

**Feature-Based Binding and Phase Theory**

A Dissertation Presented

by

**Andrei Antonenko**

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

**Doctor of Philosophy**

in

**Linguistics**

Stony Brook University

December 2012

**Stony Brook University**

The Graduate School

**Andrei Antonenko**

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

**John F. Bailyn – Dissertation Advisor**  
**Associate Professor, Department of Linguistics**

**Richard K. Larson – Chairperson of Defense**  
**Professor, Department of Linguistics**

**Daniel L. Finer**  
**Professor, Department of Linguistics**

**Marcel den Dikken**  
**Professor, Department of Linguistics**  
**CUNY Graduate Center**

**Kyle Johnson**  
**Professor, Department of Linguistics**  
**University of Massachusetts at Amherst**

This dissertation is accepted by the Graduate School.

Charles Taber  
Interim Dean of the Graduate School

Abstract of the Dissertation

**Feature-Based Binding and Phase Theory**

by

**Andrei Antonenko**

**Doctor of Philosophy**

in

**Linguistics**

Stony Brook University

2012

Current theories of binding cannot provide a uniform account for many facts associated with the distribution of anaphors, such as long-distance binding effects and the subject-orientation of monomorphemic anaphors. Further, traditional binding theory is incompatible with minimalist assumptions. In this dissertation I propose an analysis of anaphoric binding based on a feature-checking mechanism (Pesetsky and Torrego 2007), by introducing the feature  $\langle \rho \rangle$ , a formalization of the reflexivity proposal of Reinhart and Reuland 1993. I argue that the  $\langle \rho \rangle$  feature is responsible for establishing coreference between an anaphor and its antecedent, by being present and valued on reflexives while being unvalued on a higher phrasal head. Valuation of  $\langle \rho \rangle$  under Agree results in the introduction of a  $\lambda$ -operator, which binds the reflexive variable, thereby establishing the coreference between an anaphor and its antecedent. Central to the workings of this theory is a necessary revision of the definition of binding domains.

Previous definitions could not uniformly account for the possibility of long-distance binding and its correlation with subject-orientation. I reduce the notion of binding domain to a phase, a domain independently motivated in recent research. I demonstrate problems with the traditional definition of a phase, and revise this definition so that phasal domains are derivable from independent mechanisms of grammar, in particular by feature-checking under Agree. I argue that a domain becomes phasal as soon as all relevant features within this domain are valued. As a result, domains with defective tense such as infinitives and subjunctives can be closed at a late stage, permitting probing into them without violating the PIC. Having revised the definition of a phase, I show how phases can be implemented as binding domains and how this can account for cross-linguistic differences in long-distance binding as well as correctly predict the typology of subject-orientation, among other empirical advantages. Finally, I consider the interaction of  $A'$ -movement (scrambling and *wh*-movement) and anaphoric binding and show how it affects the status of binding domains. This analysis of binding has wider empirical coverage than existing analyses and makes binding theory consistent with the minimalist view on the architecture of grammar.

# Table of Contents

<b>Acknowledgements</b>	<b>x</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Introducing binding theory . . . . .	1
1.1.1 Classical binding conditions . . . . .	1
1.1.2 c-command and binding . . . . .	3
1.1.3 Binding domains . . . . .	4
1.1.3.1 Local binding domains . . . . .	4
1.1.3.2 Long-distance binding: domains revised . . . . .	9
1.1.4 Monomorphemic anaphors: subject condition and long-distance binding . .	12
1.1.5 Psych-verbs: Belletti and Rizzi 1988 . . . . .	13
1.2 Revising binding theory: Reinhart and Reuland 1993 . . . . .	15
1.2.1 Reasons for abandoning classical binding theory . . . . .	15
1.2.2 Binding theory as condition on reflexivity of predicates . . . . .	15
1.2.3 Logophors . . . . .	21
1.2.3.1 What are logophors? . . . . .	21
1.2.3.2 The treatment of logophors in Reuland 2001 . . . . .	24
1.2.4 Problems . . . . .	25
1.2.4.1 The nature of SE . . . . .	25
1.2.4.2 Possessive anaphors . . . . .	29

1.2.4.3	Subject anaphors . . . . .	30
1.2.4.4	Binding and non-argumental positions . . . . .	30
1.3	Minimalist theories of binding . . . . .	31
1.3.1	Is binding compatible with minimalism? . . . . .	31
1.3.2	Co-constituent theories of binding . . . . .	32
1.3.2.1	Kayne 2002 . . . . .	32
1.3.2.2	Zwart 2002 . . . . .	35
1.3.3	Feature theories of binding . . . . .	36
1.3.3.1	Hasegawa 2005 . . . . .	36
1.3.3.2	Reuland 2005, 2011 . . . . .	39
1.3.3.3	Hicks 2009 . . . . .	42
1.4	Conclusion . . . . .	44
<b>2</b>	<b>Deriving Phases</b>	<b>46</b>
2.1	Overview of phase theory . . . . .	46
2.2	Theoretical gaps . . . . .	49
2.2.1	What exactly is a phase? . . . . .	50
2.2.2	Spell-out gaps . . . . .	53
2.2.3	Strong and weak phases: what is the difference? . . . . .	54
2.2.4	Other syntactic categories: NPs, DPs and PPs . . . . .	55
2.3	Empirical problems . . . . .	55
2.3.1	Long-distance binding . . . . .	56
2.3.2	Negative Polarity Items . . . . .	60
2.3.3	Long-distance agreement . . . . .	64
2.3.4	Sequence of Tense . . . . .	66
2.4	Phenomena to consider . . . . .	71
2.4.1	Long distance extraction I: Wh-movement . . . . .	72
2.4.2	Long distance extraction II: Scrambling . . . . .	73

2.4.3	Long distance extraction III: Clitic-climbing . . . . .	74
2.4.4	Conclusion . . . . .	75
2.5	Deriving phases . . . . .	75
2.5.1	Theoretical assumptions about feature-checking . . . . .	75
2.5.2	Feature content of $v/V/T$ . . . . .	76
2.5.3	What is a phase? . . . . .	84
2.5.4	Deriving phases . . . . .	88
2.5.4.1	Control clauses . . . . .	90
2.5.4.2	ECM Clauses . . . . .	91
2.5.4.3	Object shift . . . . .	93
2.5.4.4	Ditransitive $vPs$ . . . . .	95
2.5.4.5	Conclusion . . . . .	98
<b>3</b>	<b>Binding in a feature-based framework</b>	<b>100</b>
3.1	Introducing relevant features . . . . .	101
3.1.1	The reflexive feature $\rho$ . . . . .	101
3.2	Deriving binding relations: Principle A . . . . .	106
3.2.1	Complex anaphors . . . . .	106
3.2.1.1	Transitive sentences . . . . .	107
3.2.1.2	Ditransitives and Barss-Lasnik effects . . . . .	109
3.2.2	Monomorphemic anaphors . . . . .	119
3.2.2.1	Transitive sentences . . . . .	119
3.2.2.2	Ditransitive constructions and subject orientation . . . . .	124
3.2.3	Locality of binding . . . . .	128
3.2.4	Intermediate summary . . . . .	129
3.2.5	Psych-verbs and binding . . . . .	130
3.2.5.1	Introducing psych-verbs . . . . .	130
3.2.5.2	The causative nature of object experiencer verbs . . . . .	132

