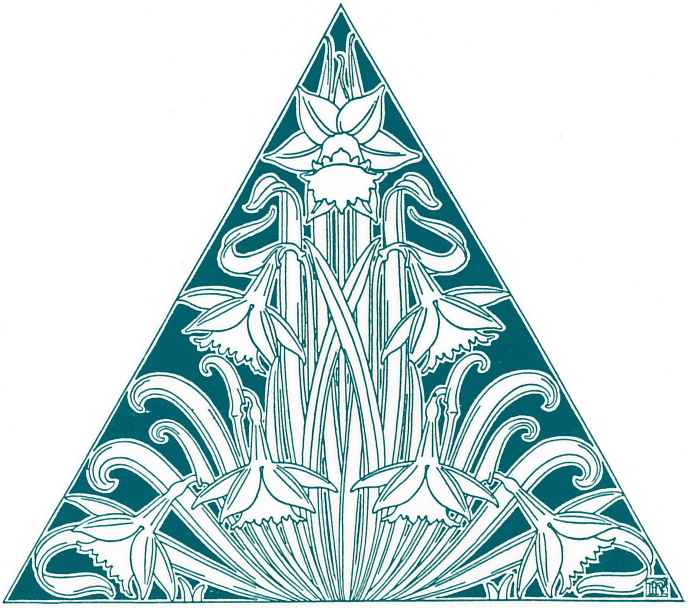
DISSERTATIONS IN LINGUISTICS

Configuring Topic and Focus in Russian

TRACY HOLLOWAY KING



Configuring Topic and Focus in Russian

DISSERTATIONS IN LINGUISTICS

A series edited by
Joan Bresnan, Sharon Inkelas, William J. Poser, and Peter Sells

The aim of this series is to make work of substantial empirical breadth and theoretical interest available to a wide audience.

Configuring Topic and Focus in Russian

TRACY HOLLOWAY KING

CSLI Publications
Center for the Study of Language and Information
Stanford, California

Printed in the United States
99 98 97 96 95 1 2 3 4 5

Library of Congress Cataloging-in-Publication Data

King, Tracy Holloway. 1966-
Configuring topic and focus in Russian / Tracy Holloway King

p. cm. – (Dissertations in Linguistics)
Includes bibliographical references and index.

ISBN: 1-881526-63-1 (cloth)
ISBN: 1-881526-62-3 (paper)

1. Russian language—Word order. 2. Russian language—Topic and comment. 3. Russian language—Grammar, Generative. I. Title. II. Series.

PG2390.K56 1995
491.75–dc20 94-36583
CIP

eISBN: 1-57586-968-3 (electronic)

CSLI Publications gratefully acknowledges a generous gift from Jill and Donald Knuth in support of scholarly publishing that has made the production of this book possible.

"Papier peint. Porillon (Fig. 30)," the drawing on the cover of the paperback edition of this book is by Maurice-Pillard Verneuil from his *Étude de la Plante: Son application aux industries d'art* (1903, Paris: Librairie centrale des beaux-arts). Courtesy Richard Manuck.

∞ The acid-free paper used in this book meets the minimum requirements of the American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

This book is dedicated to
Catherine V. Chvany
on the occasion of her official retirement.

Contents

Acknowledgements

For their tireless support of this research, I would like to thank my thesis advisors: Peter Sells, Eve Clark, and Tom Wasow. Many people provided the discussion and comments necessary for the completion of any major research project; in particular, Stanford University, CSLI, and Xerox PARC produced an ideal environment for intellectual exchange. Revision of this work benefited greatly from detailed comments by Loren Billings, Joan Bresnan, Natasha Kondrashova, Catherine Rudin, and especially Steve Franks.

Of great importance to this work were my informants. They are in no way responsible for my interpretation of the data they so patiently provided.

The National Science Foundation, the Social Science Research Council, and Stanford University provided generous financial assistance for this research.

1

Introduction

This work examines word order[1] and the encoding of topic and focus in Russian. As has long been observed, word order in Russian encodes specific discourse information: with neutral intonation, topics precede discourse-neutral constituents which precede foci. I extend this idea to show that word order encodes different types of topic and focus in a principled manner. These discourse functions appear in particular phrase structure positions and are marked intonationally. The interaction of topic and focus with the syntax and the nature of phrase structure in general has been vigorously debated in the recent linguistic literature. The discussions have focused on whether all languages have hierarchically structured VPs and how best to encode discourse function information, as distinct from scrambling. An in-depth analysis of Russian elucidates this debate since Russian contains both configurational and non-configurational characteristics. There is evidence that the underlying structure of Russian is configurational, and its surface word order, which at first appears to be very free, is constrained by the discourse context of the utterance.

1.1 The Basic Problem

Languages like English and French are considered highly configurational in that word order encodes grammatical functions, e.g., subjects precede verbs while objects follow them. Each grammatical function is associated with a distinct position in the phrase structure from which constituents rarely move. At the other extreme are languages like Warlpiri where the word order is extremely free, even allowing discontinuous constituents (Hale 1983). In these languages, the subject

[1] More accurately, it is the ordering of constituents that is discussed since prepositional phrases and most noun phrases form syntactic constituents.

and other grammatical functions may appear in any position relative to the verb.

Russian is an interesting counterpart to the languages traditionally discussed in the debate about phrase structure and scrambling since it does not belong to either extreme. Although it never uses word order alone to encode grammatical functions,[2] and the surface word order possibilities are extensive, there is an unmarked word order and the varied surface orders are predictable from discourse factors. A given word order may therefore be preferred in one context while unacceptable in another. This suggests that although grammatical functions are structurally encoded, constituents can move out of the positions which define their grammatical function into ones which define their discourse function.

Traditional descriptions of Russian have long noted that word order in Russian reflects the discourse functions of the constituents (Adamec 1966, Daneš 1974, Hajičova 1974, Isačenko 1976a, 1976b, Krylova and Khavronina 1988, Mathesius 1964, Sgall 1972, etc.). In all of these accounts, given information precedes new information in the sentence. Sgall et al. 1986 refer to this as Communicative Dynamism (CD). Constituents with more CD contain the newest, most prominent information and appear after constituents with less CD, i.e., older information. Such a theory gives the impression of having a scale of infinite gradations. Another view divides sentences into two parts: a theme and a rheme. Roughly speaking, the rheme is new information and the theme given information. In non-emotive speech, the rheme follows the theme. A third variant takes the theme-rheme analysis and allows for further subdivision of the theme into what is essentially topicalized and discourse-neutral material. The rhematic material is focused material. This third variant most closely resembles the proposal made here. Constituents can be either focused, topicalized, or discourse neutral, although these notions are notoriously difficult to define consistently.

[2]It has frequently been noted that certain sentences in which the morphology does not distinguish the subject from object tend to be interpreted as SVO. The classic example is shown in (i) (see Jakobson 1936/1971).

(i) Mat' ljubit doč'.
mother-NOM/ACC loves daughter-NOM/ACC
'Mother loves (her) daughter.'

Both NPs in (i) could be either nominative or accusative, making it potentially ambiguous as to which is the subject. When presented in isolation such a sentence is usually interpreted as SVO, which has led to the assumption that certain 'freezing' effects are operational here (Mohanan 1990). However, given an appropriate context, the above sentence can mean 'the daughter loves (her) mother', thereby eliminating the last vestige of evidence that surface linear order encodes grammatical function in Russian.

There is a parallel problem to that of defining the discourse functions of constituents: what are the syntactic structures of clauses. This problem is the basis of the configurationality debate as to whether all languages have a hierarchical phrase structure like that usually proposed for English. This is particularly a problem for languages with relatively free word order and those which exhibit few subject-object asymmetries. Some of the questions which arise are: Are all structures binary branching? What roles do specifier positions play and are they always present? At what level of representation are discourse functions encoded? How does scrambling differ from movement of topics and foci? Do languages without subject-object asymmetries have flat structures within the VP? If there are subject-object asymmetries in a language, are these always the result of a difference in structural position between the subject and object?

My assumption is that word order is a reflection of the phrase structure. The phrase structure encodes both dominance and precedence relations and at the level of S-structure this closely reflects the surface order. As such, I reject the idea of substantial reorderings at the level of PF (phonetic form). Although there are PF reorderings to account for phenomena such as clitic placement, the major ordering of the constituents is a direct reflection of their S-structure positions in the phrase structure. One of the goals of an analysis of word order is to explain why constituents move out of their D-structure positions into their S-structure positions and what precisely these S-structure positions are. To account for the word order variations found in Russian, I argue that constituents move not just to get case or inflectional features, but also to receive specific discourse function interpretations. Constituents must move into these positions to receive these discourse function interpretations, and any constituent in such a position must receive the relevant interpretation. When possible, comparisons are made with the word order and topic-focus constructions in other languages in order to determine what mechanisms are available for the encoding of discourse functions and how these interact with the syntax.

My basic proposal concerning the phrase structure of Russian is that Russian is a VSO language: SpecVP is subject position, and the tensed verb is located in I^0 where it case marks the subject in SpecVP. The underlying VSO order is obscured by the movement of constituents to receive discourse function interpretations. This movement includes the left-adjunction of topics to IP, the use of SpecIP as a contrastive

focus position, and the left-dislocation of certain topics.[3] Since movement of constituents into these discourse function positions occurs at S-structure, the word order of a clause directly reflects the discourse functions of its constituents. As a result, we can explain the classic observation that Russian word order reflects discourse functions, while employing recent proposals concerning the phrase structure.

1.2 Organization

This work comprises two parts. Part I discusses the phrase structure of Russian and its interaction with topic and focus in Government-Binding theory. Part II examines related issues in Lexical-Functional Grammar, focusing on structures which are traditionally assumed to involve movement in the phrase structure. Chapter 2 provides background information on Government-Binding Theory, Lexical-Functional Grammar, and the Russian language; readers already familiar with the basics of these topics may wish to begin directly with the main discussion in Part I.

Chapter 3 examines the evidence for configurationality in Russian. I argue that Russian projects the subject higher than the object. Evidence for this comes from traditional subject-object asymmetries such as extraction possibilities and superiority effects. Additional evidence is provided by several Russian-particular phenomena, e.g., the genitive of negation. In addition, the distribution of finite and infinitival verb forms supports the idea that there is a separate position for inflected verbs.

Chapter 4 discusses topic and focus in general, providing a working definition for them. It includes discussion of previous work on the interaction of word order and topic–focus interpretation in Russian. The final section of the chapter outlines topic and focus constructions in Russian and discusses the crucial link between intonation and word order for topic–focus interpretation in Russian.

Chapter 5 provides a syntactic account of the different surface word orders used to encode topic and focus in Russian. I propose that finite verbs in Russian raise to I^0 to receive tense and agreement features. However, subjects are case marked in SpecVP and need not move out of the VP for nominative case. This allows SpecIP to be interpreted as a discourse function position, namely a focus position. Part of the chapter looks at the structural encoding of discourse functions in other

[3] Left-dislocated elements are base generated, not the result of movement. However, they are associated with a particular phrase structure position.

languages. Finally, I discuss why Russian is considered to be an SVO and not a VSO language.

In chapter 6, a construction that explicitly mark discourse functions are analyzed: the *li* yes-no question. This constructions interact with the phrase structure and encoding of discourse functions proposed in chapter 5.

Chapters 7, 8, and 9 form Part II and explore how these phenomena can be analyzed in Lexical-Functional Grammar (LFG), a theory which does not involve movement but which separates the encoding of grammatical functions from the phrase structure. The challenge is to account for the facts that led to the conclusion that Russian is configurational, while capturing the varied surface orders and their discourse interpretations, in a framework that has only one level of phrase structure. Chapter 7 examines how case marking interacts with the (morpho)syntax. Related to this is a discussion of the genitive of negation which has often been given a structural explanation that cannot be maintained in LFG. Chapter 8 analyzes the Russian topic and focus structures in LFG, arguing for a largely X′ syntax, which avoids some of the difficulties encountered in Part I. Finally, chapter 9 discusses the phenomena which involve head-to-head movement in GB.

Chapter 10 summarizes the major proposals and discusses three areas for further research.

2

Background Assumptions

Before examining the details of the syntactic theories employed in this dissertation, it is necessary to explain why I examine the same phenomena in two relatively different frameworks, namely Government-Binding Theory (GB) and Lexical-Functional Grammar (LFG). The basic motivation is that different theories better account for different types of phenomena. Transferring and developing an account in a different theory highlights the essential factors behind the phenomena and the idiosyncrasies of a given theory. Ultimately, this allows us to adapt the theories so that they have greater explanatory power. Take, for example, the issue of phrase structure. GB depends extensively on phrase structure and as a result has a fairly developed set of constraints as to what types of structures are allowed and what types of constituents can appear where. Since LFG does not depend so heavily on its phrase structure, this structure is relatively free, and it is not obvious why certain structures never occur, while others which do occur look unusual in that they do not contain all the expected constituents. As a related issue, GB defines grammatical relations structurally, which can pose problems for languages which demonstrate relatively few subject-object asymmetries, such as Hungarian and Malayalam. This is not a problem for LFG where grammatical relations are not dependent on the phrase structure.

Part I examines how the surface orders of Russian can be accounted for in a relatively constrained way in GB. The focus is on the phrase structure, both underlying and surface, and how this structure interacts with the discourse functions of the clause. The proposal takes advantage of structures already provided by the theory. Part II examines these phenomena from an LFG perspective. Although an account of the major topic and focus phenomena is provided, some constructions which have traditionally been thought of as structural in nature

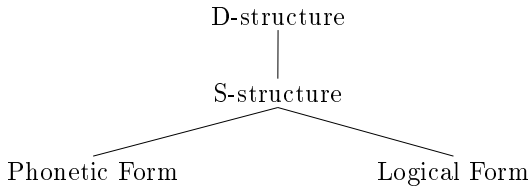
are examined in detail to determine how a theory which is not as dependent on phrase structure accounts for them. For example, the LFG analysis of the genitive of negation is discussed at length, since the genitive of negation is traditionally described as a configurational phenomenon. In addition, GB makes extensive use of head-movement to account for the distribution of verbal forms; in LFG the morphology and constraints on the phrase structure rules are formulated to make comparable statements. Thus, what at first may seem an arbitrary division of which phenomena to examine in which theories is driven by which phenomena seem more stipulative and hence more interesting to motivate in a given theory.

This chapter provides some background to Government-Binding Theory and Lexical-Functional Grammar. Section 2.3 outlines a few basics of Russian grammar for those unfamiliar with the language. More details, both theoretical and empirical, are introduced as they are needed.

2.1 Government-Binding Theory

The first part of the dissertation presumes a version of Government-Binding theory (Chomsky 1981). The grammar is organized as in (1) with constraints applying to different levels of the grammar. The differences among languages are the result of parameters whose values are set as the language is learned.

(1)



The different levels of the grammar are related by a transformation referred to as Move-α which effectively allows anything to be moved anywhere.[1] Various principles constrain which of these movements are well-formed and which are not. Two of these are the Projection Principle in (2) and the Theta Criterion in (3).

(2) Projection Principle: Representations at each syntactic level are projected from the lexicon, in that they observe the subcategorization properties of lexical items. (Chomsky 1981:29)

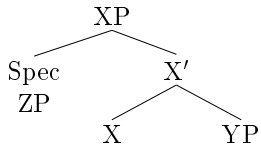
(3) Theta Criterion: Each argument bears one and only one θ-role, and each θ-role is assigned to one and only one argument.

[1] Lasnik and Saito 1984 refer to this more generally as Affect-α which states that any element can be moved, added, or deleted. See also Speas 1990:3.

Another principle is the Extended Projection Principle that stipulates that all clauses must have a subject.[2] Subjects and other grammatical function are defined structurally, and their properties are the result of differences in structural position. It is not necessarily assumed that constituents appearing in a position in one language have the same properties as constituents appearing in that position in another language. The properties may vary depending on how that position interacts with the parameters in each language.

Dominance and precedence are encoded by the phrase structure. I assume a fairly standard X′ syntax in which each projection has a head, X^0, which can be either lexical or functional in nature.[3] XPs consist of two levels of projection: complements of the X^0 head are sisters to it and appear within X′; the XP level consists of a specifier position, which can only be filled by maximal projections, and the X′ projection. The basic structure is shown in (4) where X^0 is the head, ZP is in the specifier of X^0 (referred to as SpecXP), and YP is in complement position as sister to the head.

(4)



The basic structure in (4) can be extended by adjunction. Maximal projections are assumed to adjoin only to other maximal projections, while heads adjoin to head positions through head-movement. Ad-

[2]Babby 1989 rejects the Extended Projection Principle and argues that there are subjectless constructions in some languages, including Russian and Ukrainian. In more traditional approaches, these constructions are assumed to have null expletive subjects which occupy subject position and thus satisfy the Extended Projection Principle.

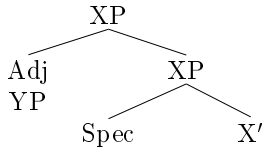
[3]Speas 1990 proposes a more general algorithm for projection which does not make a distinction between X′ and XP levels (see Kornai and Pullum 1990). Her schema is referred to as Project Alpha which creates binary branching trees; maximal and minimal projections are defined as below. The licensing of complementizers and specifiers is dependent on the nature of the head, not the X′ syntax.

> Project Alpha: A word of syntactic category X is dominated by an uninterrupted sequence of X nodes.
> Maximal Projection: $X=X^{max}$ iff for all G which dominate X, $G \neq X$.
> Minimal Projection: $X=X^0$ iff X immediately dominates a word.
> (Speas 1990:43–44)

One potentially convenient result of this system is that not all heads project specifier positions. Since the specifiers of many functional heads never appear to be utilized, Speas's view of projection could eliminate these positions.

junction to a projection results in a projection of the same type, as in (5).

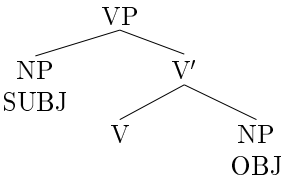
(5)



The X′ schema allows for the separation of dominance and precedence relations. The X, X′, XP representation encodes dominance but makes no claims about precedence. Precedence is now assumed to follow from independent principles needed in the grammar such as the direction of case and theta-role assignment (Stowell 1983, Travis 1984). This not only allows for cross-linguistic variation, such as that needed to capture the difference between head-final and head-initial languages, but allows for variation within a language. That is, certain projections may require their specifiers or complements to be on the right and others on the left, depending on their role in the syntax.

Following Kitagawa 1986, Koopman and Sportiche 1985, 1991, and Fukui and Speas 1986, I assume that all arguments of a verb are projected VP internally at D-structure, (6). In particular, I follow Stowell 1983 and Speas 1990 in assuming that 'external' arguments are projected into SpecVP, as opposed to a position adjoined to VP. This SpecVP subject position has properties similar to those assumed for SpecIP, although in certain languages SpecIP is still the position to which nominative case is assigned.

(6)



I assume a version of the Universal Theta Assignment Hypothesis (UTAH) (Baker 1988). The UTAH concerns itself with the projection of arguments into the syntax: which argument will be the subject and which the object? The idea behind the UTAH is that the thematic relations of the arguments, which are based on the lexical semantics of the verb, dictate which arguments are projected where. Baker's original version of the UTAH is shown in (7).

(7) The Uniformity of Theta Assignment Hypothesis (UTAH): Identical thematic relationships between items are

> represented by identical structural relationships between those items at the level of D-structure. (Baker 1988:46)

Interpreted strictly the UTAH in (7) states that arguments with identical thematic roles appear in identical positions in D-structure. However, it has been suggested that what is relevant for the syntax is not always the thematic roles, but the hierarchical relations between them. For example, Speas 1990 proposes that relative prominence is relevant for the UTAH in (8).

(8) The UTAH states that relative prominence in the Thematic Hierarchy must correspond to relative prominence in syntax. (Speas 1990:90)

I adopt the version of the UTAH in (8). Thus, arguments that are higher on the thematic hierarchy are projected higher in the syntax.[4] However, theta-roles do play a role in the syntax. In particular, I assume that the difference between unaccusative and unergative intransitive verbs is that in unaccusatives, the only argument is projected as sister to the verb, while in unergatives it is projected into SpecVP (for Russian see Chvany 1975, Pesetsky 1982a). With both types of verbs, there is only one argument and hence this argument must be the highest on the hierarchy. A theory of projection which depends solely on relative prominence incorrectly predicts that the arguments of unaccusative and unergative verbs are projected into identical positions in the syntax.[5]

The exact nature of case assignment is currently an issue of much debate. It has generally been assumed that all (overt) nouns receive case. Case can be assigned structurally, semantically, or inherently.[6] Inherent case is rather uninteresting in that it is idiosyncratically assigned by one particular verb to an argument which it theta-marks; for example, the Russian verb *upravljat'* 'govern' assigns instrumental case

[4] There is also the issue as to which thematic hierarchy to choose. All versions of the hierarchy assume that Agent is the highest role, but disagreement arises as to most other details. Efforts have been made to circumvent this problem, either by defining theta-roles in terms of prototypical roles (Dowty 1991) or by more closely examining the lexical semantics relevant for the construction of traditional thematic roles (Jackendoff 1990). Since the UTAH and thematic hierarchy play only a minor role here, I remain agnostic as to the exact nature of the thematic hierarchy.

[5] Grimshaw 1990 valiantly adopts a view of argument structure that does not refer to thematic roles but instead depends solely on relative prominence. However, she is forced to stipulate that unaccusatives are represented differently from unergatives in order to distinguish the two and project them differently.

[6] The picture is unfortunately not this simple in that some cases may have both structural and semantic components, e.g., the partitive in Russian is assigned to NPs in certain structural positions when they have a partitive meaning.

to its object, instead of accusative. Information about inherent case assignment is stored in the lexical entry of the particular predicate. Semantic case has been understudied and in some accounts is subsumed with inherent case in that it is assigned at D-structure in conjunction with theta marking (Speas 1990:16). The basic idea behind semantic case is that certain cases are assigned based on the meaning of the NP to which they are assigned. Finally, structural case is assigned to an NP in a particular structural position, e.g., the nominative case assigned to NPs in subject position. For Russian, I assume that the $V+I^0$ complex assigns nominative case to SpecVP.

Any noun which is not assigned case is ruled out by the Case Filter in (9).[7]

(9) *NP, if NP has phonetic content and no Case.

The nature of the Case Filter is currently under debate. There has been a move towards the checking of case, instead of its assignment in conjunction with the Case Filter (Chomsky 1992). The idea behind Case checking is that at some point in the derivation, the morphological case of nouns must be checked in certain positions. NPs with which checking fails are ill-formed, and the derivation is unfortuitous.

All nouns in Russian are overtly marked for case;[8] even the nominative, which in many languages is unmarked (e.g., see Mohanan 1990 on Hindi), is overtly marked. I assume that all NPs must be assigned case at some stage in the derivation.[9] An NP does not have to remain in the position where it is assigned case since this case is transmitted by the chain formed between the NP and its traces. This is a crucial assumption since many NPs are assigned case and then move to other structural positions in order to receive discourse function interpretation.

2.2 Lexical-Functional Grammar

Part II utilizes Lexical-Functional Grammar (LFG) (Bresnan and Kaplan 1982). These chapters investigate how the positioning of topic and focus and other syntactic structures often thought to involve movement operations are captured in a theory without phrase structure movement. This section outlines the basic theoretical assumptions of LFG; concrete analyses of Russian clauses can be found at the beginning of Part II. LFG separates information about grammatical functions from

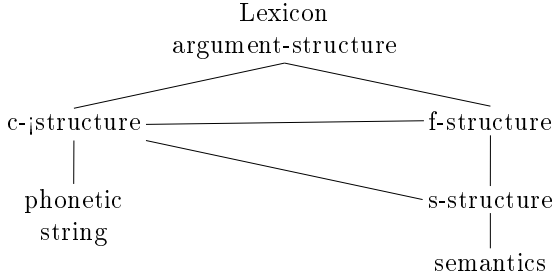
[7] Presumably *pro* and wh-traces are also affected by the Case Filter, although they have no phonetic content.

[8] Some loan words are not declined, e.g., *kenguru* 'kangaroo'.

[9] In some languages, such as Korean, a given NP may be assigned two cases, one semantic and one structural (Gerdts 1988, 1991). This does not occur in Russian.

that about the phrase structure, and grammatical function are not defined by phrase structure positions. The grammar can be represented as in (10).

(10)



Structural relationships are represented at c(onstituent)-structure, while grammatical function are represented at f(unctional)-structure. The c-structure is not derived from the f-structure, nor *vice versa*. However, c-structures determine certain properties of the f-structures by means of functional annotations on the c-structure. For more details of how the c- and f-structures interact, including example structures, see Part II.

There are relatively few formal well-formedness conditions on c-structures. C-structures represent the surface precedence and dominance relations in a language and show substantial cross-linguistic variation. The c-structure encodes only the surface form of a sentence; there are no traces and no PF movement.[10] The c-structure is built from a series of language-specific phrase structure rules, which formally need not conform to X′ structures, although the theory can constrain them in this way (see Part II for some necessary modifications). Perhaps one of the most unusual features of c-structures is that phrases can be headless; this often results where GB would claim that head-movement has occurred, e.g., a VP may dominate just an object NP and no V^0. One proposed c-structure constraint is a non-branching constraint which forbids structures like (11) in which a node dominates an identical node and nothing else; this prevents vacuous adjunction structures.

(11)

[10] There is a small amount of deviation from the c-structure to the surface form in the case of certain clitics. However, there are no major reorderings.

Unlike c-structure which encodes dominance and precedence relations, f-structures encode grammatical function and other information relevant to the syntax. Also unlike c-structures, f-structures are composed of attribute-value matrices. Attributes can be such entities as grammatical functions (SUBJ, OBJ, COMP), tense (TNS), and nominal features (CASE, NUM, GND, PRS). Each attribute must be assigned an appropriate value. Values are supplied by the lexical entries of the items and by functional annotations on the c-structure. F-structures can appear within other f-structures as the value of one of the outer f-structure's attributes. For example, the f-structure of a sentence containing a complement clause has an attribute COMP whose value is an f-structure corresponding to the complement clause.

There are three well-formedness conditions on the f-structure: functional uniqueness, completeness, and coherence.

(12) Functional Uniqueness: In a given f-structure, a particular attribute may have at most one value.

functional uniqueness guarantees that an attribute does not have more than one value. This, for example, rules out an f-structure in which the TNS attribute is specified as both PAST and PRS. This does not mean that a particular attribute may not receive its value from more than one source. As long as the values can unify, i.e., do not conflict, this is no problem. For example, if the PRS attribute of the subject is specified as first person both by the lexical entry of the subject noun and by the verb, the f-structure still satisfies functional uniqueness.

(13) Completeness: An f-structure is *locally complete* if and only if it contains all the governable grammatical function that its predicate governs. An f-structure is *complete* if and only if it and all its subsidiary f-structures are locally complete.

Completeness basically states that all of the grammatical functions for which the predicate subcatergorizes must be assigned values. This rules out clauses such as 'John likes' in which the OBJ attribute of the predicate is not assigned a value.

(14) Coherence: An f-structure is *locally coherent* if and only if all the governable grammatical function it contains are governed by a local predicate. An f-structure is *coherent* if and only if it and all its subsidiary f-structures are locally coherent.

Coherence requires every semantic form in the f-structure to be the PRED value of a grammatical function in that f-structure. This results in clauses like 'Inna goes house' being ill-formed because 'house' is not associated with any argument of the verb nor can it be interpreted as an adjunct.

The lexical entries of predicates contain information as to the arguments they take, e.g., the thematic roles of the arguments, as well as idiosyncratic information such as lexically-assigned case.[11] Thematic information is linked to grammatical functions via the Lexical Mapping Theory (LMT) which relates the argument-structure to f-structure. LMT is similar to the UTAH in GB theory in that more prominent arguments on the thematic hierarchy are associated with what are traditionally thought of as more prominent grammatical relations. However, the theory reflects the fact that there is not always a simple alignment between thematic roles and grammatical function, e.g., in locative inversion. LMT has been used to account for object asymmetries, applicatives, and locative inversion (Alsina and Mchombo 1994, Bresnan and Kanerva 1989, Bresnan and Moshi 1990). Each grammatical relation is defined by the features [±r] (semantically restricted) and [±o] (objective), as in (15).

(15)

Grammatical function	Features
SUBJ	[−r,−o]
OBJ	[−r,+o]
OBJ_θ	[+r,+o]
OBL_θ	[+r,−o]

Certain theta-roles are always associated with a given feature. For example, agents are always [−o], while themes are [−r]. This guarantees that agents are usually subjects and are never objects, and that themes are usually objects. Linking principles fill out the features specified by the argument structure in order to determine the grammatical function of the arguments. For example, when possible, the highest argument is assigned [−r] and the other arguments [+r], unless these designations conflict with the features already associated with the arguments. So, each argument has certain features specified by default rules, and the linking rules provide values which yield the final grammatical function of each argument.[12]

2.3 The Russian Language

Throughout this dissertation "Russian" refers to Contemporary Standard Russian. Contemporary Standard Russian differs in certain re-

[11] There is an additional module in LFG referred to as argument-structure (see Alsina 1993). Argument-structure is of primary importance in the formation of causatives and complex predicates and can be used to derive much of the information normally thought of as the lexical entry of a word.

[12] Grammatical functions of complements, such as COMPs and XCOMPs, are frequently ignored in discussions of linking rules. Ideally, the assignment of these functions to certain types of arguments can also be derived from the linking rules.

spects from Colloquial Russian (Yadroff 1992a, 1992b; Zemskaja 1973). For example, in Colloquial Russian constituents can raise out of finite clauses, complementizers can often be dropped, and many discontinuous constituents are allowed; these are not possible in Contemporary Standard Russian. These differences are occasionally alluded to and constructions from Colloquial Russian are used as additional evidence for certain arguments. The focus is on relatively simple declarative sentences, with some discussion of interrogatives and makes reference to complex sentences when they shed light on the structure of Russian and the interpretation of discourse functions. It is not that more complex sentences are not of interest: quite the contrary since the behavior of items in subordinate clauses often differs from that of main clauses. However, as with all work, there are many topics that must be set aside for further research.

It is impossible to go into a full excursus on the structure of Russian. However, a few basic facts are essential; for more details on the general structure of Russian see Pulkina and Zakhava-Nekrasova 1988 and Barnetová et al. 1979. As mentioned above, constituent order in Russian is not a reliable method of determining grammatical function. Subjects and objects can appear in either order with respect to one another and can either precede or follow the verb. In addition, there are no special markings on topics and foci, e.g., there is no morpheme which indicates topicalized constituents. However, there are certain regularities of phrase structure: NPs and PPs form continuous units, wh-words always appear clause initially, and material from a finite clause cannot raise into a higher clause.

Nouns are overtly case marked with one of six cases: nominative, accusative, genitive, dative, instrumental, and prepositional (certain nouns also have distinct locative and partitive forms). In addition, there are three gender classes, although the gender class distinctions are lost in the plural. gender classes are morphological divisions and do not reflect natural gender. A sample declension for feminine, masculine, and neuter nouns is shown in (16).

(16)

	Feminine	Masculine	Neuter
Singular:			
NOM	stran-a 'country'	sup 'soup'	mest-o 'place'
ACC	stran-u	sup	mest-o
GEN	stran-y	sup-a	mest-a
DAT	stran-e	sup-u	mest-u
PREP	stran-e	sup-e	mest-e
INST	stran-oj	sup-om	mest-om
Plural:			
NOM	stran-y	sup-y	mest-a
ACC	stran-y	sup-y	mest-a
GEN	stran	sup-ov	mest
DAT	stran-am	sup-am	mest-am
PREP	stran-ax	sup-ax	mest-ax
INST	stran-ami	sup-ami	mest-ami

Verbs agree with their subjects in person and number in the present and future tense and in number and gender in the past tense, shown in (17). Infinitives do not show agreement marking. In addition, verbs come in aspectual pairs, having both perfective and imperfective forms.[13] Perfectives have past and future forms; imperfectives have past and present forms, in addition to forming a compound future. The form *čitat'* in (17) is an imperfective; the perfective counterpart is *pročitat'*.

(17)

Infinitive:	čitat' 'read'		
Present:		Past:	
1 singular	čitaju	Fem. singular	čitala
2 singular	čitaeš'	Masc. singular	čital
3 singular	čitaet	Neut. singular	čitalo
1 plural	čitaem	Plural	čitali
2 plural	čitaete		
3 plural	čitajut		

Grammatical subjects of finite verbs are in the nominative case, and as a result verb agreement is only with nominatives, as in (18). Subjects can be pro-dropped, although Russian is not a pro-drop language in that the subject is usually overt, even when pronominal. Verbs agree with their subjects regardless of their relative order, i.e., pre- and post-verbal nominative subjects trigger verb agreement.

[13] Usually the perfective and imperfective forms are morphologically related. For example, perfectives are often formed by adding a prefix to the imperfective; this prefix sometimes changes the meaning of the verb. However, there are also verb pairs which are not morphologically related.

(18) a. Ja čitaju knigu.
I-NOM read-1.SG book-ACC
'I am reading a book.'
b. Deti čitajut knigu.
children-NOM.PL read-3.PL book-ACC
'The children are reading a book.'

There are a number of impersonal expressions which do not have grammatical subjects and show default agreement of third person, neuter, singular, as in (19).[14] In these sentences, the thematically highest argument appears in some case other than the nominative, usually the dative, as in (19b).

(19) a. Annu tošnilo.
Anna-ACC sick-NEUT.SG
'Anna was sick.'
b. Mne budet xolodno.
me-DAT be-3.SG cold
'I will be cold.'

In (19a) the sole argument of the verb *Annu* is in the accusative case, while in (19b) the sole argument *mne* is in the dative. These and other Russian syntactic structures are discussed in detail in Parts I and II.

[14] Alternatively, these can be thought of as having null expletive subjects with which they agree. See Babby 1989 for a subjectless account of these constructions.

Part I

3

Evidence for Configurationality

This chapter is concerned with subject-object asymmetries in Russian and other evidence that Russian does not have a flat phrase structure within the VP or above it. Russian is generally assumed to be SVO with an underlying structure like that of English to which scrambling applies. In chapter 5, I argue that Russian is VSO and that topic and focus constructions result in the common non-VSO orders (see in particular section 5.4). This VSO order is derived from leaving the subject in SpecVP and raising the verb past it to I^0. However, even though the subject follows the verb and remains in the VP, it does not behave identically to objects.[1] The relative prevalence of subject–object asymmetries in Russian suggests that the subject is in a different structural position than the object, i.e., the VP does not have a flat structure.[2] [3]

Russian has several putative properties considered diagnostic of non-configurational languages (Hale 1983): free word order, i.e., the word order does not encode grammatical functions, apparent pro-drop, no pleonastic subjects,[4] and a rich case system. These facts might

[1] There are a number of impersonal expressions in Russian. The arguments of these constructions are not grammatical subjects and are marked with non-nominative case. If these have a grammatical subject, it is a null expletive. The highest thematic argument frequently appears in initial position in impersonal constructions as a result of topicalization (section 5.4). See Bailyn 1991 on PredP.

[2] This contrasts with languages like Hungarian which have been argued to have a flat structure within the VP (Kiss 1987a, 1987b). Since subjects and objects are found in identical structural positions in flat structures, much of their behavior is identical. (See section 5.2.1 and references therein for more details.)

[3] See Bailyn 1991 for discussion of how the arguments of the verb must be ordered in Russian. He argues that subjects are projected higher than direct objects which are projected higher than indirect objects.

[4] Russian has few (overt) expletive subjects like those found in English, i.e., there are a large number of clauses with no overt nominative subject. Franks 1990 discusses the one type of clause in which overt expletives are optional found. An

lead one to presume that Russian has only limited underlying phrase structure, perhaps of the kind S $\longrightarrow$ XP*, V. However, although these phenomena are indicative of non-configurationality, they do not require it. In fact, the non-configurational properties of Russian are misleading: word order is not free since it encodes topic and focus information (this is the case with many, if not all, free word order languages), and pro-drop in Russian is quite limited, unlike in other Slavic languages, in that the unmarked case is to have overt pronominal subjects.

The connection between the underlying word order and that of the surface variants has not always been made explicit. This chapter examines the D-structure of Russian, arguing for certain structures and regularities. The S-structure, in particular the position of topics and foci, is examined in chapter 5, which shows how different word orders correspond to different discourse function interpretations. I argue that discourse-neutral arguments remain within the VP. When arguments move out of the VP, they are interpreted with different discourse functions depending on their landing site. This approach is similar to that taken for a number of relatively free word order languages in which the different word orders reflect differences in topic-focus interpretation, not in grammatical function (Aissen 1992, Diesing 1988, Kiss 1987a, *to appear*, Kroeger 1991, Rudin 1985, et al.).

Several linguists, whose work did not focus on word order, have assumed an underlying structure for Russian similar to that of English (Pesetsky 1982a, Neidle 1982, 1988). I propose that in Russian all of the verb's arguments are projected VP internally at D-structure. The basic phrase structure is shown in (1) and is justified in the ensuing sections. SpecCP is the landing site for wh-phrases, all of which front in multiple wh-questions (Rudin 1988, 1989), and for focus phrases in *li* yes-no questions (section 6). Subjects appear in SpecVP, where they c-command the object; not only are (unergative) subjects projected into this position at D-structure, but all subjects must move to SpecVP where they are assigned nominative case (Koopman and Sportiche 1985, 1991, Fukui and Speas 1986). The verb raises past the

example is given in (i). See Franks 1990 on parameterizing the occurrence of expletive subjects, and Babby 1989 and Billings 1993 on expletives in Ukrainian.

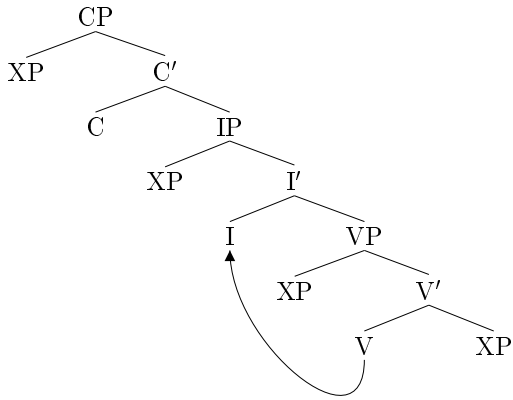
(i) (Ėto) prijatno čto my guljaem v parke.
it nice that we walk in park
'It is nice that we are walking in the park'

Loren Billings (p.c.) cites the use of *ono*, the third person neuter nominative pronoun, as an expletive in Colloquial Russian. An example is given in (ii).

(ii) Ono i ponjatno; v ix kongresse dojarok net.
it emph. clear in their congress milkmaids not
'It's plain to see; in their congress there aren't any milkmaids.'

subject in SpecVP to I^0 where it gets tense and agreement features and can case mark SpecVP with nominative case (Travis 1984, Grimshaw 1991). SpecIP is not subject position in Russian; the subject remains in SpecVP unless it moves into topic or focus position. In this chapter, I am primarily concerned with justifying the structural difference in subject and object positions for Russian and the distribution of material in the I^0 and V^0 positions.

(1)



One major issue for phrase structure and configurationality is whether a given language has a VP. Arguments for a VP generally fall into two categories: first, showing that the verb and its object act as a unit separate from the subject; second, showing that there are asymmetries in the behavior of subjects and objects that can be explained if the object is governed by the verb while the subject is not, and if the subject c-commands the object.

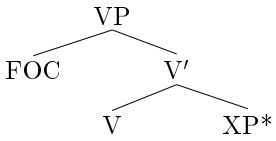
Speas 1990 develops these ideas in detail for a number of languages, including several that are often considered to be non-configurational, e.g., Navajo. She argues that all languages have a hierarchically arranged D-structure. The arguments of the verb are projected into the VP according to the UTAH so that thematically higher arguments are higher in the tree than thematically lower ones. Under this conception, the structure of the VP at D-structure is essentially identical for all languages, modulo differences in their argument structures.

Mohanan 1982 takes a different approach in his Lexical-Functional Grammar analysis of Malayalam. Although Malayalam does exhibit certain subject-object asymmetries, he argues that these are not a result of asymmetric configurational differences between subjects and objects but from a difference in grammatical function. This type of approach can be taken for Russian to account for many of the differences between subjects and objects. However, even in a theory where

grammatical functions are defined independently of phrase structure, Russian requires a configurational structure to account for syntactic groupings of constituents, albeit less than in a theory which uses structure to encode grammatical relations. This problem is the subject of chapters 7, 8, and 9 which examine the phrase structure of Russian in LFG.

Even in GB, projecting the subject higher than the object is not necessarily a universal. Kiss 1987a, 1987b, *to appear* argues that although Hungarian has a VP, the arguments of the verb are generated in a flat structure within V′.[5] There is some hierarchical structure within the VP, but it does not encode grammatical functions: SpecVP is focus position, while the arguments are within V′, as in (2) (section 5.2.1).

(2)



In (2), the arguments of the verb mutually c-command one another. This lack of structural asymmetry results in essentially no subject-object asymmetries in Hungarian.[6] For example, the subject and object can appear in either order after the verb without a difference in meaning, and both can appear in topic and focus positions. There are no VP rules which involve the verb–object, but not verb–subject: there is no VP preposing; deletion rules for the VP or V′ are insensitive to the arguments in them; idiom interpretation can be affected by subjects and by objects. Hungarian has no Empty Category Principle (ECP) effects: wh-movement out of embedded clauses and extraction out of interrogative wh-phrases is identical for subjects and objects. There are no superiority effects, no weak-crossover phenomena, and no difference in coreference possibilities which would be Principle C violations in English. This identical syntactic behavior of subjects and objects is expected if they have identical syntactic positions since these asymmetries are derived from the difference in structural position of subjects and objects (section 3.6).[7]

[5]Kiss 1987a has the verb and its arguments form an S; S′ contains the focus and S. In Kiss *to appear* this schema is recast within X′ syntax, but the generalizations concerning the projection of arguments remain the same. The X′ version is used here for ease of comparison.

[6]Not all linguists have analyzed Hungarian as having a flat structure for the arguments of the verb. See, for example, the discussion in Speas 1990:173–190 who argues for a hierarchical structure.

[7]Kiss 1987b mentions two potential problems for a flat structure account: binding of anaphors and the distribution of PRO. These show differences in the behavior

Given the VP-internal subject hypothesis, it is possible to have a language with subject-object asymmetries, but in which the verb and the object do not form a constituent exclusive of the subject. This would be expected if the subject position is internal to the VP but the finite verb is in I^0. That is, if the subject is projected into SpecVP and the object is sister to the verb, one might expect subject-object asymmetries but not perhaps the full range of those found in English.

Russian demonstrates several of the classic phenomena of a VP language. It allows verbs and their objects to be coordinated, a property that generally holds only of constituents (section 3.1). In Colloquial Russian, a verb and its objects can be scrambled, but not a verb and its subject (section 3.2).[8] Russian has subject-object asymmetries that are often explained by locating the subject outside the VP and the object inside it (Pesetsky 1982a). For example, the genitive of negation in Russian has been explained by an asymmetry between the position of the subject and the object, where NPs which are sister to the verb can be marked by the genitive of negation (section 3.3).

In the remainder of this chapter, tests for structural differences in the position of subjects and objects are applied to Russian in a fashion similar to that of van Steenbergen 1989 for Finnish. I examine the following arguments for an underlying configurational structure: coordination, scrambling, the genitive of negation, the distribution of negative markers and tensed verbs, pronominal coreference, extraction,

---

of subjects and objects: a subject can bind an object anaphor, but not *vice versa*, and PRO is only found with subjects. She argues that the conditions on anaphor binding are sensitive to an argument hierarchy: nominatives < accusatives < datives < instrumentals < locatives.

(i) An argument, p, can bind another argument, q, if
  (1) they are coindexed,
  (2) p is in an argument position,
  (3) p c-commands q,
  (4) if q also c-commands p, p precedes q on the argument hierarchy.

The definition in (i) allows the arguments to be in a flat structure where they mutually c-command one another since the argument hierarchy is a reflection of the thematic hierarchy, not the syntactic structure. Finally, concerning the distribution of PRO, in embedded infinitivals only subjects can be PRO, a restriction which in English is derived from the ungoverned nature of this position. Kiss adopts a proposal by Bouchard 1982 that the distribution of PRO is a result of its being caseless. The absence of an I^0 [+agr] node in tenseless clauses results in the subject argument not having realized case because nominative case is realized as agreement on the verb; this forces PRO to appear as the subject. These proposals concerning anaphora and the distribution of PRO allow Kiss to maintain a flat structure amongst the arguments in V′, which in turn explains the general lack of subject–object asymmetries.

[8] If the verb is intransitive, the verb and its subject can scramble since they comprise the entire IP.

weak crossover, and superiority effects. Positing an underlyingly configurational structure in a relatively free word order language requires an explanation for the range of alternative word orders (Speas 1990). The underlying structure only gives a subset of the surface word orders. The structural encoding of topic and focus functions in Russian provides the motivation for the alternative orders (chapters 4 and 5).

3.1 Coordination

Coordination is often used as a test for constituency. Unfortunately, it is a tricky test because it is not the case that only like constituents can be coordinated (Sag et al. 1985, Peterson 1981). However, Sag et al. argue that only constituents can be coordinated, using phenomena such as Right Node Raising to account for counter-examples. So, if two items can be coordinated, they are assumed to both be constituents, although their identity may not be known. Using this, I argue that finite verbs are in I^0 and infinitives in V^0 since coordination of I′s and VPs is possible in Russian. Consider the coordinate structure in (3).

(3) Deti [čitali knigi] i [smotreli fil'm].
children read books and *?pro* watched film
'The children were reading the books and watching the film.'

If finite verbs move out of their base generated position VP to I^0 in order to get tense and agreement features, then (3) makes a statement about the existence of I′, i.e., the projection that includes I^0 with the finite verb but excludes SpecIP. However, (3) might not contain coordinated I′s, i.e., the tensed verb and its object. The second conjunct, *smotreli fil'm*, could be an IP with a null subject (indicated here by *pro*). Given the pro-drop capacity of Russian, it is necessary to show that two I′s, as opposed to two sentences, are being coordinated.

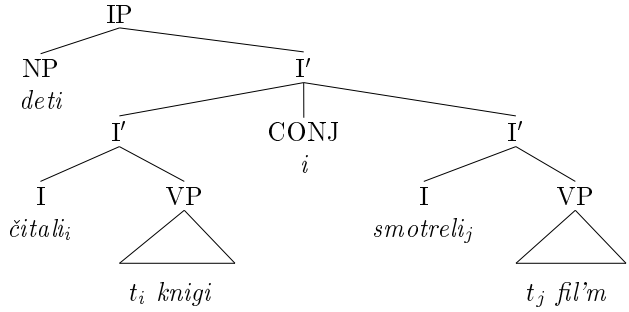
Pro-drop depends on discourse function: null subjects are permitted only when they are the topic of their clause. If a sentence with apparent coordinated I′s has a subject that is not a topic, then the second conjunct cannot have a pro-dropped subject and hence must be an I′.

(4) a. Nekotorye deti [čitali knigi] i [smotreli fil'm].
some children read books and watched film
'Some children were reading the books and watching the film.'

b. Nikto [ne čital knig] i [ne smotrel fil'mov].
no one not read books and not watch films
'No one read books and watched films.'

In (4) the subjects do not refer to specific entities and are thus unlikely to be topics. However, the sentences are grammatical, suggesting that these are coordinated I′s.[9] (5) shows the coordination structure of (3).

(5)



Consider the coordinated infinitives in (6) which show that VPs can be coordinated since infinitives remain in the VP.

(6) a. Ja budu ves' den' [pisat' pis'ma] i
I will all day write-INF letters and
[čitat' knigi].
read-INF books
'I will write letters and read books all day.'

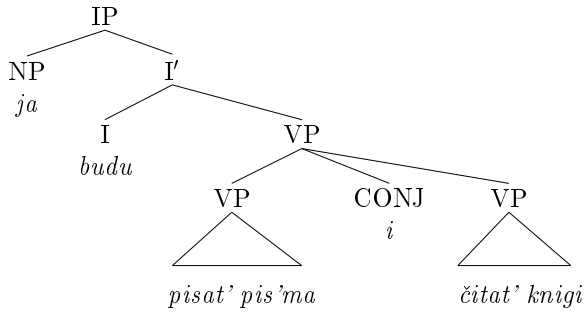
b. Mne bylo [trudno najti kartu
me-DAT was difficult-PRED.ADV find map
Moskvy] i [nevozmožno najti
Moscow and impossible-PRED.ADV find
kartu Leningrada].
map Leningrad
'It was difficult for me to find a map of Moscow and impossible to find one of Leningrad.'

In (6a) the tensed auxiliary *budu* is in I^0. The two VPs *pisat' pis'ma* and *čitat' knigi* are coordinated. Since there is only one tensed verb *budu*, (6a) cannot be an example of pro-drop. (6b) involves two modal predicate adverbs, whose tense is marked by an auxiliary, embedded under a single auxiliary. The structure for (6a) is shown in (7).[10]

[9] The repetition of the negative marker *ne* before each of the finite verbs can be used to argue for the placing of the tensed verb in I^0 (section 3.4). If *ne* is not repeated, the second conjunct will not be in the scope of negation.

[10] There are two issues with the structure in (7). The first is the placement of the subject. This subject is topicalized and thus should be adjoined to IP (section 5.1); the same is true of the subject in (5). The second is that instead of having the auxiliary in I^0, it may move to I^0 from V^0. That is, the auxiliary may head a VP

(7)



To conclude, the coordination facts in Russian suggest that an infinitive and its arguments form a constituent which does not include the finite verb, i.e., there is a VP constituent and a separate I′ with an I^0 position for inflected verbs.

3.2 Scrambling

scrambling only affects maximal projections in Russian (Yadroff 1992a, 1992b). Here, scrambling refers generally to movement of a constituent, leaving aside what motivation this movement might have. If there is a VP in Russian, we would expect to be able to scramble it as a unit.

This prediction is difficult to test in Contemporary Standard Russian because scrambling is clause-bounded. Phrases cannot be scrambled out of their minimal finite clause, although they can be scrambled out of infinitival clauses. Within a single clause, it is difficult to determine whether a VP has been scrambled, because it could be the case that the other constituents are the ones that have moved. However, Colloquial Russian allows scrambling across clause boundaries and thus provides a test for constituency (Zemskaja 1973). Although movement of constituents into higher clauses in Colloquial Russian can be to a number of landing sites, these scrambled phrases most commonly appear preverbally. Consider the sentences in (8).

(8) a. ?Ja [učit'sja v novoj škole] slyšal, on budet.
I [study-INF in new school] heard he will
'I heard that he will study in a new school.'
(Yadroff 1992b:3)

b. Mne [otpustit' Katju odnu] kažetjsa, čto
me let go-INF Katja alone seem that

which takes a VP complement whose head is the lexical verb. Such an analysis does not affect the proposals made here.

bylo by bezumiem.
would be insane
'It seems to me that it would be insane to allow Katja to go alone.' (Yadroff 1992b:7)

c. *Ja [pošel v školu] skazal, (čto) on.
I [went to school] said that he
'I said that he had gone to school.' (Yadroff 1992b:3)

The sentences in (8) have very similar meanings. However, scrambling the infinitival VP[11] in (8a)–(8b) is grammatical, but scrambling the finite I′ in (8c) is not.[12]

The difference between (8a)–(8b) and (8c) lies in the phrase structure of Russian. Notice that in (8a) the tensed auxiliary *budet* remains in the lower clause; that is, *budet* is still within the embedded CP while the infinitive and goal *učit'sja v novoj škole* has been scrambled so that they appear before the matrix verb *slyšal*. In (8b) the tensed modal *smožet* remains, while its complement *pojti v školu* has scrambled into the matrix clause.[13] In contrast, in the ungrammatical (8c) the tensed verb *pošel* is part of the scrambled constituent *pošel v školu*. This

[11]Presumably the trace of the subject in SpecVP is scrambled with the infinitive and its arguments in (8a)–(8b). This scrambling of a constituent with an initial trace may seen odd. However, McCloskey 1991 argues that such structures are permissible in English: 'I asked Fred to leave, and [$_{VP}$ leave] I think [$_{CP}$ *t* [$_{IP}$ he [did —$_{VP}$]]]' (Chomsky 1986:20, cited in McCloskey 1991:293). If only the verb and its complements are scrambled in the Russian sentences, exclusive of the subject trace, these form a V′, a non-maximal projection, and are thus ineligible for scrambling.

[12]Müller and Sternefeld 1993 cite an example which superficially contradicts this generalization in that a finite verb and its object appear in a higher clause, as in (i).

(i) No ja ix postavila pomnju, čto v škaf.
but I them put remember that in cupboard
'But I remember that I put them in the cupboard.'

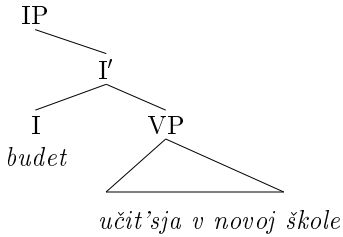
At first, it appears that the verb and its object, *ix postavila*, have scrambled into the upper clause. However, the sentence becomes ungrammatical if an overt subject, e.g., the pronoun *ja* 'I', appears in the lower clause, following the complementizer *čto*. If (i) involved scrambling of a VP or I′, as opposed to scrambling of IP, then this should be possible.

[13]It is not that case that all VPs can scramble into the higher clause. For example, most speakers find (i) relatively ungrammatical (Yadroff finds it grammatical). In (i) the tensed conditional copula, *bylo by*, remains in the lower clause with the adjective, while the infinitival small clause has moved to the higher clause.

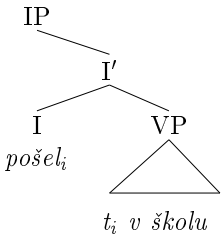
(i) ??Ja [pojti v školu] skazal, (čto) on segodnja
I go-INF to school said that he today
ne smožet.
not able
'I said that he won't be able to go to school today.' (Yadroff 1992b:3)

suggests that tensed elements in Russian, both auxiliaries and main verbs, undergo verb raising to I^0 and as a result are located not in the verb phrase, but in I^0 (Travis 1984, Grimshaw 1991). Therefore, the tensed verb does not form a maximal projection with its arguments, instead it forms an I′ and thus cannot scramble with them. On the other hand, a non-finite verb remains in the verb phrase and forms a maximal projection with its arguments.[14]

(9)



(10)



The structure in (9) corresponds to that of the lower clause in (8a). The verb *učit'sja* and the prepositional phrase *v novoj škole* form a maximal projection, a VP, while the tensed element is outside this projection. In contrast, the structure in (10) corresponds to that of the lower clause in (8c). The verb is in I^0, not V^0. Since I′ is not a maximal projection, the verb and the prepositional phrase cannot scramble. These facts differ from the coordination facts in that coordination is possible as long as the elements form a constituent, while scrambling is restricted to maximal projections (section 3.1).

The Colloquial Russian scrambling data in this section argue for the existence of a VP, in which infinitive verbs are found, and for the movement of V^0 to I^0 in tensed clauses.

3.3 The Genitive of Negation

In Russian most objects appear in the accusative case, as in (11a). However, they may be marked with the genitive when in the scope of

[14] Another possibility is that infinitives form their own IP projection with the infinitival ending as head of I^0. Even under such an analysis, infinitivals are XPs, not X′s.

verbal negation, as in (11b) (section 3.4 discusses the syntactic structure of negation). This is referred to as the genitive of negation (Babby 1980, Chvany 1975, Neidle 1982, 1988, Pesetsky 1982a; see also section 8.2). The genitive is not obligatory in Russian;[15] semantic, syntactic, and morphological factors interact to determine whether the genitive or the accusative is preferred in a given context (Timberlake 1986).[16] In general, an NP in the genitive is interpreted as indefinite, while one in the accusative is definite.

The distribution of the genitive of negation in Russian has been linked to the existence of a VP (Pesetsky 1982a), although others have argued that it is linked to the grammatical function object (Neidle 1988). The basic idea behind these accounts is that only constituents that are sister to the verb at D-structure can undergo the genitive of negation, while NPs in other positions cannot.

3.3.1 The Data

(11) shows that the direct objects of transitive verbs can appear in the genitive under sentential negation, but their subjects cannot.

(11) a. Mal'čik ne vidit knigu.
boy not see book-ACC
'The boy does not see the/(?a) book.'

b. Mal'čik ne vidit knigi.
boy not see book-GEN
'The boy does not see a/(?the) book.'

c. *Mal'čika ne videt knigu.
boy-GEN not see book
'The/a boy does not see the book.'

(11a) is a simple negated transitive sentence; the subject is in the nominative case and the object in the accusative. In (11b) the object is marked with the genitive case and the sentence is grammatical. However, in (11c) where the subject is marked with the genitive, the sentence is ungrammatical. (11c) does not improve if the object is in the genitive, nor is it improved by changes in word order. Note that the verb in (11c) is the default third person singular since Russian verbs only agree with nominative subjects.

[15] In some Slavic languages, e.g., Polish and Slovenian, the genitive of negation is obligatory.

[16] There is some disagreement as to whether the genitive of negation is being lost in Russian. Most works claim that the genitive of negation is on the decline, but Chvany 1990 reports that statistical studies show little change in use over the past century.

The genitive of negation can apply to the subjects of certain intransitive verbs, namely unaccusatives (in contrast with unergatives).

(12) a. Ne pojavilis' studenty.
not show up-PL students-NOM
'The students didn't show up/No students showed up.'
(Pesetsky 1982a:66)

b. Ne pojavilos' studentov.
not show up-SG students-GEN
'No students showed up.' (Pesetsky 1982a:66)

(13) *V pivbarax kul'turnyx ljudej ne p'et.
in beerhalls cultured-GEN people-GEN not drink-sg
'Cultured people do not drink in beerhalls.
(Pesetsky 1982a:43)

In (12b) the genitive is possible with the unaccusative verb *pojavilos'*.[17] However, the unergative verb, *p'et*, in (13) cannot appear with a genitive.

Finally, subjects of passives can be marked with the genitive of negation.

(14) Ni odnogo goroda ne bylo vzjato.
not one city-GEN not was taken
'Not one city was taken.' (Chvany 1975:184)

These arguments are not grammatical subjects when they appear in the genitive instead of the nominative (Pesetsky 1982a, Neidle 1982, 1988). The arguments for this include: in unmarked word order they come after the verb; they do not agree with the verb; they cannot control the subject of an adverbial participle; they cannot control reflexives. The grammatical function of the genitive argument is not clear, although Pesetsky places it in object position and Neidle assigns it the function object; no false predictions are made under this assumption.

There is an additional fact concerning the distribution of the genitive of negation: the genitive of negation never applies to lexically case marked NPs. Babby 1986 attributes this to a hierarchy of case assignment in which lexical case takes precedence over semantic case which takes precedence over structural case. Alternatively, the assignment of lexical case may be a requirement for theta-marking; if the lexical case is not assigned, the NP will not be theta-marked, thus violating the Projection Principle.

[17] Pesetsky describes a difference in meaning between (12a) when the subject is nominative and (12b) when the subject is genitive; this is reflected in the glosses.

(15) a. On pomogaet mal'čiku.
he helps boy-DAT
'He is helping a boy.'

b. *On ne pomogaet mal'čika.
he not help boy-GEN
'He is not helping a boy.'

c. On ne pomogaet mal'čiku.
he not help boy-DAT
'He is not helping a boy.'

So, the object of a verb like *pomogat'* which is assigned dative, as in (15a), cannot undergo the genitive of negation, as in (15b).[18] Instead, the object always appears in the dative, regardless of the presence of negation.

3.3.2 Genitive Time Adverbials

The above data might indicate that thematic roles determine the domain of the genitive of negation. Since it affects objects and unaccusative subjects, the genitive of negation may be sensitive to themes and patients, but not agents. However, accusative time adverbials may also be marked with the genitive, as in (16b) (Chvany 1975, Pesetsky 1982a).[19] The time adverbial is neither an object, as would be indicated by a grammatical function account, or a theme, as indicated by a thematic role account. Pesetsky 1982a uses this as support for his structural account of the distribution of the genitive of negation.

[18] Pesetsky 1982a attributes this to genitive of negation NPs being Case resistant: for him these genitive phrases are QPs, and QPs cannot occur in positions in which they must be assigned case. Neidle 1988 accounts for this as a conflict in case assignment where the lexical entry of the verb requires the object to be dative, while the genitive of negation requires it to be genitive; these two cases cannot be unified, and thus the assignment of the genitive of negation is impossible (section 8.2).

[19] Distance adverbials that show up in the accusative can also undergo the genitive of negation, as in (i). These distance adverbials are more argument-like than their time adverbial counterparts and may be objects. For example, the use of the prefix *pro-* on the verb in (i) essentially forms a transitive verb. See Fowler and Yadroff 1993 for discussion of accusative adverbials in Russian. They propose that these adverbials fall into three classes: arguments, quasi-arguments, and non-arguments, on the basis of the features [±θ-role] and [±referential].

(i.a) On ne proexal ni odnu milju.
he not drive not one mile-ACC
'He didn't drive a single mile.'

(i.b) On ne proexal ni odnoj mili.
he not drive not one mile-GEN
'He didn't drive a single mile.'

(16) a. Ja ni odnu minutu ne spal.
I not one minute-ACC not slept
'I did not sleep a single minute.' (Pesetsky 1982a:92)

b. Ja ni odnoj minuty ne spal.
I not one minute-GEN not slept
'I did not sleep a single minute.' (Pesetsky 1982a:92)

However, these bare accusative adverbials are not a result of the genitive of negation, but a separate phenomenon, the partitive (Franks and Dziwirek 1993).[20] Franks and Dziwirek 1993 use cross-Slavic evidence to argue that the genitive marking on adjuncts, like that in (16), is an instance of the partitive genitive. They argue that the partitive must be licensed in order to appear; that is, the partitive case is assigned by a null quantifier and this quantifier must be licensed like other null elements in the syntax.[21] Certain verbs inherently license partitives, e.g., *pit'* 'drink', while others can never occur with a partitive object, e.g., *est'* 'eat'.[22]

Verbs which inherently allow partitives, such as *pit'* in (17), cannot themselves license the null quantifier of the partitive phrase. However, the presence of negation or the perfective prefix in conjunction with such a verb can license the null quantifier.

(17) a. *On pil čaju.
he drank-IMP tea-PART
'He was drinking some tea.' (Franks and Dziwirek 1993:300)

b. On vypil čaju.
he drank-PERF tea-PART
'He drank up some tea.' (Franks and Dziwirek 1993:299)

c. On ne pil čaju.
he not drank-IMP tea-PART
'He didn't drink some tea.' (Franks and Dziwirek 1993:301)

In contrast, with verbs which do not inherently license partitives, neither the presence of negation nor a perfective prefix licenses a partitive object, as in (18).

[20] For most Russian nouns, the partitive case morpheme is identical to that of the genitive, making the two difficult to distinguish. The distinct partitive case morphology seems to be decreasing in frequency so that most partitives are marked with the 'standard' genitive.

[21] Neidle 1982, 1988 and Pesetsky 1982a also assume that the partitive involves a null quantifier.

[22] The judgments here are very subtle. See Klenin 1978 for discussion of the partitive.

