP readings distribution; (ii) Interpretive Superiority; (iii) the availability of SP readings and the lack of Interpretive Superiority effects in questions with complex wh-phrases; and (iv) the lack of PL readings across an island, unless it is a wh-island. We begin by taking a look at how compositional semantics of wh-questions works in general.

2. Covert wh-movement and semantics of questions

2.1. Interpreting wh-in-situ

Semantics of multiple questions largely depends on how wh-in-situ is interpreted. Let us begin by examining how single wh-questions are treated semantically.

Semantics of single interrogatives has been studied since Hamblin (1958). The first formal compositional semantic analysis was proposed in Hamblin (1973). Unlike the semantic value of a statement, the semantic value of a question cannot be a truth value. That is, unlike the utterance in (74a), the utterance in (74b) does not have a truth value. That is, it cannot be true or false. Rather it is a request for a statement like that in (74a).

(74) a. John left.
b. Who left?
c. {John left, Mary left, Bill left…}

Thus, Hamblin (1973) proposed that the semantic value of a question is a set of propositions which constitute all its possible answers. On this analysis, the denotation of the question in is represented as the set of propositions in (74c). A true proposition from this set is the answer to the question, providing the value for the wh-phrase.[33] The value of the wh-expression *who* proliferates through the propositions. That is why Hamblin treats wh-phrases as sets of individuals. The formal denotation of the question in (74b) is given in (75).

[33] Kartunen (1977) argues that the set of propositions denoting a question contains only the propositions that are true.

(75) λp ∃x [person(x) & p = left(x)]

Now let us see how this formula is obtained compositionally. What is the structure of (74b) that would give us the formula in (75)? More specifically, we are interested in where the wh-phrase is interpreted in the structure and where the shift from propositions to a set of propositions takes place. To see the location of the wh-phrase more clearly let us use a question with a non-subject wh-phrase. And let us use a which-phrase to obtain a more prominent restrictor.

(76) a. Which book did [$_{TP}$ John read *t*]?

b. λp ∃x [book(x) & p = read(John, x)]

It is clear from Hamblin's denotation of (76a) in (76b) that the wh-phrase is interpreted in its moved position: the wh-existential operator and the restrictor are outside the scope of the proposition variable *p* over which λ-abstraction takes place. The propositional variable is presumably introduced in C^0. It is λ-abstraction over this variable that shifts the denotation from a proposition to a set of propositions. The trace of the wh-phrase is interpreted as a variable ranging over individuals.

Note that this wh-phrase is interpreted just where it is pronounced, in the left periphery of the clause, so the mapping from overt syntax to semantics is direct. However, in multiple questions in a language like English only one wh-phrase is pronounced in that position, the other wh-phrase(s) are pronounced in situ, as in (77a). Moreover, in languages like Chinese and Japanese wh-phrases are pronounced in situ even in single wh-questions, as in the Japanese example in (77b). If we want to keep the direct mapping between syntax and semantics, how do we obtain the

denotation in (76b) for the Japanese question that has that meaning, yet has no overt wh-movement? And how do we interpret multiple questions in English where all but one wh-phrase are in-situ?

(77) a. Which student read which book (on which day)?

b. John-wa [dono hon]-o yon-da no? *Japanese*
John-Top which book-Acc read-past Q
'Which book did John read?'

This is where covert wh-movement comes in. If wh-phrases that are pronounced in-situ actually undergo covert movement to the left periphery of the clause, as in (78), the straightforward mapping is maintained. (78a) would have the denotation in (76b) and (78b) would have the denotation in (79).

(78) a. Which student which book [$_{TP}$ *t* read *t*]?

b. [dono hon]-o [$_{TP}$ John-wa *t* yon-da] no? *Japanese*
which book-Acc John-Top read-past Q
'Which book did John read?'

(79) $\lambda p \exists x \exists y$ [student(x) & book(y) & p = read(x,y)]

However, there are alternative ways to interpret wh-in-situ, which do not require covert wh-movement. One approach is that of Unselective Binding, as in Baker (1970), Pesetsky (1987) and Nishigauchi (1986, 1990), where a wh-phrase comes with a variable that is bound by the interrogative operator in C^0. This provides

an insightful parallel between wh-phrases and other existentials, since exactly the same treatment was proposed for indefinites in Lewis (1975).[34] On this approach, the structure of a multiple interrogative is as in (80a) and the denotation of a multiple interrogative is as in (80b).

(80) a. Which student$_1$ $C_{1,2}$ [$_{TP}$ t_1 read which book$_2$]?

b. $\lambda p \exists <x,y> [student(x) \& p = read(x,y) \& book(y)]$

Observe that on this analysis, the restrictor of *which book* is inside the scope of the propositional variable. Reinhart (1995, 1997) points out a problem this causes in cases where the wh-restrictor is inside an *if*-clause, as in (81a).

(81) a. Who will be offended [if we invite which philosopher]?

b. For which <x, y>, if we invite y and y is a philosopher, then x will be offended.

c. $\lambda p \exists <x,y> [p = [we\ invite(y) \& philosopher(y)] \rightarrow offended(x)]$

d. Lucie will be offended if we invite Donald Duck.

The meaning of (81a) with *which philosopher* interpreted in situ is as in (81b), which is formally stated in (81c). Then a possible answer to (81a) should be (81d): even though Donald Duck is not a philosopher, he satisfies the truth conditions of (81b) because all it says that, if he were a philosopher and we invited him, Lucie would be offended. Thus the truth conditions in (81b) are too weak. Notice that covert wh-movement does not face this problem because the wh-restrictor is interpreted in the target position, outside the *if*-clause.

[34] Similar analyses of indefinites can also be found in Kamp (1981), Heim (1982), and Kratzer (1998).

2.2. Choice Functions

To solve the *if*-clause problem, Reinhart (1995, 1997) argues for the Choice Function treatment of wh-in-situ, which still does not rely on covert wh-movement. She proposes that wh-phrases (as well as all other indefinites) do not introduce a variable ranging over individuals but rather denote a set of individuals (as on the original Hamblin's approach) and a choice function variable that applies to that set. The choice function variable is bound by the question operator and this operator also binds the trace variable left by the fronted wh-phrase, which moves overtly.

Now the meaning of the conditional sentence in (81a) is as in (82a), formally stated in (82b). Because the choice function selects a value from the set of philosophers, the values can only be from that set. Donald Duck is not in that set, so the proposition in the answer cannot be true of him.

(82) a. For which <x,f>, if we invite f(philosopher), x will be offended.

b. $\lambda p \exists <x,f>$ [p = [we invite f(philosopher)] $\rightarrow$ offended(x)]

The denotation of the multiple interrogative in (83a) can now be formulated as in (83b).

(83) a. Which student read which book?

b. $\lambda p \exists <x,f>$ [student(x) & p = x read f(book)]

Reinhart (1995, 1997) treats only the wh-in-situ with choice functions. However, under the Copy Theory of movement, nothing seems to prevent us from treating even the moved wh-phrases in that way. In fact, it seems ad hoc **not** to treat

the moved wh-phrases that way, once the mechanism of choice functions is made available in the system. A given wh-phrase would then be interpreted with the same mechanism no matter where it occurs in syntactic structure. Under the Copy Theory of movement, this would result in an algorithm for which copy of a moved wh-phrase should be interpreted by semantics. If a wh-phrase overtly moves all the way to SpecCP, we would have to interpret its lower and not the higher copy, since only in the lower position would the choice function variable be bound by a quantifier in C^0. This is under the assumption that heads cannot bind into their specifiers. On the other hand, if a wh-phrase moves to a position lower than SpecCP, both interpretive possibilities should be available.[35] Given that even moved wh-phrases can be interpreted in situ and no different mechanism is needed for those, the fact that there is overt wh-movement in language is likely to be a purely syntactic phenomenon. One possibility is that it happens for clausal typing reasons, as in Cheng (1991). Clausal typing actually seems pragmatic in that it fulfills a pragmatic function of marking a clause as interrogative. But this general function might have evolved as encoded syntactically in the familiar feature-checking mechanism.[36]

Thus, choice functions allow for the wh-in-situ to be interpreted without covert movement and without the problem in the context of conditionals. In addition, this maintains the direct mapping between syntactic structure and semantics. Given that the choice function treatment is available, and since it provides a unified analysis

[35] We will duscuss the interpretation of the copies of the moved wh-phrases in more detail in Sections 5 and 6 of this chapter.

[36] Thanks to Howard Lasnik (p.c.), who pointed out this evolutionary option to me. See also Chomsky (2000) for a similar idea about the nature of EPP.

of all indefinites, it leaves no motivation for covert wh-movement on semantic grounds. Hence, if covert movement exists, it is motivated entirely by some formal (and not semantic) requirement.

Although, we have struggled to find semantic evidence for covert movement, perhaps there is syntactic evidence for such movement. We will examine the availability of such evidence next.

2.3. Syntactic evidence for/against covert wh-movement

Since movement is a syntactic phenomenon, there maybe some syntactic evidence that covert wh-movement exists. The best kind of evidence would be if covert wh-movement were to share the properties of overt wh-movement. One of the properties of overt wh-movement is that it obeys island constraints, a well established generalization since Ross (1967). Applying this diagnostic for movement to wh-in-situ, however, brings mixed results. For the most part, wh-phrases in situ do not seem to observe syntactic islands. We have already encountered many instances of this in this chapter. Recall the data in (68) – (72), repeated below.

(84) a. Which linguist will be offended if we invite which philosopher? **PL/SP*
b. Who will be offended if we invite who? **PL/SP*
c. Who will be offended if we break what? **PL/SP*

(85) a. Which student read the book that which professor wrote? **PL/SP*
b. Who read the book that who wrote? **PL/SP*
c. Who read the book that describes what? **PL/SP*

(86) Who wonders where John bought what?

Putting aside the potential unavailability of PL readings in this context (remember that it is not clear that this is true), the sentences are acceptable. Compare the acceptability of these examples with the instances of overt wh-movement out of islands below.

(87) a. *Who will John be offended if we invite *t*?
b. *Who did John read the book that *t* wrote?
c. ??What does John wonder where Mary bought *t*?

It is also well known from the work of Huang (1982), Lasnik and Saito (1992), and Watanabe (1992), among many others, that argument wh-phrases in wh-in-situ languages like Chinese and Japanese can appear inside islands without producing unacceptability.[37] There are still ways to keep covert movement in the system and deal with its crucial asymmetry with overt movement beyond merely stipulating it, as for example in Nishigauchi (1986), Lasnik and Saito (1992), and Richards (1997). However, the fact remains that something extra has to be said in order to reconcile covert wh-movement with overt movement. Note that the problematic data just discussed are not at all problematic if wh-in-situ remains in situ and is interpreted in that position. In fact, these data are precisely as predicted on that approach.

[37] The Japanese facts are less clear than those from Chinese (cf. Nishigauchi 1986), but the overall generalization that wh-in-situ does not exhibit island effects nearly to the same extent as overtly moved wh-phrases holds in both languages.

However, there is some apparent evidence for covert wh-movement when it comes to wh-adjuncts. Adjunct wh-phrases in wh-in-situ languages, unlike argument wh-phrases, seem to obey island constraints, as was first observed by Huang (1982) for Chinese. It is also a well known fact that English adjunct wh-phrases *why* and *how* are even more restricted in their distribution: they cannot appear in situ at all.

(88) a. *Who left why?
b. *Who fixed the car how?

Both Chinese and English facts have been captured by the Empty Category Principle (ECP) in Huang (1982) and Lasnik and Saito (1992). ECP requires a trace to be properly governed. This can be obtained by lexically government or antecedent government, which is essentially like binding only subject to Subjacency.[38] In both contexts above, the traces of adjuncts are not properly governed, resulting in an ECP violation. In (88), this is because covert wh-movement of an adjunct adjoins it to the wh-phrase in SpecCP, from where it cannot c-command its trace, hence cannot antecedent govern it; and lexical government is not possible for adjuncts to start with.

However, with the elimination of government in Minimalism, ECP is eliminated as well. Of course, it can be restated, since government in the antecedent government and in the lexical government was not a uniform notion anyway. But that would remain a restatement of the actual problem.

What does the evidence from adjuncts tell us about the existence of covert wh-movement? While the behavior of Chinese wh-adjuncts suggests that they might

[38] Roughly speaking, Subjacency requires each movement step to be local, ensuring that the island boundaries are not crossed by movement. See Chomsky (1973) and Chomsky (1986) for more details.

be moving covertly since they obey island constraints, just like in overt wh-movement, the behavior of English wh-adjuncts pulls the evidence in the opposite direction. Movement within a single clause is the least overt wh-movement is capable of, as shown below, while the hypothesized covert wh-movement is not capable of it.

(89) a. What did John fix?
b. Why did John leave?
c. How did John fix the car?

Thus, we can conclude that the evidence for covert wh-movement seems to come from wh-adjuncts in Chinese. At the same time, as was discussed earlier, the semantics of wh-questions seems to favor choice functions over covert wh-movement. Hence, for now, it seems plausible that covert wh-movement exists, but solely for formal reasons. It does not affect the interpretation of questions. The semantics of questions can be computed using choice functions throughout.

3. Semantics of PL and SP readings

3.1. Previous accounts

The analysis I will develop to account for the facts presented in Section 1 is based on the analysis of PL and SP readings in Hagstrom (1998). Before we proceed with that, let me point out the other accounts of PL readings (there are virtually no accounts of SP readings as independent readings), indicating their main problems.

Besides Hagstrom (1998), there are several approaches to PL readings and most of them are developed on the basis of those readings in single wh-interrogatives

with a universal quantifier, as in (90a). Those are the approaches of Karttunen (1977) involving QR of a universal quantifier and quantification into questions; Groenendijk and Stokhof (1984), involving quantification into questions and a special interpretation of quantifiers in this particular context, referring to witness sets; Engdahl (1985) and Chierchia (1993), assimilating PL readings to functional readings, which are available only in the context of certain non-wh-quantifiers; and Krifka (2001), involving quantification into question acts.[39] Higginbotham and May (1981) examine multiple wh-questions and propose an operation Absorption, which, after covert wh-movement, turns two quantifiers into a single quantifier with the property of producing PL readings. Although, since then, Absorption was often adopted in the literature, the details of this operation were never provided.

(90) a. Which book did everyone buy?	*Indiv/Func/PL/*SP*
b. Which student bought which book?	**Indiv/*Func/PL/SP*
c. Who bought what?	**Indiv/*Func/PL/*SP*

For our purposes, providing a theory for questions with the universal quantifier may not be the best way to start because a question with a single wh-phrase and a non-wh-quantifier in (90a) does not have a SP reading, which multiple wh-questions sometimes have, as in (90b). Recall that not all multiple questions have SP readings, as shown by the contrast between (90b) and (90c). On the other hand, (90a) has "extra" readings that multiple interrogatives do not have (i.e., the individual and

[39] I will not go into the details of each of these approaches here since that will take us considerably off track, but refer the reader to the overviews in Szabolcsi (1997), Pafel (1999), Krifka (2001), and Dayal (2005).

functional readings). Although the latter asymmetry, perhaps, can be derived from the distinction between a universal quantifier and a wh-phrase, on all of these approaches, something extra has to be formally done to distinguish PL and SP readings in multiple interrogatives. Besides, there seems to be no way to capture the cross-linguistic distribution of the SP reading on these approaches. All of them predict that whenever PL reading is available, SP reading is available also, contrary to the facts in Section 1 of this chapter.

In addition, the approaches above involve QR, which translates into covert wh-movement if one tries to extend these approaches to multiple wh-questions, as in the work of Hornstein (1995), Comorovski (1996), and Dayal (1996, 2002), who extend the approach of Engdahl (1985) and Chierchia (1993). These approaches inherit the property of Engdahl (1985)'s and Chierchia (1993)'s system in not being able to capture the cross-linguistic distribution of SP readings.

All this seems to suggest that an alternative approach is needed, which distinguishes the readings in single wh-questions with a universal quantifier and in multiple interrogatives, cross-linguistically. This means that one of the analyses pointed out in this section might very well be the right analysis for the wh-phrase and universal quantifier interaction. However, it should not be overgeneralized. In the next section, I develop an alternative analysis, based on the work of Hagstrom (1998), capturing the underlying syntax and semantics of PL/SP readings in multiple interrogatives, their distribution cross-linguistically, in the context of complex wh-phrases, and in the context of islands.

3.2. PL readings as sets of questions

To understand nature of the PL and SP readings, I begin with the syntactic and semantic analysis of these readings developed by Hagstrom (1998). He proposes that wh-interrogatives with the PL reading denote a *set* of questions (i.e., a set of sets of propositions, of the type <pt,t>, where *p* stands for the semantic type of a proposition, <st>). The intuition is that a question in (91a) has the meaning of a set of questions in (91b), where each question is asking about the object bought by each individual from the set of individuals denoted by the higher wh-phrase. If the domain of individuals denoted by *who* in (91a) contained only three individuals John, Mary and Sue, there would be three questions in the set, as in (91b).

(91) a. Who bought what?
b. {What did John buy? What did Mary buy? What did Sue buy?}
c. What did John buy, what did Mary buy, and what did Sue buy?

A similar idea was also put forward by Krifka (2001). Only, instead of formalizing the PL reading a set of questions, he treats it a series of conjoined questions, where each question is a separate speech act. So technically, those are conjoined speech acts on Krifka's analysis. He focuses on single wh-questions with non-wh-quantifiers like *What dish did every guest make?* and explains many puzzling facts about those. For example, he explains why PL readings are unavailable with most of the quantifiers like *no, most, a few*, etc., as shown in (92).

(92) What dish did no/most/a few guests make? *Indiv/Func/*PL*

He argues that this is because the meaning of those quantifiers involves Boolean disjunction in one form or another, and only *every* involves Boolean conjunction; and one needs conjunction for a series of questions.[40]

The specific details of this analysis are difficult to extend to multiple wh-questions for the same reasons described in the previous section. However, the gist of the idea is very similar to the one we find in Hagstrom's analysis and we turn to Hagstrom's implementation of it next.

3.3. Syntactic and semantic contribution of the Q-morpheme

Hagstrom (1998) explores the syntactic and semantic contribution of the interrogative morpheme (Q-morpheme) to the derivation of interrogatives and concludes that it plays an important role in the derivation of the PL and SP readings.

In languages like Chinese and Japanese, there is an overt Q-morpheme. It is reasonable to assume that in languages like English, there is a phonetically null Q-morpheme as well.[41] Hagstrom examines the syntactic properties of the Q-morpheme in Japanese, Sinhala and Okinawan and concludes that this morpheme undergoes syntactic movement to C^0 from a clause internal position. The proposal is in the spirit

[40] Notice that although the PL reading is unavailable in this example, the functional reading remains available. This is one of the arguments Krifka provides for formally distinguishing functional and PL readings.

[41] Hagstrom (1998)'s account is based primarily on wh-in-situ languages, so my points about extending it to other languages may not reflect Hagstrom's view. See also Bošković (2001), who also extends this approach to non-wh-in-situ languages.

of the overt null-operator movement of Watanabe (1992). Hagstrom presents a lot of evidence for this proposal. Let me summarize some of it here.

The Q-morpheme in Sinhala, a right-branching language, can occur overtly inside a clause, next to the (lowest) wh-phrase. The affix on the verb, glossed as *E* marks the scope of the question. The Q-morpheme in this language can occur in an embedded clause in a long-distance question, as in (93a), but not when the wh-phrase is inside an island, as shown by the unacceptability of the Q-morpheme inside a relative clause in (93b). In this case, the Q-morpheme appears outside the island, as in (93c).

(93) a. Ranjit [kau **də** aawa kiyəla] dann**e**? *Sinhala*

Ranjit who Q came that know-E

'Who does Ranjit know that came?' (Kishimoto 1997:6)

b. *Oyaa [kau də liyəpu potə] kieuwe?

you who Q wrote book read-E

'Who did you read the book that wrote?'

c. Oyaa [kauru liyəpu potə] də kieuwe?

you who wrote book Q read-E

'Who did you read the book that wrote?' (Kishimoto 1992:56)

Sensitivity to islands is a property of movement and therefore a standard diagnostic for it. Hence, the Q-morpheme in Sinhala must be undergoing movement from the clause-internal position to C^0. Hagstrom suggests that it is covert head-movement of the Q-morpheme *də*.

Another piece of evidence for the movement of the Q-morpheme comes from the intervention effects exhibited by this morpheme in Japanese. Japanese *–ka* is not only a Q-morpheme, but also occurs as part of an indefinite *dare-ka* 'someone', as a disjoiner *–ka* 'or', and as part of the operator *kadooka* 'whether'. Hagstrom presents data demonstrating that elements containing –ka cannot intervene between the Q-morpheme and the wh-phrase in its scope. The paradigm in (94) shows this with the disjunctive *–ka*, and the parallel paradigm in (95) involves the indefinite *dareka*. In the (b) and (c) examples in both paradigms, the wh-phrase is higher than the potential intervener, making the sentences acceptable. The examples are from Hoji (1985).

(94) a. ??[John-ka Bill]-ga nani-o nimimasita ka? *Japanese*
John-or Bill what-ACC drank Q
'What did John or Bill drink?'

b. Nani-o_1 [John-ka Bill]-ga t_1 nimimasita ka?
what John-or Bill-NOM drank Q
'What did John or Bill drink?'

c. Darega [sake-ka biiru(ka)]-o nomimasita ka?
who-NOM sake-or beer(or)-ACC drank Q
'Who drank either sake or beer?'

(95) a. ??Dareka-ga nani-o niomimasita ka?
someone-NOM what-ACC drank Q
'What did someone drink?'

b. Nani-o_1 dareka-ga t_1 niomimasita ka?
what-ACC someone-NOM drank Q
'What did someone drink?'

c. Dare-ga nanika-o niomimasita ka?
who-NOM something-ACC drank Q
'Who drank something?'

These intervention effects also hold across a clausal boundary, as shown in (96); and embedding the intervener inside a larger constituent saves the structure, as in (97), indicating that these intervention effects are truly structural and not simply linear.[42]

(96) ?? [John-ka Bill-ga] [Mary-ga nani-o katta to] itta no?
John-or Bill-NOM Mary-NOM what-ACC bought that] said Q
'What did John or Bill say that Mary bought?'

(97) [[John-ka Bill-ga] atta hito]-ga nani-o motte kita no?
John-or Bill-NOM met person-NOM what-ACC brought Q
'What did the man John or Bill met bring?'

Based on these facts, Hagstrom (1998) argues that, if Japanese Q-morpheme is base-generated at the same position as the Sinhala Q-morpheme (next to the wh-phrase)

[42] This examples in (96) and (97) are attributed by Hagstrom (1998) to Shigeru Miyagawa (p.c. with Paul Hagstrom).

and undergoes overt movement to C^0, these intervention facts can be straightforwardly explained as minimality effects.

Although we find intervention effects with respect to *-ka*, it does not seem to be island sensitive. It is well known that wh-phrases in Japanese can appear inside certain islands (cf. Nishigauchi 1986, Pesetsky 1987, Lasnik and Saito 1992, Watanabe 1992, among others). This is demonstrated for the Complex Noun Phrase Constraint and the adjunct island below.

(98) a. Mary-wa [John-ni nani-o ageta hito-ni] atta no?
Mary-TOP John-DAT what-ACC gave man-DAT met Q
'Mary met the man who gave what to John?'

b. Mary-wa [John-ga nani-o yomu mae-ni] dekaketa no?
Mary-TOP John-NOM what-ACC read before left Q
'Mary left before John read what?' *(Pesetsky 1987:110)*

To explain this seemingly contradictory behavior of the Q-morpheme, Hagstrom once again draws the parallel with the Q-morpheme in Sinhala. Recall that in the context of an island, the Sinhala Q-morpheme appears overtly just outside the island and not inside, (93c). If the Japanese Q-morpheme can move overtly from that alternative position to C^0 (which happens covertly in Sinhala), then the lack of island effects in Japanese is expected.