3.2.5.3	The special status of accusative case in object experiencer constructions . . . . .	133
3.2.5.4	Deriving object experiencer constructions . . . . .	135
3.2.5.5	Extensions to binding . . . . .	139
3.2.5.6	Psych-verbs: conclusion . . . . .	142
<b>4</b>	<b>Long-distance binding</b>	<b>143</b>
4.1	Overview of long-distance reflexivization . . . . .	144
4.1.1	Feature-based theory of long-distance binding . . . . .	147
4.2	Interaction of binding and tense . . . . .	151
4.2.1	An analysis of the indicative/subjunctive distinction in Russian . . . . .	151
4.2.1.1	Preliminary data on Russian subjunctives . . . . .	151
4.2.1.2	Indicative clauses . . . . .	153
4.2.1.3	Subjunctive clauses . . . . .	154
4.2.2	Deriving infinitival clauses . . . . .	158
4.2.3	Reconciling the analyses . . . . .	165
4.2.4	Tense dependency and long-distance binding . . . . .	170
4.2.4.1	Anaphors within infinitival clauses . . . . .	170
4.2.4.2	Anaphors within subjunctives . . . . .	177
4.3	Anaphors within subjects . . . . .	185
4.3.1	Subject-internal reflexives in English . . . . .	187
4.3.2	Subject-internal reflexives in Russian . . . . .	190
<b>5</b>	<b>Interaction of binding and long-distance movement</b>	<b>194</b>
5.1	Interaction with long-distance <i>wh</i> -movement . . . . .	195
5.1.1	<i>Wh</i> -movement: featural approach . . . . .	197
5.1.2	Phasal status of embedded CP . . . . .	198
5.1.2.1	$\langle iQ - val \rangle$ CPs . . . . .	199



5.1.2.2	$\langle uQ - val \rangle$ CPs . . . . .	201
5.1.3	Ambiguities in binding and long-distance <i>wh</i> -movement . . . . .	206
5.1.4	Monomorphemic reflexives in Slavic: interaction of <i>wh</i> -movement and binding . . . . .	210
5.1.4.1	The case of -Q complements . . . . .	210
5.1.4.2	The case of +Q complements . . . . .	213
5.1.5	On the impossibility of A'-binding . . . . .	214
5.2	Interaction with scrambling: The case of Russian . . . . .	217
5.3	Conclusion . . . . .	223
5.4	Consequences for non-binding phenomena . . . . .	223
5.4.1	Long-distance extraction . . . . .	225
5.4.2	Parasitic gaps . . . . .	228
5.4.3	Scope . . . . .	231
5.4.4	Future research . . . . .	232
	<b>Bibliography</b>	<b>234</b>

# Acknowledgements

My path to the completion of the present dissertation was not always easy but it was nonetheless enjoyable. Many people provided guidance when I needed and helped me get to the end of it.

First and foremost I would like to thank my advisor, John Bailyn. Our numerous meetings lasted for hours and were filled not only with discussions about syntax and linguistics, but also with numerous friendly conversations about life in general. He was a constant source of ideas and supporting (and not so supporting) data for my proposals, and it was sometimes hard to catch up with his train of thought — it was so fast! His numerous comments on my work often led me to new directions I had to explore. A lot of the issues taken on in this dissertation were due to his suggestions. His constant encouragement helped me get this whole thing done.

I always treated Dan Finer as my other advisor. Since the first moment I met him in Stony Brook at his Syntax I class I was impressed by his encyclopedic knowledge of everything that has ever been done in syntax, including facts about obscure languages and often long-forgotten references. He was always open to the most crazy ideas I generated and has never turned me down. I did not understand much in Syntax I class, but as soon as I saw him, I knew I wanted to be able to think like him. Hopefully this dream will eventually come true (will growing a beard help?).

Richard Larson is one of the most brilliant teachers and linguists I have ever met. Without the ideas I collected at his seminars and classes this work would not be possible. His jokes helped me on my path to defending this dissertation.

I would also like to thank the external members of my committee. Marcel den Dikken pro-

vided numerous (and I mean numerous!) comments on various stages of my work, starting with my first conference presentation at SUNY/CUNY/NYU. While a lot of his questions were quite sophisticated, he was often helpful in providing answers to them himself. I first met Kyle Johnson when he came to Stony Brook to give a colloquium talk. After having a meeting with him, I knew I wanted him to be on my dissertation committee (even though at this time this dissertation was not even started). He was an influence for some of the main ideas of this work (even though he might not admit it).

Very special thanks go to Lori Repetti, my “linguistics godmother”. She was the first person to encourage me to take linguistics classes and recommended that I enroll in graduate Syntax and Phonology instead of just a general Intro. She also was the first one to encourage me to attend the Linguistics department parties. Without all of this I would have never become a student in the Linguistics Department. Thanks to her I was able to find the field I love.

Many other faculty members in the Linguistics Department helped me in various ways along my path to completion. Alice Harris was strict with deadlines, and now I can say rightfully so. Without enforcing them I could have been still writing my first qualifying paper. I am also thankful to her for supporting me as her research assistant and for our (unfinished) collaboration. After submitting this dissertation we will finally be able to resume our other projects. Marie Huffman was always supportive and great to talk to and is an excellent mentor. Her advice on how to survive linguistics helped me a great deal along the way. I would also like to thank other faculty members who I got to know: Christina Bethin (thanks to her, I did not fail the deadline for my second QP), Bob Hoberman (for conversations on various topics), Mark Aronoff (for being so much fun and for our regular taco nights at Mesquite), Ellen Broselow (for supporting me financially from her grant when I needed it and inviting me to stay at her and Dan’s house during Hurricane Sandy, among many other things), and John Drury (for friendly support when I was desperate). Also thanks to Paco Ordóñez for his comments on early stages of this work.

I would also like to thank numerous linguists I met at various conferences and events: David Pesetsky, Norbert Hornstein, Akira Watanabe, Hedde Zeijlstra, Željko Bošković, Sergei Tatevosov,

Heejeong Ko, Edith Aldridge, Miloje Despic, Akira Omaki, Gil Rappaport, Christina Tortora, Stephanie Harves, Steven Franks, Zaira Khalilova, Kaori Miura, Grev Corbett, and Marina Chumakina. Discussions with them were very influential for this work.

Sandra Brennan deserves a separate paragraph in these acknowledgements. Administrator, friend, and in general a great person to be around — you affected my life in this department in a very special way. I am very grateful to you for all your support and for everything you have done for me. (Who can beat driving to Niagara Falls or Mt. Washington and Portland, ME in one day?)

I got a lot of help in completing this dissertation from my fellow students in the Linguistics Department. Carlos de Cuba, Lanko Marušič, Jon MacDonald, Susana Huidobro, Tomoko Kawamura, Yukiko Asano, and Marianne Borroff all helped me feel myself at home at this department when I first came here as a student from the Applied Mathematics Department. Thanks to all of you, I was accepted here (informally!). You all helped me make my choice in what I want to do in life. Sharing an office with Mark Lindsay, Ivana Mitrovic, and Poppy Slocum was a great experience — we had the best office (and the largest amount of beer) in this department, even though some might object to the bloody red color of the walls. You all helped me tremendously and provided numerous excuses for not working that I so desperately needed. Innumerable discussions of syntax (from binding to ellipsis to phase theory to other incredibly obscure linguistic phenomena) with Marlin Taylor were a source of inspiration. It was a lot of fun! I also want to thank Chris Parles: for dinners, judgments, corrections, card games, beer tastings, rides, etc. etc. etc. This could be a very long list! Thanks also to Yuan Lu — it was a lot of fun to organize various events with you (we are the best organizers ever!). Thanks to Jisung Sun for late-night conversations and successful collaboration on our Korean-Daghestanian project. Thanks to Tanya Scott — it is always nice to chat with you about all things, linguistic or otherwise. I would also like to thank my other colleagues who I met while at Stony Brook: Chih-Hsiang Shu, Hijo Kang, Jiwon Hwang, Youngran An, Miran Kim, Hisako Takahashi, Hyun-ju Kim, Aydogan Yanilmaz, Shawn Gaffney, Svitlana Antonyuk-Yudina, and Roksolana Mykhaylyk. Thanks to the other graduate and undergraduate students in the Linguistics Department who I had the pleasure to teach some syntax to

and/or share beers and conversations with: Brigi Fodor, Christine Boucher, Megan Steenrod, Rebecca Hurley, Amanda Caffary, Amanda Yazdani, Anya Melnikova, Suzanne Goldenberg, Oxana Yudina, Vahideh Rasekhi, Sophia Kao, Chong Zhang. Thanks to our soccer teams — Loquitos de Cuba and Loquitos de Aldridge. I successfully proved that it is not necessary to score a goal to defend one's dissertation.

Thanks also to my friends from other departments in Stony Brook: Dasha Eremina, Valik Polishchuk, Ari Grossman, and Dan Woulfin.