(18) a. *On el supu.
he ate-IMP soup-PART
'He was eating some soup.'
b. *On s"el supu.
he ate-PERF soup-PART
'He ate up some soup.' (Franks and Dziwirek 1993:299)
c. *On ne (s")el supu.
he not eat-IMP(PERF) soup-PART
'He did not eat (up) some soup.'
(Franks and Dziwirek 1993:301)

Franks and Dziwirek discuss structures in which negation licenses the partitive directly. Partitive direct objects, like *supu* in (18c), cannot be licensed directly by the negative marker because the verb is an intervening head between negation and the object. However, negation itself can license the partitive,[23] or more precisely the null quantifier that forms part of the partitive phrase, if there is no intervening head between the negation and the partitive constituent, as in (16b). This is independent of the type of verb, i.e., whether the verb inherently licenses partitives or not. Franks and Dziwirek propose that time adverbials are adjoined to VP, and thus there is no head intervening between them and the negation.[24]

How do we know that the time adverbials are partitives instead of instances of the genitive of negation? None of the nouns which appear as accusative adverbials have a distinct partitive case; so, the morphology is of no help. Franks and Dziwirek provide two arguments that these are partitives, one semantic and one cross-Slavic. The semantic argument is that these genitive adverbials have meanings like partitives and not like genitives of negation. These genitive time adverbials have a less-than reading, unlike their accusative counterparts and unlike true genitives of negation. Consider the Polish examples in (19).

(19) a. Nie spałam godzinę, ale dwie (godziny).
not sleep hour-ACC but two (hours)
'I did not sleep one hour, but two.'
(Franks and Dziwirek 1993:291)

[23] For nouns which have morphologically distinct partive genitive and simple genitive forms, the partitive genitive is more common in these constructions than the simple genitive.

[24] Pesetsky 1982a analyzes these time adverbials as sister to the verb, but this account runs into difficulty when both an object and a time adverbial are present since both cannot be direct sisters to the verb, as in (21).

b. *Nie spałam godziny, ale dwie (godziny).
not sleep hour-GEN but two (hours)
'I did not sleep one hour, but two.'
(Franks and Dziwirek 1993:291)

c. Eva nie kupiła kilograma jabłek,
Eva not buy kilogram-GEN apples
a trzy (kilogramy).
but three (kilograms)
'Eva did not buy a kilogram of apples, but three.'
(Franks and Dziwirek 1993:290)

The accusative adverbial *godzinę* in (19a) allows a reading in which the sleeping may have lasted more than an hour. However, the genitive adverbial *godziny* in (19b) only allows a less than reading which is typical of partitives. A true genitive of negation, such as that of the object *kilograma* in (19c) is like the accusative adverbial in that it allows a reading in which more than one kilogram is bought; it is not constrained to a less than reading.

The cross-Slavic argument in favor of the partitive adverbials analysis is that there is a correlation between languages which have partitives and those that have genitive adverbials, while no such correlation exists between the genitive of negation and genitive adverbials. Russian and Polish have partitives, genitive adverbials, and the genitive of negation. Czech has no partitives, no genitive adverbials, and no genitive of negation. Serbo-Croatian has partitives and genitive adverbials, but no genitive of negation. Finally, Slovenian has the genitive of negation, but no partitives and no genitive adverbials. Note, however, that Franks and Dziwirek state that there is a partitive in Slovenian, although it is very restricted; this suggests that the link between the partitive and the presence of genitive adverbials may not be as direct as originally thought.

I follow Franks and Dziwirek 1993 in analyzing the genitive adverbials as an instance of the partitive, not as a result of the genitive of negation as Pesetsky 1982a proposed, although there remains the issue of how the different partitives are licensed.

3.3.3 Analysis

So, is sister to V^0 the domain of the genitive of negation in Russian? Given the evidence from Pesetsky 1982a and Neidle 1988, the answer appears to be yes, once the adverbials are factored out. Although there are certain thematic correlates as to what can undergo the genitive of negation, there is evidence that the domain is a structural, not a

thematic, one. Agents never undergo the genitive of negation because they are always projected into SpecVP and thus are not sister to V^0. Themes undergo the genitive of negation because they are projected as sister to V^0. However, if we assume a somewhat modified version of the UTAH which states that relative prominence in Argument Structure corresponds to relative prominence in the syntax (section 2.1; King 1993b, adapted from Speas 1990), then it should be possible to have arguments with thematic roles other than theme be projected as sister to V^0. This is the case. In particular, the experiencers of certain psych verbs in Russian are sister to the verb and can undergo the genitive of negation, as in (20).[25]

(20) Ni odin student ne udivil učitel'nicy.
Not one student not surprise teacher-GEN
'Not one student surprised the teacher.'

So, the object of a verb, i.e., the argument which is sister to V^0, can undergo the genitive of negation regardless of its thematic role. Not only do themes and patients appear in the genitive of negation, so can experiencers when they are projected as sister to V^0.

The next question is why the genitive of negation should be restricted to phrases which appear in this position. Neidle 1988 and Bailyn 1991 argue that the genitive is assigned to any argument under the scope of negation when the predicate is associated with a quantificational feature [+q].[26] Unlike Pesetsky's 1982a account, these accounts are independent of the presence of a null quantifier.[27] This feature marks the scope of negation and helps explain why objects under negation in the genitive frequently have an indefinite, non-specific reading, while those in the accusative do not (the semantic correlates of the

[25] In these psych verbs, the 'theme' is associated with Causation which makes it more prominent in the Argument Structure than the experiencer. As a result, the theme is in SpecVP, while the experiencer is sister to V^0.

[26] The exact wording of Bailyn's account is as follows: Assign genitive case to any case-bearing argument in SpecVP under the scope of negation whenever the higher Pr^0 is associated with the quantificational feature [+q] (Bailyn 1991:70). For Bailyn, SpecVP is the position into which themes are projected and Pr^0 is the position in which finite verbs appear. For details of Neidle's account, see section 8.2.

[27] Neidle 1988 uses null quantifiers in her analysis of partitives, but argues that the partitive and genitive of negation cannot be collapsed since they do not occur in identical situations (section 3.3.2). Bailyn 1991 rejects the use of null quantifiers for both the partitive and the genitive of negation but states that the two are not the result of a single rule, although they have similar structural domains.

Steve Franks (p.c.) points out that another problem facing the null quantifier analysis of the genitive of negation is how to account for clauses in which a constituent with an overt quantifier undergoes the genitive of negation.

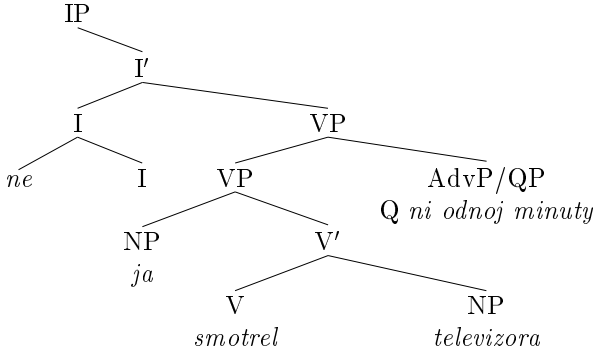
genitive of negation in Russian are extremely complex and are not explored here). What is of importance here is the difference in behavior between arguments which are projected into object position and those into subject position: the genitive of negation only affects arguments which are projected as sister to V^0.

The two types of genitive marking under negation discussed here are illustrated in (21). In (21), the direct object *televizora* and the time adverbial *odnoj minuty* are in the genitive. The object appears in the genitive due to the genitive of negation, while the genitive adverbial is a result of the partitive licensed by the negation.

(21) Ja ni odnoj minuty ne smotrel televizora.
I not one-GEN minute-GEN not watch television-GEN
'I didn't watch (any) television for a single minute.'

(21) has a D-structure approximately like that in (22). The object is sister to the verb while the adverb is adjoined to VP. The negation in I^0 licenses the genitive of negation on the object and the null quantifier in the AdvP, resulting in the partitive genitive.

(22)



In conclusion, the genitive of negation applies within a certain structural domain, namely sister to V^0, excluding material that appears outside this position.[28] There are other restrictions on its application, as well as conditions which may encourage or discourage its use in a given context, but these conditions are irrelevant to the argument made here that arguments projected into object position behave differently from those projected into subject position. There remain a number of serious

[28] The structure in (22) raises a question with respect to the genitive of negation: why do subjects never undergo the genitive of negation even though they are projected into SpecVP? In section 8.2 an account of the genitive of negation is proposed which does not make reference to the phrase structure position sister to V^0, but to underlying semantic roles and to case assignment. Under such an account, it is the underlying semantics of certain subjects which blocks their undergoing the genitive of negation.

questions, including why negation should be so intimately connected with this phenomenon, as well as with the licensing of partitives.

3.3.4 Phrases with Similar Distributions

There is additional evidence for a difference in behavior between arguments projected as sister to V^0 and those projected into SpecVP, i.e., a difference between subjects and objects. For more details on the internal structure and syntactic distribution of *po* and numeral phrases in Russian see Franks 1992. Pesetsky 1982a argues that there are two other classes of constituents whose domain is defined as sister to V^0. These are *po* distributive phrases and non-agreeing numeral phrases which take genitives. The arguments for the distribution of these phrases are similar to those for the genitive of negation. In addition to having the same D-structure distribution, these phrases are also similar to the genitive of negation in that they cannot appear in positions that are lexically case marked. However, the distribution of these phrases is not dependent on the presence of negation, as is evidenced by the examples below.

3.3.4.1 *Po* Phrases

Distributional quantification is expressed by the preposition *po* followed by an NP in the dative, as in (23).

(23) Ja dal mal'čikam po jabloku.
I gave boys PO apple-DAT
'I gave the boys an apple each.' (Pesetsky 1982a:69)

As with the genitive of negation, *po* phrases can appear as direct objects, as in (23), subjects of passives, as in (24a), subjects of unaccusatives, as in (24b), and where bare accusative adverbs would appear, as in (24c).[29]

(24) a. Každyj den' po gorodu bylo vzjato vragom.
each day PO city-DAT was taken enemy-INST
'Each day a (different) city was taken by the enemy.'
(Pesetsky 1982a:70)

b. Po jabloku upalo s každogo dereva.
PO apple-DAT fell from each tree
'An apple fell from each tree.' (Pesetsky 1982a:70)

c. Ja spal po času v den'.
I slept PO hour-DAT in day
'I slept an hour per day.' (Pesetsky 1982a:70)

[29] Pesetsky comments that the data for *po* phrases are not as clear-cut as those for the genitive of negation.

However, *po* phrases cannot appear as subjects of transitives, as in (25a), or of unergatives, as in (25b).

(25) a. ??Po studentu ubilo košku v každoj gruppe.
PO student-DAT killed cat in each group
'A (different) student killed a cat in each group.'
(Pesetsky 1982a:71)

b. ??Po sobake kusaetsja v každoj kletke.
PO dog-DAT bites in each cage
'A (different) dog bites in each cage.' (Pesetsky 1982a:71)

Pesetsky concludes from these data that distributive *po* phrases must be sister to V^0 at D-structure. In the cases where the *po* phrase corresponds to the subject of an unaccusative or passive verb, the verb is in the default third person neuter singular and the *po* phrase frequently follows the verb instead of preceding it, perhaps indicating that, as with the genitive of negation, these phrases are not grammatical subjects.

3.3.4.2 Non-agreeing Numerals

The numeral system in Russian is extremely complex (Crockett 1976). This section discusses numerals in non-oblique cases which are followed by a genitive NP.[30] These numerals have forms with which the verb agrees, as in (26a), and ones with which it does not, as in (26b).

(26) a. Šest' studentov prišli.
six students-GEN arrived-PL
'Six students arrived.'

b. Šest' studentov prišlo.
six students-GEN arrived-SG,NEUT
'Six students arrived.'

Since agreeing and non-agreeing numeral phrases are morphologically identical, they are difficult to differentiate in any position other than subject position, where the difference is apparent in the presence or absence of verb agreement. Agreeing numeral phrases can appear in all positions which would normally be associated with nominative or accusative case (see Pesetsky 1982a for data; the corresponding agreeing forms for the data in (27) are grammatical). Non-agreeing numeral phrases can appear as subjects of unaccusatives, as in (26b), and of

[30] Numbers in oblique cases act like adjectives and are followed by NPs in the same oblique case. Neidle 1988 suggests that the declension of numbers is unusual in that the nominative and accusative forms are nouns, while the other forms are adjectives.

passives, as in (27a), but not as subjects of transitives, as in (27b), or unergatives, as in (27c).

(27) a. Šest' gorodov bylo vzjato vragom.
six cities-GEN was-SG,NEUT taken-SG,NEUT enemy-INST
'Six cities were taken by the enemy.' (Pesetsky 1982a:78)

b. ??Šest' studentov ubilo košku.
six students-GEN killed-SG,NEUT cat
'Six students killed the cat.' (Pesetsky 1982a:78)

c. ??Šest' sobak kusaetsja v ètoj kletke.
six dogs-GEN bite-3,SG in this cage
'Six dogs in this cage bite.' (Pesetsky 1982a:78)

In each of these examples, the verb shows default agreement, indicative of non-agreeing numeral phrases. Thus, non-agreeing numeral phrases are permitted when the numeral phrase is projected as sister to the V^0, namely in passives and unaccusatives, but is impossible if it is projected into the external argument position, as with subjects of transitives and unergatives.

3.4 Distribution of Negation Markers

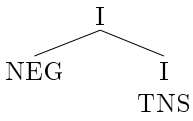
Sentential negation in Russian is associated with I^0 in that negation is never separate from the finite verb; I assume that the negative marker is adjoined to I^0 (King 1994b). This adjunction explains why the negative marker *ne* always directly precedes the inflected verb; if it is found elsewhere in the sentence it is interpreted as constituent negation. In addition, when a negated tensed verb is moved to C^0, the negative marker must move with it (section 3.5.1).

In this account, the negation marker is not in a separate NegP projection nor is it analyzed as an adverb adjoining to IP (or some other projection).[31] Instead, the negative marker forms part of a complex which contains both the information normally associated with I^0 and negation. Following Laka 1990, Piñón 1992 proposes such a position in Hungarian and Romance, labelling it ΣP. The internal structure of the Σ^0/I^0 head is shown in (28).[32]

[31] Rivero 1991, 1993 analyzes Serbo-Croatian, which appears to pattern similarly to Russian, as having a separate projection headed by Negation to which the finite head obligatorily raises. The IP analysis presented here has two main advantages: (1) there is no need to stipulate that the finite head always raises to Neg and (2) there is no additional, unused specifier position. See section 6.5.

[32] Here I have represented this as a head-adjunction structure. Research on morphology and the morphology-syntax interface should ultimately determine how best to represent this structure.

(28)



The verb moves into I^0 in order to receive tense and agreement features. If the sentence is negated, the negative marker *ne* also appears in I^0.

There is no empirical or theoretical difference between calling this node I^0 or Σ^0 as long as the head has a structure like that in (28). I use the I^0 nomenclature due to its familiarty and to emphasize the fact that the position is associated with the inflectional features. However, it must be remembered that this head also contains negation and that in Russian the specifier of this position is not a subject position, but a focus position (section 5.1; Piñón 1992 makes a similar claim for Hungarian, but not Romance).

If the negative marker is associated directly with I^0, then a prediction is made about the scope of negation in coordinated sentences (section 3.1). In sentences composed of an auxiliary followed by coordinated VPs, a single negative marker appearing before the auxiliary is sufficient to negate both VPs, as in (29a). However, in sentences with coordinated I's, in order to negate both conjuncts a negative marker must appear before each finite element, as in (29c).

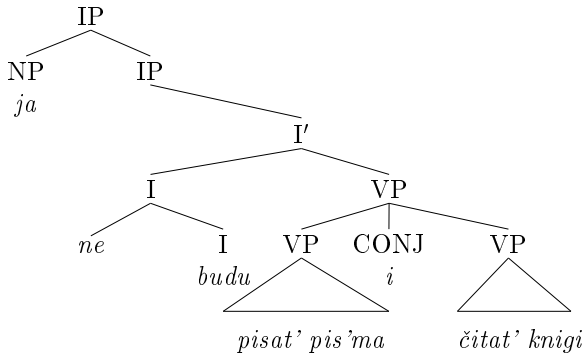
(29) a. Ja ne budu [pisat' pis'ma] i [čitat' knigi].
I not will write-INF letters and read-INF books
'I will not write letters and (will not) read books.'

b. Ja ne [pisala pis'ma] i/a [čitala knigi].
I not wrote letters and read books
'I did not write letters and (*did not) read books.'

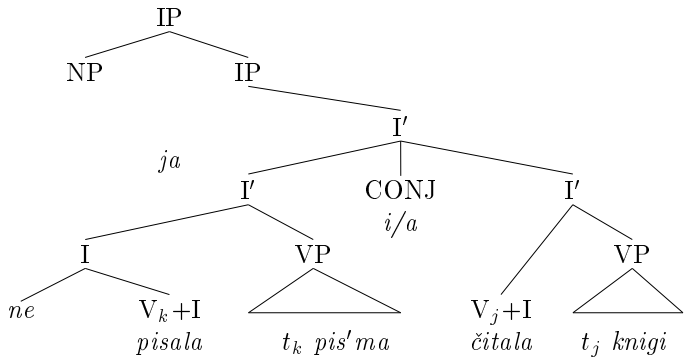
c. Ja ne [pisala pis'ma] i ne [čitala knigi].
I not wrote letters and not read books
'I did not write letters and did not read books.'

The structure for (29a) is shown in (30). There is one negative marker associated with the tensed verb *budu*. This single I^0 has scope over the conjoined VPs and as a result they are both negated. In contrast, (31) shows the structure for (29b) in which there are two conjoined I' constituents. Since only the first I^0 contains a negative marker, only the first VP is negated. In order for the second VP to be in the scope of negation, the I^0 in the second conjunct must contain a negative marker.

(30)



(31)



The distribution of the genitive of negation provides additional evidence of the scope of negation in these sentences. Only objects that are within the scope of negation can appear in the genitive (section 3.3). So, a single negative marker can license the genitive of negation in both clauses if they are infinitives, but only in one if the clauses are finite.

(32) a. Ja ne budu pisat' pisem i
I not will write-INF letters-GEN and
čitat' knig.
read-INF books-GEN
'I will not write letters and read books.'

b. *Ja ne pisala pisem i čitala knig.
I not write letters-GEN and read books-GEN
'I did not write letters and read books.'

c. Ja ne pisala pisem i ne čitala knig.
I not write letters-GEN and not read books-GEN
'I did not write letters nor read books.'

In (32a) there is only one inflected verb, the auxiliary *budu*. Subordi-

nate to this are two coordinated VPs. Material in both of these VPs is within the scope of the negated inflected element. As such, genitive objects can appear in both conjuncts. However, in (32b) there are two inflected verbs, one in each clause. It is impossible for the negation of the first verb to be interpreted on the second conjunct. A genitive object can appear after the first, negated verb, but not after the second verb which is not within the scope of the negative marker. In order to have both conjuncts within the scope of negation and have genitive objects, the negative marker *ne* must appear in front of both inflected verbs, as in (32c).

Thus, evidence from the distribution of the negative marker and the genitive of negation suggests that negation forms a unit with the tensed verb. This projection dominates the VP and as a result allows negation to scope over the material in the VP.

3.5 Head-Movement in yes-no questions

There is one construction in Russian in which the verb undergoes head-movement from I^0 to C^0. This construction is the *li* yes-no question (chapter 6). There are two ways to form matrix yes-no questions: either the intonation of the corresponding declarative can be changed or the clitic *li* can be used (Comrie 1984). When *li* is used in forming a question, the element that is questioned appears in initial position followed by the clitic. If the question is simply about the action, the verb appears in initial position, followed by the clitic, as in (33).

(33) Žil li on v Moskve?
lived Q he in Moscow
'Did he live in Moscow?'

When the verb in question is composed of a tensed auxiliary and an infinitival 'main' verb, it is the auxiliary that appears in initial position before the clitic, not the infinitive. (For some speakers, (34b) is possible with contrastive focus on the infinitive.)

(34) a. Budet li on žit' v Moskve?
will Q he live-INF in Moscow
'Will he live in Moscow?'

b. *Žit' li on budet v Moskve?
live-INF Q he will in Moscow
'Will he live in Moscow?'

In (34a) the auxiliary appears in initial position followed by the question particle *li*. (34) demonstrates that it is the material in I^0, i.e., the tensed verb, that appears in this position. Usually this is the verb

itself, but when there is an auxiliary in I^0, it is fronted and the verb remains in the VP.

3.5.1 Negated Questions

In section 3.4 I argued that the negative marker forms a unit with the inflected verb in I^0. If the formation of *li* yes-no questions involves the head-movement of the material in I^0 to C^0, then the negative marker should also move to C^0 in *li* questions. This is the case, as seen in (35).

(35) a. Ne zastupjatsja li za menja babuška ili tetuška?
not protect Q for me grandmother or aunt
'Wouldn't Grandmother or aunt speak up for me?'
(Yokoyama 1986:240)

b. Oni sprosili, ne videli li my Ivana včera večerom.
they asked not see Q we Ivan yesterday evening
'They asked if we hadn't seen Ivan yesterday evening.'

In (35a) the entire I^0 complex *ne zastupjatsja* has moved into C^0 and is followed by the clitic *li*. (35b) shows the same phenomenon in an embedded question; the I^0 complex *ne videli* is in C^0.[33]

In fact, as would be expected if negation forms a unit with the tensed verb, it is impossible to move the verb into C^0 without the negative marker.

(36) *Oni sprosili, videli li my ne Ivana včera večerom.
they asked see Q we not Ivan yesterday evening
'They asked if we hadn't seen Ivan yesterday evening.'

(36) is ungrammatical if the embedded question is interpreted as having sentential negation. The only possible interpretation is that of constituent negation of *Ivana*.

Thus, the fact that in *li* yes-no questions the negative marker moves with the tensed verb to C^0 supports the claim that the negative marker forms a unit with the tensed verb in I^0 and that the material that moves to C^0 is identical to that which is in I^0.

3.5.2 Predicate Adverbs

Additional support for positing I^0 distinct from V^0 comes from predicate adverbs.[34] The experiencer of these predicate adverbs appears in

[33] *Ne* cliticizes onto the verb to form a single phonological word.

[34] predicate adverbs are so-named because they resemble adverbs in form (some are homophonous with true adverbs). However, they function as the predicate of their clause, and their tense is marked by an auxiliary.

the dative, and they are often followed by an infinitival complement, as in (37). These datives are not grammatical subjects.[35]

(37) a. Emu možno prijti segodnja.
him-DAT can-PRED.ADV come-INF today
'He can come today.'

b. Emu bylo možno prijti včera.
him-DAT was can-PRED.ADV come-INF yesterday
'He could come yesterday.'

In predicate adverb constructions, but not in imperfective futures, the auxiliary frequently follows the predicate adverb, instead of preceding it. With many predicate adverbs, this is the preferred ordering.[36] This difference in ordering cannot be the result of Long Head Movement to C^0 because topics, which are adjoined to IP, precede the predicate adverb and the auxiliary regardless of their relative order (section 5.1.1.2). If Long Head Movement had occurred, the predicate adverb would precede the topics which would in turn precede the auxiliary.[37] For example, in (38) the dative experiencer *emu* is topicalized and precedes both the predicate adverb *nado* and the auxiliary *bylo*.

(38) Mne nado bylo prijti.
me-DAT must-PRED.ADV was come-INF
'I had to come.'

Given that the predicate adverb and auxiliary appear adjacent to one another, following topicalized constituents, I assume that the predicate adverb can head-move to I^0. This is confirmed by the behavior of predicate adverb constructions in *li* yes-no questions (section 6). In these questions either the auxiliary or the predicate adverb can appear in initial position, as in (39). In fact, both the predicate adverb and the auxiliary can appear in initial position in either order, as in (40).

(39) a. Bylo li Ivanu trudno napisat'
was Q Ivan-DAT difficult-PRED.ADV write-INF
ètu stat'ju?
this article
'Was it difficult for Ivan to write this article?'

[35] These datives do not trigger verb agreement or occur as controllees. However, as logical subjects they can antecede reflexives and control adverbial participles.

[36] Why one ordering should be preferred for some predicates and another for others remains to be explored.

[37] This would not exclude Long Head Movement to some other head, as long as the landing site were lower in the tree than the topics. For now, I assume that the predicate adverb moves to I^0.

b. Trudno li Ivanu bylo napisat'
difficult-PRED.ADV Q Ivan-DAT was write-INF
ètu stat'ju?
this article
'Was it difficult for Ivan to write this article?'

(40) a. Legko li bylo Inne napisat' ètu stat'ju?
easy-PRED.ADV Q was Inna-DAT write-INF this article
'Was it easy for Inna to write this article?'

b. Bylo li legko Inne napisat' ètu stat'ju?
was Q easy-PRED.ADV Inna-DAT write-INF this article
'Was it easy for Inna to write this article?'

The fact that both the predicate adverb and the auxiliary are in initial position is not obvious because the second position clitic *li* appears after the first phonological word of the fronted phrase (section 6.1). Since the predicate adverb plus auxiliary complex forms two phonological words, the clitic appears after the first of these. However, both appear before the topicalized dative experiencer, i.e., before *Ivanu* in (39) and before *Inne* in (40).

This behavior suggests that in predicate adverb constructions, both the auxiliary and the predicate adverb can be in I^0 at the same time.[38] This can be accomplished if the predicate adverb undergoes head movement from V^0 to I^0.[39] Once the predicate adverb has moved to I^0, the predicate adverb auxiliary complex functions as any other finite verb in Russian.[40] So, a sentence like (38) has a structure like (41).

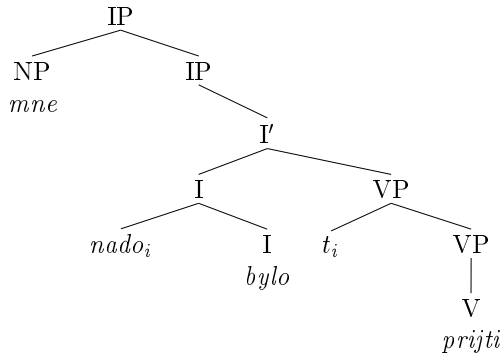
[38](42) provides the structure of a clause in which the predicate adverb has adjoined to I^0. In clauses in which the auxliary precedes the predicate adverb, such as in (37b), the predicate adverb will remain in V^0.

[39]Rivero 1991, 1993 suggests that certain predicates in Bulgarian can undergo head movement resulting in their appearing before the finite element, e.g., the participle often precedes the auxiliary. These involve Long Head Movement of V^0 to C^0 and as a result they appear only in matrix clauses. The Russian head movement differs in that the predicate adverb is moving only to the next higher maximal projection, i.e., I^0, and as a result can appear in any construction in which there is a I^0 projection, including embedded clauses.

[40]Constructions with non-finite uses of predicate adverbs are rare. The lack of a thematic grammatical subject may be the reason for this: it is hard to find semantically plausible control sentences. A possible example is shown in (i), although it is not clear that *stydno* has the same status as the predicate adverbs discussed in this section, e.g., it might be a true aderbial. (The matrix predicate in (i) *dolžno* can either agree with a third singular neuter subject or be a default form, depending on the stress (Chvany 1975; see also Babby 1989 on Ukrainian); the form in (i) is the default form since there is no grammatical subject.)

(i) Tebe dolžno byt' stydno.
you-DAT should-PRED.ADV be-INF ashamed-PRED.ADV
'You should be ashamed (of yourself).'

(41)



One issue concerning the structure in (41) is the exact structure of the material in I^0. Particularly, do the predicate adverb and auxiliary form a single unit, i.e., do they form a complex structure or a simple adjunction structure? The structure shown in (41) suggests that they are both found under I^0, in a head-adjunction structure. When a yes-no question is formed, as in (40), either or both of the elements in I^0 may appear in C^0. Another question which remains unanswered is why predicate adverbs can undergo head movement. This movement is optional: some predicate adverbs usually precede the auxiliary, while others usually follow it, but both orders can be found with any given predicate adverb. predicate adverbs differ from imperfective futures in this way since in imperfective futures the infinitive does not undergo head movement to I^0 where it would precede the auxiliary.

3.6 Traditional Subject-Object Asymmetries

Traditionally, subject-object asymmetries have been used as evidence for the existence of a VP. The idea is that if the object is governed and assigned a theta-role directly by the verb, while the subject is governed and assigned its theta-role indirectly, then subjects and objects should behave differently. In addition, the subject c-commands the object, but not *vice versa*. However, if subjects and objects are in identical structural positions, then the object does not have a special relation to the verb, and subject and objects will not exhibit such differences (see Kiss 1987b on Hungarian). The following English sentences demonstrate common subject-object asymmetries:[41]

[41]There is an additional complicating factor in (42) in that the R-expression precedes the pronominal in (42a) but follows in it (42b).

Some people find examples like (45b) grammatical. The judgments are clearer with embedded sentences. Also, superiority effects disappear when *which* phrases are used. Pesetsky 1987 argues that this is a result of D(iscourse)-linking of *which* phrases.

Pronominal Coreference:

(42) a. The woman I introduced to John$_i$ eventually left him$_i$.
b. *He$_i$ eventually left the woman I introduced to John$_i$.

ECP Effects:

(43) a. *Who$_i$ did you say that t_i went in first?
b. Who$_i$ did you say that they called in t_i first?

Weak Crossover:

(44) a. Who$_i$ t_i loves his$_i$ mother?
b. *Who$_i$ does his$_i$ mother love t_i?

Superiority Effects:

(45) a. Who$_i$ t_i gave what to Mary?
b. *What$_i$ did who give t_i to Mary?

Russian does not demonstrate the range of subject-object asymmetries found in English. This is partly because long distance extraction in Russian is ungrammatical for independent reasons (Comrie 1980). It is impossible to move any element out of its minimal finite clause in Russian.[42] In addition, the presence of reflexive possessive pronouns in Russian makes certain of the standard subject-object asymmetry tests inapplicable. However, the asymmetries that *are* found in Russian are discussed in this section.

3.6.1 Pronominal Coreference

Coreference possibilities between a lexical NP and a pronoun are different between subjects and objects. This is attributed to the Binding Theory which is stated in (46) (Chomsky 1981).

(46) Binding Theory
Principle A: An anaphor is bound in its Governing Category.
Principle B: A pronominal is free in its Governing Category.
Principle C: An R-expression is free.

Binding is defined as in (47).

(47) α binds β iff
(a) α c-commands β and
(b) α and β are coindexed.

There are many different definitions of Governing Category; (48) is relatively simple.

(48) The Governing Category for β is the smallest NP or S containing β and a governor of β. (Sells 1985:69)

[42]The examples of long-distance scrambling in section 3.2 were from Colloquial, not Contemporary Standard, Russian.

What is of immediate interest to us is Principle C of the Binding Theory which states that R-expressions must be free. R-expressions include lexical NPs and wh-traces.[43] If subjects c-command objects, then we expect a difference in behavior between them relative to Principle C.

The sentences in (49) and (50) involve coreference possibilities between an NP and a pronoun. The pronoun is an argument of the main clause while the NP is an argument of a relative clause, i.e., it is inside an NP argument of the main clause.

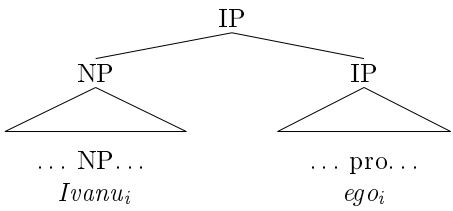
(49) Ženščina, kotoruju ja predstavila Ivanu$_i$,
woman-NOM who-ACC I introduced Ivan-DAT
pokinula ego$_i$.
deserted him-ACC
'The woman I introduced to Ivan$_i$ deserted him$_i$.'

(50) *On$_i$ pokinul ženščinu, kotoruju ja
he-NOM deserted woman-ACC who-ACC I
predstavila Ivanu$_i$.
introduced Ivan-DAT
'He$_i$ deserted the woman I introduced to Ivan$_i$.'

In (49) the pronoun *ego* is the object of the main clause. The sentence is grammatical under the reading where *Ivan* and the pronoun are coreferent. However, in (50) the pronoun *on* is the subject of the main clause and the coreferent reading is impossible, regardless of word order.

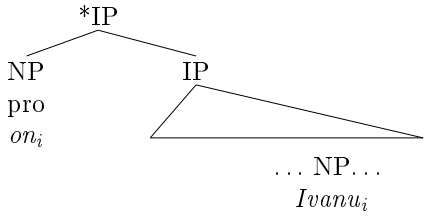
The difference between these two structures is that in (49) the NP *Ivan* is embedded within the subject NP while the pronoun is the object of the main verb, as in (51). In (50) the pronoun is the subject of the main verb while the NP *Ivan* is embedded within the object NP, as in (52).

(51)



[43] At first glance it might appear that wh-traces are bound. However, the Binding Theory is interpreted as applying to A-binding, not A′-binding. Since wh-words are in A′ positions, there is no Binding Theory violation.

(52)



If pronouns cannot c-command their antecedents by Condition C, then the behavior of (49) and (50) can be explained by the above structures. By positing a structural difference between subject and object, the pronoun in (51) does not c-command its antecedent, while the pronoun in (52) does.

Can this generalization cannot be stated simply in terms of precedence? (49) degrades when the ordering is changed, as in (53).[44] However, (50) does not improve when the R-expression precedes the pronoun, as in (54).[45]

(53) a. *Pokinula ego$_i$ ženščina, kotoruju ja
deserted him-ACC woman-NOM who-ACC I
predstavila Ivanu$_i$.
introduced Ivan-DAT
'The woman I introduced to Ivan$_i$ deserted him$_i$.'

b. *Ego$_i$ pokinula ženščina, kotoruju ja
him-ACC deserted woman-NOM who-ACC I

[44]There is a contrast between examples like (54) and those involving the A′-movement of wh-phrases into SpecCP. Consider the examples in (i).

(i.a) *Kakuju fotografiju s Ivanom$_i$ on$_i$ xočet bol'še vsego?
which picture of Ivan he want most all
'Which picture of Ivan$_i$ does he$_i$ want most of all?'

(i.b) *Kakuju stat'ju ob Ivane$_i$ on$_i$ ljubit bol'še vsego?
which article about Ivan he like most all
'Which article about Ivan$_i$ does he$_i$ like most of all?'

At S-structure *Ivan* precedes the pronoun and neither c-commands not is c-commanded by it. However, the examples are ungrammatical under the reading where the R-expression *Ivan* and the subject pronoun *on* are coreferent. This ungrammaticality can be explained if the subject c-commands the object which contains the R-expression *Ivan* in these examples. One possibility is that this c-command relation is visible after reconstruction at LF and as a result the intended coreference is impossible.

[45]Mohanan 1982:525 argues that coreference in Malayalam must be stated in terms of precedence. Note that this precedence relation is not applicable to reflexives in Malayalam; the only requirement on reflexives is that their antecedent be a grammatical subject (Mohanan 1982:566-567).

If NP$_1$ precedes NP$_2$, and NP$_1$ is a pronoun while NP$_2$ is not, then NP$_1$ and NP$_2$ are non-coreferential.

Russian differs from Malayalam. In Russian there must be an additional restriction on when coreference is allowed.

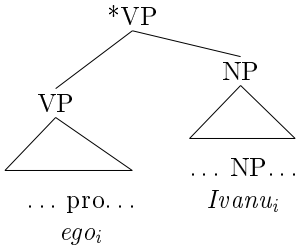
predstavila Ivanu$_i$.
introduced Ivan-DAT
'The woman I introduced to Ivan$_i$ deserted him$_i$.'

(54) Ženščinu, kotoruju ja predstavila Ivanu$_i$,
woman-ACC who-ACC I introduced Ivan-DAT
on$_i$ pokinul.
he-NOM deserted
'He$_i$ deserted the woman I introduced to Ivan$_i$.'

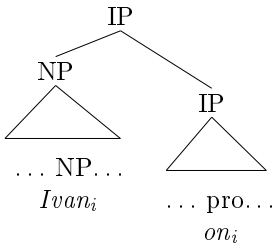
(53a) and (53b) are equivalent to (49) except that the grammatical subject *ženščina*, including the relative clause, has been focused and thus appears sentence finally (section 5.3). The sentence is ungrammatical whether the object precedes the verb, as in (53b), or follows it, as in (53a). (54) is equivalent to (50) with the direct object *ženščinu*, including the relative clause, topicalized into initial position (section 5.1.1).

The structures for (53) and (54) are shown in (55) and (56) respectively. In (55) there is a problem with precedence since the pronoun *ego* precedes its antecedent *Ivanu*; however, there is no Condition C violation since the pronoun does not c-command its antecedent. In (56) there is no problem with linear precedence since *Ivan* precedes the pronoun, and there is no trouble with c-command since the pronoun does not c-command its antecedent at S-structure.

(55)



(56)



The pronominal coreference facts of Russian can then be captured as follows:[46]

(57) A pronoun can neither precede nor c-command its antecedent (an R expression) at S-structure.

Pronominal coreference does not provide uncontroversial proof that there is a subject-object asymmetry in Russian in that the subject c-commands the object, but not *vice versa*.[47] However, positing such an asymmetry allows a relatively simple explanation for the coreference facts and argues for subject-object asymmetries in a theory which defines subjects and objects structurally.

3.6.2 Extraction

Extraction out of finite clauses in Russian is generally unacceptable. However, when extraction does occur, it shows subject-object asymmetries (Pesetsky 1982b).[48] Although these extraction sentences are uncommon and usually considered ungrammatical, there is a definite difference in grammaticality between (58a) and (58b).

(58) a. paren', kotorogo ja xotel, čtoby Maša ubila...
guy who-ACC I want that Masha-NOM kill
'The guy, who$_i$ I wanted Masha to kill t_i,...'
(Pesetsky 1982b:299)

b. *paren', kotoryj ja xotel, čtoby ubil Mašu...
guy who-NOM I want that kill Masha-ACC
'The guy, who$_i$ I wanted t_i to kill Masha,...'
(Pesetsky 1982b:299)

In (58a) the object of the verb *ubila* has been extracted and the sentence is grammatical. In (58b) the subject of the verb *ubil* has been extracted and the sentence is ungrammatical. The word order within the relative clause does not affect grammaticality. The fact that the verb governs its object but not its subject explains the difference in grammaticality. The trace of the object is properly governed by the verb, while that of the subject is not. Proper government is roughly defined as in (59).

[46] One of the differences between the behavior of Russian and English pronominal coreference is that linear precedence plays a major role in the interpretation of Russian pronominal reference.

[47] Some syntactic theories provide an alternative explanation for these facts. The condition on linear precedence can be stated under all theories. However, it is possible to capture restrictions involving c-command without reference to the phrase structure, if the phrase structure is not used to encode grammatical functions. See Bresnan 1982 for more discussion of these matters in LFG.

[48] Pesetsky's analysis of these facts is derived from the Nominative Island Condition which states that a nominative anaphor must be bound inside its clause; this is essentially a precursor of the Empty Category Principle.

(59) α properly governs β iff
(a) α governs β and α is lexical (N, V, A, or P) or
(b) α locally A′-binds β

Since the object trace in (58a) is governed by the verb, the movement of the object out of the clause will not be an Empty Category Principle (ECP) violation because the trace is properly governed. However, the subject trace in (58b) is not governed by a lexical head, nor is it locally A′-bound. As such, the subject trace is not properly governed and the resulting ECP violation renders the sentence ungrammatical. If the subject and object were in identical structural positions, e.g., both governed by the verb within the VP, then this asymmetry in extraction would not be expected. In languages where there is no structural asymmetry between the subject and object, either both the subject and the object should be able to be extracted or neither should. However, if the subject is in SpecVP and the object is sister to V^0, the difference in behavior of subjects and objects in extraction situations is predicted.

3.6.3 Weak Crossover

Weak crossover is traditionally described as a reflex of the Leftness Condition which stipulates that a pronoun cannot be coindexed with a variable to its right (however, see Bresnan 1994 on the interacting factors which account for weak crossover). This condition rules out structures like (60), while permitting those in (61).

(60) *Who does his mother love?
who$_i$ does his$_i$ mother love t_i

(61) Who loves his mother?
who$_i$ t_i loves his$_i$ mother

In (60) the pronoun *his* is coindexed with the object wh-trace to its right, which violates the Leftness Condition. However, in (61) the wh-trace is to the left of the pronoun and there is no violation of the Leftness Condition.

Unfortunately, similar contrasts cannot be found in Russian. Russian has reflexive and non-reflexive possessive pronouns, a distinction not found in English (see Rappaport 1986 and Klenin 1974 on Russian reflexives). The Russian question corresponding to (61) contains a reflexive pronoun, as in (62a); the non-reflexive counterpart in (62b) cannot be coreferential with the subject.[49]

[49]Loren Billings (p.c.) notes that Yokoyama claims that non-reflexive possessive pronouns can occassionally refer to the subject in order to reflect the speaker's point of view. However, this use does not seem to occur in wh-clauses.

(62) a. Kto ljubit svoju sobaku?
who-NOM loves self's dog
'Who$_i$ loves his$_{i,*j}$ dog?'

b. Kto ljubit ego sobaku?
who-NOM loves his dog
'Who$_i$ loves his$_{j,*i}$ dog?'

Next consider the Russian counterparts to (60) in which the object is questioned. These are ungrammatical with reflexive pronouns, as in (63), due to a violation of Binding Principle A: the reflexive must be bound in its governing category, and it is not (section 3.6.1). This can be seen in that the declarative counterpart of (63), shown in (64), is also ungrammatical.[50]

(63) *Kogo ljubit svoja sobaka?
who-ACC loves self's dog
'Who$_i$ does his$_i$ dog love?'

(64) *Svoja sobaka ljubit Ivana.
self's dog loves Ivan
'His$_i$ dog loves Ivan$_i$.'

Next consider these questions when formed with non-reflexive pronouns, as in (65). Although the preferred reading for (65) is not to have coindexation, such indexation is possible.

(65) Kogo ljubit ego sobaka?
who-ACC loves his dog
'Who$_i$ loves his$_{i,j}$ dog?'

It is interesting to see how the declarative counterpart of this question behaves. (66a) is ungrammatical because the pronoun precedes its antecedent (section 3.6.1). However, when the object precedes the subject, as in (66b), the clause is grammatical because the pronoun neither precedes nor c-commands its antecedent.

(66) a. Ego sobaka ljubit Ivana.
his dog loves Ivan
'His$_i$ dog loves Ivan$_{j,*i}$.'

b. Ivana ljubit ego sobaka, (i bol'še nikto).
Ivan loves his dog and more no one
'His$_{i,j}$ dog loves Ivan$_i$ (and no one else does).'

Thus, there is an interesting problem with weak crossover in Russian. (62b), the Russian equivalent of the grammatical (61), is ungrammatical due to the antisubject orientation of *ego* 'his'. In contrast, (65),

[50] These are also ungrammatical with *drug druga* 'each other' because Russian does not use this expression as a possessive.

the Russian equivalent of the ungrammatical (60), is grammatical. If the behavior of pronouns is taken into account, ignoring weak crossover, then the grammaticality of (65) is predicted since the pronoun neither precedes nor c-commands its antecedent, the wh-phrase *kogo*. As such, the data suggest that the motivations behind weak crossover need to be reexamined along the lines of Bresnan 1994 who proposes that weak crossover effects are the result of precedence and syntactic rank constraints. For Russian, the constraints on pronominals and reflexives is sufficient to account for the 'weak crossover' facts.

3.6.4 Superiority Effects

A fourth test of subject-object asymmetries involves ordering in multiple questions, i.e., superiority effects. In languages with superiority effects, the subject wh-phrase must precede that of the object. Superficially, Russian wh-questions appear to exhibit such asymmetries, as in (67). Unlike English, all the wh-words appear at the beginning of the clause in Russian.

(67) a. Kto kogo udaril?
who-NOM who-ACC hit
'Who hit who?'

b. *Kogo kto udaril?
who-ACC who-NOM hit
'Who did who hit?'

In (67a) the subject wh-word *kto* precedes that of the object, *kogo*, and the question is well-formed. In (67b) the object wh-word precedes that of the subject and the question is ungrammatical.

However, unlike in English, in Russian the wh-words can appear in the opposite order given sufficient context (Loren Billings, Steve Franks p.c.). That is, (67b) is a grammatical question, although under most circumstances the ordering in (67a) is preferred. This suggests that Russian does not demonstrate superiority effects, but rather orders its wh-phrases according to other principles.[51]

[51]Bulgarian is a Slavic language which exhibits complex superiority effects. The government of the traces left by wh-movement can explain the difference in grammaticality in (i). The subject trace in (i.b) cannot be properly governed at LF due to the indexing of the complementizers, as in (ii) and (iii).

(i.a.) Koj kakvo pravi? b. *Kakvo koj pravi?
who what does what who does
'Who does what?' 'Who does what?'
(Rudin 1988:481) (Rudin 1988:482)

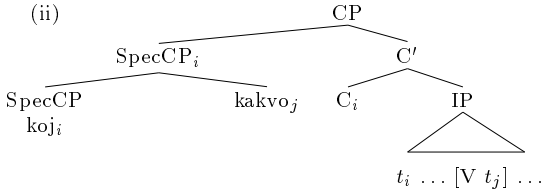
Rudin 1988 claims that the first wh-word to move into SpecCP results in C^0 indexing, as proposed by Aoun, Hornstein, and Sportiche 1980. Other wh-words

Rudin 1988, 1989 discusses multiple wh-fronting in a number of Slavic languages. Some demonstrate superiority effects, while others allow any order among the wh-words. She argues that in the languages with superiority effects, such as Bulgarian, all of the wh-words are adjoined in SpecCP.[52] In the other languages, such as Czech and Polish, although all of the wh-phrases are in initial position, only one appears in SpecCP; the rest are adjoined to IP. In these languages, the wh-trace can be governed from either SpecCP or the adjoined positions, and thus they show no superiority effects.

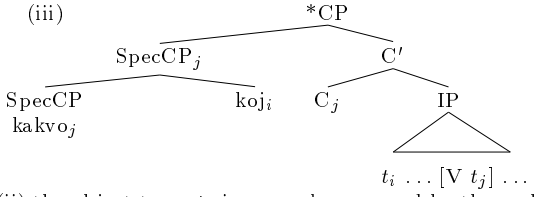
Rudin's arguments for this structural difference in placement of the wh-words involve: multiple wh-extraction, wh-islands, clitic and parenthetical placement. When all of the wh-words are in SpecCP, they act as a unit. Rudin 1989 discusses the Russian data and concludes that Russian patterns with the languages in which only one of the wh-words appears in SpecCP, while the rest are adjoined to IP (or whichever projection appears immediately below CP; section 5.2.3). Her evidence

right-adjoin to the material in SpecCP and leave no index in C^0. The resulting structures for the questions in (i.a) and (i.b) are those in (ii) and (iii) respectively.

(ii)



(iii)



In (ii) the object trace t_j is properly governed by the verb, while the subject trace t_i is licensed by the coindexed C^0 which head-governs it. Thus, when the subject wh-word is first in SpecCP, both traces are properly governed: the object by the verb and the subject by the coindexed C^0. However, this is not the case when the object wh-word is moved into SpecCP first, resulting in C^0 being coindexed with it, as in (iii). The object trace t_j is properly governed by the verb in this configuration. However, the subject trace is not properly governed since, as usual, the verb does not properly govern it and there is no coindexed C^0 to do so. As a result, the structure is ungrammatical.

[52] Rudin points out that the ordering of wh-words in Bulgarian is quite fixed. Not only must subjects appear first (fn. 51), as predicted by her account, but the other wh-words appear in a specific order. Billings and Rudin 1994, in an Optimality Theory account, argue that the ordering restrictions on Bulgarian wh-phrases are not the result of the ECP, but of the interaction of conflicting constraints.

is discussed below. Although the evidence is not as straightforward as for some of the other Slavic languages, I conclude that Russian multiple wh-questions do not resemble their Bulgarian counterparts, i.e., the Russian wh-phrases do not all adjoin to SpecCP.

Although Russian does not have second position clitics in the way Polish and Czech do, the subjunctive marker *by*, which usually appears after the verb, can appear in second position, as in (68a). In sentences with multiple wh-phrases, this marker can appear after the first of the wh-phrases, as in (68b), indicating that the wh-phrases do not form a constituent, as they would if they were all in SpecCP.

(68) a. Ja by s udovol'stviem pošel v teatr.
I BY gladly go to theater
'I would gladly go to the theater.' (Rudin 1989:12)

b. Kuda by kto pošel?
where BY who go
'Who would go where?' (Rudin 1989:12)

c. *Kuda kto by pošel?
where who BY go
'Who would go where?'

Parentheticals show the same pattern in that they can break up the fronted wh-phrases.[53]

(69) Tak, kto, vy mne skazali, kogo udaril?
so who you told me whom hit
'Who did you tell me hit whom?' (Rudin 1989:12)

[53] My informants markedly prefer that parentheticals not split up the fronted wh-phrases. Adverbials, e.g., *pervym*, and parentheticals, e.g., *po tvoemu*, rarely separate the wh-words in (iii) and (vi), although their ordering is otherwise quite free.

i. Kto kogo pervym udaril?
who-NOM who-ACC first hit
'Who hit who first?'

ii. Kto kogo udaril pervym?
who-NOM who-ACC hit first

iii. *Kto pervym kogo udaril?
who-NOM first who-ACC hit

iv. Kto kogo, po tvoemu, udaril?
who-NOM who-ACC in your opinion hit
'In your opinion, who hit who?'

v. Kto kogo udaril, po tvoemu?
who-NOM who-ACC hit in your opinion

vi. *Kto, po tvoemu, kogo udaril?
who-NOM in your opinion who-ACC hit

Russian does not allow movement out of finite clauses, thus making it difficult to test wh-islands and wh-extraction. Rudin 1989 observes that when extracting wh-phrases out of infinitival clauses, only one wh-word can be extracted, as in (70). However, as seen in (71), multiple wh-extraction out of the infinitival clause in (70b) is marginally possible with the wh-words in the opposite order.

(70) a. Komu ženščina xotela napisat'?
who-DAT woman want write-INF
'Who did the woman want to write to?' (Rudin 1989:13)

b. *Čto komu ženščina xotela napisat'?
what who-DAT woman want write-INF
'What did the woman want to write to whom?'
(Rudin 1989:12)

(71) ?Čto komu ženščina xotela napisat'?
what who-DAT woman want write-INF
'What did the woman want to write to whom?'

Finally, extraction from within a clause with a fronted wh-word is ungrammatical, as in (72). (72) is independently ungrammatical because it involves movement out of a finite clause.

(72) *Kto ty ne znaeš, gde živet?
who you not know where live
'Who don't you know where lives?' (Rudin 1989:14)

So, Russian appears to pattern like Polish and Czech which Rudin 1988 argues have only one wh-word in SpecCP while the rest are adjoined to IP.

What is interesting about this, for our purposes, is that in languages like Czech, Polish, and Russian, which normally do not show superiority effects, there is a strong preference for nominatives to precede accusatives. This preference explains the variation in judgements of (67b) in that some speakers may more freely allow contexts in which the accusative wh-phrase precedes the nominative one. Interestingly, it turns out that in Russian superiority effects appear only when a nominative and an accusative wh-phrase appear simultaneously. With wh-phrases marked with other cases, as in (73), and with non-case marked wh-phrases, e.g., adverbials, as in (74), there are no superiority effects, unlike in Bulgarian which has a strict ordering.

(73) a. Čto komu ty napisala?
what who-DAT you write
'What did you write to whom?'

b. Komu čto ty napisala?
who-DAT what you write
'Who did you write what?'

(74) a. Kogda kogo ty uvidela?
when who-ACC you saw
'When did you see who?'

b. Kogo kogda ty uvidela?
who-ACC when you saw
'Who did you see when?'

In addition, even in English it is not clear why various restrictions apply to certain wh-phrases, but not others, although there has been much discussion of these data (Pesetsky 1987). Thus, superiority effects in Russian cannot be used to argue for subject-object asymmetries in Russian, although they do not argue against such a proposal.

The data discussed in this section demonstrate that Russian has many traditional subject-object asymmetries. In order to capture these asymmetries, there must be some difference between how subjects and objects are represented. In a Government-Binding approach, this entails projecting the subject into a position in which it c-commands the object. In Lexical-Functional Grammar the notions subject and object are not defined configurationally and thus the difference between the behavior of subjects and objects in Russian need not be represented in the phrase structure (chapter 9).

This chapter discussed the possible evidence for configurationality in Russian. Two major conclusions were reached. The first is that there is a difference in the syntactic behavior of subjects and objects. This difference can be explained if subjects c-command their objects, but not the reverse. The second conclusion is that there is a structural difference between finite and nonfinite verbs. Related to this, negation is intimately associated with finite verbs, acting as a unit with the verb and taking scope over its arguments. Subjects are in SpecVP, while objects are complements to V^0, which results in the desired c-command relations. In addition, finite verbs raise out of the VP to the next higher projection, while infinitives remain in V^0; this accounts for the observed difference in syntactic behavior between infinitives and finite verbs. Note that the asymmetical c-command relations between subjects and objects are independent from the movement of finite verbs to I^0, i.e., the same c-command relations could hold whether the verb moved or not and whether the subject moved to SpecIP or not. The data presented in chapter 4 showing that word order reflects discourse

functions are used in chapter 5 to argue for particular syntactic positions being associated with discourse functions. Constituents can move out of their VP positions into these discourse function positions, resulting in the varied word orders, but maintaining a basic configurationality.

4

Linearly Ordered Topic and Focus

The linear order of the constituents in a Russian sentence does not encode grammatical information. Given a simple transitive sentence, the constituents can appear in any order without altering the interpretation as to the grammatical function of each noun phrase, e.g., which noun phrase is the subject and which the object. This information can usually be deduced from the case marking: subjects are in the nominative case, objects in the accusative, etc. In fact, such 'reorderings' are common, making it difficult to determine the unmarked word order of Russian.

So one asks: What are all these different orders for? Although the grammatical relations of the arguments to the predicate are identical regardless of the word order, can these different orders really have the same meanings? The answer to the second question is no. The answer to the first is more involved since the different meanings must be untangled; their untangling provides the answer to the first question. Essentially, the different word orders, in combination with intonation, encode discourse functions. These differences in discourse function affect the semantics and pragmatics of the sentence and restrict the environments in which a clause with a particular word order can felicitously appear. This is not a new observation. The fact that word order reflects discourse functions is the subject of countless books and articles on Russian, not to mention other languages with 'free' word order. The constituents move to certain positions because of their discourse function interpretation. This can be thought of as licensing: in order to receive a particular discourse function interpretation, an argument must move into a particular position, and any argument which moves into such a position receives the associated discourse function interpretation. These different structures result in the different linear orders found and in the interaction of intonation with the different orders. For

Russian, they give the general order of topic–discourse-neutral–focus, although the interaction of stress and contrastive focus with the phrase structure allows for certain preverbal foci.

The goal is to reconcile the more canonical D-structures with the S-structures of these surface orders, to determine what their structure is and how it is used to encode discourse information (chapter 5). Grammatical functions are encoded by the phrase structure in that constituents appearing in certain positions have certain grammatical functions. Constituents can move out of these positions and receive different discourse function interpretations depending on the position to which they move: for example, adjunction to IP results in the constituent being interpreted as a topic. In essence, D-structure reflects the argument structure of the predicate, while S-structure encodes grammatical functions, and the final orderings reflect structurally encoded discourse functions; this final ordering is exclusive of PF level reorderings.[1] This strong correlation between phrase structure and discourse function interpretation provides a basis for future research into the more complicated details of Russian word order and into the syntactic behavior of other phenomenon, such as case marking and impersonal expressions.

This chapter provides the background for chapter 5 in which the phrase structure of these different word orders is explored. The first part of this chapter outlines definitions of topic and focus found in the literature. These notions have been used in numerous ways, especially topic, and the result has necessarily been some confusion. The focus is on those ideas and definitions which prove useful in the description of Russian. Section 4.3 examines topic and focus in Russian and delineates the data described here. The divisions of the data are based primarily on traditional works, which are concerned not with the phrase structure of sentences but with the surface word orders and their interaction with the discourse. In particular, the difference between emotive and non-emotive sentences is examined, as this distinction proves vital in determining the association of word order, intonation, and topic–focus interpretation in Russian sentences.

4.1 Notions of Topic

'Topic' seems to have as many meanings as there are linguists who write about it, although it is usually loosely defined as what the sentence is about (Prince 1981). This has led to confusion and inconsistencies in

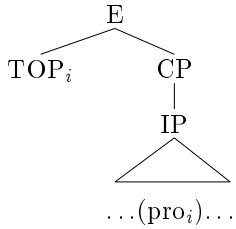
[1] There are relatively few PF reorderings under this account since most movement is motivated by the syntax and discourse functions. An example of such a reordering would be the placement of clitics, e.g., of *li* in yes-no questions (chapter 6).

the literature. This section describes the previous definitions of topic which are most relevant for describing Russian word order.

4.1.1 External vs. Internal Topics

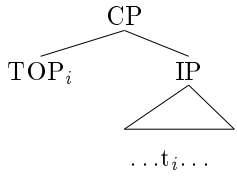
Aissen 1992 makes a distinction between external and internal topics in Mayan (section ??). This distinction has syntactic and semantic reflexes. External topics are essentially dislocated structures. Although they may be coreferent with an argument of the predicate, they are base generated in topic position and are not arguments of the predicate. Aissen labels the projection under which external topics are found E(xpression), using a term from Banfield 1973. This projection is a root phenomenon, and as a result external topics are only found in root clauses. Semantically, external topics identify a referent in the discourse about which an assertion will be made. This entity remains the topic until a new topic is introduced, i.e., external topics are 'new' or 'shifted' topics (Aissen 1992:51; section 4.1.3). The structure she proposes for external topics is shown in (1).

(1)



In contrast to external topics, internal topics are arguments of the verb which move into topic position. They bind a trace in the clause and are not associated with resumptive pronouns. In addition, internal topics can appear in embedded clauses since they are not associated with the node E. Internal topics also differ from external topics semantically because they need not introduce a new topic, although they can do so. Instead, they can refer to the current or continuing topic (Aissen 1992:77). The structure Aissen proposes for internal topics in Tz'utujil is shown in (2). Cross-linguistically, the landing site of the internal topic may vary, but the basic configuration is identical to that in (2).

(2)



Rudin 1985, *to appear* makes a similar distinction for Bulgarian in which she proposes that external topics are located under an E node, while internal topics are adjoined to S′ or CP (section 5.2.3). As with Aissen's analysis of Mayan, external topics, which Rudin terms Left-dislocated, are not arguments of the verb, although they are often coreferent with pronominal arguments; on the other hand, internal topics are arguments of the verb and move into topic position. As discussed in chapter 5, Russian has both internal and external topics.

4.1.2 Subject-of-Predication

Subject-of-predication is what Kiss *to appear*, 1993 uses to describe the role of topics in Hungarian. The subject-of-predication "functions semantically as a notional subject, foregrounding a referent about which the VP will predicate something" (Kiss *to appear*:4). The subject-of-predication can bear any grammatical relation to the predicate, i.e., it is not restricted to subjects.[2] In fact, it is possible to have more than one subject-of-predication in a single clause, although this is relatively rare. Two examples of subjects-of-predication are shown in (3); these involve the same basic sentence with different choices of subject-of-predication, i.e., the object is the subject-of-predication in (3a) and the subject in (3b).

(3) a. [Évát] János várta a mozi előtt.
Eve-ACC John waited the cinema in-front-of
'Eve-S.PRED was waited for in front of the cinema by John.'

b. [János] Évát várta a mozi előtt.
'John-S.PRED waited for Eve in front of the cinema.'

This is a more restricted notion than the one discussed in section 4.1.3, used in analyzing Russian word order. However, the subject-of-predication helps to explain why Russian is often thought of as an SVO language. Most clauses have a subject-of-predication, which like all topics in Russian is preverbal, and these subjects-of-predication are often grammatical subjects (section 5.4).

[2]Kiss 1993:4 defines primary predication as a syntactic notion (she claims that in Hungarian the VP bears a primary predication relation to the topicalized constituents). This predication structure is the result of NP-movement, so that the trace in the VP is a non-variable trace.

A VP bears a primary predication relation to an XP iff
i. XP c-commands VP,
ii. XP binds an empty argument position in VP, and
iii. (disregarding the functional projections spelling out the L-features of the V), XP and VP are not separated by a maximal projection excluding XP.

4.1.3 Shared Current Concern

Yokoyama 1986 provides a schema for dividing information into classes relevant for discourse. Although she does not divide the information directly into topic and focus, she mentions that her ontology could be used to more accurately define the notions theme/rheme and topic/focus. The basic sets used in her divisions are shown in (4). The information in these sets can overlap, yielding new sets which are relevant to the organization of sentences and their interpretation.

(4) A A's knowledge set
B B's knowledge set
C_a A's current matter of concern
C_b B's current matter of concern

The set of primary importance here is $C_a \cap C_b$ which is the shared matter of current concern. This set is effectively the set of topics of an utterance.[3] The placement of items in $C_a \cap C_b$ is dependent on the speaker's assessment of what constitutes material of interest to both the speaker and the addressee. For example, a speaker may place items with which she is empathetic in this set, with the presumption that the addressee will likewise take an interest in them.

In discourse-initial situations, Yokoyama 1986:32-33 states that this set includes certain deictic material (I, you, here, now) and the minimal proposition-like knowledge that the interlocutors share prior to the first utterance. So, even in utterance-initial contexts, in which no topic of discussion has necessarily been agreed upon, deictic elements can appear in initial position, as in (5) and (6).

(5) [Vam] prislali Ivanovy priglašenie.
you sent Ivanovs invitation
'The Ivanovs have sent you an invitation.'
(Yokoyama 1986:277)

(6) S kem [vy] ezdili v London?
with whom you went to London
'Who did you go to London with?' (Yokoyama 1986:299)

In (5) the indirect object *vam* appears before the verb instead of after it. If *vam* can be in $C_a \cap C_b$ due to its deictic nature, this is to be expected since topic material appears before the verb in Russian. In (6) the second person pronoun once again appears before the verb, although after the question word which must appear in initial position. Once

[3] Yokoyama defines other sets which are relevant to the organization of the discourse. However, these sets are tangential to the discussion of topic.

again this position is expected if deictic items can be shared matters of concern to the interlocators, and these appear preverbally.

In non-discourse-initial contexts, $C_a \cap C_b$ can contain a wider variety of elements than in discourse-initial contexts. The degree of freedom the speaker has in choosing the material to be placed in $C_a \cap C_b$ depends on the context of the utterance. In answers to questions, $C_a \cap C_b$ is relatively set by the context of the question. That is, items mentioned in the question and deictic items (unless they answer the question) are in $C_a \cap C_b$.

(7) Q: S kem [vy] ezdili v London?
with whom you went to London
'Who did you go to London with?' (Yokoyama 1986:299)

A: [V London] [ja] ezdila s moim sotrudnikom
to London I went with my colleague
po kafedre, Xlestakovym.
by department Xlestakovyj
'I went to London with a colleague from my department, Xlestakovyj.'

(7Q) was discussed previously in (6). The discourse division of material in the answer is fixed by the question. The subject pronoun *ja* is a member of the deictic set and thus can be in $C_a \cap C_b$; this placement is reinforced by the fact that it was a member of $C_a \cap C_b$ in the question. In addition, the adverbial *v London* is a member of $C_a \cap C_b$ due to its presence in the previous question. Since both the subject and adverbial are members of $C_a \cap C_b$, they both appear before the verb.

There is more freedom of choice as to the material in $C_a \cap C_b$ in non-discourse-initial statements which are not answers to questions. Yokoyama 1986:327 outlines three basic strategies in such a situation, exemplified below. The first she terms Repeated Topic. This retains in $C_a \cap C_b$ the information that was there in the previous utterance and is sometimes referred to as Continuing Topic. In contrast to Repeated Topic, Shift Topic replaces the material in $C_a \cap C_b$ in the previous utterance with a recently promoted item. This item could come from elsewhere in the previous utterance or from the deictic elements. Finally, Extended Topic contains both the material in $C_a \cap C_b$ from the previous utterance and items recently promoted to $C_a \cap C_b$. An example of these is shown in the short monologue in (8).