Hagstrom presents further evidence in support of this analysis based on the fact that the familiar interveners for the movement of the Q-morpheme are ineffective when they are inside islands, as in (99a). The intervention effect comes back if the intervener is outside the island, as in (99b).

(99) a. Mary-wa [[John-**ka** Bill]-ga **nani-**o katta ato de] dekaketa **no**?
Mary-TOP John-or Bill-NOM what bought after left Q
'Mary left after John or Bill bought what?

b. ??[John-**ka** Bill]-wa Mary-ga **nani-**o katta ato de] dekaketa **no**?
John-or Bill-NOM Mary-NOM what-ACC bought after left Q
'John or Bill left after Mary bought what?'

Hagstrom (1998) suggests that the Q-morpheme moves from the position inside an island to the position just outside it. However, since this movement would have to be insensitive to islands and interveners, it seems plausible that the position outside an island is the alternative position where the Q-morpheme can actually be base-generated.

Thus, there are two different positions in which the Q-morpheme can originate: next to a wh-phrase or in some position outside an island. Hagstrom shows that in multiple questions in Sinhala, the Q-morpheme overtly appears next to the lowest wh-phrase and not the higher one.

(100) a. Kauru mokak də kieuwe? *Sinhala*
who what Q read-E
'Who read what?'
b. *Kau də mokak kieuwe?
who Q what read-E
'Who read what?'

Based on this fact, Hagstrom generalizes that this is a general property of the Q-morpheme. I will return to this issue after discussing the Hagstrom's compositional

semantics of multiple questions and argue that selecting the lower wh-phrase is not a general property of the Q-morpheme and that the Q-morpheme can select any wh-phrase in any position. But for now, let us assume along with Hagstrom, that the Sinhala pattern reflects the general selectional specification of the Q-morpheme. According to Hagstrom, the other selectional option for the Q-morpheme is merging with TP. That is the option we saw used when there was an island. This is illustrated in (101).

(101) a. $[_{CP}\ \mathbf{Q}_j\text{-}C^0 \ldots[_{TP} \ldots \text{WH1} \ldots V \ldots\ [\boldsymbol{t_j}\text{-WH2}]]]$ *PL*

b. $[_{CP}\ \mathbf{Q}_j\text{-}C^0 \ldots[_{QP}\ \boldsymbol{t_j}\ [_{TP} \ldots \text{WH1} \ldots V \ldots \text{WH2}]]]$ *SP*

Hagstrom further proposes that the PL and SP readings are derived from these two structural possibilities: the PL reading arises when the Q-morpheme selects the lower wh-phrase, as in (101a), and the SP reading is a result of selecting TP, as in (101b).[43] The Q-morpheme in both cases moves to the interrogative C^0, where it checks the $+Q$ feature of C^0. Hagstrom proposes that the Q-morpheme is interpreted as a quantifier over choice functions and a wh-expression denotes merely a set of individuals, as on Hamblin (1973) original proposal. It is the Q-morpheme that provides the choice function variable in the position it originates and then closes that variable from the position it moves to (C^0). This approach seems to not only consistent with the Choice Functions approach of Reinhart (1995, 1997), but it also

[43] Hagstrom (1998) actually hypothesizes some head F^0 above TP as the place of generating the Q-morpheme. However, there seems to be nothing wrong with the Q-morpheme merging directly with TP, as I do in (101b).

rationalizes it and provides additional support for it by identifying a specific morpheme responsible for contributing the choice function variable.

3.4. Compositional semantics for PL and SP readings

Recall that, on the Hagstrom (1998) approach, a multiple wh-question with the SP reading denotes a single set of propositions (i.e., the semantic type <pt>), while a multiple wh-question with the PL reading denotes a set of single questions (i.e., a set of sets of propositions; the semantic type <pt,t>). Let us consider how the PL and SP readings of a simple question like *Who bought what?* are derived compositionally. The LF representations of the two readings are given below.[44]

(102) a. $[_{CP}\ \mathbf{Q}_j\text{-}C^0\ [_{TP}$ who bought $[\boldsymbol{t_j}$ -what]]] *PL*

b. $[_{CP}\ \mathbf{Q}_j\text{-}C^0\ [_{QP}\ \boldsymbol{t_j}\ [_{TP}$ who bought what]]] *SP*

By movement to C^0, the Q-morpheme leaves behind a variable ($\boldsymbol{t_j}$) whose value ranges over generalized choice functions. A generalized choice function (type <αt,α>) basically picks a member out of the set it is merged with.

In the derivation of the PL reading, the choice function variable takes *what* (denoting a set of individuals) as its argument returning an individual (<e>). Further, applying the function denoted by the verb *bought* to this individual produces a property or, in other words, a function from individuals to truth values (<et>). In order to apply this function to the set of individuals denoted by *who*, Flexible Functional Application (FFA), defined as in Rullmann and Beck (1997), is needed.

[44] Although, English does not have the SP reading in this simple context, other languages do. I will discuss the nature of this parameterization in the next section.

FFA allows for a function that applies to an individual to apply to a set of individuals by applying to each member of that set and putting the result into a set.

Thus, the function denoted by the VP applies to every individual in the set denoted by *who* and the result is put into a set. This set is a set of propositions (i.e., <pt>, where *p* stands for the type of a single proposition <st>). The movement of the Q-morpheme to C^0 evokes λ-abstraction over this set, turning it into a set of propositions abstracted over choice functions (<cp,t>), where *c* stands for a choice function.[45] The function denoted by the complex head [Q-C^0], of type <cp,pt>, then applies to this set of unsaturated propositions via FFA, producing a set of sets of propositions <pt,t>.[46] Each set of propositions in this set represents a question about each individual in the set of individuals denoted by *who*, as in (103) below.

(103) {What did John buy?, What did Mary buy?, What did Sue buy?}

In the SP reading derivation, the choice function variable is not there to reduce the set of individuals denoted by *what* because the Q-morpheme moves from the position above TP. The denotation of the VP in this case is a set of properties <et,t>. Then the set of individuals denoted by *who* is taken as an argument via FFA, giving back a set of propositions (type <pt>), pairing each individual in the set of *who* with each property. The choice function variable then applies to this set and picks one of its members, a single proposition (type <p>). Via λ-abstraction over choice functions, we get an unsaturated proposition (<cp>). Combining it with the complex head [Q-

[45] This requires 'flexible-lambda-abstraction'. See Hagstrom (1998:169) for details.

[46] Internally to [Q-C^0], the denotation of C^0 takes the denotation of *Q* as an argument.

C^0] results in a set of propositions and, crucially, **not** a set of sets of propositions, as we saw in the PL derivation.

The major difference between the PL and SP derivations is that, in the PL derivation, there is no choice function variable immediately above TP, due to the Q-morpheme merging with the wh-phrase *what*. This allows the set of individuals denoted by *who* to propagate through the derivation, producing in the end a set of sets of propositions. However, this is not possible in the SP derivation due to the choice function variable above TP reducing the set of propositions denoted by TP to a single proposition. This proposition becomes the input to further computation, producing in the end just a set of propositions and not a set of sets of propositions.[47]

Hagstrom (1998) provides this semantics for languages where wh-phrases do not undergo overt movement (e.g., Japanese and Sinhala). Extending this analysis to the languages with overt wh-movement could be done as follows. One can interpret the variable left by wh-movement as an entity of type <e>, with further λ–abstraction over individuals, and interpret the wh-phrase in its target position. The set of individuals denoted by the moved wh-phrase will propagate through the function denoted by *C'*. In the PL reading derivation, *C'* will denote a function from individuals to a set of propositions. In the SP reading derivation it will be the function from individuals to a single proposition.[48]

---

[47] For more explicit formal details of the two derivations, see Hagstrom (1998):136-145.

[48] In the derivation with two wh-phrases inside a VP, like in a double object construction, in a language like Japanese, where the wh-phrases would remain unmoved, the question arises how the combination of an external argument of type <e> with the set of properties <et,t> (the denotation of the VP) procedes. I suggest that Flexible Function Application can be used here: each property in the set will be saturated by a single individual and the result will be put into a set, producing a set of propositions.

Note that, given this semantics, nothing actually changes if the Q-morpheme were to merge with either of the wh-phrases. That is, it does not necessarily have to merge with the lower wh-phrase, contrary to Hagstrom's proposal. There are three reasons to think that the Q-morpheme indeed simply selects a wh-phrase and does not differentiate between higher and lower wh-phrases.

First, it is unclear how one can implement the selectional restriction where a given lexical item selects some category only when this category is in a certain position.

Second, the truth-conditions come out correct no matter which wh-phrase the Q-morpheme merges with. The only difference between the two derivations is that, in the derivation where the Q-morpheme merges with the higher wh-phrase, it is the lower wh-phrase that propagates through the derivation, in the end producing a set of questions of the sort in (104a). Compare this with the set we got in the derivation where the Q-morpheme merged with the lower wh-phrase, repeated in (104b).

(104) a. {Who bought the cheese?, Who bought the wine?, Who bought the cake?}
b. {What did John buy?, What did Mary buy?, What did Sue buy?}

This might be relevant to the issue of exhaustivity in questions. Comorovski (1996) claims that, in an answer to a multiple wh-question, the set of individuals denoted by the higher wh-phrase must be exhausted, while the set denoted by the lower wh-

---

It will then be an input to further computation as described above. In a language with overt wh-movement, the issue does not even arise since the external argument will be combined with just a single property (type <et>), since wh-movement of one of the non-subject wh-phrases leaves a variable of type <e> inside the VP.

phrase does not have to be exhausted. However, Hornstein (1995) reports a different intuition, namely, that both sets must be exhausted. Given that the judgments are quite delicate and are in need of further study, it is difficult to conclude whether the difference between (104a) and (104b) is indicative of any detectable semantic effect or just vagueness.

Finally, there is empirical data from Navajo and Okinawan, presented in Hagstrom (1998), which seems to indicate that the Q-morpheme can indeed be merged with the higher as well as the lower wh-phrase.

(105) a. Háí-**lá** ha'át'íí nayiisnii? *Navajo*
who-**Q** what bought
'Who bought what?' (Barss et al. 1991:34)

b. Háí ha'át'íí-**lá** nayiisnii?
who what-**Q** bought
'Who bought what?' (Peggy Speas, p.c. with Paul Hagstrom)

(106) a. Taa-ga-**GA** nuu kam-ta-ra? *Okinawan*
who-NOM-**Q** what eat-past-M
'Who ate what?'

b. Taa-ga nuu-**GA** kam-ta-ra?
who-NOM what-**Q** eat-past-M
'Who ate what?' (~Sugahara 1996:246)[49]

[49] The morpheme glossed with *M* corresponds to the Sinhala morpheme glossed with *E*. These morphemes appear on the verbs in clauses where the Q-morpheme takes scope.

Hagstrom (1998) also reports that several consultants judge (106a) as requiring a PL reading, enumerating for each food, who ate that food. And this is not the case for (106b). That is just as our analysis above has predicted. The Okinawan-speakers are sensitive to the fact that, when the Q-morpheme merges with the higher wh-phrase, the question denotes a set of questions of the form in (104a) and not of the form in (104b).[50]

Hagstrom (1998) also proposes a theory how the speakers answer questions with the PL/SP readings. For the SP reading, it works in a familiar way: the speaker selects one proposition out of the set of propositions as the answer. For the PL reading, the task is a bit more complex because the speaker is confronted with a set of sets of propositions. Hagstrom suggests that the semantic value of this utterance (of type <pt,t>) allows the speaker to recognize it as pair-list question and respond by selecting one proposition from each member set of propositions. This seems quite plausible: a speaker must provide an answer to each question in the set, so he or she selects a true proposition from each set of propositions.[51]

We are now ready to explain the cross-linguistic variation with respect to the availability of the SP reading.

[50] However, see Section 6 for more discussion of this and some new evidence to the effect that, in some languages, the Q-morpheme still must merge with the lower and not the higher wh-phrase.

[51] For more details on this, see Hagstrom (1998:148).

4. Capturing cross-linguistic variation with respect to the SP reading

4.1. Bošković (2003)

On Hagstrom's approach, what licenses the SP reading is the presence of the Q-morpheme above TP. Bošković (2003) observes that SP readings seem to be unavailable in multiple wh-questions where overt syntactic wh-movement (i.e., movement of a wh-phrase to SpecCP to check the uninterpretable *+wh* feature on C^0) takes place.

Using Superiority effects as a diagnostic for syntactic wh-movement, Bošković identifies English, German and Bulgarian wh-questions as languages involving such movement.[52] On the other hand, all contexts in Japanese and main clauses with null C^0 in Serbo-Croatian do not involve wh-movement to SpecCP. In previous work, Bošković (1997a, 1998) argues that Serbo-Croatian involves covert C^0 insertion in this context and multiple wh-fronting is viewed as focus movement to a position lower than C^0, triggered by an attract-all *+focus* feature. Bošković (2003) concludes that it is in these contexts, that lack syntactic wh-movement to SpecCP, the SP readings are allowed freely.

Particularly, Bošković argues that syntactic wh-movement in a SP reading derivation creates a Relativized Minimality violation. That is, the movement of a wh-phrase to SpecCP in a language like English violates Relativized Minimality by crossing the Q-morpheme. Bošković suggests that the Q-morpheme carries a *+wh*

[52] German is actually known to not exhibit Superitrity effects, suggesting that Superiority is not a sufficient diagnostic for wh-movement to SpecCP.

feature, like the wh-interrogative C^0 and wh-phrases do. The proposed SP reading derivation of the question in (107) is given in (108).

(107) Who bought what? **SP/PL*

(108) *[$_{CP}$ *Who*$_j$ C^0 [$_{QP}$ **Q** [$_{TP}$ t_j …bought *what*]]]

The derivation in (108) creates a Relativized Minimality violation and that is why the SP reading is unavailable in English multiple questions. This is generalized to other languages with overt wh-movement to SpecCP, including German. It is also assumed here that, in a language with overt wh-movement, the wh-phrases are interpreted in the base-generated position and the Q-morpheme moves to C^0 covertly. If it moved overtly, it would be crossing the subject wh-phrase in the PL reading derivation and incorrectly ruling out the only reading a bare multiple interrogative has in a language like English. However, given our conclusion that the Q-morpheme can be merged with either of the wh-phrases, this problem goes away and the Q-morpheme could be moving overtly in English, as it does in Japanese.

In Grebenyova (2004), I point out that there might be a conceptual problem with this Relativized Minimality account and that this analysis is not sufficient to rule out SP readings in a language like Russian.

The conceptual problem has to do with the fact that the Q-morpheme carries a *+wh* feature. Since this feature never seems to be checked against another *+wh* feature, it must be an interpretable feature. However, it is not clear what it means for a Q-morpheme to have an interpretable *+wh* feature. Whatever the precise nature of

that feature might be, it must match a certain feature on a wh-phrase in order to be in competition with it; and that seems difficult to instantiate.

Moreover, if the Q-morpheme carries a *+wh* feature, this morpheme eventually ends up in C^0, it is not clear why it cannot check the strong *+wh* feature of C^0. Of course, that would take away the motivation for the wh-phrases to move in a language like English, producing ungrammatical results of the kind in (109).

(109) *Did John give who what?

The crash of the SP reading derivation would then seem to be rather a violation of Last Resort (i.e., a wh-phrase moves to SpecCP for no reason) and not a Relativized Minimality violation. However, this leaves us with (109) being acceptable under the PL reading, which is not the case.[53] Of course, the covertness of the Q-morpheme movement avoids this problem because a wh-phrase would always be attracted in overt syntax in English, before the Q-morpheme is attracted at LF. But this overt-covert distinction is rather difficult to implement, given that the Q-morpheme has the *+wh* feature at the time when C^0 attracts a *+wh* feature.

Besides these technical difficulties, there are some empirical limitations of the Relativized Minimality analysis. Below, I present data from Russian, Icelandic, and Serbo-Croatian, showing that the Relativized Minimality account is not sufficient to rule out SP readings in Russian and it incorrectly predicts the absence of SP readings

[53] This problem might be avoided though if we assume the exact specification of whether a feature can be checked in a head-head or a spec-head relation. See Bošković (2001b) for some arguments for the necessity of such specification.

in Icelandic and embedded clauses in Serbo-Croatian. First, consider the facts from Russian in (110).

(110) Kto kogo priglasil na užin? *PL/*SP* *Russian*
who whom invited to dinner
'Who invited who to dinner?'

According to all my informants and myself, only the PL reading is available in (110); the SP reading is disallowed.[54] The SP reading is also disallowed when the object wh-phrase is fronted over the subject wh-phrase, as in (111). This is a context where Interpretive Superiority effects are attested in a languages that allow SP readings, as was shown for Japanese and Serbo-Croatian in Chapter 2. In those languages, the PL reading disappears in this context and only the SP reading remains available. [55]

[54] These judgements contrast with those of Stepanov (1998), who reports that (i) can have a SP reading. The sentence in (i) has a potentially interfering factor in that Superiority effects emerge with *who/what* combination in Russian as shown in (ii), with other combinations of wh-phrases being insensitive to Superiority. This is important because Superiority effects are used as a diagnostic for syntactic wh-movement. Hence, I modified the questions and corresponding scenarios by using a *who/who* combination.

(i) Kto čto kupil? *Russian*
who what bought
'Who bought what?'

(ii) *Čto kto kupil?

[55] English d-linked wh-questions allow both PL and SP readings whether the object is fronted over the subject or not. I will discuss this issue in Section 6 of this chapter.

(111) Kogo kto priglasil na užin? *PL/*SP* Russian
whom who invited to dinner
'Who invited who to the dinner?'

I will explain how this fact is captured in my system in Section 5 of this chapter.

There is some additional evidence for the lack of SP readings in Russian. Multiple sluicing (i.e., sluicing with multiple remnants) in Russian depends on interpretative properties of multiple interrogatives in this language. The example in (112) is unacceptable in Russian.

(112) *Kto-to priglasil kogo-to na tanec, no ja ne znaju kto kogo.
someone invited someone to dance but I not know who whom
'Someone invited someone to a dance but I don't know who invited whom.'

The antecedent clause forces the SP reading in the embedded clause. The sentence is bad, as predicted if the SP reading is unavailable in Russian. The corresponding example in Serbo-Croatian is fine, as reported in Stjepanović (2003). This is not surprising since Serbo-Croatian allows SP readings.

(113) Neko je video nekog, ali ne znam ko koga. *Serbo-Croatian*
somebody is seen somebody but not know who whom
'Somebody saw someone, but I don't know who whom.'

Multiple sluicing is permitted in Russian if the antecedent imposes a PL reading in the sluice, as in (114). Recall that the PL reading is available in Russian.

(114) Každyj priglasil kogo-to na tanec, no ja ne znaju kto kogo
everyone invited someone to dance but I not know who whom
'Everyone invited someone to a dance but I don't know who invited who'

Thus, multiple sluicing provides a new diagnostic for the availability of certain readings in multiple questions across languages. Of course, this applies only to languages that allow multiple sluicing to start with.

Let us see how we can explain the lack of SP readings in Russian. Russian does not involve syntactic wh-movement to SpecCP, as argued extensively in Stepanov (1998), and Bošković (2002a) based on the fact that Russian does not exhibit Superiority effects in main or embedded clauses. On these analyses, Russian C^0 has a weak *+wh* feature and the wh-phrases undergo focus movement to some position lower than C^0. Thus, a separate explanation is needed for why the SP reading is unavailable in Russian, since there is no wh-movement to SpecCP in this language.

Željko Bošković (p.c.) suggests that the base-position of Q-morpheme in a SP reading structure in Russian might be lower than the target position of the focus movement of wh-phrases. In that case, fronting of wh-phrases will still cross the Q-morpheme. However, for this to hold, it can no longer be a *+wh* feature that triggers the Relativized Minimality violation since wh-phrases in Russian do not front in order to check the *+wh* feature of C^0 but rather they move in order to check the *+focus* feature. Hence, it is not clear why the Q-morpheme would intervene for the purposes of focus-triggered wh-fronting.

Taking the Relativized Minimality to be insensitive to features but rather sensitive to the Head/A/A' distinction, as in the original Rizzi's proposal, is another

option. However, then it is not clear why a head would intervene for the purposes of phrasal movement. The only way to maintain the Relativized Minimality account of Russian is by assuming that there is no distinction between heads and XPs, which is made possible in Bare Phrase Structure.[56] Thus, Russian turns out to fit into the generalization about the interaction of overt wh-movement and the availability of the SP reading quite well.

Let explore this generalization further by considering the data from Icelandic. Recall the Icelandic paradigm from Chapter 2 showing that Icelandic multiple interrogatives allow SP readings in both main and embedded clauses. It is repeated below.

(115) a. *PL/SP*

Hver bauð hverjum í veisluna? *Icelandic*
who invited whom in the-dinner
'Who invited who to the dinner?'

b. ?PL/SP

Hverjum bauð hver í veisluna?
whom invited who in the-dinner
'Who invited who to the dinner?'

(116) a. *PL/SP*

Jón veit ekki hver bauð hverjum í veisluna.
John knows not who invited whom in the-dinner
'John does not know who invited who to the dinner.'

[56] Thanks to Željko Bošković for suggesting this to me.

b. ?PL/SP

Jón veit ekki hverjum bauð hver í veisluna.
John knows not whom invited who in the-dinner
'John does not know who invited who to the dinner.'

It is not clear, however, whether Icelandic has the Interpretive Superiority effect. My informant, Kjartan Ottosson (p.c.), tells me that the PL reading might still be available in (115b) and (116b) but it is harder to get than in the examples in (115a) and (116a). So there is a contrast pointing in the direction of Interpretive Superiority, but it is not as strong as has been reported for Japanese and Serbo-Croatian.[57]

What is a much clearer judgment is that Icelandic allows SP readings in multiple questions. It is standardly assumed that wh-phrases move to SpecCP in this language. According to the Relativized Minimality account, this should prevent the possibility of a SP reading in Icelandic, contrary to the fact.[58]

Another interesting set of data comes from Serbo-Croatian, which has been argued to involve overt wh-movement to SpecCP in embedded clauses. Thus, although Serbo-Croatian allows the SP reading in matrix multiple questions, as reported in Bošković (2003), we should not expect to find the SP reading in embedded clauses in Serbo-Croatian. My informants, however, do get the SP reading

[57] Some of my Japanese informants do not get the robust Interpretive Superiority effect either (even when presented with pictures and explicit scenarios). However, several informants did get the effect. Thus, the judgments seem quite delicate here.

[58] The scenarios presented with the test sentences were carefully controlled as to avoid the possibility of getting the Order reading, which is similar to the SP reading, yet not the same. Thus, we can be pretty certain that the reading that was obtained is indeed the SP reading. However, in future work, it would be interesting to see if the result is replicable with a non-reversible predicate like *buy* or *say*, where the Order reading is inherently unavailable.

in embedded clauses, as in (117), as well as in matrix clauses, as in (118) in Serbo-Croatian.

(117) Pavle je pitao ***ko šta*** o njemu govori. *PL/SP*
Pavle Aux asked who what about him says
'Pavle asked who says what about him'

(118) ***Ko šta*** o njemu govori? *PL/SP*
who what about him says
'Who says what about him?'

Note that *him* is used in (117) to ensure the truly embedded status of the subordinate clause. And the scenarios given to the informants for the two readings were as follows. On the single-pair reading, Pavle knows that there is *one* person saying something about him. On the pair-list reading, Pavle knows that there are several people saying different things about him.