Very special thanks go to my family — my mother Elena, my father Vladimir, and my sister Natasha — for all your support during all these years. You never stopped believing that this moment would come. Your encouragement at all stages of my graduate career helped me keep moving along. I had no right to disappoint you.

Finally my thanks go to my wife, Yunju. Thank you so much for all your encouragement and for all the help and support you gave me along this path. I am not sure I would have been able to finish this work without you.

# Chapter 1

## Introduction

### 1.1 Introducing binding theory

#### 1.1.1 Classical binding conditions

In human language, referential noun phrases are generally thought of to be of three different types. The first type, *R-expressions*, are expressions which have their reference fixed and this reference is independent on other noun phrases in the sentence. Examples of such noun phrases are *the man*, *John*, *John's sister*, etc. The reference of R-expressions is either inherent as for proper names, or determined by the context, like in the case of *the man*.

The second type of referential noun phrases is *pronominals*. These elements do not have inherent referential specification, and are only specified for their  $\phi$ -features, i.e. number, gender, person, or their subset. In English, such expressions are *I*, *you*, *he*, etc. There is no requirement on pronominals to be dependent on other expressions, as their reference can be also taken from the context or deictically, i.e. from a pointing gesture alone.

The third type of expressions is *anaphors*, such as *himself*, *each other*, etc. The main difference between anaphors and other types of referential expressions is that anaphors *must* depend on other (linguistic) elements for their reference. They cannot be used without a proper linguistic

*antecedent*, the element from which the anaphors get their reference. Anaphors also cannot be accompanied by a pointing gesture.

Classical binding theory is concerned with the rules and restrictions governing how referential properties of anaphors and pronominals depend on the referential properties of other expressions within a sentence, and under which circumstances coreference is possible or impossible.

In the binding literature coreference is usually denoted by indices, such as *i, j, k* (starting from Chomsky 1981/1993). Two expressions are co-indexed, i.e. share the same index, if they are coreferential. For example, in (1) both *John* and *himself* refer to the same individual, and therefore bare the same index, *i*.

(1) John<sub>*i*</sub> loves himself<sub>*i*</sub>.

Notice that this is the only possible indexation in this case: *himself* must be coreferential with *John*. Any other indexation would lead to an ungrammatical sentence, as the sentence (2) cannot mean that John loves anybody but himself:

(2) \*John<sub>*i*</sub> loves himself<sub>*j*</sub>.

Similarly, in case of pronominals, not any coindexation is possible. If one replaces *himself* with *him* in examples above, the reverse situation would be true:

(3) a. \*John<sub>*i*</sub> loves him<sub>*i*</sub>.  
b. John<sub>*i*</sub> loves him<sub>*j*</sub>.

The only possible meaning of the sentence above is that John loves a certain masculine individual other than himself, and cannot be understood in a way that John loves himself, i.e. John.

The structural conditions on possible coreference between two referential expressions are the domain of binding theory.

## 1.1.2 c-command and binding

The first condition in classical binding theory on coreference between an antecedent and an anaphor is a condition on the structural position of the anaphor with respect to its antecedent.

Consider the following two examples.

- (4)
- a. John<sub>i</sub> saw [pictures of himself<sub>i</sub>].
  - b. \*[Friends of himself<sub>i</sub>] saw John<sub>i</sub>.
  - c. \*[John<sub>i</sub>'s friends] saw [pictures of himself<sub>i</sub>].

As the indicated coindexation shows, the first sentence is grammatical, while the second and the third are not. (4-b) shows that an anaphor within the subject position of a sentence cannot be anteceded by a direct object. Similarly, while the subject itself can serve as an antecedent for an anaphor within the object position (4-a), a nominal embedded into the subject cannot, see (4-c). These facts can be captured by postulating that anaphors must be *c-commanded* by their antecedents, where c-command is defined as in (5).

- (5)  $\alpha$  *c-commands*  $\beta$  iff  $\alpha$  does not dominate  $\beta$  and every phrase  $\gamma$  dominating  $\alpha$  also dominates  $\beta$ .

This requirement takes care of the ungrammaticality of both (4-b) and (4-c). Not only must antecedents precede anaphors according to this definition, but they also cannot be embedded into larger phrases, as c-command out of phrases is impossible.

The classical definition of binding refers to the notion of c-command.

- (6)  $\alpha$  *binds*  $\beta$  iff
- a.  $\alpha$  is coindexed with  $\beta$ , and
  - b.  $\alpha$  *c-commands*  $\beta$ .



### 1.1.3 Binding domains

The c-command condition on binding was rarely questioned; the question of locality and how far an antecedent can be away from an anaphor has, on the other hand, been a source of controversy, and was much harder to define generally, even for English alone.

#### 1.1.3.1 Local binding domains

In order to delimit binding domains, Chomsky 1973 proposes the *Tensed-S Condition*:

(7) **Tensed-S Condition:** No rule can involve X, Y in the structure

... X ... [ $\alpha$  ... Y ...] ...

where  $\alpha$  is a tensed sentence.

According to this condition coreference cannot hold between an antecedent and an anaphor if they are separated by a finite clause boundary. The ungrammaticality of the following sentences would therefore be accounted for:

- (8) a. \*John<sub>i</sub> believes that Mary loves himself<sub>i</sub>.  
b. John<sub>i</sub> expects himself<sub>i</sub> to win the lottery.

In the first example, the boundary which separates the anaphor from its antecedent is a finite clause boundary. In the second example, binding is possible since an ECM clause boundary does not prevent coreference between an anaphor and its antecedent according to Tensed-S Condition.

However, the Tensed-S Condition is not enough to account for the full distribution of anaphors. For instance, anaphors in the object positions of untensed clauses still cannot be bound by an antecedent in the matrix clause, as shown in (9):

- (9) a. \*John<sub>i</sub> expects Mary to love himself<sub>i</sub>.  
b. \*John<sub>i</sub> saw Mary's pictures of himself<sub>i</sub>.

- c. \*John<sub>i</sub> convinced Mary<sub>j</sub> PRO<sub>j</sub> to kiss himself<sub>i</sub>.

In the cases above *John* and *himself* are not separated by a tensed clause boundary, however the coreference between them is impossible. In order to explain cases like this, Chomsky 1973 introduces another condition, the Specified Subject Condition, which prevents application of operations over a specified subject:

- (10) **Specified Subject Condition:** No rule can involve X, Y in the structure

...X... [ $\alpha$  ...Z... - WYV ...] ...

where Z is the specified subject of WYV in  $\alpha$ .

The examples in (9) are ungrammatical because of this condition. In (9-a) there is a specified subject *Mary* of the embedded ECM clause, and it prevents the relation between the anaphor and its antecedent. Similarly, in (9-b), *Mary* is considered to be the specified subject of the DP *Mary's pictures of himself*, and therefore prevent the relation between *John* and *himself*.

In control sentences, Chomsky argues that PRO is unspecified only with respect to its controller. Therefore, in (9-c), PRO is controlled by *Mary*, and therefore specified with respect to *John*, resulting in coreference between *John* and *himself* being impossible.

This condition also explains why with subject control verbs similar coreference is possible:

- (11) a. John<sub>i</sub> promised Mary PRO<sub>i</sub> to shave himself<sub>i</sub>.  
 b. \*John<sub>i</sub> promised Mary<sub>j</sub> PRO<sub>j</sub> to kiss herself<sub>j</sub>.

In both examples in (11), PRO is controlled by *John*, and therefore is unspecified with respect to *John* but specified with respect to *Mary*. Therefore the object reflexive can be coreferential with *John*, as shown in (11-a), but cannot be coreferential with *Mary* as in (11-b).

The notion of bound and free anaphors was first introduced in Chomsky 1980.

- (12) An anaphor  $\alpha$  is *bound* in  $\beta$  if there is a category c-commanding it and co-indexed with

it in  $\beta$ ; otherwise,  $\alpha$  is *free* in  $\beta$ .