(8) a. [U menja] est' odna znakomaja.
at me is one friend
'I-TOP have a friend.'

b. [Ej] let 35.
she-DAT years 35
'She-TOP is about 35.'
c. [Ona] filosof.
she philosopher
'She-TOP is a philosopher.'
d. [∅] živet v Siètle.
lives in Seattle
'(and) she-TOP lives in Seattle.'
e. [U nee] dvoe detej, mal'čik i devočka.
at her two children boy and girl
'She-TOP has two children, a boy and a girl.'
f. [Det'mi] [ona] ne zanimaetsja.
children she not occupies
'She-TOP doesn't care for her children-TOP.'
g. [Gotovit'] (ona) tože ne umeet.
cooking she also not know how
'She-TOP also doesn't know how to cook-TOP.'
(Yokoyama 1986:306)

(8) is a series of utterances by a single speaker. In the first utterance (8a), the topic is *u menja*. This is a viable topic in a discourse-initial utterance because it is a member of the set of deictic elements. The next utterance (8b) exhibits Shift Topic. Material which was mentioned in the previous utterance, *odna znakomaja*, is now reevaluated as a matter of shared current concern and as a result becomes the topic. (8c)–(8e) exhibit Repeated Topics. In each utterance *odna znakomaja*, in various pronominal forms, including being the pro-dropped subject in (8d), is the topic. In fact, all pro-dropped subjects in Russian must be topics. The conditions which govern pro-drop in Russian warrant further research. However, it appears that pro-dropped subjects in Russian are always topics. The referent of the subject must be established previously in the discourse, and if a new topic is introduced and the subject is no longer a topic, then an overt subject pronominal must be used for the subject. (8f) has two elements in $C_a \cap C_b$, i.e., there are two topics of the utterance. This is an example of Extended Topic; one of the topics, *ona*, is a Repeated Topic because it was in $C_a \cap C_b$ in the previous utterance, while the other topic, *det'mi*, is a Shift Topic because it was mentioned in the previous utterance, but not as a member of $C_a \cap C_b$. Note the order of the two topics in (8f): the material more recently placed in $C_a \cap C_b$ appears before the material which was there in the previous utterance, i.e., the Switch Topic *det'mi* appears

before the Repeated Topic *ona*. This ordering is the most common in utterances with multiple topics. Finally consider (8g) which also contains an Extended Topic. Once again, *odna znakomaja*, this time represented by the pronoun *ona*, is a member of $C_a \cap C_b$. However, there is a new member of this set, *gotovit'*. At first glance, it might appear that this topic is unrelated to the previous utterance, counter to what has been claimed about topic choice. However, *gotovit'* is a permissible topic because in this context it forms a set with 'caring for children' of activities women are traditionally involved in. Yokoyama argues that members of this type of set are possible topics if a member of that set was mentioned in the previous utterance. Thus, in (8g) there is a Repeated Topic and a Shift Topic, resulting in an Extended Topic. Once again, the ordering of the elements within the Extended Topic reflects the order in which they were added to the set $C_a \cap C_b$.

There is obviously much freedom in a speaker's deciding what belongs in $C_a \cap C_b$. Yokoyama 1986:253 discusses how items towards which the speaker is empathetic can be promoted 'egocentrically' into $C_a \cap C_b$, e.g., family members. Another form of promotion has to do with what she terms 'anthropological empathy.' Anthropological empathy promotes thematically prominent items to $C_a \cap C_b$.[4] This promotion accounts for why subjects and other thematically prominent arguments, such as dative experiencers, often precede verbs even when they are not canonical items of $C_a \cap C_b$. This is discussed in greater detail in section 5.4.

So, one possible definition of topic, and the one which I adopt here, is the material in the set $C_a \cap C_b$. This is a broader notion of topic than some in that it allows multiple items to be topics simultaneously.[5] Although this may at first appear to be less constrained than other definitions of topic, it is the most practical for describing the Russian data (as is not surprising given that the majority of Yokoyama's data are from Russian).

4.2 Notions of Focus

Focus can be roughly defined as new information, relative to some state in the discourse. Determining a precise definition of 'new' is more

[4] Yokoyama refers to thematic relations as grammatical relations, but makes clear that she does not mean notions such as subject and object by this term, but rather agents, patients, etc. (Yokoyama 1986:274). This difference is important in impersonal expressions and certain psych verb constructions in which the thematically highest argument is not the grammatical subject.

[5] Even more constrained notions of topic, such as in Kiss 1993, who defines topic as the element of which the sentence is predicated, allow for multiple topics (section 4.1.2).

complicated. What is new is relative to what the speaker believes the addressee knew prior to the relevant utterance. In addition, although it is generally argued that all new information must be focus, it is not the case that all focused items must be, in the most immediate sense of the word, new information. Take the following example discussed by Rochemont and Culicover 1990:21.

(9) Q: Who does John's mother like?
A1: John's mother likes MARY.
A2: John's mother likes JOHN/HIM.

The focus in A1, *Mary*, is new information which answers the previous question. As such, it poses no problem for the simplest view of focus being new information. However, consider answer A2 to the same question. Here, the focus *John* is not strictly speaking new information because it is mentioned in the question and in the first part of the answer. Rochemont and Culicover claim that the focus in A1 is both contrastive and presentational; it is presentational in that it introduces new material into the discourse. In contrast, the focus in A2 is only contrastive; it is not presentational since *John/him* is already in the domain of discourse.

4.2.1 Question-Answer Pairs

Question-answer pairs are often used to determine the focus of a clause. Rochemont and Culicover 1990:19 make the generalization in (10), a commonly observed fact that supports this classic test for focused material.

(10) In a well-formed simple question/answer sequence, all and only the information provided in the response that is not contained in the question is focused.

(10) does not make any claim as to whether the information provided in the answer is c(ontext)-construable,[6] i.e., is already part of

[6] The definition of c-construable is given in (i).

(i) *a* is c-construable if (i) *a* is under discussion, or
(ii) *a* is an indexical expression.

C-construability becomes an issue for the following reason. Rochemont 1986:172 defines focus as follows:

(ii) If *a* is not c-construable, then *a* is focused.

There is a problem with the definition in (ii) in that some c-construable items are focused, as (9:A2). To avoid this problem, Rochemont 1986:175 has contrastive focus defined as in (iii).

(iii) If *a* is a focus and *a* is c-construable, then *a* is a Contrastive focus.

However, this makes the definition of focus in (ii) incomplete since the word 'focus' in (iii) does not fall under its definition.

the discourse.[7] However, it is new information within the immediate context of the question, since, even if the existence of the contrastively focused entity is known, its participation in the particular situation described is unknown, new information. Rochemont and Culicover's 1990:20 more generalized statement of focus is given in (11).

(11) If *a*, a phrase in a sentence S, constitutes new information in C, the utterance context to which S is being added, then *a* must be a focus in S.

This definition, which makes explicit reference to new information, brings us to the next section which discusses new-information focus. A distinction is made between simple new-information focus (section 4.2.2) and contrastive focus (section 4.2.3). As discussed in section 4.3.4.2, the word order possibilities in Russian differ with new-information and contrastive focus.

4.2.2 New Information

Jackendoff 1972 divides clauses into focus and presupposition.[8] Roughly speaking, the focus of the sentence denotes the information in the sentence that is assumed by the speaker not to be shared by the speaker and the hearer, while the presupposition denotes the information that is assumed by the speaker to be shared by the hearer (Jackendoff 1972:230). Focus assignment is a process which divides the sentence into focus and presupposition based on the syntactic representation: certain parts of the sentence are mapped onto the focus, others onto the presupposition (all information in the sentence is one or the other; in this way the division is similar to the theme-rheme division of the sentence).

He defines a one-place predicate $\text{Presupp}_s(\text{x})$ which is formed by replacing the focus of the sentence by an appropriate semantic variable (see Jackendoff 1972:242–245 for details on what constitutes an appropriate semantic variable for these purposes). The presuppositional set is then obtained via lambda extraction, giving $\lambda\text{x}\text{Presupp}_s(\text{x})$. Jackendoff gives the following conditions on this set:

(12) $\lambda\text{x}\text{Presupp}_s(\text{x})$ is a coherent set in the present discourse
is well-defined in the present discourse
is amenable to discussion
is under discussion

[7] Topics, using a definition of topic similar to that of Yokoyama 1986, must be c-construable, although all c-construable material is not topic. C-construable material more exactly maps onto the notion theme, which includes discourse-neutral items.

[8] Jackendoff 1972:229–230 bases this analysis on an intuition in Chomsky 1970.

What is useful in this is that in asserting a declarative sentence, a statement is being made that the focus is an element of the presuppositional set in that sentence. If the focus is not an element of that set, then the sentence will be false.

(13) Focus $\in \lambda$xPresupp$_s$(x)

The above definitions of presupposition and focus allow Jackendoff to propose the definition of focus assignment in (14). He assumes that the focused constituent is associated with a feature [F] in the syntax. This feature, or more precisely its association with a particular phrase structure node, is also relevant to the assignment of intonation and emphatic stress.

(14) In a sentence S, with otherwise determined semantic representation SR, the semantic material associated with surface structure nodes dominated by F is the Focus of S. Substitute an appropriate semantic variable x for Focus in SR to form the function Presupp$_s$(x). The presupposition of S is then formed as (12) and the assertion is (13). (Jackendoff 1972:247)

Jackendoff is particularly interested in the interaction of emphatic stress and intonation with focus (all of his examples are from English). In particular, he notes that if a phrase P is chosen as the focus of a sentence S, the highest stress in S will be on the syllable of P that is assigned stress by the regular stress rules (Jackendoff 1972:237). For example, if the indirect object NP of a sentence is focused, then the greatest stress of the sentence will fall on whichever syllable of that NP would normally be stressed.[9]

There is an interesting class of sentences which form a subset of sentences with new-information focus. These are presentational sentences whose purpose is to introduce the new information into the discourse. Niño 1993 argues that in Spanish the arguments of presentational sentences remain within the VP.[10] The motivation behind this is that,

[9] Jackendoff 1972:241 formalizes this rule as in (i); emphatic stress is applied to the vowel which would receive primary stress in the [+F] constituent.

(i) V $\longrightarrow$ [emph stress] / [X [1stress] Y]$_F$

[10] Some sample Spanish VSO sentences are shown below (see Niño 1993 for further data; the data presented here are hers).

(i) Rodean montañas el valle.
surround mountains the valley
'Mountains surround the valley.'

(ii) No comprendía el buen sacerdote los malos corazones.
not understand the good priest the hard hearts
'The good priest did not understand hard hearts.'

especially with transitive predicates, these sentences present the argument structure of the verb and that the argument structure is preserved by the arguments remaining within the VP. The fact that the subject does not move to SpecIP in Spanish presentational sentences results in VS(O) word order in these sentences, in contrast to the usual SVO order. This same generalization holds for Russian presentational sentences, so that in presentational sentences the arguments of the verb remain within the VP, with the subject appearing after the verb.

4.2.3 Contrastive Foci

There is a distinction between simple new-information focus and contrastive focus. Kiss 1993 argues that the difference between the two is whether the set over which the focus operates is open or closed (Kiss 1993:26). In particular, contrastive focus occurs when the set is closed, i.e., when the interlocators know the members of the set. If the set is closed, then the focus does not simply pick out an entity for which the sentence is true. It also designates a complement, all of the entities in the closed set which were not chosen as the focus, with which the focused entity is contrasted.

In both Hungarian and Russian, contrastive foci are marked by emphatic stress (section 4.3.4.2). In Kiss's 1993 analysis of Hungarian, both contrastive and non-contrastive foci appear in SpecVP, the focus position. The difference between the two is signalled by stress, so that preverbal contrastive foci are emphatically stressed, while non-contrastive foci are not. The difference in meaning between the two types of focus is reflected by the phonology, not by the syntactic position in which the focus is found. That is, the Hungarian facts show that the difference between contrastive and non-contrastive is not one of syntactic scope since both types of focus can have scope over the VP. Examples of each type of focus in Hungarian are shown in (15).

(15) a. [János] ette meg a süteményt.
John ate perf the cookie
'John-c.foc ate the cookie.' (Kiss 1993:25)

b. [Egy autó] állt meg a ház előtt.
a car stopped perf the house in-front
'A car-foc stopped in front of the house.' (Kiss 1993:25)

In (15a) and (15b) the focus appears before the verb. However, in (15a) the focus is emphatically stressed and hence a contrastive focus; (15a)

(iii) Cruzó una sombra la calle.
crossed a shadow the street
'A shadow crossed the street.'

answers a question like 'who ate the cookie' in which a set of possible cookie eaters is known. In contrast, the focus in (15b) is not contrastive and can answer a question such as 'what happened', as well as 'what stopped in front of the house'. In either case, the car is simply being identified as something which stopped in front of the house; there is no implication that it was the only thing which did this or that there was a set of possible entities which could have stopped in front of the house. See Kiss 1993 for more discussion of these and related examples.

Thus, the difference between contrastive and new-information or non-contrastive focus depends on the nature of the set from which the focus is chosen. With contrastive focus, the focus is chosen in contrast to the other, known members of the set which were not chosen, while with non-contrastive focus, no information is provided about these other members.

4.2.4 The Relevant Notions

One question is whether these different types of focus need to be represented in different ways or whether some of them can be unified. Dik et al. 1980 suggest that there are several different types or subdivisions of focus used in different situations.[11] However, each language does not have a unique way of encoding each type, e.g., separate morphemes or positions. Although Russian uses position and intonation to mark focus, not all of Dik et al.'s types of focus are uniquely marked. In addition, not all possible combinations of foci can exist in a single sentence since some are incompatible with others.

For present purposes, three types of focus are distinguished. The first is contrastive focus (section 4.2.3). The basic idea behind contrastive focus is that it provides information which contrasts with that which might be expected and emphasizes that a certain item, and not others which belong to a similar class, is the information in question. Contrastive focus is often associated with an exhaustive listing reading. In Russian this type of focus is marked by sentence stress and can appear in a number of positions in the clause, although it usually appears immediately before the verb (section 4.3.4). The second

[11] Dik et al. 1980:44–47 describe five parameters that determine the types of focus available cross-linguistically. The first is the scope of the focus, namely the predicate as a whole or some constituent. The second is whether it is emphatic or contrastive. The next is its relation to pragmatic information known to the addressee, namely does it limit, add to, fill in, etc. a piece of information. The fourth is whether the focus is new or given. Finally, focus can be either exhaustive or not. The fourth parameter, that of new versus given, essentially distinguishes focus from topic, which refers to given information, although it might distinguish whether the focus is introducing a new referent into the discourse.

type is new-information focus (section 4.2.2). This is the most common type of focus and can be indicated by the ever-popular question-answer test for focus (section 4.2.1). New-information focus appears clause-finally in Russian and is usually marked by neutral intonation in which a falling tone marks the right edge of the focused constituent (section 4.3.4). The final type of focus which we are concerned with is presentational focus whose main purpose is to introduce new referents into the discourse. As with new-information focus, presentational foci appear clause-finally. They remain within the VP and often have a slightly different intonational pattern, requiring a following clause to complete the scene created by the presentational sentence.

4.3 Russian Topic and Focus

This section sets out the data to be analyzed. As mentioned above, linear order in Russian is used to encode discourse functions, not grammatical functions. In the most basic cases, initial elements are interpreted as topics and final ones as foci, although intonation and stress interact with these, especially with the assignment of focus. With neutral intonation, there is a sentence-final falling tone which falls on the right edge of the focused constituent. However, sentenceor emphatic stress can appear anywhere in the sentence, and the element on which it falls is contrastively focused. Traditionally, Russian sentences are divided into two parts: the theme and the rheme. With neutral intonation, the theme, which is generally described as given information, always precedes the rheme, which is new information. Once sentence stress is introduced into the picture, the rheme can precede the theme, and discontinuous themes and rhemes are permitted. A three-way division of the Russian sentence provides a more explanatory account of Russian word order. Such a three-way division splits the theme into topicalized and discourse-neutral information. Topics always precede discourse-neutral material, and given the proper intonation and sentence stress, foci can appear in different positions relative to the non-focused items.

Before discussing the Russian data, section 4.3.1 briefly outlines the traditional descriptions of Russian word order since these provide an overview of the general issues and intuitions to be captured by any analysis of Russian word order. The data are discussed in the remaining sections of this chapter.

4.3.1 Traditional Divisions of the Russian Sentence

The Russian linguistic literature discusses the basic word orders in Russian and other Slavic languages in terms of Functional Sentence

Perspective (FSP) or Topic Focus Articulation (TFA) (Adamec 1966, Daneš 1974, Firbas 1964a, 1964b, Hajičová 1974, Krylo-va and Khavronina 1988,[12] Mathesius 1964, Sgall 1972). According to FSP, each sentence is divided into two parts: a theme and a rheme. The theme constitutes given information and the rheme new information. In some versions of this approach, the material internal to the theme and the rheme is also arranged along a hierarchy.[13] This, for example, is roughly what Sgall et al. 1986 refer to as Communicative Dynamism; newer, more important information is more dynamic than older, more predictable information.

Under this schema, all information must fall into either the theme or the rheme. The rheme, being new information, is equivalent to focus. However, the theme consists of all non-focused material, regardless of whether it is topicalized or discourse-neutral. The varying identities of the items in the theme is recognized by those working with this system. For example, Krylova and Khavronina 1988 discuss the order of constituents within a complex theme. From their discussion, it is clear that topicalized material, i.e., the items which are of immediate interest to both speakers, precedes the thematic material which does not fit this description, i.e., non-topicalized but also non-rhematic material.

Another difficulty with the traditional two-way division of the Russian sentence, be it into topic–comment or theme–rheme, is that of determining what role the verb plays in the division of the sentence. When the verb is part of the rheme, there is no problem with the bipartite division. However, in many sentences, the verb is not focused but at the same time has a different role in the discourse function of the sentence than the other thematic material. This led Firbas 1965 to propose that the verb plays a transitional role between the theme and the rheme. Thus a three-way division arises of theme, transition, rheme. If this three-way division is further expanded to allow elements other than the verb to appear in the transition, then this division closely resembles the division into topic, discourse-neutral material, and focus: the theme corresponds to the topic, the transition to discourse-neutral material, and the rheme, as before, to the focus.

All versions of these theories agree that in non-emotive sentences, i.e., ones with neutral intonation, the theme precedes the rheme and

[12] Krylova and Khavronina 1988 is written for upper level students of Russian and is based on traditional word order studies.

[13] This continuum is useful in ordering multiple topics in Russian. As detailed by Yokoyama 1986, topics more recently promoted to the set of topics appear before those which are already in the topic set (section 4.1.3). However, topics are distinct from discourse-neutral information and from foci.

that the sentence is organized along a scale from given to new information (section 4.3.4.1). This is true of theories with only a two-way division and those with a three-way division. However, as often noted, this given-to-new generalization holds only of sentences with neutral intonation. Once emphatic stress is introduced into the picture, the word order variation increases dramatically (section 4.3.4.2; Yokoyama 1986, Krylova and Khavronina 1988). With emphatic stress to mark the contrastively focused material, word order *per se* is no longer needed for this purpose.

The Functional Sentence Perspective division of the sentence provides valuable insights into the communicative nature of a given utterance and its corresponding word order. However, this division can be furthered by utilizing the notions of topic and focus (Dik et al. 1980, Li 1976 and articles therein, Rochemont 1986, Yokoyama 1986). A working definition of topic and focus, as discussed in sections 4.1 and 4.2, allows us to examine more closely the role of discourse functions in the word order of Russian and in turn how the syntax encodes these discourse functions.

4.3.2 Russian Topics

The sentences in (16) exemplify Russian topics as shared matters of current concern (section 4.1.3; Yokoyama 1986).

(16) a. [Na stole] stojala lampa.
on table stood lamp
'There was a lamp on the table-TOP.' (Chvany 1973:266)

b. [Lampa] stojala na stole.
lamp stood on table
'The lamp-TOP is on the/a table.'

c. [Rasskazov] [ja] pročitala mnogo.
stories-GEN I read many
'Stories-TOP, I-TOP read many of.'

d. [Ivan], [ja] [ego] ne ljublju.
Ivan-NOM I him-ACC not like
'Ivan-TOP, I don't like (him).' (Gundel 1988:185)

e. [Včera] priexala mama.
yesterday came mother
'Yesterday-TOP, mother came.'

The bracketed constituents in (16) are all topics. One of the first things to notice is that they all appear in initial position, before the verb. In addition, they tend to be definite and are often pronominal. This is particularly striking in the classic examples in (16a) and (16b).

With the word order in (16a), the topic *na stole* is definite while the subject *lampa* is focused and as a result is frequently interpreted as indefinite. When the order of these phrases is reversed so that the subject appears before the verb and the PP after it, as in (16b), it is the subject that is definite, while the PP is definite or indefinite depending on the context. (16c) has two topics, both of which precede the verb.[14] (16d) contains an example of an external topic, *Ivan* (section 4.1.1, 5.1.1.1). The external topic is a left-dislocated item and is coreferential with an argument of the clause, the pronoun *ego*. There are two internal topics in this sentence which appear after the external topic. Finally, in (16e) the adverb *včera* is the topic, giving the period of time of concern to the speakers.

One question to ask is: what can be topicalized? Only constituents, specifically maximal projections, can be topicalized. An NP can be a topic, but not a noun exclusive of its modifiers. This is the result of independent syntactic constraints on what types of constituents can be moved.[15] A second question is: how many topics can a given utterance have? multiple topics are allowed and in fact relatively frequent. With external topics, such as that in (16d), there is only one topic position which can be filled.[16] However, the more common internal topics are found in adjuntion structures, and multiple adjuntion is common (section 5.1.1). Neither of these topic positions need be filled, either overtly or with a pro-dropped subject. So, it is possible to have a sentence without a topic (section 5.1). This naturally leads one to ask: how many topic positions are there and how many types of topic?

I argue that there are essentially two types of topic in Russian. These correspond roughly to Aissen's 1992 internal and external topic (section 4.1.1, 5.2.2). The external topic is a left-dislocation structure outside the CP. In section 5.1.1.1, I argue, following Rudin 1985 for Bulgarian and Aissen 1992 for Mayan, that these topics are under a

[14]The first topic *rasskazov* differs slightly from many topics in that it is not an argument of the verb. Franks and House 1982 provide a detailed discussion of genitive themes. They provide evidence from mismatches in number and in the case of modifying adjectives, as well as from lexically idiomatic quantifiers, that these genitives are not moved from an argument position. Instead, these constituents involve a null quantifier, which is responsible for the genitive case, and the interpretation of this quantifier is the result of quantifier raising of a constituent of the sentence, e.g., in (16c) *mnogo* would raise at LF.

[15]Russian is freer than English in allowing agreeing adjectives to move out of the NP; in contrast, unlike in English, prepositions in Russian cannot be stranded.

[16]Loren Billings (p.c.) suggests that there may be sentences which contain more than one external topic, a not implausible suggestion. Unfortunately, we have not been able to find a solid example of multiple external topics, but the reader should bear in mind that further research may show that this is the case.

node E(xpression) (Banfield 1973, 1982). Although external topics are not arguments of the verb, they can be coreferential with one. These topics have traditionally been called strong topics in Russian (Comrie 1980, Gundel 1988). Unlike external topics, internal topics are within the clause and are often arguments. In section 5.1.1.2, I propose that internal topics are adjoined to IP. The adjuntion structure predicts that given an appropriate context multiple topics are allowed; this prediction is borne out.

4.3.3 Russian Foci

As mentioned in section 4.2, linguists have discussed a number of different types of focus. The interpretation of focus is often connected to the intonational patterns of the sentence, with the intonational peak marking the focus. This is very much the case for fixed word order languages like English, where there are limited ways in which to mark focus constructions. Russian allows intonation and word order to encode focus. Russian encodes contrastive focus, new-information focus, and presentational focus, which can be considered a type of new information focus. The sentences in (17) demonstrate how these focus phenomena are encoded.

(17) a. Word-order:
Čitaet knigu [otec].
reads book father
'Father-FOC is reading a book.'

b. Sentence/emphatic stress:
Boris [VYPIL] vodku.
Boris drank vodka
Boris drank-FOC the vodka.

c. Clefting:
Èto [Boris] vypil vodku.
it Boris drank vodka
It is Boris-FOC (who) drank the vodka. (Gundel 1988:1)

d. *li* Questions:
[Knigu] li ona čitaet?
book Q she read
Is it a book-FOC that she is reading?

e. Morphological Association with Focus:
On uedet [segodnja že].
he will leave today ŽE
He will leave today-FOC.

Of primary interest in this dissertation is focus signalled by word order. The signalling of focus by word order is partially dependent on the intonation of the clause and corresponds to particular structural positions in which the focused material is located. (17a) is the most common pattern: the focused constituent is in final position, marked by the falling tone of neutral intonation, and remains within the VP. In (17b) the contrastive focus is marked by sentence stress which ultimately interacts with word order; although items with sentence stress can remain *in situ*, contrastively focused maximal projections usually appear immediately preverbally. (17c) and (17d) represent two constructions which involve focused constituents. In the cleft construction in (17c) the constituent following the pronoun *èto* is focused. In the *li* yes-no question construction the constituent to which *li* cliticizes is focused (chapter 6). Finally, in (17e) the particle *že* interacts with focused items, making them more emphatic; this particle is one way of indicating contrastive focus in written Russian where sentence stress is not available as a marker of contrastive focus.

4.3.4 (Non-) Emotive Sentences

One of the greatest difficulties in discussing word order in relatively free word order languages is deciding which data to concentrate on. First, I examine what Krylova and Khavronina 1988 refer to as non-emotive speech. Non-emotive speech is used particularly in writing and academic discourse but is also found in standard speech. It contrasts with emotive speech which is prevalent in the colloquial language and in fiction imitating this language. Yokoyama 1986 proposes that the difference between emotive and non-emotive speech is one of sentence stress: non-emotive sentences have no sentence stress, while emotive sentences have sentence stress. Her division of sentences is discussed in detail because the presence of sentence stress plays a crucial role in the possible structures and interpretations of Russian sentences.

Before discussing the details of Krylova and Khavronina's description of Russian word order, it is necessary to define non-emotive speech. In addition to the vague intuition that a sentence is non-emotive, there are two other indications of non-emotive speech. One of these is the theme-rheme division of the sentence. Krylova and Khavronina divide each sentence into theme and rheme. As far as word order is concerned, the principal difference between emotive and non-emotive speech is that in non-emotive speech the theme precedes the rheme while in emotive

speech the rheme precedes the theme.[17] [18] In non-emotive sentences the theme always precedes the rheme.[19] If the rhematic, focused material is found in any other position in the sentence, the sentence is considered to be emotive. Also, it is never the case that thematic and rhematic material are discontinuous in these sentences; for example, in non-emotive speech it is impossible to have theme–rheme–theme.

The second diagnostic of non-emotive sentences is the intonational pattern of the sentence, discussed in greater detail below. Yokoyama 1986 proposes that non-emotive sentences have a particular intonational pattern but no sentence stress; in contrast, emotive sentences have sentence stress, and the placement of this stress interacts with the discourse interpretation of the sentence.

Traditionally, the intonation patterns of Russian sentences have been divided into seven intonational contours (called IKs for their Russian name; see Bryzgunova 1980 for a detailed discussion of the IK system). Each of these contours is associated with a particular type of sentence, e.g., questions, exclamations, emotive speech, although there is much semantic overlap among some of the contours. In general, non-emotive speech is associated with IK-1 which has a falling tone at the end. In a non-emotive sentence, the theme precedes the rheme, and the theme is marked by a slight rise of tone,[20] while the rheme is

[17] Other characteristics of emotive speech are that the structure of constituents can be altered and that the order of constituents may be different in sentences without themes. Finally, it is possible for the theme or rheme to be discontinuous; frequently this occurs when a focused element, i.e., the rheme, appears between the subject and verb (Krylova and Khavronina 1988:113). Some of these differences are only used in Colloquial Russian, while others are found in Contemporary Standard Russian as well. Unfortunately, a complete examination of emotive speech is impossible here. However, section 5.3 examines the most frequent structures found in emotive sentences.

[18] There are sentences composed entirely of the rheme, i.e., the entire sentence is focused (Krylova and Khavronina 1988:24–27). However, there are no sentences composed entirely of a theme; such sentences would have no discourse value because they provide no new information.

[19] This does not mean that in academic writing, which is traditionally non-emotive in nature, orders other than theme–rheme are not found. However, these other orders only occur when there is overt (morpho)syntactic marking of the rheme or focus. For example, *li* yes-no questions sometimes appear rhetorically in writing and in these the focus appears before the clitic *li* (chapter 6).

[20] Krylova and Khavronina 1988:12 provide a sentence in which the rise in tone appears before the end of the theme. On more careful examination of the sentence, it appears that the rise is associated with the topic of the sentence and that non-topicalized elements of the theme follow it. Their sentence is in (i).

(i) Roman Vojna i Mír napisal Lev Tolstòj.
novel war and peace wrote Lev Tolstoj
Lev Tolstoj-RH wrote "War and Peace".

marked by a fall of tone. The falling tone of the rheme is placed on whichever element in the phrase would normally receive the stress, i.e., the stressed syllable of the last word.

However, Yokoyama 1986 argues that this traditional division into seven intonational patterns is misleading and that what separates emotive from non-emotive speech is the presence or absence of sentence stress, which corresponds to emphatic stress. Emotive speech occurs in sentences with sentence stress, non-emotive speech in those without. Yokoyama labels the intonation pattern found with non-emotive sentences Type I. She argues, on the basis of descriptions in the literature and phonetic studies of her own, that Type I intonation consists of a series of LH tones followed by a single HL tone.[21] The series of LH contours are realized with a downstep and their number depends on the number of constituents in the sentence and how quickly and carefully the sentence is pronounced. Consider the sentence in (18), which Yokoyama discusses in detail.

(18) Nad Krakovom nakrapyval doždiček.
over Krakow drizzled rain
'It was drizzling over Krakow.'

(18) is a non-emotive sentence since the thematic part of the sentence, *nad Krakovom nakrapyval*, precedes the rhematic part, *doždiček*, and there is no sentence stress. The Type I intonation pattern for the sentence is shown in (19). The sentence is divided into three syntagms: the first two are marked with a LH contour tone, while the third has a falling HL tone associated with it. The precise location of each contour tone is dependent on the stress pattern of the words in the syntagm. Syntagms are separated by a slash (/).[22]

(19) /nad krakovom /nakrapyval /doždiček
LH LH HL
[downstep]

Since the rheme always comprises the last element (or elements) in a non-emotive sentence, the falling HL tone always falls on some portion

In this sentence, the theme is *roman vojna i mir napisal*, i.e., all of the sentence except the subject. However, the rise in intonation is on the last element of the topicalized object, i.e., on *mir*.

[21] Yokoyama 1986 points out that this is roughly equivalent to IK-1 preceded by a series of other IKs, where IK-1 is the intonational contour found at the end of non-emotive sentences.

[22] The difficulty is then to define what a syntagm is, especially since the syntagmatic division can vary with how the sentence is uttered. For our purposes, what is important is the difference between Type I and Type II intonation. For more details on these two types of intonation see Yokoyama 1986 chapters 6-9, especially the introductory material in chapter 6.

of the rhematic or focused material. However, parts of the focused material may be marked with a rising tone since the focused material need not form a single syntagm to which the falling tone is assigned.

In contrast to Type I non-emotive sentences, emotive sentences, which Yokoyama terms Type II, have sentence stress associated with them. Yokoyama 1986:192 states that sentence stress is not signalled by one particular phonetic factor, but by a combination of factors such as amplitude, pitch, syllable duration, pitch contour, etc. She concludes that sentence stress cannot be defined with reference to its own prosody, intensity, etc., but relative to the intonational character of the rest of the utterance. The sentence stress falls on the focused material, and this element forms the last intonational center of the utterance. Since it is the last intonational center, all the material following the sentence stress forms a single syntagm with the element bearing the sentence stress. The focused material in Type II sentences is frequently found directly preceding the verb.

Consider (20) which is an emotive Type II counterpart to (18). The focused material, *doždiček*, which followed the verb in the non-emotive (18), now precedes it. The syntagmatic and tone contours of (20) are shown in (21) (from Yokoyama 1986:192–195). Sentence stress is marked with an asterisk (*), which falls on the stressed syllable of the focused item.

(20) Nad Krakovom [**doždiček**] nakrapyval.
over Krakow rain drizzled
'It was drizzling-FOC over Krakow.'

(21) /nad krakovom /doždiček nakrapyval
*
LH HL

Instead of the three syntagms found in (19), (21) consists of only two syntagms. The material following the sentence stress, *nakrapyval*, cannot form its own syntagm; instead it forms part of the syntagm with *doždiček*. The first syntagm, *nad krakovom*, has a rising LH contour tone, as it did in the Type I sentence in (19). The falling HL tone is implemented on the stressed syllable of *doždiček*, the word on which the sentence stress falls. There is no contour tone on the verb *nakrapyval*, as expected since it does not form its own syntagm. This observation as to the nature of the material following the sentence stress is similar to that made by Krylova and Khavronina 1988:116 that in emotive sentences the rheme "is made more prominent by emphatic stress, while the theme is pronounced in a low tone almost without stress and it

nearly 'disappears' in pronunciation". Here, theme refers to the material appearing after the focus, i.e., after the stressed *dožd'iček*.

Thus, the discourse function interpretation of a sentence is dependent on both word order and intonation. With Type I intonation, i.e., in non-emotive sentences, topicalized constituents appear before the verb while focused ones appear sentence-finally. In contrast, with Type II emotive sentences, the placement of the sentence stress corresponds to the focused element. This focused element is most commonly found in preverbal position, preceded by any topicalized elements. These generalizations are discussed below and in chapter 5 in which the phrase structure of these constructions is analyzed.

4.3.4.1 Non-emotive Sentences

Although Krylova and Khavronina's analysis must be refined to take into account the subdivision of the theme into topic and discourse-neutral material, their description of Russian word order in non-emotive sentences provides a concise and thorough overview of the data in question. It is important to keep in mind that their generalizations are based solely on linear order; the phrase structure of the sentence is not considered.[23] The structures corresponding to these linear orderings are discussed in chapter 5.

Krylova and Khavronina's discussion is divided into two main parts: patterns involving subjects and their predicates and patterns involving direct objects. The word orders are summarized in (22) and (23); example sentences and contexts are provided in the discussion following.

(22) Non-emotive Speech: Main Variants

	THEME	RHEME
I	SUBJ	PRED
II	PRED	SUBJ
III	ADV	PRED SUBJ
IV	ADV SUBJ	PRED
Va	SUBJ PRED	ADV
Vb	PRED SUBJ	ADV

[23] The closest Krylova and Khavronina come to discussing syntactic structure is to mention that certain types of constituents cannot be broken up, e.g., prepositional phrases.

(23) Non-emotive Speech: Objects

	THEME	RHEME
I	SUBJ	PRED OBJ
IIa	PRED OBJ	SUBJ
IIb	OBJ PRED	SUBJ
III	OBJ	PRED SUBJ
IV	OBJ SUBJ	PRED
V	SUBJ PRED	OBJ

Before providing examples, a few comments should be made concerning the divisions in (22) and (23). The first is that a given linear string may have more than one possible theme-rheme division. For example, in (23:I) and (23:V) the subject is followed by the verb and its object; however, this SVO order has two interpretations, one in which the verb together with its object form the rheme, the other in which the object alone constitutes the rheme. Also, the complex themes found in (22:V) and (23:II) have different discourse functions depending on the order of the elements within the theme. In Krylova and Khavronina's analysis these are subtypes of a theme-rheme division. However, their discussion of these subtypes supports the idea of dividing the theme into topicalized and discourse-neutral material. So, in (22:Va) the subject is the topic of the sentence, while in (22:Vb), when it appears after the verb, it is discourse-neutral. The same holds of the pair in (23). In (23:IIb) the object is the topic of the sentence, while in (23:IIa) the object follows the verb and is discourse-neutral. One final note on the patterns in (22): in these divisions ADV refers to all types of sentence modifiers, not just syntactic adverbs, e.g., locative PPs.

An example of each pattern is shown below (all examples are from Krylova and Khavronina 1988). Each example is introduced by a question to show the theme-rheme division. Using question-answer pairs introduces two biases. First, information which is repeated in the answer of a question would frequently be designated by a pronoun; however, for sake of explication, the noun itself is repeated in the answer. In addition, if a noun phrase was mentioned in the question, there is a tendency for it to be topicalized in the answer. Constituents which form part of the rheme, i.e., the focused constituents, are marked as such in the gloss and are divided from the rest of the Russian by square brackets. All sentences in this section are Type I, non-emotive sentences and have the corresponding Type I intonation with a final falling tone. Examples of emotive sentences are given in section 4.3.4.2.

(24) Main:I

Q: Čto delajut deti?
what do children
'What are the children doing?'
A: Deti [igrajut].
children play
'The children are playing-RH.'

(25) Main:II
Q: Kto igraet?
who plays
'Who is playing?'
A: Igrajut [deti].
play children
'(The) children-RH are playing.'

(26) Main:III
Q: Čto proizošlo včera?
what happened yesterday
'What happened yesterday?'
A: Včera [sostojalos' sobranie].
yesterday take place meeting
'Yesterday a meeting-RH was held-RH.'

(27) Main:IV
Q: Čto oni delali zdes'?
what they did here
'What did they do here?'
A: Zdes' oni [poznakomilis' s rabotami studentov].
here they get acquainted with work students
'Here they got acquainted-RH with the students' work-RH.'

(28) Main:Va/b
Q: Gde naxoditsja pionerskij lager'?
where located pioneer camp
'Where is the pioneers' camp located?'
A: Pionerskij lager' naxoditsja [v Krymu].
pioneers' camp located in Crimea
'The pioneers' camp is located in the Crimea-RH.'

In (24) the verb is focused (or the entire VP since the verb is intransitive). This focusing is forced by the preceding question which asks about the actions of the children. (25) is the reverse of (24). Here, the question asks about the subject, requiring it to be the focus of the sentence, while the action is known, in this case that someone arrived.[24]

[24]The predicate in this situation is the theme of the sentence since it is not focused.

(26) is a presentational sentence, introducing the occurrence of the meeting into the discourse. (27) has a complex theme consisting of the adverb *zdes'* and the subject *oni*, both of which were mentioned in the preceding question. The VP forms the rheme of the sentence. Finally, in (28) the rheme consists of the adverbial *v Krymu*. Once again, the theme is complex and depending on the order of the elements in the theme, there may or may not be a topicalized constituent. (28) has a topic, *pionerskij lager'*, which precedes the verb.

The object patterns are a subset of the main patterns in that they allow the predicate to be divided into a verb and object. Thus, it is not surprising that there is overlap in the patterns. For example, Main I is identical to Object I. This identity is also apparent in the similarity among the questions which provide the theme-rheme context of the sentences.

(29) Object:I
Q: Čto sdelal kolxoz?
what did kolxoz
'What did the *kolxoz* do?'
A: Kolxoz [zakončil uborku urožaja].
kolxoz finished harvest crop
'The *kolxoz* finished-RH the crop harvest-RH.'

(30) Object:IIa
Q: Kto šil?
who sewed
'Who sewed?'
A: Šila mne èto plat'e [Inna].
sewed me that dress Inna
'Inna-RH sewed me that dress.'

(31) Object:IIb
Q: Kto šil tebe èto plat'e?
who sew you that dress
'Who sewed you that dress?'
A: Èto plat'e mne šila [Inna].
that dress me sew Inna
'Inna-RH sewed me that dress.'

(32) Object:III

However, it is not a topic. Verbs in Russian, and in general cross-linguistically, are rarely topicalized.

Q: Čto proizošlo s kartinami?
what happened with paintings
'What happened to the paintings?'
A: Neskol'ko kartin [priobrel mestnyj muzej].
a few paintings acquired local museum
'The local museum-RH acquired-RH a few of the paintings.'

(33) Object:IV
Q: Čto vy sdelali so staroj lodkoj?
what you do with old boat
'What did you do with the old boat?'
A: Staruju lodku my [prodali].
old boat we sold
'We sold-RH the old boat.'

(34) Object:V
Q: Čto ty kupila?
what you buy
'What did you buy?'
A: Ja kupila [šarf].
I bought scarf-RH.
'I bought a scarf.'

(29) involves a complex rheme corresponding to the VP. The subject *kolxoz*, which is mentioned in the question, is topicalized and appears before the verb. (30) and (31) are similar in that they both involve focusing the subject which must appear in final position in non-emotive sentences. The difference between the two is in the order of elements in the complex theme. In (30) there is no topic and the discourse neutral material *mne èto plat'e* appears after the verb but before the rheme; note that (30Q) might be analysed as containing an elided object. However, in (31) the theme contains topicalized information, namely *èto plat'e mne*, which precedes the verb. In (32) the rheme consists of the subject and verb, exclusive of the object. The order of elements in this complex rheme is fixed under Type I intonation: the subject must follow the verb. Otherwise the sentence will be incorrectly divided such that the rheme consists only of the verb while the object and the subject form a complex theme. This is what happens in (33) where only the verb *prodali* is focused; the object and subject are both part of the theme. Finally, (34) is an example of a focused object, *šarf*.

The part of each sentence which answers the preceding question is the rheme. For this reason, it is equivalent to the notion focus, since the answer to a question is new information (section 4.2), and will be referred to as such in the following discussion. In each example,

the focused material occurs in final position, following the theme. The falling HL tone of Type I intonation marks the right edge of the focused material in the VP (section 5.3).

The theme, as demonstrated in the above examples, is not identical to most notions of topic. The theme contains both discourse-neutral information and topic information. In section 5.1.1, I propose that topics are adjoined to IP, while discourse-neutral information remains *in situ*. This means that topics appear to the left of the finite verb. This ordering is seen most readily in the sentences which have complex themes. In these sentences, the topicalized items precede those which are discourse-neutral. Material in the theme which appears before the verb is topicalized, while that which appears after the verb is discourse-neutral.

4.3.4.2 Emotive Sentences

The previous section discussed how word order interacted with the topic and focus interpretation of non-emotive sentences. This section is concerned with these interactions in emotive sentences. One of the primary differences between non-emotive (Type I) and emotive (Type II) sentences is the presence or absence of sentence stress. However, there is also a difference in word order between the two. In its most basic form, this difference is that in non-emotive sentences the theme precedes the rheme, while in emotive sentences the rheme precedes the theme. This is an over-simplification. In part this is because the notion theme is not fine-grained enough: in emotive sentences the topicalized portion of the theme often precedes the rheme or focus, while the discourse-neutral portion of the theme is sentence final. In addition, these sentences have a contrastive focus reading (section 4.2.3, 4.2.4). The most common patterns for emotive sentences are discussed here.

In emotive sentences, the item on which the sentence stress falls is the focus of the sentence. This is true regardless of its position. However, the focused item is most commonly found directly before the verb. Consider first the sentence which was discussed earlier in regard to the intonation contours in Russian, repeated below as (35). (Sentence stress is indicated by bold faced type.)

(35) a. Nad Krakovom nakrapyval [doždiček].
over Krakow drizzled rain
'It was drizzling-FOC over Krakow.' (Yokoyama 1986:184)
(Non-emotive)

b. [Nad Krakovom] [**doždiček**] nakrapyval.
over Krakow rain drizzled
'It was drizzling-FOC over Krakow.' (Yokoyama 1986:192)
(Emotive)

c. [**Doždiček**] nakrapyval nad Krakovom.
rain drizzled over Krakow
'It was drizzling-FOC over Krakow.' (Yokoyama 1986:194)
(Emotive)

All three sentences in (35) have the same focus interpretation. In (35a) there is no sentence stress and the noun *doždiček* is focused as would be expected in a non-emotive sentence since it is in final position. (35b) and (35c) are emotive variants of (35a) in which the sentence stress falls on the focused *doždiček*. The principal difference between the two is the placement of the adverbial *nad Krakovom*. In (35b) it appears before the focused item and is topicalized, while in (35c) it appears after the verb and is discourse-neutral.

(36)–(39) provide further examples of the word order in emotive sentences.

(36) Q: Kto napisal 'Evgenija Onegina'?
who wrote Eugene Onegin
'Who wrote Eugene Onegin?'
A: ['Evgenija Onegina'] [**Puškin**] napisal.
Eugene Onegin Pushkin wrote
'Pushkin-FOC wrote Eugene Onegin-TOP.'
(Yokoyama 1986:191)

(37) Q: Kto priexal k vam?
who came to you
'Who visited you?'
A: [K nam] [**Anna**] priexala.
to us Anna came
'Anna-FOC visited us-TOP.'

(38) A: Kuda ty ideš'?
where you go
'Where are you going?'
Q: [Ja] [**k Anne**] idu.
I to Anna go
'I-TOP am going to Anna's-FOC.'

(39) Počemu [Boris] často [**na rabotu**] opazdyval?
why Boris often to work was late
'Why was Boris-TOP often late for work-FOC?'

In (36) the focus of the answer is *Puškin*. Unlike in a non-emotive sentence, the focus is found pre-verbally and is marked with sentence stress; the topic, *Evgenija Onegina*, is sentence initial, as in a non-emotive sentence. (37) is a similar example in which a focused subject, *Anna*, appears after the topic and before the verb. In contrast, in (38) the subject is topicalized and appears in initial position; the sentence stress falls on the goal, *k Anne*, which appears directly before the verb. Finally, (39) is a question in which the goal, *na rabotu*, is focused with sentence stress and appears before the verb.

However, although the orders shown above for emotive sentences are the most frequent, i.e., the focused item usually occurs before the verb but after the topic, other orders are possible. In particular, the focused phrase can occur before the topic. This greater freedom of word order is more common in more colloquial speech. In addition, the item with sentence stress can only precede the topic in non-discourse initial utterances (Yokoyama 1986:222). So, the word orders found in the emotive sentences in (40) and (41) are also possible, although they could not occur in discourse initial situations.

(40) Q: Who wrote Eugene Onegin?
A: [**Puškin**] 'Evgenija Onegina' napisal.
Pushkin Eugene Onegin wrote
'Pushkin-FOC wrote Eugene Onegin.' (Yokoyama 1986:191)

(41) Q: Who broke the window?
A: [**Mal'čik von tot v sinej kurtke**] ego razbil.
boy there that in blue jacket it broke
'That boy-FOC over there in a blue jacket broke it.'
(Yokoyama 1986:198)

In (40) and (41) the subject is focused and appears in initial position with sentence stress. The topicalized object follows the focused subject and, as expected, precedes the verb. Earlier it was seen that the other order is possible for (40): the topic can precede the focus with its sentence stress. (41) is less likely to have the alternative order, partially due to the weak stress of the object pronoun, *ego*, in comparison with the complex subject NP with sentence stress; pronouns tend to appear near the verb in Russian, although they are not clitics.[25]

[25]In this way, Russian differs from many of the Slavic languages which have substantial inventories of pronominal and auxiliary clitics (for example, see Hauge 1976 on Bulgarian clitics). However, Russian pronouns do behave somewhat differently than their non-pronominal counterparts, tending to appear adjacent to the verb, especially if non-pronominal arguments are also present in the clause. Further research on the behavior of pronominals should ultimately help elucidate Russian phrase structure.

Thus, although the majority of emotive sentences involve placing the focused item with sentence stress immediately before the verb, this position is not obligatory and other orders are possible, including ones in which the focused item appears initially, preceding the topics. These orderings are discussed in chapter 5.

4.3.5 Conclusions

This chapter discussed the types of topic and focus relevant to Russian word order and how intonation interacts with the interpretation of topic and focus and with word order possibilities. Under neutral intonation, i.e., in non-emotive sentences, the word order is strictly determined by the discourse function organization of the clause: topics precede the verb and discourse-neutral material, which precede the focus. In emotive sentences, i.e., sentences with emphatic or sentence stress, the word order at first appears to be less constrained because focused constituents need not appear sentence finally. However, these non-final foci, which are contrastive in meaning, are always marked with sentence stress and generally occur immediately before the verb, following the preverbal topics.

Following Yokoyama 1986, topic was defined as material of shared current concern; a definition which relies heavily on the speaker's assessment of their addressee. This relatively loose notion of topic allows for multiple topics, as is frequently the case in Russian, i.e., there can be several preverbal constituents, all of which are interpreted as topics. In general, multiple topics are arranged so that the more recently topicalized elements precede older topics. If topics form a set, the items more recently added to the set precede those which were already in the set from previous utterances. There is an additional, more constrained notion of topic: subject-of-predication. A subject-of-predication is the item about which the rest of the clause is predicated; it need not be a grammatical subject. These appear preverbally amongst the other topics. It is important to note that not all sentences have topics; some sentences may be composed entirely of focused material, while others contain both a focus and discourse neutral items, but no topic. Topicless sentences play an important role in determining the underlying phrase structure of Russian (chapter 5).

In non-emotive speech, foci appear clause finally. The size of the focused material varies substantially, although it always forms a continuous unit whose right edge is the final material in the clause. For example, the focus can be a single noun phrase in clause-final position or it can include the verb with all of its arguments following it. However, any arguments which appear before the verb are interpreted as

topics, not as foci, even if the verb itself is in the domain of the focus. This clause-final focus was termed new-information focus because it provides the new information in the clause, usually information about the topic of the clause. Following Jackendoff 1972, among others, it is assumed that all new information in an utterance is focused (but not all non-new information is topic, i.e., old information is not necessarily part of the topic). A further subdivision of focus is needed for the emotive sentences in which sentence stress falls on the focused item. This focused item is a contrastive focus (section 4.2.3; contrastive focus is being used as a convenient term to describe a number of related types of focus, not just those that are strictly speaking contrastive). With contrastive focus, the focused item is a member of a set of related items. The sentence is true of just that contrastively focused item; it contrasts the focused item with the other items in that set. Since contrastive focus involves a known or inferable set of items, these foci sometimes appear to be topical in nature. However, their behavior patterns with foci, not with topics; as such, the new information which they provide, defining them as foci, takes precedence over the fact that they may have already been present in the discourse, i.e., not all items which are already present in the discourse are necessarily topics.

Finally, it is important to recognize that not all constituents are topics or foci. This material was termed discourse-neutral. Discourse-neutral material appears between the initial topics and the clause-final foci. Allowing for discourse-neutral material solves the problem of how to treat non-focused verbs. In systems which only allow a binary division of the sentence, these verbs are often forced to be considered topics. However, they do not have the same interpretation as topics. In addition, it explains why non-focused constituents which appear after the verb have different meanings than the topicalized constituents which appear before the verb. The postverbal constituents are, under this view, discourse-neutral and thus not expected to have a topic interpretation.

In the next chapter, these observations about the discourse function interpretations of different word orders are used to determine the phrase structure of Russian. The relative positions of different discourse functions to the verb are pivotal in this discussion. Of particular importance is the observation that preverbal constituents are topics while postverbal ones are either discourse-neutral or focused. I argue that underlyingly all arguments appear postverbally and that there are preverbal topic positionXSs into which these arguments can move. Similar proposals are made for the interpretation of focused items.

5

S-Structure and Encoding Discourse Functions

Section 4.3.4 discussed the topic and focus interpretations that the various word orders in a sentence have. Several generalizations were made:

(1) Topics appear clause initially.

(2) In non-emotive sentences foci appear clause finally.

(3) Sentence stress marks contrastive foci which usually appear immediately before the finite verb.

(4) Discourse-neutral information follows topicalized information.

The question then arises as to the phrase structure of these sentences. Chapter 3 argued that Russian has an underlying structure similar to that usually proposed for English in that the subject is projected higher than the object.

This chapter discusses the surface phrase structure of Russian, i.e., where topicalized and focused phrases are located in order to derive the correct linear orderings and interpretations. The assumption is that word order is a reflection of the phrase structure. The phrase structure encodes both dominance and precedence relations, and at S-structure this closely reflects the surface order. As such, there are no substantial reorderings at the level of PF (phonetic form). Although there may be slight PF reorderings to account for clitic placement, the major ordering of the constituents is a direct reflection of their S-structure positions. As such, one of the goals of an analysis of word order is to explain why constituents move out of their D-structure positions into their S-structure positions and what these S-structure positions are. To account for the word order variations in Russian, I argue that constituents move not just to get Case or inflectional features, but also to receive discourse function interpretations.

Section 5.1 examines topic and focus in non-emotive sentences, first arguing for a basic VSO order. I then argue that topics are left-adjoined to IP, while SpecIP is reserved for contrastive foci; motivation for these positions can be found in their relative ordering. Section 5.2 discusses several proposals for structurally encoding topic and focus in other languages, namely Hungarian, Mayan, and Bulgarian; this allows for a comparison with word order phenomena found in these languages. The next section discusses sentences in which the verb's arguments remain in the VP and how intonation and focus interact in these constructions. In particular, all new-information focused arguments are within the VP or right-adjoined to it; the right-edge of the focused constituent is in final position, marked by the falling tone of neutral intonation. Finally, the last section examines why Russian is usually considered an SVO language and how the VP internal subject analysis proposed here is compatible with the data often assumed to support an SVO analysis.

5.1 Positionally Marked Discourse Functions

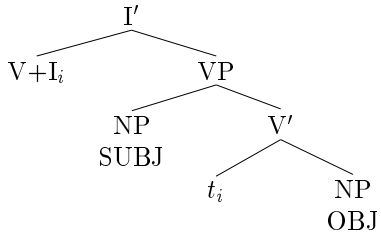
Before discussing the structural positions in which topics and foci are found, I argue that contrary to common assumptions Russian is a VSO language, not an SVO one. The Russian data support the VP-internal subject hypothesis (Koopman and Sportiche 1985, Fukui and Speas 1986) which projects all arguments of the verb, including the subject, in the VP at D-structure. Unlike in English, SpecIP in Russian is not subject position, but one associated with discourse functions, namely contrastive focus. In fact, all movement out of the VP in Russian results in either topic or focus interpretation.[1] [2] Related proposals have been made for a variety of languages, where different specifier positions are associated with different discourse functions: see Alsagoff 1992 for Malay, Diesing 1988 for Yiddish, Kiss *to appear* for Hungarian, Kroeger 1991 for Tagalog, Ouhalla 1992 for Standard Arabic, among others.

In this structure, V^0 raises to I^0 in finite clauses. Thus, an *in situ* subject will follow, not precede, a finite verb, as in (5). This means that the discourse-neutral position for the subject is after, not before, the verb and it appears in this position unless it is moved into topic or focus position.

[1] The restriction on movement applies to arguments of the verb or predicate, not to movement of the verb to I^0.

[2] I am making this claim specifically for Russian. However, it appears that many languages which encode discourse functions structurally have the same restriction. It would be interesting to determine under what circumstances movement out of the VP is *not* correlated with discourse functions.

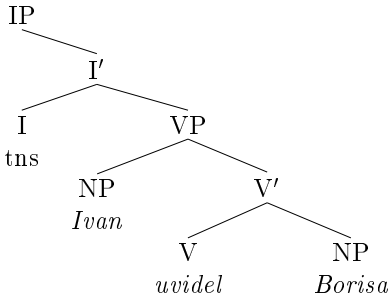
(5)



The D-structure of a sentence with a transitive verb, as in (6), is shown in (7). Since the subject in (6) precedes the verb, it must ultimately move out of the VP and adjoin to IP where it is interpreted as a topic (see section 5.1.2 for why topics are not in SpecIP). The verb moves to I^0 to get inflectional features, and from there it assigns nominative case to SpecVP.

(6) Ivan uvidel Borisa.
Ivan saw Boris
'Ivan saw Boris.'

(7)



There are two major types of evidence that support the VP-internal subject analysis over one in which the subject is either base generated in SpecIP or obligatorily raises there: first, the word order and interpretation of sentences which are composed entirely of focused material, and second, the interpretation of sentences in which the subject follows the verb or main predicate, as opposed to those in which the subject precedes the verb. Additional evidence is found in the word order of questions, in which some of these contrasts are more easily seen.

Sentences composed entirely of focused material are relatively rare, but in non-emotive speech they occur invariably with the subject following the verb (Krylova and Khavronina 1988:24–27).

(8) a. [Šel dožd'.]
go rain
'Rain was falling.'

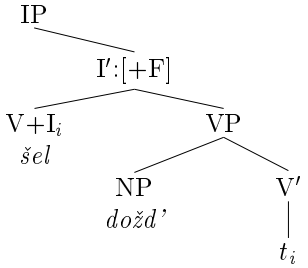
b. [Neslyšno proletela kakaja-to neizvestnaja ptica.]
noiselessly fly past some unknown bird
'Some type of unknown bird flew noiselessly past.'

c. [Prislal muž den'gi.]
sent husband money
'My husband sent (me) the money.'

In (8) the entire sentence is focused and provides the answer to a question such as 'What was happening?'. The question is how the appropriate interpretation is assigned to this, and only this, word order. In particular, the subject cannot precede the verb and have the same focused reading. Assuming a VP-internal subject, the answer to this question is simple. The intonation marks the right-edge of the focus constituent with a falling tone; this falling tone is always sentence final, forcing the final constituent to be contained in the focused constituent. With sentences such as those in (8) the subject is focused. However, the subject marks only the right-edge of the focused constituent, which can be the entire sentence. If subjects were obligatorily raised to SpecIP, we would expect the order subject–predicate when the entire sentence was focused, not the actual predicate–subject order (see below).

The structure of (8a) is that in (9), with the scope of the intonationally marked focus indicated by the feature [+F].

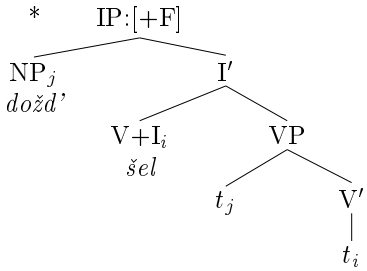
(9)



If the subject were obligatorily in SpecIP, as in an SVO account, we would expect it to be possible to focus the sentence with the subject preceding the verb, with the falling tone on the verb. However, this ordering does not occur, and it is not obvious how to block it in favor of the order in (8). In order for an SVO account to produce the correct word order, the subject would have to move out of SpecIP into a final focus position, perhaps right-adjoining to IP, and then focus of the sentence would be marked intonationally. This is a problem since the spread of focus leftward in this manner generally applies only when the arguments are still in the VP, not when they have moved out of the VP, e.g., preverbal arguments are not within the domain of focus but

instead receive topic interpretation. The structure which the SpecIP subject analysis incorrectly predicts to be possible is shown in (10).

(10) *



Additional confirmation of the VP-internal subject comes from word order in wh-questions. In wh-questions, all of the wh-phrases are fronted to SpecCP (Rudin 1988, 1989, section 3.6.4).[3] When some phrase other than the subject is questioned, the unmarked word order is wh-phrase – verb – subject, which is to be expected if the verb has moved into I^0, thus appearing before the subject, as in (11). As shown below, the verb does not move to C^0 since topics appear between the wh-phrases and the verb.

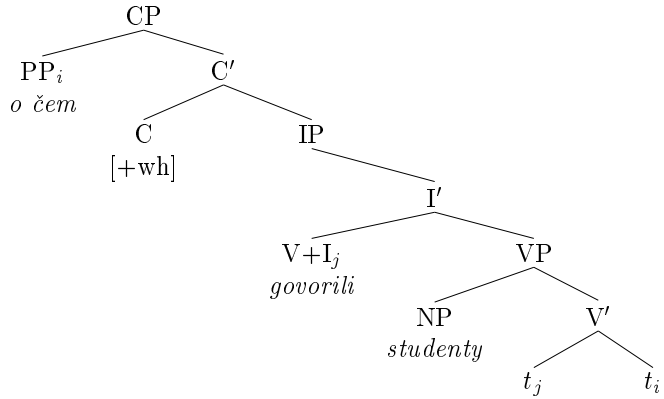
(11) a. Čto videli deti?
what see children
'What did the children see?'

b. Na kakom zavode rabotaet vaša sestra?
at what factory works your sister
'What factory does your sister work at?'

c. O čem govorili studenty?
about what talked students
'What did the students talk about?'

[3]There is reason to believe that SpecCP is also interpreted as a focus position, but constituents can appear there only when licensed, i.e., in wh- or certain yes-no questions. See sections 6, 9.3, and 10.2 for further discussion.

(12)



(12) is the structure of (11c). The wh-phrase *čem* has moved to SpecCP. Since the subject, *studenty*, is not a topic, it remains in SpecVP. The verb undergoes head-movement to I^0 over the subject in SpecVP. This results in the discourse-neutral subject following the finite verb.

It is possible for the subject to precede the verb in questions. However, this only occurs when the subject is topicalized and would thus be adjoined to IP, preceding the inflected verb. Pronominal subjects frequently precede the verb since they are usually topicalized (Krylova and Khavronina 1988:94-95). The fact that topicalized constituents can appear between the wh-phrase and the verb indicates that wh-questions in Russian cannot be analyzed as involving movement of the verb in I^0 to C^0. If such movement did occur, no constituent would be able to separate the wh-phrase and the inflected verb.

(13) a. Čto [on] slyšit?
what he hear
'What does he-TOP hear?'

b. A vaša sestra, gde [ona] rabotaet?
and your sister where she work
'And your sister, where does she-TOP work?'

c. Kogda [Lermontov] rodilsja?
when Lermontov born
'When was Lermontov-TOP born?'

(13a) and (13b) have pronominal subjects which have been topicalized and appear before the verb. (13c) is a possible word order only when *Lermontov* is the topic of the sentence, e.g., if there was a discussion going on about Lermontov's life. If *Lermontov* were not the topic, it

would have to appear after the verb, in SpecVP. The structure of these sentences is discussed in more detail in section 5.1.1.2.

Finally, consider sentences with focus phrases that are composed of more than one element, in particular of the subject and verb.

(14) Q: What happened yesterday?
A1: Včera [priexal brat].
yesterday arrived brother
'Yesterday (my) brother-FOC arrived-FOC.'
A2: Včera [prislal muž den'gi].
yesterday sent husband money
'Yesterday (my) husband-FOC sent-FOC (me) money-FOC.'

The question in (14) ensures that the entire proposition other than the adverb *včera* is focused since it is new information. The answer to such a question must have the word order found in (14); in particular, the subject must follow the verb. Also, in (14A2) the order of arguments after the verb is subject then object, as expected if both arguments are in their base positions. Consider briefly how the focus is correctly assigned.

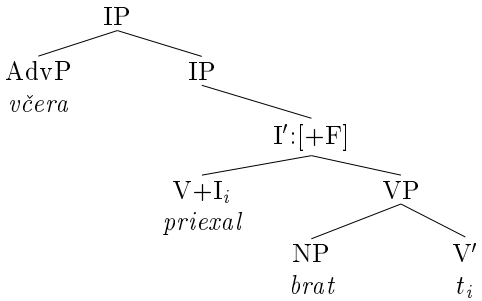
First take the analysis in which subjects are in SpecIP. Under such an analysis, the subject in (14) must be in a right-edge position, for example right-adjoined to IP, because there is no other way for it to appear after the verb (there is no position for the verb to move leftward into, except C^0, and such head-movement is not licensed in this construction, and makes incorrect predictions about topic placement). This explains the focus interpretation that the subject receives, but not that of the verb. The verb itself must be included in the intonational marking of the focus. This is a conceivable situation, but does not explain why the entire subject-verb complex could not be marked as focus intonationally with subject–verb ordering where the subject remains in SpecIP. That is, we would predict that (15) is as appropriate an answer as (14), but this is not the case. (# indicates that the sentence is grammatical, but not under the given reading.)

(15) # Včera [brat priexal].
yesterday brother arrived
'Yesterday my brother-FOC arrived-FOC.'