Thus, the correlation between overt wh-movement to SpecCP and the availability of the SP reading seems to be more of a tendency since Icelandic and embedded clauses in Serbo-Croatian present potential exceptions to this generalization. In what follows, I present an account of the cross-linguistic distribution of the PL and SP that can be viewed as an alternative account to that of the Relativized Minimality account of Bošković (2003) presented above. However, it can also be viewed as an addition to the Relativized Minimality account, if we want to also capture the tendency of overt movement to SpecCP ruling out the SP reading. For more discussion of these two accounts, see the next subsection, as well as

Subsection 5.2 of Chapter 5, where I discuss the learnability issues associated with the two accounts.

4.2 Parameterized Selectional Restrictions of the Q-morpheme

In Grebenyova (2004), I propose that the distinction between the languages with and without SP readings lies in the selectional restrictions of the Q-morpheme. Recall the structures for the PL and SP readings from Hagstrom (1998), repeated below.

(119) a. $[_{CP}\ Q_j\text{-}C^0 \ldots[_{TP} \ldots WH1 \ldots V \ldots\ t_j\ \ WH2]]$ *PL*

b. $[_{CP}\ Q_j\text{-}C^0 \ldots[_{QP}\ \ t_j\ [_{TP} \ldots\ WH1 \ldots V \ldots\ WH2]]]$ *SP*

If a Q-morpheme cannot be merged with TP in some languages, those languages would not have the option of licensing the SP reading. That is what I propose happens in Bulgarian, English, and Russian. The Q-morpheme in these languages only selects a wh-phrase and never TP. That is why these languages lack the SP reading in multiple questions with bare wh-phrases. As for the questions with complex wh-phrases, we will discuss those in Section 8.

Supporting evidence for this analysis comes from Serbo-Croatian multiple wh-questions with an overt Q-morpheme *li*. Recall that Serbo-Croatian is a language allowing both PL and SP readings in multiple interrogatives. However, when *li* is

used in a multiple wh-question in this language, it forces the SP reading on the question, as shown in (120a) and (120a).[59]

(120) a. Ko li koga pozva na večeru? *SP/??PL*
who Q whom invited to dinner
'Who (on earth) invited who to the dinner?'

b. Ko li koga tuče? *SP/??PL*
who Q whom beat
'Who (on earth) is beating whom?'

Based on these facts, I suggest that Serbo-Croatian has two Q-morphemes. One is phonetically null and selects either a wh-phrase or TP, resulting in ambiguity between the PL and SP readings. The other Q-morpheme is phonetically realized as [li] and selects only TP, producing only the SP reading. The use of the latter morpheme in wh-questions seems to be on its way out in Serbo-Croatian.[60]

This analysis does not postulate anything new in the system. Particularly, it does not posit that the Q-morpheme carries a *+wh* feature. Hence, the problems

[59] *Li* is primarily used in Yes/No questions in Serbo-Croatian. When used in wh-questions, it adds an emphatic force to a question. This additional semantic property of *li* should not prevent us from analyzing it as a legitimate Q-morpheme, for such "fusion" of functional and lexical information is a common property of Slavic morphology (e.g., Slavic aspectual affixes often carry additional lexical meaning along with grammatical information). For a detailed study of the behavior of *li* in Serbo-Croatian and other Slavic languages, see Bošković (2001b).

[60] Bulgarian allows *li* in multiple wh-questions even more freely than Serbo-Croatian. Unlike Serbo-Croatian *li*, Bulgarian *li* is compatible with the PL reading. I suspect that this difference between Bulgarian and Serbian/Croatian stems from a more general difference between *li* in those two languages, as discussed in Bošković (2001b). Russian *li* is only allowed in Yes/No questions and is disallowed in wh-questions.

associated with Relativized Minimality do not arise. The theoretical foundation for this parameterization is already present in the analysis of Hagstrom (1998) in that there are two structural positions for the Q-morpheme. Hence, it seems only natural to derive the cross-linguistic distribution of the SP readings from the selectional restrictions of this morpheme. This account also captures the Icelandic facts, which are problematic for the Relativized Minimality account. In addition, it is consistent with the minimalist spirit in that cross-linguistic variation is restricted to the properties of individual items in the lexicon.

The account has predictions for learnability of PL/SP readings. We will discuss those in Chapter 5, where I present the experimental data from Russian-, English-, and Malayalam-speaking children, further supporting this theory of parametric variation.

5. Interpretive Superiority

Let us now examine the nature of Interpretive Superiority. My analysis of it will depend on the interpretation of wh-phrases under the Copy Theory of movement. As was discussed in Section 2, combining Reinhart (1995, 1997)'s approach to interpreting wh-in-situ via choice functions, we can interpret even moved wh-phrases that way. Whether a given copy of a moved wh-phrase will be interpreted will depend on whether that copy is in the scope of C^0, which existentially binds the choice function variable. On this approach every wh-phrase comes with a choice function.

On Hagstrom (1998)'s approach, however, wh-phrases do not introduce their own choice function variables, but rather it is a Q-morpheme that introduces a single

choice function variable. Given, flexible-functional application (FFA), we can potentially interpret any copy of any wh-phrase in any position. The existential closure by C^0 is not at issue here with respect to the copies of wh-phrases. The question arises then, which copy out of several copies of a moved wh-phrase is interpreted by semantics? I suggest that an analogy with PF can be useful here. It is (almost) always the higher copy that is interpreted by phonology. Pronouncing the lower copy is permitted only when pronouncing the higher copy creates a phonological problem, as discussed in Bošković and Franks (2002). If similar logic is applied to semantic interpretation, a higher copy of a moved wh-phrase must be chosen over a lower one for the semantic interpretation (unless some semantic problem arises).[61] This overall preference for the interpretation of copies in the target position of movement by both PF and LF might be due to the general constraint against vacuousness of movement.[62]

On Hagstrom's approach, this means that the higher copy of a wh-phrase would be interpreted as a set of individuals. I assume that the lower copy in an argument position can simply be treated as a variable over individuals, over which lambda-abstraction would take place.

Now, recall that Interpretive Superiority characterizes the loss of the PL reading in contexts where a lower wh-phrase moves over a wh-phrase higher in the

[61] I will discuss the issues concerning recunstruction in the next section, where we will discuss complex wh-phrases.

[62] See also Lasnik (2001) for the evidence that DPs in A-movement are interpreted in the landing site of movement.

structure, as we observed in Serbo-Croatian and Japanese. Consider (121), which demonstrates abstractly what happens in such PL derivation.

(121) $[_{CP}$ WH2$_j$ Q$_i$-C^0 …$[_{TP}$ … WH1 …V… t_i t_j]] *PL*

The choice-function variable left by Q-movement needs a set to apply to. However, if the object wh-phrase moves, as in (121), the Q-morpheme is left without a set to apply to. Hence, one cannot derive the PL reading in context of object wh-fronting. In languages like Serbo-Croatian and Japanese, where the Q-morpheme selects either a wh-phrase or TP, if the PL derivation is unavailable, an alternative derivation is available, namely, the SP reading derivation with the Q-morpheme originating above TP. The Q-morpheme in this derivation will not be affected by object fronting since it takes a different set as its argument (i.e., a set of propositions denoted by TP).

What about languages that do not have the option of generating the structure with the Q-morpheme above TP, like English, Russian, and Bulgarian? In these languages, we would expect the result of fronting the lower wh-phrase over the higher one to be simply unacceptable, since the alternative SP reading derivation is unavailable in these languages. That fits perfectly with our analysis in Chapter 2. The degraded status of an English matrix question in (122a), is then the result of not having the right Q-morpheme needed for the SP reading derivation when such derivation is being forced. Recall that T-to-C movement 'obviates' Attract Closest in (122a) in the way described in Chapter 2. Thus, the only source of unacceptability here is Interpretive Superiority.

(122) a. ??What$_1$ did who buy t_1?

b. *John wonders what$_1$ who bought t_1

However, there is an additional factor involved in embedded clauses that is not there in the main clauses, namely, the absence of T-to-C movement, which prevents C^0 from attracting the lower wh-phrase over the higher one (a minimality effect).[63]

There is one remaining issue to address here. Recall that in previous discussion, we found no problem with allowing the Q-morpheme to merge with any wh-phrase in any position. Here we do find a problem with that kind of freedom. If, in the derivation in (121), a wh-phrase merged with the higher wh-phrase instead of the lower one, the resulting sentence would be predicted to be fine and get a PL reading. This creates a problem for English since the resulting sentence is degraded. It also creates a problem for Japanese and Serbo-Croatian, where the corresponding sentences fine but only get a SP and not the PL reading. On this problematic derivation, however, they are expected to be ambiguous between the two readings.

This leads us back to the conclusion of Hagstrom (1998) that the Q-morpheme must always merge with the lowest wh-phrase and conclude with Hagstrom that, at least for these languages, it seems to be true. The intriguing question is then why it is true and how we can implement a selectional restriction such that an item selects a lower and not the higher instance of the same category. I suggest that it can be done by hypothesizing a timing restriction on the merger of the Q-morpheme where it must merge as soon as possible. If Merge involves feature-checking, as was suggested, for

[63] See Hagstrom (1998) for an alternative analysis of Interpretive Superiority based on the parametric variation with respect to stranding the Q-morpheme when a wh-phrase that is merged with it moves.

example, in Hornstein (2001), then this can be viewed as a result of a strong (viral) selectional feature on the Q-morpheme. Such a feature would require for it to merge as soon as the first wh-phrase is introduced into the derivation. In a bottom-up derivation, this requires merging with the lowest wh-phrase. Hence, this is a potentially interesting outcome, providing a potential insight into the nature of selectional features and their sensitivity to the timing of the derivation.

The only viral features we are familiar with so far are the ones that affect (or 'infect' and potentially crash) the derivation; and those features cannot stay in the derivation but have to be licensed immediately. However, these features can remain in the numeration without causing any viral problem there. My proposal essentially extends the same logic to the selectional features. These features do not affect the derivation but rather affect the numeration, causing the same viral effect there. Such a feature then must be licensed as soon as possible (by merging as soon as possible). This suggests that numeration has at least some structure to it. This is in line with what seems already implicitly assumed in Chomsky (2000, 2001) and in line with the recent work of Uriagereka (To appear).

It appears that the Q-morpheme has this viral property only in some languages. Recall the data from Navajo, Okinawan, and Russian, which do not exhibit Interpretive Superiority effects and, in case of Navajo and Okinawan, we can see the overt Q-morpheme being capable of merging with the higher as well as the lower wh-phrase. The Q-morpheme in these languages must then be 'weak' (or non-viral) in that it does not have the requirement of being merged as soon as possible.

6. Complex Wh-phrases

As was demonstrated in the beginning of this chapter, in languages that lack SP readings in questions with bare wh-phrases, both SP and PL readings are available in questions with complex wh-phrases, as in (123) from English and Russian.

(123) a. *PL/SP*
Which diplomat invited which journalist to the dinner?

b. Kakoj diplomat kakogo žurnalista priglasil na užin? *Russian*
which diplomat which journalist invited to dinner
'Which diplomat invited which journalist to the dinner?'

Russian allows for the second complex wh-phrase to optionally remain in situ. The judgments with respect to PL/SP readings remain the same in this case.

When the lower complex wh-phrase is fronted over the higher one, the Interpretive Superiority effect does not arise. That is, both PL and SP readings remain available, as shown below.

(124) a. *PL/SP*
Which journalist did which diplomat invite to the dinner?

b. Kakogo žurnalista kakoj diplomat priglasil na užin? *Russian*
which journalist which diplomat invited to dinner
'Which journalist did which diplomat invite to the dinner?'

To account for this asymmetry between the bare and complex wh-phrases, I propose that unlike bare-wh-phrases, complex wh-phrases come with their own choice functions. One can view it as a special contribution of *which* in *which girl*. In

this case, both wh-phrases in a multiple question can be interpreted with choice functions. This reduces the sets of individuals denoted by these wh-phrases to singleton sets and neither wh-phrase proliferates through the derivation, in the end producing only a set of propositions and not a set of sets of propositions. This allows questions with complex wh-phrases to have SP readings.

It also seems that complex wh-phrases come in different flavors. For example, in Russian, there are certain complex wh-phrases that require the SP reading, like the ones in (125); and then there are those that allow both PL and SP readings, as we saw in (123) and (124).

(125) *PL/SP

Kotoryj diplomat kotorogo žurnalista priglasil na užin? *Russian*
which diplomat which journalist invited to dinner
'Which diplomat invited which journalist to the dinner?'

Thus, I suggest that some wh-determiners come with obligatory choice function variables, while others are ambiguous in having an option to either activate their choice function variable or not. It is this optionality that allows them to have either a PL or a SP reading. And the bare wh-phrases are on the other side of the spectrum in that they do not have a choice function variable at all. For these phrases, the PL/SP possibilities are limited to those provided by the Q-morpheme.

It is interesting what happens to the Q-morpheme in the derivation with complex wh-phrases. Since it is not needed there, one might wonder if it is present there at all. So far, nothing seems to go wrong whether it is there or not. If it is there, it can apply vacuously to the output of the choice function application that takes place

within the wh-phrase. In the next section, we will see some evidence that the Q-morpheme is indeed there even in questions with complex wh-phrases.

In our discussion of the copy-interpretation in the previous section, I suggested that interpreting the higher copy of the moved element is the default option. It is then important to address the fact that complex wh-phrases allow and sometimes require reconstruction. Keeping everything else constant, reconstruction would either have to be an on-line process (i.e., binding conditions would have to be met in the process of the derivation) or there is something about a restrictor being separate in complex wh-phrases (unlike in bare wh-phrases) that allows it to reconstruct.

7. Locality Effects

The final set of phenomena to address in this chapter has to do the obligatory switch from PL readings to SP readings in multiple questions with an island boundary between the two wh-phrases, as in (126) and (127). The switch takes place unless it is a wh-island. The question in (128) is known to have a PL reading, as was first observed Baker (1970).

(126) a. Which linguist will be offended if we invite which philosopher? **PL/SP*
b. Who will be offended if we invite who? **PL/SP*
c. Who will be offended if we break what? **PL/SP*

(127) a. Which student read the book that which professor wrote? **PL/SP*
b. Who read the book that who wrote? **PL/SP*
c. Who read the book that describes what? **PL/SP*

(128) Who wonders where John bought what?

However, keep in mind that the facts in (126) - (127) are rather controversial, as we discussed in Section 1 of this chapter.

Let us begin by filling out the paradigm with long-distance multiple questions without islands. However, here the picture is not very clear. Dayal (2002) reports the judgments of an anonymous reviewer that (129a), does not have a PL reading but only has a SP reading. My informants are able to get the PL reading but point out that it is easier to get without the complementizer *that* in the embedded clause, as in (129b). Because of the ability of the complex wh-phrases to generate SP readings, let us also consider a bare wh-question, as in (129c). Here, my informants are able to get only the PL reading.

(129) a. Which student believes that Mary read which book?
b. Which student believes Mary read which book? *PL/SP*
c. Who believes Mary read what? *PL/*SP*

A related fact is from Hagstrom (1998), who presents data from Chinese and Japanese showing that, when both wh-phrases are inside an island in these languages, the matrix reading of a question can only be the SP reading and not the PL reading. This is shown below.

(130) a. *PL/SP
Ta keneng hui [yinwei Li jiao shei mai shenme] shengqi ne? *Chinese*
he maybe will because Li ask who buy what angry Q
'For what x and y, he might be angry because Li asked if x bought y?'

b. **PL/SP*

Tarooga [darega nanio katta tokini] okotta no? *Japanese*

$\text{Taroo}_{\text{NOM}}$ who_{NOM} what_{ACC} bought when got-angry Q

'For what x and y, Taroo got angry when x bought y?'

Following Hagstrom (1998)'s proposal that the Q-morpheme undergoes movement from the clause internal position, let us consider the latter facts first. Chinese and Japanese normally allow both PL and SP readings in bare multiple interrogatives, hence the Q-morpheme in these languages selects both a wh-phrase and a TP. If in the examples in (130), it were to merge with a wh-phrase to generate the PL reading, it would have to move across an island to get to the matrix C^0. To avoid the locality violation, it merges with the TP outside an island instead. However, this can only produce the SP reading. This is essentially, the logic behind Hagstrom's analysis of these facts.

Let us apply similar logic to the English cases. First, notice that the interpretive island effects are there regardless of whether it is a question with bare or complex wh-phrases. This seems to settle to issue raised in the end of the previous section on whether the Q-morpheme is present at all in questions with complex-wh-phrases. The island effects, which on our theory are diagnostic of the movement of the Q-morpheme, suggest that the Q-morpheme is present in these questions as well.

The crucial difference between Chinese and Japanese on one hand and English on the other is that English does not have an option of merging the Q-morpheme with TP. We concluded this from the fact that English questions like *Who bought what?* only have PL readings. What are the remaining options? Merging the Q-morpheme with the higher wh-phrase is prohibited by the timing restriction we

discussed in Section 5 of this chapter; and it cannot be merged with the wh-phrase inside an island because it wouldn't be able to move out of it. However, the Q-morpheme must be present in the derivation as the questions with complex wh-phrases suggest. It is in this particular context, when the movement of the Q-morpheme is precluded, I suggest that a resumptive strategy kicks in and the Q-morpheme is base-generated directly in C^0.

As for the wh-island case, the fact that we get the PL reading there can be viewed as an extra piece of evidence that it is the Q-morpheme that is moving and not the wh-phrase-in-situ. That is why the wh-phrase in the embedded SpecCP does not intervene in this case. This can be achieved under the assumption that wh-island is different from other islands in that, unlike other islands, it is an instance of Relativized Minimality. Alternatively, we could say that wh-island is like other islands and the Q-morpheme moves successive-cyclically, stopping in the intermediate C^0 (as opposed to the intermediate SpecCP).

The peculiar *that*-effect, where it is easier for some speakers to get the PL readings without *that* in the embedded C^0, might be telling us something as well. It could be a Relativized Minimality effect: although a wh-phrase in SpecCP does not intervene for the head-movement of the Q-morpheme, *that* in C^0 does, precisely because it is a head. Or again, it could be the case of successive cyclicity of the Q-movement. *That* in the intermediate C^0 then prevents the Q-morpheme from stopping there.

8. Summary

In this chapter, we have examined semantics of multiple interrogatives, focusing on various aspects of deriving PL and SP readings in these constructions. First, we explored various approaches to semantics of interrogatives and discovered that there is no semantic evidence for the existence of covert wh-movement. Hence, we concluded that, if covert wh-movement exists, it must be driven by a purely formal requirement.

After reviewing previous accounts of PL and SP readings and pointing out their empirical and conceptual problems, we adopted the basis of the analysis of Hagstrom (1998) of the syntax and semantics of the PL and SP readings. We also provided an account of the cross-linguistic variation with respect to the availability of the SP reading.

Further, by combining Hagstrom's account with the Copy Theory of movement, we developed an account of Interpretive Superiority. This lead us to some conclusions about the nature of the features inside the numeration, namely, that the selectional features can also have viral effects, analogous to the derivational features.

We then argued that the morphological distinction between bare and complex-wh-phrases results in a semantic distinction such that complex wh-phrases are equipped with their own choice function variables. This allowed us to capture the asymmetry between bare and complex wh-phrases with respect to PL and SP readings.

Finally, we explored the semantics of long-distance multiple questions and questions with islands. Here we found further evidence for the movement of the Q-

morpheme and a resumptive strategy employed by the Q-morpheme whenever movement from a clause-internal position is precluded.

In the next chapter, we will explore how multiple interrogatives behave under Sluicing.

Chapter 4: Multiple Interrogatives and Ellipsis

1. Introduction

Having examined the syntactic and semantic properties of multiple interrogatives in the previous chapters, we will now explore how these properties manifest themselves in structures involving ellipsis. The type of ellipsis that can be found in wh-clauses is TP-deletion (Sluicing). Hence, the discussion will focus primarily on sluicing, although the analysis will be extended to VP-ellipsis whenever relevant.

First, I determine what positions the remnant wh-phrases occupy in the sluices (i.e., the clauses undergoing sluicing) cross-linguistically. The nature of these positions is important for understanding what configurations license sluicing and why. I argue that contrastive focus is capable of licensing sluicing in languages like Russian, Polish, Hungarian, and Chinese. To support this conclusion, I show that contrastively focused R-expressions can be remnants of sluicing in these languages. However, it has been argued by Lobeck (1995) and Merchant (2001) that the [+wh] feature on the interrogative C^0 licenses sluicing in English. To avoid the stipulation that two different features (i.e., *+focus* and *+wh* features) can license sluicing in different languages, I propose that even in English, it is the *+focus* feature, and not the *+wh* feature that licenses sluicing. Wh-movement in this language simply happens to be the operation that gets a wh-phrase to the Spec of the projection that bears a (weak) *+focus* feature.

Second, I demonstrate how semantic properties of multiple interrogatives affect the availability of sluicing in certain contexts. Specifically, the semantic restrictions on Single-pair readings in Russian multiple interrogatives constrain the nature of the antecedent clauses required in multiple sluicing in these languages. This presents a new argument to the effect that the sluice (i.e., a clause where sluicing takes place) contains a full clausal structure at least by LF.

Finally, I explore how Superiority effects are manifested under sluicing. I demonstrate that although Superiority effects are not generally present in Russian, they emerge in sluicing contexts. A similar situation has been observed in Serbo-Croatian by Stjepanović (2003). I will derive these puzzling effects from Parallelism, an independently motivated property of ellipsis.

2. The Phenomenon of Sluicing

Sluicing refers to a phenomenon of clausal ellipsis, which was first discovered and explored by Ross (1969). A typical instance of sluicing can be found in an interrogative clause with only a wh-element pronounced, as in (131). The crossed out text indicates the unpronounced yet interpreted part of the structure.

(131) a. John will buy something but I don't know what ~~[John will buy *t*]~~.

Both the subject *John* and the modal auxiliary *will* are elided in (131). The fact that modals, located in T^0, and subjects, occupying SpecTP, are elided in sluicing

constructions suggests that we are dealing with TP-ellipsis. Sluicing occurs in main clauses as well, as can be seen in (132).[64]

(132) *Speaker A:* John loves somebody.
Speaker B: Who ~~[John loves *t*]~~?

I adopt the basic analysis of sluicing as in Ross (1969), Lasnik (2001) and Merchant (2001), where the derivation proceeds as in (133): a wh-phrase undergoes wh-movement to SpecCP and then TP is deleted at PF.[65]

(133) *Step 1:* John bought something. I wonder [$_{CP}$ *what* [$_{TP}$ John bought *t*]
Step 2: John bought something. I wonder [$_{CP}$ *what* [$_{TP}$ ~~John bought *t*~~]

There are alternative analyses of ellipsis, in which an empty category is present in the position of the elided TP and is replaced by copying the antecedent TP at LF. In this case, no deletion takes place since there is no clausal structure in the sluice to start with. Such analyses have been developed in Williams (1977), Lobeck (1991, 1995), and Chung et al. (1995). There are also strictly semantic approaches, as developed in Dalrymple et al. (1991), Jacobson (1992), and Hardt (1993, 1999). However, extensive arguments against the non-deletion approaches can be found in

[64] See Bechhofer (1976 and 1977), Lasnik (2001), and Merchant (2001) for extensive arguments that sluicing in main clauses is indeed an instance of clausal ellipsis and is different from fragment questions.

[65] Ross (1969) actually argues for the deletion taking place at S-structure. However, with the elimination of S-structure as a level of representation, the deletion can be viewed as taking place at PF or at the point of Spell-out.

Ross (1969), Merchant (2001) and Stjepanović (2003). Thus, in what follows, I will assume the PF-deletion analysis of sluicing.