Further, in order to eliminate the Specified Subject and Tensed-S conditions, Chomsky 1980 proposes that the domain  $\beta$  can be made opaque in the presence of the subject or tense, effectively combining both conditions into one:

- (13) If  $\alpha$  is an anaphor in the domain of the tense or the subject of  $\beta$ ,  $\beta$  is minimal, then  $\alpha$  cannot be free in  $\beta$ ,  $\beta = \text{NP}$  or  $S'$ .

The Tensed-S condition however is still problematic, whether treated separately or as a part of a more general condition. Consider example (14):

- (14) a. John<sub>i</sub> said that [the pictures of himself<sub>i</sub>] are published in the newspaper.  
b. \*John<sub>i</sub> said that [himself<sub>i</sub> likes Mary]

It is clear that according to the Tensed-S Condition the examples in (14) must be ungrammatical, since a tensed clause boundary separates an anaphor from its antecedent. However, it is clear that only anaphors in nominative positions are ungrammatical, as in (14-b), while anaphors can be embedded into nominative subjects without leading to ungrammaticality, (14-a).

As a result, Chomsky completely dispenses with the “tensed” part of the condition on domains, as too restricting. In fact, not only does it predict the ungrammaticality of (14-a), but it also overlaps with the Specified Subject Condition: it is only subject anaphors which are covered by the Tensed-S Condition but are not covered by the Specified Subject Condition.

In order to account for all the problematic examples above, Chomsky 1980 proposes the Nominative Island Condition:

- (15) **Nominative Island Condition:** A nominative anaphor cannot be free in  $S''$

and replaces his condition on domains with the following version:

- (16) If  $\alpha$  is [an anaphor] in the domain of the subject of  $\beta$ ,  $\beta$  is minimal, then  $\alpha$  cannot be free in  $\beta$  [,  $\beta = \text{NP}$  or  $\text{S}'$ ].

Replacement of a model of binding involving TSC and SSC with the model of binding involving TSC and NIC eliminates overlap between the conditions responsible for the behavior of anaphors, and correctly captures the facts for English: instances of long-distance binding are ruled out by the TSC while subject anaphors are ruled out by the NIC.

The classical formulation of binding theory principles is given in Chomsky 1981/1993, where he defines binding domains as *governing categories*:

- (17) a. **Principle A:** An anaphor is bound in its governing category  
b. **Principle B:** A pronominal is free in its governing category  
c. **Principle C:** An R-expression is free

The notion of *governing category* combines two seemingly different constraints on binding, and is introduced in Chomsky 1981/1993 (the following definition is taken from Harbert 1995). In order to combine these constraints, a notion of *accessible SUBJECT* is now necessary and replaces the TSC and NIC.

- (18) a.  $\beta$  is a *Governing Category* for  $\alpha$  iff  $\beta$  is the minimal category containing  $\alpha$ , a governor for  $\alpha$ , and a SUBJECT accessible to  $\alpha$ .  
b. SUBJECT = AGR where present, a subject NP otherwise.  
c.  $\alpha$  is accessible to  $\beta$  iff  $\alpha$  is in the c-command domain of  $\beta$  and the assignment to  $\alpha$  of the index of  $\beta$  would not violate *i*-within-*i* condition, (18-d).  
d. ***i*-within-*i* condition:** \* $[\delta \dots \gamma \dots]$ , where  $\delta$  and  $\gamma$  bear the same index.

While the definition of SUBJECT still has the disjunctive nature of the previous binding model (TSC/SSC or TSC/NIC), AGR and subject NP are not unrelated: AGR stands for subject-verb (aux) agreement  $\phi$ -features, and therefore SUBJECT can be taken as the locus of these  $\phi$ -features.

(19) John<sub>i</sub> said that [the pictures of himself<sub>i</sub>] are published in the newspaper.

Further, notice that the notion of accessibility and *i*-within-*i* condition allow us to take care of cases when an anaphor is embedded within the subject of a sentence. For instance, in example (14-a), repeated below as (19), AGR is related to the subject of the embedded clause, *pictures of himself*. Coindexation between the anaphor and subject would therefore result in a violation of the *i*-within-*i* condition, and as a result, the embedded AGR is irrelevant for the definition of the Governing Category of *himself* in (20); the anaphor embedded within the subject of the embedded clause is allowed to be free in the embedded clause. Intuitively, the domain in which an anaphor is bound is the minimal domain with specified  $\phi$ -features, and if this is impossible by accessibility restriction, the domain is extended until the next higher locus of  $\phi$ -features.

(20) \*John<sub>i</sub> said that [himself<sub>i</sub> likes Mary]

On the other hand, coindexation of AGR and the subject anaphor as in (20) does not violate the *i*-within-*i* condition, and therefore the binding domain of the reflexive *himself* is just the embedded clause. Coindexation with *John* is prohibited.

The cases with possessives can be handled in the same way using accessibility:

- (21) a. John<sub>i</sub> likes [books about himself<sub>i</sub>]  
b. \*John<sub>i</sub> likes [Mary's books about himself<sub>i</sub>].

In the former case, the object DP lacks a SUBJECT, and therefore the governing category for the anaphor coincides with the entire sentence. In the latter case, the object DP has a SUBJECT, *Mary's*, and therefore *himself* must be bound within the object phrase, and thus *himself* cannot be coreferential with *John*.

Further, the condition on government category to include a governor allows Chomsky to take care of ECM cases:

(22) John<sub>i</sub> expects himself<sub>i</sub> to win the lottery.

The governor of *himself* is the matrix verb *expects*, and therefore the binding domain includes the matrix subject, which can bind the reflexive, and the sentence is correctly predicted to be grammatical.

### 1.1.3.2 Long-distance binding: domains revised

All the cases considered in the previous section are cases of either local binding, when the antecedent and the reflexive are located within the case clause, or cases of reflexives embedded within subjects — and in the latter case the matrix subject can be the antecedent of the reflexive because there are no other alternative binders which can satisfy the accessibility condition on SUBJECT.

Anaphoric binding is, however, not always local. A number of languages allow anaphors to be anteceded by DPs in the higher clauses — in some cases there are virtually no locality restrictions on binding, while in some cases a long-distance relationship between an anaphor and its antecedent must satisfy certain conditions.

The following cases from Russian exemplify such situation:

- (23) a. Ivan<sub>i</sub> poprosil menja<sub>j</sub> PRO<sub>j</sub> narisovat' svoj<sub>i</sub> portret.  
Ivan asked me draw<sub>INF</sub> self's portrait  
'Ivan asked me to draw his portrait'
- b. Ivan<sub>i</sub> poprosil menja<sub>j</sub> PRO<sub>j</sub> narisovat' sebja<sub>i</sub>.  
Ivan asked me draw<sub>INF</sub> self  
'Ivan asked me to draw him.'

Both of these cases have an embedded infinitival clause and an anaphor in the embedded object position (either the anaphor is the object, or it is embedded into the object). While the reflexive is bound by the matrix subject avoiding binding by the infinitival PRO, the sentences are grammatical. In order to revise the government-based approach to account for possibility of (23), one would have to redefine the SUBJECT, possibly parametrizing it among various languages whether PRO

is accessible or not. This however seems like an ad hoc solution to the problem, considering that this situation is quite common cross-linguistically, and not only in case of an infinitival embedded clause.

The following examples from Italian and Icelandic show that binding of a reflexive is possible across a subjunctive clause boundary:

- (24) a. Jón<sub>i</sub> segir að María elski sig<sub>i</sub>. (Icelandic)  
 Jon said that Maria loves<sub>SUBJ</sub> self  
 ‘John said that Maria loves him’ (Sigurðsson 1990)
- b. Gianni pensava che quella casa appartenesse ancora alla propria famiglia. (Italian)  
 Gianni<sub>i</sub> thought that this house belonged still to self’s<sub>i</sub> family.  
 ‘Gianni thought that this house still belonged<sub>SUBJ</sub> to his family.’ (Giorgi 1983)

Subjects of the embedded subjunctive clauses are marked as nominative, and also show agreement with embedded verbs, and therefore qualify as SUBJECTs, regardless of definition. At the same time, binding across them is possible in these languages.