Now consider the situation if subjects remain in SpecVP while the verb raises to I^0. In this case, the verb-subject complex can be marked for focus intonationally because it forms a constituent. The alternate order in (15) does not arise because the only way for the subject to appear before the verb is for it to adjoin to IP where it will receive topic,

not focus, interpretation.[4] The structure I advocate for (14A2) is shown in (16), where [+F] is used to indicate the scope of focus as demarcated by the intonational contour. It is difficult to determine whether this is a purely syntactic constituent marked by a particular intonational contour or a prosodic constituent whose left-edge is constrained by the mappings between prosodic and syntactic constituents.

(16)



The same type of argumentation holds with transitive sentences in which the subject and verb are focused and the object topicalized, as in (17A). The word order in these sentences is object–verb–subject, not object–subject–verb.

(17) Q: What happened to the bird?
A: Pticu [ranil mal'čik.]
bird wounded boy
'A boy-FOC wounded-FOC the bird-TOP.'

If subjects were in SpecIP, there would be no way to explain why the order object–subject–verb is not possible. However, if the subject is in SpecVP, there is a straightforward analysis of these sentences. The object is adjoined to IP where it is interpreted as a topic. The verb and subject remain *in situ* and the domain of focus is marked intonationally by the falling tone on the subject *mal'čik*. Thus, the structure of (17) is identical to that in (16) except that the direct object *pticu* is adjoined to IP, having moved there from the underlying object position.

To summarize, if Russian is VSO, there is an explanation for why subjects which appear preverbally are always interpreted as topics and

[4] In theory, the subject could be right-adjoined to VP, although this movement is vacuous as far as this surface string is concerned. Further research needs to be done to determine whether there are other reasons to prevent such an analysis. It may be that there is no case in which focus spreads leftward onto more than one constituent. This would suggest that there is no need to permit such movement. Ideally, all and only those constituents which remain *in situ* would be discourse neutral; however, it appears that this generalization does not hold for Russian (section 5.3).

why when the entire clause or verb–subject complex is focused the subject follows the verb.

5.1.1 Topic Position

As seen in section 4.1 and 4.3.2, two types of topic are relevant for Russian: external topics and internal topics. The structural position of these topics is discussed below. I argue that external topics are external to CP, while internal topics are adjoined to IP. Although internal topics are more common than external topics, external topics are discussed first because their syntactic distribution is more readily apparent than that of internal topics.

5.1.1.1 External Topics

As seen in section 4.3.2, the topic of an utterance is usually an argument of the verb. However, Russian has a left-dislocation structure in which the topic is not an argument of the verb, although it may be coreferential with one.[5]

(18) a. [Moskva], ona gorodam mat'.
Moscow-NOM she cities mother
'Moscow-E.TOP, she is the mother of cities.'
(Gundel 1988:185)

b. [Boris(a)], ja ego ne ljublju.
Boris-NOM(-ACC) I him not like
'Boris-E.TOP, I don't like him.' (Gundel 1988:185)

c. [Milicionery], na stole ležalo dve furažki.
policemen-NOM on table lay two service caps
'Policemen-E.TOP, on the table there lay two service caps.'
(Franks and House 1982:161)

In (18), the first word forms a distinct phonological phrase. In (18a) and (18b), the topic is coreferent with a pronoun in the sentence: *ona* in (18a) and *ego* in (18b).[6] However, as seen in (18c), the external topic

[5]There are constructions in Russian which resemble the right-dislocated structures found in English. Although these form their own phonological phrase, their case marking matches that of the argument with which they are coreferent, as in (i). This difference in case marking from the left-dislocated constructions indicates a difference in structure of right- and left-dislocation constructions.

(i) A kuda ee ubrat', mašinu?
and where it-acc take, car-acc
And where should it be taken, the car? (Gundel 1988:186)

Chvany 1973:271 says of these constructions in both English and Russian: "It seems to be a correction of a sentence with an infelicitously used pronoun." and claims that they are not topicalization structures. Bolinger 1967 makes a similar claim for the English constructions.

[6]This pronoun is in turn an internal topic of the following clause.

need not be coreferential with an argument of the verb. The external topic can always appear in the nominative case, regardless of the case or grammatical function of the pronoun with which it is coreferent; if it is coreferential with an accusative direct object, the external topic can appear in the accusative, as in (18b).[7]

As expected if these phrases are topics, the dislocated noun phrase must uniquely identify an entity or set of entities.

(19) *[Mnogo televizorov], v ètom magazine ix est'.
many televisions in this store them be
'Many televisions-E.TOP, they are in this store.'
(Gundel 1988:188)

In (19), the topic phrase, *mnogo televizorov*, does not designate a unique set of entities. The same restriction holds for the English left-dislocated sentence.

Due to their distinct phonological and non-argument status, the left-dislocated topics are represented differently from internal topics. I follow Rudin 1985 for Bulgarian and Aissen 1992 for Mayan in placing these dislocated structures under a node E(xpression) which dominates CP (Banfield 1982). By placing external topics under this node, there is an explanation for why these topics never appear in embedded clauses, even in languages like Bulgarian and Russian which allow internal topicalization in embedded clauses. embedded clauses are maximally projected to CP, while the E projection is only found in root clauses. If there is no E projection in subordinate clauses, then there is no position into which an external topic could be projected.

The structure for the external topic construction in (18b) is shown in (20). In addition to the external topic, *Boris*, there are two internal

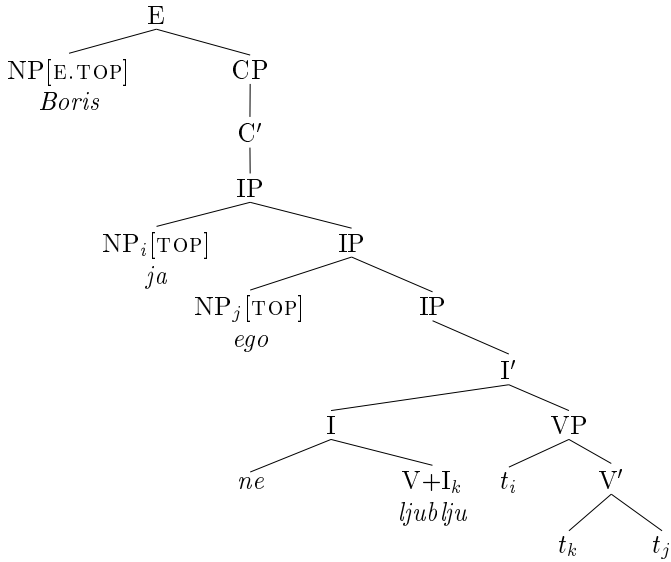
[7] There is an additional construction which can be considered with the external topics. These involve a constituent in the genitive plural whose scope is determined by an argument of the clause, as in (i).

(i) Romanov na stole bylo dva.
novels-GEN.PL on table were two
'The novels on the table were two.' (Franks and House 1982:157)

The topic *romanov* cannot have been extracted from the argument *dva*, because *dva* governs the genitive singular, not the genitive plural. Franks and House 1982 propose that the genitive forms a constituent with a covert quantifier, which accounts for the genitive case marking. Len Babby (p.c.) suggests that this quantifier phrase is assigned nominative case, thus unifying it with external topics. Franks and House point out some differences between external topics and these genitives. One is that the genitives are not followed by a pause, unlike with external topics. A second difference is that there must be a quantifying expression in the clause that refers to the genitive, i.e., there are no genitive topic constructions analogous to (18c); they suggest that the overt quantifier raises at LF, licensing the null quantifier of the genitive constituent.

topics, *ja* and *ego*, which are discussed in section 5.1.1.2. The internal topic *ego* is coreferential with the external topic.

(20)



Interestingly, the E projection does not conform to the usual X′ schema: there is no head, no specifier. There are two ways to look at this, both related to the fact that E only occurs in matrix clauses. The first possibility is that E should be an anomalous projection because it occurs in such a limited environment. Under Banfield's 1973 conception, this node dominates a number of items and constructions found only in matrix clauses, e.g., vocatives. The second possibility is to make E conform to X′ syntax so that there is a, presumably null, head E^0 which projects a specifier in which external topics and certain parentheticals and other constituents appear.[8] The head E^0 would carry some semantic features which license the appearance of these particular items in its specifier and would take a CP as its complement. This view would potentially allow an embedded EP projection. However, such constructions are not found, i.e., E^0 is not subcategorized for. A syntactic reason why this EP projection does not appear in embedded clauses would be that no head in the language takes an EP complement; so, for example, all embedded clauses are (morpho)syntactically

[8]Having this be a specifier position conceivably allows for movement into SpecEP. However, all of the items which appear in this position are base generated. If E is not part of the the X′ syntax and hence does not have a specifier, it is predicted that items will not move into this position.

required to be maximally CPs. Another reason would be that the semantics of the E^0 head and specifier are such that they are incompatible with embedded contexts. Either or both of these reasons would allow parameterization so that in some dialects or with certain subordinating verbs, EPs could appear in embedded contexts. However, although certain matrix phenomena, such as subject-aux inversion, do occur in certain dialects, there are a number of items argued by Banfield to occur under E which never occur in these embedded environments. As such, I suggest that E remain as a anomalous projection which cannot appear in embedded contexts and that embedded main clause phenomena result from some other source (see McCloskey 1992 for some discussion of these phenomena).

5.1.1.2 Internal Topics

The more common type of topic in Russian is the internal topic. As seen in section 4.3.2 and above, internal topics appear before the verb. In addition, they appear after complementizers, wh-phrases in SpecCP, and verbs which have undergone head-movement to C^0.[9] So, topics must be in a structural position lower than C^0, but higher than I^0, in which finite verbs appear. Thus, topics must be either in SpecIP or adjoined to IP. The fact that multiple topics are relatively common suggests that they should be in an adjoined position. However, Slavic languages are known to allow multiple adjunction structures within Spec positions; for example, Rudin 1988 argues for such structures in SpecCP for multiple wh-questions in certain of the Slavic languages. In this section, I argue that the adjunction to IP analysis is preferred over the branching SpecIP analysis, although evidence for this involves the analysis of focus in emotive sentences (section 4.3.4.2, 5.1.2). This adjunction position is interpreted as topic regardless of whether the sentence is emotive (having sentence stress) or non-emotive (without sentence stress).

Topicalized NPs are exemplified in (21) and (22). These preverbal NPs must be interpreted as topics unless they receive sentences stress, in which case they are focused. NPs appearing after the verb cannot receive topic interpretation.

(21) [È̀to plat'e] šila Inna.
this dress sewed Inna
'This dress-TOP, Inna sewed.'

[9] This contrasts with Bulgarian in which topics, both external and internal, can appear to the left of material in C^0 and SpecCP (section 5.2.3).

(22) [Staruju lodku] [my] prodali.
old boat we sold
'We-TOP sold the old boat-TOP.'

In (21) the object *èto plat'e* must be a topic. In (22) there are two topics, both of which appear before the verb. The relative ordering of these topics is based largely on pragmatic factors (section 4.1); basically, more recently topicalized material precedes older topicalized material.

One argument for adjunction of topics to IP, instead of to CP, is that in Russian topics appear in embedded clauses and after wh-phrases. Since the complementizers of embedded clauses and wh-phrases both occur in the projection of C^0, topic position must be below this position, so that topics occur to the right of the material in C^0 and SpecCP.

(23) Ivan skazal, čto [emu] nužna kniga.
Ivan said that he-I.OBJ need book-SUBJ
'Ivan said that he-TOP needed a book.'

(24) O čem [vy] govorite?
about what you talk
'What are you-TOP talking about?'

In (23) the topicalized pronoun *emu* appears after the complementizer *čto* and before the predicate *nužna*. Note that the English translation of (23) is misleading: the topicalized pronoun is not the subject of the predicate; the subject *kniga* remains after the verb since it is not topicalized. In this way, embedded clauses resemble their main clause counterparts. (24) exemplifies a topic within a wh-question. The topic, *vy*, appears after the wh-phrase, *o čem*, in SpecCP, but before the verb in I^0.

Additional support for adjunction to IP is provided by the fact that topics appear after verbs fronted in C^0, i.e., in *li* yes-no questions (section 3.5 and 6). The clitic *li* is a complementizer, and under certain conditions the finite verb head-adjoins to C^0. In these questions, topicalized phrases appear after the verb and complementizer complex in C^0.

(25) Čitaet li [ètot mal'čik] novuju knigu?
read Q this boy new book
'Is this boy-TOP reading a new book?'

(26) Uvidel li [tebja] professor posle uroka?
saw Q you professor after lesson
'Did the professor see you-TOP after the lesson?'

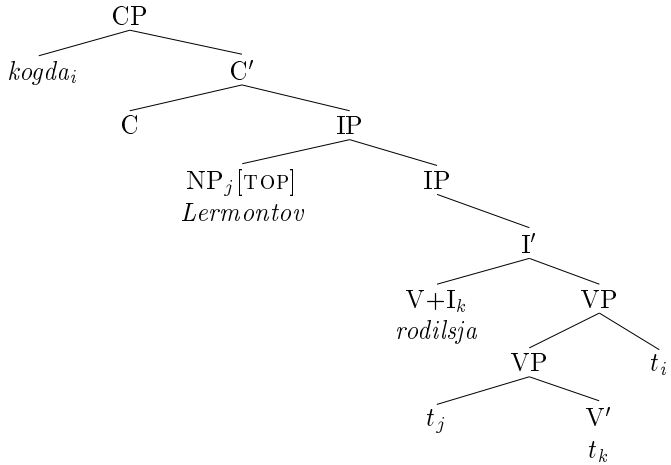
The question in (25) is ambiguous as to whether the subject *ètot mal'čik*

is topicalized or not. This is because once the verb has moved to C^0, the order subject–object is indicative both of their order *in situ* in the VP and their order when the subject has adjoined to IP where it must be interpreted as a topic. However, the order in (26) is unambiguous in that the object *tebja* must be a topic. The only way for the direct object to appear before the subject *professor* and the adverbial *posle uroka* is for it to move out of the VP and adjoin to IP where it receives topic interpretation.

The data discussed above provide evidence that topic position must be below the projection of C^0, since topics appear after material in C^0 and SpecCP, and higher than I′, since topics precede the finite verb in I^0. (28) gives the structure proposed for a sentence like (27) in which the topic *Lermontov* appears to the right of the wh-word in SpecCP.

(27) Kogda [Lermontov] rodilsja?
when Lermontov born
'When was Lermontov-TOP born?'

(28)



There is another possible analysis for topic position, namely, internal topics could be in SpecIP. This was the proposal for Yiddish made by Diesing 1988 and for Hungarian by Kiss *to appear*.[10]

An indirect piece of evidence for having topics in an adjoined position is the relative frequency of multiple topics, as in (29).

[10]Diesing made the interesting claim that SpecIP in Yiddish was both topic and subject position. When subjects move to SpecIP they can receive topic interpretation. However, whenever a non-subject appears in SpecIP it must be a topic. Thus, SpecIP is both an A and A′ position.

(29) a. [Èto plat'e] [mne] šila Inna.
this dress me sewed Inna
'Inna sewed me-TOP this dress-TOP.'

b. [Deti] [ee] prodali.
children it sold
'The children-TOP sold it-TOP.'

To maintain the SpecIP as topic analysis for Russian, either multiple adjunction structures would have to be allowed within the SpecIP position or there would have to be an additional topic position adjoined to IP. This second option is the one chosen by Kiss *to appear*. Neither of these is an insurmountable problem, but there is additional evidence that SpecIP is a focus position.

In non-emotive sentences, foci always appear in final position. However, as discussed in section 4.3.4.2 and 5.1.2, in emotive sentences the focused item receives sentence stress and can appear preverbally. The most frequent place in which these foci are found is immediately before the finite verb, following the topics. In fact, Yokoyama 1986:222 claims that the focused item in these utterances cannot precede the topicalized phrases in discourse-initial contexts. This suggests that SpecIP is a focus position. The difference between this position and that of right-edge foci is that the items in SpecIP are contrastive foci and as a result are associated with sentence stress (section 5.1.2). In contrast, right-edge foci are within the VP and have a new-information focus reading (section 5.3).

The sentences in (30) show the most common relative ordering between topics and foci in emotive sentences. The constituent on which the sentence stress falls and which is the focus of the sentence is in bold face.

(30) a. K nam [**Anna**] priexala.
to us Anna came
'Anna-FOC visited us.'

b. Ja [**k Pavlu**] prišel.
I to Pavel came
'I visited Pavel-FOC.'

c. [**K Ivanu**] priexali gosti.
to Ivan came guests
'Ivan-FOC had guests.'

In (30) the topicalized constituents precede the focus which in turn precedes the verb. When there is no topicalized constituent, as in (30c), the focus is sentence-initial but still immediately preverbal.

The position which captures this distribution most effectively is

SpecIP, a position which cross-linguistically is often associated with focus.[11] If SpecIP is a focus position and topics are adjoined to IP, then the desired ordering of topic–focus–verb falls out from the syntax. Thus, I conclude that in Russian, topics are adjoined to IP. The next section discusses the syntactic position of contrastively focused elements in more detail.

5.1.2 Focus Position: SpecIP and Sentence Stress

In Russian there are two positions associated with focus. These correspond to the canonical linear position of focus in emotive and non-emotive sentences. In emotive sentences, the focus is the constituent marked by sentence stress. The most common position for this focused constituent is immediately preverbal (section 4.3.4.2). I propose that the focus in these emotive sentences is in SpecIP. In non-emotive sentences, focused phrases occur in final position (section 4.3.4.1). The scope of the focus is marked intonationally and scrambling in the form of adjunction to VP allows the falling tone of non-emotive sentences to be placed on the appropriate constituent (section 5.3)

Emotive sentences have sentence stress which falls on the contrastively focused item, similar to how stress marks focus in English (Jackendoff 1972, Rochemont and Culicover 1990). The sentence stress can fall on any position in the sentence. However, there is a strong tendency for the stressed item to be immediately preverbal, and in discourse-initial utterances it cannot precede topicalized information. Examples of foci with sentence stress are shown below.

(31) U menja [**golova**] bolit.
at me head hurt
'I have a headache-FOC.'

(32) Nad Krakovom [**doždiček**] nakrapyval.
over Krakow rain drizzled
'It was drizzling-FOC over Krakow.' (Yokoyama 1986:192)

(33) [**Zavtra**] pozvonit muž.
tomorrow call husband
Tomorrow-FOC my husband will call.

(34) [**Ni s kem**] ne sovetovalsja.
with nobody not consult
'I didn't consult with anybody-FOC.'

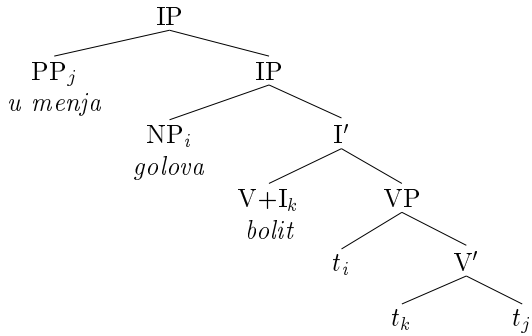
In (31)–(34) the focus is immediately preverbal and must be marked

[11] This statement depends on the position of the finite verb in the phrase structure. Focus position corresponds to the specifier of whichever projection contains the finite verb. See fn. 23 for further discussion.

with sentence stress. If the focused item is not marked with sentence stress, the sentence will be interpreted as non-emotive, and the preverbal constituents will be topicalized, not focused. In (31) and (32) the preverbal focus is preceded by the sentence topic, while in (33) and (34) there is no topic and the focus is sentence initial. In particular, note that when contrastive foci and topics co-occur, the topics precede the contrastive focus.

Thus we have the ordering topic–focus(+stress)–V–neutral. This suggests that topics be adjoined to IP, while foci are in SpecIP. The structure for (31) is shown in (35).

(35)



The topicalized PP *u menja* is adjoined to IP and as a result precedes all other constituents in the clause. The subject *golova* is focused and appears in SpecIP; the implication is that one's head, as opposed to some other part of the body, is hurting.[12] Finally, the finite verb has raised to I^0 in order to get tense and agreement features.

SpecIP licenses contrastive focus. Foci in this position can only occur with sentence stress because they must have a contrastive focus reading, which is associated with sentence stress. Although these foci can appear in other positions, the contrastive focus nature of SpecIP makes movement of foci to this position optimal. The conditions under which contrastive foci can appear elsewhere in the clause and the structure of these constructions remain to be investigated (however, see section 9.2).

5.2 Configurational Topic and Focus in Other Languages

In this chapter I have proposed that movement of constituents out of the VP is linked to their discourse function. The idea of associating

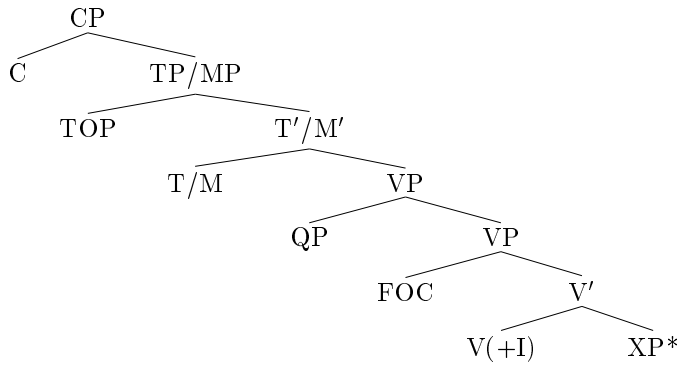
[12]With this type of expression the part of the body that hurts usually follows the verb of hurting.

certain syntactic positions with particular discourse functions has been proposed to account for the varying word orders and their interpretations in a number of languages. Here I discuss proposals made for three such languages: Hungarian (Kiss 1987a, *to appear*), Mayan (Tzotzil, Jakaltek, and Tz'utujil) (Aissen 1992), and Bulgarian (Rudin 1985, *to appear*). This is by no means an exhaustive list of languages which use the syntax to encode discourse functions as well as, or instead of, grammatical relations, nor is it a complete list even of the languages which have been examined in the linguistic literature. These three languages were chosen for a variety of reasons: Hungarian because it has been extensively analyzed and is one of the most widely recognized language of this type; Bulgarian because it, like Russian, is a Slavic language, but one which differs from Russian in interesting ways; Mayan because it is not Indo-European and because the three languages described demonstrate subtle but important differences.

5.2.1 Hungarian

Hungarian is one of the best-known examples in the current linguistic literature of a language in which topic and focus appear in different structural positions in the clauses (Brody 1990, Horvath 1986, *to appear*, Kiss 1987a, *to appear*, Marácz 1990). For the purposes of this section, I discuss the proposal in Kiss *to appear*.[13] The basic structure she proposes for Hungarian is shown in (36).

(36)



Topics appear in SpecTP; when a sentence has more than one topic, the additional topics are adjoined to TP. Quantifier phrases are adjoined

[13]Kiss *to appear* contains a very brief description of some other proposals, concentrating on how they differ from hers. One major difference is whether the arguments of the verb are sisters or are hierarchically arranged. Another is whether there is a separate projection in which topics appear.

to VP.[14] Foci appear in SpecVP. Finally, as argued extensively in Kiss 1987b, the arguments of the verb are projected in a flat structure after the verb in V′.

Before discussing the word order, intonation, and scope facts which support this structure, a few words must be said about the notion of topic in Hungarian. Kiss states that the Hungarian topic is unlike its English and German counterparts. Essentially, the VP bears a predication relation to the topic, similar to the relation in English between the VP and the subject. The sentence makes a statement about the topic, which is the subject-of-predication (section 4.1.2). Topics are either definite or generic entities which can be presupposed to exist; neither operators nor idiom chunks can be topics (Kiss *to appear*:4). Not all sentences have topics: 'when a situation is described not as a statement about an entity the existence of which is presupposed, but is conceived as a whole, as in presentative sentences, the specifier of TP/MP remains empty' (Kiss *to appear*:4).

Two examples of topics are shown in (37) (all data are from Kiss *to appear*). In both sentences, the topic is the initial element to which the VP bears a predication relation. There is no case restriction on the topic: in (37a) it is accusative and in (37b) nominative.

(37) a. [Évát] [János] várta a mozi előtt.
Eve-ACC John waited the cinema in-front-of
'Eve-TOP was waited for in front of the cinema by John-FOC.'

b. [János] [Évát] várta a mozi előtt.
'John-TOP waited for Eve-FOC in front of the cinema.'

SpecVP, the immediately preverbal position, is a focus position. For example, in (37a) *János* is focused, while in (37b) *Évát* is focused. In each case the focused constituent follows the topic. Wh-operators and negated constituents also appear in this position. If this position is not filled by a maximal projection, the preverb appears there.[15] As a result, the relative position of the verb and preverb serve as a diagnostic for whether a constituent is in SpecVP, as illustrated (38).

(38) a. [János] el ment.
John away-PREV went
'John-TOP went away.'

[14]This is a simplification of the intricacies of quantifier phrase placement. Kiss argues that some QPs are adjoined to VP, some are in SpecVP, and some remain *in situ*. Their position is predictable based on their meaning and scope. Universal quantifiers are adjoined to VP, negative quantifiers are in SpecVP, and non-universal positive quantifiers can appear in either position.

[15]Kiss refers to this as an incorporated constituent.

b. [János] ment el.
John went away-PREV
'John-FOC went away.'

In (38a) the subject *János* is the topic and hence in SpecTP. Since SpecVP is not filled, the preverb *el* appears before the verb. In contrast, in (38b) the subject is focused and appears in SpecVP, and the preverb remains after the verb because SpecVP is already filled.

Quantifiers undergo quantifier raising at S-structure and adjoin to VP. As a result, they appear after topicalized phrases in SpecTP and before material in SpecVP, as in (39). The order of the quantifiers reflects their scope. So, (39) is unambiguous in meaning.

(39) János [kétszer is] [minden tanárt] meg kérdezett.
John twice even every teacher PERF asked
'John asked every teacher twice.'

In (39), the quantifier phrases *kétszer is* and *minden tanárt* follow the topic *János* and precede SpecVP. The fact that they are not in SpecVP is supported by the presence of the preverb *meg* before, and not after, the verb.

The word order facts shown above give the order: topic–QP–focus–V–XP*. This order is captured by the structure shown at the beginning of this section where topics are in SpecTP, quantifier phrases adjoined to VP, and foci in SpecVP. The grouping of these constituents receives additional support from intonation and placement of adverbials. The preverbal quantifier phrases and foci are obligatorily stressed, with the leftmost receiving the greatest stress. In contrast, topics are usually unstressed, and when they are stressed, their stress is less than that of the leftmost quantifier phrase or focus (Kiss *to appear*:10–12). This can be accomplished if phrasal stress is assigned to the initial element of each constituent. Foci and wide scope quantifiers are obligatorily stressed. If the stress rule is applied to the VP, then the decreasing stress pattern on the quantifier phrases and foci is predicted and the topics can be excluded from this pattern.[16]

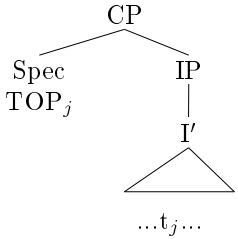
To summarize, the preverbal positions in Hungarian are associated with particular discourse functions. Topics are in SpecTP and adjoined to TP and foci are in SpecVP if they have scope over the VP. The arguments of the verb are arranged in a flat structure in V′ and appear after the verb unless they move into one of the preverbal discourse function positions.

[16]Kiss uses this as evidence against an analysis in which the topics are adjoined to the same node as foci and quantifiers.

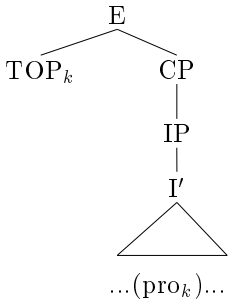
5.2.2 Mayan

Aissen 1992 argues that Tz'utujil, Tzotzil, and Jakaltek (all Mayan languages) are underlyingly verb initial, in particular VOS,[17] and that constituents appearing before the verb receive particular discourse interpretations. She argues for three distinct preverbal positions which are shown below.

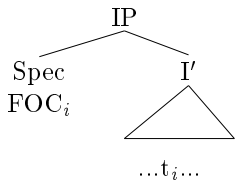
(40) a. Internal Topic:



b. External Topic:



c. Focus:



In all of these languages, the focus is in the immediate preverbal position while the topic precedes the focus, as in (41a). Aissen's arguments for placing focus in SpecIP include the focus being immediately preverbal, its appearing after the negative marker which adjoins to IP, as in (41b), and after the question marker which appears in C^0, as in (41c). (All data in this section are from Aissen 1992; square brackets indicating topic and focus constituents are mine.)

(41) a. [a ti prove tzeb-e] [sovra] ch'ak'bat
TOP DET poor girl-ENC leftovers was.given
'It was the leftovers-FOC that the poor girl-TOP was given.'
(Tzotzil)

[17] Aissen derives VOS word order by projecting SpecVP to the right, which results in subjects being the final element in the VP. The verb remains in V^0, with the object as its right sister.

b. mu [chobtik-uk] tztz'un
NEG corn-UK he.plants
'It wasn't corn-FOC that he was planting.' (Tzotzil)

c. mi [vo'ot] batz'i xapas mantal
Q you really you.do order
'Are you-FOC the one who gives all the orders?' (Tzotzil)

One item of particular interest is the difference between internal and external topics, shown in (40a) and (40b) respectively. Internal topics are in SpecCP and bind a trace, while external topics are outside the main clause structure under E (following Banfield 1973, Emonds 1985) and can be coreferent with a pronominal in the clause. So, internal topics and foci move into their respective positions, while external topics are base generated under E. Related to this, external topics can be coindexed with resumptive pronouns but need not be associated directly with an argument of the sentence; on the other hand, foci and internal topics bind a trace in the sentence.

Aissen argues that Tzotzil and Jakaltek have only external topics.[18] In Tzotzil topics are preceded by a topic marker *a*, followed by an enclitic *e*, and are usually marked with a definite determiner, indicative of their topic function. Evidence that these topics are external to CP comes from the fact that they appear before negation which is adjoined to IP, as in (42a), and the question marker which is in C^0, as in (42b).

(42) a. pero [li vo'on-e] mu xixanav
but DET I-ENC NEG I.walk
'But me-TOP, don't walk.' (Tzotzil)

b. [a li vo'ot-e] mi mu k'usi xana' un
TOP DET you-ENC Q NEG what you.know ENC
'As for you-TOP, don't you know anything?' (Tzotzil)

In addition, Tzotzil and Jakaltek topics cannot occur in subordinate clauses, as in (43) and (44). If topics were located in a position internal to the projection of C^0, it should be possible for them to appear in embedded clauses, but this is not the case.

(43) a. liyalbe li xun-e ti taxtal
he.told.me DET Xun-ENC COMP comes
li petul-e
DET Petul-ENC
'Xun told me that Petul was coming.' (Tzotzil)

[18] Tzotzil and Jakaltek appear to be very similar in their behavior. Most of the data reproduced here are from Tzotzil, but similar examples can be found in Jakaltek.

b. *liyalbe li xun-e ti [a
he.told.me DET Xun-ENC COMP TOP
li petul(-e)] taxtal(-e)
DET Petul-(ENC) comes-(ENC)
'Xun told me that Petul-TOP was coming.'

(44) a. xvinaj ti taxtal li petul-e
appears COMP comes DET Petul-ENC
'It appears that Petul is coming.' (Tzotzil)

b. *xvinaj ti [a li petul(-e)] taxtal-e
appears COMP TOP DET Petul(-ENC) comes-ENC
'It appears that Petul-TOP is coming.'

(43b) and (44b) show that the subject of an embedded clause cannot be topicalized. When the topicalized constituent is placed before the verb the sentence is ungrammatical, and there is no other position in which the topic could potentially appear. Presumably the sentence is ungrammatical even if the topicalized phrase is placed before the complementizer *ti*.

However, these data are also compatible with the topic being in SpecCP. More conclusive evidence comes from intonational phrasing and the placement of certain enclitics (see Aissen 1992:52-68 for extensive details and involved argumentation). Thus, projecting these topics external to CP explains why they cannot appear in embedded contexts and why they appear to the left of material adjoined to IP, in addition to the clitic placement facts. Finally, since external topics are base generated, they can be coreferent with pronouns, even ones in syntactic islands. All of this contrasts with foci which are moved into focus position in SpecIP from their D-structure position and as a result respect island constraints and are not coreferential with pronouns.

In contrast to Tzotzil and Jakaltek, Aissen argues that Tz'utujil is a language with internal topics. Like external topics, these topics appear to the left of focused phrases. However, these topics are not separated by a pause from the rest of the sentence and can appear in embedded contexts. The question then arises as to where these internal topics are located in the phrase structure.

Focus phrases in Tz'utujil precede the verb and follow the negative and question markers indicating that they are in SpecIP, the same position that focused material appears in in Tzotzil and Jakaltek. Topics appear before focused phrases, e.g., in (45) the topic *ja tzyaq* precedes the focused *ch'ooyaa'*.

(45) [ja tzyaq] [ch'ooyaa'] x-ee-tij-ow-i
the clothes rats ate
'The rats-FOC ate the clothes-TOP.' (Tz'utujil)

Topics precede the negative marker *ma* and the question marker *la*, as in (46). In contrast, foci follow the negative and question markers, as in (47). Aissen assumes that the negative marker is adjoined to IP while the question marker is in C^0, as in Tzotzil and Jakaltek.

(46) a. [ja ch'ooj] ma x-uu-tij ta ja kéeso
the rat NEG ate IRREAL the cheese
'The rat-TOP didn't eat the cheese.' (Tz'utujil)

b. [aa teeko] la x-uu-ch'ey aa li'p
youth Diego Q hit youth Felipe
'Did Diego hit Felipe?' (Tz'utujil)

(47) a. ma [ch'ooy] ta x-tij-ow-i ja kéeso
NEG rat IRREAL ate the cheese
'It wasn't a rat-FOC that ate the cheese.' (Tz'utujil)

b. la [aa teeko] x-ch'ey-o aa li'p
Q youth Diego hit youth Felipe
'Was it Diego-FOC who hit Felipe?' (Tz'utujil)

In contrast to the behavior of Tzotzil and Jakaltek topics, Tz'utujil topics can occur in embedded clauses after the complementizer *chi*, as in (48).

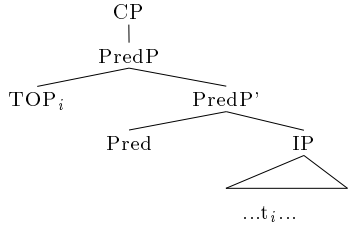
(48) aa xwaan n-0-b'ij chi [ta mari'y] ma t-r-aajo'
youth Juan says that miss Maria NEG want
'Juan says that Maria-TOP doesn't want it.' (Tz'utujil)

These data indicate that topics must be within the CP but external to IP since they follow complementizers but precede the focus which is in SpecIP. Since sentential adverbs, which are adjoined to IP, follow the topic, Aissen concludes that internal topics cannot be adjoined to IP but instead are in SpecCP. As she points out in footnote 33, one of the major flaws of this analysis is that it predicts that topics should appear to the left of complementizers, while in actual fact they appear to the right.[19]

[19]Interestingly, these internal topics do appear to the left of the question marker *la* which in Tzotzil was analyzed as being in C^0. However, Aissen points out that the question marker in Tz'utujil, unlike in Tzotzil, cannot occur in embedded clauses. This suggests that the question marker is syntactically similar to the negative marker in distribution and might be better analyzed as adjoining to IP. The generalization about topics occurring after complementizers has to do with the subordinating complementizer *chi*.

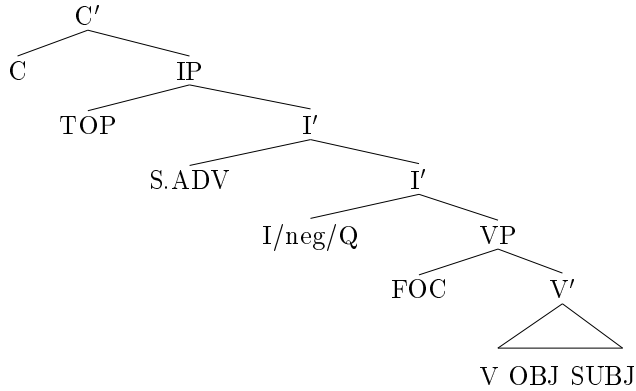
One solution is to adapt Kiss's *to appear* analysis of Hungarian to Tz'utujil.[20] The idea is that focus position would be SpecVP, instead of SpecIP, and that SpecIP would instead be a topic position. As with Aissen's 1992 account and Kiss's account of Hungarian, the verb would not raise out of the VP. The negative marker and question marker would be associated with I^0, which would correctly predict that they would occur after the topic and before the focus. These markers could be part of I^0. Sentence adverbials would be adjoined to I' to give them scope over the sentence but ensure that they appear after topics. Adjunction to non-maximal projections is in general considered undesirable, but has been suggested for a number of languages. The greatest problem with such an account is that unlike Hungarian, Tz'utujil exhibits a number of subject-object asymmetries, e.g., differences in extraction possibilities (Aissen p.c.). This suggests that the arguments of the verb are not projected in a flat structure. One way around this difficulty would be to provide a nested V' structure to posit sufficient differences between subjects and objects to account for these similarities, although this must be investigated further. The proposed structure is shown in (49).

[20] Another possibility would be to posit an additional projection between IP and CP, perhaps Pred(ication)P (see Bailyn 1991 for discussion of PredP in Russian). The specifier of this projection would be the landing site of internal topics. This would guarantee that they occurred before adverbs and particles adjoining to IP, as well as before foci in SpecIP, but that they would follow any complementizers. Thus, the structure for internal topics would be as below.



The major problem with such an analysis is what heads PredP. The verb in Tz'utujil cannot move into this position because this would erroneously predict that the verb would appear between topicalized and focused phrases.

(49)



Russian patterns more closely with Tz'utujil than with Tzotzil and Jakaltek in that it has both external topics and internal ones. The semantics of these two types of Russian topics coincides with that described by Aissen for Mayan. In addition, the topic–focus–verb order is similar in the two languages.

5.2.3 Bulgarian

Russian is not the only Slavic language in which the preverbal elements have specific discourse functions. Izvorski 1993 and Rudin 1985 analyze the syntax of preverbal constituents in Bulgarian.[21]

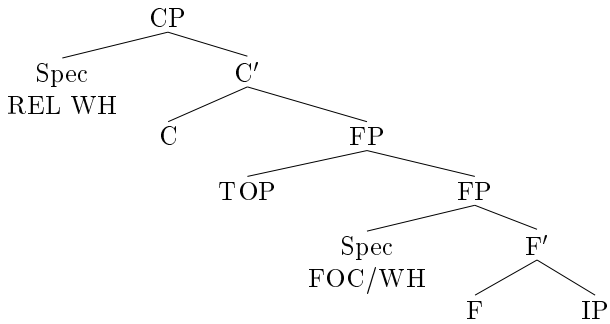
5.2.3.1 Izvorski 1993

Izvorski 1993 makes the following proposal concerning the phrase structure and encoding of discourse functions in Bulgarian. Bulgarian has a F(ocus) Projection below CP and above IP (or the expanded Infl). Focused phrases and wh-phrases appear in SpecFP, while relative wh-phrases appear in SpecCP. Topics adjoin to FP, or to IP when FP is not projected. Finally, as proposed here for Russian, the finite verb appears in I^0 at S-structure. This structure is shown in (50).[22]

[21] As for the order of constituents following the verb, Rudin 1985:38 claims that all orders are possible and commonly used. As such, she proposes that S be composed of the verb followed by XPs in a flat structure, with no particular structural prominence given to the subject. However, certain "freezing" effects that she describes suggest that subjects immediately follow the verb. For example, if the verb is followed by two NPs, either of which would be appropriate for both the subject and the object, the first is interpreted as the subject and the second as the object. The discussion in this section does not crucially depend on whether Bulgarian is underlyingly VSO or VOS or flat.

[22] In addition to topic and focus positions, Rudin 1985 argues that there are two dislocated positions, to the left and right of S′, daughter to E (Banfield 1973). That is, these dislocated positions are sister to CP, but not adjoined to it (Rudin 1993b,

(50)



Topics follow complementizers in C^0, but precede the verb and any focused constituents. For example, in (51a) the topic *na Maria* follows the complementizer *če*, but precedes the preverbal subject and the verbal complex *e posvetil*. In (51b) the clitic *gi* doubles the topic *cvetjata*.

(51) a. Kazaha mi, če [na Maria] Ivan e posvetil
told me that to Maria Ivan is dedicated
tri ot knigite si.
three of books refl
'I was told that Ivan dedicated three of his books
to Maria-TOP.' (Izvorski 1993:10)

b. Ana razbra če [cvetjata] Ivan gi e
Ana learned that the-flowers Ivan them is
kupil za Maria.
bought for Maria
'Ana learned that Ivan bought the flowers-TOP for Maria.'
(Izvorski 1993:10)

---

to appear utilizes X′ syntax to account for the data discussed in Rudin 1985). Rudin 1985:37-38 states that right-dislocated phrases share all the characteristics of left-dislocated phrases except position. In Russian, right-dislocated phrases appear to be corrections of incorrectly used pronouns; it is unclear whether the same is true of Bulgarian. Whichever the case, the right-dislocated phrases are located outside of CP.

The left-dislocated phrase is similar to the external topic in Russian (Rudin 1985:33-37; section 5.1.1.1) and Mayan (Aissen 1992; section 5.2.2). These left-dislocated phrases are always definite and refer to given information. They must be NPs, are in the nominative case, and are associated with a resumptive pronoun, as in (i). (Rudin does not make it clear whether a resumptive pronoun must be used if the left-dislocated phrase is coreferent with the subject. Given that Bulgarian has pro-drop, it is predicted that subject resumptive pronouns are not always necessary.) In addition, this phrase is followed by a pause and can only appear in matrix clauses.

(i) Ivan — nego (go) vidjah včera.
Ivan him him saw-1s yesterday
'Ivan-LD, I saw him yesterday.' (Rudin 1985:33)

However, topics precede interrogative wh-phrases, which Izvorski 1993, 1994 argues are in SpecFP. In (52) the topic *novata si kniga* precedes the wh-phrase *na kogo*.

(52) Popitah go [novata si kniga] na kogo šte posveti.
asked him the-new refl book to whom will dedicate
'I asked him to whom he would dedicate his new book-TOP.'
(Izvorski 1993:11)

As with Russian, multiple topic constructions are possible. In particular, note that subjects can appear in any order among the topics, i.e., in (53b) the topic phrases, including the subject *nie* can appear in any order.

(53) a. [S Maria] [po telefona] govorih včera.
with Maria by phone talked yesterday
'I talked with Maria-TOP by phone-TOP yesterday.'
(Izvorski 1993:17)

b. [Pismoto] [nie] na Ivan šte mu go pusnem utre.
the-letter we to Ivan will him it mail tomorrow
'We-TOP will mail the letter-TOP to Ivan tomorrow.'
(Izvorski 1993:17)

Foci appear immediately before the inflected verb, as in (54), suggesting that foci either adjoin to IP or move to SpecIP. Rudin 1993b assumes an adjunction analysis; here, in accordance with Izvorski 1993, I propose an analysis in which focus position is SpecIP.[23] In (54) the subject *Ivan* is focused and appears before the verb *e kupil*.

(54) a. Mislja če [Ivan] e kupil knigata.
think-1s that Ivan bought book-the
'I think that Ivan-FOC bought the book.' (Rudin 1985:24)

[23] Rivero 1993 has an analysis of Bulgarian phrase structure which involves a number of functional projections above the VP (section 6.5). Since the focused constituent appears to the left of the upper functional projection, excluding the projection of C^0, she would be forced to make one of two statements about focus in Bulgarian. One possibility would be to have a focus projection whose head would contain a phonetically unrealized focus feature. The other would be to state that the specifier of the highest functional projection in a given clause is the focus position. This could prove particularly amenable if the unused specifiers of the lower functional projections are not projected and, as a result, only the highest projection has a specifier into which the focused constituent can move. For example, in analyses of Bulgarian like Rivero's that posit NegP above TnsP, SpecTnsP could serve as focus position in declarative clauses, but in negative clause, i.e., when NegP is projected, SpecNegP will act as focus position. However, see King 1994b which proposes that tense and negation in Bulgarian form a complex head, as proposed here for Russian (section 3.4).

b. Dali [knigata] šte kupi?
Q book will buy
'Will she buy the book-FOC?' (Rudin 1993a:4)

c. Ne znaeh, če [knigata] šteše da kupi.
not know that book will buy
'I didn't know that she was going to buy the book-FOC.'
(Rudin 1993a:5)

Although Rudin describes constructions with two fronted focus phrases as "slightly odd", they are possible in the appropriate context (Rudin 1985:19-23). This indicates that multiple adjunction structures are allowed in SpecIP. This is not unreasonable given that Bulgarian allows multiple wh-phrases to appear in a single specifier (Rudin 1988).[24] In this case, the constraints on multiple foci in SpecIP would be semantic and pragmatic.

Next consider the behavior of interrogative wh-phrases and relative pronouns. Izvorski 1993 argues that interrogative wh-phrases are in SpecFP, while relative pronouns are in SpecCP. This accounts for the fact that relative pronouns can co-occur with focused phrases, while wh-phrases cannot since both the wh-phrase and the focused phrase would have to be in SpecFP. In addition, topics follow relative pronouns, but precede interrogative wh-phrases. In fact, no phrase can intervene between an interrogative wh-phrase and the verb, as in (55). In contrast, a topicalized or focused phrase can appear between a relative pronoun and the verb, as in (56).

(55) a. Koe pismo napisa deteto?
which letter wrote the-child
'Which letter did the child write?' (Izvorski 1993:2)

b. *Koe pismo deteto napisa?
which letter the-child wrote
(Izvorski 1993:2)

(56) Pismoto, koeto deteto napisa, e na masata.
the-letter which the-child wrote is on table
'The letter which the child wrote is on the table.'
(Izvorski 1993:2)

Thus, Izvorski concludes that while relative pro-nouns are in SpecCP, foci and interrogative wh-phrases are in SpecFP, while topics adjoin to FP/IP.

[24] SpecVP differs from SpecIP and SpecCP in this way. This correlates with the difference between A and A′ positions and with restrictions on the projection of arguments into D-structure. SpecVP is an A position which is assigned a theta-role, while SpecIP and SpecCP are A′ positions.

5.2.3.2 Rudin 1985

The topicalization data described by Rudin 1985 differ from those of Izvorski 1993, 1994. Although both agree that foci appear immediately before the verb, Rudin 1985 claims that topics appear to the left of complementizers and are distinct from left-dislocated phrases in that left-dislocated phrases cannot appear in embedded clauses and are always nominative. This claim is counter to that of Izvorski 1993, 1994 discussed in the previous section; Izvorski argues that topics appear to the right of material in CP, i.e., they are adjoined to IP. It appears that the majority of Bulgarian speakers have adjunction to IP of topics; however, certain speakers, including those described in Rudin 1985 allow topics to appear higher in the clause (Catherine Rudin, Roumi Izvorski p.c.). As such, I discuss Rudin's 1985 proposal in more detail since it poses some interesting problems for syntactic theory.

In (57a) the topic *Ivan*, which is the subject of the lower clause, precedes the complementizer *če*. The same is true of the topic *knigata* in (57b) which is the object of the lower clause. In (57c) the topic *Ivan* precedes the question word *dali* which is in C^0, and the focus *na vas* appears before the verb. Note that *da* is not a complementizer (Rudin 1985).

(57) a. Znaeh, [Ivan] če šte hodi na kino.
knew-1s Ivan that will go-3s to movies
'I knew that Ivan-TOP was going to the movies.'
(Rudin 1985:30)

b. Ne znaeh, [knigata] če šteše da kupi.
not know book that will buy
'I didn't know that she was going to buy the book-TOP.'
(Rudin 1993a:5)

c. Ivan dali na vas da se obadi?
Ivan whether to you to refl call-3s
'As for Ivan-TOP, should he call you-FOC?' (Rudin 1985:21)

Rudin (1985) proposed that these topics are adjoined to S′, where S′ is a projection containing C^0 and S. Using an X′ formalism, there are two positions before the complementizer into which maximal projections can move: SpecCP and adjoined to CP. Either possibility is compatible with the data in (58) since Bulgarian wh-phrases are in SpecFP, not SpecCP (section 6.5.2.4; Izvorski 1993).

(58) a. [Tja] dali napisa pismoto?
she whether wrote letter-the
'Did she-TOP write the letter?' (Rudin 1985:30)

b. [Ivan] na kogo dade knigite?
Ivan to whom gave books-the
'Who did Ivan-TOP give the books to?' (Rudin 1985:92)

c. [Knigata] koj e kupil?
book who has bought
'Who bought the book-TOP.' (Rudin 1993a:5)

The topics in (58) are not left-dislocated elements. Left-dislocated elements are accompanied by a resumptive pronoun, and no such pronoun is present in the questions in (58).

This leaves two options. The first is that SpecCP is a topic position and multiple topics are adjoined within the position.[25] The other is that topics are adjoined to CP, which is essentially Rudin's original analysis in which topics are adjoined to S′. This is the structure proposed in Rudin 1993, *to appear* and the one I follow here.[26]

However, there is a (Government-Binding) theoretical problem with this structure: adjunction to CP is generally not allowed. This is a result of a prohibition against adjunction to argument categories (Chomsky 1986). Although this prohibition does not affect adjunction to matrix CPs and relative clauses, it should rule out adjunction to CP in complement clauses.[27] Why then can adjunction to CP occur

[25]Evidence against topics occurring in SpecCP comes from the fact that they can occur in relative clauses, in which case the topicalized element precedes the relative pronoun. In relative clauses, the relative pronoun is in SpecCP (Izvorski 1993). In addition, these topics cannot be left dislocated phrases since such phrases cannot occur in relative clauses. (Note that Izvorski 1993 and many other speakers find these data ungrammatical; for such speakers, the topic must appear to the right of the relative pronoun. However, much of Rudin's data come from naturally occurring speech and are accepted by certain speakers.)

(i.a) lekarkata, [knigata] kojato e kupila...
doctor book who has bought
'the doctor who bought the book-TOP...' (Rudin 1993a:5)

(i.b) ženata, [naj-složnite pesni] kojato peeše...
woman-the most-complex-the songs who sang-3S
'the woman who sang the most complex songs-TOP...'
(Rudin 1985:127)

(i.c) čoveka, [Boris] kogoto vidja...
person Boris who saw
'the person who Boris saw-TOP...' (Rudin 1985:126)

(i.d) ulicata, [pazara] kŭdeto e...
street market-the where is
'the street where the market-TOP is...' (Rudin 1985:125)

[26]There is an essential difference between being adjoined to CP, as topics are, and being outside of CP, under E, as left-dislocated phrases are. If left-dislocated phrases were adjoined to CP, they would be indistinguishable from topics, which is not the case.

[27]McCloskey 1992 notes that adjunction to relative clauses is ungrammatical in

in Bulgarian complement clauses? McCloskey 1992, in his discussion of adjunction to clauses in Hiberno-English and Spanish, suggests that certain verbs have double CP complements; since the second of these CPs is not an argument of the verb, it is a possible adjunction site. He then claims that a certain semantic class of verbs whose complements are indirect questions, e.g., 'ask' or 'wonder', allow double CP complements in these languages, while others do not, i.e., 'know' under most circumstances.[28] Whenever the double CP construction appears, the complement has the semantics of a true (matrix) question.

Regardless of whether McCloskey's double CP analysis is correct for Hiberno-English and Spanish, it does not work for Bulgarian topicalization (see Izvorski 1993 for a somewhat different discussion of double CP structures in Bulgarian). The reason is that, although the distribution of double CPs is governed by c-selection, adjunction to CP in Bulgarian occurs in precisely those clauses in which the double CP construction never occurs. For example, adjunction is possible with the complement of 'know' and similar verbs, as in (59).

(59) a. Znaeh, [Ivan] če šte hodi na kino.
knew-1s Ivan that will go-3s to movies
'I knew that Ivan-TOP was going to the movies.'
(Rudin 1985:30)

b. Mislja, [tvojata sestra] če vidjah.
think-1s your sister that saw-1s
'I think that I saw your sister-TOP.' (Rudin 1985:32)

c. Kazah, [knigite] če Ivan trjabva da donese.
said-1s books-the that Ivan should to bring
'I said that Ivan should bring the books-TOP.'
(Rudin 1985:47)

So, there seems to be a contradiction. The syntactic evidence from topicalization in Bulgarian suggests that adjunction to CP is possible. Chomsky's Adjunction Prohibition prohibits adjunction to CPs in argument position. McCloskey's double CP structure, which 'saves' the Adjunction Prohibition in Hiberno-English and Spanish, is not applicable to the Bulgarian data because topics can appear with the complements of all types of verbs. One solution is to claim that the Adjunction Prohibition is not a universal, at least not in its current form, and that it does not hold in Bulgarian.

English, despite its not being a violation of the Adjunction Prohibition. He explores the idea of extending the Adjunction Prohibition to the attachment of all modifying elements, not just to adjunction derived by movement.

[28] McCloskey's divisions and discussion is more involved than this, but the basic idea is all that is necessary here.

This non-application does not mean that adjunction applies freely in Bulgarian; rather, the constraints on adjunction have some other source. Even in English, the Adjunction Prohibition alone cannot account for the distribution of adjoined constituents. For example, the placement of adverbials in English was one of the arguments for the prohibition on adjunction to arguments. Adverbs can adjoin to VPs and IPs, but not CPs. Adjunction to CP is allowed when that CP is not an argument; for example, adjunction to matrix CPs can occur. However, as McCloskey points out but does not offer a solution to, the Adjunction Prohibition predicts that adjunction of adverbs to relative clauses and adjunct clauses should be possible. However, it is not. This suggests that there are other constraints on adjunction in addition to the Adjunction Prohibition and that further work must be done on the distribution of adverbs and adjunction in general. Müller and Sternefeld 1993 provide a detailed discussion of adjunction possibilities as relates to wh-movement and scrambling, and how these can be parameterized.[29] In particular, they argue that Colloquial Russian allows scrambled constituents to left-adjoin to CP, while German only allows adjunction of scrambled constituents to VP and IP. Bulgarian is similar to Colloquial Russian in this regard.

As concerns the distribution of Bulgarian topics, in addition to adjoining to IP, as in Russian, they can appear: adjoined to matrix CPs, similar to the English adverbials; adjoined to embedded CPs, both interrogative and declarative, unlike the English adverbials; and adjoined to relative clauses, as permitted for English by the Adjunction Prohibition but not found. For now, I maintain that at least some dialects of Bulgarian allow adjunction of topics to CP. Restrictions on adjunction in general in Bulgarian, e.g., restrictions governing the placement of adverbials, must derive from other principles (Izvorski 1993). Further investigation into the distribution of these adverbials is necessary to determine these restrictions, which may depend on factors such as case marking, scoping, and differences between movement to adjoined positions and base-generation there.

To conclude this section on the phrase structure of Bulgarian, I follow Izvorski 1993, Rudin 1985, 1993b, *to appear* in arguing that Bulgarian is essentially VSO and that preverbal constituents are the result of movement into focus and topic positions. The major difference between Russian and Bulgarian in this regard is that Bulgarian allows

[29]Steve Franks (p.c.) poses the interesting question of how children learn the different systems for different languages. This issue remains to be explored, although the parameter possibility.

topics to adjoin to CP and as a result they can appear to the left of complementizers.

5.3 Intonationally Signaled Focus

Section 5.1 examined how certain phrase structure positions are associated with discourse functions in Russian. Any constituent which moves to one of these positions must have the associated discourse function or the sentence will be ungrammatical. So, phrases adjoined to IP must be topics, while those in SpecIP can only be contrastive foci. It was assumed that constituents which remain *in situ* are discourse-neutral. However, this is not always the case. The statement in (60) is true, but that in (61) is not.

(60) Any constituent which moves out of the VP will receive a particular discourse function interpretation dependent on its landing site.

(61) Any constituent which does not move out of the VP will be discourse-neutral.

This section examines sentences in which constituents within the VP receive focus interpretation, in particular that of new-information focus. These tend to be non-emotive sentences and include presentational sentences, especially when there is no topicalized element and the verb thus appears in initial position. In non-emotive sentences, the focus is always sentence-final and the falling tone associated with Type I intonation falls on the focused item (section 4.3.4).

First consider V-initial sentences, as in (62)–(64). The question-answer sequences below demonstrate the sentence-final nature of focus in non-emotive sentences. Yokoyama 1986 and Niño 1993 point out that VSO sentences of this type are often found at the beginning of stories to set the scene and introduce the participants or in newspaper headlines. Kraska 1993, in a discussion of word order in Polish, mentions that V-initial utterances are often best when followed by a sentence which continues the discourse, i.e., the V-initial utterance serves primarily to introduce the participants, and it is only in the following utterances that the events relevant to these participants are described. However, not all V-initial sentences of this type are presentational. Type I non-emotive sentences also provide new information focus in answer to questions, etc.

(62) Posadil ded repku.
planted old man turnip
'An old man planted a turnip.'

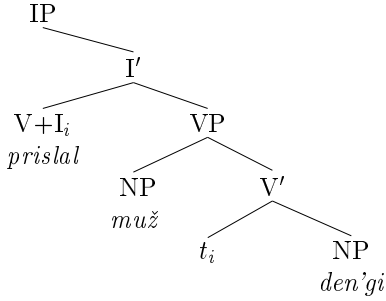
(63) Vyšla baba na pole žat'...
went woman to field reap
'A peasant woman went out into a field to reap...'
(Yokoyama 1986)

(64) Prislal muž den'gi.
sent husband money
'My husband sent (me) money.'

(62) and (63) are beginnings of stories. (64) can be analyzed as a non-discourse-initial utterance in which the goal, presumably *mne* 'me', is an understood topic.[30] However, given the proper context it too can have a more presentational reading.

The intonation of these sentences is that of Type I non-emotive sentences with a falling tone on the last syntagm. The VSO order suggests that all of the constituents have remained in the VP, except the verb which raises to I^0. The structure of (64) is shown in (65); both the subject and the object remain *in situ* within the VP.

(65)



The sentences in (62)–(64) are not composed entirely of discourse-neutral information despite all of the arguments in their base-generated positions. Instead, the entire sentence is focused. For example, in (62) and (63) not only is the action new, so are the participants, as evidenced by the use of the indefinite article in the English glosses.

This focus is not encoded by movement to a particular position. Nor can it be a direct result of the arguments remaining *in situ* since many constituents which remain *in situ* are discourse-neutral. Instead, focus is marked intonationally: not by sentence stress which yields a contrastive reading, but by the interpretation of the falling tone that appears clause-finally. In a manner similar to that of English, the falling tone signals the scope of the focus. If we think of focus as being represented by a feature [+F] which appears on a phrase structure node

[30] Even in discourse-initial utterances the pronoun *mne*, overt or understood, could be the topic since it is a deictic element (section 4.1.3).

over which it has scope, the falling tone demarcates the right edge of the constituent marked with the feature. In the case of the sentences mentioned above, the scope of the feature [+F] is the entire IP.[31]

Sentences (62)–(64) involved focusing of the entire sentence. There are also sentences in which more than one constituent is focused while the others are either discourse-neutral or topicalized. Most of these involve the focusing of a verb and its subject or a verb and its object.

(66) Q: What happened to the paintings?
A: Neskol'ko kartin [priobrel mestnyj muzej].
a few paintings acquired-FOC local museum-FOC
'The local museum acquired a few of the paintings.'

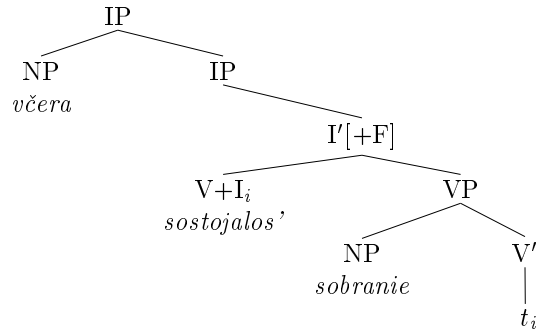
(67) Q: What did the *kolxoz* do?
A: Kolxoz [zakončil uborku urožaja].
kolxoz finished harvest crop
'The *kolxoz* finished-FOC the crop harvest-FOC.'

(68) Q: What happened yesterday?
A: Včera [sostojalos' sobranie].
yesterday took place meeting
'Yesterday a meeting-FOC took place-FOC.'

In (66)–(68) the first constituent is the topic adjoined to IP. The verb and the final constituent are both focused, as is apparent from the context provided by the question. How is this focus interpretation encoded? The verb has moved to I^0, while the focused subject or object remains within the VP. The scope of the focus is then marked by the falling tone on the final constituent. In (66)–(68) the focus feature [+F] can be thought of as taking I′ as its scope. Thus, the structure for (68) is shown in (69).

[31] The scope cannot be higher than the lowest IP because any constituents higher in the tree will necessarily be topics and thus incompatible with the focus feature. Contrastive stress, especially in more colloquial speech, can override the topic interpretation, but non-emotive Type I intonation, which is indicative of new information focus, cannot.

(69)



Each of the sentences in (66)–(68) has another possible discourse function interpretation, with the same Type I intonation. The preverbal element must be the topic. However, the verb need not be in the scope of the focus. In the contexts provided below, only the final constituent of each sentence is focused.

(70) Q: Who acquired the paintings?
A: Neskol'ko kartin priobrel [mestnyj muzej].
a few paintings acquired local museum
'The local museum-FOC acquired a few of the paintings.'

(71) Q: What did the *kolxoz* finish?
A: Kolxoz zakončil [uborku urožaja].
kolxoz finished harvest crop
'The kolxoz finished the crop harvest-FOC.'

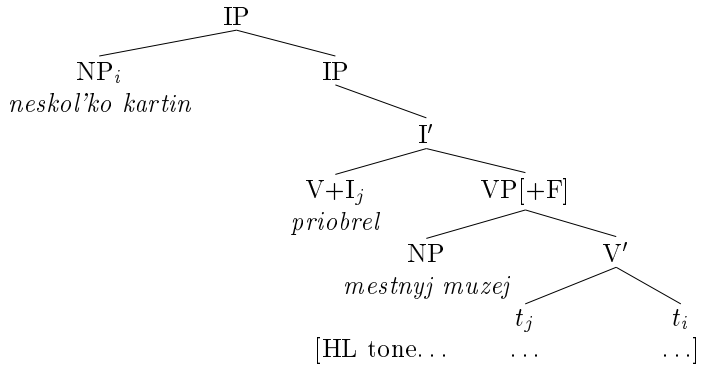
(72) Q: What took place yesterday?
A: Včera sostojalos' [sobranie].
yesterday took place meeting
'Yesterday a meeting-FOC took place.'

There are two possible structures for these sentences. One is that the focused item, e.g., the subject *sobranie* in (72), is adjoined to VP. The other is that the focused constituent is *in situ* in the VP and that it is the falling tone alone which encodes the focus interpretation. Instead of the focus feature [+F] marking the scope of focus as I′, it marks only the focused NP as its scope.

(73) shows the structure of (70). Since the the subject of (70) *mestnyj muzej* is focused, this forces, minimally, the scope of the focus feature to be the entire VP.[32]

[32]Consider the traces in structures like (73). The trace of the topicalized object *neskol'ko kartin* and of the raised verb *priobrel* fall within the scope of the focus feature since the falling tone is assigned to the right edge of the clause. The traces

(73)



Adjunction to VP is sometimes necessary in order for the correct constituents to be in the scope of the focus. Their focus interpretation is not the result of adjunction to VP being a focus position; instead, the focus interpretation results from the right-adjoined constituent being clause final and thus falling within the scope of the intonationally signaled focus. Consider (74).