Sluicing is quite common across languages and is very productive in Slavic. I will primarily focus on Russian and Polish, and draw parallels with other Slavic languages whenever relevant. Consider the sluicing examples from Russian and Polish in (134) and (135) respectively, where the (a) examples demonstrate embedded sluicing and the (b) examples demonstrate matrix sluicing.[66]

(134) a. Ivan budet davat' komu-to podarki, no ja ne znaju komu/*kto *Russian*
Ivan will give someone$_{DAT}$ presents but I not know who$_{DAT/NOM}$
'Ivan will be giving someone presents but I don't know who.'

b. *Speaker A:* Ivan budet davat' komu-to podarki.
Ivan will give someone$_{DAT}$ presents
'Ivan will be calling someone.'
Speaker B: Komu/*Kto?
who$_{DAT}$/who$_{NOM}$
'Who?'

(135) a. Jan bedzie dawac komuś prezenty ale nie wiem komu/*kto. *Polish*
Jan will give someone$_{DAT}$ presents but not know who$_{DAT/NOM}$
'Jan will be giving someone presents but I don't know who.'

b. *Speaker A:* Jan bedzie dawac komuś prezenty.
Jan will give someone$_{DAT}$ presents
'Jan will be giving someone presents.'

[66] For the corresponding examples from Bulgarian and Serbo-Croatian, see Merchant (2001) and Stjepanović (2003) respectively.

Speaker B: Komu/*Kto?
who$_{DAT}$/who$_{NOM}$
'Who?'

Observe that the remnant wh-phrases in these examples are obligatorily marked with overt dative case morphology and match the case of the indefinites in the antecedent clauses. The Russian verb *davat'* and the Polish verb *davac*, corresponding to the English verb *give*, obligatorily assign dative case to the indirect object. The fact that switching the case of the remnant wh-phrases to nominative produces unacceptability argues that these wh-phrases have indeed moved from a position inside TP, where the dative case was assigned. This strongly suggests that we are, in fact, dealing with sluicing. A potential alternative is Pseudo-sluicing, which would have a cleft structure in the sluice, as in (136).

(136) John called someone on the phone but I don't know who ~~[it was]~~.

Clefted elements in Slavic obligatorily bear nominative case, as shown in (137) from Russian and (138) from Polish.

(137) Ivan podaril komu-to podarok, no ja ne znaju kto/*komu eto byl. *Russian*
Ivan gave someone present but I not know who$_{NOM/DAT}$ it was
'Ivan called gave someone a present but I don't know who it was.'

(138) Jan dal komuś prezent ale nie wiem kto/*komu to byl. *Polish*
Jan gave someone present but not I-know who$_{NOM/DAT}$ it was
'Ivan called gave someone a present but I don't know who it was.'

It is the opposite of what we find in the paradigm in (134) - (135). Thus, we can conclude that the examples in (134) - (135) are indeed instances of sluicing.

Besides sluicing with a single wh-remnant, Slavic also permits sluicing with multiple wh-remnants, as in (139) and (140). Following Takahashi (1994), I will refer to this phenomenon as *multiple sluicing*. Like single sluicing, multiple sluicing is available in embedded clause, as in the (a) examples, and in main clauses, as in the (b) examples below.

(139) a. Každyj priglasil kogo-to na tanec, no ja ne znaju kto kogo. *Russian*
everyone invited someone to dance but I not know who whom
'Everyone invited someone to dance but I don't know who whom.'

b. *Speaker A:* Každyj priglasil kogo-to na tanec.
everyone invited someone to dance
'Everyone invited someone to dance.'

Speaker B: Kto kogo?
who whom
'Who whom?'

(140) a. Kazdy zaprosil kogoś do tanca, ale nie pamietam kto kogo. *Polish*
everyone invited someone to dance but not know who whom
'Everyone invited someone to dance but I don't know who whom.'

b. *Speaker A:* Kazdy zaprosil kogoś do tanca.
everyone invited someone to dance
'Everyone invited someone to dance.'

Speaker B: Kto kogo?
who whom
'Who whom?'

It is this construction that is most relevant for our task of exploring how syntactic and semantic properties of multiple interrogatives are manifested under ellipsis.

The availability of multiple sluicing in Slavic is not surprising since it is well known that Slavic languages are multiple wh-fronting languages. That is, all wh-phrases are typically fronted in non-elliptical multiple questions in Slavic. This is shown below with a representative paradigm from Russian, although similar paradigms for other Slavic languages can be found in Rudin (1988), Bošković (1997a, 1998, 2002a), Richards (1997), among others.

(141) a. Kto$_1$ kogo$_2$ [t_1 ljubit t_2]?
who whom loves
'Who loves who?'

b. *Kto$_1$ [t_1 ljubit kogo]?
who loves whom

Since there is an independent way for multiple wh-remnants to move out of TP in Slavic, it is reasonable to assume that the same happens in multiple sluicing. This line of reasoning has implications for languages that might have something resembling multiple sluicing found in Slavic, yet no multiple wh-fronting. Japanese, Hindi, and certain contexts in English have been reported to allow structures that look like multiple sluicing (see Takahashi (1994) for Japanese, Merchant (2001) and Mahajan

(to appear) for Hindi, and Richards (2001) and Lasnik (2005) for English). In the most straightforward scenario, these cases would have to be analyzed as involving a different derivation from the one operative in Slavic. And many researchers have gone precisely in that direction, attributing the rise of these structures to Pseudo-clefting (Takahashi, 1994), Gapping (Mahajan, to appear), or Extraposition (Lasnik, 2005).

In the following sections, we will examine how the syntactic and semantic properties of multiple interrogatives are manifested in the context of multiple sluicing and what these properties can tell us about the nature of sluicing.

3. Licensing TP-deletion

One of the central issues in ellipsis is what categories license the elision of their complements. Beginning with Ross (1969), researchers have been identifying the interrogative *+wh* complementizer as the head licensing the deletion of its complement TP. This conclusion is largely based on the fact that sluicing in Germanic is restricted to the interrogative clauses with a wh-phrase in SpecCP. Lobeck (1995) and Merchant (2001) examine various contexts in English where one might expect TP-ellipsis to be possible, yet it is not. These contexts include finite declarative clauses, lexically governed TP-s, and relative clauses (including clefts and free relatives). Thus, Merchant (2001) concludes that the complementizer bearing the *+Q* and the *+wh* features licenses the deletion of its complement TP. This is illustrated in (142).

(142) John bought something. I wonder [$_{CP}$ what C^0 [$_{TP}$ ~~John bought *t*~~].
+*Q*
+*wh*

However, it is not clear how this analysis can be straightforwardly extended to the Slavic languages that exhibit a rather different pattern of wh-movement from the one found in Germanic. Stjepanović (1998) and Bošković (1998, 2002a) extensively argue that wh-fronting in Slavic languages like Russian, Polish, and some contexts in Serbo-Croatian involves focus-movement of the wh-phrases to a position below CP. In some languages, like Bulgarian, the +*focus* feature is located on the interrogative C^0, along with the strong +*wh* feature. The target position of wh-movement in Bulgarian is then SpecCP, just as in English, presenting no problem for C^0 being the licenser of sluicing. However, sluicing in Russian, Polish and certain contexts in Serbo-Croatian is in need of explanation. How do the remnants of sluicing survive deletion if their target position of movement is part of the complement of C^0? Why are they not deleted along with the complement of C^0?

3.1. Multiple Wh-fronting and Contrastive Focus

Let me describe the focus-movement analysis of wh-fronting. I will concentrate on Russian but the same logic extends to Polish. Stepanov (1998) argues that wh-movement in Russian is not driven by the +*wh* feature of C^0 and, therefore, the wh-phrases, even though they move, do not end up in SpecCP in overt syntax. He uses superiority as a diagnostic of a strong feature triggering movement and assumes the Economy approach to superiority, where C^0 with a strong +*wh* feature attracts the closest element with a +wh feature to SpecCP for feature checking, as in the Minimal

Link Condition of Chomsky (1995). This approach explains the presence of superiority effects in English. Consider the paradigm from English in (143), repeated from Chapter 2 of this thesis. In both (143b) and (143d), C^0 attracts *what*, which is not the closest wh-phrase to C^0. The closer wh-phrase is *who*, hence the instances of wh-movement in (143b) and (143d) are not economical.[67]

(143) a. Who bought what?
b. ??What did who buy *t*?
c. Who did John persuade *t* to do what?
d. *What did John persuade who to do *t*?

Notice that only one wh-phrase is fronted in English. Some multiple wh-fronting languages also exhibit superiority effects. Bulgarian is a language like that. The order of the fronted wh-phrases is fixed in Bulgarian, such that the wh-phrase which is the closest to C^0 prior to wh-movement precedes other wh-phrases after all wh-phrases move. This is shown in (144) for main and embedded clauses.[68]

(144) a. Koj kogo e pokanil na večeriata? *Bulgarian*
who whom Aux invited to party
'Who invited who to the party?'

[67] I am abstracting away from the effect of T-to-C movement on superiority, which makes (143b) less degraded than (143d), as discussed in Chapter 2.

[68] I am using '*who'* for both subject and object wh-phrases for the Slavic paradigms in order to avoid the homophony created by the *'who-what'* combination in these languages. Homophony tends to interfere with superiority effects, as was observed by Stepanov (1998) and Bošković (2002). The accusative *who* is different enough from the nominative *who,* allowing us to control for this interfering factor.

b. *Kogo koj e pokanil na večeriata?
whom who Aux invited to party

c. Tja me popita koj kogo e pokanil na večeriata.
she me asked who whom Aux invited to party
'She asked me who invited who to the party.'

d. *Tja me popita kogo koj e pokanil na večeriata.
she me asked whom who Aux invited to party

As was discussed in Chapter 2, in order to extend the Economy analysis of superiority to Bulgarian successfully, it is not sufficient for C^0 to attract the closest wh-phrase to its Spec first. It must be insured that either the next wh-phrase tucks-in underneath the first one, as in Richards (1997), or that it necessarily right-adjoins to the first wh-phrase, as in Rudin (1998) and Bošković (1998).

Unlike English and Bulgarian, Russian multiple wh-questions do not exhibit superiority effects in virtually any contexts. This is illustrated in (145) for main and embedded clauses.

(145) a. Kto kogo priglasil na večer? *Russian*
who whom invited to party
'Who invited who to the party?'

b. Kogo kto priglasil na večer?
whom who invited to party

c. Ja ne znaju kto kogo priglasil na večer.
I not know who whom invited to dinner
'I don't know who invited who to the party.'

d. Ja ne znaju kogo kto priglasil na večer?
I not know whom who invited to party

How can these facts be reconciled with the Economy approach to superiority? Stepanov (1998) proposes that Russian has a weak *+wh* feature, like in the wh-in-situ languages (e.g., Japanese, Korean, etc.). Thus, the *+wh* feature in Russian does not trigger overt wh-movement and hence we do not find superiority effects.

This raises the question as to why wh-phrases obligatorily front in Russian. Stepanov attributes such fronting to contrastive focalization. The idea is based on the correlation between *wh*-fronting and fronting of contrastively focused R-expressions in Slavic, first observed by Stjepanović (1998). Just like wh-phrases, contrastively focused R-expressions are fronted in Slavic, as demonstrated in (146).[69]

(146) a. IVANA ja vstretila *t*. Russian
Ivan$_{ACC}$ I met$_{1.FEM.SG}$
'I met IVAN'

[69] It is also possible to front the focused phrases to the immediately preverbal position in Russian as in (i). This suggests that there might be two focus positions in Russian: one is TP internal and the other is TP external. Interestingly, wh-phrases can use the lower focus position as well, as in (ii).

(i) Ja IVANA vstretila.
I Ivan$_{ACC}$ met$_{1.FEM.SG}$
'I met IVAN'
(ii) Komu Ivan čto dal?
who$_{DAT}$ Ivan what$_{ACC}$ gave
'To whom did Ivan give what?'

d. ??Ja vstretila IVANA.
I met IVAN_{ACC}

Thus, Stepanov (1998) concludes that wh-phrases in Russian are fronted to a *focus* position below CP. As mentioned before, the same argument can be made for Polish, since superiority effects are absent in Polish in the same contexts as in Russian. Stepanov (1998) further explains the insensitivity of such focalization to superiority by suggesting, following Bošković (1998), that each wh-phrase itself carries a strong +*focus* feature and therefore the wh-phrases do not compete with each other with respect to the closeness to C^0. See also Bošković (2002a) for the purely Attract-based version of this analysis.

3.2. A Note on Multiple Foci

The analysis of multiple wh-fronting above strongly relies on the correlation between wh-fronting and contrastive focus. However, there is a long-standing puzzle that this correlation faces: although multiple wh-fronting is very productive, multiple contrastive foci in general are not easy to find. That is, it is hard to contrastively focus more than one R-expression in a clause, as shown in (147).

(147) *IVANU KNIGU on podaril. *Russian*
Ivan_{DAT} book_{ACC} he gave
'He gave IVAN A BOOK.'

To solve this puzzle, I suggest that multiple wh-questions and clauses with multiple foci share certain semantic properties. Recall from Chapter 2 that multiple wh-questions in English and Russian require pair-list and not single-pair readings:

(148) Who bought what? *PL/*SP*

(149) Kto čto kupil? *PL/*SP* *Russian*
who what bought
'Who bought what?'

If the clauses with multiple foci behave similarly to the clauses with multiple wh-phrases syntactically, it is not unreasonable they behave similarly semantically as well. Specifically, what if the structures with multiple foci in general require pair-list readings? The only way to get a pair-list interpretation with R-expressions is by literally enumerating the pairs of participants in an event, as in (150).

(150) IVANU VELOSIPED Ded Moroz podaril, LENE KNIGU on podaril, a
Ivan_{DAT} bike_{ACC} Santa Claus gave Lena_{DAT} book_{ACC} he gave and

SAŠE ČASY on podaril.
Saša_{DAT} $\text{watch}_{\text{ACC}}$ he gave
'Santa Claus gave Ivan a bike, he gave Lena a book, and he gave Saša a watch.

It is unacceptable to leave all the foci in situ, as shown in (151), just as expected since these are contrastive foci. However, it is possible to leave those in situ if it is done only in the first conjunct, as in (152). The nature of this effect is unclear, but it might have something to do with processing. Only after processing the first

conjunct and at the point of beginning to process the second conjunct, it is clear that one is dealing with a list of propositions and not just a single proposition. And only a list of propositions can generate a pair-list reading, which is, by hypothesis, required for multiple foci.

(151) *Ded Moroz podaril IVANU VELOSIPED, on podaril LENE KNIGU, a
Santa Claus gave $Ivan_{DAT}$ $bike_{ACC}$ he gave $Lena_{DAT}$ $book_{ACC}$ and

on podaril SAŠE ČASY.
he gave $Saša_{DAT}$ $watch_{ACC}$

(152) Ded Moroz podaril IVANU VELOSIPED, LENE KNIGU on podaril, a
Santa Claus gave $Ivan_{DAT}$ $bike_{ACC}$ $Lena_{DAT}$ $book_{ACC}$ he gave and

SAŠE ČASY on podaril.
$Saša_{DAT}$ $watch_{ACC}$ he gave

3.3. Focus-licensed Sluicing

Now returning to sluicing, we must explain how the remnant wh-phrases in Russian and Polish sluicing survive the deletion if they are not in SpecCP. I propose that any functional category bearing a *+focus feature* can license the deletion of its complement, as illustrated in (153) below.

(153) Ivan kupil čto-to, no ja ne znaju [**čto** X^0 [$_{TP}$ ~~Ivan kupil **t**~~]]?
+*focus*
Ivan bought something but I not know what Ivan bought
'Ivan bought something but I don't know what.'

This allows for the wh-phrases in Russian and Polish to survive TP-deletion.

A direct implication of this proposal is that sluicing should be possible with contrastively focused R-expressions as remnants. The data from Russian below shows that contrastively focused R-expressions can in fact be remnants of sluicing. In (154), the remnant is *Mašu* and, in (155), we have three remnants: a wh-phrase and two R-expressions. This further strengthens the parallelism between wh-fronting and contrastive-focus-fronting in Slavic.

(154) *Speaker A:* Ty skazala čto on budet uvažat' Mašu? *Russian*
you said that he will respect Maša_{ACC}
'Did you say that he will respect Maša?'

Speaker B: Net. Ja skazala čto IVANA ~~[on budet uvažat' *t*]~~
no I said that Ivan_{ACC} he will respect
'No. I said that (he will respect) IVAN.'

(155) *Speaker A:* Ty ne pomniš kogda Ivan vstretil Mašu?
you not remember when Ivan_{NOM} met Maša_{ACC}
'You don't remember when Ivan met Maša?'

Speaker B: Net, ja ne pomnju GDE SERGEY LENU
no I not remember where $\text{Sergey}_{\text{NOM}}$ Lena_{ACC}
'No, I don't remember WHERE SERGEY (met) LENA.'

Polish shows the same behavior, as demonstrated in (156) and (157).

(156) *Speaker A:* Powiedzialas, że szanujesz Marie? *Polish*
you-said that he-will-respect Maria$_{ACC}$
'Did you say that he will respect Maria?'

Speaker B: Nie, powiedzialam że Jana ~~[szanujesz *t*]~~.
no I-said that Jan$_{ACC}$ he-will-respect
'No. I said that (he will respect) JAN.'

(157) *Speaker A:* Nie pamietasz, kiedy Jan spotkal Marie?
not you-remember when Jan$_{NOM}$ met Maria$_{ACC}$
'You don't remember when Ivan met Maria?'

Speraker B: Nie. Nie pamietam GDZIE BARBARA ZOSIE.
no. not I-remember where Barbara$_{NOM}$ Zosia$_{ACC}$
'No. I don't remember WHERE BARBARA (met) ZOSIA.'

Let us examine the properties of this construction in detail. First, it is important to make sure that we are actually dealing with sluicing. Alternative derivations could involve pseudo-gapping or gapping.

It is quite unlikely that the data above are the instances of pseudo-gapping, which has been analyzed VP-ellipsis in much of the literature (e.g., Sag (1976), Jayaseelan (1990), and Lasnik (1995)). Notice that, in (154), the auxiliary *budet 'will'* is elided, indicating that a larger constituent than VP is elided (under the standard assumption that such auxiliaries are generated in T^0). In addition, pseudo-gapping is not readily available in Slavic in general, as shown in Russian (158).

(158) *Maša budet čitat' knigu, a Ivan budet gazetu ~~[čitat' *t*]~~.
Maša$_{NOM}$ will read book$_{ACC}$ and Ivan$_{NOM}$ will newspaper$_{ACC}$ read
'Maša will read a book and Ivan will a newspaper.'

Another possibility to consider is that the data above are derived through gapping. However, given the well known properties of gapping, it too cannot account for the cases under consideration. First, similarly to English, gapping in Slavic is largely restricted to local coordinations with conjunctions corresponding to the English *and* and *or*; the conjunction corresponding to *but* cannot occur in gapping structures, as demonstrated in (159).

(159) a. Maša budet čitat' knigu, a Ivan ~~budet čitat'~~ gazetu.
Maša$_{NOM}$ will read book$_{ACC}$ and Ivan$_{NOM}$ will read newspaper$_{ACC}$
'Maša will be reading a book and Ivan a newspaper'

b. Ili Maša budet čitat' knigu, ili Ivan ~~budet čitat'~~ gazetu.
either Maša$_{NOM}$ will read book$_{ACC}$ or Ivan$_{NOM}$ will read newspaper$_{ACC}$
'Either Maša will be reading a book or Ivan a newspaper'

c. *Maša budet čitat' knigu, no Ivan ~~budet čitat'~~ gazetu.
Maša$_{NOM}$ will read book$_{ACC}$ but Ivan$_{NOM}$ will read newspaper$_{ACC}$
'Maša will be reading a book and Ivan a newspaper'

This is not the case in (154) - (157) since these can easily contain *but*, as demonstrated below.

(160) Ty skazala čto on budet uvažat' Mašu, **no** ja dumaju čto IVANA.
you said that he will respect Maša$_{ACC}$ **but** I think that Ivan$_{ACC}$
'You said that he will respect Maša, but I think that he will respect Ivan.'

Second, as in English, gapping cannot take place in an embedded clause in Russian, as shown by the contrast between (161a) and (161b).

(161) a. Maša budet čitat' knigu, a Ivan ~~budet čitat'~~ gazetu.
Maša$_{NOM}$ will read book$_{ACC}$ and Ivan$_{NOM}$ will read newspaper$_{ACC}$
'Maša will be reading a book and Ivan a newspaper'

b. *Maša budet čitat' knigu, a Lena dumala, čto Ivan gazetu.
Maša will read book and Lena$_{NOM}$ thought that Ivan newspaper
'Maša will be reading a book and Lena thought that Ivan a newspaper.'

Moreover, gapping cannot seek an antecedent in en embedded clause, as the contrast between (162a) and (162b) illustrates.

(162) a. Ili Maša budet čitat' knigu, ili Ivan ~~budet čitat'~~ gazetu.
either Maša$_{NOM}$ will read book$_{ACC}$ or Ivan$_{NOM}$ will read newspaper$_{ACC}$
'Either Maša will be reading a book or Ivan a newspaper'

b. *Ili Lena dumala, čto Maša budet čitat' knigu,ili Ivan budet čitat' gazetu.
either Lena thought that Maša will read book or Ivan will read newspaper
'Either Lena thought that Maša will be reading a book, or Ivan a newspaper'

None of these basic requirements for gapping are met in (154) - (157), leaving sluicing as the only plausible derivation for these data.

3.4. Overt material in Comp

Lobeck (1995), Chung et al. (1995), Lasnik (1999a) and Merchant (2001) among others have observed an interesting fact that nothing besides the overt material in SpecCP can survive sluicing. That is, no overt material in C^0 itself survives sluicing. The data for this generalization comes from a number of languages, such as English, Danish, Dutch, Frisian, German, Norwegian, Slovene, among others (see Merchant (2001) for the data from all these languages). Here is a representative paradigm from English, showing that, although T-to-C movement is obligatory in the main clauses in English, the auxiliary cannot remain undeleted under sluicing:

(163) a. What will John buy?
b. *What John will buy?
c. John will buy something but I don't know what.
d. *John will buy something but I don't know what will.

The generalization extends not only to the elements that move to C^0 but also to those that are base generated there, as, for example, the case in Slovene. The analyses of the moved elements in Lasnik (1999a) and Merchant (2001) rely on Economy and feature-movement. The logic of it is such that, if the element in T^0 does not move to C^0 overtly, this material in T will cause a PF crash (either because its strong feature will remain unchecked or because this material will be unpronounceable at PF, on the feature-movement account). However, if ellipsis deletes the structure with the inadequacy at PF, the problem goes away. On this analysis, *will* in (163) does not move to C^0, creating a problem that is later eliminated by sluicing. The base-generated elements, on the other hand, are analyzed as clitics

which must cliticize to the right, and hence cannot remain unsupported in C^0 under sluicing.

However, Russian allows the base-generated particle *li*, an interrogative *yes/no* question complementizer, to be a remnant of sluicing, as long as there is a focused element in SpecCP, as shown in (164).

(164) Ivan vstretil kogo-to, no ja ne znaju LENU li
Ivan met someone$_{ACC}$ but I not know Lena$_{ACC}$ li$_C$
'Ivan met someone but I don't know whether he met LENA.'

This suggests two things about Russian. First, it becomes clear that C^0 can carry *+focus* feature in Russian. Thus, there seem to be two focus positions in Russian above TP: one below CP and one in CP. The position below CP is justified by examples like (165), where the focused element follows the declarative complementizer *čto*.

(165) Maria ne znala čto IVANA ona dolžna vstrečat'.
Maria not knew that Ivan$_{ACC}$ she must meet
'Maria didn't know that it was Ivan who she was supposed to meet.'

Second, to account for the fact that Russian *li* can be a remnant of sluicing, I suggest that *li* is a clitic that cliticizes to the left, and therefore can remain in C^0 under sluicing. And the Economy considerations do not apply to *li* because it is base-generated in C^0. Thus, sluicing provides a diagnostic for the properties of certain clitics.