The most extreme cases are however demonstrated by examples from Chinese, Japanese, and Korean:

- (25) a. Zhangsan<sub>i</sub> renwei Lisi<sub>j</sub> hai-le ziji<sub>i</sub> (Icelandic)  
 Zhangsan thought Lisi hurt-ASP self  
 ‘Zhangsan<sub>i</sub> thought that Lisi hurt him<sub>i</sub>’ (Huang and Tang 1991)
- b. John<sub>i</sub>-wa [ Bill-ga zibun<sub>i</sub>-o nikundeiru ] -to omotteiru (Japanese)  
 John-TOP Bill-NOM self-ACC hates -THAT thinks  
 ‘John<sub>i</sub> thinks that Bill hates him<sub>i</sub>’. (Manzini and Wexler 1987)

In (25), reflexives are located in indicative embedded clauses, and are bound by an antecedent located outside of their clause. Clearly, the definition of binding domain needs to be revised. Harbert 1995 proposes that there is a possible correlation between the lack of overt agreement features on the verb and availability of long-distance binding.

There are several ways to deal with examples similar to the ones above. First, it is possible to

claim that long-distance reflexives do not need to follow traditional binding theory, i.e. they do not need to be subject to Principle A. Such an approach was advocated by Giorgi 1983 and Reinhart and Reuland 1993 among others. According to them, these elements while phonologically identical to reflexives, are in fact *logophors*. I will discuss this approach and problems associated with it below in section 1.2.3.

The second approach stems from the Principles and Parameters framework and its major proposal is to parametrize the definition of binding domain or government category. In English government category is defined by presence of a SUBJECT, while in other languages government category can be defined as a larger domain, i.e. it can for instance “skip” over infinitival or subjunctive clauses, or be as large as a root clause. Such an approach was advocated by Manzini and Wexler 1987 who proposed the following definition of government category.

- (26)  $\gamma$  is a governing category for  $\alpha$  iff  $\gamma$  is the minimal category that contains  $\alpha$  and a governor of  $\alpha$  and
- a. can have a subject, or, for  $\alpha$  =anaphor, has a subject  $\beta$ ,  $\beta \neq \alpha$ ; or
  - b. has an INFL; or
  - c. has a Tense; or
  - d. has a “referential” Tense; or
  - e. has a “root” Tense.

According to Manzini and Wexler 1987, languages differ in which definition of government category is chosen. English chooses (26-a), Russian, which allows long-distance binding out of infinitival clauses, chooses (26-c), while Icelandic and Italian, which allow binding into subjunctive clauses (claimed to have deficient Tense), choose option (26-d). East Asian languages with virtually unconstrained binding possibilities choose option (26-e).

An obvious problem with this approach is that the choice of this parameter is arbitrary: it is unclear whether this parameter is correlated with any other parameter of language. While this approach provides a description of what is happening in a particular language, it fails to explain



why a language behaves in this particular way. Further, it cannot capture certain generalizations: for instance, there is no explanation of why only monomorphemic reflexives can choose a non-local binding option.

The third approach to long-distance anaphors is to reduce it to a local relationship between an anaphor and its antecedent by postulating anaphor movement (at LF) to enter in a local configuration with its antecedent. In this case the difference between English on the one hand and Japanese on the other lies in constraints on this LF-movement of anaphors. For some reasons this movement is not allowed in English, but possible in Japanese. I discuss this approach in greater details in the next section.

#### **1.1.4 Monomorphemic anaphors: subject condition and long-distance binding**

The idea of LF movement of anaphors was originally explored by Pica 1987 in order to explain the following properties of long-distance anaphors, which could not be explained under the Manzini and Wexler 1987 approach:

- (27) a. Monomorphemicity  
b. Subject-orientation

Monomorphemicity of anaphors refers to the contrast between English-type anaphors like *himself*, which are composed of a pronominal element *him* and an anaphoric element *self*, and anaphors of the Russian or Danish style, such as *sig*, which does not have a pronominal element and has no  $\phi$ -features.

Subject-orientation of anaphors refers to the fact that only subjects can be their antecedents, and not objects. This contrast is shown in (28); Russian *sebe* can only be bound by the subject, while English *himself* can be bound by either subject or object, making a corresponding example ambiguous:

- (28) a. John<sub>i</sub> told Bill<sub>j</sub> about himself<sub>i/j</sub>.  
 b. Ivan<sub>i</sub> rasskazal Borisu<sub>j</sub> o sebe<sub>i/\*j</sub>.  
 Ivan.NOM told Boris.DAT about self  
 ‘Ivan<sub>i</sub> told Boris<sub>j</sub> about himself<sub>i/\*j</sub>.’

Pica 1987 proposes that monomorphemic anaphors are heads ( $X^0$ ), and can undergo LF head-movement to INFL. Considering that binding is evaluated after that raising, the only element which c-commands the anaphor is now the subject, and therefore binding by the object is not allowed.

Complex anaphors cannot undergo such head movement because they are phrases (XP) and not heads. Considering that all anaphors have to undergo LF-raising, Pica proposes that they adjoin to the category they are contained in, i.e. VP. Adjunction to VP does not prevent them from remaining c-commanded by the object, and therefore binding by an object is possible.

LF head-raising of monomorphemic anaphors explains why only they can participate in long-distance binding. According to Pica’s analysis, long-distance binding is possible only if a reflexive has raised to a position in a local configuration with its antecedent. Only monomorphemic reflexives can undergo such long head-movement (under certain conditions, potentially depending on the properties of the complementizer), and therefore only they can have an antecedent outside of their base clause. The fact that only subjects can be antecedents of long-distance anaphors follows from the fact that monomorphemic anaphors must be interpreted at INFL, and therefore even if they raise to a higher clause, they still have to go to at least INFL, and cannot stay, say in C/V position.

### 1.1.5 Psych-verbs: Belletti and Rizzi 1988

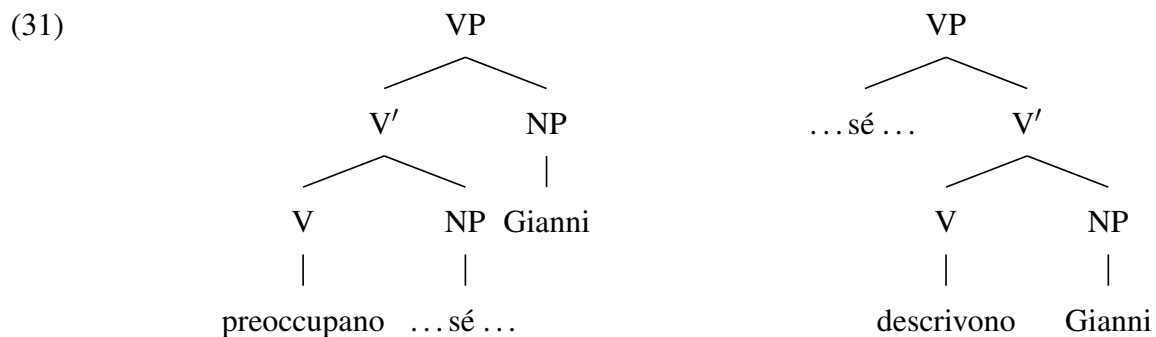
Another property of anaphors which any theory must be able to account for is their interaction with psych-verbs. Belletti and Rizzi 1988 notice that the object “experiencer of psych verbs can bind an anaphor contained within the subject,” even though the c-command relation does not hold. This property is particular to psych-verbs, as the following examples from Belletti and Rizzi 1988 show.

- (29) a. Questi pettegolezzi su di sé<sub>i</sub> preoccupano Gianni<sub>i</sub> più di ogni altra cosa  
 These gossips about self worry Gianni more than any other thing  
 ‘These gossips about himself worry Gianni more than anything else.’
- b. \*Questi pettegolezzi su di sé<sub>i</sub> descrivono Gianni<sub>i</sub> meglio di ogni biografia ufficiale  
 These gossips about self describe Gianni better than any biography official  
 ‘These gossips about himself describe Gianni better than any official biography.’