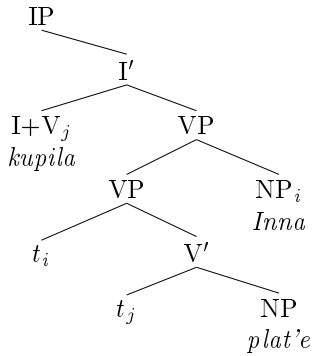
(74) Q: Who bought a dress?
A: Kupila plat'e [Inna].
bought dress Inna
'Inna-FOC bought a dress.'

In (74) the subject is focused and the object is discourse-neutral. The order of the subject and object must be that in (74); the focused subject cannot precede the discourse-neutral object. If the subject is in SpecVP and the object sister to V^0, the only way to obtain the order in (74) is for the subject to move out of SpecVP. This suggests right-adjunction to VP. The structure proposed for (74) is shown in (75).[33]

falling within the scope of the HL tone do not result in the antecedent of that trace being interpreted as focused. Note that although adjunction of the subject to VP can 'save' the traces in (70), it cannot do so in all cases. Consider (70), only with both the verb and the subject as focus, as in (66). (66) has the phrase structure in (73). However, the [+F] feature is on I′ since the finite verb *priobrel* is within the scope of the focus. Even if the subject *mestnyj muzej* right-adjoined to VP, the trace of the object *neskol'ko kartin* would be in the scope of the [+F] feature. As such, I conclude that traces are irrelevant for the interpretation of new-information focus; that is, an argument whose trace is in the scope of [+F] is not necessarily focused.

[33]One possibility would be that in these structures a specifier position is projected to the right, providing a right-edge position. First consider the projection of SpecIP to the right. Neither the focus interpretation nor the intonation of the right-edge constituents is consistent with this analysis, i.e., they do not pattern with contrastive foci (section 5.1.2). Next consider projecting SpecVP to the right. In this case, only subjects would appear in this position, and only when they are new-

(75)



In conclusion, in non-emotive sentences, the final falling tone marks the right edge of the focused constituent. No unique phrase structure position was proposed for these foci, i.e., *in situ* constituents may be within the domain of the new-information focus. All of the material in the focused constituent, excepting the verb which has raised to I^0, is within the VP: the arguments are either in their base generated position or right-adjoined to VP. right-adjunction to VP is necessary when the subject is the only focus and must appear in final position after the verb and the object; this right-adjunction may extend to other arguments. Since the falling tone marks only the right edge of the focus constituent, a sentence may be ambiguous as to the extent of the focus constituent, i.e., the context determines the left edge of the focus.

5.4 SVO vs. VSO

Why has it usually been assumed that SpecIP is subject position in Russian despite the evidence to the contrary presented here? I know of no accounts, among those which makes claims on this matter, which analyze Russian as other than having a preverbal subject position (however, see Rudin 1985 on Bulgarian (section 5.2.3)). It is easy to see why one might be led to the assumption that Russian has a preverbal subject position since in many clauses this is the order found. In addition, preverbal subjects frequently do not seem to be as markedly topicalized as other preverbal elements (Yokoyama 1986:270-296).[34]

There at first seems to be a problem in that much of the data in-

information focus. It would probably be possible to make such an analysis work. However, since other material can adjoin to VP as new-information focus, such a stipulation seems unwarranted.

[34]Yokoyama's discussion of grammatical relations is actually one of thematic roles; the relations she speaks of have to do, not with subjects and objects, but with more fundamental roles, which she likens to case roles in the Fillmorean sense. In particular, she states that word order is not to be formulated in terms of surface

dicates that SpecVP is subject position, yet other data suggests that subjects are preverbal. The answer to this lies in the fact that sentences usually have a topic, and that subjects are among the types of arguments commonly chosen as topics. In particular, evidence from word order in impersonal expressions and constructions with dative experiencers suggests that it is not grammatical subjects which tend to appear preverbally in topic position, but the thematically highest argument of a predicate.[35]

Since the behavior of the thematically highest argument of impersonal expressions plays an important role in the discussion of preverbal elements, I provide some examples and further discussion. Here I use impersonal expression to mean predicates which have no (nominative) grammatical subject, as in (76).[36]

(76) a. [Mne] bylo xolodno.
me-DAT was cold
'I-TOP was cold.'

b. [Innu] tošnilo.
Inna-ACC sick
'Inna-TOP was sick.'

In (76) the thematically highest argument appears before the verb. For example, in (76b) the accusative experiencer *Innu* appears before the verb *tošnilo*. This is the usual or 'unmarked' order for these sentences.[37] The order found with the predicates in (76) is due to the fact that sentences frequently have topics, and there is only one argument available for topicalization.

In constructions in which the grammatical subject is not the the-

grammatical notions (Yokoyama 1986:274). The importance of this generalization is discussed below.

[35]Bailyn 1991 makes a similar claim in a more indirect fashion. In his discussion of case assignment in Russian, he suggests that there is a position in the syntax, SpecPredP, into which the highest argument must move. This is not a subject position since dative experiencers frequently occur there. One trouble with Bailyn's analysis is that certain sentences do not appear to fill this position, in particular the types of sentences used to argue that SpecIP is not subject position in Russian. One of these types, presentational sentences, may have an explanation if presentational sentences are not predicated of anything, their only role being to introduce new entities into the discourse.

[36]I leave aside the possibility of null expletives (see Babby 1989 who argues against null expletives); since null expletives have neither phonetic nor thematic value they do not interact with the surface word order.

[37]A similar pattern is found when transitive verbs are used with null third person plural subjects. Although the accusative object in these constructions is not a subject, it often appears preverbally, patterning similarly to subjects of passive sentences.

matically highest argument, e.g., dative experiencer verbs, the thematically highest argument frequently appears before the verb, as in (77).[38]

(77) a. [Ej] očen' nravilas' èta kniga.
her-DAT very like that book-NOM
'She-TOP really liked that book.'

b. [Mne] nužen novyj slovar'.
me-DAT need new dictionary-NOM
'I-TOP need a new dictionary.'

In (77a) the theme *èta kniga* is the grammatical subject, but the experiencer *ej* is thematically higher than the theme. This experiencer appears before the verb, while the grammatical subject appears after it. The same holds true of the predicate *nužen* in (77b).

What does this tell us? The common preverbal subjects are not a result of a privileged preverbal syntactic position for subjects, e.g., SpecIP. Instead, subjects tend to appear preverbally because thematically highest arguments tend to appear preverbally, as was seen with the impersonal expressions in (76) and the dative experiencer verbs in (77). There is only an indirect syntactic explanation for this phenomenon, i.e., it is not the result of a 'subject' position which must be filled by some argument.[39] Sentences are generally predicated of some entity; that is, there is usually a subject-of-predication for the sentence (Kiss 1993). In Russian, this subject-of-predication is found in preverbal position since it is subsumed under the notion of topic employed here. However, there is no syntactic requirement that some element move into topic position. As for the observation that subjects often appear preverbally, this is a result of the fact that for pragmatic reasons, thematically prominent arguments are excellent candidates for topics and as a result often appear preverbally.

In this chapter, I argued that Russian is not an SVO language in that there is not a preverbal position in which the subject always

[38] The existence of constructions in which the thematically highest argument is not the subject at first seems to violate the UTAH (Baker 1988, Speas 1990) which states that relative prominence in the Thematic Hierarchy must correspond to relative prominence in the syntax. However, in psych verb constructions with theme subjects and experiencer objects, which are a seeming violation of the UTAH, the theme is associated with Causation, resulting in its being more prominent in Argument Structure (Grimshaw 1990, Jackendoff 1990, Pesetsky 1992). Truly impersonal expressions do not pose a problem since they have no thematic subject argument.

[39] Here I differ from Bailyn 1991 who posits a projection whose Specifier is always filled with the thematically highest argument (fn. 35).

appears. Instead, the subject remains in SpecVP, and the object as sister to the verb, unless they move into discourse function positions. In particular, topics appear left-adjoined to IP, while SpecIP licenses contrastive foci. Since constituents must adjoin to IP to receive topic interpretation, they appear clause-initially. The assignment of clause final new-information focus differs from topic and contrastive focus in that it is not licensed by a particular phrase structure position. Instead, it is the result of the interaction of the intonation with the phrase structure. The final falling tone of neutral intonation marks the right edge of the focused constituent. This constituent can be of any size as long as its focus interpretation does not conflict with the discourse functions of the items it contains, e.g., topics cannot appear in this constituent. Finally, there is an external topic position outside the CP which is a type of left-dislocated constituent. Unlike the other topics and foci, external topics are base generated and are not arguments of the verb, as can be determined by their case marking and the fact that they can, but need not, be coreferent with arguments of the verb. Since external topics are a main clause phenomenon, they are generated under a node E which characterizes main clauses.

The discussion of the encoding of topic and focus in Russian included a comparison with other languages which structurally encode discourse functions, namely Hungarian, Mayan, and Bulgarian. Despite the varying syntactic structures of these languages, the association of phrase structure positions with particular discourse functions explains the link between word order and discourse function interpretation, employing structures already available to the language. As a result, a tighter link is established between surface word order and syntactic structure, one which helps to explain how discourse functions are encoded by the language.

6

Focus in Yes-No Questions

This chapter examines the syntactic distribution of focused constituents in Russian yes-no questions formed with the clitic *li*.[1] In these questions, the questioned element is focused and appears in left-most position.[2] The initial element can be either a maximal projection or the finite verb. When it is a maximal projection, that maximal projection is focused and questioned. When the verb is in initial position, the reading is that of questioning the entire clause. These are shown in (1).

(1) a. [Knigu] li Anna pročitala?
book Q Anna read
'Did Anna read a book-FOC?'

b. [Pročitala] li Anna knigu?
read Q Anna book
'Did Anna read a book?'

In (1a) the direct object *knigu* appears in initial position, followed by the clitic. It is the focus of the question. That is, the speaker is asking about the identity of what was read, and it is presupposed that Anna read something.[3] The reading is similar to that which arises when 'book' is stressed in the corresponding English question. In contrast, in (1b) the verb appears in initial position followed by *li*, and the entire clause is questioned. That is, the question is asking about whether

[1]This chapter appeared in slightly modified form in King 1994a. The original version of this chapter included a section on Russian *èto* 'clefts'; having found no insightful analysis of this construction, it is not included here. A description of the problem can be found in King 1993b, chapter 5.

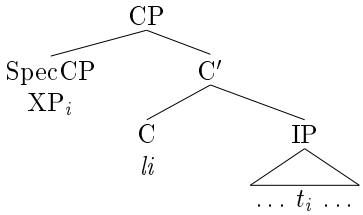
[2]One of the most complete compilations of *li* questions is in Restan 1972 which contains examples from literature and non-fiction divided into categories according to the syntax of the questions. Readers who would like further examples should consult this work.

[3]See Jackendoff 1972 for discussion of the interaction of focus and presupposition.

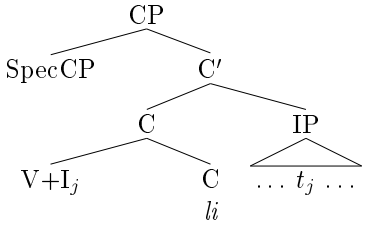
Anna's reading of a book took place and carries no presupposition about the existence of this event. This reading is similar to the neutral reading of the corresponding English question.

I argue that these two types of *li* questions have different syntactic structures, shown in (2). In both, *li* is in C^0 (section 6.1). However, the structure in (2a) contains a focused maximal projection which is in SpecCP, while that in (2b) involves a fronted finite verb which is adjoined to C^0 (section 6.2). The assignment of an inherent focus feature by *li* results in the observed distribution of focused elements (section 6.3).

(2) a.



b.



This chapter is organized as follows. First, I argue that *li* is an interrogative complementizer which can license subordinate clauses, i.e., embedded questions. The second section is concerned with the syntactic position of the pre-*li* constituents. The third section discusses the focus reading that the pre-*li* constituents receive. The final section includes a discussion of *li* constructions in Serbo-Croatian and Bulgarian.

6.1 What is *li*?

Before discussing the structure of *li* questions, the behavior of Russian subordinate clauses is examined since it sheds light on the syntactic category of the clitic. All finite subordinate clauses must have C^0 or SpecCP filled. Unlike in English, complementizers cannot be deleted.

(3) Ja znaju, [$_{CP}$ [$_{C}$ *(čto)] [$_{IP}$ on priexal]].
I know that he arrived
'I know (that) he arrived.'

(4) a. Ja ne znaju, [$_{CP}$ [$_{SpecCP}$ kogda] [$_{IP}$ on priexal]].
I not know when he came
'I don't know when he came.'

b. *Ja ne znaju, [$_{CP}$ [$_{IP}$ on priexal kogda]].
I not know he came when

Sentence (3) demonstrates that complementizers in finite subordinate clauses are obligatory. When *čto* is deleted, the sentence is ungrammatical. Sentence (4) is an example of an embedded question. The subordinate clause begins with the wh-word in SpecCP; if the wh-word is elsewhere in the clause, it is ungrammatical.

The next question to ask is how the embedding of yes-no questions works. There are two ways to ask matrix yes-no questions. The most common is to change the intonation of the corresponding declarative sentence, as in (5). The second is to use the clitic *li*, as in (6).

(5) On živet zdes'?
he live here
'Does he live here?' (with appropriate intonation)

(6) Živet li on zdes'?
live Q he here
'Does he live here?'

However, in embedded yes-no questions the clitic must be present.

(7) a. Ja ne znaju, živet li on zdes'.
I not know live Q he here
'I don't know whether he lives here.'

b. *Ja ne znaju, živet on zdes'.
I not know live he here

c. *Ja ne znaju, on živet zdes'.
I not know he live here

Sentence (7a) is a well-formed embedded yes-no question. The verb is in initial position in the clause, followed by the clitic *li*. However, if the clitic does not appear in the embedded clause, the sentence is ungrammatical, regardless of word order or intonation, as in (7b) and (7c). Since it licenses finite complement clauses, I assume that the clitic must be in either C^0 or SpecCP.

In particular, I propose that *li* is an interrogative complementizer, similar to English 'whether'; however, *li* can appear in matrix clauses. A similar claim is made for Serbo-Croatian *li* by Rivero 1993 and for Bulgarian *li* by Rivero 1993 and Rudin 1993 (section 6.5). The clitic is essentially a phonological realization of the abstract feature used to indicate that a clause is a question (see Nishigauchi 1990 on question

marking in Japanese). In particular, *li* is indicative of yes-no questions and as a result is incompatible with wh-words, as in (8).[4]

(8) *Čto li ona delaet?
what Q she do

The surface position of the clitic is dictated by the phonology. The clitic always appears after the first phonological word (see Nespor and Vogel 1986 on the notion phonological word). This cannot be seen when a verb is fronted because verbs comprise a single phonological word; as such, the first phonological word corresponds exactly to the first constituent. However, when a maximal projection comprising more than one phonological word is fronted, the clitic must appear after the first phonological word, as in (9a). It cannot appear after the entire fronted constituent, as in (9b). (The relevant phonological words are labeled with a ϕ.)

(9) a. [[Na ètom]$_\phi$ li [zavode]$_\phi$]$_{PP}$ on rabotaet?
at this Q factory he works
'Does he work at this factory-FOC?'

b. *[[Na ètom]$_\phi$ [zavode]$_\phi$]$_{PP}$ li on rabotaet?
at this factory Q he works

In (9) the PP *na ètom zavode* has been fronted. This PP is composed of two phonological words, *na ètom* and *zavode* (*na* is a proclitic). The clitic *li* must appear after the first phonological word, i.e., after *na ètom*; it cannot appear after the entire fronted PP or elsewhere in the sentence.

6.2 Location of the Focused Constituents

6.2.1 Fronted XPs

If the clitic *li* is in C^0, the question arises as to the location of the focused items, i.e., the initial verbs and maximal projections. Consider first the sentences in which a maximal projection is focused, as in (10).

(10) a. [$_{NP}$ Vodku] li ona kupila?
vodka Q she bought
'Did she buy vodka-FOC?'

[4]This contrasts with Bulgarian in which *li* is a general interrogative (and focus) marker and can appear in wh-questions, as in (i). *Li* focuses the wh-word to which it cliticizes; note the English translation in (i).

(i) Kakvo li nameri?
what Q found
'What on earth did s/he find?'

b. [$_{PP}$ V magazin] li pošel Ivan?
to store Q go Ivan
'Did Ivan go to the store-FOC?'

c. [$_{AP}$ Doroguju] li ona kupila knigu?
expensive Q she bought book
'Did she buy an expensive-FOC book?'

In (10a), the direct object NP *vodku* appears in initial position. In (10b), a PP *v magazin* is in this position, while in (10c) an AP *doroguju* has been fronted.[5] In each case, the fronted constituent is the focus of the question, while the remainder of the question is presupposed. For example, in (10b) it is presupposed that Ivan has gone somewhere; the speaker is asking whether it is to the store that he has gone.

These focused elements appear in initial position to the left of any topicalized constituents, as in (11).

(11) a. [Ivan] li ètu poèmu čital?
Ivan Q this poem read
'Did Ivan-FOC read this poem-TOP?'

b. [Kniga] li tebe nužna?
book Q you need
'Did you-TOP need a book-FOC?'

In (11a) the object *ètu poèmu* is topicalized and appears before the verb, following the focused subject *Ivan* and the clitic *li* in C^0. (11b) is a similar example, only the subject *kniga* is focused and appears before the clitic, while the dative experiencer *tebe* is topicalized.

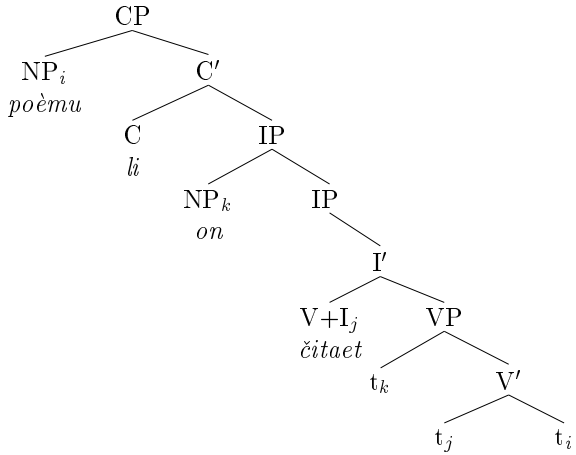
The focused phrases must move either to a specifier or adjunction position since these are appropriate landing sites for maximal projections (in contrast, heads move only to other head positions (section 6.2.2)). I propose that the focused constituent moves to SpecCP (see section 6.3.1 for how this movement is licensed). This correctly places it before the preverbal topics, which are adjoined to IP. In addition, movement to a specifier position explains why there is only one fronted maximal projection since there is only one SpecCP position; if the maximal projections moved to an adjoined position, e.g., adjoined to CP, we would expect multiple frontings to be possible. Finally, since these maximal projections are in SpecCP, Spec-head agreement can occur between *li* in the head C^0 and the maximal projection in SpecCP (sec-

[5] In (10c) the AP *doroguju* has moved out of its NP into initial position. This movement is permissible because in Russian, APs can more generally move out of NPs. What is important is that the moved phrase is a maximal projection (see example 14).

tion 6.3.1). The structure of a yes-no question like (12) is shown in (13).

(12) [Poèmu] li on čitaet?
poem Q he reads
'Is he reading a poem-FOC?'

(13)



In (13) the direct object *poèmu* has moved to SpecCP and thus appears in initial position, followed by the clitic in C^0, which encliticizes to the item in SpecCP. The verb moves to I^0, where it remains. The verb cannot have moved into C^0 because it appears after, not before, the topicalized subject *on* which has adjoined to IP.

6.2.2 Fronted Verbs

Now consider the constructions with fronted verbs. The verb is not a maximal projection and is thus predicted not to appear in SpecCP; so, *li* questions with fronted verbs cannot have the same account as the sentences with fronted maximal projections. I propose that material in I^0 can undergo head-movement to C^0, where it head-adjoins to C^0 and host the clitic *li*. This movement differs from that described in section 6.2.1 in that it involves syntactic heads; unlike maximal projections, which move to specifier and adjunction positions, heads move to other head positions (Baker 1988; Travis 1984; Grimshaw 1991).

6.2.2.1 Verb-Movement

As was seen in section 6.2.1, several different types of maximal projections can move to SpecCP in the structures discussed above (e.g., NPs, PPs, and APs; subjects, objects, and adjuncts). However, no non-maximal projection, other than material in I^0, can appear in ini-

tial position. For example, a noun cannot be questioned out of its NP, as in (14a). Instead the entire NP must be fronted, as in (14b).

(14) a. *[$_N$ Knigu$_i$] li on čitaet [$_{NP}$ interesnuju t_i
book-ACC Q he reads interesting-ACC
Ivana]?
Ivan-GEN

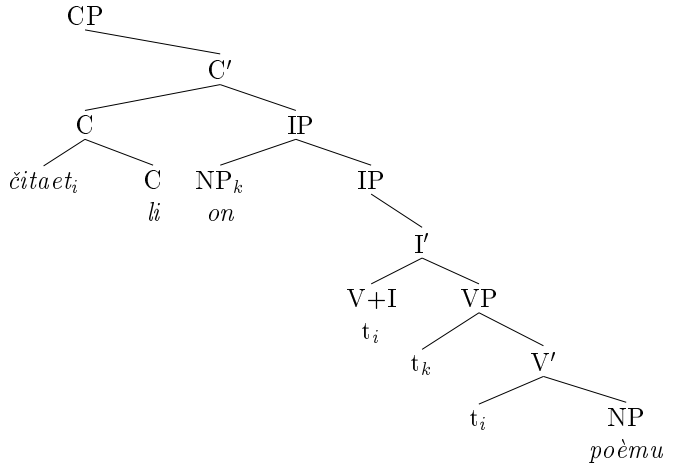
b. [$_{NP}$ Interesnuju li knigu Ivana] on
interesting-ACC Q book-ACC Ivan-GEN he
čitaet?
reads
'Is he reading an interesting book-FOC of Ivan's?'

The fact that only heads in I^0 can move to C^0 is a result of more general constraints on head-movement. In particular, the Head-Movement Constraint (HMC) restricts movement of heads to head positions which properly govern them (Travis 1984). For example, a V^0 can move to I^0 and then to C^0, but the P^0 head of a PP cannot move to C^0. This is what rules out the movement in (14a) since C^0 does not properly govern the N^0 head of the NP.[6]

The structure of a sentence like (15) is shown in (16).

(15) [Čitaet] li on poèmu?
read Q he poem
'Is he reading the poem?'

(16)



[6]Following Grimshaw 1991, heads move along extended projections. In this case, C^0 and I^0 are of the same basic type of functional, verbal projection, and so they form an extended projection. As a result, movement of the verb from I^0 to C^0 is possible (section 10.3).

In (16) the verb *čitaet* first moves to I^0, as it does in all finite clauses. From I^0 it undergoes head-movement to C^0. Since C^0 is already filled by *li*, the verb head-adjoins to it. Note that the subject NP *on* is adjoined to IP, i.e., it is in topic position; it can remain in SpecVP, if it is not interpreted as a topic.

6.2.2.2 Restriction to I^0

Only material in I^0 is eligible for head-movement to C^0 in the *li* construction. That is, infinitives in V^0 cannot undergo head-movement to C^0. Consider the imperfective future in (17).

(17) a. [Budet] li on žit' v Moskve?
will Q he live-INF in Moscow
'Will he live in Moscow?'

b. *[Žit'] li on budet v Moskve?
live-INF Q he will in Moscow

Imperfective futures comprise an inflected auxiliary in I^0 and an infinitive in V^0 (chapter 3; section 10.1.2). In (17a), the auxiliary *budet* is in C^0, while the infinitive remains in V^0, and the sentence is grammatical. However, in (17b) the infinitive *žit'* has moved to C^0, and the structure is ungrammatical.[7]

6.2.2.3 Negated questions

Further evidence for the fact that it is material in I^0 that moves to C^0 comes from the structure of negated *li* questions. In these constructions, the negative marker *ne* must move with the tensed verb to C^0 where they appear before *li*, as in (18).[8]

(18) a. Oni sprosili, [ne videli] li my Ivana včera večerom.
they asked not see Q we Ivan yesterday evening
'They asked if we hadn't seen Ivan yesterday evening.'

[7] Catherine Chvany (p.c.) points out that certain infinitives are relatively acceptable in this construction, as in (i).

(i) ?Ne lodyrničat' li on budet v Moskve?
not loaf around-INF Q he will in Moscow
'Won't he loaf around in Moscow?'

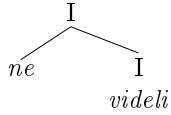
For speakers who find such constructions acceptable, the analysis presented here can be extended in one of two ways, the first of which is quite tenable on the basis of certain scrambling facts. First, these speakers might allow movement of infinitives in V^0 to I^0, from whence they can then move to C^0 in a manner similar to that of predicate adverbs (section 6.2.2.4). Secondly, these speakers may allow Long-Head Movement of the infinitive in V^0 directly to C^0. See Rivero 1991 for discussion of Long-Head Movement in Slavic; the potential parameterization of Long-Head Movement remains an issue for further research.

[8] The negative marker *ne* is itself a proclitic. As such, it forms a single phonological word with the verb, which in turn hosts the complementizer *li*.

b. [Ne zastupjatsja] li za menja babuška ili tetuška?
not protect Q for me grandmother or aunt
'Wouldn't Grandmother or aunt speak up for me?'
(Yokoyama 1986:240)

This movement is predicted if the negative marker and the tensed verb are both in I^0, as in the configuration in (19).[9]

(19)



In (19) both the negative marker *ne* and the finite verb *videli* are in I^0. Consequently, they act as a unit. When I^0 undergoes head-movement to C^0, the entire I^0 complex moves. It is impossible to move just the finite verb, as shown in (20).

(20) *Oni sprosili, [videli] li my [ne] Ivana
they asked see Q we not Ivan
včera večerom.
yesterday evening

In (20) the finite verb *videli* has been moved to C^0, leaving the negative marker *ne* in I^0. The resulting structure is ungrammatical, in contrast to (18a) in which both the finite verb and the negative marker move as a unit.

6.2.2.4 Predicate Adverbs

A third piece of evidence that material in I^0 can undergo head-movement to C^0 in the *li* construction comes from predicate adverbs, like *možno* 'can'.[10] Predicate adverbs are predicates whose tense is marked by an auxiliary and whose experiencer appears in the dative case. I analyze the auxiliary as being in I^0 (section 3.5.2; Schoorlemmer 1991), while the predicate adverb is in V^0.[11] Pre-auxiliary dative experiencers are in topic position, adjoined to IP. However, unlike the

[9]The structure in (19) is an adjunction structure; nothing hinges on this choice of structure. The important point is that the negative marker and the finite verb form a single head which is an inseparable unit (section 3.4).

[10]Predicate adverbs are so-named because they resemble adverbs in form; however, they function as the predicate of their clause.

[11]Schoorlemmer 1991 analyzes these and several other constructions as involving a Modal Projection above TP, in place of AgrP, since these modals characteristically lack agreement. This analysis predicts the opposite 'default' order of auxiliary and predicate adverb from the one here. However, since predicate adverb constructions allow either the predicate adverb or the auxiliary to move to C^0 in *li* questions and since even in declaratives both orders are possible, her analysis must take this into account. In more recent work, Schoorlemmer (p.c.) provides an analysis similar to the one presented here where predicate adverbs are categorically verbs. Her

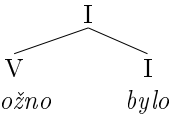
infinitives in imperfective futures, predicate adverbs can head-adjoin to I^0, so that they precede the auxiliary, as in (21).

(21) a. Emu [bylo možno] prijti segodnja.
he-DAT was can-PR.ADV come-INF today
'He could come today.'

b. Emu [možno bylo] prijti segodnja.
he-DAT can-PR.ADV was come-INF today
'He could come today.'

In (21a) the auxiliary precedes the predicate adverb, while in (21b) the predicate adverb *možno* has head-adjoined to I^0 and precedes the auxiliary *bylo*. Most predicate adverbs have a 'preferred' order with respect to the auxiliary, although both orders are usually possible. I assume that the order in (21b) is the result of the predicate adverb head-adjoining to I^0, shown in (22).[12]

(22)



Consider the behavior of predicate adverbs in the *li* construction. As expected, the auxiliary in I^0 can move to C^0, as in (23). However, since the predicate adverb can adjoin to I^0, as in (22), we predict that it can then move to C^0 in the *li* construction.[13] This prediction is borne out, as seen in (24). In fact, both the predicate adverb and the auxiliary can appear in C^0, although this movement is difficult to

account differs from the one here in that all predicate adverbs raise to I^0, unifying their behavior with that of finite verbs, although they differ morphologically from verbs in that the predicate adverbs require an auxiliary to carry tense.

[12]Unfortunately, at this time there is no explanation as to why predicate adverbs freely allow this movement, while infinitives do not (correspondingly, predicate adverbs can move to C^0 in *li* questions, while infinitives cannot). Perhaps an account like Schoorlemmer's, in which predicate adverbs occupy a different structural position from infinitives, can provide a motivation for this difference. Relatedly, one reviewer suggested that predicate adverbs may be being reanalyzed as auxiliaries.

[13]As sketched here, the structure of the negative+finite-verb complex in (19) is superficially similar to that of the predicate adverb adjoined to this complex. However, the negative complex forms an inviolable unit, i.e., the negative marker plus the finite verb comprise a single head, while the adjoined predicate adverb does not. Since the negative and the finite verb are segments of a single head, they undergo head-movement as a unit. In contrast, the predicate adverb is only adjoined to the I^0 head and can thus move independently of it, although the constituency shown in (22) allows for head-movement of the entire adjoined I^0 complex. Technically, movement of the predicate adverb out of I^0, stranding the auxiliary, is a violation of the Head-Movement Constraint, even though the predicate adverb and auxiliary do not form a morphological unit. Hopefully, a better understanding of the predicate adverb construction as a whole will elucidate this problem.

see since the clitic *li* appears after the first phonological word, as in (25) or (26).[14] The fact that the topicalized experiencer, which like all topics is adjoined to IP, appears after both the predicate adverb and the auxiliary indicates that they both must have moved.[15]

(23) [Bylo] li Ivanu trudno napisat' ètu statju?
was Q Ivan-DAT difficult-PR.ADV write this article
'Was it difficult for Ivan-TOP to write this article?'

(24) [Trudno] li Ivanu bylo napisat' ètu statju?
difficult-PR.ADV Q Ivan-DAT was write this article
'Was it difficult for Ivan-TOP to write this article?'

(25) [Legko li bylo] Inne napisat' ètu statju?
easy-PR.ADV Q was Inna-DAT write this article
'Was it easy for Inna to write this article?'

(26) [Bylo li legko] Inne napisat' ètu statju?
was Q easy-PR.ADV Inna-DAT write this article
'Was it easy for Inna to write this article?'

Thus, in *li* yes-no questions either a maximal projection moves into SpecCP or the material in I^0 undergoes head-movement to C^0, where *li* is. *Li* then cliticizes onto this constituent. Maximal projections from a variety of syntactic positions and syntactic categories can move to SpecCP in this construction. However, only heads which can appear in I^0 are eligible for head-movement to C^0; these heads include not just finite verbs and auxiliaries, but also predicate adverbs.

6.3 Licensing Movement and Focus Interpretation

The structures proposed above for *li* questions explain the left-edge position of the focused elements. In addition, this placement explains why these are not interpreted as topics. Topics, although a left-edge phenomenon, are adjoined to IP, while the focus phrases are within the projection of C^0. However, the proposed structures do not explain why the fronted maximal projections are obligatorily interpreted as focused. I propose that the clitic contains an inherent focus feature [+F] that it can assign via Spec-head agreement. These *li* questions are similar to English clefts in that they involve constructional focus. The difference in interpretation between the two types of *li* questions

[14] The structure in (26) is the result of right adjoining the predicate adverb to I^0 and then moving the I^0 complex to C^0. Predicate adverbs do not form a morphological unit with the auxiliary, i.e., they are neither clitics nor affixes, and left and right adjunction to I^0 are possible.

[15] With different predicate adverbs, certain of the four possible orders are preferred, just as certain orders are preferred in the declarative counterparts.

is due to the different syntactic positions of the fronted constituents: the [+F] feature is discharged onto elements in SpecCP, but not onto heads adjoined to C^0.

6.3.1 Spec-head Agreement

Spec-head agreement can occur between the complementizer *li* and SpecCP (Rizzi 1990).[16] Once SpecCP is marked with the focus feature via Spec-head agreement, a constituent must move to this position to license the [+F] feature. In turn, this constituent receives focus interpretation, while the remainder of the clause is presupposed.

(27) [Poèmu] li on čitaet?
poem Q he read
'Is he reading a poem-FOC?'

(28) $[_{CP}\ poèmu_i\ [_{C'}\ [_{C[+WH]}\ li[+F]\ \ [_{IP}\ \ldots\ t_i\ \ldots\ ]]]]$

In (28) Spec-head agreement has occurred between the clitic in C^0 and SpecCP. This forces a maximal projection, in this case the direct object *poèmu*, to move to SpecCP where it then receives focus interpretation. In addition, the non-focused part of the sentence is presupposed to be true.[17] For example, in (27), it is assumed that the person in question is reading something; the issue is what precisely is being read. The focus interpretation is often similar to a contrastive focus reading, i.e., the question is asking about one particular item in a set to the exclusion of other items which might belong to the same set (section 4.2.3). However, the interpretation of the focus need not be strictly contrastive. That is, the question may simply ask whether the focused element is a (potentially non-unique) correct value for the presupposed portion of the question.[18]

[16] SpecCP cannot be analyzed as always being a focus position because focused phrases can only appear there in this one construction. For example, focused phrases do not appear before complementizers in embedded clauses, as such an analysis would predict.

[17] Kiefer 1980 points out that this is similar to the assumptions behind wh-questions, only in yes-no questions it is assumed that both the speaker and the hearer know that the non-questioned, non-focused part of the question is true. Kiefer 1980 and Hajičová 1983 discuss the degree to which the non-focused portion of the question can be presupposed, in the technical sense; I leave this issue aside. The important fact is the focus reading which the pre-*li* constituent obligatorily receives.

[18] I would like to thank an anonymous reviewer for bringing this to my attention. A similar contrast can be seen in the English examples in (i) and (ii).

(i) Was it a book that Ann read?
(ii) Did Ann read a BOOK?

This analysis predicts that if the focus feature is assigned to SpecCP, inherently non-focused elements should be, minimally, odd in these constructions. These include *kto-nibud'* 'someone' and *čto-nibud'* 'something'. As predicted, although these elements are possible in *li* questions, they cannot appear in SpecCP followed by *li*, where they have a focused interpretation forced on them.

(29) a. Oni sprosili, [kupil] li kto-nibud' xleba.
they asked bought Q someone bread
'They asked if someone had bought bread.'

b. *Oni sprosili, [kto-nibud'] li kupil xleba.
they asked someone Q bought bread

(30) a. Oni sprosili, [kupila] li ona čto-nibud'.
they asked bought Q she something
'They asked if she bought something.'

b. *Oni sprosili, [čto-nibud'] li ona kupila.
they asked something Q she bought

In (29a), the embedded question contains the phrase *kto-nibud'* 'someone'. Since the verb has head-adjoined to C^0, there is no focused phrase in SpecCP, and the sentence is grammatical. However, if *kto-nibud'* moves to SpecCP, as in (29b), the sentence is ungrammatical because the usually unfocused *kto-nibud'* is incompatible with the focus position. Such questions are pragmatically odd because the non-focused part of the clause is presupposed in the *li* construction, but the placement of an item like *kto-nibud'* in focus position potentially denies the truth of this presupposition. For example, in (29b) it is presupposed that someone bought bread, but the focusing of *kto-nibud'* potentially allows for a negative answer, which would indicate that no one bought bread, contrary to the presupposition. The sentences in (30) show the same incompatibility, with the accusative *čto-nibud'* 'something'.[19]

In (i) the presupposition is that Ann read a unique something, while in (ii) the presupposition is that Ann read something, but not necessarily a unique something. The Russian *li* questions have readings similar to those in (ii).

[19]It has been suggested that 'everyone' is also incompatible with focus. However, the Russian equivalent *vse* can appear in SpecCP. In (i) *vse* is the subject of the verb and appears focused in SpecCP; (ii) shows the object form *vsex*. There has been some disagreement about whether 'everyone' is inherently unfocused. These sentences may be evidence that it is not.

(i) Ja ne znaju, [vse] li priexali.
I not know everyone(all) Q arrived
'I don't know if everyone arrived.'

(ii) Oni sprosili, [vsex] li ty videl.
they asked everyone(all) Q you saw
'They asked if you saw everyone.'

Since the non-fronted material is obligatorily presupposed and an item cannot be simultaneously focused and presupposed, it is predicted that no material other than that in SpecCP can be focused. As seen in (31), this prediction is borne out, and emphatic stress cannot be used to focus an item which is not in SpecCP (Chvany 1973).[20] (Stress, and hence contrastive focus, is indicated by bold face.)

(31) a. Oni sprosili, [Ivan] li ušel včera.
they asked Ivan Q left yesterday
'They asked if Ivan-FOC had left yesterday.'

b. *Oni sprosili, [Ivan] li **ušel** včera.
they asked Ivan Q left yesterday

c. *Oni sprosili, [Ivan] li ušel **včera**.
they asked Ivan Q left yesterday

In (31a) the subject *Ivan* has been fronted and is the focus of the question. It is impossible to focus any other element of the sentence when the subject is fronted. So, (31b) in which the verb is stressed and thus must be interpreted as contrastively focused, is ungrammatical. The same holds for (31c) in which the adverb must be focused. Thus, when a maximal projection moves to SpecCP in a *li* question, it must be focused due to the assignment of the focus feature, and no other constituent can be focused because it would be within the presupposed portion of the question.

Finally, note that a part of the maximal projection may be the material that is focused and questioned. For example, if an NP containing an AP moves into SpecCP, it is possible to interpret the AP as being questioned, as in (32).[21] Which part of the maximal projection is focused is indicated by stress. So, if the AP is focused, it will be stressed, while if the head noun is focused, it is stressed.

(32) [Doroguju li knigu] ona kupila?
expensive Q book she bought
'Did she buy an expensive-FOC book?'

[20] Multiple foci are possible in very limited situations, i.e., in corrections of previously uttered questions. Consider (i) in which *knigu* is focused by virtue of being in SpecCP and *Inna* is focused by emphatic stress.

(i) Knigu li kupila Inna?
book Q bought Inna
'Was it a book-FOC that Inna-FOC bought?'

Under usual conditions, (i) would be ungrammatical, as (31b) and (31c) are. However, if the speaker had first asked the question without focusing *Inna*, and the hearer had misinterpreted the question, e.g., thinking the speaker was asking about *Anna*, then the speaker might repeat the question as in (i).

[21] As seen in example (10c), it is also possible to move the AP *doroguju* to SpecCP, stranding the NP.

Basically, SpecCP defines the domain in which the focus must be found. A similar phenomenon is discussed in Jackendoff 1972 with respect to English cleft constructions. The generalization he makes is that although the focus of the cleft construction must be found in the upper clause, i.e., within the constituent following 'it is', the constituent itself need not be focused. One possibility, following Chomsky 1992, is that SpecCP defines the checking domain for the question's focus, and a focused item must appear in this domain for the sentence to be well-formed.

6.3.2 Head-Adjunction

Next consider the clauses in which the verb has head-adjoined to C^0. If no maximal projection moves to SpecCP, *li* still needs a phonological host, and so head-to-head movement of the material in I^0 to C^0 occurs.[22] Once this material head-adjoins to C^0, it is in a position to host the clitic complementizer.

(33) Čitaet li on poèmu?
read Q he poem
'Is he reading a poem?'

(34) $[_{CP}\ [_{C'}\ [_{C}\ \textit{čitaet}_i\ [_{C[+WH]}\ \textit{li}[+F]]]\ [_{IP}\ \textit{on}\ [_{IP}\ [_{I'}\ [_{I}\ t_i\ ]\ \ldots]]]]$

In (34) the verb *čitaet* has head-adjoined to C^0, providing a lexical host for the complementizer.

Unlike the maximal projections in SpecCP, the head-adjoined verb does not necessarily have a contrastive focus reading of the verb. In fact, the usual reading, at least in embedded contexts, is one in which the question carries no presupposition. That is, the question concerns whether the relevant event occurred. Correlated with this is the fact that when the verb is fronted, another element can be focused by stress (Chvany 1973).

(35) a. Oni sprosili, [ušel] li Ivan včera.
they asked left Q Ivan yesterday
'They asked if Ivan had left yesterday.'

[22] Under this analysis, the [+F] feature of *li* would not be assigned when head-movement to C^0 occurs. That is, the focus feature is not assigned to the adjoined head (see below). One way in which to account for this distribution is to claim that the [+F] feature is optional. When it appears, Spec-head agreement occurs and the focus feature is licensed by movement of a maximal projection to SpecCP (section 6.3.1). When the focus feature is absent, no Spec-head agreement occurs and movement to SpecCP is not licensed.

b. Oni sprosili, [**ušel**] li Ivan včera.
they asked left Q Ivan yesterday
'They asked if Ivan had left-FOC yesterday.'

c. Oni sprosili, [ušel] li **Ivan** včera.
they asked left Q Ivan yesterday
'They asked if Ivan-FOC had left yesterday.'

d. Oni sprosili, [ušel] li Ivan **včera**.
they asked left Q Ivan yesterday
'They asked if Ivan had left yesterday-FOC.'

The 'neutral' reading of the question in seen in (35a) which has no emphatic stress. However, if emphatic stress is assigned to a constituent, that constituent is the focus of the question. This stress and focusing can fall on any item, even though it is the verb that appears in C^0. First, the verb itself can be contrastively focused, as in (35b). The stress on the verb forces a focused reading in which the implication of the question is that Ivan did something yesterday, but the speaker is not sure what, perhaps Ivan left. This contrasts with (35a) in which the question has no such implication and merely asks whether Ivan left yesterday or not. In (35c) the subject *Ivan* is the focus of the question, and in (35d) the adverb *včera* receives focus interpretation.

The explanation for the data in (35) is that when the verb moves to C^0, the focus feature is not assigned, and correspondingly there is no obligatorily presupposed portion of the clause. The resulting interpretation is that the entire content of the clause is the focus of the question. This accounts for the fact that *li* questions with fronted verbs usually have a reading in which there is no presupposition, i.e., no constituent within that clause is contrastively focused. So, the entire action denoted by the verb is negated by a negative answer (Hajičová 1983:92–94). For example, in (35a) the event of Ivan's leaving yesterday is questioned; a negative answer will deny the fact that he left yesterday, and positive answer will affirm it.[23]

However, the clause whose content is being questioned can contain a contrastively focused constituent. This focus is marked by stress and the remainder of the clause is presupposed, as it would be in a declarative sentence, shown in (36).

(36) a. **Ivan** ušel včera.
Ivan left yesterday
'Ivan-FOC left yesterday.'

[23] See Kiefer 1980 on the felicity of responses to different types of yes-no questions.

b. 'Evgenija Onegina' **Puškin** napisal.
Eugene Onegin Pushkin wrote
'Pushkin-FOC wrote 'Eugene Onegin'.' (Yokoyama 1986:191)

c. Ja **k Anne** idu.
I to Anna go
'I am going to Anna's-FOC.'

In the verb-initial interrogative counterparts to clauses like those in (36), the contrastive focus is also the focus of the interrogative, resulting in a reading similar to that of the structures with fronted maximal projections. If there is no such contrastive focus in the assertion, the question carries no presuppositions and hence the entire clause is the focus of the question. Consider (35c), which is the interrogative counterpart of (36a). (36a) presupposes that someone left yesterday and focuses Ivan as the one who left. The interrogative retains this interpretation, only now it is within a yes-no question. As such, the non-focus part of the question is presupposed, as was seen with the *li* questions containing fronted maximal projections, while the contrastively focused item which is the focus of the question. So, in (35c) it is assumed that someone left yesterday, and the speaker wants to know whether it was Ivan or someone else who did the leaving. Thus, when the verb is in C^0, *li*'s focus feature is not assigned, and the entire clause is questioned. If that clause contains a contrastively focused element, it becomes the focus of the question, resulting in a meaning similar to that of fronting that contrastively focused item to SpecCP.

6.4 Conclusion

In sum, *li* is a complementizer which assigns a focus feature. If Spec-head agreement occurs, a maximal projection moves to SpecCP where it is the focus of the question and hosts the clitic. If no maximal projection moves to SpecCP, then the verb in I^0 undergoes head-movement to C^0 in order to host the clitic. In these verb-initial structures, the entire clause is questioned. If the clause contains a focused constituent, as indicated by stress, then that focused constituent will be the focus of the question, and the resulting reading is similar to if the focused constituent had moved to SpecCP. However, if there is no contrastively focused constituent, the result is a 'simple' yes-no question.

6.5 Cross-Slavic Comparison

The clitic *li* appears in a number of Slavic languages, including Serbo-Croatian and Bulgarian.[24] In this section I present data from these languages, utilizing the analyses of Rivero 1993, Rudin 1993, and Izvorski 1994.[25] I follow Rudin 1993 and Rivero 1993 in claiming that Serbo-Croatian and Bulgarian *li*, like their Russian counterpart, are [+wh] complementizers. *Li* contains a focus feature that can be discharged onto a maximal projection in SpecCP. If no constituent moves to SpecCP, a verbal head undergoes head-movement to C^0 to host the clitic. There are slight differences among the *li* constructions in the three languages, concerning which heads can move to C^0 to host the clitic and whether *li* appears after the first phonological word or the first constituent. Despite these differences I suggest that my analysis of Russian *li* can be extended to Serbo-Croatian and Bulgarian.

---

[24] *li* also resembles the particle used to form Finnish yes-no questions, and Ouhalla 1992 contains an interesting discussion of focus in Standard Arabic yes-no questions and the two question particles found in that language. Examples of Finnish yes-no questions with the question particle *ko* are given below (data from Arto Anttila, p.c.; thanks to Erika Mitchell for bringing this Finnish construction to my attention).

(i) Juo-tt-i-ko Jussi Marja-lle vodka-a?
drink-CAUS-PAST-Q John Mary-ALLA vodka-PART
'Did John make Mary drink vodka?'

(ii) [Jussi-ko] Marja-lle vodka-a juotti?
John-Q Mary-ALLA vodka-PART drink-CAUS-PAST
'Did John-FOC make Mary drink vodka?'

(iii) [Marja-lle-ko] Jussi vodka-a juotti?
Mary-ALLA-Q John vodka-PART drink-CAUS-PAST
'Did John make Mary-FOC drink vodka?'

(iv) [Vodka-a-ko] Jussi Marja-lle juotti?
vodka-PART-Q John Mary-ALLA drink-CAUS-PAST
'Did John make Mary drink vodka-FOC?'

[25] SC has an additional complementizer that introduces matrix and embedded yes-no questions *da li*, as in (i). This complementizer is not a clitic and does not license movement into either SpecCP or C^0.

(i) Da li Marko studira medicinu?
COMP Marko study medicine
Is Marko studying medicine?

Bulgarian also has a non-clitic complementizer for yes-no questions: *dali*. *Dali* does not license the movement of either the verb or maximal projections, possibly because it lacks a focus feature, and as such can only be interpreted as questioning the entire sentence. An example is shown in (ii).

(ii) Pitaha me dali e kupil vestnika.
asked-3P me whether have bought-3S newspaper
They asked me whether he had bought a newspaper. (Rudin 1985:63)

6.5.1 Serbo-Croatian

Serbo-Croatian (SC) *li* questions are very similar to Russian ones in that either a focused maximal projection in SpecCP or an I^0 head adjoined to C^0 can precede the clitic complementizer *li*, as in (37).

(37) a. Marko li studira medicinu?
Marko Q studies medicine
'Is it Marko-FOC who studies medicine?' (Rivero 1993:568)

b. Studira li Marko medicinu?
study Q Marko medicine
'Is Marko studying medicine?' (Rivero 1993:568)

In (37a) the subject NP *Marko* is in SpecCP and is obligatorily the focus of the question. Rivero 1991, 1993 focuses on head-movement constraints and as a result does not discuss in detail the constructions in which a maximal projection hosts the clitic.[26] I propose that SC *li* can assign a focus feature via Spec-head agreement and thus license the movement of a maximal projection to SpecCP. However, as in Russian, if there is no maximal projection to host the clitic, the finite verb moves to C^0, as in (37b), in which the finite verb *studira* is adjoined to C^0.[27]

SC has compound tenses which involve an auxiliary and a non-finite verb. These auxiliaries are often clitics. In declarative clauses, if there is no maximal projection to their left to act as a phonological host, Rivero 1991 argues that the verb raises to C^0 to host the clitic auxiliary, as in (38). However, in *li* questions this long-distance head movement is impossible.[28] The only head which can move to C^0 in *li* questions is that in I^0, a restriction similar to that proposed for Russian. In these cases, the non-clitic form of the auxiliary appears since only it can host the clitic, as in (39).

(38) Čitao sam knjigu.
read have book
'I have read a book.' (Rivero 1993:572)

(39) Jesam li čitao knjigu?
have Q read book
'Have I read a book?' (Rivero 1993:572)

[26] In a footnote, Rivero 1993 suggests that the pre-*li* maximal projections are in SpecCP, although she does not explore these data.

[27] SC has a number of clitic pronouns and auxiliaries, whose precise distribution does not concern us here. However, the clitics follow *li*; Rivero 1993 suggests that clitic pronouns are within the IP complex which explains why they appear after *li* in C^0.

[28] The reason for this restriction is unclear and remains to be investigated.

When a sentence is negated, the negative marker directly precedes the finite verb, as in (40a). In negative *li* questions, both the finite verb and the negative marker move to C^0, shown in (40b), as was the case in Russian.

(40) a. (Ja) ne čitam.
I not read
'I do not read/I am not reading.' (Rivero 1991:334)

b. Ne vidim li ga?
not see Q him
'Don't I see him?' (Rivero 1993:572)

Rivero 1993 uses this behavior as evidence that the finite verb incorporates into the head of NegP, resulting in the negative marker and the finite verb behaving as a unit. An alternative view is that the negative marker and finite verb form a complex head from the beginning (King 1994b), i.e., they are both in I^0, as was proposed for Russian. This approach eliminates the need for incorporation and the unused specifier in these constructions.

Thus, I adopt Rivero's analysis for SC *li*, with two modifications: first, using a complex I^0 instead of incorporation of I^0/Tns^0 to Neg^0; second, the specification that via Spec-head agreement with *li*, SpecCP can be filled by a maximal projection which is the focus of the question. If no maximal projection moves to SpecCP, then the verb moves to C^0 to host the clitic.

6.5.2 Bulgarian

Bulgarian *li* is similar to its Russian and SC counterparts in that it can question either the entire clause when the verb is in C^0 or a maximal projection in SpecCP (Rudin 1993). This is illustrated in (41).[29]

(41) a. (Vie) namerihte li kŭštata?
you found Q house-the
'Did you find the house?' (Rudin 1985:64)

b. Kŭštata li namerihte (vie)?
house-the Q found you
'Was it the house that you found?' (Rudin 1985:64)

In (41a) *li* appears after the verb, *namerihte*; note that the subject pronoun can appear adjoined to CP in topic position[30] and hence pre-

[29] Unlike in Russian, yes-no questions in Bulgarian are almost always formed with the clitic, although questions signaled solely by intonation are possible (Rudin 1985:63).

[30] This contrasts with Russian in which topics are adjoined to IP; the topicalization constructions in Bulgarian and Russian are in other respects similar (section 5.2.3).

cede the verb. In (41b) a maximal projection has moved to SpecCP and is followed by the clitic. In these cases, as with Russian and SC, the maximal projection is the focus of the question.[31]

Consider briefly the fronting of a maximal projection to SpecCP. Unlike in Russian, Bulgarian *li* does not appear after the first phonological word in the clause. Instead, it cliticizes directly to the constituent to its left, regardless of how many phonological words this constituent comprises, as in (42). As a result, the surface word order of Bulgarian more directly reflects the phrase structure than that of Russian. Regardless, the focus of the question must be contained in this constituent, as was the case with Russian.

(42) $[_{NP}$[Novata]$_{\phi}$ [zelena]$_{\phi}$ [riza]$_{\phi}$] li ti podari Krasi?
new-the green shirt Q to-you give Krasi
'Did Krasi give you a new green shirt?' (Ewen 1979:111)

Like Russian, Bulgarian has the option of moving an inflected verb in I^0 to C^0 to host *li*. However, as discussed by Rivero 1993, the interaction of the other Bulgarian clitics, negation, and the future marker *šte* result in a more complicated description. I describe the basic problem below.

6.5.2.1 Bulgarian Clitics

First, a brief discussion of the other Bulgarian clitics is necessary. The clitics always appear as a group in the order AUX(except 3sg)–IO–DO–AUX(3sg) and are adjacent to the verb. If a non-clitic precedes the verb, it hosts the clitics, as in (43). If there is no host for the clitics, they follow the verb, as in (44). (Clitics are italicized.)

(43) Az *sŭm mu go* dal.
I have to-him it given
'I have given it to him.' (Hauge 1976:6)

[31]There may be uses in which the clitic is not in C^0. These remain to be investigated in order to determine whether they can be subsumed by the analysis of *li* described here. See Rudin 1985, 1993 for a more complete description of Bulgarian *li*.

For example, there is a construction in which *li* appears after an entire sentence; the construction questions a previous utterance in the discourse, as shown by the context and question below. Rudin 1985:67 describes the *li* final constructions as having an echo status, which would indicate that they are a different construction.

S1: Znaeš li dali e v kŭšti?
know-2s Q whether is in house
Do you know if s/he is at home? (Rudin 1985:65)

S2: Dali e v kŭšti li? Ne znaja.
whether is in house Q not know-1s
If s/he's at home? I don't know (Rudin 1985:65).

(44) a. Kazvam *mu* vsičko.
tell him everything
'I tell him everything.'
b. Viždal *go e*.
seen him have
'He has seen him.' (Rivero 1993:570)

In (43) there is a non-clitic *az* in initial position which hosts the clitics. In contrast, in (44) a verbal form moves so that it can host the clitics. In (44a) the host is a finite verb *kazvam*. In (44b) the finite element *e* is itself a clitic; so, the participle *viždal* moves from V^0 to C^0 (Rivero 1991).

The negative marker *ne* precedes the finite verb, as in (45). The clitics follow the negative marker, as in (46). Hauge 1976 points out that the clitic following *ne* is stressed, and the *ne* plus stressed clitic complex forms a non-clitic which can host any remaining clitics. Under no other circumstances can clitics bear stress. So, in (46) the auxiliary *sŭm* is stressed, but not the other two clitics. This unusual behavior ultimately plays a role in the placement of *li*.

(45) Ivančo ne običa devojki.
Ivancho not like girls
'Ivancho doesn't like girls.' (Ewen 1979:128)

(46) Ne *sŭm mu go* dal.
not am to-him it given
'I didn't give it to him.' (Hauge 1976:18)

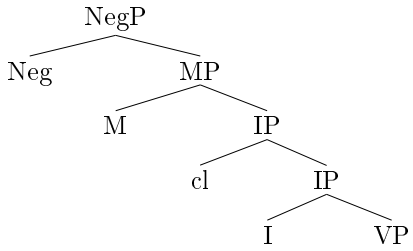
The final item to interact with clitic placement is the future marker *šte*. This marker is followed by a finite verb, and Rivero 1993 argues that it heads a projection M(odal)P. Clitics appear between *šte* and the finite verb, as in (47).

(47) (Vie) šte *mu go* dadete.
you fut to-him it give
'You will give it to him.' (Hauge 1976:17)

Rivero suggests that Bulgarian phrase structure is roughly like that in (48). Unlike Russian and SC, the Bulgarian structure has a separate projection for Negation.[32] Rivero is not explicit as to where the clitics are located, although it can be assumed that they are in a position similar to that in SC, which she describes as within the IP complex. Here I have represented them as adjoined to IP.

[32]However, see King 1994b for an analysis of Bulgarian in which Neg^0 and Tns^0 form a complex head.

(48)



6.5.2.2 Head-Movement in *li* Questions

To return to the main issue at hand: which heads can move to C^0 to support *li*. In the usual case, the finite verb, or participle if the finite verb is a clitic auxiliary, moves to C^0, as in (49).

(49) a. Izpratih li *mu* kniga?
send Q to-him book
'Did I send him a book?' (Rivero 1993:569)

b. Vižd*a*l li *go* *e*?
seen Q him has
'Has he seen him?' (Rivero 1993:570)

This pattern changes when negation or *šte* are present. If the clause is negated, *li* appears after the negative marker and any other clitics, as in (50).

(50) Ne *mu* li izpratih kniga?
not to-him Q send book
'Didn't I send him a book?' (Rivero 1993:573)

The pattern with *šte* is more unexpected. *li* appears after both *šte* and the finite verb, as in (51). The modal *šte* and any clitics precede the finite verb as they do in non-*li* clauses.

(51) Šte *go* viždaš li?
fut him see Q
'Will you see him?' (Rivero 1993:574)

These patterns with the negative and future modal, led Rivero 1993 to propose that Neg^0 and M^0 are barriers for head-movement to C^0. If they are present, a verbal head cannot move to C^0, and thus there is no host for the clitic. As a result, *li* must lower ('hopping' in her terms). The landing site depends on the construction. With the negative, Rivero claims that *li* left-adjoins to I^0 so that it follows the clitics but precedes the finite verb and participles. With the future, *li* right-adjoins to I^0 and thus follows the finite verb. Although Rivero is correct in that, as in Russian and SC, the finite verb raises to C^0 to support the clitic, her description of *li* lowering is empirically inadequate. In

section 5.2.3 I present the data which her analysis does not account for and propose an analysis which does not involve syntactic lowering.

6.5.2.3 Empirical Problems

Putting aside the possible theoretical objections to a lowering analysis, there are empirical problems with Rivero's analysis. Consider the negated sentences. Under Rivero's analysis, when no clitics are present, *li* should follow *ne* and precede the finite verb because *li* will left-adjoin to I^0. This is not the case; instead, *li* follows both *ne* and the verb, as in (52). Note that the ungrammaticality of (52b) is not phonological since *ne* can host other clitics, as was seen in (46).

(52) a. Ne običa li devojki?
not like Q girls
'Doesn't he like girls?'

b. *Ne li običa devojki?
not Q like girls
'Doesn't he like girls?'

A second problem is that Rivero predicts that if there are several clitics, *li* will follow all of them. However, when more than one clitic follows the negative marker, *li* appears after the first clitic and is followed by the others, as in (53). This happens regardless of which clitics are present.

(53) a. Ne *go* li *e* viždal?
not him Q is seen
'Didn't he see him?'

b. Ne *mu* li *go* dadohte?
not to-him Q it gave
'Didn't you give it to him?'

Why should *li* appear in these positions? A lowering analysis becomes ungainly because the landing site must be specified differently for a number of situations.

Consider instead the following possibility. Under usual conditions, *li* can be hosted by a maximal projection in SpecCP or a verbal head adjoined to C^0. In addition, following Rivero 1993, *ne* in Neg^0 and *šte* in M^0 are barriers for head-movement. However, even when *ne* or *šte* are present and there is no maximal projection in SpecCP, *li* needs a phonological host. Under such conditions, prosodic inversion can occur at PF as a last resort mechanism. Prosodic inversion allows a clitic to encliticize to the right-edge of a following phonological word if no constituent precedes the clitic (Halpern 1992). Unlike lowering, prosodic inversion occurs at PF, not in the syntax. So, when there is

no constituent to its left, *li* undergoes prosodic inversion and cliticizes to the first stressed element in the clause. In negated clauses, this stressed element will be the clitic or verb which immediately follows *ne*. With *šte*, this stressed element will be the finite verb since neither *šte* nor any clitics following it bear stress (Hauge 1976). As was seen in the above discussion, these are the desired orderings.[33]

This account predicts that when *ne* and *šte* co-occur, *li* will appear after *šte* because the presence of *ne* results in *šte*, the element immediately following *ne*, being stressed. Although the *ne šte* combination is 'bookish' and hence rarely used, this prediction is borne out, as seen in (54).

(54) Ne šte li *ste* *mu* *go* dali?
not will Q aux to-him it give
'Won't you give it to him?' (Hauge 1976:20)

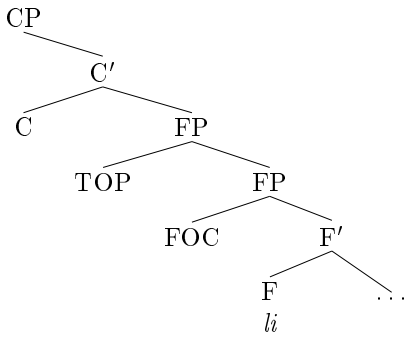
6.5.2.4 FocusP: Izvorski 1993, 1994

The previous discussion has assumed that *li* in Bulgarian, like *li* in Russian, is in C^0. However, Izvorski 1993, 1994 has argued that in Bulgarian *li* heads a projection below CP, but above the verbal functional projections. She terms this projection F(ocus)P. Focused phrases can appear in SpecFP both when *li* heads FP and when FP has a null head (section 5.2.3.1). Izvorski cites evidence from the order of topicalized and focused phrases in *li* questions, from the distribution of focused phrases in *li* questions, and from the contrast in behavior between questions formed with *li* and those formed with *dali*, which is a true complementizer. I briefly outline Izvorski's arguments below, highlighting the basic similarity between Russian and Bulgarian *li*, despite the difference in syntactic position of *li* in the two languages.

The basic structure Izvorski posits is shown in (55) (see section 5.2.3 for further discussion of topic and focus positions in Bulgarian).

[33] Unfortunately, what this account does not explain is the difference in behavior of *ne* and *šte* with respect to their ability to place stress on the following element. Clearly, the complex behavior of the Bulgarian clitic system warrants further investigation.

(55)



First, notice that in Bulgarian *li* questions, the order of constituents is topic–focus–*li*, as in (56). If topics are adjoined to the highest functional projection below CP, namely FP, then this is the order predicted.

(56) a. Ivan na Maria li dade knigata?
Ivan to Maria Q gave the-book
'Was it to Maria-FOC that Ivan-TOP gave the book?'
(Izvorski 1993:28)

b. Na Maria cvetja li podari Ivan?
to Maria flowers Q gave Ivan
'Was it flowers-FOC that Ivan gave to Maria-TOP?'

If *li* were in C^0, as it is in Russian, topics would appear after *li* and the focused constituent. However, this order is not possible, as seen in (57) in which the preverbal position of the subject *decata* indicates it topic interpretation.

(57) *Na kino li decata bjaha?
at cinema Q the-children were
'Was it to the movies-FOC that the children-TOP went?'
(Izvorski 1993:22)

In addition, placing *li* in F^0 predicts that topics can precede *li*, even when *li* is hosted by the verb. This is seen in (58) in which the topic *decata* precedes the verb *bjaha* which hosts *li*.

(58) Decata bjaha li na kino?
the-children were Q at cinema
'Did the children-TOP go to the movies?'

The second piece of evidence Izvorski cites in favor of *li* heading FP is that if a focused phrase appears before *li*, no other focused phrase can precede the verb. This is predicted by her account since the focused phrase must be in SpecFP and there is no additional preverbal focus position into which a maximal projection can move. This is demonstrated in (59) in which *Vojna i mir* occupies SpecFP; however, as

Izvorski 1993 argues, negative pronominals, such as *nikoj*, also occupy focus position. When *nikoj* is moved to preverbal position in (59), the result is ungrammatical since the pre-*li* focused phrase is in this position.