There is some independent evidence that Russian *li* cliticizes to the left. It comes from a construction where *li* attaches to several focused constituents, and it always follows these constituents, as shown in (166). If, however, *li* is supported by another morpheme from the left, as by *to-* in (167), then it can precede the focused constituents. The fact that, *to-* doesn't add any extra meaning to the sentence argues that we are observing a dummy 'do-support'-like process.

(166) a. Ivan **li**, Maša **li** prijedet, mne vsjo ravno.
Ivan li Maša li will-come, to-me all equal
'Whether Ivan or Maša will come, doesn't matter to me.'

b. *Ivan **li**, **li** Maša prijedet, mne vsjo ravno.
c. ***Li** Ivan, **li** Maša prijedet, mne vsjo ravno.

(167) **To-li** Ivan, **to-li** Maša prijedet, mne vsjo ravno.
'Whether Ivan or Maša will come, doesn't matter to me.'

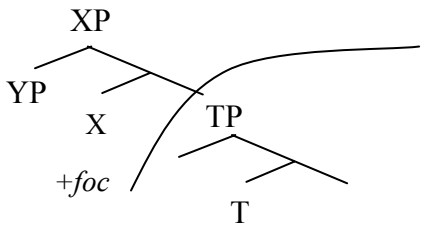
Note that the existence of (167) means that Russian *li* cannot be analyzed as merely a second position clitic.

From the data we have examined in this section, we can conclude that contrastive focus licenses sluicing in Russian and Polish. The idea that focus can license the deletion of its complement is also used in the analysis of fragment answers in English by Merchant (2004) and in Korean by Park (2005). A similar conclusion is also reached in the analysis of ellipsis in relative clauses in Hungarian by van Craenenbroeck and Lipták (2005). Thus, we can conclude that focus has an ellipsis-licensing capability in a number of languages.

3.5. Unifying the theory of licensing TP-deletion

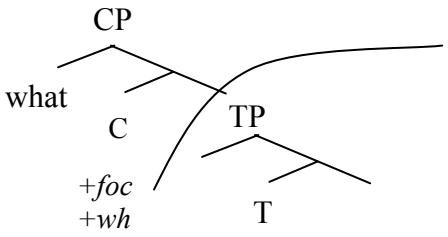
However, recall the conclusions reached by Merchant (2001) for English, namely, that it is the *+wh* feature that licenses sluicing in this language. How can we reconcile these with our conclusions reached in the previous section? Are *+wh* and *+focus* features both capable of licensing TP-deletion or is the *+focus* feature the licenser of TP-deletion in general. The latter option is the stronger one and therefore is more difficult to maintain, especially in a language like English, where contrastively focused phrases always remain in situ. However, this is the direction I would like to pursue. I propose that sluicing is licensed by the *+focus* feature with an overtly realized specifier of the head carrying this feature. This is illustrated in (168).

(168)



The +focus feature can be weak, as in English, or strong, as in Russian. If we try to unify the sluicing licensing mechanism in both types of languages, the feature strength should not matter for licensing sluicing. Given this, let us consider what the CP layer looks like in English:

(169)



What this means is that wh-movement in English simply happens to be the operation that creates the needed configuration for licensing TP-deletion. The *+wh* feature itself, however, has nothing to with licensing TP-deletion. This seems to be a promising hypothesis, especially since the environments that do not permit sluicing in English tend to contain elements that cannot be focused, such as the relative pronouns in relative clauses and complementizers like *that* and *if*.

4. Multiple sluicing and semantics of multiple interrogatives

In this section, I examine how the interpretive properties of multiple interrogatives are manifested under sluicing. Consider the contrast between (170) and (171) from Russian.

(170) Každyj priglasil kogo-to na tanec, no ja ne pomnju kto kogo.
everyone invited someone to dance but I not remember who whom
'Everyone invited someone to a dance but I don't remember who whom.'

(171) ??Kto-to priglasil kogo-to na tanec, no ja ne pomnju kto kogo.
someone invited someone to dance but I not remember who whom
'Someone invited someone to a dance but I don't remember who whom.'

The contexts which allow multiple sluicing in Russian seem to depend on the interpretation of multiple interrogatives in this language. Russian, unlike languages like Serbo-Croatian or Japanese, lacks single-pair readings in multiple interrogatives, as was discussed in Chapter 2. Let me briefly summarize the crucial facts. Multiple interrogatives in general can have a Pair-List (PL) or a Single-Pair (SP) reading, with the SP reading being more restricted crosslinguistically, as discussed by Wachowicz (1974), Hagstrom (1998), Bošković (2003) and Grebenyova (2004). The readings are demonstrated in the scenarios in (172) and (173) with respect to the English question in (174), which is infelicitous on the SP scenario in (173) since English also lacks SP readings.

(172) *Scenario 1 (PL):* John is at a formal dinner where there are diplomats and journalists. Each journalist was invited by a different diplomat. John wants to find out all the details, so he asks the host: (174)

(173) *Scenario 2 (SP):* John knows that a very important diplomat invited a very important journalist to a private dinner. John wants to find out all the details, so he asks the caterer:

(174) Who invited who to the dinner?
*PL/*SP*

Bulgarian and Russian pattern with English in lacking the SP reading in multiple interrogatives, as demonstrated in (175). Languages like Serbo-Croatian and Japanese, on the other hand, allow both PL and SP readings.

(175) a. [Bulgarian]

Koj kogo e pokanil na večerjata? *PL/*SP*

who whom Aux invited to dinner

'Who invited who to the dinner?'

b. [Russian]

Kto kogo priglasil na užin? *PL/*SP*

who whom invited to dinner

'Who invited who to the dinner?'

(176) a. [Serbo-Croatian]

Ko je koga pozvao na večeru? *PL/SP*

who Aux whom invited to dinner

'Who invited who to the dinner?'

b. [Japanese]

Dare-ga dare-o syokuzi-ni manekimasita-ka? *PL/SP*

who_{NOM} who_{ACC} $dinner_{DAT}$ invited-Q

'Who invited who to the dinner?'

Therefore, it seems plausible to analyze the degraded status of the Russian multiple sluicing example in (171) as the result of the antecedent clause imposing a single-pair reading on the interrogative clause in the sluice, since this is a reading which a multiple wh-question cannot have in Russian.[70]

There is another reading, sometimes not easily distinguished from the SP reading, namely, the *Order* reading, as in (177) from English. Multiple sluicing is

[70] For specific accounts of what prohibits SP readings in certain languages, see Chapter 3.

available with this reading in Russian if the antecedent provides the relevant context, as in (178).

(177) John and Bill were fighting. Who hit who first?

(178) *Maša i Ivan pošli na večer. Kto-to iz nix priglasil drugogo na*
Maša and Ivan went to party. One of them invited the-other to

tanec, no ja ne znaju kto kogo.
dance but I not know who whom.

'Maša and Ivan went to a party. One of them invited the other to dance but I don't know who invited who.'

Thus, we arrive at the rather straightforward generalization that the only interpretations of wh-interrogatives available under sluicing in a given language are the interpretations generally available to wh-interrogatives in that language. This presents another argument for the analysis of the sluices as full interrogative clauses.

One of the predictions of this outcome is that multiple sluicing should not be available with adjunct wh-questions since the order reading is impossible with adjuncts. The prediction is borne out, as shown in (39).

(179) *Kto-to sprjatal gde-to zdes' klad, no ja ne znaju kto gde.
someone hid somewhere here treasure but I not know who where
'Someone hid the treasure somewhere here but I don't know who hid it where.'

Another control test for the generalization above comes from Serbo-Croatian, a language allowing SP readings in multiple interrogatives. The Serbo-Croatian

equivalent, from Stjepanović (2003), of the unacceptable Russian example in (171), is fine, as expected:

(180) [Serbo-Croatian]
Neko je video nekog, ali ne znam ko koga.
somebody is seen somebody but not know who whom
'Somebody saw someone, but I don't know who whom.'

5. Superiority under Sluicing

In this section, we will examine another property of sluicing in Russian. The main generalization here is that sluicing enforces superiority effects in contexts where parallel non-elliptical structures do not exhibit any superiority effects. This was observed for Serbo-Croatian multiple sluicing in main clauses with null C^0 by Stjepanović (2003). The same is true of Russian multiple sluicing in both main and embedded clauses.

First, consider the data in (181) and (182) (slightly modified examples from Bošković (1998)), demonstrating that superiority effects in Serbo-Croatian are present in embedded but not in main clauses.

(181) a. Ko šta$_1$ o njemu govori t_1?
who what about him says
'Who says what about him?'

b. Šta$_1$ ko o njemu govori t_1?

(182) a. Pavle je pitao *ko* *šta*$_1$ o njemu govori t_1.
Pavle aux asked who what about him says
'Pavle asked who says what about him.'

b. ??Pavle je pitao *šta*$_1$ *ko* o njemu govori t_1.

However, as Stjepanović (2003) points out, superiority effects emerge in Serbo-Croatian in main clauses under sluicing:

(183) *Speaker A:* Neko voli nekog.
somebody loves somebody
'Somebody loves somebody.'

Speaker B1: Ko koga?
who whom

Speaker B2: *Koga ko?
whom who

The same effects hold under sluicing in embedded clauses in Serbo-Croatian, but that is of no relevance since this corresponds to the facts in the parallel non-elliptical structures.

Let us now examine the same contexts in Russian, a language without any superiority effects in either main or embedded clauses in non-elliptical structures, as we recall from Stepanov (1998). Like in Serbo-Croatian, superiority effects emerge in Russian under Sluicing in both main in embedded clauses, as demonstrated in (184) and (185).

(184) a. *Speaker A:* Každyj priglasil kogo-to na tanec.
everyone invited someone to dance
'Everyone invited someone to a dance.'

b. *Speaker B:* Kto kogo?
who whom

c. *Speaker B:* *Kogo kto?
whom who

(185) a. Každyj priglasil kogo-to na tanec, no ja ne pomnju *kto kogo*.
everyone invited someone to dance but I not remember who who
'Everyone invited someone to a dance but I don't remember who who.'

b. *Každyj priglasil kogo-to na tanec, no ja ne pomnju *kogo kto*.

These are rather surprising facts, given that sluicing is known to sometimes repair the derivation (e.g., amelioration of island effects under sluicing investigated by Ross (1969), Lasnik (2000) and Merchant (2001)). It is surprising that, in the cases above, sluicing seems to destroy it. Of course, if superiority effects are essentially minimality effects and minimality is encoded into the definition of Attract (Chomsky 1995), such violations cannot technically exist in any derivation and therefore cannot be repaired by deletion. This means that we would not expect superiority effects in non-elliptical structures in a language like Bulgarian to disappear under sluicing. Merchant (2001) reports data demonstrating that this is indeed the case in Bulgarian. This, as Merchant points out, presents additional evidence for the deletion approach to ellipsis, since superiority is a diagnostic of movement and movement could have taken place out of

the ellipsis site only if a full clause is present in the structure from the beginning and is deleted at PF. But why would sluicing invoke superiority effects in languages and contexts that lack superiority effects without ellipsis, as in Serbo-Croatian and Russian?

Stjepanović (2003) attempts to explain the Serbo-Croatian data as follows. Assuming that the feature licensing TP-deletion must be on C^0, she concludes that C^0 must be merged in overt syntax in sluicing constructions. The strong *+wh* feature of C^0 then triggers superiority effects in Serbo-Croatian matrix sluices.

However, it is difficult to extend this analysis to Russian. Since the *+wh* feature is weak in 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 elliptical structures, namely, quantifier parallelism.

I adopt the notion of parallelism of Fiengo and May (1994), further developed by Fox and Lasnik (2003), which requires that variables in the elided and antecedent clauses be bound from parallel positions. I also assume that the variable introduced by an indefinite in the antecedent clause is bound by existential closure (Kratzer 1997) and that wh-words like *who* and *what* are quantifiers over individuals.

Let us now consider the LF of the antecedent in Russian multiple sluicing in (186a), given in (187).

(186) a. *Speaker A:* Každyj priglasil kogo-to na tanec.
everyone invited someone to dance
'Everyone invited someone to a dance.'

b. *Speaker B:* Kto kogo [priglasil na tanec]?
who whom invited to dance

c. *Speaker B:* *Kogo kto [priglasil na tanec]?

(187) $\forall x \exists y$ [*x* priglasil *y* na tanec]
invited to dance

This is the only reading available in (186a), since surface quantifier scope is preserved in Russian. This can be seen in (188) and even more clearly in the unacceptable (189), based on an English example in Fox (2000:70). For similar observations, see also Ionin (2001), Pereltsvaig (in press), and Bailyn (2006).

(188) Kakoj-to paren' poceloval každuju devušku. $\exists x \ \forall y$ / $*\forall y \ \exists x$
some guy_{NOM} kissed every girl_{ACC}
'Some guy kissed every girl.'

(189) #Odin/kakoj-to časovoj stoit naprotiv každogo zdanija.
one/some guard is-standing in-front-of every building
'One/some guard is standing in front of every building.'

Now consider the LF representations of the acceptable sluice in (186b) and the unacceptable one in (186c), given in (190b) and (190c) respectively. Do they meet the parallelism requirement? That is, are the variables in these sluices and in the LF of the antecedent (repeated as (190a)) bound from parallel positions?

(190) a. $\forall x \exists y$ [*x* priglasil *y* na tanec] ← *LF (antecedent)*
invited to dance

b. kto *x* kogo *y* [*x* priglasil *y* na tanec] ← *LF (wh1 > wh2)*
who whom invited to dance

c. kogo *y* kto *x* [*x* priglasil *y* na tanec] ← *LF (wh2 > wh1)*
whom who invited to dance

The parallelism in variable binding is met between (190a) and (190b), but it is not met between (190a) and (190c). That is, the quantifier binding the object variable is inside the scope of the quantifier binding the subject variable in the antecedent clause, while it is outside the scope of the parallel quantifier in the sluice in (190c).

To test this further, let us scramble the object quantifier over the subject in the antecedent clause, as in (191a). This results in an acceptable sluice with the *wh2>wh1* order in (191b), as predicted by the parallelism account, since now the object quantifier is outside the scope of the subject quantifier in both the antecedent and the sluice.[71]

(191) a. *Speaker A:* Každogo$_1$ kto-to priglasil t_1 na tanec
everyone$_{ACC}$ someone$_{NOM}$ invited to dance
'Someone invited everyone to a dance.' *(with $\forall x \exists y$)*

b. *Speaker B:* Kogo kto?
whom who

[71] The universal quantifier is used as the object here to maintain the pair-list reading requirement in Russian multiple interrogatives.

c. *Speaker B:* *Kto kogo?
who whom

And the subject>object order of the wh-phrases in (9c) is unacceptable now, which strengthens the parallelism account proposed above.[72]

Thus, the source of the apparent superiority effects under sluicing in Russian turns out to be parallelism and not minimality.

The next step is to see if this analysis can be extended to Serbo-Croatian, the language exhibiting similar effects under sluicing. Unfortunately, there is an interfering factor in Serbo-Croatian. According to Sandra Stjepanović (p.c.), scrambling an object over the subject prohibits sluicing all together in Serbo-Croatian. This is true even with single sluicing, as can be seen in (192).

(192) *Speaker A:* Nekog je Petar volio.
somebody$_{ACC}$ is Petar$_{NOM}$ loved
'Petar loved somebody.'

Speaker B: *Koga?
whom

[72] Steven Franks (p.c.) reports of a Russian informant who does not share the judgments in (191). The same informant, however, is sensitive to superiority effects in Russian (i.e., not allowing the lower wh-phrase to be fronted over the higher one even in non-elliptical contexts.) As Merchant (2001) shows for Bulgarian, a language with robust superiority effects, such effects do not go away under sluicing if they are present in non-elliptical contexts. Thus, parallelism and superiority are independent properties of grammar and can be distinguished from each other under ellipsis only if a speaker is insensitive to superiority in non-elliptical contexts (as my Russian informants and myself are). The attested variation with respect to superiority effects is itself an interesting puzzle for syntactic theory and is in need of further exploration.

Thus, running the diagnostic with scrambling, as in Russian (9), is problematic in Serbo-Croatian. When I attempted to run it with a number of Serbo-Croatian speakers, as in (193), the judgment was as expected: scrambling does improve the *wh2>wh1* order in the sluice but it does not make it perfect.

(193) *Speaker A:* Nekog neko voli

somebody$_{ACC}$ someone$_{NOM}$ love

'Petar loves somebody'

Speaker B: ?/??Koga ko?

whom who

Although, identifying of the source of the mysterious effect in (192) is beyond the scope of this chapter, I will point out a few directions for further research. One plausible direction would be to identify the position where the scrambled indefinite moves in the antecedent clause and the position where the wh-phrase moves in the sluice. These positions might be different in such a way that the parallelism is violated.

Another potential source of this effect is the specificity effect produced by scrambling in Serbo-Croatian, as brought to my attention by Sandra Stjepanović (p.c.). It is known that an indefinite that is a correlate of the remnant of sluicing already has a specificity requirement on it. That is, it is already interpreted as specific. Now, if scrambling an indefinite object over the subject has its own specificity effect in Serbo-Croatian, it might be incompatible with sluicing, where the indefinite is

already specific to start with. Of course, this matter needs more exploration before a more solid conclusion can be reached.

6. Summary

To summarize, we have examined how the syntactic and semantic properties of multiple interrogatives are manifested in sluicing and reached the following results.

First, given the movement of wh-phrases to a focus position in Russian and Polish, I proposed that contrastive focus licenses TP-deletion deletion in these languages. As a correct prediction of this proposal, I showed that contrastively focused R-expressions can also be the remnants of sluicing in Russian and Polish. I further extended this analysis to English by arguing that wh-movement to SpecCP only gets a potential remnant of sluicing into the right position (the specifier of the projection carrying +*focus* feature) and it is the +*focus* feature with the overt material in its Spec that licenses sluicing.

Second, we have seen that sluicing licensing contexts depend on the interpretation of multiple interrogatives in a given language. That is, sluicing is prohibited in Russian if an antecedent imposes the SP reading on the interrogative in the sluice, just as non-elliptical multiple interrogatives are unacceptable under the SP reading in this language.

Finally, considering the quantifier parallelism requirement in ellipsis allowed us to analyze apparent superiority effects under sluicing as parallelism effects. That is, the unacceptability of certain sluices is caused by the lack of parallelism in quantifier-variable binding between the antecedent and the sluice. This analysis

provides a prediction for further research, namely, that there is no language with fixed isomorphic scope that allows for free ordering of wh-phrases under sluicing.

This concludes our theoretical investigation in the syntax and semantics of multiple interrogatives. The next chapter explores acquisition of the syntactic and semantic properties of multiple interrogatives by Russian-, English-, and Malayalam-speaking children.

Chapter 5: Acquisition of Multiple Interrogatives

1. Introduction

Having examined the syntactic and semantic properties of multiple interrogatives in the previous chapters, let us now explore how children acquire these properties. Because the majority of the studies on acquisition of interrogatives focus only on single interrogatives, my goal here is to make the initial steps in approaching the learning issues in multiple interrogatives. I specifically aim to (i) find out how much evidence for the syntactic and semantic properties of multiple interrogatives children get in the linguistic input; (ii) investigate at what age children exhibit the knowledge of the language-specific syntax and semantics of multiple interrogatives; and (iii) develop certain explanations of how the learners acquire that knowledge on the basis of the available evidence in the input.

Consider the basic examples of multiple interrogatives from English, Russian, and Malayalam in (194) – (196).

(194) a. Who did John give what?
b. *John gave who what?
c. *Who what did John give?

(195) a. Komu čto Ivan dal? *Russian*
who$_{DAT}$ what$_{ACC}$ Ivan gave
'Who bought what?'

b. *Komu Ivan dal čto?
who$_{DAT}$ Ivan gave what$_{ACC}$

c. *Ivan dal komu čto?
Ivan gave who$_{DAT}$ what$_{ACC}$

(196) Gibu aaRko-okke ent-okke kodutu? *Malayalam*
Gibu who-each what-each gave
'Who did Gibu give what?'

Already, from these simple mono-clausal multiple wh-questions, we can observe some major points of syntactic variation across languages that a learner of these languages must acquire. While only one wh-phrase is fronted in English, all wh-phrases are fronted in Russian, and no wh-phrase needs to be fronted in Malayalam.[73]

Moreover, there is some cross-linguistic variation with respect to semantics of multiple interrogatives. As we discussed in Chapters 2 and 3, languages differ in whether they allow SP readings in multiple interrogatives.

These language-specific syntactic and semantic properties must be acquired by the children on the basis of the available evidence in the input. Hence, I examine the availability of the relevant positive evidence in the input by conducting a search of the parental speech in CHILDES database for the utterances containing multiple interrogatives. I report the results of this search in Section 2, concluding that there is very little direct evidence in the input that can be used by the child in order to acquire language-specific syntactic and semantic properties of multiple interrogatives. This

[73] Malayalam, like other wh-in-situ langauges, allows scrambling, so alternative orders of the wh-phrases are possible but are irrelevant for our purposes since the operations producing those orders apply to non-wh-elements as well as wh-phrases.

makes the area of multiple interrogatives a rather fruitful area for the study of language acquisition, since it should allow us to observe what hypotheses language learners make on the basis of the available input and what evidence they rely on in the process of acquiring their target grammar.

In Section 3, I present the new methodology for the elicitation of multiple interrogatives and report the results of a study eliciting multiple interrogatives from English- and Russian-speaking children and adults. The contexts in which the subjects produced multiple interrogatives indicated what interpretation they assigned to those constructions, and their utterances themselves allowed me to examine the syntactic structure underlying these expressions.

While no deviations were found in children's semantic knowledge of multiple interrogatives in either English or Russian, the syntactic behavior of Russian-speaking children was somewhat different from that of Russian-speaking adults. Unlike adults, Russian-speaking children (mean age 4;7) left some bare wh-phrases in situ in multiple wh-questions.

To investigate this finding, I conducted a follow-up experiment on Russian, described in Section 4. In this experiment, the original finding that Russian-speaking children go through a phase of not fronting all wh-phrases was confirmed in a large variety of syntactic contexts. I explore two possibilities for the underlying syntactic structure that Russian children assign to multiple wh-questions. The first is based on the syntax of contrastive focus and the second relates children's behavior to the asymmetry between bare and complex wh-phrases in Slavic, namely, that complex wh-phrases, unlike bare wh-phrases, can remain in situ in these languages.

In Section 5, I present the results of a parallel study on Malayalam, a wh-in-situ language allowing SP readings in multiple wh-questions. Malayalam-speaking children produced multiple interrogatives in both PL and SP scenarios. I propose a learning algorithm for the PL/SP readings, where children rely on an independent property of language, namely, existence of an independent Focus projection above TP in their language. The positive evidence available to learners consists then of the information about the distribution of the focus morphemes, complementizers, and focus-fronted expressions. This algorithm is consistent with the theoretical analysis of semantics of multiple questions given in Chapter 3. Section 6 is a summary of the overall results.

2. Multiple interrogatives and the nature of the input

2.1. What is there to learn?

There have been many studies conducted on the acquisition of single interrogatives (i.e., interrogatives with a single wh-phrase), like the one in (197a). The point of cross-linguistic variation in these constructions has to do with the obligatoriness of the overt wh-movement to the left periphery of the clause. In English (197a), *what* moves from the position of the object of the verb *buy* to the clause initial position known as SpecCP. This movement is obligatory, as shown by the unacceptability of (197b) and (197c).

(197) a. What did John buy *t*?
b. *Did John buy what?

c. *John bought what? *(as a non-echo question)*

In other languages, like Japanese, Chinese, Malayalam, overt wh-movement does not take place, as shown in (198). Of course, scrambling is available in Japanese, making wh-fronting possible but this is irrelevant for our purposes since the operations producing those orders apply to non-wh-elements as well as wh-phrases. The crucial point here is that movement of a wh-phrase is not required in this language, unlike in English.