The sentence with the verb *describe* behaves as expected: it is ungrammatical because there the anaphor is not bound by its antecedent. The grammaticality of the former sentence with the verb *worry* is unexpected: the absence of c-command (and even precedence) does not result in an unacceptable sentence. Similar results hold for English *himself*:

- (30) a. Rumors about himself<sub>i</sub> worry John<sub>i</sub>.  
 b. \*Rumors about himself<sub>i</sub> describe John<sub>i</sub> in details.

Belletti and Rizzi (1988) argue that this asymmetry is due to a difference between the deep structures of these two sentences. According to their analysis, the following two structures explain this difference:



Belletti and Rizzi propose that Principle A can be satisfied at deep structure, and does not have to be satisfied at surface structure. Therefore, in the case of psych-verbs, Principle A is satisfied at the level of VP, before a constituent with the anaphor raises to the subject position; this further movement of the subject will not trigger the sentence ungrammatical. On the other hand in the case

of a regular transitive verb there is no moment of the derivation when the reflexive is c-commanded by its antecedent. Therefore the sentence is ungrammatical.

They further propose a view of Principle A as an “anywhere” principle: “it suffices for principle A to be met somewhere, either at D-structure or S-structure or, perhaps, LF.” While this approach is supported by a number of cases, including *wh*-movement, (32), and raising, (33), below in my discussion of example (25) on page 115 I explore why an “anywhere” application of principle A must in fact be limited.

(32) Which picture of himself<sub>i</sub> do you think that Bill<sub>i</sub> likes  $\tau$  best]?

(33) Replicants of themselves<sub>i</sub> seemed to the boys<sub>i</sub> [ $\tau$  to be ugly].

## **1.2 Revising binding theory: Reinhart and Reuland 1993**

### **1.2.1 Reasons for abandoning classical binding theory**

An approach different from the traditional syntactic-based approach to binding theory was advocated for by Reinhart and Reuland 1993 in their “Reflexivity” paper. They propose that binding theory is not a condition on particular elements, such as anaphors and pronominals, but it is a theory about the reflexivity of predicates. According to them, “Conditions A and B govern the well-formedness and the interpretation of reflexive predicates in natural language.” Considering that binding theory Principles A and B are conditions on interpretation, it is therefore tempting to find a natural semantic explanation for them. Further, locality conditions on binding can potentially be viewed as conditions on coarguments of a predicate. Below I describe the approach by Reinhart and Reuland 1993.

### **1.2.2 Binding theory as condition on reflexivity of predicates**

Reinhart and Reuland 1993 start off by making a distinction between two major types of anaphors: what they call SE-type anaphors, which are monomorphemic (such as Dutch *zich*), and complex SELF-type anaphors, such as English *himself* or Dutch *zichzelf*. Anaphors of the former type lack  $\phi$ -feature specification (at least for number and gender), while anaphors of the latter type consist of a pronominal or SE-type anaphor combined with the nominal head SELF. The corresponding structures are given in (34).

- (34) a. **SE-type anaphors:** [<sub>NP</sub> SE [<sub>N'</sub> ... e ... ]]  
 b. **SELF-type anaphors:** [<sub>NP</sub> PRON/SE [<sub>N'</sub> ... SELF ... ]]

Binding is viewed as assigning a content to otherwise referentially defective NPs. Only the reflexivizing function of anaphors is governed by binding theory, where by reflexivizing Reinhart and Reuland understand a restriction on coarguments of a predicate. All other aspects of anaphoric expressions do not belong to binding theory, and fall under other modules of grammar. As a result of this approach, the problem of defining a binding domain is solved: only coarguments participate in binding theory, and any long-distance coreference between an antecedent and a reflexive no longer needs to be explained within binding theory. Any use of reflexives for reasons other than establishing coreference between coarguments of a predicate is called *logophoric*, and is not subject to binding theory (see section 1.2.3 below to discussion of logophors).

In order to state their conditions A and B, Reinhart and Reuland propose the following definition of a reflexive predicate and reflexively-marked predicate

- (35) a. A predicate is *reflexive* iff (at least) two of its arguments are coindexed.  
 b. A predicate is *reflexively marked* iff either  
 (i) one of its arguments is a *self*-anaphor; or  
 (ii) a predicate is lexically reflexive.

The idea behind reflexive marking of a predicate is the following: the first option is for a predicate to come from the lexicon with a requirement on its arguments to be coindexed (like English *behave* or Dutch *shaamt* ‘shame’); this can be viewed as an operation on the  $\theta$ -grid of a verb. The second option is to turn a regular transitive predicate into a reflexive, and for that one of the predicate’s arguments must be a SELF-anaphor.

Binding theory puts these two definitions together by requiring the following conditions to hold:

- (36) a. **Principle A:** A reflexive-marked (syntactic) predicate is reflexive.  
b. **Principle B:** A reflexive (semantic) predicate is reflexively marked.

Abstracting away for now from predicate types (semantic vs. syntactic), these principles can account for the following.

(37) \*John<sub>i</sub> loves himself<sub>j</sub>.

In (37), the predicate is reflexively marked by the presence of the SELF-anaphor, and therefore it must be reflexive, i.e. its arguments must be coindexed. As this is not the case in (37), the sentence is ungrammatical under the given coindexation.

(38) \*John<sub>i</sub> loves him<sub>i</sub>.

In (38), the predicate is reflexive since its arguments are coindexed. It is however not reflexively marked: neither is the verb *love* inherently reflexive nor is any of its arguments a SELF-anaphor. Example (38) is therefore ungrammatical under the given coindexation.

While conditions A and B defined as above can take account for basic binding cases, nothing prevents anaphors in subject position, as no reference is made to c-command or precedence in these principles. For instance, neither condition predicts the ungrammaticality of example (39):

(39) Himself<sub>i</sub> loves John<sub>i</sub>.

The predicate in (39) is both reflexive, since its coarguments are coindexed, and reflexively marked, since one of the arguments is a SELF-anaphor. According to Reinhart and Reuland 1993, the problem with (39) is not a result of violating binding theory. In order to account for the ungrammaticality of (39), Reinhart and Reuland 1993 introduce the *Chain Condition*:

- (40) a. An *A-chain*, is any sequence of coindexation that is headed by an A-position and satisfies antecedent government; that is, each coindexed link, except for the head, is c-commanded (i.e., m-commanded) by another link, and there is no barrier between any two of the links.
- b. **Chain Condition:** A maximal A-chain  $(\alpha_1, \dots, \alpha_n)$  contains exactly one link —  $\alpha_1$  — that is both +R and Case-marked.
- c. An NP is +R iff it carries a full specification for  $\phi$ -features [(specifically, number and gender)] and structural Case.

In (39) the chain consisting of two links, (himself<sub>i</sub>, John<sub>i</sub>) violates the Chain Condition: the first link of this chain is not +R. Similarly, the chain condition would also predict the ungrammaticality of the following examples:

- (41) a. \*John<sub>i</sub> loves him<sub>i</sub>
- b. \*He<sub>i</sub> assigned him<sub>i</sub> to himself<sub>i</sub>.
- c. \*Henk<sub>i</sub> wees hem<sub>i</sub> aan zichzelf<sub>i</sub> toe. (Dutch)  
Henk assigned him to himself  
'Henk<sub>i</sub> assigned himself to himself<sub>i</sub>'

In all these cases, the pronouns are +R, and therefore the chain condition is violated. Notice that only (41-a) violates condition B. Both (41-b) and (41-c) are ok with respect to Principle B: the predicates in these examples are reflexive, and they are also reflexively marked because one of the arguments is a SELF-reflexive.

Above I showed that Reinhart and Reuland must crucially rely on the Chain Condition in order to explain facts which are usually explained by binding conditions. The next important concept for Reinhart and Reuland 1993 is the difference between syntactic and semantic predicates.

- (42) a. The *syntactic predicate* formed of (a head) P is P, all its syntactic arguments, and an external argument of P (subject). The *syntactic arguments* of P are the projections assigned  $\theta$ -role or Case by P.
- b. The *semantic predicate* formed of P is P and all its arguments at the relevant semantic level.

The importance of these definition stems from the fact that different binding theory principles apply to different types of predicates.