(59) *Vojna i mir li nikoj ne pročete?
War and Peace Q nobody not read
'Was it *War and Peace*-FOC that nobody read?'

Finally, the contrast of behavior between *li* and *dali* show that *li* is not in C^0. *Dali* is a complementizer which marks yes-no questions. In the structure in (55), since *dali* is in C^0 it will precede both topicalized and focused phrases, as in (60).

(60) a. Dali na Maria cvetja podari Ivan?
Q to Maria flowers gave Ivan
'Was it flowers-FOC that Ivan gave to Maria-TOP?'

b. Dali knigata vidjaha te?
Q the-book saw they
'Was it the book-FOC that they saw?' (Izvorski 1993:23)

Thus, following Izvorski, Bulgarian *li* is not in C^0. Instead it heads a functional projection below CP. As with Russian *li*, a maximal projection in the specifier of this position is the focus of the question. In addition, if there is no maximal projection to host *li*, prosodic inversion occurs, in which case the material heading the functional projections, usually the finite verb, hosts the clitic (Izvorski 1994).

To conclude, in Bulgarian and SC, as in Russian, a maximal projection can move to the Specifier of the projection headed by *li*, where it is obligatorily focused. If no maximal projection moves to this Specifier, a verbal head adjoins to the head containing *li* to host *li*. In SC, this head is C^0; in Bulgarian, following Izvorski, it is F^0. Contra Rivero's analysis, I propose that when there is no host for the clitic, prosodic inversion can occur at PF, so that *li* will follow the first stressed element in the clause.

Part II

7

An Overview of LFG

The first part of this dissertation argued for a specific underlying structure for Russian and then discussed how the varying surface word orders were derived and interpreted. In particular, the basic word order was argued to be VSO, with other orders representing particular discourse function interpretations. The discussion presumed a syntactic theory that involved the projection of arguments into base-generated positions and then their movement into certain surface positions. However, not all syntactic theories presume this type of organization. The question then arises as to how the observations and analyses in the first part can be captured in a theory without movement in the phrase structure and whether some of the problems that arose with the movement account can be avoided when analyzed in a theory without movement.

Chapters 8–10 discuss how these phenomena are captured in Lexical-Functional Grammar (LFG), a 'unification' based theory (Bresnan 1982 and articles therein). I argue that certain phrase structure positions in Russian are associated with particular discourse function information in LFG, as was the case in GB. In fact, the phrase structure proposed in chapter 9 closely resembles that proposed in chapter 5. One of the differences between the theories is that grammatical functions are not assigned via phrase structure positions in LFG; the distribution of grammatical functions in phrase structure is regulated by the interaction of functional uncertainty with well-formedness conditions on the functional-structure of the sentence. The distribution of verbal heads is regulated by the morphology and strictly constrained head-expansion rules for the phrase structure. Wherever there would have been traces in the structures in Part I, there is no node in the phrase structure of LFG. By looking at these structures, it is easier to determine which syntactic generalizations are about the phrase structure positions of items, which about their grammatical functions, and which about their

morphological distribution. Further discussion of these issues is found in the ensuing chapters.

In LFG, the surface word order is encoded by c(onstituent)-structure. C-structure encodes dominance and precedence relations amongst the words via phrase structure trees. C-structure is not used to encode grammatical functions, unlike S-structure in GB. Instead, grammatical functions and relations are encoded in the f(unctional)-structure. F-structures are composed of attribute–value matrices, where the attributes are entities such as Subject or Tense and the values range over a set appropriate to the attribute. Annotations on the c-structure, as well as information from the lexical entries of the words, provide the information needed to form the f-structure (see Bresnan 1982 and Sells 1985).[1] There are other modules of the grammar such as a(rgument)-structure and s(emantic)-structure whose properties are beginning to be explored in depth (Halvorsen 1983, Halvorsen and Kaplan 1988, Dalrymple et al. 1993). These different parts of the grammar interact with one another, and the information in one module must be compatible with that in another. I am primarily concerned with c- and f-structures. The idea is to properly constrain the related c- and f-structures to produce all and only those which capture the Russian data. To do this, a set of annotated c-structure rules and constraints on the f-structure is proposed.

In order to demonstrate how c- and f-structures interact, simplified representations of the c- and f-structures for an English sentence are discussed below. The c-structure rules are shown in (1). The annotations under the phrase structure nodes state how this information maps onto information in the f-structure. The $\downarrow$ indicates the information corresponding to that node, while the $\uparrow$ indicates the f-structure corresponding to the mother node. (This is intuitively clearer if the annotation is thought of as appearing above the phrase structure node in the tree so that the $\downarrow$ points to that node and the $\uparrow$ to its mother.)

[1] Annotations on c-structures are extremely powerful and perhaps insufficiently constrained. There has been some effort to decrease the extent to which the f-structure is 'derived' from the c-structure, a dependency which was not intended in LFG (Butt 1993).

(1) Annotated C-structure Rules:

a. S ⟶ NP (↑SUBJ)=↓ VP ↑=↓

b. VP ⟶ V ↑=↓ NP (↑OBJ)=↓

(1a) states that a sentence, S, is composed of an NP and a VP. The annotation (↑SUBJ)=↓ under the NP states that the NP therein is the subject of the f-structure corresponding to its mother, i.e., of the f-structure corresponding to S. The annotation ↑=↓ under the VP states that the VP is the head of its f-structure, which is identical to that of S. So, there is one f-structure corresponding to two c-structure nodes. (1b) expands the VP into a V, which is the head of the VP as seen by the ↑=↓annotation, and an NP, which is the object of the f-structure of its mother, i.e., of the f-structure corresponding to the VP and in turn the S. So, in (4) there are three f-structures: one outer f-structure containing two inner f-structures; the f-structures are delineated by square brackets.

The sample English sentence *I saw Ivan* is presented in (2)–(4), which provide the relevant lexical entries, c-structure, and f-structure respectively.

(2) Lexical Entries:

I	*saw*
CAT:NP	CAT:V
(↑PRED)='PRO'	(↑PRED)='see<SUBJ,OBJ>'
(↑PRS)=1	(↑TNS)=PAST
(↑NUM)=SNG	
(↑CASE)=NOM	

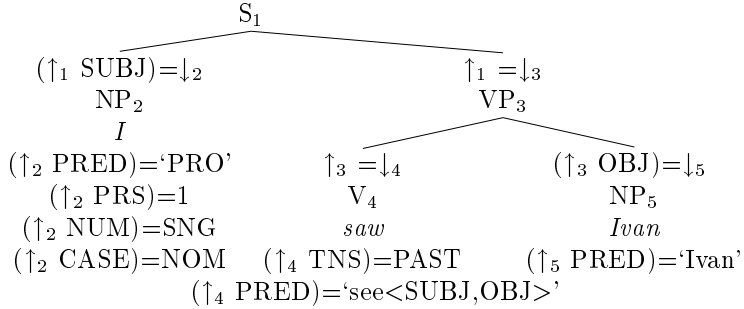
Ivan
CAT:NP
(↑PRED)='Ivan'
(↑PRS)=3
(↑NUM)=SNG

The information to the left of the equations in the lexical entries, i.e., that which is preceded by an ↑ specifying which f-structure it belongs to, forms the attribute of an f-structure. The material to the right of the equations designates the value of the attribute for that item. The lexical entries in (2) specify the PRED(icate) value of the word along

with other information that is relevant to the (morpho)syntax. The CAT(egory) information is used for insertion into the c-structure. In addition to the PRED which contains information as to its argument structure, the verb *saw* states that its TNS (tense) is past. For the pronoun person, number, and case specifications are included as well as the PRED value.

(3) is an annotated phrase structure tree produced by combining the lexical entries in (2) with the phrase structure rules in (1). The annotations on the c-structure map between c- and f-structures.

(3) Annotated C-structure Tree:



The f-structure corresponding to the annotated c-structure tree in (3) is shown in (4). The numbers in (3) correspond to the numbered f-structures in (4), collapsing the labels 1=3=4 which are identical. The value of an attribute in an f-structure can itself be an f-structure, e.g., the value of SUBJ is an f-structure containing information about the subject. The outermost f-structure has the PRED value provided by V^0, which is the head of VP, which in turn is head of S.

(4) F-structure:

$$1{=}3{=}4\begin{bmatrix} \text{PRED} & \text{'see<SUBJ,OBJ>'} \\ \text{TNS} & \text{PAST} \\ \text{SUBJ} & 2\begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{PRS} & \text{1st} \\ \text{NUM} & \text{SNG} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\ \text{OBJ} & 5\begin{bmatrix} \text{PRED} & \text{'Ivan'} \\ \text{PRS} & \text{3rd} \\ \text{NUM} & \text{SNG} \end{bmatrix} \end{bmatrix}$$

What do Russian c-structures and their corresponding f-structures look like? In the phrase structure (c-structure) rules, I assume that

all nodes, both complements and heads, are optional (this optionality is not explicitly marked, except where confusion might arise). This optionality overgenerates c-structures in that many of the producible c-structures have no counterparts in the language. However, well-formedness conditions on the f-structure, such as functional uniqueness, completeness, and coherence (section 2.2), rule out many of these c-structures since they have no corresponding well-formed f-structure. In addition, this optionality, especially the optionality of heads, results in c-structures that at first appear rather bizarre. However, it must be remembered that c-structure is not identical to GB S- or D-structure and as a result does not correspond to either of these structures.

There is one assumed exception to this optionality: adjunction structures. In adjunction structures, the complement must be present for the rule to apply. This prevents vacuously branching nodes. (Although having such nodes does not result in ill-formed f-structures, it can be thought of as a violation of minimality.) In addition, the distribution of null nodes is constrained by a requirement that a projection in a language exists only if there is at least one construction in which that projection has a head; the same holds of complements.

The basic c-structure rules are in (5) and (6): (5) encodes dominance relations, (6) precedence. These are compatible with the underlying structures argued for in the first part of this dissertation; for example, they separate finite from non-finite verbs, place complementizers at the beginning of the clause.

(5) a. $\begin{array}{lcl} \text{CP} & \longrightarrow & \begin{array}{cc} \text{XP}, & \text{C}' \\ (\uparrow\text{GF})=\downarrow & \uparrow=\downarrow \end{array} \end{array}$

b. $\begin{array}{lcl} \text{C}' & \longrightarrow & \begin{array}{cc} \text{C}, & \text{IP} \\ \uparrow=\downarrow & \uparrow=\downarrow \end{array} \end{array}$

c. $\begin{array}{lcl} \text{IP} & \longrightarrow & \begin{array}{cc} \text{XP}, & \text{I}' \\ (\uparrow\text{GF})=\downarrow & \uparrow=\downarrow \end{array} \end{array}$

d. $\begin{array}{lcl} \text{I}' & \longrightarrow & \begin{array}{cc} \text{I}, & \text{VP} \\ \uparrow=\downarrow & \uparrow=\downarrow \end{array} \end{array}$

e. $\begin{array}{lcl} \text{VP} & \longrightarrow & \begin{array}{cc} \text{XP}, & \text{V}' \\ (\uparrow\text{SUBJ})=\downarrow & \uparrow=\downarrow \end{array} \end{array}$

f. $\begin{array}{lcl} \text{V}' & \longrightarrow & \begin{array}{cc} \text{V}, & \text{XP} \\ \uparrow=\downarrow & (\uparrow\text{GF})=\downarrow \end{array} \end{array}$

(6) a. $\text{XP} > \text{Y}'$
b. $\text{X}^0 > \text{YP}$

(5a–b) provide a basic structure for the CP: a specifier position that can contain a maximal projection with any grammatical function (GF) (section 9.1) and a complementizer position that can take an IP com-

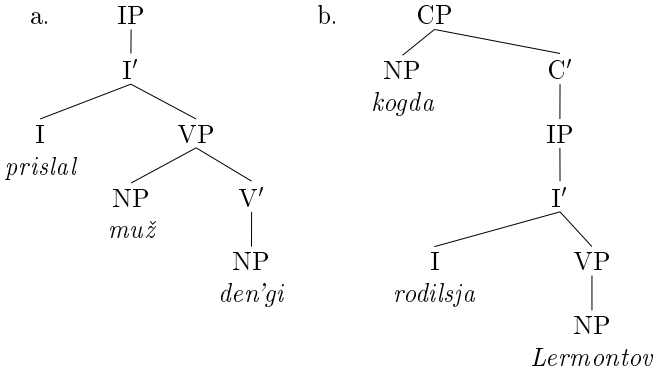
plement. (5c–d) do the same for IP. C′ and I′ are double headed structures, as witnessed by the ↑=↓ annotation (Grimshaw 1991; see section 10.3). The VP rule in (5e) states that the subject precedes the V′ which contains the verb and a maximal projection associated with any grammatical function, (5f); this allows the subject to follow the finite verb in I^0. The precedence rules in (6) state that maximal projections precede intermediate bar level projections and that heads precede maximal projections. These precedence rules in conjunction with the dominance rules in (5) produce a basic X′ syntax, taking into account the optionality of the nodes in (5).

Two now familiar Russian examples and their c-structures, as described by the above rules, are demonstrated briefly below. (The annotations are not included.)

(7) a. Prislal muž den'gi.
sent husband money
'My husband sent (me) money.'

b. Kogda rodilsja Lermontov?
when born Lermontov
'When was Lermontov born?'

(8)



In both (8a) and (8b) the finite verb is in I^0. In addition, in both the subject a daughter of the VP, as permitted by (5e). In (8a), the VP also contains a V′ with the direct object, *den'gi*, created by (5f). In (8b), the wh-phrase *kogda* is in SpecCP due to the application of (5a) (section 9.3). The c-structure rules and the structures they produce are discussed in more detail in the following chapters.

Organization The second part of this dissertation comprises three chapters, each addressing a different issue concerning the implementation and role of phrase structure in LFG. As was discussed above and

in chapter 2, the dissociation of grammatical functions from phrase structure, i.e., constituency and dominance of items, allows for structures which would be impossible under an approach in which the phrase structure is used to encode grammatical relations. The first chapter in Part II examines case assignment in LFG. In GB, the chain composed of an argument and its traces is assigned case once and only once. This case can be inherent, in which case its assignment is linked to theta role assignment, or it can be structural, in which case it is assigned in a certain configuration. I argue that there are a few Russian constructions which assign case to a particular phrase structure position in LFG; however, most of what is structural case in GB is an association of grammatical functions with a particular case in LFG. As such, a particular grammatical function can be case marked with the desired case regardless of where it is located in the phrase structure. Since this assignment is essential for understanding how noun phrases can appear anywhere in the phrase structure, without being associated with any traces, case assignment is discussed before delving into the correlations between discourse functions and phrase structure in LFG. A particular case in point is the genitive of negation which has generally been assumed to have a structural description. This structural description is difficult to capture in LFG where the only available phrase structure is that of the surface string. Instead, I argue that the genitive of negation interacts with the other case assignment rules and the realization of grammatical functions, without making reference to phrase structure.

This dissociation of phrase structure and grammatical functions does not mean that certain phrase structure positions may not be associated with certain grammatical functions; such association is possible, but not obligatory. In addition, just as I argued in chapter 5 that certain phrase structure positions in GB are associated with particular discourse functions, the same argument can be made for the phrase structure in LFG. This is the topic of chapter 9. Functional uncertainty is introduced to account for the interpretation of discourse functions and wh-questions. Functional uncertainty allows the items in a particular phrase structure position to range over a set of discourse and grammatical functions; this set can be defined to constrain the types and interpretations of the constituents in each position. Although the phrase structure rules proposed follow a simple X′ schema and incorporate the configurationality argued for in chapter 3, the instantiation of these rules involves no empty categories or place holders, in accordance with the constraints on LFG phrase structure.

Finally, chapter 10 discusses one particular aspect of Russian phrase

structure: how to capture the phenomena generally attributed to head-movement. In particular, the distribution of *li* questions and predicate adverbs, as well as that of finite and non-finite verbs, is examined. The differences between finite verbs and infinitives is reduced to a morphological difference: finite verbs are type I^0, while infinitives are type V^0. This morphological difference results in a difference in phrase structure distribution since certain phrase structure rules make reference to I^0, while others refer to V^0. The behavior of predicate adverbs and *li* questions suggests that there are phrase structure rules which expand heads into two or more heads. I argue that these rules are constrained by the category information of the head, so that only certain types of expansions are possible. Thus, the interaction of morphological information with the phrase structure rules allows us to capture all of the phenomena attributed to head-movement without such movement taking place.

8

Case Assignment and the Genitive of Negation

The first section of this chapter examines case assignment in LFG, both from a general perspective and as concerns Russian. The second section investigates the genitive of negation from a different perspective than that in Part I.

8.1 Case Assignment in LFG

This section provides a discussion of case marking in LFG. All Russian nouns are inflected for case.[1] Even the nominative, which is considered unmarked in some languages (e.g., Hindi), has distinct case endings in two of the three gender classes. The nominative is not the absence of case. The first part of this section discusses the structure of case markers. The second part examines the types of case assignment, outlining some of the more important rules for Russian case assignment. There are essentially four[2] types of case assignment in LFG: configurationa, grammatical, lexical, and semantic. These potentially overlap in that certain cases may have both semantic and grammatical or configurationa requirements.

I assume that the case endings indicate the case assigned to the noun.[3] Each case-marked noun (hence each noun, since all nouns are case-marked) must be licensed. A case-marked noun that is not as-

[1] See section 2.3 for sample case paradigms. For more details on case assignment in Russian see Babby 1986, 1993, Brecht and Levine 1986, Fowler 1987a, 1987b, Jakobson 1936/1971, 1958, among others.

[2] GB has only three types due to the fact that grammatical relations are defined in terms of structural position. In LFG these are separate, and as a result, case assignment sometimes depends on one, sometimes on the other.

[3] Nominative case masculine nouns and genitive plural feminine and neuter nouns do not have overt case endings. However, these are still constrained to be assigned case. One way of doing this is to have a zero morpheme. This morpheme combines

signed the case shown by the case affix is ill-formed. Within LFG this is accomplished by constraint equations associated with the morphological endings (Andrews 1982, Neidle 1988). Constraint equations do not assign a value to an attribute; instead, they require that the specified value be assigned to the attribute in order for the f-structure to be well-formed.[4]

(1) *-u* $(\uparrow \text{CASE}) =_c \text{ACC}$
$(\uparrow \text{NUM}) = \text{SNG}$
$(\uparrow \text{GND}) =_c \text{FEM}$

(1) is a lexical entry of a Russian case marker. The marker *-u* consists of two constraint equations and an assignment equation. The assignment equation indicates that the noun is singular. The constraint equations require that the noun be feminine and be assigned accusative case. The information as to the gender class of the noun is part of the lexical entry of the noun stem. The case of the noun is assigned as discussed later in this section.

A noun stem entry is shown in (2). The morphology ensures that all nouns have case endings since noun stems are bound morphemes and must be combined with a case suffix to form an X^0 category that can be inserted into the syntax (Selkirk 1982). In X′ terms, we can say that noun stems are of category N^{-1} and only become category N^0 when combined with a case marker.

(2) *knig-* $(\uparrow \text{PRED}) = \text{'book'}$
$(\uparrow \text{GND}) = \text{FEM}$

Combining the noun stem entry in (2) with the case marker entry in (1) gives (3).

---

with the noun stem in the lexicon just as any other case marker. Compare the zero morpheme below to the overt one in (1).

-∅ $(\uparrow \text{CASE}) =_c \text{NOM}$
$(\uparrow \text{NUM}) = \text{SNG}$
$(\uparrow \text{GND}) =_c \text{MASC}$

There will have to be several such entries for the morpheme: one for masculine nominative singulars, one for feminine genitive plurals, one for neuter genitive plurals. However, this is a problem for overt case markers as well. For example, the case ending *-u* marks neuter and masculine dative singulars, as well as feminine accusative singulars. Much of this redundancy can be captured by using featural decomposition of the cases and genders (Jakobson 1958; Neidle 1988), although some of it, such as the multiple uses of *-u*, is simply arbitrary.

[4]Constraint equations can be compared to checking in the Minimalist program (Chomsky 1992). Under this approach, case on nouns must be checked at some point in the derivation or the derivation will crash, resulting in a malformed structure.

(3) *knigu* (↑PRED) = 'book'
(↑GND) = FEM
(↑NUM) = SNG
(↑CASE) $=_c$ ACC

(3) states that *knigu* is a feminine noun whose predication value is that of 'book'. *Knigu* itself is not an accusative case noun, rather it is a form that must be assigned accusative case.

Before beginning the main proposal, a brief point about Russian morpho-syntax must be discussed. There is a problem with gender mismatches in Russian. (This is a problem for all theories of morphosyntax.) Certain morphologically feminine nouns are masculine in referent, e.g., *mužčin-a* 'man'. For purposes of the syntax, i.e., for verb and adjective agreement, these act as masculine nouns, but they take morphologically feminine case endings. (See Walinska 1992 for discussion of this in Polish.)

(4) Tolstyj mužčina byl v magazine.
fat-MASC.SNG man-FEM.SNG was-MASC.SNG in store
'The fat man was in the store.'

In (4) the subject is morphologically feminine, but it is masculine in reference. Both the modifying adjective and the verb show masculine agreement. One way to implement this in LFG is to have two types of gender for each noun: morpholocigal gender (M-GND) and syntacticosemantic gender (S-GND). The default is for these to be identical. However, certain noun stems are of one M-GND but a different S-GND. The case endings are only specified as to M-GND. Verb and adjective agreement, which apply in the syntax, are sensitive to S-GND.

The remainder of this section discusses four different types of case assignment. The first two are configurationa case assignment and case assignment based on grammatical functions. Configurational case assignment is linked to the c-structure, i.e., phrase structure, position of an NP. In contrast, case assignment can also be based on the grammatical function of an NP, regardless of its c-structure position. The second two types are lexical and semantic case assignment. Lexical case assignment is idiosyncratically assigned by a given predicate to one of its arguments; information about lexically assigned case must be included in the lexical entry of a predicate. Finally, semantic case is assigned to an argument based on its meaning and role in the clause.

8.1.1 Configurational Case Assignment

Case can be assigned configurationally. Any noun appearing in a certain phrase structure position is assigned the case in question. This method of case assignment is quite limited in Russian, although it may be more prelevant in 'fixed' word order languages. Two examples of configurationally assigned case in Russian are complements of nouns and external topics. The complements of nouns receive genitive case, as in (5); an example is given in (6), in which *Ivana* is assigned genitive case.

(5) $\quad \text{NP} \longrightarrow \text{N} \quad (\text{NP})$
$\qquad\qquad\qquad\qquad\qquad\quad ((\downarrow\text{CASE}) = \text{GEN})$

(6) [kniga]$_N$ [Ivana]$_{NP}$
book Ivan-GEN
'Ivan's book'

This case assignment is optional since a few nouns assign lexical or idiosyncratic case to their complements. If the assignment of genitive case in (5) was not optional, then the complements of lexically case assigning nouns would receive case twice: the lexical case and the genitive case. The resulting case clash would result in a ill-formed f-structure.[5]

External topics appear in the nominative case (section 5.1.1.1). This case is associated with the position in which external topics appear. (7) states that the XP in the external topic position is assigned nominative case (see section 9.2.1 on the placement of external topics in c-structure); an example is shown in (8).[6]

(7) $\quad \text{E} \longrightarrow$ XP CP
$\qquad\qquad\quad (\uparrow\text{E-TOP})=\downarrow \quad \uparrow=\downarrow$
$\qquad\qquad\quad (\downarrow\text{CASE})=\text{NOM}$

(8) [Milicionery,]$_{XP}$ [na stole ležalo dve furažki.]$_{CP}$
policemen-NOM on table lie two service caps
'Policemen, on the table there lay two service caps.'
(Franks and House 1982:161)

[5]For Russian, this view can be maintained. However, there are languages in which some nouns have double case marking, i.e., the noun bears two case forms, one semantic and one based on grammatical function (e.g., Korean (Gerdts 1988, 1991)). For these languages, some modification must be made as to how case is represented in f-structure (Cho and Sells 1993).

[6]As it stands, the rule in (7) would assign nominative case to PPs in this position. Pending further investigation, I assume that such an assignment is vacuous, but not ungrammatical. It may turn out that certain categories are case resistant, in which case this rule would have to be modified. One simple modification would be to make it optional; since there is no other rule which would assign case to external topics, it would apply to all and only NPs.

In (8) the external topic *milicionery* is assigned nominative case due to its position under E.

8.1.2 Grammatical Functions and Case

Case can be associated with certain grammatical functions. In GB, this is collapsed with configurationally assigned case in that certain positions are associated with certain grammatical functions and thus with certain cases. However, in LFG these nouns are assigned case in direct association with grammatical functions. Three such rules are discussed below. The first assigns nominative case to the subjects of tensed verbs, the second accusative to objects, and the third dative to oblique goals. These rules are associated with the lexical entries of predicates in the lexicon. (9) assigns nominative case to the subjects of finite verbs, i.e., any verb with a TNS attribute (infinitives have no TNS attribute, chapter 10).[7] An example is shown in (10). The subject *vrač* is in the nominative regardless of its position relative to the verb.

(9) $(\uparrow \text{TNS}) \longrightarrow [(\uparrow \text{SUBJ CASE}) = \text{NOM}]$

(10) Prišel [vrač]$_{SUBJ}$.
arrived doctor-NOM
'A/the doctor arrived.'

There are no oblique case subjects in Russian, unlike languages such as Hindi (Mohanan 1990) and Icelandic (Zaenen et al. 1990). The exception to this may be subjects of infinitival clauses that have been argued to assign dative case to their subjects (Neidle 1988).[8] However, these are tenseless, and rule (9) will not apply.

Accusative case is assigned to direct objects by (11). In the example in (12), the object *sobaku* is assigned accusative case, regardless of its phrase structure position.

(11) $((\uparrow \text{OBJ CASE}) = \text{ACC})$

[7] One potential exception is sentential and infinitival subjects, which are not marked for case. It may be that these are assigned case which is not morphologically realized.

See also Greenberg and Franks 1991 who argue that the datives in constructions with an infinitive and a dative, as in (i) are in fact subjects. In these constructions, tense is indicated by the copula, which shows default third singular neuter agreement.

(i) Mne bylo uxodit'.
me-DAT was-NEUT.SNG leave-INF
'I had to leave.' (Greenberg and Franks 1991:72)

[8] There are a number of impersonal expressions in Russian. If these contain an expletive subject, then (9) will assign nominative to that expletive. However, if these constructions do not have expletive subjects, then (9) is vacuous since there is no subject.

(12) Inna videla [sobaku]$_{OBJ}$.
Inna saw dog-ACC
'Inna saw the dog.'

Unlike the nominative case assignment rule, the accusative rule is optional, as indicated by the parentheses. There are two reasons for this. The first is that the genitive of negation can mark the objects of negated verbs with genitive, instead of accusative case. The second is that certain verbs assign lexical or idiosyncratic case to their objects which is not overridden.[9]

The rule in (13) assigns dative to oblique goals. Once again, in the example in (14) the position of the oblique goal is immaterial to the application of the case assignment rule.

(13) (↑OBL$_{GO}$ CASE) = DAT

(14) On dal [Inne]$_{OBL.GO}$ knigu.
he gave Inna-DAT a book
'He gave Inna a book.'

(13) assigns the dative case to oblique goals, e.g., indirect objects. This rule might extend to cover oblique experiencers, such as the logical subjects of predicate adverbs, which appear in the dative case.

These rules are associated with the appropriate lexical items before lexical insertion. (12) is associated with any lexical item bearing a TNS specification, e.g., finite verbs. (13) and (14) are associated with the entries of all verbs. The rules do not apply to verbs without OBJ or OBL$_{GO}$ arguments. In fact, all of the rules discussed in this section could be associated with any lexical item, although in many cases their application will be vacuous or will optionally not apply.

8.1.3 Lexical Case Assignment

Lexical or idiosyncratic case is assigned by certain prepositions and verbs and a few nouns derived from these verbs. lexical case assignment rules refer to grammatical functions associated with the lexical entry of these prepositions and verbs, as in (15) and (16).

(15) a. *u* 'at/near' PREP $< \theta >$
(↑OBJ CASE) = GEN

b. u [okna]$_{OBJ}$
at window-GEN
'at the window'

[9]Some of these are not objects, but obliques, in which case the rule in (12) would not apply (Fowler 1987a). However, as discussed in section 8.1.3, some objects are assigned cases other than accusative, and these cases are not overridden by (11).

(16) a. *upravljat'* 'govern' V $<$ AGT TH $>$
(↑OBJ CASE) = INST

b. On ploxo upravljal [armiej]$_{OBJ}$.
he badly governed army-INST
'He ran the army badly.'

In (15) the object of the preposition *u* is in the genitive case. In (16) the object of the verb *upravljat'* is assigned the instrumental case. Although theta-role information, and not grammatical function information, is specified in the lexical entry, the annotation assigning lexical case refers to grammatical functions. For example, in (16) the lexical case annotation refers to OBJ, but the argument structure of the verb specifies that it takes an agent and a theme, saying nothing about their grammatical functions. However, Fowler 1987 argues from passivization facts that the instrumental (and genitive) objects of verbs are in fact objects. Interestingly, when these verbs are passivized, the subject appears in the nominative, not the instrumental, as in (17).[10]

(17) [Russkaja armija]$_{SUBJ}$ upravljalas' Kutuzovym,
Russian army-NOM govern-PASS Kutuzov
s ego štabom, i gosudarem iz Peterburga.
with his staff and sovereign from Petersburg
'The Russian army was run by Kutuzov, with his staff,
and by the sovereign from Petersburg' (Fowler 1987:7)

In (17) the subject *russkaja armija* is assigned nominative case, as with any subject in Russian. Since the case assignment rule in (16) only makes reference to OBJ, it will not apply in passive constructions, like that in (17).[11] Had the instrumental case been associate directly with the argument, i.e., with the thematic role, then it would be impossible to have nominative subjects of passive verbs of this type.

Consider the consequences of associating case directly with the theta-role. The argument associated with that theta-role would be marked with the relevant case regardless of the grammatical function it receives. Evidence from passivization in Icelandic indicates that this

[10] Case government by prepositions is another matter. Most likely, this government is lexical and hence should be assigned via grammatical functions, as in (15), to unify it with lexical case government by verbs. However, if it is possible to have prepositions assign case semantically, then direct association with the thematic role would be desirable (section 8.1.4).

[11] Neidle 1988 presumes that when passive occurs, all relevant mentions of OBJ are replaced by SUBJ. However, the mechanisms of such an across-the-board substitution are not discussed, and her analysis was developed before Lexical Mapping Theory. The data presented in Fowler 1987 demonstrate that such across-the-board substitution is untenable empirically, as well as theoretically.

is the correct approach for non-accusative objects in Icelandic. When an Icelandic verb that assigns lexical case to its object is passivized, the same lexical case is assigned to the subject (Zaenen et al. 1990). This contrasts with the situation in Russian in which passive subjects are always nominative. Thus, the mechanism of lexical case assignment differs in Icelandic and Russian: in Icelandic it is associated with a particular argument of a predicate, while in Russian it is associated with a particular grammatical function.

8.1.4 Semantic Case

The final type of case assignment is semantic case. With semantic case a particular case is associated with a particular semantic meaning. For example in Urdu the ergative is associated with volitionality (Butt and King 1991), and Simpson 1983 discusses semantic case in Warlpiri. For Russian, Babby 1986 argues that the genitive of negation is a type of semantic case, although it also has a structural component (section 8.2), and Babby 1993 discusses adversity passives in Russian. The instrumental case marking instruments without a preposition is another candidate for semantic case in Russian.[12]

Unlike lexical case, semantic case generalizes across a particular semantic meaning, just as case can generalize across a particular grammatical function. Since this meaning is constant, the case will always be associated with the argument bearing that particular meaning, regardless of its thematic role or grammatical function. For example, Fowler 1987 argues that dative "objects" of verbs never undergo passivization, unlike their accusative, instrumental, and genitive counterparts (section 8.1.3), and hence are not true objects. There are two, possibly interrelated ways to represent this. One is to claim that these datives are not objects, but instead bear some other grammatical relation; this was suggested in section 8.1.2. The second is to try to unite these datives by their semantics so that any argument with the appropriate semantic function is assigned the dative case. Since this case is semantic and is not associated with a particular grammatical function, it cannot be overridden and these arguments will be unable to passivize.

[12]The instrumental case is used in a variety of contexts unrelated to the notion instrument. The use relevant to semantic case is demonstrated in (i).

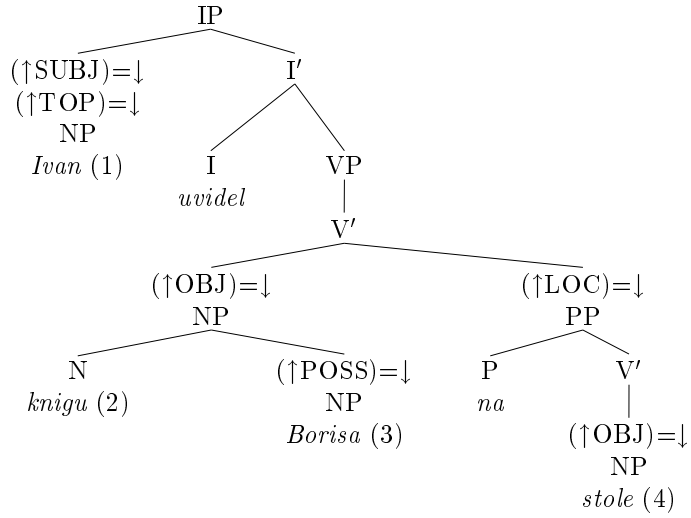
(i) Ona napisala pis'mo karandašom.
she wrote letter pencil-INST
'She wrote the letter with a pencil.'

8.1.5 An Example

The following is a simple example of how case assignment works in Russian. The sentence in (18) has an f-structure like (20) and a c-structure like (19).

(18) Ivan uvidel knigu Borisa na stole.
Ivan-NOM saw book-ACC Boris-GEN on table-PREP
'Ivan saw Boris's book on the table.'

(19)



(20)

$$\begin{bmatrix} \text{PRED} & \text{'see<SUBJ,OBJ>'} \\ \text{TNS} & \text{PAST} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'Ivan'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'book'} \\ \text{CASE} & \text{ACC} \\ \text{POSS} & \begin{bmatrix} \text{PRED} & \text{'Boris'} \\ \text{CASE} & \text{GEN} \end{bmatrix} \end{bmatrix} \\ \text{LOC} & \begin{bmatrix} \text{PRED} & \text{'on<OBJ>'} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'table'} \\ \text{CASE} & \text{PREP} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

How are the nouns in (18) assigned case? (1) Consider first the subject of the sentence *Ivan*. Since the TNS of the sentence is past, i.e., finite, the rule that assigns nominative case to the subject applies (rule 9). (2) The object *knigu* is assigned accusative case by the object rule (rule 11) since the verb *uvidet'* does not assign lexical case to its object. (3) The possessor of the object, *Borisa*, is assigned structural case by the head noun *knigu* (rule 5). (4) Finally, *stole*, the object of the preposition *na*, is lexically assigned prepositional case by the preposition (*na* 'on$< \theta >$' (↑OBJ CASE)=PREP).

8.2 The Genitive of Negation

The genitive of negation was discussed in section 3.3 as an example of a subject-object asymmetry. The basic generalization is shown in (21).

(21) **Genitive of Negation:** Any NP sister to V^0 at D-structure may be assigned genitive case when the sentence is negated, if that NP does not receive inherent case.

This domain was chosen because the genitive of negation affects direct objects, subjects of passives and unaccusative, but not subjects of transitive verbs or unergatives. Examples are repeated below.

(22) a. Mal'čik ne vidit knigi.
boy not see book-GEN
'The boy does not see a/(?the) book.'

b. Ni odnogo goroda ne bylo vzjato.
not one city-GEN not was taken
'Not one city was taken.' (Chvany 1975:184)

c. Ne pojavilos' studentov.
not show up-SG students-GEN
'No students showed up.' (Pesetsky 1982a:66)

d. *Mal'čika ne vidit knigu.
boy-GEN not see book
'The/a boy does not see the book.'

e. *V pivbarax kul'turnyx ljudej ne p'jet.
in beerhalls cultured-GEN people-GEN not drink-sg
'Cultured people do not drink in beerhalls.'
(Pesetsky 1982a:43)

Although there are many unanswered questions as to why the genitive of negation exists and what mechanism results in the genitive marking, this section is concerned with how to capture the distribution in (22) in LFG. The challenge is that there is no way to capture a structural domain which includes direct objects and subjects of certain

intransitives in a theory that has no underlying structure from which elements move.

8.2.1 Neidle's Account

Neidle's 1988 account of the genitive of negation is that the genitive of negation marks objects with the genitive when the sentence is in the scope of a Q marker, which is provided by the negative marker. Certain subjects are affected by this rule after being demoted to objects. The annotated phrase structure rules in (23) and (24) are proposed by Neidle to account for the assignment and distribution of the genitive of negation.[13]

(23) S $\longrightarrow$ (NP) VP
($\uparrow$SUBJ) $= \downarrow$ $\uparrow = \downarrow$

(24) VP $\longrightarrow$ (ne) V (NP)
($\uparrow$Q)=+ $\uparrow = \downarrow$ ($\uparrow$OBJ)=$\downarrow$
(($\downarrow$CASE)=ACC)
(($\uparrow$Q)=+ $\longrightarrow$($\downarrow$CASE)=GEN &
(($\downarrow$Q)=+))

The rule in (23) states that a sentence is composed of a VP, which is its head, and an optional NP subject. The case of the subject, i.e., nominative in finite clauses, is supplied by a redundancy rule (section 8.1.2). The head of the VP is the verb. The optional particle preceding the verb marks negation. The particle's presence marks the VP as being within the scope of the operator Q (it is unclear what Q represents other than the possible scope of negation). The final line of the annotation on the direct object is optional and indicates that if the mother of the NP is marked by the scope of Q, then the direct object can be marked with the genitive.

Neidle derives the different semantic effects of the genitive of negation from the spread of the Q feature. The optional (($\downarrow$Q) = +) allows the noun phrase itself to be marked with the feature Q. If the direct object is marked with the accusative, then the object cannot be marked with the Q feature, since the only way to assign Q requires assigning

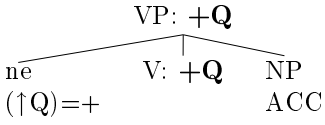
[13] In (24) I have simplified Neidle's rules. She states them in terms of her featural decomposition system. The actual version is as below.

VP	$\longrightarrow$	(ne)	V	(NP)
line 1		($\uparrow$Q)=+	$\uparrow = \downarrow$	($\uparrow$OBJ)=$\downarrow$
line 2				(($\downarrow$CASE)=[–,(–),+])
line 3				(($\uparrow$Q)=+ $\longrightarrow$($\downarrow$CASE)=[,+,] & (($\downarrow$Q)=+))

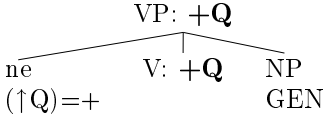
The case of the direct object is either accusative, if the option [–, –, +] is supplied by default (line 2), or genitive if the quantifying value is specified as + (line 3). These binary features correspond to locational, quantitative and directional respective; genitive and accusative differ in that the former is + quantitative.

the genitive. However, if the object is marked with the genitive, it is morphologically ambiguous whether the Q marker itself has spread onto the object, as in (27) (in which case an indefinite reading is applicable) or whether the genitive is a result of the scope of the Q marker on the VP, as in (26) (in which case a definite reading is applicable).

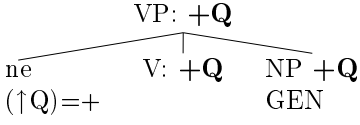
(25)



(26)



(27)



So, the Q scope-marking feature is passed by the negation to the VP node at which point it can be transmitted optionally to the direct object (27).[14]

The general principles of case assignment ensure that the genitive of negation does not apply to verbs with lexically marked objects. For these objects, the case is already specified in the lexical entry of the verb, and the case annotation on the phrase structure rules is not needed. This annotation is optional, and if it were to apply, the resulting f-structure would be ill-formed since the conflicting case features could not unify.

As stated, Neidle's genitive of negation rule affects only OBJ(ects). However, the genitive of negation also appears on the subjects of unaccusative intransitives and passives. Neidle formulates the lexical redundancy rule in (28).

(28) $(\uparrow \text{SUBJ}) \longrightarrow (\uparrow \text{OBJ})$
$(\uparrow \text{OBJ Q}) =_c +$

The first line of (28) converts subjects into objects. It is blocked in transitive verbs since no verb can have two direct objects due to Function-

[14]The existential verb 'to be' shows an interesting pattern in the negative. Its argument must occur in the genitive; there is no option of either nominative or accusative case (Chvany 1975). Neidle claims that the existential verb idiosyncratically states that if the verb is [+Q], e.g. negated, then the theme is an object within the scope of negation. This stipulation is not as idiosyncratic as might first be thought since the negation of an existential forces a non-definite interpretation (Neidle 1988:248).

Argument Biuniqueness; unergatives do not undergo the rule because the role of Agent is incompatible with the OBJ function. The second line is a constraint equation which requires the object to be marked with the feature Q. This guarantees that the object occurs in the genitive case since the marking of the object node with Q only occurs if the sentence is negated and the genitive case is chosen.[15] Thus, it is by use of this constraint equation that the argument must occur in the genitive if it is demoted, i.e., if it would have been a subject in the non-genitive of negation form. The passive examples are particularly interesting in that they show the interaction of two lexical rules, both of which affect the argument structure of the predicate. The passive demotes the subject and promotes the object to subject. Since the single argument of the verb is now a non-agent subject, it can undergo the rule in (28) which re-demotes it to an object and adds the constraint equation.

Thus, Neidle's proposal works as follows. The negative marker projects a marker Q which demarcates its scope. If an OBJ is in the scope of negation, it can appear in the genitive case. In addition, there is a rule that demotes SUBJ to OBJ with the constraint that the demoted argument be in the scope of Q. This accounts for the occurrence of the genitive of negation on the objects of transitive verbs and on the single argument of unaccusatives and passives.

8.2.1.1 Difficulties with the Account

Although I retain some of her basic ideas (section 8.2.2), there are two difficulties with Neidle's account: the first concerns genitive time adverbials and the second, the subject demotion rule.

The genitive time adverbs are discussed in section 3.3.2. Chvany 1975 points out that certain accusative time adverbials can occur in the genitive within the scope of negation, as in (29). Pesetsky 1982a used this as evidence that the domain of the genitive of negation must be structural, not thematic.

(29) a. Ja ni odnu minutu ne spal.
I not one minute-ACC not sleep
'I did not sleep a single minute.' (Pesetsky 1982a:92)

b. Ja ni odnoj minuty ne spal.
I not one minute-GEN not sleep
'I did not sleep a single minute.' (Pesetsky 1982a:92)

These examples posed a problem to Neidle's 1988 account because these time adverbials are not OBJ(ects), just as they posed a prob-

[15] The relevant annotation is:
$(\uparrow Q) = + \longrightarrow (\downarrow \text{CASE}) = \text{GEN} \ \& \ ((\downarrow Q) = +)$

lem for the GB account since they are not sisters to V^0.[16] Since they are not OBJ, they are not affected by the genitive of negation case marking rule. However, given Franks and Dziwirek's 1993 analysis of genitive time adverbials as partitives (section 3.3.2), this problem is avoided. Genitive time adverbials are not part of the genitive of negation phenomenon and as a result are not expected to be case-marked by the genitive of negation rule. This, of course, leaves the problem of accounting for the distribution of the partitive.

The real difficulty with Neidle's account is the demotion rule which turns SUBJ into OBJ. Although such a rule may have been technically possible given the machinery available in LFG at the time Neidle analyzed these data,[17] it is impossible once Lexical Mapping Theory (LMT) is incorporated into the system (section 2.2). Using LMT, each verb is specified as taking a set of thematic roles as its arguments. These thematic roles are specified as to whether they are r(estricted) or o(bjective). In turn, this information constrains the grammatical function of the argument. The lexical entries do not contain fully specified information as to the grammatical function of the arguments. This makes statement of Neidle's demotion rule impossible in the lexicon, since the grammatical function information is not there, and in the f-structure, since such a rule would be altering the structure of the predicate.

There is an additional difficulty with the demotion rule. Neidle's demotion rule in (28) requires that the demoted NP be marked with Q, i.e., be directly in the scope of Q. Since Neidle's rule assigning case to objects marks the object with Q only if that object is in the genitive, her constraining equation results in all demoted objects being in the genitive.[18] Although this association with Q correctly drives the case marking, it has undesirable consequences for the meaning of the object marked with Q. Neidle 1988:59-60 suggests that there is a semantic reflex of being marked with Q. The existence of the referent of lexical items marked with the feature Q is not presupposed. This accounts for

[16] It might be suggested that time adverbials could arbitrarily be assigned the function OBJ to derive the facts. However, this ignores the co-occurrence of time adverbials with real OBJ (section 3.3.3), which would be unexpected since grammatical functions are unique in their clause.

[17] The demotion rule may not have worked even before the advent of LMT. For example, which instances of SUBJ should be replaced by OBJ? Would this include mention of SUBJ in control rules? The answer appears to be no. Even if the demotion rule applies only in the lexicon, which might eliminate the control problem, there may be other rules in the lexicon with which it would incorrectly interact.

[18] This prevents the object from occuring in the accusative; this demoted argument is always in the nominative or the genitive, never the accusative.

the fact that many genitives under negation have an indefinite reading, in contrast to accusatives which usually receive a definite reading. This correlation is quite interesting. However, the demotion rule predicts that the arguments of intransitives should never be presupposed when they appear in the genitive. The data in (30) involving existential sentences demonstrate the falsity of this predication¡.

(30) a. Ego tam ne bylo.
he-GEN there not was
'He was not there.' (Neidle 1988:50)

b. (K sožaleniju) menja ne budet na vašem koncerte.
unfortunately me-GEN not is at your concert
'Unfortunately, I will not be at your concert.'

In (30) the argument which would have been the subject is in the genitive case. Under Neidle's account, these arguments must be marked with the feature Q. However, these arguments' existence must be presupposed since they are pronominal forms; this is particularly salient in (30b) in which the pronoun *menja* refers to the speaker.[19] Thus, the elegant correlation which Neidle made between the scope possibilities of the feature Q and semantics of the genitive of negation is undermined by the constraint equation in the subject demotion rule.

As such, I abandon Neidle's demotion rule. However, the basic insight behind this rule is the one that underlies unaccusativity and that was also utilized by Pesetsky 1982a in his work on the genitive of negation. The arguments which undergo the genitive of negation are identical in their theme-like nature. In GB this indirectly translates into identity in structural position due to the UTAH. In LFG this correspondence is not structural, although it has certain grammatical function correlations (next section).

8.2.2 LMT and the Genitive of Negation

The major difficulty with Neidle's 1988 analysis of the genitive of negation was that it made direct reference to grammatical functions. In this section I propose an alternative account using Lexical Mapping Theory. The intuition behind the genitive of negation is that it only affects arguments which can be objects. This description includes themes, the canonical undergoers of the genitive of negation, and certain experiencers. These experiencers have the same LMT specifications as

[19] Identical examples can be found with transitive verbs. However, these are not a problem for Neidle's analysis because with transitives, genitive marking on the object does not necessarily indicate that Q has spread to the object itself, which is what triggers this semantic difference.

themes, allowing them to be realized as objects and to appear in the genitive under negation.

8.2.2.1 Analysis

The fact that none of the genitive NPs are subjects suggests that the case marking rule associated with the genitive of negation interacts with the rules which link arguments to grammatical functions. Since grammatical functions are related to thematic roles via the linking rules, the fact that the genitive of negation only appears with certain thematic roles falls out.

The correlation between arguments and grammatical functions is captured by Lexical Mapping Theory (LMT).[20] Default rules associate certain theta-roles are associated with features from the set [±o(bject)] and [±r(estricted)]. LMT provides the additional features for each theta-role, and these feature sets define the grammatical functions as in (31).

(31)

SUBJ	[−r,−o]
OBJ	[−r,+o]
OBJ_θ	[+r,+o]
OBL_θ	[+r,−o]

Of concern here are the grammatical functions SUBJ and OBJ and the features associated with the theta-roles agent and theme.[21] In particular, agents are always [−o], which prevents their being objects, and themes are [−r], which results in their being objects of transitive verbs and subjects of intransitives. These features are always associated with the thematic roles in question and can be assigned via default rules in the argument-structure. Linking rules fill in their specification to derive the corresponding grammatical functions. There is an additional default rule as in (32). This rule assigns the [−r] feature to the highest argument.[22] In many cases, this argument is an agent; since agents are also assigned the [−o] feature by a default rule, they will be fully specified as subjects, i.e., [−r,−o]. With transitive verbs, (32) ensures that the agent, and not the theme, will be the grammatical subject.

(32) Assign [−r] to the highest argument.

[20] See the references in section 2.2 for details and motivation of LMT.

[21] Many of the other theta-roles have restrictions on them which result in their being realized either as OBJ_θ or OBL_θ. For details on argument-structure and LMT see Alsina 1993.

[22] (32) is a default which does not apply if it conflicts with a lexically specified feature assignment. This caveat is necessary because some impersonal verbs have a single, non-subject argument which is [+r]. An alternate way around this is to specify in the lexical entry of these verbs that they have a [−r] expletive subject.

(33) is a modified version of the Subject Condition which states that all clauses have a subject. (33) states that if it is possible to create a subject, this is done. In particular, (33) results in the [−r] theme of unaccusatives being realized as the subject of the sentence ((32) does not cause these arguments to be subjects since it does not provide the necessary [−o] feature).

(33) If possible, create a subject ([−r,−o]).

What does this have to do with the genitive of negation? The rule for the genitive of negation allows genitive case to be assigned to an argument if that argument is located in a negated f-structure, i.e., an f-structure whose NEG value is +. However, the linking rules effectively restrict the application of this rule to arguments which are [−r] in the lexicon, i.e., those arguments which are always [−r]. The rule is stated in (34).

(34) $(\text{CASE} = \text{GEN}) \wedge ((\text{GF} \uparrow)\ \text{NEG}) =_c +$

The rule in (34) allows arguments to be assigned genitive case under the condition that they appear under the scope of negation. This is similar to lexical case assignment in that a case is associated with a particular argument. The constraint equation requires that the argument be within a negated structure.[23] This rule does not directly refer to the scope of the negation. However, constituent negation is not marked as an attribute of the main f-structure; so, it cannot trigger (34). Since the rule makes no reference to thematic roles, there is no difficulty with the fact that experiencer objects, as well as themes, can undergo the genitive of negation. Below I discuss why inherently [−o] arguments never undergo (34).

Remember that genitives of negation are never subjects.[24] In fact, there are no genitive subjects in Russian. The nominative case assignment rule requires that all subjects of finite verbs be in the nominative (section 8.1.2). The application of (34) requires that the argument be in the genitive. If this argument is linked to the SUBJ function, the values of the SUBJ CASE attribute in the f-structure cannot unify, and the f-structure will be ill-formed. However, if the argument is linked to the OBJ function, there is no conflict, since the assignment

[23] Neidle 1988 uses the negation to introduce a feature Q which is the syntactic trigger for the genitive of negation. Introducing a feature could prove useful if additional elements trigger (34). For example, Neidle 1988:31-33 suggests that certain verbs, such as *ždat'* 'wait-for', inherently introduce the Q feature. The use of a feature introduced by negation, as opposed to the negation itself, will not alter the application of (34).

[24] Subjects can, however, be in the scope of negation, as witnessed by the distribution of negative polarity items which appear in both object and subject position.

of accusative case to objects is optional, and the genitive case can be realized. Once (34) has applied, the genitive argument cannot be a subject, and (33) cannot apply.[25] [26]

The genitive of negation rule applies to objects of transitives, unaccusatives, subjects of transitives, and unergatives, as in (35)–(38) respectively. Only the first two derivations result in a well-formed structure. The last two fail because the agent is assigned both nominative and genitive case.

(35) Transitive Object:

	Agent	Theme	
	[−o]	[−r]	Inherent Assignment
	[−r]		Rule 32
		gen	Rule 34
	nom		Nom. Subj. Rule
√	[−r,−o]:nom	[−r,+o]:gen	LMT

(36) Unaccusative:

	Theme	
	[−r]	Inherent Assignment, Rule 32
	gen	Rule 34
√	[−r,+o]:gen	LMT

(37) Transitive Subject:

	Agent	Theme	
	[−o]	[−r]	Inherent Assignment
	[−r]		Rule 32
	gen		Rule 34
	nom		Nom. Subj. Rule
*	[−r,−o]:nom/gen	[−r,+o]:acc	LMT

[25] This contrasts with the rule which creates ordinary direct objects. This rule also takes a [−r] argument and specifies it as [+o]. However, it does not apply in unaccusatives and passives, because the modified Subject Condition rule has precedence. One way in which to create direct objects which does not make reference to ordering among the linking rules is to have objects formed as a result of Function-Argument Biuniqueness, i.e., if there is already a subject, there cannot be another one; so, these [−r] arguments must be objects (Alex Alsina p.c.).

[26] There are other constructions in Russian which result in the non-application of rule (33). Among these are adversity impersonals which alternate with structures which have grammatical subjects (Babby 1994), suggesting that a lexical rule is in operation.

(38) Unergative:

	AGENT	
	[−o]	Inherent Assignment
	[−r]	Rule 32
	GEN	Rule 34
	NOM	Nom. Subj. Rule
*	[−r,−o]:NOM/GEN	

In addition to subjects and objects which are [−r], there are two other types of case-marked arguments: restricted objects, which are [+r,+o], and restricted obliques, which are [+r,−o]. In general, these restricted arguments are assigned case lexically (section 8.1.3) or semantically (section 8.1.4). In either situation, their case marking will conflict with that of the genitive of negation. That is, these arguments fail to undergo the genitive of negation in a fashion similar to that of the transitive and unergative agent arguments in (37) and (38). Lexically case marked objects suffer the same fate.

Thus, the genitive of negation affects certain types of arguments, namely those which are inherently [−r], when they are in the scope of negation. These arguments are always realized as objects since their case marking is incompatible with subject specifications. Subjects of transitive and unergative verbs do not undergo the genitive of negation because they are inherently [−o] and are the highest thematic argument; this forces them to be grammatical subjects and thus be incompatible with the assignment of genitive case. In particular, note that this account avoids one of the difficulties of the analysis presented in section 3.3, in which both subjects and objects were within the scope of negation, but only constituents which were sister to the verb at D-structure could undergo the genitive of negation. Reference to thematic information, instead of phrase structure position, circumvents this problem.

8.2.2.2 Semantic Correlates

There are certain semantic correlates associated with the genitive of negation, as discussed above.[27] Roughly speaking, an accusative object under negation is interpreted as definite, while a genitive object can be either definite or indefinite, although the indefinite reading is often preferred. Neidle 1988:60 suggests that with the genitive, the object

[27] Neidle 1988 suggests that part of the problem is that the genitive of negation may be changing and that there is variation among speakers. The change may be towards using the genitive of negation as a morphological marker of a particular semantics and away from allowing the genitive under sentential negation regardless of the particular semantics.

is not presupposed, hence the tendency for arguments affected by the genitive of negation to be indefinite, while the accusatives are more 'individuated'. Similarly, what is important for the genitive of negation is whether the specification of the object is within the scope of negation. Neidle tries to capture these semantic correlations with the spreading of a feature Q, introduced by the negation, where the spread of negation onto the object correlates with the indefinite, non-presupposed reading (section 8.2.1). I have not attempted to provide a syntactic account for these correlations due to the difficulty in determining what they are. For example, the correlation between genitive arguments of intransitive verb which Neidle seems to predict, i.e., that these genitives must be indefinite or not presupposed, does not hold since pronouns can occur in these constructions (section 8.2.1.1). There is a link between negation, genitive case-marking, and indefinites and presupposition. However, whether this connection is a result of the genitive of negation rule or is a set of separate phenomena which center around the same triggers is a matter for further investigation, one which will crucially depend on the semantics and interpretation of arguments in negative sentences.

The purpose of this chapter was to outline how LFG case assignment works for Russian. In addition to lexical and semantic case, case can be assigned to arguments in particular phrase structure positions or to arguments in particular grammatical functions. These are predictable, non-idiosyncratic case. Structural case, particularly in Russian, is relatively rare: here it was proposed for external topics, which are not arguments of the verb, and for nominal arguments. In contrast, the case marking associated with grammatical functions accounts for the distribution of several cases in Russian, e.g., subjects are nominative and objects are accusative. These cases are assigned to noun phrases with the appropriate grammatical function regardless of their position in the phrase structure, e.g., it is immaterial whether the subject is within the VP or sister to I′, it is still assigned nominative case. Since all nouns must be assigned case, as witnessed by the constraining equations associated with the lexical entry of each case-marked noun, this separation of case marking from phrase structure position is essential for a language like Russian in which a given phrase structure position may be associated with nouns marked with a variety of cases. This association of phrase structure positions with a variety of grammatical functions, especially in the preverbal field, becomes apparent in the next chapter. If case marking were dependent on the c-structure, it would prove impossible to account for the varied surface orders in

Russian without making reference to phonologically null positions in the phrase structure.

9

In a Theory Without Movement

Although grammatical functions are not universally defined by their phrase structure positions in LFG, certain phrase structure positions may be associated with particular grammatical functions or restricted to a certain subset of grammatical functions. For example, in English the preverbal position might be restricted to grammatical subjects; the annotation on this position would result in any constituent in that position corresponding to the subject in the f-structure. However, even in such cases, the phrase structure position does not define the grammatical function of the argument which appears there. In this chapter, I am primarily interested in the ways in which phrase structure positions are associated with discourse functions in Russian. Although there are few positions restricted to particular grammatical functions, associating specific phrase structure positions with discourse functions accounts for the interaction of word order and discourse function interpretation discussed in chapter 4. The phrase structure rules and associated interpretations proposed here are compatible with the other, independently motivated, phrase structure rules.

The basic Russian c-structure rules are in (1a–f). All nodes, both complements and heads, are optional, with the exception of adjunction structures. In adjunction structures, the complement must be present for the rule to apply. In addition, the nodes are unordered, i.e., the rules only encode dominance, not precedence. Linear precedence is stated as two general principles in (2).

(1) a. $\text{CP} \longrightarrow \underset{(\uparrow \text{GF}) = \downarrow}{\text{XP}}, \underset{\uparrow = \downarrow}{\text{C}'}$

b. $\text{C}' \longrightarrow \underset{\uparrow = \downarrow}{\text{C}}, \underset{\uparrow = \downarrow}{\text{IP}}$

c. $\text{IP} \longrightarrow \underset{(\uparrow \text{GF}) = \downarrow}{\text{XP}}, \underset{\uparrow = \downarrow}{\text{I}'}$

d. $I' \longrightarrow$ I, VP
$\uparrow=\downarrow \quad \uparrow=\downarrow$

e. $VP \longrightarrow$ XP, V'
$(\uparrow SUBJ)=\downarrow \quad \uparrow=\downarrow$

f. $V' \longrightarrow$ V, XP
$\uparrow=\downarrow \quad (\uparrow GF)=\downarrow$

(2) a. $XP > X'$
b. $X^0 > XP$

Section 9.1 discusses functional uncertainty, a notion that is essential for the interpretation of discourse functions and wh-questions. The ensuing sections demonstrate how this works for the Russian question and topic-focus data. The final section discusses the placement of discourse function information in the grammar.

9.1 Functional Uncertainty

Since LFG does not have movement and hence links between traces and surface positions, it needs a way to describe long-distance dependencies. Functional uncertainty captures such dependencies between discourse functions and clause internal grammatical functions (Kaplan and Zaenen 1988a, 1988b). These relationships are specified by uncertainty annotations of regular expressions.

The wh-question in (3) and the topicalization structure in (4) provide examples of when functional uncertainty is necessary in English.

(3) Who did John say that Mary saw?

(4) Inna, John claimed (that) he saw at the beach.

The problem with sentences like (3) and (4) is that a constituent which is located in non-canonical position logically belongs to the f-structure of a lower clause (Kaplan and Zaenen 1988b, Kroeger 1991, Matsumoto 1992). For example, in (4), 'Inna' is in initial position and is the topic of the sentence. However, it is also the object of the verb 'saw'. The mapping between c-structure and f-structure must specify that the topic of the sentence is identical to the object of the verb 'saw'. This is accomplished via functional uncertainty.

Let us assume the annotated c-structure rule in (5) for topic position in English. This rule states that the XP daughter to S' is mapped onto the TOPIC attribute of the f-structure.[1]

[1]Topic information is traditionally placed in the f-structure. However, there is reason to place topic and focus information in the semantic-structure as well since these discourse functions interact with the semantics and pragmatics (section 9.4).

(5) $S' \longrightarrow$ XP (↑TOP)=↓ S ↑=↓

(5) captures the fact that the XP is interpreted as the topic of the sentence, but it does not solve the problem of identifying the XP with a grammatical function in S. For a sentence like (4) this could be captured by an additional annotation under the XP like the one in (6). (6) states that the topic is identical to the object of the COMP, the desired statement for the facts in (4).

(6) (↑TOP)=(↑COMP OBJ)

The f-structure for (4) is shown in (7). The functional identity stipulated by (6) is manifested by the sharing of values for the TOP and the COMP OBJ attributes. The value is not provided separately for both attributes; instead, the one value, e.g., the f-structure corresponding to the NP 'Inna', is shared by both attributes.

(7)
$$\begin{bmatrix} \text{TOP} & [\text{PRED} \quad \text{'Inna'}] \\ \text{PRED} & \text{'<SUBJ,COMP>'} \\ \text{SUBJ} & [\text{PRED} \quad \text{'John'}] \\ \text{COMP} & \begin{bmatrix} \text{PRED} & \text{'<SUBJ,OBJ>'} \\ \text{SUBJ} & [\text{PRED} \quad \text{'PRO'}] \\ \text{OBJ} & [\ ] \end{bmatrix} \end{bmatrix}$$

However, topics are not always objects of complement clauses. They can also be objects of XCOMPs or of the same clause or they can bear other grammatical relations. The possibilities are essentially endless and as such it is undesirable, if not impossible, to list them by a series of annotations.

This is where functional uncertainty comes into play. Functional uncertainty allows annotations of the general type in (8).

(8) (↑DF)=(↑BODY BOTTOM)

DF stands for discourse function, e.g., topic or focus. BODY is a path through the f-structure defined by a regular expression of grammatical functions, and BOTTOM is a grammatical function or set of grammatical functions.

To be more concrete, (9) provides the functional uncertainty equation needed for English topics (Huang 1992a, 1992b).

(9) (↑TOP)=(↑{COMP,XCOMP}* (GF-COMP))

The DF in (9) is restricted to topics since the position in (5) is only used for topics. The body is any number of COMPs and/or XCOMPs, i.e.,

the grammatical function can be part of the mother f-structure or part of any COMP or XCOMP nested within it. The bottom specifies that the topic can be any grammatical function (GF), except for COMP.[2] This blocks the topicalization of constituents corresponding to COMPs but allows SUBJ, OBJ, etc. So, (5) with the annotation in (9) allows sentences like those in (10).