(198) John-wa nani-o kaimasita ka? *Japanese*
John$_{NOM}$ what$_{ACC}$ bought Q
'What did John buy?'

The studies of Clahsen, Kurasawe and Penke (1995), Santelmann (1998), Guasti (2000), and Seidl et. al. (2003), among others, show that the parameter with respect to wh-movement in single wh-questions is set by the time the child begins producing wh-questions (by the age of 1;8).

However, multiple interrogatives (i.e., interrogatives with more than one wh-phrase) involve additional layers of parameterization. Thus, additional learning issues arise. Languages employ three syntactic strategies with respect to formation of multiple wh-questions. In some languages, only one wh-phrase is fronted in such questions, as in English. In others, all wh-phrases are fronted, as in Russian. And there are also languages in which no wh-phrases are fronted, as in Japanese. This is demonstrated for all three varieties of languages in the examples (199)-(201).

(199) a. What did Smurf put *t* where? *English*

b. *What where did Smurf put?

(200) a. Čto kuda Ivan položil *t t*? *Russian*

what where Ivan put

'What did Ivan put where?

b. *Čto Ivan položil *t* kuda?

(201) Smurf-wa dokoni nani-o oitano? *Japanese*

Smurf where what put-Q

'What did Smurf put where?

These facts pose questions as to when and how these language-specific properties are acquired by children.

In addition to syntactic variation in multiple interrogatives that we just observed, languages also vary in semantics of these constructions. Multiple interrogatives can potentially have a pair-list (PL) or a single-pair (SP) reading. This was extensively discussed in Chapters 2 and 3; therefore I will only briefly demonstrate the two readings. The question in (203) has the PL reading and is felicitous in the scenario in (202). An expected response to such a question involves a list of propositions with ordered pairs as in (204).

(202) *PL Scenario:* John is at a formal dinner where there are diplomats and journalists. Each journalist was invited by a different diplomat. John wants to find out all the details, so he asks the host:

(203) Who invited who to the dinner?

(204) Mr. Smith invited Mr. Jones, Ms. Black invited Mr. Green, etc.

A scenario corresponding to the SP reading is given in (205). English lacks SP readings in questions with bare wh-phrases. In SP contexts, English speakers use either a conjoined question (e.g., *Who invited somebody to the dinner and who did they invite?*) or a question with complex wh-phrases, where the SP reading is available in English, as in (206). A felicitous response to a single-pair question is given in (207).

(205) *SP Scenario:* John knows that a very important diplomat invited a very important journalist to a private dinner. John wants to find out all the details, so he asks the caterer:

(206) Which diplomat invited which journalist to the dinner?

(207) Mr. Black invited Ms. Smith.

The distribution of the SP readings is subject to cross-linguistic variation, as reported in Hagstrom (1998), Bošković (2003) and Grebenyova (2004). Recall the paradigm from Chapter 3, repeated below.

(208) a. *PL/*SP*

Who invited who to the dinner? *English*

b. *PL/*SP*

Koj kogo e pokanil na večerjata? *Bulgarian*
who whom Aux invited to dinner
'Who invited who to the dinner?'

c. *PL/*SP*

Kto kogo priglasil na užin? *Russian*
who whom invited to dinner
'Who invited who to the dinner?'

d. *PL/*SP*

Quem convidou quem para (o) jantar? *Brazilian Portuguese*
who invited whom to (the) dinner
'Who invited who to (the) dinner?'

(209) a. *PL/SP*

Ko je koga pozvao na večeru? *Serbo-Croatian*
who aux whom invited to dinner
'Who invited who to the dinner?

b. *PL/SP*

Dare-ga dare-o syokuzi-ni manekimasita-ka? *Japanese*
who-Nom who-Acc dinner-Dat invited-Q
'Who invited who to the dinner?

c. *PL/SP*

Hver bauð hverjum í veisluna? *Icelandic*
who invited whom in the-dinner
'Who invited who to the dinner?'

This paradigm demonstrates that, while multiple questions in English, Bulgarian, Russian, and Brazilian Portuguese have only PL reading, the corresponding multiple questions in Serbo-Croatian, Japanese, and Icelandic have both PL and SP readings. This cross-linguistic variation raises questions as to when and how this kind of semantic knowledge is acquired by children.

Thus, there are aspects of syntax and semantics of multiple interrogatives which need to be acquired by children on the basis of the evidence available in the input and it will be our goal to explore how this acquisition proceeds.

2.2. Previous Studies

Compare to the studies on the children's acquisition of single wh-questions, studies on acquisition of multiple wh-questions are quite rare. Roeper and de Villiers (1991) and Yamakoshi (2002) conducted studies on the acquisition of pair-list readings in questions containing a wh-phrase and a universal quantifier. Such questions can be ambiguous between the PL reading and the *group* reading, as demonstrated in (210a).[74]

(210) a. What did everyone take *t*? *Group/PL*
b. Who *t* took every vegetable? *Group/*PL*

The goal of these studies was to find out whether children know the constraint on the availability of the PL reading in structures where the universal quantifier is in a lower position than the wh-phrase, as in (210b).

These constructions, however, are different from multiple wh-questions in that they contain only one wh-phrase, which does not allow us to investigate children's knowledge of syntax of questions with multiple wh-phrases. Moreover, structures of this kind disallow SP readings due to an interfering factor, namely, the presence of

[74] In Chapter 3, we have referred to the Group reading as the Individual reading, based on the nature of the expected answer to such a question which contains only a single individual instead of a list of pairs of individuals. Here, to avoid confusion, I am using the terminology used in the acquisition literature.

the universal quantifier. The learning issues we set out to explore are independent of those involved in questions with the universal quantifier. Therefore, a new study seems to be needed for our purposes: a study that targets questions with multiple wh-phrases.

2.3. Evidence in the Input

Before we turn to the experiments testing what children actually know, let us first explore how much positive evidence is available to children in the linguistic input. In order to find out how frequent multiple questions are in the adult speech, I conducted a search of the CHILDES database for single and multiple wh-questions in the parental speech in Russian and English.

First, I searched the corpus for the Russian-speaking child Varvara, which contains 7 recorded sessions between Varvara's ages 1;7 - 2;11. In this corpus, I found 138 single questions containing the wh-phrase *kto 'who'*, 412 single questions containing *čto 'what'* and 147 single questions containing *kak 'how'*. The total for the questions with these wh-phrases is 697. What about multiple interrogatives with any combination of those three wh-phrases? I found only one multiple interrogative, given in (211).[75]

[75] The question was asked by Varvara's mother and it referred to Varvara's birthday party. Varvara struggled to answer it. She answered only after her mother changed it into a series of single questions about what each guest at Varvara's party gave her for her birthday.

(211) Kto tebe čto podaril *t*? *Russian*

who$_{NOM}$ you$_{DAT}$ what$_{ACC}$ gave-as-present

'Who gave you what?

In addition to wh-movement of *kto*, the indirect object *tebe* in (211) has undergone scrambling to a position between the two wh-phrases. Scrambling is a common process in Russian. This raises a question as to what position the wh-phrase *čto* occupies in the structure. Given that Russian is an SVO language, it is clear that *čto* has undergone some movement to a pre-verbal position. Whether this is an instance of focus movement or of scrambling is not crucial for our purposes. The crucial point is that *čto* is not in situ, unlike its English counterpart in the translation of (211), and that is something a child has to learn.

In the English CHILDES, I searched the first 5000 wh-questions in the parental speech, and found only 3 multiple interrogatives.

As we can see from the drastic difference in the rate of occurrence of single vs. multiple questions in the parental speech, the linguistic input that the child receives provides much less positive evidence for the acquisition of multiple wh-questions, as compared to single wh-questions. This presents a learning puzzle as to how children converge on the correct adult grammar.

It is important to point out that the reported results based on CHILDES must be taken with a grain of salt since the discourse situations recorded there typically involve a dialogue between a single child and a single adult. Such settings are hardly compatible with scenarios needed for multiple interrogatives to be produced, especially in Russian and English, where the interpretation of bare multiple wh-

questions is restricted to PL readings. PL readings require pairs of individuals, which is most easily done when the more individuals are participating in the discourse. Consider a scenario where John takes out his three children for ice-cream, so it is natural for him to ask *Who wants what?* However, it is not natural for John to ask this question if he takes out only one child for ice-cream. CHILDES, unfortunately, is based on the situations of the latter and not the former type.

However, until we confirm that there are a lot of multiple wh-questions in the input children get, we need to look for the independent observable properties of the language from which learners could deduce the grammar of multiple interrogatives. There seem to be two possibilities: either this sparse direct evidence from multiple wh-questions is enough for children to eventually acquire their properties, or these properties are determined by other observable properties of the language. After examining the results of the experiments showing what children actually know, I will discuss in some detail what those independent observable properties that guide the learners toward the grammar of multiple interrogatives might be.

Thus, we arrive at the following set of questions we will attempt to answer with respect to multiple interrogatives:

(212) a. At what age do children acquire language-specific syntactic properties of multiple interrogatives?
b. At what age do children acquire language-specific semantic properties of multiple interrogatives?
c. How do they come to know those properties, given the nature of the input?

3. Eliciting Multiple Interrogatives

3.1. Experimental Schema

As we have observed in the previous section, multiple questions are not very frequent in spontaneous speech. Therefore, analyzing spontaneous speech of children would not be productive for our purposes. Truth Value Judgment Task would not be useful here either because only propositions can be evaluated with respect to truth values. Interrogatives, on the other hand, are sets of propositions and cannot be true or false. That is why Elicited Production Task, based on the one developed in Thornton (1990), was selected in order to unveil children's competence in syntax and semantics of multiple interrogatives. The produced utterances will allow us to examine children's syntactic knowledge, while eliciting those utterances in controlled PL and SP contexts will allow us to examine children's semantic knowledge.[76]

The experimental schema is as follows. Kermit, the puppet, is learning how to be a magician and must guess about what happened in a story without watching the story. The stories are acted out with toys. Kermit is blind-folded and hides under the table during the relevant parts of the stories. After each story, the experimenter gives a lead-in prompting the subject to ask Kermit a question about the story. There are stories with PL and SP contexts.

The stories were designed in such a way as to prompt the subject to produce questions that are felicitous in certain contexts and are of the syntactic form that is relevant to our study. That is, it was important to provide the subjects with contexts

[76] Under the analysis developed in Chapter 3, where the availability of SP readings in multiple questions is the consequence of the syntactic selection restrictions of the Q-morpheme, we would actually be examining children's syntactic knowledge in both aspects of the study.

supporting the PL and SP readings of multiple interrogatives and ensure as much as possible that they use multiple wh-questions and not some other constructions compatible with those contexts.

Let me demonstrate this with the PL context for eliciting the question in (213).

(213) Who hid what?

In this context, we have three characters each hiding a different object and one character who does not hide anything, as in Figure 1.

Figure 1. Pair-list Context (*Who hid what?*)

Besides the actual question we would like to elicit (*Who hid what?*), there are other utterances that are felicitous in this context. One possible utterance is a question with complex wh-phrases, where both or one of the wh-phrases is a complex one, as shown in (214).

(214) a. Which x hid which y?

b. Who hid which x?

c. Which x hid what?

This potentially complicates the picture since complex wh-phrases behave differently syntactically and semantically, as was discussed in Chapter 3. In order to reduce the possibility of such utterances being produced, it was important to choose the characters and the objects that do not constitute some obvious category. For example, they could not be all animals or all fruits, so that the subject could not easily refer to the obvious category by saying *which animal* or *which fruit*.

Another type of utterance that is felicitous in the PL context is a question with a single wh-phrase and the universal quantifier:

(215) What did everyone hide?

To avoid this type of utterance, an extra character was added to the story who does not hide anything. In addition, it was brought to the attention of the subject in the lead-in right before the question is elicited that not everyone hid something. Of course, this cannot completely eliminate the possibility of such utterances since the subject can set the domain of *every* to be the individuals who did hide something, but at least it can reduce the number of such utterances.

Another interfering utterance turned out to be the question involving the pronoun *they*, as in (216).

(216) What did they hide?

To reduce the possibility of getting such responces, it was ensured that the names of the characters are not mentioned in the lead-in to the question, so that a subject could not easily refer to those with the 3rd person plural pronoun.

Yet another felicitous way to ask about the PL reading context is by producing a single wh-question, such as *What did Snow White hide?,* wait for the puppet's answer, then ask about the next character *What did the Rabbit hide?* and so on. Another version of this is a series of conjoined single questions about each character without waiting for the answer in between each question, as in (217).

(217) What did Snow White hide, what did Rabbit hide and what did Horse hide?

To avoid that and increase the number of multiple interrogatives in the utterances, the preceding story was targeting a single wh-question and the puppet gave the correct response to that question. Therefore, the experimenter and the subject decide to ask a more difficult question next time. In addition, the lead-in is also designed to prompt for a multiple and not a single wh-question.[77]

Now, consider how the all of these elements of design work together. The four characters and the objects to be hidden are introduced. The puppet is blindfolded and hides under the table. Three characters each hide a different object behind them. The fourth character considers hiding something but decides not to hide anything after all. The puppet comes back. The experimenter presents the lead-in, which is addressed to the puppet as a clue about the story. It is given in (218).

[77] The series of single questions were still produced by a number of subjects and were often used by particular children who did not produce any multiple wh-questions, as what seemed to be a last resort strategy. I will return to this issue with more details in section 5.

(218) Kermit, we can tell you that the dog didn't hide anything. But the rest of them hid something and each hid a different thing. Now JOHNNY will ask you about it.

The last sentence of the lead-in prompts the child to ask a question in an indirect way by telling the puppet that the child will ask him a question now. It proved to be more effective than prompting the child directly.

PL contexts were interchanged with SP contexts, where only one character hid one particular object, as in Figure 2 below. The lead-in in such contexts was of the form in (219).

Figure 2. Single-pair Context (*Who hid what?*)

(219) Kermit, we can tell you that someone hid something here. Now JOHNNY will ask you about it.

There were also two warm-up stories eliciting single wh-questions (one subject wh-question and one object wh-question) and the fillers on single subject and object wh-questions after each story.

3.2. Experiment 1

The idea was to conduct this study in languages that employ different syntactic strategies in multiple question formation. English and Russian were selected for this reason, where English is a language with single wh-fronting and Russian is the language with multiple wh-fronting. It is also important to examine languages that have only PL readings in bare multiple questions and compare the results with those from the corresponding study in languages that have both PL and SP readings in these contexts. Both English and Russian share the property of allowing only the PL readings in bare multiple interrogatives. I also conducted a study on Malayalam, which is different from Russian and English both syntactically and semantically in that it is a wh-in-situ language and it allows both PL and SP readings in bare multiple interrogatives. The Experiment 1 was devoted to English and Russian.

The participants were 20 English-speaking children (ages 3;7–6;2, mean 4;9), 20 Russian-speaking children (ages 3;5–6;5 mean 4;7), and 20 adult controls for each group. There were 2 test stories per subject, with the target questions *Who hid what?* and *Who won what?*. The stories were given in PL and SP contexts, mixed with fillers, which were targeting single subject and object wh-questions.

Since this was the first study using this particular methodology, one of the main results was that we actually got quite a number of multiple interrogatives from

children and adults, suggesting that this methodology is on the right track. The distribution of multiple wh-questions in PL and SP contexts in children and adults is given in Tables 1 and 2 respectively.

Table 1. Production of multiple interrogatives (Children)

	PL	**SP**
ENG	32%	0
RUS	45%	0

Table 2. Production of multiple interrogatives (Adults)

	PL	**SP**
ENG	38%	0
RUS	50%	0

The effect of context is clear from these results: neither adults nor children produced multiple interrogatives in SP contexts. Children were uniformly producing single wh-questions in SP contexts (e.g., *What did Snow White hide?*). Adults produced some single questions as well but they also produced conjoined questions in these contexts (e.g., *Who hid something and what did they hide?*). These numbers reflect the production of questions containing only bare wh-phrases. Adults also produced some multiple questions with complex wh-phrases but those were not counted since they are ambiguous between PL and SP readings. Interestingly, children never produced complex wh-phrases when bare wh-phrases were elicited.

Thus, it seems clear that by the age 4;9, both English- and Russian-children exhibit perfect knowledge of the interpretive possibilities of multiple interrogatives in

their language. However, multiple interrogatives themselves are not very frequent in the parental speech, as we saw in the previous section. What evidence in the input do children use in learning these facts? I will discuss some possibilities in Section 4 after presenting the corresponding results from Malayalam.

Because in this initial experiment there were only two test stories per subject, it is difficult to see the effect of age statistically. Below is a graph suggesting there is some effect of age, however, we will see a much clearer picture in the results of the follow up experiment, where there were 5 test stories per subject.

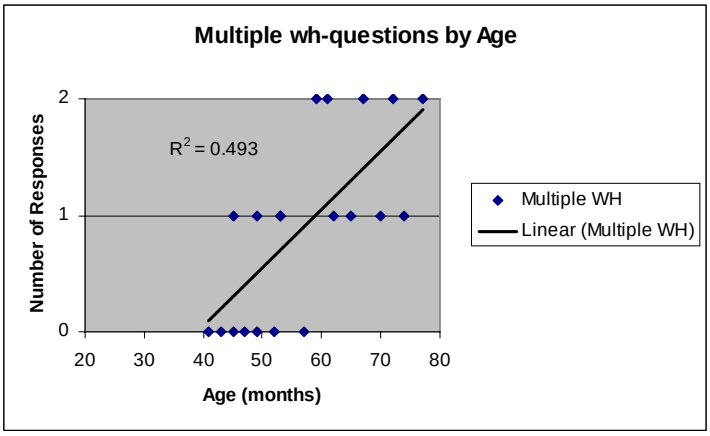


Figure 3. Multiple Questions of Russian-speaking children by age

As for the syntax of multiple interrogatives, the questions produced by the English-speaking children had adult-like syntax throughout: the first wh-phrase was always fronted, while the second one remained in situ. Based on this result, we can conclude that by the age of 4;9, children's knowledge of the syntax of multiple questions in English matches that of adults.

However, Russian-speaking children exhibited certain deviations from the syntax of multiple interrogatives in adult Russian. Specifically, 15% of the time, Russian children produced questions with only one wh-phrase fronted and the other wh-phrase remaining in situ, as in (220). This never occurred in the utterances of adults and is unacceptable in Russian.[78]

(220) *Kto sprjatal čto? *Russian*
who hid what
'Who hid what?'

Russian children's production of wh-in-situ raises certain learnability questions. What syntax do Russian-speaking children assign to multiple wh-questions and why is it different from the adult syntax? I will develop potential answers to these questions in section 4. Before we proceed to that, notice that the target multiple wh-questions in Experiment 1 were all of the subject-object type. Now that we discovered that Russian children sometimes leave a bare wh-phrase in situ, it is important to test other contexts where the higher wh-phrase is not a subject. By doing so, we would confirm the validity of the finding in Experiment 1 and obtain more data from more varied wh-contexts. Thus, I conducted a follow-up experiment on Russian, which is described in the next section.

[78] Both Russian-speaking children and adults sometimes violated Superiority (i.e., fronting the lower wh-phrases over the higher ones, but that is a general property of Russian and therefore this behavior is not surprising. It only further confirms the observation from the theoretical literature.

4. Experiment 2: Child wh-in-situ in Russian

In this experiment, several contexts with non-wh-subjects were added in order to determine where exactly the higher and the lower wh-phrases are located. I also added an argument/adjunct asymmetry with respect to the lower wh-phrase, since it is an important linguistic contrast, especially when it comes to wh-in-situ (i.e., adjuncts tend to resist being left in situ).

The participants were 20 Russian-speaking children (ages 4;1-6;3, mean 5;2) and 20 adult controls. The type and the number of target questions were as follows.

(221) a. 1 subject-object question, as part of the warm up (i.e., *Who hid what?*)
b. 2 double-object questions (i.e., *Who did Lizard give what?)*
c. 2 direct object – adjunct questions (i.e., *Who did the dog find where?*)

4.1. Results

The rate of production of multiple questions by children vs. adults is given in Table 3 and Figure 4 below. There is no significant difference between the two groups in this respect: $t(38) = 0.51$, $p = 0.62$.

Table 3. Multiple interrogatives of Russian-speaking children vs. adults

	Multiple Qs
Adults	(56) 56%
Children	(60) 60%

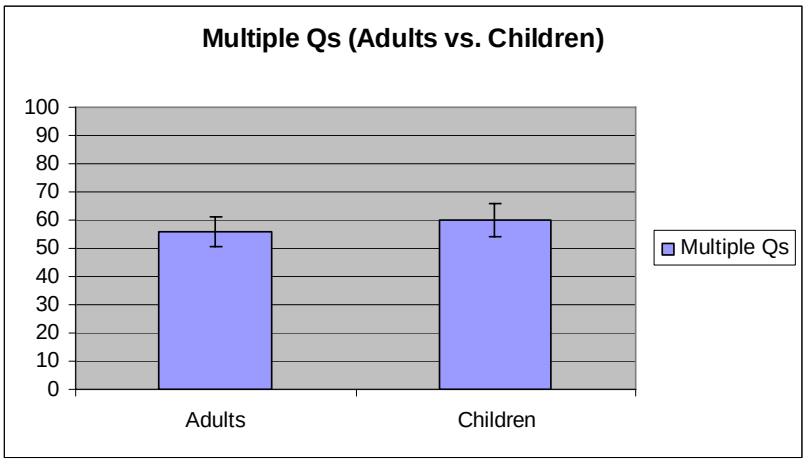


Figure 4. Multiple interrogatives of Russian-speaking children vs. adults

Let us now consider the distribution of the different types of wh-fronting in multiple questions produced by Russian children vs. adults, which is shown in Table 4.

Table 4. Distribution of types of wh-fronting in Russian-speaking children vs. adults

	Multiple wh-fronting	Partial wh-fronting	Wh-in-situ
Adults	(37/56) 66%	(19/56) 34%	(0) 0%
Children	(33/60) 55%	(16/60) 27%	(11/60) 18%

There were three specific patterns produced with respect to the position of wh-phrases in multiple interrogatives. The first pattern was where all wh-phrases are fronted to the left periphery of the clause, as in (222), produced at almost the same rate by children and adults: $t(38) = 0.68$; $p = 0.5$.

(222) a. Komu čto jaščerica podarila?
who$_{DAT}$ what$_{ACC}$ lizard gave-as-present
'Who did the lizard give what?'

b. Kogo gde sobaka našla?
who$_{ACC}$ where dog$_{NOM}$ found
'Who did the dog find where?'

In addition to fronting all wh-phrases, children and adults also produced questions where the first wh-phrase was fronted completely to the left edge of the clause while the second wh-phrase was only fronted to the immediately preverbal position, as in (223). I refer to this pattern as *Partial Wh-fronting*.

(223) a. Komu jašjerica čto podarila?
who$_{DAT}$ lizard what$_{ACC}$ gave-as-present
'Who did the lizard give what?'

b. Kogo sobaka gde našla?
who$_{ACC}$ dog$_{NOM}$ where found
'Who did the dog find where?'

This is an acceptable pattern of wh-movement in adult Russian. Recall that we even observed partial wh-fronting in the example from parental speech in (211). Because children and adults behave similarly with respect to this pattern, I conclude that it does not create any learnability issues in need of explanation.

Partial wh-fronting does, however, have theoretical consequences for the overall theory of multiple wh-fronting. Since Rudin (1988), it is often claimed that

Slavic languages require fronting of all bare wh-phrases at least as high as TP. Partial wh-fronting in Russian seems to suggest something slightly different. At least for Russian, the generalization is that bare wh-phrase cannot remain in situ but they do not have to both be fronted all the way. One of the wh-phrases can be fronted partially, to the immediately preverbal position.