Condition A applies to syntactic predicates and Condition B applies to semantic predicates. Consider the following sentences:

- (43) a. \*The queen invited myself for tea.
- b. The queen invited both Max and myself/me for tea.
- c. Max said that the queen invited both Lucie and himself/him for tea.

Only (43-a) violates condition A, as the predicate is reflexively marked, but not reflexive. In examples (43-b)-(43-c) there are no reflexive (syntactic) predicates, and also neither of the predicates is reflexively marked, and the reflexives are used logophorically, i.e. they are not subject to binding theory (see section 1.2.3 for more discussion of logophors).

Now consider the sentences in (44).

- (44) a. The queen<sub>i</sub> invited both Max and herself<sub>i</sub> to our party.
- b. \*The queen<sub>i</sub> invited both Max and her<sub>i</sub> to our party.



Again, in both (44-a)-(44-b) neither (syntactic) predicate is reflexive nor reflexively marked. There are no violations of binding theory principles. Why is (44-b) ungrammatical? Reinhart and Reuland's assumption that Principle B operates on semantic predicates takes care of this example. The semantic representation of (44-b) shown in (45) contains a reflexive predicate:

(45) the queen ( $\lambda x (x \text{ invited Max \& } x \text{ invited } x)$ )

This reflexive predicate must be reflexively marked, but there is no SELF-anaphor present in (44-b). The sentence violates Principle B, and is therefore ungrammatical.

Condition A on the other hand does not care about semantic predicates: sentence (43-b) has the following semantic representation:

(46) the queen ( $\lambda x (x \text{ invited Max \& } x \text{ invited myself})$ ).

One of the predicates in (46) is reflexively marked, however none of them is reflexive. If Principle A applied to semantic predicates it would incorrectly rule this sentence out. Therefore, Reinhart and Reuland conclude that Principle A applies to syntactic predicates.

A further important difference between semantic and syntactic predicates comes from consideration of ECM sentences:

- (47) a. Lucie expects [herself to entertain herself].  
b. \*Lucie expects [myself to entertain myself].  
c. \*Lucie<sub>i</sub> expects [her<sub>i</sub> to entertain herself].

Sentence (47-a) satisfies Principle A since it applies to syntactic predicates. The use of logophors is not allowed as shown in (47-b): an anaphor would reflexive-mark the matrix verb while there is no reflexive predicate. Lastly and crucially, Principle B does not rule out (47-c) since it only applies to semantic predicates. It is only the Chain Condition which explains ungrammaticality of (47-c).

In my review of Reinhart and Reuland 1993 I mostly concentrated on how their theory can be applied to English and complex reflexives. I will take on their treatment on SE-type reflexives below in section 1.2.4.1, while now I explore their treatment of logophors.

### 1.2.3 Logophors

#### 1.2.3.1 What are logophors?

Above I described three types of nominal elements: R-expressions, pronominals, and anaphors. According to Principle A of the binding theory (whether in its traditional formulation or in the reflexivity-based framework of Reinhart and Reuland 1993) anaphors must be locally bound by their antecedent (or antecedents and anaphors must be coarguments of the same predicate). Clements 1975 discusses another particular class of nominal elements in Ewe. These forms, which are different from pronouns and reflexives “distinguish reference to the individual whose speech, thoughts, or feelings are reported or reflected in a given linguistic context, from reference to other individuals.” These items do not need to be c-commanded by a coreferring R-expression, and rely for their identification on point of view, discourse and other non-syntactic notions.

- (48) *Kɔmi*<sub>i</sub> xɔ agbalẽ tso *Kofi*<sub>j</sub> gboɔ be wò-a-va me kpe na yè\*<sub>i/j</sub> (Ewe)  
 Kwami receive letter from Kofi side that Pro-T-come case block for LOG  
 ‘Kwami got a letter from Kofi<sub>j</sub> saying that he should case some blocks for him<sub>j</sub>.’

Even though the subject *Kɔmi* c-commands the logophoric pronoun *yè*, and *Kofi* does not, as Clements notes, for some speakers the only possible referent of the logophoric pronoun in *Kofi*. Further, antecedent of the logophor does not even need to be present in the same sentence. Therefore there seem to be no syntactic requirements on logophoric pronouns, or at least no binding theory requirements.

While there are no universal syntactic constraints on logophors, their distribution can still be restricted by a set of constraints particular to a language.

For Reinhart and Reuland 1993, reflexive pronouns, which do not satisfy the reflexivity version of Principle A, are also logophors. For instance, any instance of long-distance reflexive binding involves logophors, even if they are phonologically the same as anaphors. For instance, as we have seen, in Icelandic binding is normally possible into a subjunctive embedded clause, see (49).

- (49) a. Jón<sub>i</sub> segir að Pétur<sub>j</sub> raki sig<sub>i/j</sub>.  
 John say.3SG that Peter shave.SUBJ.3SG self  
 ‘John says that Peter shaves himself.’
- b. Jón<sub>i</sub> veit að Pétur<sub>j</sub> rakar sig<sub>\*i/j</sub>.  
 John know.3SG that Peter shave.IND.3SG self  
 ‘John knows that Peter shaves himself.’

The pronoun *sig* can have its antecedent outside of its clause, if this clause is subjunctive; this is impossible and results in ungrammaticality if the embedded clause is indicative. Since *sig* and its antecedents are not coarguments, in this case *sig* must be a logophor, and binding theory does not have to explain its behavior. Binding theory turns into a theory of logophors.

I summarize some contexts where logophors appear for Reinhart and Reuland 1993 in (50):

- (50) a. Anaphors in non-argument positions;  
 b. Anaphors with an antecedent outside of their clause;  
 c. Anaphors which are not c-commanded by their antecedents;  
 d. Anaphors which lack an antecedent within the same sentence.

While all of these cases are problematic for the reflexivity-based theory of binding, traditional binding theory can handle the first environment and after some modifications (e.g. allowing LF-movement of anaphors) can handle the second environment. The last two environments are the most problematic for a syntactic approach to binding.

Sells 1987 proposes the following three discourse roles which are needed to describe the reference of a logophor.

- (51) a. SOURCE: the intentional agent in a communication  
 b. SELF: one whose mental state is described in the content of the proposition  
 c. PIVOT: one with respect to whose location the content of the proposition is evaluated.

SOURCE can be diagnosed by usage of “speaker-evaluative” phrases such as “that fool”. in (52) Takasi is the SOURCE because it is he who calls Yosiko “that fool.”

- (52) Takasi<sub>i</sub>-wa Taroo-ni [baka-no Yosiko-ga zibun<sub>i</sub>-o oikake-mawasiteriru koto] -o  
 Takasi-TOP Taroo-DAT [fool-GEN Yosiko-NOM self-ACC chase-around-be COMP] -ACC  
 hanasita  
 told  
 ‘Takasi<sub>i</sub> told Taroo that that fool Yosiko was following him<sub>i</sub>.’

The diagnostics for SELF is the adverb “mysteriously.”

- (53) Takasi-wa [Yosiko-ga hukakainimo ato o take-mawasiteitu to] itta.  
 Takasi-TOP [Yosiko-NOM mysteriously be-following COMP] said  
 ‘Takasi<sub>i</sub> said that Yosiko was mysteriously following him<sub>i</sub>.’

Here Takasi is the *self*, since the action described in the embedded clause is mysterious *to him*.

Lastly, PIVOT is shown in the following example:

- (54) Takasi<sub>i</sub>-wa [Yosiko-ga zibun<sub>i</sub>-o tazunete-kita node] uresigatta.  
 Takasi-TOP [Yosiko-NOM self-ACC visit-came because] happy  
 ‘Takasi<sub>i</sub> was happy because Yosiko came to visit him<sub>i</sub>.’

In this case, Takasi is a PIVOT, and the sentence is grammatical. At the same time, when the verb is changed to *visit-went* instead of *visit-came*, as in (55), the sentence is ungrammatical:

- (55) \*Takasi<sub>i</sub>-wa [Yosiko-ga zibun<sub>i</sub>-o tazunete-itta node] uresigatta.  
 Takasi-TOP [Yosiko-NOM self-ACC visit-went because] happy  
 ‘Takasi<sub>i</sub> was happy because Yosiko went to visit him<sub>i</sub>.’