(10)		Sentence:	Body/bottom:
	a.	Inna, Boris loves.	(↑OBJ)
	b.	Inna, Ivan says that Boris loves.	(↑COMP OBJ)
	c.	Inna, Ivan wants to love.	(↑XCOMP OBJ)
	d.	Inna, Ivan says that Boris wants to love.	(↑COMP XCOMP OBJ)
	e.	Inna, Ivan claims loves Boris.	(↑COMP SUBJ)
	f.	That box, Ivan put the cake into.	(↑OBL)

Thus, functional uncertainty allows us to state the relationship between the discourse functions of certain constituents and their grammatical relations.[3] This relationship is constrained by the choice of DF, BODY, and BOTTOM which varies from language to language and within a language, e.g., the annotation necessary for focused items could be different from that of topics. This dependency allows the DF corresponding to a particular constituent to be located in a different f-structure than its GF, and more generally for the discourse function not to be tied to a particular grammatical function. For further details on functional uncertainty see Kaplan and Zaenen 1988a, 1988b, Kaplan and Maxwell 1988, Dalrymple 1990, and Huang 1992a, 1992b.

9.2 Positional Topic and Focus

The basic data for Russian topic and focus and their interaction with word order are discussed in chapter 4. The main generalizations are as follows:

- External topics are outside and to the left of CP.

[2] (9) is Huang's rule. Tom Wasow (p.c.) points out that COMPs can be topicalized in English. For example, 'That Clinton is cautious, I don't think anyone has ever claimed'. These constructions are further discussed in Bresnan and Kaplan 1982. Huang's rule can be adapted to include these examples by removing the restriction on the bottom.

[3] Functional uncertainty is generally used to associate positionally signaled discourse functions, usually topics, with their grammatical function and f-structure. Stress can also be used to mark focus. One idea is that FOC can be associated with any node in the c-structure. The association of FOC with this node constrains the phonological form so that the node must be stressed. See Fenstad et al. 1985 for discussion of the interaction of focus with the phonology in LFG.

- Internal topics appear to the left of the finite verb.
- Internal topics follow material in the projection of C^0.
- Foci with sentence stress usually appear preverbally, following topics.
- Items in sentence-final position are focused R $\longrightarrow$ L with the falling tone of unmarked intonation on the rightmost item.

There are two major concerns for representing the correlations between discourse function and Russian word order in LFG. The first is what the annotated c-structure rules should look like for these constructions. The second is where the information as to discourse function interpretation is stored (section 9.4).

9.2.1 External Topics

The first discourse function structure to be considered is that of external topics since they are syntactically the simplest to deal with (section 4.3.2, 5.1.1.1). The analysis of external topics, or left-dislocation constructions, is fundamentally the same in LFG as it is in GB. The external topic is sister to CP under a node E(xpression), shown in (11). Since this is the only c-structure rule which mentions E, this node can only appear in main clauses, and correspondingly external topics can only appear in main clauses, as desired. Also, the rule is not recursive, allowing for only one external topic per clause.

(11) $\quad \text{E} \longrightarrow \quad \underset{(\uparrow \text{E-TOP})=\downarrow}{\text{XP}} \quad \underset{\uparrow=\downarrow}{\text{CP}}$

There is one particularly interesting feature of the rule in (11): there is only one functional annotation associated with the node corresponding to the external topic.[4] That is, there is an annotation that states that the node corresponds to the E-TOP of the f-structure, but none for grammatical function. The maximal projection in this position plays no part in satisfying the argument structure requirements of the clause's predicate. So, for example, this XP cannot be the subject of the clause, although it can be coreferent with it. This contrasts sharply with internal topics and foci, which have not only a discourse function role, but are also assigned a grammatical function relating them to the predicate.

An example of an external topic is shown in (12). The external topic *Gleb* is not an argument of the predicate *ljublju*. The arguments of the verb are the subject *ja* and object *ego*. However, the external topic is coreferent with the object pronoun. This coreference is not

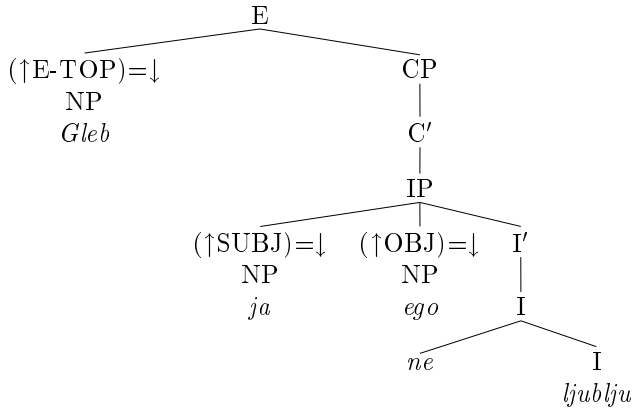
[4] An additional annotation might be desired to assign nominative case to the external topic (section 8.1), but none is needed for grammatical function.

forced by the grammar, and external topics appear in constructions without a coreferent pronoun.[5]

(12) [Gleb], ja ego ne ljublju.
Gleb I him not like
'Gleb-E.TOP, I don't like him.'

The c-structure and f-structure corresponding to (12) are shown in (13) and (14) respectively. The first node of the tree in (13) introduces the external topic into the phrase structure.

(13)



[5] An example of this is shown in (i).

(i) [Opera], net drugogo vida muzykal'nogo iskusstva, kotoryj
opera not other type musical art which
privlekal by k sebe takoe vnimanie.
would attract to itself such attention
'Opera-E.TOP, there's no other form of musical art which would attract such attention.' (Gundel 1988:189)

(14)
$$\begin{bmatrix} \text{E-TOP} & [\text{PRED} \quad \text{'Gleb'}] \\ \text{PRED} & \text{'love<SUBJ,OBJ>'} \\ \text{NEG} & + \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{PRO} \\ \text{PRS} & \text{1st} \\ \text{NUM} & \text{SNG} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{PRO} \\ \text{PRS} & \text{3rd} \\ \text{NUM} & \text{SNG} \\ \text{GND} & \text{MASC} \end{bmatrix} \end{bmatrix}$$

In (13) the external topic appears under E as sister to CP. The annotation above the NP daughter of E states that it is the E-TOP of the f-structure; there is no annotation for grammatical function since the NP has none. Correspondingly, in the f-structure in (14) the value of the E-TOP attribute is the f-structure corresponding to the information provided by that NP node; shown here as 'Gleb'. This f-structure is *not* the value for any other attribute, e.g., it is not the value for the object, even though it is coreferent with the object. The subject and object pronouns are sister to I′ in (13). Since they are internal topics, the annotations on these rules should result in this information being passed to the f-structure as well: that is the subject of the next section.

9.2.2 Internal Topics

Consider the now-familiar constructions in (15) which contain internal topics.

(15) a. [Na stole] stojala lampa.
on table stood lamp
'There was a lamp on the table-TOP.' (Chvany 1973:266)

b. Kogda [Lermontov] rodilsja?
when Lermontov born
'When was Lermontov-TOP born?'

These topics must appear after the elements in the projection of C^0, such as the wh-phrase *kogda* in (15b), and before the finite verb which heads I′ and precedes the VP. There are two places these topics could appear in the phrase structure. The first is to adjoin them iteratively to IP, which is the account proposed in section 5.1.1. The second is

to arrange them in a flat structure as sister to I′, an option which is viable in LFG c-structure, but which would be unexpected in GB since it runs contrary to common conceptions of adjunction and the binary branching of X′ syntax.[6] These rules are shown in (16) and (17) respectively. Note that the equation ↓∈(↑TOP) designates the XP as a member of the set of topics, i.e., there could be more than one topic per clause.

(16) IP ⟶ XP (↑GF)=↓ ↓∈(↑TOP) IP ↑=↓

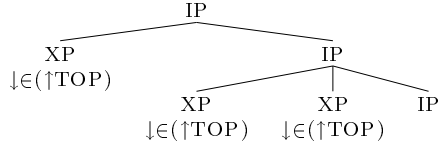
(17) IP ⟶ XP* (↑GF)=↓ ↓∈(↑TOP) I′ ↑=↓

The c-structure trees produced are shown in (18). (18a) is the only type of tree produced by (15). That is, (15) results in a binary branching structure above IP in which each topic is sister to IP.[7] (18b) is the type of tree resulting from (16): all of the topics are under the same IP node, sister to I′.

[6] There is a third possibility which allows topics to adjoin to IP (or I′) either in a hierarchical structure or a flat structure or some mixture thereof; this rule is shown in (i). The difference between the output of this rule and the one which produces a strictly hierarchical structure is only apparent when there is more than one topic. Topics can be in a binary branching structure if only one XP is sister to IP, or several topics can be sisters to one another under the same IP if the * option is taken; (ii) is one of the possible structures produced by (i). (i) can also produce the structures in (18a) and (18b). There seems to be no empirical reason to have some of the topics appear as sisters to one another while others are in a hierarchical structure. As such, there is little justification for a rule like (i) which allows these varying structures.

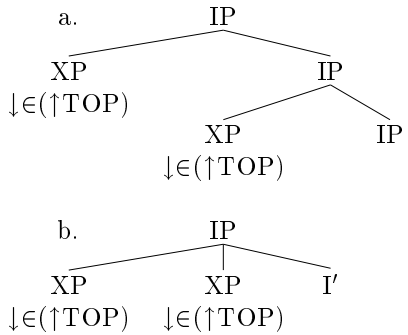
(i) IP ⟶ XP* (↑GF)=↓ ↓∈(↑TOP) IP ↑=↓

(ii)



[7] It would also be possible to produce a binary branching structure with adjunction to I′, instead of to IP. The c-structure rule for this would be: I′ ⟶ XP I′. This rule would allow for a SpecIP position, whose nature would be specified by an IP expansion rule. This structure does not conform with those expected in X′ syntax.

(18)



Thus, there is a choice between (15) in which the topics are hierarchically arranged and (16) in which they are sisters. At present, there is no empirical reason to have a hierarchical structure for topics. For example, they do not have scope effects relative to one another which might reflect a c-command relation. (See Kiss *to appear* for a similar observation concerning Hungarian topics; the structure here collapses her structure in which one topic appears in SpecTP while the rest adjoin to TP.)[8]

So, pending evidence to the contrary, I assume that topics are sisters and are introduced into the c-structure by (16). However, this is a relatively arbitrary choice, and further evidence may argue in favor of (15).

[8]The choice among these structures is analogous to more general current debate about phrase structure. Although there are relatively few proposals arguing for flat structures, there are a number of issues concerning the nature of X′ syntax. For example, Kiss *to appear* places topics in SpecTP/MP; however, when additional topics are present, they are adjoined to TP/MP. This suggests that Specifier positions need not be privileged and differentiated from adjoined positions. A similar, formal proposal that X′ and XP levels are indistinguished is found in Speas 1990 (section **??**). Kayne's 1994 proposes another approach to this problem: there are only binary right branching structures and no adjunction. Although this approach eliminates the issue of whether specifiers are differentiated from adjunction, it results in the positing of numerous functional projections and counterintuitive structures for head-final languages.

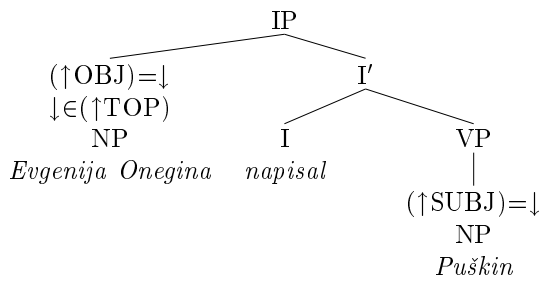
The issue of flat vs. branching structures is somewhat tangential to the positing of X′ syntax. A basic X′ syntax requires that a maximal projection XP be headed by a head of type X. However, the configuration of the complements need not be hierarchical, i.e., positing only binary branching structures is a theoretical choice. Positing only binary branching structures is a stronger hypothesis, but it remains to be seen whether the hierarchical structures that it requires can always be supported by the data, i.e., whether c-command relations exist between all constituents. For example, Bresnan 1994 argues that weak crossover effects usually attributed to c-command may be the result of other factors such as linear precedence. However, in the realm of discourse functions, it may be that the order of multiple topics (or multiple foci) reflects their relative scope, e.g., the outermost topic might have widest scope; in turn, these scope relations, or their absence, may be reflected as relative hierarchical or linear ordering.

These flat structures bring up an issue which deserves further research: what theoretically and empirically motivated constraints hold on the types of LFG phrase structure rules. For example, flat structures might have a limited distribution. One way to limit this would be that only in certain places could flat structure be introduced, e.g., under S. Another would be to restrict flat structures to situations in which adjunction structures occur in X′ syntax.[9] Regardless of how this limitation is accomplished, there are two constraints discussed here which ideally hold for any account. The first is the optionality of nodes introduced in the phrase structure (with the restriction on adjunction structures); any apparently obligatory positions must be independently required to appear, e.g., by well-formedness constraints on the f-structure. The other is that all nodes must be overt in at least one construction in the language; this prevents abuse of the optionality in an effort to force the phrase structure rules to resemble X′ syntax.

To demonstrate the placement of internal topics, a sentence like (19) has the c- and f-structures in (20) and (21). Since there is only one c-structure rule annotation which introduces internal topics, all topicalized constituents appear in this position. This contrasts with foci which are not dependent in this fashion on the annotations of c-structure positions (section 9.2.4). The verb in (20) is in I^0, not V^0, because it is finite and finite verbs are of category I^0 (section 10.1).

(19) ['Evgenija Onegina'] napisal Puškin.
Eugene Onegin wrote Pushkin
'Pushkin wrote Eugene Onegin-TOP.'

(20)



[9] Exactly which structures would be permitted under this account would depend on the precise nature of the X′ projections. For example, depending on the way in which arguments of the verb are projected into the phrase structure, this structure might include adjunction to V′ or not.

(21)
$$\begin{bmatrix} \text{PRED} & \text{`write<SUBJ,OBJ>'} \\ \text{TOP} & \left\{ \begin{bmatrix} \text{PRED} & \text{`E.O.'} \end{bmatrix} \right\} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{`Pushkin'} \end{bmatrix} \\ \text{OBJ} & [\ ] \end{bmatrix}$$

9.2.3 Preverbal Foci

As seen in chapter 5, there are two ways of encoding focus. One is via the use of sentence stress to mark the focused element; this results in a largely contrastive reading (section 4.2.3).

(22) a. Nad Krakovom [**dožd**i**ček**] nakrapyval.
over Krakow rain drizzled.
'It was drizzling-FOC over Krakow.'

b. Ja [**k Anne**] prišel.
I to Anna arrived
'I visited Anna-FOC.'

Both sentences in (22) have the focused element before the verb, following topicalized constituents, i.e., following *nad Krakovom* and *ja*. Constituents with sentence stress can occur in any position in non-discourse-initial utterances, although they tend to appear preverbally (section 4.3.4.2). As such, it was proposed that SpecIP licensed contrastively focused constituents, and as a result of this interpretation any material in SpecIP bears sentence stress.

Since these focused constituents appear before the verb but after topics, the rule introducing internal topics must be modified. This rule is repeated as (23).

(23) IP $\longrightarrow$ XP*, I′
$(\uparrow \text{GF}) = \downarrow$ $\uparrow = \downarrow$
$(\uparrow \text{TOP}) = \downarrow$

(23) can be modified to introduce a non-topicalized XP between the topics and I′. A constraint must be added to ensure that the XP in this position receives contrastive focus interpretation. This position is assigned a particular type of focus, C-FOC. The new annotated c-structure rule is shown in (24).

(24) IP $\longrightarrow$ XP* XP I′
$(\uparrow \text{GF}) = \downarrow$ $(\uparrow \text{GF}) = \downarrow$ $\uparrow = \downarrow$
$\downarrow \in (\uparrow \text{TOP})$ $\downarrow \in (\uparrow \text{C-FOC})$

However, there is another possibility which abstracts away from the direct annotation of c-structure rules and allows for sentences in which

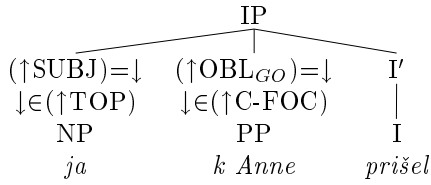
the contrastive focus precedes the topic. The c-structure rule in (23) can be annotated with the functional uncertainty expression DF which encompasses TOP and C-FOC. The ordering among the preverbal elements is captured by a precedence relation between topics and foci; this relation holds for the relative ordering of topics and foci found in the languages examined in section 5.2. However, certain other constraints may override the precedence rule, resulting in the attested structures in which C-FOC precede topics, e.g., when pronominal topics interact with contrastive focus. The annotated c-structure rule and precedence relation are shown in (25) and (26) respectively.

(25) $\text{IP} \longrightarrow \underset{(\uparrow \text{GF})=\downarrow}{\underset{\downarrow \in (\uparrow \text{DF})}{\text{XP*,}}} \quad \underset{\uparrow=\downarrow}{\text{I}'}$

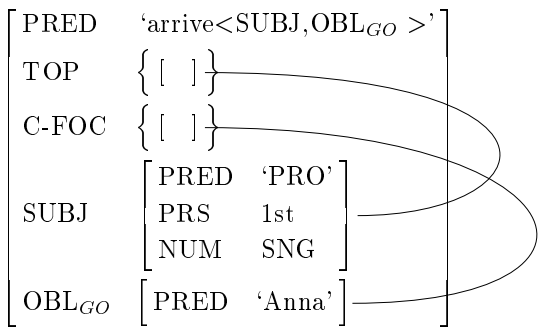
(26) TOP > C-FOC

Implementing these rules, the c-structure of (22b) is shown in (27) and the f-structure in (28). The value of TOP is identical to that of SUBJ, while the value of C-FOC is that of the OBL_{GO}.

(27)



(28)



9.2.4 Right-edge Foci

Examples of right-edge new-information focus are shown in (29). The sentences have Type I non-emotive intonation and the focus appears at the right edge of the clause (section 4.3.4.1).[10]

[10]The final falling tone of neutral intonation marks the right edge of the focused constituent. The left edge of this constituent is not marked, and the scope of the focus is determined by context.

(29) a. Kupila plat'e [Inna].
bought dress Inna
'Inna-FOC bought a dress.'

b. Kolxoz [zakončil uborku urožaja].
kolxoz finished crop harvest
'The kolxoz finished-FOC the crop harvest-FOC.'

c. [Posadil ded repku].
planted old man turnip
'An old man-FOC planted-FOC a turnip-FOC.'

How should we account for the distribution of focused phrases in sentences like those in (29)? The intonation marks the domain of focus (section 5.3). The falling tone of the neutral intonation appears on the last word of the clause. The item marked by this falling tone must be included in the new-information focus. In addition, any items to the left of the final one can be focused as long as they are adjacent to a focused item, i.e., constituents cannot be 'skipped'. For example, in (29b), either *uborku urožaja* or *zakončil uborku urožaja* can be the focus of the sentence; *kolxoz* cannot be included in the focus since it is topicalized.

The VP rules are of importance to this type of focus since the constituents introduced by the VP rule are in clause-final position. These rules are restated below as (30) and (31).

(30) $\text{VP} \longrightarrow \underset{(\uparrow\text{SUBJ})=\downarrow}{\text{XP}},\ \underset{\uparrow=\downarrow}{\text{V}'}$

(31) $\text{V}' \longrightarrow \underset{\uparrow=\downarrow}{\text{V}},\ \underset{(\uparrow\text{GF})=\downarrow}{\text{XP}}$

(31) allows the arguments of the verb to appear in any order after the verb, although there is still a privileged preverbal position, introduced by (29), in which only subjects can appear and which c-commands other VP-internal arguments. The unmarked order for arguments has the subject precede the object. This ordering can be accomplished by a 'rule of thumb' that arguments should be ordered according to their thematic prominence; this proposal is compatible with those made by Rudin 1985 for Bulgarian and Yokoyama 1986 for Russian.

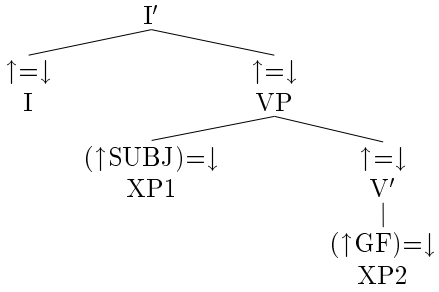
How is the focus interpretation assigned? The focused elements must form a continuous unit from the right edge of the clause. One way to accomplish this is by specifying the right-edge constituents as focused in the c-structure rule, similar to topics and contrastive foci. This is shown in (32).

(32) $V' \longrightarrow$ V (↑=↓) XP* (↑GF)=↓ XP* (↑GF)=↓ (↑FOC)=↓

The rule in (32) correctly places all focused phrases following the non-focused ones. However, although (32) can account for the data in which arguments of the verb are focused, it fails to account for sentences in which the verb is part of the focused material. In fact, there is no intuitively obvious way to modify the VP rule to allow this; the rule would have to guarantee that the verb could be part of the focus only when all of the phrases following it were focused.[11]

This suggests that this type of focus is not assigned via annotations to the c-structure rules. Instead focus is read off the phrase structure as a result of the intonation pattern of the sentence, as was argued for in section 5.3. So, reconsider the c-structure rule in (31) which gave the correct word order but said nothing about which of the items were focused. (31) produces structures like (33) when combined the other c-structure rules. (With finite clauses, the relevant structure begins at I′.)

(33)



For ease of explication, take the focused material in a clause to be marked [+F] and the intonation to mark the right edge of this focused material. This material can be of varying sizes, e.g., just the NP on which the stress falls or both the NP and the I^0. With this in mind, consider (33). The falling tone of neutral intonation falls on XP2. This can correspond to: XP2 being marked [+F]; XP2 and XP1 being marked [+F], so that both arguments are focused; XP2, XP1,

[11] One way to capture these facts is to have two similar VP rules. One would be identical to (32). The other would be similar to (32), but would specify that all of the material in the VP was focused, including the verb, as in (i). Although this is technically possible, the fact that two almost identical rules are needed suggests that a generalization is being missed.

(i) $V' \longrightarrow$ V (↑=↓ (↑FOC)=↓) XP* (↑GF)=↓ (↑FOC)=↓

and I^0 being marked [+F], in which case the verb and its arguments are focused.

Nowhere in the c-structure rules for I′, VP, and V′ is a specification made as to the assignment of focus. The focus interpretation that these sentences receive is dependent on the c-structure in that right-edge constituents can be focused; however, it is not a result of a c-structure annotation in the way that topic and contrastive focus were. Since the scope of the focus depends on the c-structure, it would be convenient to have a way to provide this information to the f-structure. A first approximation might be to say that any mother node can be labeled $\downarrow\in(\uparrow\text{FOC})$. However, this incorrectly allows non-final constituents to be new-information focus.[12] A second possibility would be state that the rightmost node in a phrase structure rule can be annotated $\downarrow\in(\uparrow\text{FOC})$, but this allows NPs within PPs to be focused. A third possibility would be to assume that any mother node of type V^0 or I^0 can be annotated $\downarrow\in(\uparrow\text{FOC})$ (see section 10.3 on how the distinction between V^0 and I^0 can be collapsed). However, consider the c-structure in (33). What happens if XP1 and XP2 are both focused? If the VP constituent which contains the XPs is marked $\downarrow\in(\uparrow\text{FOC})$, then the XPs will be focused. However, the I′ will also be focused because the annotation on the VP is $\uparrow=\downarrow$, which means that any annotation on the VP is equivalent to an annotation on the I′. Thus, if the VP is part of FOC, then I′ is, which means that the verb is also, incorrectly, part of FOC.

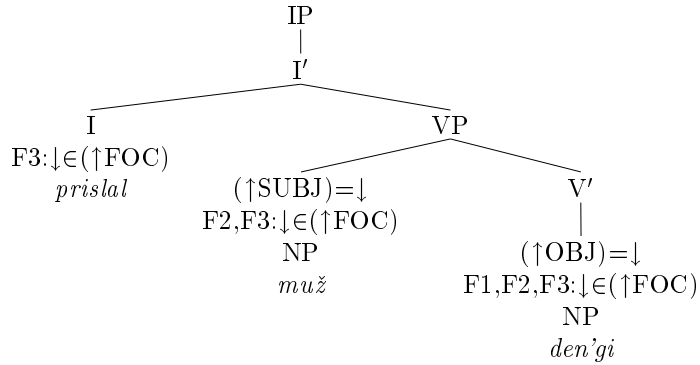
Instead, the scope of the focus depends on the prosody. Leaf nodes in the c-structure can be marked $\downarrow\in(\uparrow\text{FOC})$ from right to left, starting at the right-edge constituent with the falling intonation characteristic of this type of focus. The left edge of the focused material is limited to material within I′, because material outside of I′ is assigned either TOP of C-FOC discourse functions; these discourse functions are incompatible with FOC and have their own characteristic intonation. This account is largely prosodic in nature and relatively independent of syntactic constituency, although the mapping between the prosody and the syntax results in substantial overlap as to what material falls within the domain of the focus.

The c-structure of a sentence like (34) is shown in (35). The object *den'gi* is focused under all possible interpretations of the clause. The subject *muž* can also be focused, resulting in the f-structure in (37) and (38). Finally, the verb can only be focused if both the subject and object are focused, shown in (38).

[12] Such a rule might work for contrastive focus (C-FOC) in more colloquial Russian.

(34) [Prislal muž den'gi.]
sent husband money
'My husband sent (me) the money.'

(35)



(36) F1: $\begin{bmatrix} \text{PRED} & \text{'send<SUBJ,OBJ>'} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED 'husband'} \end{bmatrix} \\ \text{OBJ} & [\ ] \\ \text{FOC} & \left\{ \begin{bmatrix} \text{PRED} & \text{'money'} \end{bmatrix} \right\} \end{bmatrix}$

(37) F2: $\begin{bmatrix} \text{PRED} & \text{'send<SUBJ,OBJ>'} \\ \text{SUBJ} & [\ ] \\ \text{OBJ} & [\ ] \\ \text{FOC} & \left\{ \begin{bmatrix} \text{PRED} & \text{'husband'} \end{bmatrix}, \begin{bmatrix} \text{PRED} & \text{'money'} \end{bmatrix} \right\} \end{bmatrix}$

(38) F3: $\begin{bmatrix} \text{PRED} & \text{'send<SUBJ,OBJ>'} \\ \text{SUBJ} & [\ ] \\ \text{OBJ} & [\ ] \\ \text{FOC} & \left\{ —, \begin{bmatrix} \text{PRED} & \text{'husband'} \end{bmatrix}, \begin{bmatrix} \text{PRED} & \text{'money'} \end{bmatrix} \right\} \end{bmatrix}$

The focusing of a constituent with multiple PREDs is done by creating a set for the value of FOC. The sentence in (36) could involve focusing of just the object *repku*, as indicated by the focus scope marker F1. In addition, focus of the object and subject exclusive of the verb

is possible, as seen by the scope marker F2 and the corresponding f-structure in (37), although such readings with two arguments focused without the verb are relatively rare and often involve sentence stress, resulting in a contrastive reading. (38) shows the f-structure in which the entire sentence is focused.[13] Thus, the scope of new-information focus is determined by annotations to the c-structure, but the position of this annotation within the c-structure is determined by the prosody, not the annotated c-structure rules.

9.3 Wh-Questions

This section provides a sketch of Russian wh-questions. The semantics of these questions, including how scopal differences, which in Russian are often reflected in word order, should be captured is not discussed; this is a result of the interaction of c- and s(emantic)-structure (Halvorsen 1983).[14] Constraints on the order of wh-phrases in Russian are also not analyzed since these restrictions not are syntactically motivated (section 3.6.4).[15]

This section first describes the basic facts to be explained. As in English, Russian wh-phrases appear in initial position, both in matrix and embedded questions.

(39) a. Kogo udaril Boris?
who-ACC hit Boris-NOM
'Who did Boris hit?'

b. Ja ne znaju, kogo Boris udaril.
I not know who Boris hit
'I don't know who Boris hit.'

However, there are several differences between English and Russian wh-questions. There is no subject-aux inversion in Russian, and no difference in word order in matrix and subordinate clauses; compare (39a) and (39b). In addition, wh-phrases can be moved out of infinitival clauses, but not out of finite subordinate clauses, as in (40).

[13] There is a technical difficulty with the f-structure in (38). Since the PRED 'send' is the head of the f-structure, when it is focused, the entire f-structure is focused. This problem arises even if the verb is the only focused item in the f-structure; it is not a result of the way in which the focus interacts with the c-structure. Perhaps this difficulty could be avoided by the mapping between f-structure and semantic-structure.

[14] This is similar to how LF encodes the scoping of wh-phrases.

[15] Rudin 1985 suggests that the ordering of wh-phrases in Bulgarian, which is more fixed than that of Russian, is the result of several 'rules of thumb,' giving priority to animates, to NPs over PPs, etc. Billings and Rudin 1994 make a similar proposal using Optimality Theory to formally rank the constraints. Such a tack might be taken for Russian.

(40) a. Čto vy xoteli kupit'?
what-ACC you want buy-INF
'What did you want to buy?'

b. *Kogo vy skazali čto on uvidel?
who-ACC you say that he saw
'Who did you say (that) he saw?'

Finally, in Russian matrix and embedded questions all the wh-phrases appear in initial position; none of them remain *in situ*. (Contrast the English glosses with the Russian.)

(41) a. Kto kogo udaril?
who-NOM who-ACC hit
'Who hit whom?'

b. Kogda kto prišel?
when who arrived
'Who arrived when?'

The wh-phrases are in SpecCP. Since more than one wh-phrase can appear in initial position, they can either be sisters to one another in the SpecCP position or have SpecCP optionally produce a nested adjunction structure.[16] Although the nested adjunction structure may ultimately prove useful in mapping the relative scope of wh-phrases into the semantics, for our purposes the two approaches are identical, and I adopt the 'sisters' structure, as in (42).

(42) $\text{CP} \longrightarrow \underset{\substack{(\uparrow \text{XCOMP}^* \text{ GF}) = \downarrow \\ \downarrow \in (\uparrow \text{Q-FOC})}}{\text{XP}^*}, \quad \text{C}'$

The rule in (42) states that the CP expands into a series of maximal projections and a C′. Each of the maximal projections is assigned a grammatical function and a discourse function in that it becomes a member of the question's focus (Q-FOC). The Q-FOC feature takes a set as its value; in questions with only one wh-phrase, this set has only one element. Having the value of Q-FOC be a set allows for multiple wh-questions.

(42) allows questions to be formed from infinitival subordinate clauses,[17] as in (40a) where the question word is the OBJ of the XCOMP; this assumes that finite subordinate clauses correspond to

[16] Rudin 1989 concludes that only one wh-phrase is in SpecCP while the rest are adjoined to IP (section 3.6.4). At this point, in addition to the fact that the proper analysis remains an issue in GB, it is not clear whether the arguments for doing this in GB hold for LFG.

[17] (42) allows for questions to be formed out of doubly embedded infinitives, as in (i).

COMPs in f-structure, while infinitivals correspond to XCOMPs.[18] [19] This process can apply iteratively down through nested XCOMPs. As a result, the wh-phrase appears as c-structure sister to the matrix C′, although it can be an argument of an embedded predicate.

Finally, stating in the lexical entry of wh-phrases that they are required to be in a Q-FOC set insures that all wh-phrases appear in sentence initial position. A sample entry is shown in (43).

(43) *kto* 'who'
($\uparrow$PRED) = 'who'
($\uparrow$CASE) $=_c$ NOM
$\uparrow \in_c$ (($\uparrow$XCOMP* $\uparrow$) Q-FOC)

The PRED value for the wh-phrase is 'who' and the constraint equation requires that it be assigned nominative case (other forms have other constraint equations; for example, *komu* is associated with a constraint equation ($\uparrow$CASE) $=_c$ DAT). The last line states that *kto* must be an element of the f-structure corresponding to the mother f-structure's Q-FOC.

Thus, a question like (44) has a c-structure as in (45) and an f-structure as in (46).

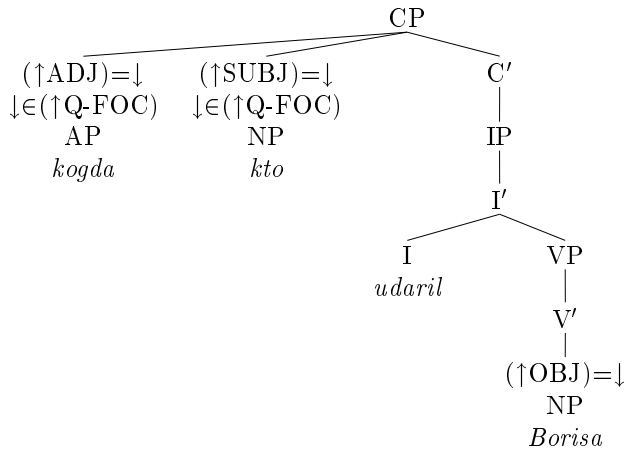
(44) Kogda kto udaril Borisa?
when who hit Boris
'Who hit Boris when?'

(i) Kakie knigi ty xočeš' naučit' čitat'?
which books you want learn-INF read-INF
'What kind of books do you want to learn to read?'

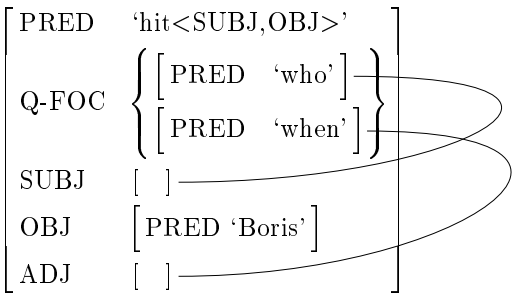
[18] The main difference between XCOMPs and COMPs is that COMPs contain all of the arguments of their predicate, while XCOMPs depend on the outer f-structures for the value of some of their arguments, usually the subject (Bresnan 1982). Neidle 1988 argues that some infinitives in Russian, namely ones in which the object can be coreferent with the subject of the infinitive, are COMPs, not XCOMPs. The examples discussed here are of the XCOMP variety, but the rule in (43) could be modified to take her analysis into account.

[19] The fact that finite subordinate clauses require complementizers while infinitives do not is irrelevant to the grammatical function, i.e., COMP or XCOMP, of subordinate clauses. This can be seen in languages like Icelandic and French in which infinitives can appear with complementizers.

(45)



(46)



The wh-phrases *kto* and *kogda* appear in SpecCP where they are assigned to the Q-FOC. In addition, they are assigned the grammatical functions subject and adjunct respectively; if they were assigned any other GF, the resulting f-structure would be ill-formed since the verb's argument structure would not be satisfied, although the c-structure would be licit. Thus, the use of functional uncertainty correctly constrains the occurrence of wh-phrases so that they always appear in initial position and so that only wh-phrases appear in this position (but see section 10.2).

9.4 Where is DF Information?

Discourse function information is included in the f-structure, as seen in the previous sections (Bresnan 1982, Huang 1992a, 1992b, Kaplan and Zaenen 1988a, 1988b). TOPIC and FOCUS are treated as f-structure attributes which take f-structures corresponding to particular items as their values, similar to the values of SUBJ and OBJ. In fact, the value of TOPIC or FOCUS in a given f-structure is usually identical to the value of one of the grammatical functions. This reflects the fact that when a

constituent has a particular discourse function in the clause, it usually has a particular grammatical function as well. Technically, the value is not represented twice in the f-structure; instead, it is simultaneously the value of both attributes. This TOPIC and FOCUS information then participates in the semantics, which is often assumed to be read off of the f-structure.

However, although discourse function information interacts with the f-structure, it must also appear in semantic-structure. In languages like Malay, which are truly topic-prominent in that many syntactic properties refer to topic and not subject, there is empirical evidence that topic information is relevant to the f-structure (Alsagoff 1992). In addition, Bresnan and Mchombo 1987 discuss how topic and focus interact with the agreement system in Chicheŵa. (See also Kuno 1972 on how the notion topic/theme is of importance in English syntax.) Not all languages make such wide reference to topic in the syntax. This is not an argument against this information being in the f-structure for all languages, but it requires further consideration. Placing focus information in the f-structure is perhaps less justified since there are few syntactic processes which refer to foci;[20] one does not hear of focus-prominent languages. The syntactic position of focus is often important, but this is properly a statement about the c-structure, not the f-structure. However, some types of focus, such as that associated with wh-questions have syntactic consequences and as such belong in the f-structure.

Halvorsen 1983 and Kaplan 1987 discuss how the semantics interacts with the other components of the grammar in LFG. Under their conception, there is a s(emantic)-structure which, like f-structures, is composed of attribute-value matrices.[21] This s-structure interfaces with the formal semantics to derive the meaning of the sentence. S-structure is built off of the f-structure of the clause. C-structure is only indirectly relevant for the description of s-structure in that c-structure provides constraints on f-structure; however, there is no direct mapping between c- and s-structures. Halvorsen and Kaplan 1988 modify this conception so that both the c- and f-structures are relevant for the s-structure. This allows c-structure information which is not present in f-structure, such as the positioning of adverbs, to play a role in

[20] The interaction of focus with the 'syntax' is usually with the phrase structure. For example, the requirement that the focus of a cleft construction be within the postcopular constituent is a statement about phrase structure for LFG, not about the syntax as it relates to f-structure information.

[21] This s-structure is different from a(rgument)-structure. S-structure includes the information found in LF in GB and perhaps other semantic information.

the s-structure. Under such a system, certain information about topic and focus need not be present in the f-structure in order to appear in the s-structure. Kaplan 1987 also posits a discourse-structure module which, presumably, would contain topic and focus information. This discourse-structure is parallel to the s-structure and both contribute information to the meaning of the sentence. Dalrymple (p.c.) has also suggested that discourse function information might comprise a separate module. Thus, topic and focus information in the c-structure is directly relevant to s-structure (and/or discourse-structure) which in turn defines the meaning of the sentence. As such, only the topic and focus information relevant to the f-structure need be present there.

This conception exploits the separation of phrase structure and grammatical relations in LFG. Discourse functions are structurally encoded in a number of languages. This indicates that, as usually assumed, discourse function information can be annotated on the c-structure. In addition, it is commonly assumed that discourse functions, particularly focus, interact with the semantics to indicate their relative scopes over the sentence. As such, some, if not all, discourse function information must be included in semantic-structure. When necessary, this information can be mapped from the c-structure directly to the s-structure via σ-projections which, for our purposes, work like the mapping between c- and f-structures. However, the inclusion of this information in c- and s-structure and the mapping between them does not imply that discourse function information also must be present in the f-structure. Only that information relevant to the syntax, e.g., information concerning topics in Malay, is also contained in the f-structure.

The question of where the topic and focus information is stored is tangential to determining the phrase structure of these constructions. Given a particular c-structure for discourse function constructions, the annotations can be written in such a way as to ensure that the relevant topic and focus information is represented where it should be. That is, the annotations are not restricted to statements about the f-structure but can also be made to refer to the s-structure. As the interaction of topic and focus with the syntax, semantics, and phrase structure are studied in more detail, the proper relations between c-structure, f-structure, and s-structure will become more obvious, and it will be possible to posit principled constraints on their interaction in both LFG and GB.

To conclude, this chapter examined the interaction of the phrase structure and discourse functions in LFG. In particular, it was con-

cerned with constructions in which a particular position is associated with a particular discourse function or functions. The basic phrase structure proposed follows an X′ schema in which all nodes are potentially optional; the ordering of these nodes is accomplished by independent linear precedence rules. Annotations on the phrase structure require that constituents in these positions bear particular grammatical functions and discourse functions. In many cases, these annotations employ functional uncertainty to allow for a range of possible functions. For example, the preverbal constituents under IP can be of any grammatical function and either topic or contrastive focus. However, not all topic and focus are determined by annotated c-structure rules, i.e., new-information focus is constrained only by the right-edge constituency of the clause; the c-structure rules do not make reference to it. The structure of wh-questions was also discussed. The interaction of wh-questions with other Russian constructions is discussed in more detail in section 10.2. In this chapter it was proposed that wh-phrases are constrained to appear in SpecCP. This constraint is the result of the interaction of the lexical entry for wh-phrases and the annotation on the specifier position of CP. Finally, it was suggested that some discourse function information plays a role in the f-structure, while other discourse function information may not. Many of the syntactic aspects of discourse functions appear to be constraints on their position in the phrase structure, as witnessed by the c-structure annotations, and when necessary the discourse function information itself can be passed directly to the semantics and generally to the structure of the discourse. The following chapter continues to explore the nature of and constraints on Russian phrase structure in LFG.

10

Capturing Head-Movement

In section 3.5, head-movement accounted for the distribution of finite and infinitival verbs and that of items in C^0 in *li* yes-no questions. In a theory with movement, the verb can be base-generated in V^0 where it theta-marks its arguments and then moved under certain conditions: for example, to I^0 in order to combine with the inflectional features or, having undergone V^0-to-I^0 movement, to C^0 to host the clitic *li*. However, in LFG the verb is base generated in its position in the c-structure, and for each position in which the verb appears, we must insure that that position is only compatible with the appropriate f-structures.

The intuition behind how to do this is identical to that behind head-movement. The type of features that triggered head-movement are the ones used to constrain the c- and f-structures. The difference between what appears in I^0 and what in V^0 is morphosyntactic (Kroeger 1991): that of an inflected verb and that of an infinitival (there are a few subtleties involving predicate adverbs (section 10.1.4)). What appears in C^0 is an inflected verb that is also the focus of a *li* question. The implementation of this idea is discussed in the remainder of this chapter.

10.1 What Can be in I^0

At an intuitive level, finite verbs appear in I^0, while infinitives appear within the VP in V^0. This distribution is overt in imperfective futures in which a finite auxiliary precedes the infinitive 'main' verb, as in (1).

(1) Ja budu čitat' knigu ves' den'.
I will-AUX read-INF book all day
'I will read the book all day.'

The imperfect future in (1) is formed by a tensed auxiliary *budu* and the infinitive of the main verb *čitat'*. The auxiliary is in I^0 and the infinitive in V^0.[1]

The problem is as follows: tensed verbs only appear in I^0 (or C^0), while untensed forms only appear in V^0; imperfective futures involve a single clause and single f-structure, but a tensed and an untensed verb form. The obvious difference between the material that appears in I^0 and that which appears in V^0 is that the former involves one class of verbal suffixes while the latter another. This difference can be exploited to change the category of the underlying verb stem. After the addition of verbal suffixes, verbs with finite suffixes are of category I^0, while those with infinitival stems are of category V^0. This suffixation and category specification occurs in the lexicon as part of the morphological processes that form verbs. (2) shows the lexical entries for a verb stem and two relevant verbal suffixes. Verb stems are of category V^{-1}. These stems must combine with either an infinitival suffix or a tense suffix to be of a category suitable for insertion into the c-structure. The infinitival suffix is of type V^0/V^{-1}: it takes a V^{-1} item, i.e., a verb stem, and turns it into a lexical item of category V^0. In contrast, the tense suffixes are I^0/V^{-1}; they take a V^{-1} item and turn it into a lexical item of category I^0.

(2)

čitaj-	*-t'*
CAT:V^{-1}	CAT:V^0/V^{-1}
(↑PRED) = 'read<SUBJ OBJ>'	

-u
CAT:I^0/V^{-1}
(↑TNS) = PRS
(↑SUBJ PRS) = 1st
(↑SUBJ NUM) = SNG

The root, *čitaj-*, provides the PRED value for the form (as well as information about aspect and morphological information of use in the lexicon, which are not shown here). The morphological category information is used within the lexicon to insure that it combines correctly with affixes.[2] The suffix *-t'* forms an infinitive from a verb stem; the resulting infinitive are CAT:V^0, shown in (3a). The suffix *-u* forms a first

[1] The data presented in sections 3.1, 3.2, and 3.4 support this distribution.

[2] The category notation in (2) is used purely for explanatory purposes; it is not intended to be a developed morphological theory. There has been recent discussion of developing morphological-structure to encode morphologically relevant information, thus freeing up f-structure, which currently contains such information, and allowing for mismatches between the syntax and morphology.

person singular present tense form of the verb which is CAT:I^0, shown in (3b). Here, we are concerned only with the final verb form used in lexical insertion. Note that '=' denotes information relevant to the f-structure, while ':' denotes information relevant to the c-structure.

(3) a. *čitat'* (↑PRED) = 'read<SUBJ OBJ>'
CAT:V

b. *čitaju* (↑PRED) = 'read<SUBJ OBJ>'
CAT:I
(↑TNS) = PRES
(↑SUBJ PRS) = 1st
(↑SUBJ NUM) = SNG

The category information forces finite verbs to appear in I^0 and infinitives to appear in V^0, presuming that the condition on lexical insertion into the c-structure is that category types match. The interesting case is that of the imperfective futures, which involve both a finite auxiliary and an infinitive main verb (section 10.1.2). The main idea is that the auxiliary appears in I^0 and the main verb in V^0 by virtue of their CAT feature. This results in the desired ordering since I^0 precedes V^0 in the c-structure rules. There is no difficulty with conflict of TNS values if infinitives have no value for TNS.[3]

10.1.1 Simplex Verb Forms

Before looking at specific c- and f-structures for imperfective futures, the structures of simplex verb forms is discussed in more detail. (4) is a simple past tense sentence. Present and perfective future sentences are similar to the past tense ones.

(4) Ona pročitala knigu.
she read book
'She read the book.'

The sentence consists of a pronominal subject *ona*, a past tense perfective verb *pročitala*, and a direct object *knigu*. The entries for these lexical items are shown in (5).

[3]Steve Franks (p.c.) suggests that this solution may pose a problem for Polish. In Polish there are two ways to form these futures. The first is with an infinitive, as in Russian. The second involves using an *-l* participle in place of the infinitive. Initially, this appears to be a problem since *-l* participles would appear to bear a TNS value of their own, e.g., in Russian these forms mark past tense. One possibility is that in Polish these participles are, perhaps optionally, unmarked for TNS. However, the this proposal is tentative and thorough study of the Polish tense system is necessary both to solve this problem and to shed light on the structure of compound tenses in general.

(5)

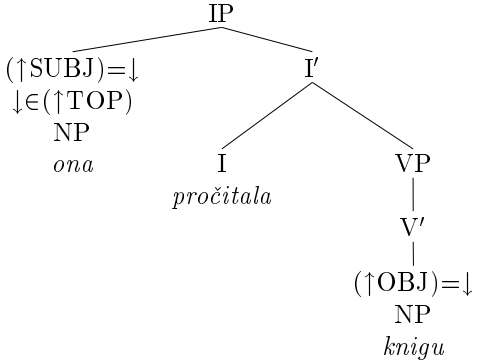
pročitala	*ona*
CAT:I	CAT:NP
(↑PRED)='read<SUBJ OBJ>'	(↑PRED)='PRO'
(↑TNS)=PAST	(↑CASE)=$_c$ NOM

knigu
CAT:N
(↑PRED)='book'
(↑CASE)=$_c$ ACC

All nouns are required to be assigned a specific case: NOM for *ona* and ACC for *knigu* (section 8.1). The verb is designated as CAT:I^0. This information is relevant for lexical insertion into the c-structure, but plays no role in the f-structure.

The c-structure in (6) corresponds to the sentence in (4). Nodes without annotations are understood as being labeled ↑=↓.

(6)



Since the verb is tensed and thus CAT:I^0, it appears in I^0. The headless VP is the equivalent of V^0 to I^0 movement having occurred. If a verb were to appear in V^0 in addition to the one in I^0, the resulting f-structure would have two PRED values, one provided by the verb in I^0 and one by the verb in V^0. These PRED values cannot unify, and the f-structure would be ill-formed. So, no verb can appear in V^0 in (6) (see section 10.1.2 for instances where verbs can appear in both I^0 and V^0).

The structure in (6) at first appears odd in that there are headless projections, namely there is no V^0 within the VP (Zaenen and Kaplan n.d.). However, this is a necessary type of construction. The appearance of headless projections is constrained so that any given projection in a language may exist only if there is at least one construction in which that projection has a head. So, since there are constructions in Russian where V^0 heads the VP, e.g., the imperfective futures, it is

possible to posit a VP. For most constructions, the head of the XP is necessary to produce a well-formed f-structure. (7) is the f-structure corresponding to the information in (5) and (6).

(7)
$$\begin{bmatrix} \text{PRED} & \text{`read<SUBJ,OBJ>'} \\ \text{TNS} & \text{PAST} \\ \text{TOP} & \left\{ \begin{bmatrix} \text{PRED `PRO'} \end{bmatrix} \right\} \\ \text{SUBJ} & [\ ] \\ \text{OBJ} & \begin{bmatrix} \text{PRED `book'} \end{bmatrix} \end{bmatrix}$$

As a side note, it is possible to pro-drop the subject in a sentence like (4). The pro-dropped sentence has a different c-structure from (6) because there is no subject NP. However, the f-structure for such a sentence is identical to that in (7). The difference is that all of the values for the SUBJ come from the verb itself, including the PRED value.[4] Tensed verbs have an optional specification of ((↑SUBJ PRED)='PRO'), in addition to the person, number, and gender specifications that are always instantiated.[5]

10.1.2 The Imperfective Future

Having seen how simplex inflected verbs work, consider the imperfective future shown in (8).

(8) Ja budu čitat' knigu.
I will read-INF book
'I will be reading a book.'

First consider what type of f-structure this sentence must have. Both the main verb and the auxiliary provide information for the f-structure. The main verb provides the PRED value, in this case 'read', while the auxiliary provides information about the tense (TNS) of the sentence. Thus, both the I^0 and the V^0 node are labeled ↑=↓, as would be expected if these positions are heads of the sentence.

This means that the information provided by the auxiliary and the main verb must unify since both are heads of the clause and provide information for the same f-structure. The potential difficulty comes when considering the TNS values of the two forms: the auxiliary's TNS value is future (FUT), while that of the main verb would appear to be infinitive or untensed. This creates a conflict in the construction

[4] There is another difference in that pro-dropped subjects are always topics, while pronominal subjects need not be (section 4.1.3).

[5] Unlike Irish (McCloskey and Hale 1984), pro-drop in Russian is not obligatory in neutral contexts; in fact, overt subject pronouns are the rule, not the exception.

of a well-formed f-structure. To avoid this, it might seem desirable to have the infinitive form an XCOMP, with the auxiliary providing the outer f-structure and the infinitive providing the XCOMP (Bresnan and Kaplan 1982). If this were the case, the two TNS values would be in separate f-structures and thus would not conflict. (9) is an f-structure of this type.

(9)
$$\begin{bmatrix} \text{PRED} & \text{`PROGRESSIVE<SUBJ,XCOMP>'} \\ \text{TNS} & \text{FUT} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{`PRO'} \\ \text{NUM} & \text{SNG} \\ \text{PRS} & \text{1st} \end{bmatrix} \\ \text{XCOMP} & \begin{bmatrix} \text{PRED} & \text{`read<SUBJ,OBJ>'} \\ \text{TNS} & - \\ \text{SUBJ} & [\quad] \\ \text{OBJ} & [\text{PRED} \quad \text{`book'}] \end{bmatrix} \end{bmatrix}$$

However, this XCOMP structure is unnecessary for Russian auxiliary constructions. Consider the perfective future in (10).

(10) Ja pročitaju knigu.
I read book
'I will read the book.'

The future of perfective verbs is a simplex form; there is no auxiliary like that found in the future of imperfective verbs, as in (8). It seems undesirable to posit an XCOMP structure for perfective futures when the apparent c-structure is a simplex form. However, since the perfective and imperfective futures are identical, except for aspect, one could argue that if the imperfective future involves an XCOMP, so should the perfective future.

Also, in the past tense, there is no apparent c-structure difference between perfectives and imperfectives as far as auxiliaries are concerned: both are simplex forms.

(11) a. Ja čitala knigu.
I read-IMP book
'I was reading the book.'

b. Ja pročitala knigu.
I read-PERF book
'I read the book.'

The past imperfective is a simplex form, as in (11a), as is the past perfective, as in (11b). For these reasons, I do not posit differing f-structures for perfectives and imperfectives, i.e., I do not posit an XCOMP for the auxiliary constructions (Andrews 1990).

The sentence in (12) consists of a subject *Anna*, an auxiliary *budet*, an infinitive *čitat'*, and a direct object *knigu*. The relevant lexical entries are in (13).

(12) Anna budet čitat' knigu.
Anna will read-INF book
'Anna will be reading a book.'

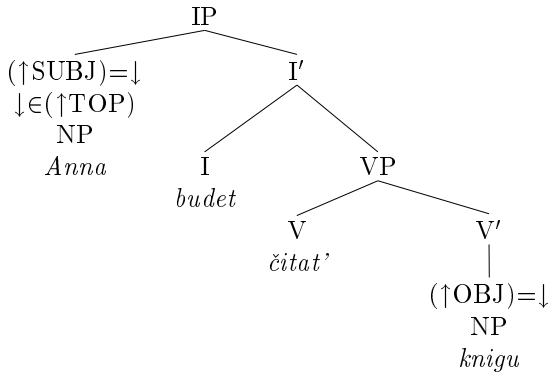
(13)

Anna	*budet*
CAT:NP	CAT:I
(↑PRED)='Anna'	(↑TNS)=FUT
čitat'	*knigu*
CAT:V	CAT:N
(↑PRED)='read<SUBJ OBJ>'	(↑PRED)='book'

The auxiliary *budet* has no PRED value; its only purpose is to mark tense. In contrast, the infinitive *čitat'* has no specification for TNS, i.e., infinitives are tenseless, but it does have a PRED value. So, there is no problem with unification.

(14) shows the c-structure for (12). Since *budet* is CAT:I^0, it must appear in I^0. The infinitive *čitat'* is CAT:V^0 and thus appears in V^0. This structure differs from that of the simplex verbs in that the V^0 head of VP is filled, as well as I^0.

(14)



The annotated c-structure in (14) and the lexical entries in (13) correspond to the f-structure in (15).

(15)
$$\begin{bmatrix} \text{PRED} & \text{`read<SUBJ,OBJ>'} \\ \text{TNS} & \text{FUT} \\ \text{TOP} & \left\{ \begin{bmatrix} \text{PRED `Anna'} \end{bmatrix} \right\} \\ \text{SUBJ} & [\ ] \\ \text{OBJ} & \begin{bmatrix} \text{PRED `book'} \end{bmatrix} \end{bmatrix}$$

Despite having two verbal forms in the c-structure, an auxiliary and an infinitive, the f-structure has only one verbal predicate; that is, there is no XCOMP containing the infinitival predicate, as has sometimes been proposed for English auxiliaries.[6] The only difference in this f-structure and the one for perfective futures is that of aspect, which is not indicated here.

10.1.3 Blocking

The proposal discussed above correctly derives the imperfective futures. However, it would seem to derive the structures in (16), in addition to the grammatical structures in (17).[7]

(16) a. *Ja budu pročitat' knigu.
I will read-INF book
'I will read the book.' (perfective)

b. *Ja byla (pro)čitat' knigu.
I was read-INF book
'I was reading (read) the book.' (perfective/imperfective)

(17) a. Ja pročitaju knigu.
I read-1.SG book
'I will read the book.' (perfective)

b. Ja (pro)čitala knigu.
I read-SG.FEM book
'I read the book.' (perfective/imperfective)

In (16a) the future auxiliary has been combined with a perfective infinitive, and the resulting form is ungrammatical. However, the f-structure is well-formed since the auxiliary would provide the TNS feature and

[6] Andrews 1990 posits such a structure for English 'do' to prevent two tensed verbs from appearing in 'do' support sentences. However, this would not be necessary if TNS is assigned an index (see below).

[7] Steven Franks (p.c.) notes that the imperfective, periphrastic, future is a problem for most accounts and suggests that the fact that *byt'* 'be' only takes imperfective complements may in part be a mater of selection. Further study of the copular use of *byt'* may help to clarify these problems (Kondrashova 1994). See fn. 3 on the future in Polish.

the infinitive the PRED value. In (16b) the past auxiliary has been combined with an (im)perfective infinitive and the result is ungrammatical. However, it should be grammatical for the same reason that (16a) should be. How can forms like those in (16) be prevented? Andrews' 1990 Morphological Blocking Principle can account for this. He states the principle as follows:

> Suppose the structure S has a preterminal node P occupied by a lexical item l_1, and there is another lexical item l_2 such that the f-structure determined by the lexical entry of l_1 properly subsumes that determined by the lexical entry of l_2, and that of l_2 subsumes the f-structure associated with P in S (the complete structure, after all unifications have been carried out). Then S is blocked.

There is a single lexical entry that corresponds to the auxiliary-infinitive constructions, e.g., those in (17). This lexical entry blocks the auxiliary-infinitive constructions. Blocking depends on the auxiliary's not taking an XCOMP because, if the auxiliary-infinitive construction formed an XCOMP in the f-structure, the f-structures would not be subsumed and blocking would not apply. In contrast to the forms in (16), the imperfective futures are not blocked because there is no single lexical item corresponding to the auxiliary-infinitive construction.

10.1.4 Predicate Adverbs

A construction related to the distribution of verbal elements is that of predicate adverbs (section 3.5.2). Predicate adverbs are impersonal constructions whose tenses are marked by auxiliaries. This tense marking is overt in the past and future and phonologically null in the present. Their experiencer, when overt, appears in the dative case. Many predicate adverbs are modal and take infinitival complements, as in (18a).

(18) a. Ivanu trudno bylo čitat'.
Ivan-DAT difficult-PRED.ADV was read-INF
'It was hard for Ivan to read.'

b. Mne xolodno.
me-DAT cold-PRED.ADV
'I am cold.'

Interestingly, both the predicate adverb and the auxiliary can be in I^0. This can be seen via yes-no question formation in which the only heads that appear in initial position are those that can be in I^0 (sections 6 and 10.2). When yes-no questions are formed from predicate adverb constructions, the auxiliary can appear in initial position (section 3.5.2). What is more surprising is that the predicate adverb can appear in this position.

The lexical entries for the predicate adverb construction in (18a) are shown in (19).[8]

(19)

trudno	*Ivanu*
CAT:I	CAT:NP
(↑PRED)='difficult<OBL$_{GO}$ XCOMP>'	(↑PRED)='Ivan'
(↑OBL$_{GO}$)=(↑XCOMP SUBJ)	

bylo	*čitat'*
CAT:I	CAT:V
(↑TNS) = PAST	(↑PRED)='read<SUBJ>'

Both the auxiliary and the predicate adverb are CAT:I^0. This allows them to appear in I^0 simultaneously via the I^0 expansion rule in (20). Only the auxiliary *bylo* is specified for TNS; so, there is no conflict in the f-structure. In theory, the expansion rule in (20) allows for infinite recursion of I^0 elements. However, whenever more than two such elements appear, there are two TNS values or two PRED values. These cannot unify, and the resulting f-structure will be ill-formed. As a result, the rule in (20) is effectively non-iterative and applies at most once in well-formed sentences.[9]

(20) $\quad \mathrm{I} \longrightarrow \underset{\uparrow=\downarrow}{\mathrm{I}} \quad \underset{\uparrow=\downarrow}{\mathrm{I}}$

The c-structure for (18) appears in (21).

[8] The experiencer must bear some grammatical relation to the predicate adverb; I have chosen OBL$_{GO}$ since these are also marked with dative case (section 8.1.2). The reason for considering these arguments of the predicate adverb, and not of the XCOMP's predicate, is that even predicate adverbs without infinitival complements can have dative experiencers, as in (18b). Thus, it would be difficult to argue that the dative is an argument of the XCOMP. The infinitive and any of its objects form a VP and correspond to the XCOMP of the predicate adverb.

[9] Consider constructions like that in (i).

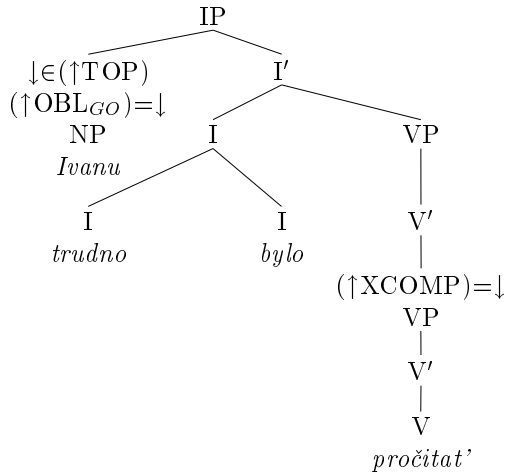
(i)

*Ivanu	trudno	pročitalo	knigu.
Ivan-DAT	hard-PRED.ADV	read-PAST.3.NEUT	book

'It was hard for Ivan to read the book.'

Descriptively, the problem with (i) is that predicate adverbs like *trudno* occur only with infinitival complements, not with finite ones ((i) is ungrammatical regardless of what number and gender are chosen for the finite verb). Unlike with imperfective futures (section 10.1.2), the explanation cannot be that there is a conflict with PRED or TNS values, because the infinitive complement forms its own XCOMP, e.g., see the f-structure in (21). Instead, this is part of a larger problem of subcategorization. For example, the verb *xotet'* 'want' takes infinitival and sentential complements, but cannot occur with finite complements unless they are contained within a CP. Tentatively, I assume that these predicate adverbs simply subcategorize for VPs, which under the approach here can only be infinitives. However, see Pesetsky 1982a on reducing subcategorization restrictions by reference to semantics.

(21)



The predicate adverb *trudno* and the auxiliary *bylo* are in I^0. They can appear in the opposite order, with the auxiliary preceding the predicate adverb.[10] The XCOMP function is associated with a VP argument of the matrix VP; the XPs produced by the V′ expansion rules are usually NPs and PPs, but in certain cases, such as with infinitival and sentential complements, these XPs are of other categories. The experiencer is in topic position under IP and is assigned the grammatical function OBL_{GO} by functional uncertainty (section 9.2). (22) is the f-structure for the sentence in (18a).

(22)
$$\begin{bmatrix} \text{PRED} & \text{'difficult<OBL}_{GO}\text{,XCOMP>'} \\ \text{TNS} & \text{PAST} \\ \text{TOP} & \left\{ \begin{bmatrix} \text{PRED} & \text{'Ivan'} \end{bmatrix} \right\} \\ \text{OBL}_{GO} & [\ ] \\ \text{XCOMP} & \begin{bmatrix} \text{PRED} & \text{'read<SUBJ>'} \\ \text{SUBJ} & [\ ] \end{bmatrix} \end{bmatrix}$$

The TNS is provided by the auxiliary *bylo*, while the PRED value is provided by the predicate adverb. Since the auxiliary has no PRED value and the predicate adverb no TNS value, there is no conflict in the f-structure. The experiencer is licensed by the lexical entry of the predicate adverb. The subject of the XCOMP is identical to the

[10] Some predicate adverbs prefer one order, some the other. I know of no syntactic reason for this preference.

OBL_{GO} of the main f-structure. This identity is captured by a control equation, as in (19).

To conclude this section, tensed and untensed verbs are distributed between I^0 and V^0 in such a way that only tensed forms appear under I^0 in the c-structure, while untensed infinitives appear under V^0. Imperfective futures contain both tensed and untensed forms. However, this poses no problem because the tensed forms in these constructions have no PRED value to conflict with that of the infinitive, and the infinitive has no TNS value. Predicate adverbs can contain two category I^0 forms in a single clause, but one of these, the predicate adverb, does not contain a TNS feature, and as a result there is no conflict in the f-structure.

10.2 *Li* Yes-No Questions

Another interesting distribution of finite elements appears in *li* yes-no questions (section 6). The clitic appears in second position and forms yes-no questions, primarily in embedded contexts.

(23) Pročitala li ona knigu?
read Q she book
'Did she read the book?'

In (23) the verb *pročitala* appears in initial position followed by the clitic *li*. When the verb is in initial position followed by *li*, it has undergone head-movement to C^0. This section discusses how these sentences are derived without such movement.

The basic idea is that *li* requires its mother's f-structure to have a Q-FOC attribute which must then be assigned a value (see the lexical entry in (25)). The clitic never assigns a value to Q-FOC, but assures the attribute's existence. There are two ways this attribute can be assigned a value. One is for a maximal projection to occur in SpecCP. The other is for a finite verb or other I^0 element to appear in C^0 via the rule in (24). The appearance of a finite verb in C^0 requires an addition to the c-structure rules, which previously only allowed complementizers in C^0 (section 10.3).