The position a partially-fronted wh-phrase occupies in the structure could be the position to which contrastively focused elements move. This means that there is a focus position between TP and *v*P in Russian. This is confirmed by the fact that contrastively focused R-expressions in Russian can occur in the same position, as in (224).

(224) Sobaka ZA DEREVOM našla ego.
dog$_{NOM}$ behind tree found him
'The dog found him BEHIND THE TREE.'

Such focus position above *v*P has been identified in other languages as well. Izvorski (1993) shows some evidence for its existence in Bulgarian and Jayaseelan (1999) argues for its existence in Malayalam.

Partial wh-fronting in Russian is also supports the view of focus-driven wh-fronting in Slavic, as developed in Bošković (1998, 2002a) and Stjepanović (1998).

An important result of Experiment 2 is that Russian-speaking children once again produced a number of wh-questions with the second bare wh-phrase remaining in situ, as in (224). In the previous experiment, they produced those structures 15% of the time and, in the follow-up experiment, they produced it 18% of the time.

(225) a. Komu jašjerica podarila čto?
who$_{DAT}$ lizard gave what$_{ACC}$
'Who did the lizard give what?'

b. Kogo sobaka našla gde?
who$_{ACC}$ dog$_{NOM}$ found where
'Who did the dog find where?'

There was no age effect found across children with respect to the rate of production of either multiple wh-questions in general or wh-questions with wh-in-situ. This is shown in the graphs below.

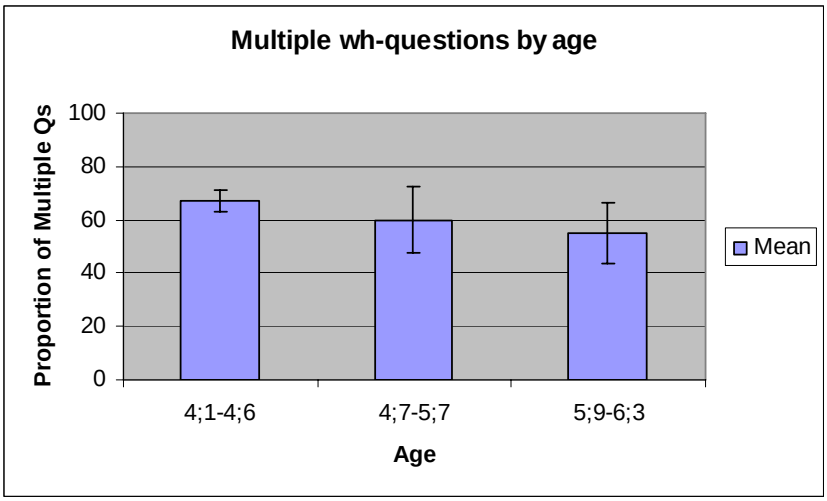


Figure 5. Multiple wh-questions by age

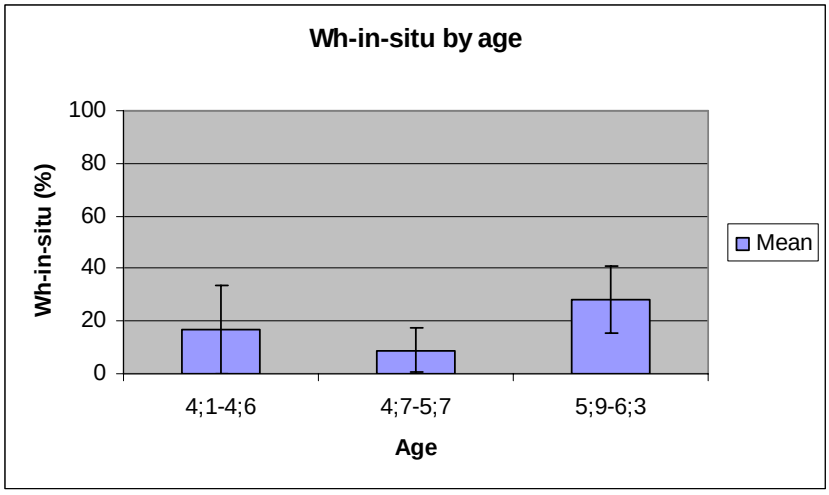


Figure 6. Wh-in-situ by age

Thus, our finding from Experiment 1, namely, that Russian-speaking children go through a stage of producing non-adult-like questions with wh-in-situ, was confirmed in Experiment 2. I propose an explanation of these results next.

4.2. The source of Russian child wh-in-situ

I propose two possibilities for the syntactic representation Russian-speaking children assign to their questions containing wh-in-situ. The first possibility is that child wh-in-situ in Russian is a 'side-effect' of children acquiring the properties of contrastive focus in this language. Contrastive focus plays an important role in multiple wh-questions in Slavic, as argued extensively by Bošković (1998, 2002a), Stjepanović (1998), and Stepanov (1998), among others. We have previously mentioned this analysis with respect to partial wh-fronting.

Contrastively focused R-expressions cannot remain in situ in Slavic, as shown in Russian (226).[79]

(226) a. IVANA ja uvidela *Russian*
Ivan$_{ACC}$ I saw
'I saw IVAN'

b. ??Ja uvidela IVANA

Likewise, bare wh-phrases cannot remain in situ in Slavic. Based on this correlation, Bošković (1998, 2002a), Stjepanović (1998) and Stepanov (1998) analyze multiple wh-fronting in Slavic as a result of the wh-phrases undergoing focus movement. Under this theory of wh-fronting in Slavic, the children can rely on the positive evidence from contrastively focused R-expressions in acquiring how many wh-phrases must be fronted in multiple questions in their language. It is plausible then that Russian children's wh-in-situ is a result of either their not having acquired the fact that wh-phrases behave like contrastively focused R-expressions in Russian, or that contrastive focus in Russian is such that it prohibits the focused expressions to stay in situ. We can tease these two options apart by further testing children's knowledge of the grammar of contrastively focused R-expressions in Russian and whether the course of its acquisition is parallel to that of multiple wh-questions. If it is parallel, it would suggest that the former option is more plausible, and if children

[79] The degree of badness of (226b) varies among Russian speakers, but most speakers get some contrast between (226a) and (226b).

acquire contrastive focus fronting before they acquire multiple wh-fronting, that would argue in favor of the latter option.

An alternative source (and, quite possibly, an additional source) of child wh-in-situ is related to the asymmetry between complex and bare wh-phrases in Slavic. Unlike bare wh-phrases, complex wh-phrases are optionally multiply fronted in Slavic. That is, only one complex wh-phrase must be fronted; the other may remain in situ, as demonstrated in (227).

(227) a. Kakuju knigu Ivan dal kakomu studentu?
which book Ivan gave which student
'Which book did Ivan give to which student?'

b. Kakuju knigu [kakomu studentu]$_1$ Ivan dal t_1?

Child wh-in-situ may then be also a result of "confusing" evidence children are getting in the input and having to sort out which wh-elements obligatorily front and which may remain in situ.

The bare vs. complex wh-asymmetry with respect to obligatoriness of fronting in itself presents a learnability puzzle, given how rare multiple wh-questions are in the input compare to single questions. The asymmetry is not well understood yet in the theoretical literature. In Grebenyova (In press), I suggest that an idea of Boeckx and Grohmann (2004) can be used to explain this asymmetry. They argue that D-linking is very closely related to scrambling. Adopting this view, and extending not only to D-linked wh-phrases, but to complex wh-phrses in general, seems to provide an answer to the learnability issue in question. That is, Russian children can rely on

the positive evidence from scrambling in acquiring the grammar of complex wh-phrases. And the syntax of multiple wh-fronting of bare wh-phrases can be deduced from the behavior of contrastively focused R-expressions, as was suggested above.

Thus, we have arrived at two potential factors that may affect acquisition of wh-fronting in Russian: contrastive focus and the asymmetry between bare and complex wh-phrases. Future research will, hopefully, show which of these possibilities is correct, or, perhaps, both of these factors affect the acquisition process at the same time.

5. Experiment 3: Malayalam

Recall that, although English and Russian differ with respect to wh-fronting in multiple questions, they do not, however, differ in the PL/SP readings distribution. That is, both English and Russian prohibit SP readings in the core contexts we have examined; and the acquisition data showed that children obey this constraint as well as adults do. However, since some languages do allow SP readings in the very same contexts, it will be interesting to examine at what age children who acquire one of those languages exhibit the knowledge of this fact. We will then discuss how they might acquire it.

Malayalam is a language precisely of this type. It is a Dravidian language spoken primarily in Kerala, one of the southern regions of India. Malayalam is an SOV wh-in-situ language that allows both PL and SP readings in multiple interrogatives, as in (228).

(228) Gibu aaRko ent kodutu? *PL/SP* *Malayalam*
Gibu who what gave
'Who did Gibu give what?'

Experiment 3 was conducted in order to examine at what age Malayalam-speaking children show the knowledge of the distribution of the PL and SP readings in multiple questions in their language. The participants of the experiment in this language were 18 monolingual Malayalam-speaking children (ages 4;5 - 5;4, mean 5;1) and 18 adult controls. The nature and the number of target scenarios were as follows.

(229) a. 2 PL scenarios for double-object questions (i.e., *Who did Lizard give what?)*
b. 2 SP scenarios for double-object questions
c. 2 PL scenarios for object–adjunct questions (i.e., *Who did the dog find where?*)
d. 2 SP scenarios for object–adjunct questions

The target scenarios were mixed with fillers targeting single wh-questions.

5.1. Results

Unlike in Russian and English, both Malayalam-speaking adults and children produced multiple interrogatives in SP scenarios as well as in PL scenarios. The rate of production of multiple questions by Malayalam-speaking adults in PL vs. SP scenarios is shown in Figure 7, and the corresponding behavior of children is represented in Figure 8.

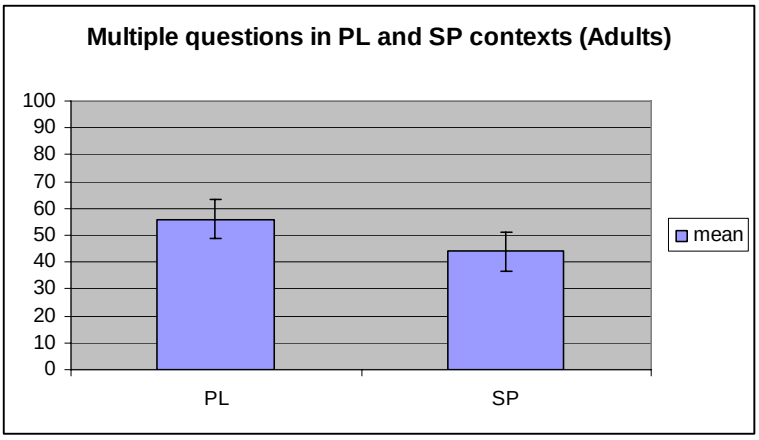


Figure 7. Multiple questions of Malayalam-speaking adults in PL/SP contexts

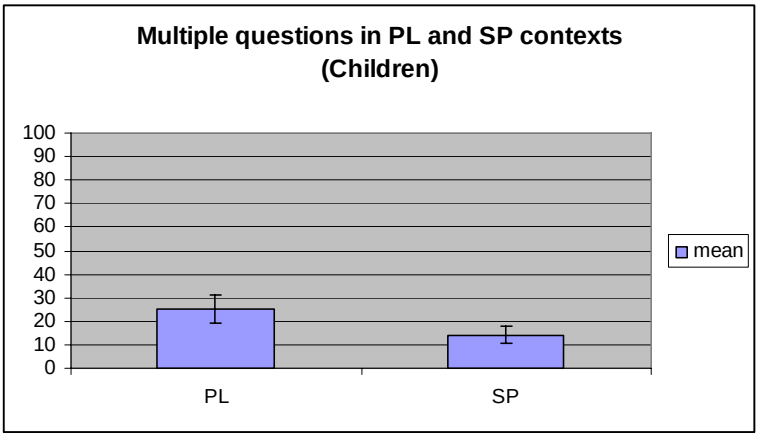


Figure 8. Multiple questions of Malayalam-speaking children in PL vs. SP contexts

Table 5 and Table 6 show the contrast with the results from the experiments on English and Russian.

Table 5. Multiple questions (Adults)

	PL	**SP**
ENG	38%	0
RUS	50%	0
MAL	56%	44%

Table 6. Multiple questions (Children)

	PL	SP
ENG	32%	0
RUS	45%	0
MAL	25%	14%

Recall that Malayalam allows both PL and SP readings in multiple interrogatives, hence the observed adult behavior is as expected. Malayalam children also produced multiple questions in both PL and SP contexts. This is consistent with our overall theory of syntax and semantics of the PL/SP readings, developed in Chapter 3.[80]

5.2. The evidence learners use to acquire PL/SP readings

On the basis of the results from English, Russian, and Malayalam, we can draw a general conclusion that children show a considerably high success rate at learning the language-specific semantic properties of multiple interrogatives. However, recall from Section 2 that multiple interrogatives themselves are not frequent in the linguistic input. This presents a question of how the children manage to acquire this knowledge on the basis of the available input.

Recall that we considered two theoretical accounts of cross-linguistic variation that in Chapter 3. The Relativized Minimality account of Bošković (2003) allowed to capture the tendency that overt wh-movement to SpecCP cancels the SP reading in a number of languages, leaving only the PL reading available, as in English. This account is very learnability-friendly in the sense that plenty of positive

[80] No age effects were found within the age range that was tested.

evidence is available to children in the form of single wh-questions, where they can observe overt wh-movement to SpceCP.

However, the generalization that the Relativized Minimality account was capturing had some puzzling exceptions (i.e., Icelandic multiple questions, and Serbo-Croatian multiple questions in embedded clauses). The account of Grebenyova (2004), based on the selectional properties of the Q-morpheme was then provided either as an alternative to the Relativized Minimality account, or as at least an addition to the Relativized Minimality account. Let us explore what positive data is available to the children in order to learn the selectional properties of the Q-morpheme in their language.

I suggest that the language-specific semantics of multiple interrogatives can be deduced from an independent observable property of the language, namely, the existence of an independent contrastive focus projection. Languages that allow SP readings in multiple questions, such as Japanese, Malayalam, Serbo-Croatian, and Icelandic seem to be related to one another by virtue of having an independent Focus projection above TP that is different from CP or any other projection in the left periphery. I schematize such a structure in (230).

(230) $[_{CP}\ \underset{+wh}{C^0}\ [_{FocP}\ \underset{+foc}{Foc^0}\ [_{TP}\ T^0 \ldots]]]$

Evidence for the existence of such independent Focus projection in these languages can be found in Hiraiwa and Ishihara (2002) for Japanese, Jayaseelan (1999, 2001) for Malayalam, and Bošković (1998, 2002a) for Serbo-Croatian. Icelandic facts seem

to support this structure as well, given its property of allowing V2 in embedded clauses. However, more work needs to be done on Icelandic in this regard.

Putting it in line with syntax and semantics of multiple wh-questions developed in Chapters 2 and 3, it is this FocP, that the Q-mopheme selects in these languages, producing the SP reading. Thus, if a language has this kind of projection, it will allow SP readings in bare multiple wh-questions.[81]

Not all languages have an independent Focus projection. In such languages, the *+foc* feature is located on some other projection hosting other features. For example, it seems plausible that *+wh* and *+foc* features are both located on C^0 in languages like English, Bulgarian, and Brazilian Portuguese, as schematized in (231).

(231) $[_{CP}\ C^0_{+wh\ +foc}\quad [_{TP}\ T^0 \ldots]]$

I have argued that this is the case in English in Chapter 4 on the basis of the evidence from Sluicing. For the evidence that this also holds in Bulgarian, see Bošković (2002a). Brazilian Portuguese facts, as described in Pires (2004), are also compatible with this system.

Given this typology, children acquiring the interpretive possibilities of multiple interrogatives in their language, can rely on the evidence even from non-wh-constructions. What sort of evidence is it? In languages like Japanese and Malayalam, the focus head is overtly realized, hence its independence from the complementizer is

[81] Recall from Chapter 3 that complex wh-phrases have their own way of obtaining a SP interpretation, namely, by introducing their own choice function variables.

easy to observe. In languages like Serbo-Croatian, where the focus head is phonetically null, children have to rely on the distribution of focus-fronted expressions with respect to the complementizer and the occurrence of the intervening lexical material between focus-fronted expressions.

5.3. More results from Experiment 3

The data from child Malayalam shows that children produce multiple interrogatives more frequently in PL contexts than in SP contexts: $t(32) = 1.6$, $p = 0.1$ This result can be attributed to the fact that they might still be in the process of acquiring the properties of the focus projection in Malayalam, needed for the SP reading. The PL reading is easier in this respect because, in this case, the Q-morpheme selects a wh-phrase, and all the properties of wh-phrases in this language can be observed from single interrogatives alone.

Another result from Experiment 3 is that the rate of production of multiple questions by Malayalam-speaking children is overall lower than that of Malayalam-speaking adults. Consider the graphs in Figures 9 and 10.

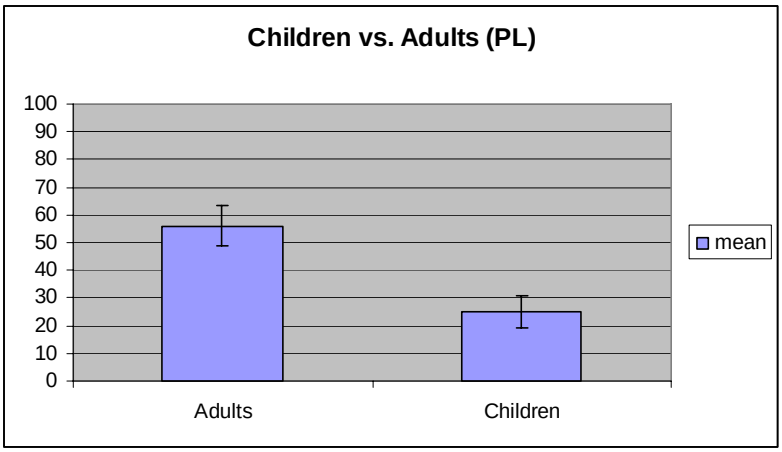


Figure 9. Multiple questions of Malayalam-speaking children vs. adults in PL contexts

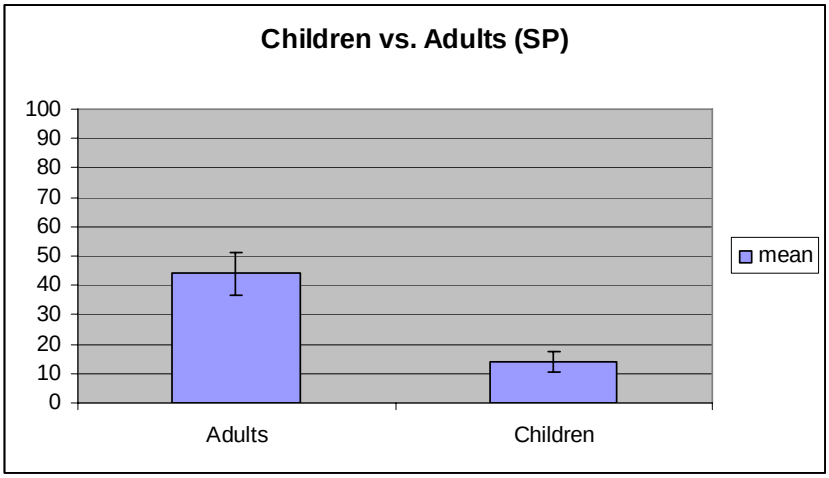


Figure 10. Multiple questions of Malayalam-speaking children vs. adults in SP contexts

Note that this is true of both SP and PL contexts. No such contrasts were found in the experiments on English and Russian. That is, neither English nor Russian children struggled with the PL reading, which is the only reading available in those languages due to the lack of the independent FocP above TP, by hypothesis. This can be interpreted as indicating that the absence of an independent Focus projection is the default option. That is, the grammar seems to prefer to host multiple features on a single head rather than distributing them across multiple heads. For a similar

intuition, see Uriagereka (To appear). The learners then always need positive evidence for the existence of a certain independent projection in their target language, but they do not need evidence for the lack of an independent projection. It is important to clarify that this concerns only the independence status of a given projection (i.e., whether a certain feature forms a head of its own or forms one head with another feature). The presence or absence of certain features in a language is a different matter.

There are other explanations of the adult-child asymmetry in Malayalam that are worth considering. One is that the source of this contrast lies in some non-linguistic factors. For instance, the Malayalam-speaking children were overall much more shy in participating in the task, much more so than the English- or Russian-speaking children. Only several children in the English and Russian experiments could not handle the task from the very beginning. Such subjects typically, instead of asking questions about the story, were telling the puppet about the story. On the other hand, 14 out of 32 potential Malayalam subjects could not handle the task, and not because they misunderstood the task but rather because they seemed too shy to say anything at all in the experimental setting. The same amount of time was spent on getting to know the children in all three experiments, suggesting that there might be a cultural factor involved.

Another factor that seems relevant here is the complexity of a basic wh-question in Malayalam, something that children need to learn in addition to everything else. Wh-questions in Malayalam tend to be clefts more often than in other languages, as discussed in Mohanan (1984) and Jayaseelan (1999, 2001). Malayalam

wh-phrases host optional copula affixes in clefted single and multiple wh-questions, as demonstrated in (232).[82] The examples are taken from the utterances elicited from adult Malayalam-speakers.

(232) a. Palli raajakumaari-kku ent-**aaNu** kodutt-atu
lizard princess-DAT what-be give-NOMINALIZER
What did the lizard give to the princess? (What is it that the lizard gave…?)

b. TavaLa ent-**aaNu** eviDe-**yaaNu** oLi-ppi-ccu vecc-atu?
frog what-be where-be hide-CAUS-PST put-NOMINALIZER
'What did the frog hide where?'

Malayalam wh-questions also use optional distributive markers in PL contexts. These markers can sometimes be the only affixes on wh-phrases, as in (233a), and sometimes they can be combined with the copula affixes, as in (233b).

(233) a. TavaLa ent-**okke** eviDe-**yokke** oLi-ppi-ccu vecc-atu?
frog what-each where-each hide-CAUS-PST put-NOMINALIZER
'What all did hide where all?'

b. TavaLa ent-**okke-yaaNu** eviDe-**yokke-aaNu** oLi-ppi-ccu vecc-atu?
frog what-each-be where-each-be hide-CAUS-PST put-NOMINAL
'What did the frog hide where?'

[82] In multiple wh-questions, the copula morpheme can appear on both, on one, or none of the wh-phrases.

These additional complexities of the wh-questions in Malayalam might be making the overall learning task harder, producing the differences between the rate of production of multiple questions by children and adults.

6. Summary

To summarize, in this chapter, we have explored at what age Russian-, English- and Malayalam-speaking children acquire syntax and semantics of multiple interrogatives and what evidence in the input they use in making hypotheses about their target grammar.

First, we explored the frequency of multiple interrogatives in parental speech and concluded that those are quite rare in the input available to children. That raised a question as to how the acquisition of the grammar of multiple interrogatives proceeds.

Results from Experiments 1 and 2 showed that both English- and Russian-speaking children, like adults, produce multiple interrogatives only in PL contexts, exhibiting robust knowledge of the semantics of multiple questions in their languages. Results from Experiment 3 showed that Malayalam-speaking children, like adults, produce multiple questions in both PL and SP contexts, exhibiting knowledge of the language-specific interpretive properties of multiple interrogatives in Malayalam. In explaining how children acquire these properties on the basis of the available input, I proposed that they rely on an independent property of language, namely, the presence of an independent Focus projection above TP, the evidence for which comes from the distribution of the overt focus morphology, complementizers, and focus-fronted expressions.

It was also discovered that children exhibit perfect knowledge of syntax of multiple interrogatives in most cases, except for the Russian-speaking children's lack of fronting of one of the wh-phrases in a question. I attributed this to two factors. One factor is the acquisition of contrastive focus. Specifically, on the theory where multiple wh-fronting (MWF) is driven by contrastive focus, children learn whether a given language has MWF or not on the basis of the evidence from sentences with contrastively focused R-expressions. The other factor that I pointed out as playing a role in the acquisition of MWF is the acquisition of the asymmetry between complex and bare wh-phrases in Russian. That is, children have to figure out that only bare wh-phrases are obligatorily fronted in Russian. I further suggested that the bare-complex wh-phrase asymmetry itself can be acquired on the basis of the evidence from scrambling because complex wh-phrases behave just like scrambled R-expressions in Russian. Thus, children have the evidence available to them in sentences containing scrambled R-expressions, from which they can deduce the properties of complex wh-phrases.

References

An, Duk-Ho. To appear. Clauses in Non-canonical Positions at the Syntax-Phonology Interface. *Syntax*.

Aoun, Joseph and Yen-hui Audrey Li. 2003. *Essays on the representational and derivational nature of grammar: the diversity of Wh-constructions*. Cambridge, Mass., MIT Press.

Bailyn, John. 2006. On the Scrambling Anti-Movement Movement. *Formal Approaches to Slavic Linguistics: The Princeton Meeting*, 2005, H. Filip, S. Franks, J. Lavine and M. Tasseva-Kurktchieva (eds), Ann Arbor, MI: Michigan Slavic Publications.

Baker, C.L. 1970. Notes on the description of English questions: the role of an abstract question morpheme. *Fopundations of Langauge* 6: 197-219.

Barss, Ansrew. 2000. Minimalism and Asymmetric wh-interpretation. *Step by Step.* Roger Martin, David Michaels, and Juan Uriagereka (eds.), Cambridge, Mass.: MIT Press.

Barss, Andrew, Ken Hale, Ellavina Tsosie Perkins, and Margaret Speas. 1991. Logical Form and barriers in Navajo. In James Huang and Robert May (eds.), *Logical structure and lingusitic structure*. Dordrecht: Kluwer.

Bechhofer, Robin. 1976. Reduced Wh-Questions. *Harvard Studies in Syntax and Semantics*, eds. J. Hankamer and J. Aissen. Cambridge, MA, Harvard University**:** 31-67.

Bechhofer, Robin. 1977. A Double Analysis of Reduced Wh-questions. *Proceedings of the 7th North East Linguistic Society Meeting*, Amherst, MA: 19-31.

Bobaljik, Jonathan David 1995. *Morphosyntax: The syntax of verbal inflection*, Doctoral Dissertation, MIT.

Bobaljik, Jonathan David and Dianne Jonas. 1996. Subject Positions and the Roles of TP. *Linguistic Inquiry* 27: 195-236.

Boeckx, Cedric. 2003. *Islands and chains: Resumption as derivational residue*. Amsterdam/Philadelphia, John Benjamins.

Boeckx, Cedric and Kleanthes Grohmann. 2004. "Sub-Move: Towards a Unified Account of Scrambling and D-Linking." In David Adger, Cécile de Cat & George Tsoulas, eds. *Peripheries: Syntactic Edges and Their Effects*. Dordrecht: Kluwer, 341-357.

Bošković, Željko. 1997a. Fronting wh-phrases in Serbo-Croatian. In *Formal Approaches to Slavic Linguistics: The Indiana Meeting*, 1996, M. Lindseth and S. Franks (eds), 86-107. Ann Arbor, MI: Michigan Slavic Publications.

Bošković, Željko. 1997b. On certain violations of the superiority condition, Agr_O and economy of derivation. *Journal of Linguistics* 33: 227-254.

Boskovic, Zeljko. 1997c. *The syntax of non-finite complementation: An economy approach*. Cambridge, Mass., MIT Press.

Bošković, Željko. 1998. Multiple wh-fronting and economy of derivation. *Proceedings of the 16th West Coast Conference on Formal Linguistics*. E. Curtis, J. Lyle and G. Webster. (eds.), Stanford, CSLI Publications**:** 49-63.

Bošković, Željko. 2001. *On the nature of the syntax-phonology interface: Cliticization and related phenomena,*. Amsterdam, Elsevier.

Bošković, Željko. 2002a. On multiple wh-fronting. *Linguistic Inquiry* 33: 351-384.

Bošković, Željko. 2002b. A-movement and the EPP. *Syntax* 5: 167-218.

Bošković, Željko. 2003. On the Interpretation of Multiple Questions. *Linguistic Variation Yearbook*. John Benjamins, Amsterdam, 1-15.

Bošković, Željko. 2005. On the Locality of Move and Agree: Eliminating the Activation Condition, Generalized EPP, Inverse Case Filter, and Phase-Impenetrability Condition. *The University of Connecticut Occasional Papers in Linguistics.*

Bošković, Željko and Howard Lasnik. 1999. How Strict Is the Cycle? *Linguistic Inquiry* 30: 691-703.

Cheng, Lisa. (1991). *On the Typology of Wh-questions*, Ph.D. Dissertation, MIT, Cambridge, Mass.

Chierchia, Gennaro. 1993. Questions with Quantifiers. *Natural Language Semantics* 1/**2**: 181-234.

Chomsky, Noam. 1955. The logcal structure of linguistic theory. Ms., Harvard Unoversity, Cambridge, Mass. [Revised 1956 version published in part by Plenum New York, 1975; University of Chicago Press, 1985.]

Chomsky, Noam. 1957. *Syntactic structures*. The Hague, Mouton.

Chomsky, Noam. 1973. Conditions on transformations. *A Festschrift for Morris Halle*. S. Anderson and P. Kiparsky (eds.), New York, Holt Rinehart & Winston**:** 232-286.

Chomsky, Noam. 1986. *Barriers*. Cambridge, Mass., MIT Press.

Chomsky, Noam. 1991. Some notes on economy of derivation and representation. *Principles and parameters in comparative grammar*. R. Freidin. Cambridge, Mass., MIT Press**:** 417-54.

Chomsky, Noam. 1995. Categories and Transformations, in *The Minimalist Program*. Cambridge, Mass.: MIT Press, 219-394.

Chomsky, Noam. 1999. *Derivation by phase. MIT Occasional Papers in Linguistics 18*. Cambridge, Mass., MITWPL.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. *Step by step: Essays on minimalism in honor of Howard Lasnik*. R. Martin, D. Michaels and J. Uriagereka. Cambridge, Mass., MIT Press**:** 89-155.

Chomsky, Noam. 2001. *Beyond explanatory adequacy. MIT Occasional Papers in Linguistics 20*. Cambridge, Mass., MITWPL.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36(1): 1-22.

Chung, Sandra. 1982. Unbounded dependencies in Chamorro grammar. *Linguistic Inquiry* 13(1): 39-77.

Chung, Sandra, William Ladusaw, and James McCloskey. 1995. Sluicing and logical form. *Natural Language Semantics,* 3. Kluwer Academic Publishers, 239-82.

Clahsen, Harald, Claudia Kursawe and Martina Penke. 1995. Introducing CP: Wh-questions and subordinate clauses in German child language. In Essex Research Reports in Linguistics 7, 1-28. Essex, England: University of Essex, Department of Linguistics.

Comorovski, Ileana. 1996. *Interrogative phrases and the syntax-semantics interface*. Dordrecht: Kluwer Academic Publications.

Craenenbroeck, Jeroen. van and Anikó Lipták. 2005. Ellipsis in Hungarian and the typology of Sluicing, in Kiyong Choi and Changguk Yim (eds.), *Ellipsis in Minimalism*, *Proceedings of the 7th Seoul International Conference on Generative Grammar*. Deoul: Hankook, 103-133.

Dalrymple, Mary, Stuart Schiebe and Fernando Pereira. 1991. Ellipsis and Higher-Order Unification. *Linguistics and Philosophy*, 14. Kluwer Academic Publishers, 339-452.

Dayal, Veneeta. 1996. *Locality in Wh quantification : questions and relative clauses in Hindi*. Dordrecht ; Boston, Kluwer Academic Publishers.

Dayal, Veneeta. 2002. Single-pair versus multiple pair answer: wh-in-situ and scope. *Linguistic Inquiry* 33(3): 512-520.

Dayal, Veneeta. 2005. Multiple Wh-Questions. In M. Everaert and H. van Riemsdijk, *Syntax Companion 3*, Case #44, Blackwell Publishers.

Engdahl, Elisabet. 1985. *Interpreting Questions*. Cambridge: Cambridge University Press.

Fiengo, Robert and Robert May. 1994. *Indices and identity*. Cambridge, Mass., MIT Press.

Fox, Danny. 1995. Economy and scope. *Natural Language Semantics* 3: 283-341.

Fox, Danny. 2000. *Economy and Semantic Interpretation*. Cambridge, Mass.: MIT Press.

Fox, Danny and Howard Lasnik. 2003. Successive cyclic movement and island repair: the difference between sluicing and VP ellipsis. *Linguistic Inquiry* 34(1): 143-154.

Fox, Danny and David Pesetsky. 2005. Cyclic linearization of syntactic structure. *Theoretical Linguistics* 31(1-2): 1-45.

George, L. M. (1980). Analogical generalization in natural language syntax**.** Ph.D. Dissertation, MIT, Cambridge, Mass.

Grebenyova, Lydia. 2004. Interpretation of Slavic Multiple Wh-Questions, in Olga Arnaudova, Wayles Browne, Maria Luisa Rivero, and Danijela Stojanović (eds.), *Formal Approaches to Slavic Linguistics: The Ottawa Meeting*, 2005. Ann Arbor, MI: Michigan Slavic Publications, 169-186.

Grebenyova, Lydia. In press. Multiple Left-branch Extraction under Sluicing. Proceedings of the 41st Meeting of the Chicago Linguistic Society.

Groenendijk, Jeroen and Stokhof, Martin. 1984. *Studies on the semantics of questions and the pragmatics of answers.* Doctoral Dissertation, University of Amsterdam.

Guasti, Maria Teresa. 2000. An excursion into interrogatives in early English and Italian. In M.-A. Friedemann and Luigi Rizzi, eds. *The acquisition of syntax*. Harlow, England: Longman.

Hagstrom, Paul. 1998. *Decomposing Questions.* Doctoral Dissertation, Massachusetts Institute of Technology.

Hamblin, C.L. 1958. Questions. *Australasian Journal of Philosophy* 36(3): 159-168.

Hamblin, C.L. 1973. Questions in Montague English. *Foundations of Language* 10:41-53.

Hardt, Daniel. 1993. *Verb Phrase Ellipsis: Form, Meaning and Processing*. Ph.D. Dissertation, University of Pennsilvania.

Hardt, Daniel. 1999. Dynamic Interpretation of Verb Phrase Ellipsis, *Linguistics and Philosophy,* 22. Kluwer Academic Publishers, 185-219.

Heim, Irene. 1982. *The semantics of definite and indefinite noun phrases.* Doctoral Dissertation, University of Massachusetts, Amherst.

Higginbotham, James and Robert May. 1981. Questions, quantifiers and crossing. *The Linguistic Review* 1: 41-79.

Hoji, Hajime 1985. *Logical form constraints and configurational structures in Japanese*. Doctoral Dissertaton, University of Washington, Seattle.

Hornstein, Norbert. 2001. *Move! A minimalist theory of construal*. Malden, Mass., Blackwell.

Ionin, Tania. 2001. The One Girl who was Kissed by Every Boy: Scope, Scrambling and Discourse Function in Russian, in *Marjo van Koppen, Joanna Sio, and Mark de Vos (eds.), Proceedings of Console X.* Student Organization of Linguistics in Europe. Leiden University. Leiden, The Netherlands, 79-94.

Izvorski, Roumyana. 1993. On wh-movement and focus movement in Bulgarian. Paper presented at *CONSOLE 2*, University of Tübingen, December 1993.

Jacobson, Pauline. 1992. Antecedent-Contained Deletion in a Variable Free Semantics. In C.L. Baker and David Dowty (eds.), *Proceedings from the 2nd*

Conference on Semantics and Linguistic Theory. Columbus, OH: Ohio State University, 193-214.

Jayaseelan, Karattuparambil A. 1990. Incomplete VP deletion and gapping. *Linguistic Analysis* 20: 64-81.

Jayaseelan, Karattuparambil A. 1999. A focus phrase above vP. In Proceedings of the Nanzan GLOW, 1995-212. Nanzan University, Nagoya, Japan.

Jayaseelan, Karattuparambil A. 2001. Questions and question-word incorporating quantifiers in Malayalam. *Syntax* 4:2, 63-93.

Kamp, Hans. 1981. A theory of truth and semantic representation. *Formal methods in the study of language*. J. Groenendijk, T. M. V. Janssen and M. Stokhof (eds.), Amsterdam, Mathematical Centre: 277-322.

Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics and Philosophy* 1: 3-44.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, Mass., MIT Press.

Kishimoto, Hideki. 1992. LF pied-piping: Evidence from Sinhala. *Gengo Kenkyu* 102:46-87.

Kishimoto, Hideki. 1997. Wh-in-situ and null operator movement in Sinhala questions. Ms., Hyogo University of Teacher Education.

Kitahara, Hisatsugu. 1993. Deducing strict cyclicity from derivational economy. Paper presented at the 16th GLOW Colloquium, Lund.

Kitahara, Hisatsugu. 1994. *Target ALPHA: A unified theory of movement and structure building*. Cambridge, Mass., Harvard University.

Kitahara, Hisatsugu. 1997. *Elementary operations and optimal derivations*. Cambridge, Mass., MIT Press.

Koizumi, Masatoshi. 1995. *Phrase Structure in Minimalist Syntax*. Doctoral Dissertation, Massachusetts Institute of Technology.

Kratzer, Angelika. 1998. Scope or pseudoscope? Are there wide scope indefinites? *Events in Grammar*. S. Rothstein. Dordrecht (Ed.), Kluwer: 163-196.

Kratzer, Angelika. forthcoming. *The Event Argument and the Semantics of Voice*. Cambridge, MA, MIT Press.

Krifka, Manfred. 2001. Quantifying into question acts. *Natural Language Semantics* 9: 1-40.

Lasnik, Howard. 1995. Notes on pseudogapping. *Papers on minimalism. MIT Working Papers in Linguistics* 27. R. Pensalfini and H. Ura (eds.), Cambridge, Mass., MITWPL**:** 143-163.

Lasnik, Howard. 1999a. On Feature Strength: Three minimalist approaches to overt movement. *Linguistic Inquiry* 30: 197-217.

Lasnik, Howard. 1999b. Chains of arguments. *Working minimalism*. D. S. Epstein and N. Hornstein. Cambridge, Mass., MIT Press**:** 189-215.

Lasnik, Howard. 2001. Subjects, objects, and the EPP. In W. D. Davies and S. Dubinsky (eds.) *Objects and Other Subjects: Grammatical Functions, Functional Categories, and Configurationality.* Kluwer, pp. 103-121.

Lasnik, Howard. 2001b. When can you save a structure by destroying it? *Proceedings of the 31st North East Linguistic Society Meeting*, eds. M. Kim and U. Strauss. Amherst, Massachusetts, GLSA**:** 301-320.

Lasnik, Howard. 2005. More Sluicing Repairs. Handout for the LSA 208 course at the Summer Institute of the Linguistic Society of America.

Lasnik, Howard and Juan Uriagereka. 2004. *A course in minimalist syntax*. Oxford, Blackwell.

Lewis, David. 1975. Adverbs of Quantification. In *Formal Semantics of Natural Language*. Edward L. Keenan (Ed.), 3-15. Cambridge: Cambridge University Press.

Lobeck, Anne. 1991. The Phrase Structure of Ellipsis. *Perspectives in Phrase Structure (Syntax and Semantics)*. S. Rothstein. San Diego, Academic Press**:** 81-103.

Lobeck, Anne. 1995. *Ellipsis: functional heads, identification, and licensing*. Oxford, Oxford University Press.

Mahajan, Anoop. 1990. The A/A-Bar Distinction and Movement Theory. Doctoral Dissertation. Massachusetts Institute of Technology.

Mahajan, Anoop. To appear. Gapping and Sluicing in Hindi. *Proceedings of Glow in Asia V*, New Delhi.

McCloskey, James. 1990. Resumptive Pronouns, A'-binding, and Levels of Representation in Irish. In *Syntax and Semantics*, Volume 23: The Syntax of the Modern Celtic Languages.

McCloskey, James. 2000. Quantifier float and wh-movement in an Irish English. *Linguistic Inquiry* 31(1): 57-84.

Merchant, Jason. 2001. *The Syntax of Silence: Sluicing, Islands, and the Theory of Ellipsis*. Oxford: Oxford University Press.

Merchant, Jason. 2004. Fragments and Ellipsis. *Linguistics and Philosophy* 27.6: 661-738.

Mohanan, K. P. 1984. Operator binding and the path containment condition. *Natural Language & LinguisticTheory* 2: 357-396.

Nishigauchi, Taisuke. 1986. *Quantification in Syntax*, Doctoral Dissertation. University of Massachusetts, Amherst.

Nunes, Jairo. 2004. *Linearization of chains and sideward movement*. Cambridge, Mass, MIT Press.

Oka, Toshifusa. 1993. Shallowness. *MIT Working Papers in Linguistics* 19.

Pafel, Jürgen. 1999. Interrogative quantifiers within scope. *Linguistics and Philosophy* 22: 255-310.

Park, Bum-Sik. 2005. *Locality and Identity in Ellipsis*. Doctoral Dissertation, University of Connecticut.

Pereltsvaig, Asya. In press. Small Nominals. *Natural Language and Linguistic Theory,* 24. Kluwer Academic Publishers.

Pesetsky, David. 1987. Wh-in-situ: movement and unselective binding. *The representation of (in)definiteness*. E. Reuland and A. t. Meulen. Cambridge (eds.), Mass., MIT Press: 98-129.

Pesetsky, David. 1989. The Earliness Principle. Paper presented at GLOW, available at http://web.mit.edu/linguistics/www/pesetsky/earliness.pdf.

Pesetsky, David. 2000. *Phrasal movement and its kin.* Cambridge, Mass.: MIT Press.

Pesetsky, David and Esther Torrego. 2001. T-to-C Movement: Causes and Consequences. In Michael Kenstowicz (ed.), *Ken Hale: A Life on Language*. Cambridge, MA: MIT Press.

Pires, Acrisio. 2004. Verb movement and clitics: Variation and change in Portuguese. *Selected Papers from the DIGS VII – Diachronic Generative Syntax Conference*, University of Girona, Spain. F. Roca and M. Batlori (eds.), Oxford University Press.

Postal, Paul Martin. 1974. *On raising: One rule of English grammar and its implication*. Cambridge, Mass., MIT Press.

Reinhart, Tahya. 1995. *Interface strategies. OTS Working Papers*, Research Institute for Language and Speech, Utrecht University.

Reinhart, Tanya. 1997. Quantifier scope - How labor is divided between QR and choice functions. *Linguistics and Philosophy* 20: 335-397.

Richards, Norvin. 1997. *What moves where when in which language?* Ph.D. dissertation, MIT, Cambridge, Mass.

Richards, Norvin. 1999. Featural cyclicity and the ordering of multiple specifiers. *Working minimalism*. S. D. Epstein and N. Hornstein. Cambridge, Mass., MIT Press**:** 127-158.

Richards, Norvin. 2001. *Movement in language*. Oxford/New York, Oxford University Press.

Rivero, Maria-Luisa. 1994. Clause structure and V-movement in the languages of the Balkans. *Natural Language and Linguistic Theory* 12: 63-120.

Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, Mass., MIT Press.

Rizzi, Luigi. 1996. Risidual verb second and the wh-criterion. In Parameters and functional heads: Essays in comparative syntax, Eds. A. Belletti and L. Rizzi: 63-90. Oxford: Oxford University Press.

Roeper, Tom and Jill de Villiers. 1991. The emergence of bound variable structures. In T. Maxfield and B. Plunkett, eds., *Papers in the acquisition of WH*. Amherst: University of Massachusetts, GLSA.

Ross, John Robert. 1969. Guess who? In *Papers from the Fifth Regional Meeting of the Chicago Linguistic Society*, ed. Robert I. Binnick, Alice Davison, Georgia M. Green and Jerry L. Morgan, 252-286. Chicago Linguistic Society, University of Chicago, Chicago, Ill.

Rudin, Catherine. 1988. On multiple questions and multiple WH fronting. *Natural Language and Linguistic Theory* 6(4): 445-501.

Sag, Ivan. 1976. *Deletion and logical form*. Doctoral dissertation, MIT.

Santelmann, Lynn. 1998. The acquisition of verb movement and spec-head relationships in child Swedish. In David Adger, Susan Pintzuk, Bernadette Plunkett,, and George Tsoulas, eds., Specifiers: Minimalist Approaches. Oxford: Oxford University Press.

Seidl, A., G. Hollich, and P. Jusczyk. 2003. Infants' and toddlers' comprehension of subject and object wh-questions. *Infancy* 4(3).

Silva, G. V. 2001. *Word order in Brazilian Portuguese*. Berlin: Mouton.

Stepanov, Arthur. 1998. On Wh-Fronting in Russian, in *Proceedings of the 30th Meeting of the North-Eastern Linguistic Society*.

Stepanov, Arthur and Dylan Tsai. submitted. The Hows and Whys of *How* and *Why*. Ms.

Stjepanović, Sandra. 2003. Multiple wh-fronting in Serbo-Croatian matrix questions and the matrix sluicing construction. *Multiple Wh-fronting*, C. Boeckx and K. Grohmann (Eds.), Amsterdam, John Benjamins**.**

Sugahara, Mariko. 1996. Shuri Okinawan *kakari-musubi* and movement. In Masatoshi Koisumi, Masayuki Oishi, and Uli Sauerland (eds.), *Formal Approaches to Japanese Linguistics* 2, MITWPL 29. Cambridge, MA: MIT Working Papers in Linguistics.

Szabolcsi, Anna. 1997. Quantifiers in pair-list readings. In *Ways of scope taking*. Anna Szabolcsi (Ed.), Dordrecht, Kluwer**:** 109-154

Takahashi, Daiko 1994a. *Minimality of movement.* Doctoral Dissertation, University of Connecticut, Storrs, CT.

Takahashi, Daiko. 1994b. Sluicing in Japanese. *Journal of East Asian Linguistics* 3: 265-300.

Thornton, Rosalind. 1990. Adventures in long-distance moving: The acquisition of complex *wh*-questions. Doctoral dissertation, University of Connecticut, Storrs.

Torrego, Esther. 1984. On inversion in Spanish and some of its effects. *Linguistic Inquiry* 15(1): 103-29.

Travis, Lisa. 1991. Parameters of Phrase Structure and Verb-Second Phenomena. In R. Freidin (ed.), *Principle and Parameters in Comparative Grammar*. MIT Press, Cambridge, MA: 339-364.

Wachowicz, Kristina. A. 1974. Against the universality of a single wh-question movement. *Foundations of Language* 11:155-166.

Williams, Edwin. 1977. Discourse and logical form. *Linguistic Inquiry* 8: 101-139.

Yamakoshi, Kyoko. 2002. Children's Understanding of the Universal Quantifier *Wh*+mo in Japanese. In Alejna Brugos, Linnea Micciulla, and Christine E. Smith, eds., *Proceedings of the 28th annual Boston University Conference on Language Development*. Cascadilla Press.

Zwart, C. Jan Wouter. 1991. Clitics in Dutch: Evidence for the Position of INFL. *Groninger Arbeiten zur Germanistischen Linguistik* 33, 71-92.

Zwart, C. Jan Wouter. 1993. *Morphosyntax of verb movement: A minimalist approach to the syntax of Dutch*. Dordrecht, Kluwer.