(24) $\text{C} \longrightarrow$ I C
$\uparrow=\downarrow$ $\uparrow=\downarrow$
$(\downarrow\text{PRED})\in(\uparrow\text{Q-FOC})$

The rule in (24) states that C^0 can be expanded to contain a CAT:I^0 element that is the Q-FOC of its s-structure.[11] In this structure, nei-

[11]The C^0 expansion rule can apply more than once so that there will be two I^0 heads and a C^0 head under the C^0 node; this occurs in *li* yes-no questions when

ther constituent is optional, e.g., I^0 cannot appear in C^0 without C^0 being filled. In general, rules involving the expansion of lexical heads, like (24) and the I^0 rule in (20), behave differently from other phrase structure rules; one of these differences is that they do not contain optional expansions.

(24) states that both C^0 and I^0 are labeled $\uparrow=\downarrow$, in addition to I^0's PRED being marked as the Q-FOC. There are several occasions in which a node is labeled both for grammatical function and for discourse function (section 9.2). The only difference in (24) is that the grammatical function marking is $\uparrow=\downarrow$, instead of $(\uparrow GF)=\downarrow$ or a specific grammatical function (see section 9.1 on functional uncertainty). Thus, I^0 is the head of the f-structure and as a result the material under I^0 in C^0 provides the PRED value.

Lexical entries for the relevant lexical items in (23) are provided in (25).

(25)

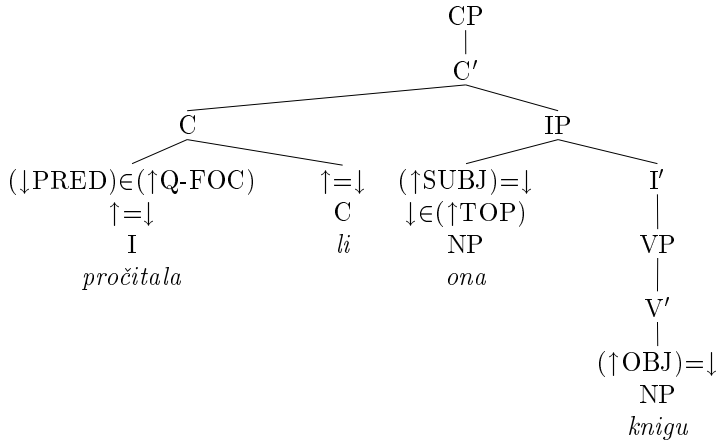
pročitala	*li*
CAT:I	CAT:C
(↑PRED)='read<SUBJ OBJ>'	(↑Q-FOC)
(↑TNS)=PAST	
ona	*knigu*
CAT:NP	CAT:N
(↑PRED)='PRO'	(↑PRED)='book'
(↑PRS)=3rd	
(↑NUM)=SNG	
(↑GND)=FEM	

The c-structure for (23) is shown in (26) and the f-structure in (27).

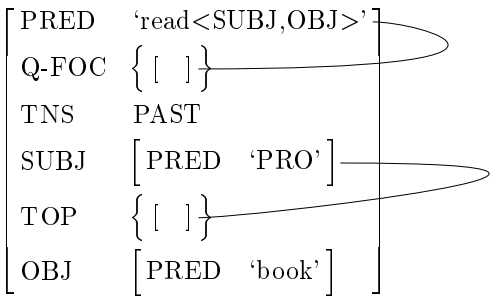
---

both the predicate adverb and the auxiliary appear in C^0 (section 3.5.2). In this construction, IP does not contain an I^0 node since the I^0 elements are in C^0.

(26)



(27)



The f-structure has a Q-FOC as designated by the lexical entry of the clitic *li*. Although the verb appears in C^0, no other tensed verb appears in I^0 or V^0. If one did, the c-structure would be well-formed but there would be a clash of PRED and TNS values in the f-structure. In (26) the PRED of the material in C^0, namely the verb, is designated as the f-structure's Q-FOC. However, due to the fact it is the head, the entire sentence is interpreted as the focus of the question. This results in the desired reading of having the entire clause questioned (section 6.3).

There is one construction in which both I^0 in C^0 and I^0 in I' are filled simultaneously. This occurs when predicate adverbs appear in the *li* yes-no question construction (see section 10.1.4 on predicate adverbs). The lexical entries for a question like (28) are shown in (29).

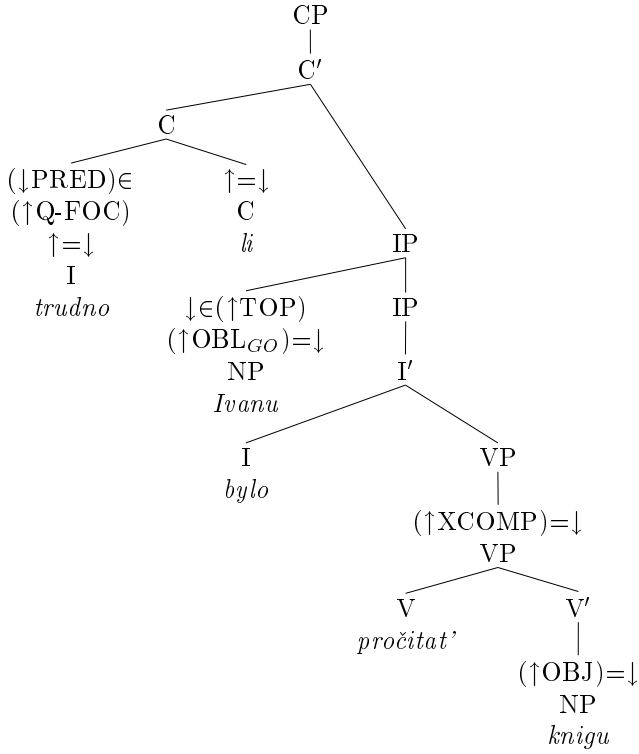
(28) Trudno li Ivanu bylo pročitat' knigu?
difficult-PRED.ADV Q Ivan-DAT was read-INF book
'Was it hard for Ivan to read the book?'

(29)

trudno	*li*
CAT:I	CAT:COMP
(↑PRED)='difficult<OBL$_{GO}$,XCOMP>'	(↑Q-FOC)
bylo	*knigu*
CAT:I	CAT:NP
(↑TNS) = PAST	(↑PRED)='book'
pročitat'	*Ivanu*
CAT:V	CAT:NP
(↑PRED)='read<SUBJ OBJ>'	(↑PRED)='Ivan'

Both the auxiliary and the predicate adverb are CAT:I^0. This allows them to appear in either I^0 within C^0 or I^0 within I′. However, only the auxiliary *bylo* is specified for TNS; so, there is no conflict in the f-structure. The c-structure for (27) appears in (30).

(30)



The predicate adverb *trudno* is in C^0, while the auxiliary *bylo* is in I^0. Both C^0 and I^0 allow CAT:I^0 elements and both the predicate adverb and the auxiliary are specified CAT:I^0. If neither the auxiliary nor the main verb were in the I^0 position in C^0, the resulting f-structure would be ill-formed because the Q-FOC attribute required by *li* would not have a value assigned to it.[12] (31) is the f-structure for the question in (27). This f-structure is essentially identical to the one provided for the declarative predicate adverb construction in section 10.1.4. The primary difference is the presence of the Q-FOC attribute whose value is identical to that of the PRED attribute.

(31)
$$\begin{bmatrix} \text{PRED} & \text{`difficult<OBL}_{GO}\text{,XCOMP>'} \\ \text{Q-FOC} & \{ [\ ] \} \\ \text{TNS} & \text{PAST} \\ \text{TOP} & \{ [\ \text{PRED}\quad \text{`Ivan'}\ ] \} \\ \text{OBL}_{GO} & [\ ] \\ \text{XCOMP} & \begin{bmatrix} \text{PRED} & \text{`read<SUBJ,OBJ>'} \\ \text{SUBJ} & [\ ] \\ \text{OBJ} & [\ \text{PRED}\quad \text{`book'}\ ] \end{bmatrix} \end{bmatrix}$$

10.2.1 'Fronted' XPs

The yes-no questions just discussed had the verb in initial position. However, *li* questions can also be formed with an initial XP. This XP is interpreted as the focus of the question. Section 6.2.1 argued that this focused XP is in SpecCP.

(32) Knigu li ona pročitala?
book Q she read
'Was it a book that she read?'

The lexical entries for the items in (32) are identical to those for (25). The c-structure rule for the expansion of CP is repeated below. This rule was proposed for wh-questions, allowing for multiple elements to appear before C′ as long as they are all elements of the question's Q-FOC set (section 9.3).

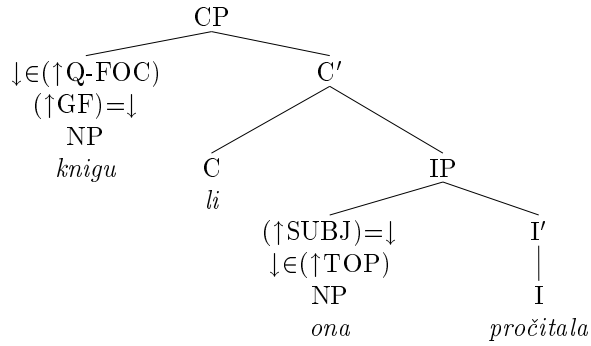
(33) $\text{CP} \longrightarrow \quad \underset{\substack{\downarrow\in(\uparrow\text{Q-FOC}) \\ (\uparrow\text{XCOMP* GF})=\downarrow}}{\text{XP}^*}, \quad \text{C}'$

[12] The I^0 position in C^0 must be empty if some XP is in SpecCP and thus becomes the f-structure's Q-FOC value (see below).

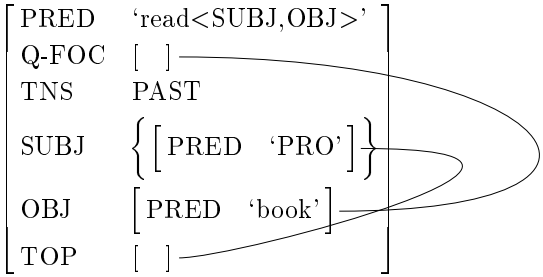
(33) states that any XP in SpecCP must be interpreted as the focus of the question in f-structure. This applies to any material in this position, be it a wh-word or not. In addition, it is assigned a grammatical function. The rule allows for multiple XPs to appear in this position, as with wh-questions. However, this does not happen in *li* questions. In section 6 the restriction on having multiple foci in yes-no questions was discussed: yes-no questions, especially those formed with *li*, rarely contain multiple contrastive foci. This restriction appears to be semantic in nature. As such, I assume that although c-structures in which multiple XPs appear before C^0 in *li* questions are possible, the semantic structure corresponding to such a c-structure will be ill-formed and thus the construction will be ungrammatical.

The c- and f-structures for (32) are shown in (34) and (35) respectively.

(34)



(35)



10.2.2 Q-TYPE and Co-occurrence Restrictions

As discussed in the previous section, the yes-no clitic *li* allows a maximal projection to co-occur with it in SpecCP, and this maximal projection is a member of the sentence's Q-FOC. The relevant annotated c-structure rule is (33). This is the same rule proposed for wh-questions in Russian. Thus it would seem possible to have a wh-phrase co-occur with the clitic *li*; *li* simply states that its f-structure must contain a

Q-FOC, and the wh-phrase in SpecCP provides the value for Q-FOC. However, in Russian wh-phrases never co-occur with *li*, as in (36).

(36) *Kto li priexal včera?
who Q arrived yesterday
'Who came yesterday?'

The clitic *li* is used to mark yes-no questions, while wh-phrases mark wh-questions, and a question cannot simultaneously be both a yes-no and a wh-question. Technically this can be accomplished by using the semantic attribute Q-TYPE (Huang 1992b). The lexical entry of the clitic *li* specifies that its Q-TYPE is YES-NO, as in (37a), while the entries of the wh-phrases specify that they are of Q-TYPE WH, as in (37b).

(37) a. *li*
CAT:C
(↑Q-FOC)
(↑Q-TYPE)=YES-NO

b. *kto* 'who'
CAT:NP
(↑PRED)='who'
(↑Q-TYPE)=WH

So, if *li* and a wh-phrase like *kto* were to co-occur in a sentence, the Q-TYPE values could not unify and the structure is predicted, correctly, to be ill-formed. This method of dealing with the problem is rather stipulative. Ideally, the co-occurrence restriction should result from a more basic semantic difference between the meanings of wh- and yes-no questions. However, for now this difference is represented overtly by the use of the attribute Q-TYPE.[13]

10.3 Why These Heads?

One issue that the previous discussion failed to address is why these particular lexical heads can appear in the specified positions. Of particular concern is the rule in (24) that expands the complementizer position into C^0 and I^0. Why couldn't the there be a rule that expanded C^0 into C^0 and N^0 or ADV^0?

Grimshaw 1991 provides an analysis for head-movement that addresses precisely these issues. Her suggestion is that C^0, I^0, and V^0 are all the same category and only differ in a feature F that designates the functional specification of the category. V^0 is of functional type F0, I^0 of type F1, and C^0 of type F2. However, they are all of category

[13]The fact that Russian *li* only appears in yes-no questions must be stipulated since Bulgarian *li* can appear in wh-questions (section 6.5). According to Rudin 1985, the addition of the particle *li* emphasizes the question, essentially placing contrastive focus on the wh-word (this can be detected in a change in the presupposition of the question: when *li* appears, it is presupposed that the action occurred (Hajičová 1983)).

[+V,–N]. Different bar levels are designated by a third variable, L; L0, L1, and L2 represent X, X′, and XP respectively. So, for example, CP represents a set of values, namely {[+V,–N],L2,F2}. Together, these projections form an extended projection, and lexical heads can move along this extended projection into positions of higher F type. The definition of extended projection is in (38) (Grimshaw 1991:4).

(38) x is the *extended head* of y,
and y is an *extended projection* of x iff:
(a) y dominates x,
(b) y and x share all categorial features,
(c) all nodes intervening between x and y share all categorial features.
(d) If x and y are not in the same perfect projection, the F value of y is higher than the F value of x
where n *intervenes* between x and y
if y dominates x and n;
n dominates x, and n does not dominate y.

The idea behind extended projections can be adapted to a non-movement analysis as a condition on head expansion rules. Such rules can only expand to include heads of the same category and of certain F values, e.g., the F value can never be more than one less than that of the expanded position.[14] Assuming that V^0, I^0, and C^0 are all of category type [+V,–N], the new phrase structure rules are as follows.

(39) {[+V,–N],L2,F2} ⟶ XP*, (↑GF)=↓ ↓∈(↑Q-FOC) {[+V,–N],L1,F2}

(40) {[+V,–N],L1,F2} ⟶ {[+V,–N],L0,F2}, {[+V,–N],L2,F1}

(41) {[+V,–N],L2,F1} ⟶ XP*, (↑GF)=↓ ↓∈(↑DF) {[+V,–N],L1,F1}

(42) {[+V,–N],L1,F1} ⟶ {[+V,–N],L0,F1}, {[+V,–N],L2,F0}

(43) {[+V,–N],L2,F0} ⟶ XP, (↑SUBJ)=↓ {[+V,–N],L1,F0}

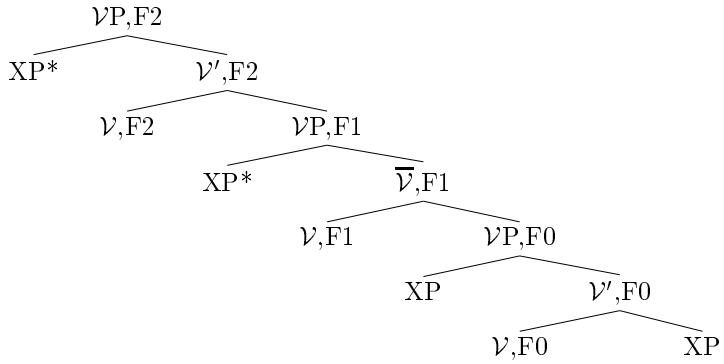
(44) {[+V,–N],L1,F0} ⟶ {[+V,–N],L0,F0}, XP* (↑GF)=↓

(45) a. L0 > L2
b. L2 > L1

[14] This is the equivalent of saying that heads must move from head to head and cannot skip projections, e.g., the Head Movement Constraint (Baker 1988; Travis 1984). Whether this is the desired generalization must be determined empirically.

These produce the same tree structure as before, only the similarity between the categories is now apparent. For ease of exposition, $\mathcal{V}$ is used to represent [+V,–N] and X′ notation is retained. What is of primary importance is the identity of the category information in contrast to the varying functional type. The structure produced is shown in (46).

(46)



Under this schema, finite verbs differ from infinitives not in category type, but in functional type. Finite verbs are of category $\mathcal{V}$ and functional type F1, while infinitives are of category $\mathcal{V}$ and functional type F0. Lexical insertion matches both category type and functional type, resulting in the desired distribution of finite and infinitival verbal elements.

This leaves the appearance of finite verbs in C^0 to be accounted for. The new C^0 expansion rule is as in (47).

(47) $\mathcal{V}$,F2 $\longrightarrow$ $\mathcal{V}$,F1, $\uparrow=\downarrow$ $(\downarrow\text{PRED})\in(\uparrow\text{Q-FOC})$ $\mathcal{V}$,F2 $\uparrow=\downarrow$

The rule expands into the same category; the only difference is that the second element is of a different functional type, namely one lower than that of the expanded node. The lower functional type is a reflection of what in GB would be movement from a lower position in the extended projection. Exactly what types of expansions of L0 categories (heads) are permissible and desirable remains to be investigated, but restricting them to within a single category type whose functional type differs maximally by one is an initial step in constraining the otherwise arbitrary expansion of heads.

To conclude, the phenomena which are often attributed to head-movement can be captured by the interaction of the morphology with

the phrase structure rules. When these phrase structure rules follow a basic X′ syntax, as the ones proposed in chapter 9 for Russian do, this interaction results in a correlate of head-movement constructions. Since head-movement is often proposed to account for related lexical items having different distributions depending on their morphology, e.g., inflection, it is not surprising that in a theory without movement the same facts can be accounted for by the morphology. So, while in GB a verb moves to I^0 to receive inflectional features, in LFG the finite verb is of type I^0 due to its inflectional features, i.e., the suffixation of inflectional features changes the category of the verb, and as a result it appears only in phrase structure positions which make reference to I^0. This could be viewed either as a direct application of morphology in the syntax in which the inflectional suffix is added after the movement to I^0 or as a means of licensing or checking the inflectional features added to the verb prior to lexical insertion. See Janda 1993 on the difficulty of performing the morphology by moving the verb from one head to another where it acquires the relevant affixes. A similar approach to that proposed here could be taken in GB, as in the 'minimalist' theory (Chomsky 1992).

In this last section, I argued that the distribution of heads, in particular the occurrence of more than one head in a single position (head expansion rules), is constrained in a fashion similar to that proposed elsewhere for head-movement. For example, in head expansion rules, the heads must be related to one another in type, differing minimally. Also, unlike the proposal made for the major phrase structure rules, head-expansion rules do not contain optional elements; if the rule is employed, all of its parts are instantiated.

This exploration of 'head-movement' in LFG completes the discussion of how to capture the correlations between word order, phrase structure, and discourse functions in Russian. The phrase structure rules proposed here follow a basic X′ syntax, although their instantiation sometimes appears to violate the fundamentals of such a schema, e.g., the majority of VPs in finite clauses contain no V^0. However, these apparent aberrations result from performing the morphology in the lexicon and inserting lexical items into the syntax according to their post-inflectional-morphology type. I suggested that for each phrase structure rule, there is at least one construction in the language which results in the complete instantiation of the rule, i.e., there will be constructions in which the head is present in the phrase structure, and the same for complements. However, there is no formal principle in LFG which requires phrase structure rules to conform to X′ syntax, although there may be theoretical reasons to adhere to this. If there is

a language or language particular construction which shows evidence against the hierarchical structure imposed by the X′ syntax, then other structures are allowed. For example, there would be no formal difficulty in having a flat VP for Hungarian, the configuration argued for by Kiss *to appear*.

11

Conclusion

This work has investigated word order in Contemporary Standard Russian. It is concerned with what the different word orders mean and how they should be represented in the syntax. These two concerns are related in that one of the goals has been to show how syntax interacts with topic–focus interpretation. The claim here is that Russian is essentially a configurational language and that the varied surface word orders are a result of specific phrase structure positions being associated with specific discourse functions. The appearance of constituents in these positions marks their displacement from a canonical or underlying position.

I first present a brief summary of the conclusions reached. This is followed by a discussion of several areas for further research based on issues raised here.

11.1 Summary

As has long been observed, the different word orders of a Russian clause represent the discourse function organization of the clause. In general, given information precedes new information. I propose that a three-way division into topic, focus, and discourse-neutral material is necessary to account for the possible word orders (chapter 4). In addition, different types of topics and foci must be recognized since they have different syntactic and phonological realizations, as well as different semantic and pragmatic uses.

The association of particular phrase structure positions with specific discourse functions captures the intuition that word order reflects the discourse functions of constituents (chapter 5). These syntactic positions act as licensers: in order for a constituent to be interpreted as having a particular discourse function, it must appear in the appropriate position. In some cases, only a single constituent can be licensed

in a given position (sections 5.1.1.1, 5.1.2). In others, these positions allow multiple constituents to bear the licensed discourse function (section 5.1.1.2). These multiple constituents form a topic or focus field, and the internal ordering of the field is largely determined by pragmatic factors. Since the word order of a clause is derivative from the phrase structure, the motivated movement of constituents into these positions results in the desired orderings, without resorting to extensive 'stylistic' PF reorderings.

Intonation also plays an important role in the interpretation of discourse functions, particularly that of focus. As a result, the permissible intonation patterns are considered concurrently with the study of word order. In Russian, marked intonations correspond to different possible word orders than clauses with unmarked, neutral intonation (section 4.3.4). A constituent receiving the marked sentence stress is always the focus of the clause and can appear in non-clause final position. The underlying phrase structure of the sentence, which was determined by examining neutral sentences, is identical regardless of the type of intonation. Thus, the 'free' word order of Russian, even when non-neutral intonation is considered, is not free, but is instead determined by the discourse functions of the constituents and their corresponding position in the phrase structure.

The basic phrase structure proposed, including the syntactic encoding of discourse functions, conforms to a basic X′ syntax. I argue that Russian is configurational in nature (chapter 3). The multiplicity of word orders is not indicative of a flat structure. The arguments of the verb remain within the VP unless they move to receive certain discourse function interpretations. Since finite verbs do not remain in the VP, unlike infinitives, the default word order is VSO. However, most clauses have preverbal topics, disguising this basic order and frequently resulting in SVO order. Impersonal constructions suggest that the thematically highest argument tends to be topicalized and that there is no privileged preverbal subject position (section 5.4).

Finally, the interaction between phrase structure, discourse functions, and grammatical relations was investigated in LFG. Certain phrase structure positions are associated with particular discourse functions via functional uncertainty (chapter 9). In general, the phrase structure rules follow an X′ schema. Although all nodes are optional, the rules are constrained so that each node must be overtly realized in at least one construction. The distribution of verbal heads is governed by the interaction of morphology and lexical insertion, often resulting in headless projections in the phrase structure (chapter 10). In addition, heads can be expanded to contain other heads, as long as they differ

minimally from the expanded head. In a few cases, a flat structure occurs where adjunction was proposed in GB. This structure avoids the projection of undermotivated null-headed projections. I suggest that flat structures be restricted in distribution: one possibility is to restrict them to certain nodes, e.g., S; the other is that flat structures correspond to certain adjunction structures.

11.2 Areas for Further Research

Three possible extensions of this research are discussed here. The first is that other 'free' word order languages should be analyzed similarly to Russian in that the word orders reflect the discourse functions of the constituents. The second is that topic and focus need to be further defined, both within Russian and cross-linguistically, as does their role in the grammar as a whole. Finally, the analysis proposed here for Russian should be applied to naturalistic corpora to see how it accounts for sustained discourse.

11.2.1 'Free' Word Order

One of the more prominent claims made here is that free word order is not in fact free. This is a return to the idea that if there are different forms of a clause, they correspond to different meanings. In particular, different word orders correspond to different topic and focus interpretations. These interpretations are further constrained by the intonation of the sentence. For example, any preverbal constituent in Russian is either a topic or a focus; preverbal material cannot be discourse-neutral. I argue that the different orderings are reflected in the syntax of the phrase structure of the language. As a result, topic and focus can interact with other phenomena which are sensitive to the phrase structure. In the literature, it appears that most 'free' word order languages are similar to Russian in that the different orders reflect different discourse functions. The syntactic structure of these languages varies substantially, as does the syntactic realization of topic and focus. However, the regularities which can be found suggest that there is an interaction of word order, and, I would hypothesize, phrase structure, with discourse functions. Although languages unrelated to Russian may have only minimal syntactic similarities, it is hoped that the basic ideas behind this analysis of Russian can be adapted to explain these free word order phenomena. Work done on languages such as Hungarian (Kiss *to appear*, Horvath 1986), Korean (Jo 1986), Mayan (Aissen 1992), and Modern Greek (Tsimpli 1992) suggests that there is such a commonality.

Although not explored in depth here, I distinguished between

scrambling and movement into topic and focus positions. Scrambling can be thought of as movement affecting the ordering of constituents within a particular domain. For example, constituents can right-adjoin to VP. Although these right-adjoined constituents tend to be interpreted as focused because they usually appear on the right-edge of the clause, this adjunction is not movement to a focus position *per se*. The permuting of constituents within the VP, especially if the entire VP is within the domain of focus, is scrambling. Another example of what might be thought of as scrambling is the ordering of multiple topics. All topics are adjoined to the same position in Russian, i.e., they undergo movement to a discourse function position. However, when there are multiple topics, an ordering must be established between them. Although this ordering can be determined by successive adjunction of the topicalized elements, conceptually it can be thought of as scrambling within the topic field. Given that movement out of the VP results in different discourse function interpretations, a natural question is what motivates scrambling within these discourse function fields. I suggest that this ordering or scrambling within a given field also reflects the discourse function organization of the sentence. So, within a set of multiple topics, the first one tends to be the one most recently added to the set of topics, while the last one is the oldest of the topics; this is consistent with Yokoyama's 1986 generalizations about the organization of preverbal constituents. Unlike the differences between the topic and focus positions, these reorderings do not reflect movement into different syntactic positions, each associated with a different semantics or pragmatics.

Colloquial Russian Here I focused on Contemporary Standard Russian. However, the relation between Contemporary Standard Russian (CSR) and Colloquial Russian remains to be investigated. In fact, the structure of Colloquial Russian in general deserves further investigation (Zemskaja 1973). The differences between Colloquial and CSR can be divided into differences in lexical items and differences in word order and related syntactic processes. One question of interest to syntacticians is how the differences in word order are to be accounted for. One hypothesis is that Colloquial Russian is basically CSR with certain constraints relaxed. Alternatively, the basic structure of the two language could be fundamentally different. If we explore the first idea, it could be posited that the phrase structure of the two languages, including the interpretation of discourse function positions, is identical. This seems to be the case. Consider one of the most striking differences between Colloquial and CSR: the ability to move arguments out of their mini-

mal finite clause in wh-questions and in topic and focus constructions. Yadroff 1992a, 1992b notes that these moved constituents usually appear immediately before the verb of the higher clause. If this preverbal position is a focus position, as it is in CSR, this fact is explained since the moved constituents are often contrastively focused. So, here the difference between Colloquial and CSR is whether movement is possible out of a finite clause; the motivation for such movement and the discourse function interpretation of the moved constituent are identical. Ideally, other differences between Colloquial and CSR, such as the ability to move adjective phrases out of NPs stranding the noun, will have similar explanations and motivations. With only minor modification to independent syntactic parameters, the analysis of CSR could then be extended to Colloquial Russian.

Slavic Rudin's 1985, 1993 analyses of Bulgarian raise the issue of cross-Slavic realizations of topic, focus, and subject position. Since the Slavic languages are closely related to one another and their syntax has been relatively well studied, the ways in which they encode topic and focus should have interesting consequences. For example, it could be taken as a basic hypothesis that all the Slavic languages are similar in that all arguments are generated within the VP. SpecVP is subject position and the verb raises out of the VP, giving VSO word order. In addition, wh-words appear in SpecCP, complementizers in C^0. With the assumption that all Slavic languages are right-branching, this provides a common syntactic basis for comparison. Then the question arises as to the syntactic differences among the languages and how these affect word order and the encoding of discourse functions.

One obvious difference is the existence of pronominal and auxiliary clitics in many of the Slavic languages. In some languages, these clitics appear in second position, while in others they cluster around the verb. These interact with word order not only because of their restricted position, but also because preverbal constituents, such as foci in Bulgarian, can host clitics. The syntactic category and placement of clitics remain to be determined (Halpern 1992, Rivero 1991, 1993). Their placement interacts with the syntactic structure of the clause above the VP, e.g., functional projections and projections containing the preverbal focused and topicalized constituents.

In addition to assuming that the Slavic languages have similar underlying structures, one hypothesis is that, as with Russian, all movement out of the VP results in particular discourse function interpretations. The question then arises as to which syntactic positions are associated with which discourse functions and whether the interpreta-

tion of these functions is identical to that of Russian. For example, in matrix clauses, Bulgarian bears a close resemblance to Russian in that topics precede foci which in turn precede the verb and its other arguments. Bulgarian poses the same problem as Russian in that preverbal subjects are less marked as topics and foci than other preverbal arguments. Focus in Bulgarian appears to be similar to Russian in that it appears immediately preverbally. However, there are some striking differences between Bulgarian and Russian topics since in Bulgarian topics can be adjoined to CP, as can be seen in questions and subordinate clauses. The full extent of these similarities and differences remains to be investigated. By observing possible differences among closely related systems, it should be possible to determine a common basis for this analysis and to better understand the details of the individual systems.

11.2.2 Discourse Functions

Different types of topic and focus must be recognized in order to capture the basic system responsible for the different word orders. A more finely grained typology of topic and focus allows more accurate description of the data and hence a better analysis. In addition, the more concrete the descriptions of these discourse functions, the easier it is to make cross-linguistic comparisons.

First, it is important to recognize a distinction between material which is topic or focus and that which is discourse-neutral. This three-way distinction, as opposed to analyses which make two-way distinctions such as topic–comment or given–new, allows an explanation for why the non-focused information does not behave uniformly. That is, within the domain of what is traditionally referred to as given information, some items behave differently than others and occur in different positions. However, not only must a distinction be made between topic, focus, and discourse-neutral material, but between the different types of topics and foci. The extent to which the differences between the types of topic and focus play a role in the semantics is a topic for further research. It is possible that there is one underlying property of focus and of topic and that this property is what is relevant to the semantics. The subdivisions of focus and topic would then be pragmatic in nature, not semantic, but either could potentially have reflexes in the syntax and phonology. Regardless of where their meaning is represented in the grammar, these subdivisions are important for the word order of the clause, and the better they are understood, the more accurate our analysis of phrase structure and word order in Russian and cross-linguistically.

Focus Focus is usually defined as new information. One question to raise is whether all new information is focus. The answer is traditionally taken to be fundamental to the definition of focus (Jackendoff 1972). Defining focus as new information seems to be a correct and useful generalization. However, it opens two areas for exploration.

The first is how one determines what is new information in a given context. A better understanding of this will in turn help to explain how different types of focus are defined. Perhaps the best known example of the difficulty in defining new information is that of contrastive focus. With contrastive focus, the focused item has already been introduced into the discourse or is a member of a well defined set of objects. As such, contrastive foci frequently appear topical in character. However, they are foci since, within the smaller domain of the utterance, they are new information. It is this new-information aspect which results in their patterning like foci, and not like topics.

The second area to be explored is whether there are further subdivisions among the kinds of new information useful either to the interpretation of the sentence or to the syntax. This appears to be the case. In languages like Hungarian, in which only one constituent can move into the structural focus position, other new information, which in the broad sense of the term must be focus, can remain within the VP. The question is how the item in focus position is different from the other new information and then what types of focus reading are possible for constituents in this position, as opposed to constituents in other positions.

Topic The semantics of topic and subject-of-predication have begun to be explored in greater detail. Being able to differentiate among the different types of constituents which appear in the topic field in Russian will lead to more detailed accounts of restrictions on the preverbal word order. For example, it may be that the subject-of-predication always appears in a certain position relative to other topics and preverbal pronouns. If such is the case, then a syntactic position for subject-of-predication would be reasonable, and the other topic and focus positions would be differentiated from this position. Such an analysis might help untangle the positioning of adverbs in preverbal position since there would be several possible adjunction sites in which they can appear.

Not all clauses need to have a topic or subject-of-predication. Such topicless sentences clearly exist, and as such, no account can require that there always be a subject-of-predication for each clause. There is no restriction as to which argument of the verb can be the subject-of-

predication, i.e., it is not restricted to being the thematically highest argument or the grammatical subject, although there may be pragmatic reasons why these arguments are frequently subjects-of-predication. In order to provide such an analysis, reliable means must be found for determining which constituent, if any, is the subject-of-predication of a clause. The current account predicts that this subject-of-predication must appear preverbally in the topic field. However, there should be further semantic restrictions on the subject-of-predication which would be more constricting than those for topics, i.e., subjects-of-predication would have all of the properties of topics plus some additional ones. An interesting possibility of such an account would be that a sentence could have a topic without having a subject-of-predication. Languages could differ as to which types of topic they encode in which ways and whether, in fact, they have topics separate from subjects-of-predication.

The Representation of DF Information The representation of discourse function information as a whole still remains to be investigated. Some aspects of focus, and probably topic, need to be represented in the semantics proper, in addition to the pragmatics. However, discourse function information has syntactic and phonological reflexes, as well as interacting with the representation of the discourse. One issue touched on here is where discourse function information belongs in LFG. (This is also relevant to GB, although it is more easily avoided there since topic and focus information do not have to be passed to a particular module, as in LFG.) The fact that certain phrase structure positions are associated with discourse functions requires phrase structure to make reference to discourse functions. For example, the annotation on a position may state that that constituent is the topic of the clause.

Where is this information realized? Topic and focus information has been placed in the functional-structure from which a semantic-structure is projected. However, it has been suggested that constituent-structure information is directly relevant to semantic-structure, and as such, the topic and focus information can be directly relevant to the semantics and pragmatics, without appearing in the functional-structure. As such, the type of discourse function information relevant to the syntax, such as topic in topic-prominent languages like Malay, will appear in the functional-structure, while discourse function information which does not interact with the functional-structure need not appear there. That is, this information can be shared directly between the constituent-structure and the semantic-structure. Note that in GB this particular issue does not arise since phrase structure and grammat-

ical function structure are not independent. However, like LFG, these structures feed LF, which is a rough equivalent to semantic-structure, which in turn interfaces with the semantics and presumably the pragmatics and discourse structure.

Another, broader, issue concerns how discourse as a whole is to be represented. One possibility would be to have it as a part of the semantics, or perhaps what is usually thought of as semantics and pragmatics is a part of the larger representation of the discourse. A second possibility is that there is a separate discourse structure and that topic and focus information are part of this representation. Since the continuity of the discourse must be represented, this structure should provide information about whether a constituent is a viable topic or focus. These possibilities are not mutually exclusive: there could be separate semantic and discourse modules or representations, which share certain information.

11.2.3 Corpora

The representation of discourse structure relates to the final issue addressed here: the application of the proposed analyses to corpora of sustained discourse. Most of the data discussed here were judged in relative isolation, e.g., in two sentence 'discourses'. However, the analyses must be tested on naturally occurring discourse in order to determine their accuracy and to provide further refinements.

By analyzing larger corpora, it should be possible to determine whether the difficulties the proposals here encountered are the result of flaws in the analysis or of certain motivated exceptions. For example, the fact that Russian pronouns tend to occur adjacent to the verb could be the result of the types of topics they usually represent or of their relatively weak prosodic status. In addition, contrastive foci, which usually appear immediately preverbally, sometimes occur in initial position. Yokoyama 1986 notes that these foci cannot appear initially in discourse initial contexts. By observing their naturally occurring distribution, it may be possible to determine what contexts permit initial contrastive foci, e.g., they might only occur as direct rebuttals or as answers to questions.

By analyzing more extensive, connected data, it should also be possible to further refine the proposed analyses. For example, different phrase structure positions were not proposed within the topic and focus fields; that is, all topics adjoin to the same position and all new-information foci are within the VP. However, the ordering of the constituents within these fields is not random, i.e., it is not the result of unmotivated scrambling. Instead, the ordering is also driven, at least

in part, by pragmatic discourse factors, which could be untangled given sufficient data. Another construction which would benefit from such an endeavor is external topics. Although the syntax of external topics is relatively straightforward, their semantics and pragmatics is less understood. Since this construction occurs primarily in spoken Russian, it is principally through the examination of large quantities of spoken data that the precise discourse role of external topics will be revealed.

Bibliography

Bibliography

Adamec, P. 1966. Porjadok slov v sovremennom russkom jazyke. *Rozpravy Československé Akademie Věd, Řada společenských věd* 76.

Aissen, J. 1992. Topic and focus in Mayan. *Language* 68:43–80.

Alsagoff, L. 1992. *Topic in Malay: The Other Subject.* PhD thesis, Stanford University.

Alsina, A. 1993. *Predicate Composition: A Theory of Syntactic Function Alternations.* PhD thesis, Stanford University.

Alsina, A., and S. Mchombo. 1994. Object asymmetries and the Chichewa applicative construction. In S. Mchombo (Ed.), *Theoretical Aspects of Bantu Grammar.* Stanford, CA: CSLI Publications, CSLI lecture notes; no. 38.

Andrews, A. 1982. The representation of case in Modern Icelandic. In J. Bresnan (Ed.), *The Mental Representation of Grammatical Relations*, 427–503. Cambridge, MA: MIT Press.

Andrews, A. D. 1990. Unification and morphological blocking. *Natural Language and Linguistic Theory* 8:507–558.

Aoun, J., N. Hornstein, and D. Sportiche. 1980. Some apects of wide scope quantification. *Journal of Linguistic Research* 1:69–96.

Babby, L. 1980. *Existential Sentences and Negation in Russian.* Ann Arbor: MI: Karoma.

Babby, L. 1986. The locus of case assignment and the direction of percolation: Case theory and Russian. In R. Brecht and J. Levine (Eds.), *Case in Slavic*, 170–219. Columbus, OH: Slavica.

Babby, L. 1989. Subjectlessness, external subcategorizaion, and the projection principle. *Zbornik za Filologiju i Lingvistiku* XXXIII:7–40.

Babby, L. 1993. Adversity impersonal sentences in Russian. Presented at the 2nd Workshop on Formal Approaches to Slavic Linguistics, MIT.

Bailyn, J. 1991. The configurationality of case assignment in Russian. In *Cornell Working Papers in Linguistics* 9, 56–98.

Baker, M. 1988. *Incorporation: A Theory of Grammatical Function Changing.* Chicago, IL: University of Chicago Press.

Banfield, A. 1973. Narrative style and the grammar of direct and indirect speech. *Foundations of Language* 10:1–39.

Banfield, A. 1982. *Unspeakable Sentences.* Boston, MA: Routledge and Kegan Paul.

Barnetová, V., et al. 1979. *Russkaja Grammatika.* Vol. 2. Prague, Czechoslovakia: Academia.

Billings, L. 1993. A note on expletives in Ukrainian -NO/-TO passives. *The Slavic Syntax Newsletter* 3:1–8.

Billings, L., and C. Rudin. 1994. Optimality and superiority: A new approach to multiple wh-ordering. Presented at Formal Approaches to Slavic Linguistics 3, Univ. of Maryland; to appear in J. Toman (ed.) Annual Workshop on Formal Approaches to Slavic Linguistics: The Maryland Meeting.

Bolinger, D. 1967. Apparent constituents in surface structure. *Word* 23:47–56.

Bouchard, D. 1982. *On the Content of Empty Categories.* PhD thesis, Massachusetts Institute of Technology.

Brecht, R., and J. Levine (Eds.). 1986. *Case in Slavic.* Columbus, OH: Slavica.

Bresnan, J. (Ed.). 1982. *The Mental Representation of Grammatical Relations.* Cambridge, MA: MIT Press.

Bresnan, J. 1994. Linear order vs. syntactic rank: Evidence from weak crossover. In *Papers from the 30th Regional Meeting of the Chicago Linguistic Society.*

Bresnan, J., and J. Kanerva. 1989. On locative inversion in Chichewa. *Linguistic Inquiry* 20:1–50.

Bresnan, J., and R. Kaplan. 1982. Lexical-Functional Grammar: A formal system for grammatical representation. In J. Bresnan (Ed.), *The Mental Representation of Grammatical Relations*, 173–281. Cambridge, MA: MIT Press.

Bresnan, J., and S. Mchombo. n.d. Topic, pronoun, and agreement in Chichewa. *Language* 63:741–783.

Bresnan, J., and L. Moshi. 1990. Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry* 21:147–186.

Brody, M. 1990. Some remarks on the focus field in Hungarian. UCL Working Papers in Linguistics 2.

Bryzgunova, E. 1980. In Shedova, N. et al. (Ed.), *Russkaja Grammatika*. Moscow, USSR: Akademija Nauk SSSR.

Butt, M. 1995. *The Structure of Complex Predicates in Urdu*. Stanford, CA: CSLI Publications, Dissertations in Linguistics.

Butt, M., and T. King. 1991. Semantic case in Urdu. In *Papers from the 27th Regional Meeting of the Chicago Linguistic Society*, 31–45.

Cho, Y., and P. Sells. 1993. A lexical account of phrasal affixation in Korean. Unpublished manuscript, Stanford University.

Chomsky, N. 1970. Deep structure, surface structure, and semantic interpretation. In R. Jakobson and S. Kawamoto (Eds.), *Studies in General and Oriental Linguistics*, 183–216. Tokyo, Japan: TEC Corporation.

Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht, Holland: Foris.

Chomsky, N. 1986. *Barriers*. Cambridge, MA: MIT Press.

Chomsky, N. 1992. *A Minimalist Program for Linguistic Theory*. Cambridge, MA: MIT Occasional Papers in Linguistics 1.

Chvany, C. 1973. Notes on 'root' and 'structure-preserving' in Russian. In *You Take the High Node and I'll Take the Low Node*, 252–290. Papers from the 9th Regional Meeting of the Chicago Linguistic Society.

Chvany, C. 1975. *On the Syntax of BE-Sentences in Russian*. Columbus, OH: Slavica Publishers.

Chvany, C. 1990. A note on Svetomir Ivančev's major contributions. *Contrastive Linguistics* 15:17–22.

Comrie, B. 1980. Clause structure and movement constraints in Russian. In C. Chvany and R. Brecht (Eds.), *Morphosyntax in Slavic*, 98–113. Columbus, OH: Slavica Publications.

Comrie, B. 1984. Interrogativity in Russian. In W. Chisholm (Ed.), *Interrogativity*, 7–46. Amsterdam, The Netherlands: John Benjamins.

Crockett, D. 1976. *Agreement in Contemporary Standard Russian*. Columbus, OH: Slavica.

Dalrymple, M. 1993. *The Syntax of Anaphoric Binding*. Stanford, CA: CSLI Publications, CSLI lecture notes; no. 36.

Dalrymple, M., J. Lamping, and V. Saraswat. 1993. LFG semantics via constraints. Preliminary abstract, Xerox PARC, Palo Alto, CA.

Daneš, F. 1974. Functional sentence perspective and the organization of the text. In Daneš (Ed.), *Papers on Functional Sentence Perspective*, 106–128. Prague, Czechoslovakia.

Deising, M. 1988. Verb movement and the subject position in Yiddish. *Natural Language and Linguistic Theory* 8:4–79.

Dik, S., et al. 1980. On the typology of focus phenomena. In T. Hoekstra, H. van der Hulst, and M. Moortgat (Eds.), *Perspectives on Functional Grammar*. Dordrecht, Holland: Foris.

Dowty, D. 1991. Thematic proto-roles and argument selection. *Language* 67:547–619.

Emonds, J. 1985. *A Unified Theory of Syntactic Categories*. Dordercht, The Netherlands: Foris.

Ewen, R. 1979. *A Grammar of Bulgarian Clitics*. PhD thesis, University of Washington.

Fenstad, J., P.-K. Halvorsen, T. Langholm, and J. van Benthem. 1985. Equations, schemata and situations: A framework for linguistic semantics. CSLI Report, CSLI Publications, Stanford University.

Firbas, J. 1964a. From comparative word–order studies. *Brno Studies in English* 4:111–128.

Firbas, J. 1964b. On defining the theme in functional sentence analysis. *Travaux Linguistiques de Prague* 1:267–280.

Firbas, J. 1965. A note on transition proper in functional sentence analysis. *Philologica Pragensia* 8:170–176.

Fowler, G. 1987a. Oblique passivization in Russian. Presented at AATSEEL 1987.

Fowler, G. 1987b. *The Syntax of the Genitive Case in Russian*. PhD thesis, University of Chicago.

Fowler, G., and M. Yadroff. 1993. The argument status of measure nominals in Russian. *Journal of Slavic Linguistics* 1:251–279.

Franks, S., and R. House. 1982. Genitive themes in russian. In *Papers from the 18th Regional Meeting of the Chicago Linguistic Society*, 156–168.

Franks, S. 1990. On the status of null expletives. *Lingua* 81:1–24.

Franks, S. 1992. Parametric properties of numeral phrases in Slavic. To appear in *Natural Language and Linguistic Theory*.

Franks, S., and K. Dziwirek. 1993. Negated adjunct phrases are really partitive. *Journal of Slavic Linguistics* 1:280–305.

Fukui, and P. Speas. 1986. Specifiers and projection. *MIT Working Papers in Linguistics* 8:128–172.

Gerdts, D. 1988. Case stacking in universal grammar. Presented at Linguistic Society of America 62, New Orleans.

Gerdts, D. 1991. An outline of a relational theory of case. In P. McFetridge (Ed.), *Working Papers in Linguistics*, 1:25–53. Simon Fraser University.

Greenberg, G., and S. Franks. 1991. A parametric approach to dative subjects and the second dative in Slavic. *Slavic and East European Journal* 35:71–97.

Grimshaw, J. 1990. *Argument Structure.* Cambridge, MA: MIT Press.

Grimshaw, J. 1991. Extended projection. Unpublished manuscript, Brandeis University.

Gundel, J. 1988. *The Role of Topic and Comment in Linguistic Theory.* New York, NY: Garland.

Hajičová, E. 1974. Meaning, presupposition, and allegation. *Philologica Pragensia* 17:18–25.

Hajičová, E. 1983. On some aspects of presuppositions of questions. In F. Kiefer (Ed.), *Questions and Answers*, 85–96. Dordrecht, Holland: Reidel.

Hale, K. 1983. Warlbiri and the grammar of nonconfigurational languages. *Natural Language and Linguistic Theory* 1:5–47.

Halpern, A. 1995. *On the Placement and Morphology of Clitics.* Stanford, CA: CSLI Publications, Dissertations in Linguistics.

Halvorsen, P.-K. 1983. Semantics for lexical-functional grammar. *Linguistic Inquiry* 14:567–615.

Halvorsen, P.-K., and R. Kaplan. 1988. Projections and semantic description in Lexical-Functional Grammar. *Proceedings of the International Conference on Fifth Generation Computer Systems* 1116–1122.

Hauge, K. 1976. The word oder of predicate clitics in Bulgarian. Meddelelser 10, Slavisk-Baltisk Institutt, Universitetet i Oslo.

Horvath, J. 1986. *FOCUS in the Theory of Grammar and the Syntax of Hungarian.* Dordrecht, Holland: Foris.

Horvath, J. *to appear.* Structural focus, structural case, and the theory of feature-assignment. In K. Kiss (Ed.), *Discourse Configurational Languages.* Oxford, England: Oxford University Press.

Huang, C.-R. 1992a. Certainty in functional uncertainty. To appear in *Journal of Chinese Linguistics.*

Huang, C.-R. 1992b. Interpretation of Mandarin Chinese questions and reverse functional uncertainty. Unpublished manuscript, Academia Sinica.

Isačenko, A. 1976a. Frazovoe udarenie i porjadok slov. In *Opera Selecta*, 373–382. Verlag.

Isačenko, A. 1976b. O grammaticheskom porjadke slov. In *Opera Selecta*, 365–372. Verlag.

Izvorski, R. 1993. On wh-movement and focus-movement in Bulgarian. Presented at CONSOLE 2, Tübingen.

Izvorski, R. 1994. Yes-no questions in bulgarian: Implications for phrase structure. Presented at the 9th Conference of Balkan and South Slavic Linguistics, Literature, and Folklore.

Jackendoff, R. 1972. *Semantic Interpretation in Generative Grammar.* Cambridge, MA: MIT Press.

Jackendoff, R. 1990. *Semantic Structures.* Cambridge, MA: MIT Press.

Jakobson, R. 1936/1971. *Beitrag zur allgemeinen Kasuslehre. Gesamtbedeutungen der russischen Kasus.* The Hague: Mouton. 23–71.

Jakobson, R. 1958. Morfologičeskie nabludenija nad slavjanskim sklo-neniem. In *American Contributions to the IVth International Congress of Slavists.* The Hauge, The Netherlands: Mouton.

Janda, R. 1994. Checking-Theory, syntactic feature geometry, and the structure of IP. In Erin Duncan, Donka Farkas, and Philip Spaelti (Eds.), *The Proceedings of the Twelfth West Coast Conference on Formal Linguisitics.* Stanford, CA: CSLI Publications.

Jo, M.-J. 1986. *Fixed Word Order and the Theory of the Pre-Verbal Focus Position in Korean.* PhD thesis, University of Washington.

Kaplan, R. 1987. Three seductions of computational psycholinguistics. In P. Whitelock, M. Wood, H. Somers, R. Johnson, and P. Ben-net (Eds.), *Linguistic Theory and Computer Applications.* London, England: Academic Press.

Kaplan, R., and J. Maxwell. 1988. An algorithm for functional uncer-tainty. In *COLING*, 12:297–302.

Kaplan, R., and A. Zaenen. 1988a. Functional uncertainty and func-tional precedence in Continental West Germanic. Presented at WWWS, Vienna.

Kaplan, R., and A. Zaenen. 1988b. Long-distance dependencies, con-stituent structure, and functional uncertainty. In M. Baltin and A. Kroch (Eds.), *Alternative Conceptions of Phrase Structure.* Chicago, IL: Chicago University Press.

Kayne, R. 1994. The antisymmetry of syntax. Unpublished manuscript, CUNY.

Kiefer, F. 1980. Yes-no questions as wh-questions. In J. Searle, F. Kiefer, and M. Bierwisch (Eds.), *Speech Act Theory and Pragmatics*, 97–120. Dordrecht: Reidel.

King, T. H. 1995a. *Configuring Topic and Focus in Russian.* Stanford, CA: CSLI Publications, Dissertations in Linguistics.

King, T. H. 1993b. The UTAH and causation in Russian psych verbs. In P. Ackema and M. Schoorlemmer (Eds.), *Proceedings of CONSOLE 1*, 115–130. Leiden: SOLE Publications. Utrecht, The Netherlands.

King, T. H. 1994a. Focus in Russian yes-no questions. *Journal of Slavic Linguistics* 2:92–120.

King, T. H. 1994b. Structuring negation in Slavic. In J. Toman (Ed.), *Annual Workshop on Formal Approaches to Slavic Lingusitics: The Maryland Meeting.*

Kiss, K. 1987a. *Configurationality in Hungarian.* Dordrecht, Holland: Reidel.

Kiss, K. 1987b. Is the VP universal? In I. Kenesei (Ed.), *Approaches to Hungarian II.* Szeged, Hungary: JAT.

Kiss, K. 1992. Scrambling as the base-generation of a random complement order. To appear in N. Corver and H. van Riemsdijk (Eds.). *Scrambling.* Dordrecht, Holland: Foris.

Kiss, K. 1993. Genericity, predication, and focus. Unpublished manuscript, Center for Advanced Study in the Behavioral Sciences, Stanford University.

Kiss, K. to appear. NP-movement, operator movement, and scrambling in Hungarian. In K. Kiss (Ed.), *Discourse Configurational Languages.* Oxford, England: Oxford University Press.

Kitagawa, Y. 1986. *Subjects in Japanese and English.* PhD thesis, University of Massachusetts, Amherst.

Klenin, E. 1974. *Russian Reflexive Pronouns and the Semantic Roles of Noun Phrases in Sentences.* PhD thesis, Princeton University.

Klenin, E. 1978. Quantification, partitivity, and the genitive of negation in Russian. *International Review of Slavic Linguistics* 3:163–182.

Kondrashova, N. 1994. The Russian copula: A unified approach. In J. Toman (Ed.), *Annual Workshop on Formal Approaches to Slavic Lingusitics: The Maryland Meeting.*

Koopman, H., and D. Sportiche. 1985. Theta-theory and extraction. *GLOW Newsletter* 57–58.

Koopman, H., and D. Sportiche. 1991. The position of subjects. *Lingua* 85:211–258.

Kornai, A., and G. Pullum. 1990. The X-bar theory of phrase structure. *Language* 66:24–50.

Kraska, I. 1993. Some observations about Polish word order. Unpublished manuscript, University of Illinois, Urbana-Champaign.

Kroeger, P. 1993. *Phrase Structure and Grammatical Relations in Tagalog.* Stanford, CA: CSLI Publications, Dissertations in Linguistics.

Krylova, O., and S. Khavronina. 1988. *Word Order in Russian.* Moscow, USSR: Russky Yazyk Publishers.

Kuno, S. 1972. Three perspectives on the functional approach to syntax. In L. Metjka (Ed.), *Sound, Sign, and Meaning*, 119–190. Ann Arbor, MI: Michigan Slavic Contributions.

Laka, I. 1990. *Negation in Syntax: on the Nature of Functional Categories and Projections.* PhD thesis, Massachusetts Institute of Technology.

Lasnik, H., and M. Saito. 1984. On the nature of proper government. *Linguistic Inquiry.*

Li, C. (Ed.). 1976. *Subject and Topic.* New York, NY: Academic Press.

Marácz, L. 1990. V-movement in Hungarian: A case of minimality. In I. Kenesei (Ed.), *Approached to Hungarian III.* Szeged, Hungary: JAT.

Mathesius, V. 1964. On some problems of the systematic analysis of grammar. In J. Vachek (Ed.), *A Prague School Reader in Lingusitics*, 306–319. Bloomington, IN: Indiana University Press.

Matsumoto, Y. 1992. *On the Wordhood of Complex Predicates in Japanese.* PhD thesis, Stanford University.

McCloskey, J. 1991. Clause structure, ellipsis, and proper government in Irish. *Lingua* 85:259–302.

McCloskey, J. 1992. Adjunction, selection and embedded verb second. Unpublished manuscript, UC Santa Cruz.

McCloskey, J., and K. Hale. 1984. On the syntax of preson-number inflection in Modern Irish. *Natural Language and Linguistic Theory* 1:487–533.

Mohanan, K. P. 1982. Grammatical relations and clause structure in Malayalam. In J. Bresnan (Ed.), *The Mental Representation of Grammatical Relations*, 504–589. Cambridge, MA: MIT Press.

Mohanan, T. 1994. *Argument Structure in Hindi.* Stanford, CA: CSLI Publications, Dissertations in Linguistics.

Muller, G., and W. Sternefeld. 1993. Improper movement and unambiguous binding. *Linguistic Inquiry* 461–507.

Neidle, C. 1982. *The Role of Case in Russian Syntax.* PhD thesis, Massachusetts Institute of Technology.

Neidle, C. 1988. *The Role of Case in Russian Syntax.* Dordrecht, Holland: Kluwer Academic Publishers.

Nespor, M., and I. Vogel. 1986. *Prosodic Phonology.* Dordrecht: Foris.

Niño, M.-E. 1993. VSO order in Spanish declaratives. Presented at the 67th annual meeting of the Linguistic Society of America, Los Angeles.

Nishigauchi, T. 1990. *Quantification in the Theory of Grammar.* Dordrecht: Kluwer.

Ouhala, J. 1992. Focus in Standard Arabic: The identification requirement and the principles of economy. Unpublished manuscript, Queen Mary and Westfield College, London.

Pesetsky, D. 1982a. *Paths and Categories.* PhD thesis, Massachusetts Institute of Technology.

Pesetsky, D. 1982b. Complementizer-trace phenomena and the Nominative Island Condidition. *Linguistic Review* 1:297–343.

Pesetsky, D. 1987. *Wh*-in-situ: Movement and unselective binding. In E. Reuland and A. ter Meulen (Eds.), *The Representation of (In)definiteness*, 98–129. Cambridge, MA: MIT Press.

Pesetsky, D. 1992. Zero syntx I: Experiencers and cascades. Unpublished manuscript, MIT.

Peterson, P. 1981. Problems with constraints on coordination. *Linguistic Analysis* 8:449–460.

Piñón, C. 1992. ΣP and Hungarian. In *Proceedings of the 11th West Coast Conference on Formal Linguistics.*

Prince, E. 1981. Toward a taxonomy of given-new information. In P. Cole (Ed.), *Radical Pragmatics*, 223–256. New York, NY: Academic Press.

Pulkina, I., and E. Zakhava-Nekrasova. 1988. *Russian: A Practical Grammar with Exercises.* Moscow, USSR: Russky Jazyk.

Rappaport, G. 1986. On anaphor binding in Russian. *Natural Language and Linguistic Theory* 4:97–120.

Restan, P. 1972. *Syntax of the Interrogative Sentence.* Oslo, Norway: Universitetsforlaget.

Rivero, M.-L. 1991. Long head movement and negation: Serbo-Croatian vs. Slovak and Czech. *Linguistic Review* 8:319–351.

Rivero, M.-L. 1993. Bulgarian and Serbo-Croatian yes-no questions: V^0-raising to *-li* vs. *li*-hopping. *Linguistic Inquiry* 24:567–575.

Rizzi, L. 1990. *Relativized Minimality.* Cambridge, MA: MIT Press.

Rochemont, M., and P. Culicover. 1990. *English Focus Constructions and the Theory of Grammar.* Cambridge, England: Cambridge University Press.

Rochemont, M. 1986. *Focus in Generative Grammar.* Amsterdam, The Netherlands: John Benjamins.

Rudin, C. 1985. *Aspects of Bulgarian Syntax: Complementizers and WH Constructions.* Columbus, OH: Slavica.

Rudin, C. 1988. On multiple questions and multiple WH fronting. *Natural Language and Linguistic Theory* 6:445–502.

Rudin, C. 1989. Multiple questions south, west and east: A Government–Binding approach to the typology of wh-movement in Slavic languages. *(to appear) International Journal of Slavic Linguistics and Poetics* 39–40.

Rudin, C. 1993. On focus position and focus marking in Bulgarian questions. Presented at Formal Linguistic Society of Mid-America IV, Iowa City, Iowa.

Rudin, C. *to appear.* Topic and focus in Bulgarian. To appear in *Acta Linguistica Academiae Scientarum Hungaricae.*

Sag, I., G. Gazdar, T. Wasow, and S. Weisler. 1985. Coordination and how to distinguish categories. *Natural Language and Linguistic Theory* 3:117–172.

Schoorlemmer, M. 1991. Syntactic modality in Russian. *Proceedings of Going Romance and Beyond* 113–130.

Selkirk, E. 1982. *The Structure of Words.* Cambridge, MA: MIT Press.

Sells, P. 1985. *Lectures on Contemporary Syntactic Theories.* Stanford, CA: CSLI Publications, CSLI lecture notes; no. 3.

Sgall, P., E. Hajičová, and J. Panevová. 1986. *The Meaning of the Sentence in Its Semantic and Pragmatic Aspects.* Dordrecht, Holland: Reidel.

Sgall, P. 1972. Topic, focus, and the ordering of elements of semantic representations. *Philologica Pragensia* 15:1–14.

Simpson, J. 1983. *Aspects of Warlpiri Morhpology and Syntax.* PhD thesis, Massachusetts Institute of Technology.

Speas, M. 1990. *Phrase Structure in Natural Language.* Dordrecht, Holland: Kluwer Academic Publishers.

Steenbergen, M. v. 1989. Finnish: Configurational or not? In L. Marácz and P. Muysken (Eds.), *Configurationality*, 143–158. Dordrecht, Holland: Foris.

Stowell, T. 1983. Subjects across categories. *Linguistic Review* 2:285–312.

Timberlake, A. 1986. Hierarchies in the genitive of negation. In R. Brecht and J. Levine (Eds.), *Case in Slavic*, 338–360. Columbus, OH: Slavica.

Travis, L. 1984. *Parameters and Effects of Word Order Variation.* PhD thesis, Massachusetts Institute of Technology.

Tsimpli, I. 1992. Focussing in Modern Greek. Unpublished manuscript, University College London.

Walinska, H. 1992. Non-major functional categories. Presented at XVIII Incontro di Grammatica Generativa, Ferrara.

Yadroff, M. 1992a. Scrambling and relativized minimality. Unpublished manuscript, Indiana University, Bloomington.

Yadroff, M. 1992b. The syntactic properties of adjunction (scrambling in Russian). Unpublished manuscript, Indiana University, Bloomington.

Yokoyama, O. 1986. *Discourse and Word Order.* Amsterdam, The Netherlands: John Benjamins.

Zaenen, A., and R. Kaplan. n.d. Formal devices for linguistic generalizations: West Germanic word order in LFG. Manuscript, Xerox PARC, Palo Alto, CA.

Zaenen, A., J. Maling, and H. Thráinsson. 1990. Case and grammatical functions: The Icelandic passive. In J. Maling and A. Zaenen (Eds.), *Modern Icelandic Syntax*, 95–136. San Diego, CA: Academic Press.

Zemskaja, E. 1973. *Russkaja Razgovornaja Reč'.* Moscow: Nauka.

Index

This book discusses the syntactic structure of Russian, traditionally thought of as a 'free' word order language in which word order reflects discourse functions. Tracy Holloway King argues that Russian is a configurational language, but that the expected orderings are marked by preverbal topic and focus positions. In other words, she asserts that although surface word order in Russian is quite free, these different orderings reflect different structures. With an in-depth sytactic analysis of a free word order language, King discusses the syntactic representation of discourse functions. The first part of the book presents this topic using Government-Binding Theory, while the second presents the topic employing Lexical-Functional Grammar.

Specifically, King proposes a tripartite division among topical-ized, focused, and discourse-neutral material and defines several types of topic and focus. Her distinctions are motivated by the different syntactic and phonological, as well as semantic, reflexes of the interpretation assigned to a constituent. Specific phrase structure positions act as licensers for discourse functions: in order for a constituent to be interpreted as having a particular discourse function, it must appear in ththe appropriate position. Since the word order of a clause is derivative from the phrase structure, the motivated occurrence of constituents in these positions results in the desired orderings and interpretations, without resorting to scrambling or stylistic PF recordings. Ultimately, King suggests that this type of analysis can be extended to other free word order language.

"King's work brings the study of Russian syntax into the modern age. She argues convincingly for a configurational view of Russian syntax and shows how many discoveries of traditional Russian grammar fit beautifully into that view. The book is full of new discoveries as well as syntheses of previous work. For all these reasons, King's book will surely become a starting point for all future work on Slavic syntax and on the place of Russian in the theory of universal grammar."

—David Presetsky, MIT.

TRACY HOLLOWAY KING is a Visiting Research Associate at Indiana University, Bloomington.

CSLI Publications
Center for the Study of Language and Information
Stanford, California

ISBN-10 1-88152-662-3
9 781881 526629