I will not go into further details of how Sells's system works, but according to him languages differ in which of these discourse categories can or cannot be an antecedent for a logophoric pronoun.

### 1.2.3.2 The treatment of logophors in Reuland 2001

The last thing which needs to be discussed with respect to logophors is why they are sensitive to syntactic conditions. For example, as I showed above, in Icelandic logophors are only possible in embedded subjunctive clauses, but not in embedded indicatives.

This is a question taken on by Reuland 2001:

(56) Why, under certain conditions, is logophoric interpretation of anaphors blocked?

Reuland proposes that logophoric interpretation is blocked just in case there is some other possible mechanism to interpret it. Formally, he proposes the following Rule of Logophoric interpretation:

(57) **Rule L – Logophoric Interpretation:** NP A cannot be used logophorically if there is a B such that an A-CHAIN  $\langle B, A \rangle$  can be formed.

The relevant definitions of CHAIN are given in (58).

- (58) a. **Chain:**  $(\alpha, \beta)$  for a *Chain* if (a)  $\beta$ 's features have been (deleted by and) recovered from  $\alpha$ , and (b)  $(\alpha, \beta)$  meets standard conditions on chains such as uniformity, c-command, and locality.
- b. **CHAIN:** If  $(\alpha_1, \alpha_2)$  is a Chain and  $(\beta_1, \beta_2)$  is a chain and  $\alpha_2 = \beta_1$ , then  $(\alpha_1, \alpha_2 / \beta_1, \beta_2)$  is a *CHAIN*.

How can this definition account for availability of logophors in subjunctive clauses in Icelandic? According to Reuland 2001, subjunctives must block chain formation between anaphor and antecedent. He assumes that subjunctives are licensed by an operator which attracts features of V/I-complex. The structures before and after feature movement are shown in (59):

- (59) a. Op [<sub>IP</sub> Oscar [[<sub>I,I</sub> FF<sub>sig</sub> [<sub>I,I</sub> V<sub>subj</sub> I]] [t<sub>Oscar</sub> [t<sub>v</sub> [sig ... ]]]]]  
 b. [<sub>Op,Op</sub> [<sub>I,I</sub> FF<sub>sig</sub> [<sub>I,I</sub> FF<sub>vsubj</sub> I]]<sub>i</sub> Op] [<sub>IP</sub> Oscar<sub>φ</sub> ... [[δ] V<sub>subj-i</sub>] ... sig<sub>φ</sub> ...]

As a result of this feature movement, there are no formal features left between *Oscar* and *sig* at position  $\delta$ . A CHAIN cannot be formed between *Oscar* and *sig*, and therefore *sig* can get its interpretation from the discourse.

Alternatively, only the features FF<sub>vsubj</sub> can move to Op, and then a CHAIN between *Oscar* and *sig* would be possible to establish; *sig* would be bound by the embedded subject.

There are a few problems with this approach. First, it is unclear why both derivations are equally possible. It seems like the derivation in which less movement is made, i.e. the second derivation, would be more economical than the first derivation.

Second, it is unclear how one can guarantee that the antecedent of the logophor is the matrix *subject*, and not any other element in the matrix clause. Of course, one can possibly answer this question by referring to discourse prominence of the subject as opposed to other matrix elements, similar to the theory proposed in Sells 1987. This answer however misses the important fact that logophors in languages with subject oriented anaphors are also subject-oriented. It seems like if one proposes distinct theories for anaphors and logophors, subject-orientation would have to be derived separately, and not necessarily in the same way, creating a significant theoretical redundancy.

## 1.2.4 Problems

### 1.2.4.1 The nature of SE

As opposed to SELF-type reflexives, SE-type reflexives are not subject to binding theory for Reinhart and Reuland. According to them, SE-type reflexives do not have reflexivizing function, and in that they are similar to pronouns, nor are they referentially independent (i.e. SE-type reflexives are specified as -R), in which they are similar to reflexives. This is summarized in table (60).

	SELF	SE	Pronominals
(60) Reflexivizer	✓	✗	✗
Referential	✗	✗	✓

The fact that SE-reflexives do not reflexivize their predicates, means that they behave similar to pronouns with respect to Condition B.

- (61) a. \*Max haat zich. (Dutch)  
 Max hates SE  
 ‘Max hates himself’
- b. \*Max praat met zich.  
 Max speaks with SE  
 ‘Max speaks with himself’

SE can be used with some verbs, but as Reinhart and Reuland 1993 claim, only with the ones that can be marked as intrinsically reflexive in the lexicon, such as *schamen* ‘shame’ and *wassen* ‘wash’. Here the former verb is intrinsically reflexive, while the latter verb is listed in lexicon twice: as a reflexive verb and as a transitive verb.

- (62) a. Max wast zich.  
 Max washes SE  
 ‘Max washes himself’
- b. Max schaamt zich.  
 Max shames SE  
 ‘Max is ashamed.’

In (62), both verbs are intrinsically marked as reflexive, and therefore there is no need to use a SELF-anaphor: condition B is not violated, the sentences are grammatical with SE. The verbs in (61) cannot be marked intrinsically as reflexive, and therefore the resulting reflexive predicate is not reflexively marked, which is a violation of condition B. Notice that even though the sentences in (62) are grammatical with SE-type anaphors, the corresponding sentences with pronominals (63)

would be ungrammatical, but not because of a binding theory violation, but because of the Chain Condition.

- (63) a. \*Max<sub>i</sub> wast hem<sub>i</sub>.  
Max washes him  
'Max washes himself'
- b. \*Max<sub>i</sub> schaamt hem<sub>i</sub>.  
Max shames him  
'Max is ashamed.'

Another confirmation for their theory comes from the fact that when one of the arguments of the verb is a SELF-anaphor, the other argument could be a SE-anaphor. SELF marks a predicate as reflexive, therefore there will be no condition B violation. Further, since SE is not a +R element, the Chain Condition would also hold:

- (64) Henk<sub>i</sub> wees zich<sub>i</sub> aan zichzelf<sub>i</sub> toe.  
Henk assigned SE to himself  
'Henk<sub>i</sub> assigned himself<sub>i</sub> to himself<sub>i</sub>.'

SE-type anaphors also show interesting behavior in ECM contexts, and Reinhart and Reuland 1993 discuss it based on their distinction between syntactic and semantic predicates.

- (65) a. \*Jan<sub>i</sub> hoorde zich<sub>i</sub> / hem<sub>i</sub>.  
Jan heard SE / him  
'John heard himself'
- b. Jan<sub>i</sub> hoorde [zich<sub>i</sub> zingen ].  
Jan heard [SE sing ]  
'John heard himself sing'
- c. \*Jan<sub>i</sub> hoorde [hem<sub>i</sub> zingen ].  
Jan<sub>i</sub> heard [him<sub>i</sub> sing ]  
'John heard himself sing'
- d. Jan<sub>i</sub> hoorde [zichzelf<sub>i</sub> zingen ].  
Jan heard [himself sing ]  
'John heard himself sing'



Example (65-a) is ungrammatical because of a Principle B violation. Neither of the following sentences, however, violates Principle B, since it operates on semantic predicates. Example (65-c) violates the Chain Condition, since the pronoun *hem* is +R. Such problems do not arise with respect to example (65-b), since SE is -R. Lastly, Condition A is relevant for (65-d), since it operates on syntactic predicates: matrix syntactic predicate is reflexive and is also reflexively marked.

The theory of Reinhart and Reuland 1993 accounts for Germanic facts about SE-anaphors, however *se*-anaphors do not behave in the same way in other languages, such as Russian. Russian anaphor *sebjja* lacks  $\phi$ -features, and is referentially dependent. As a result, one would expect it to behave in a way similar to Dutch reflexive.

Consider the following Russian data:

- (66) a. Ivan nenavidit sebjja  
 Ivan hates SE  
 ‘Ivan hates himself’
- b. Ivan razgovarivaet s soboj  
 Ivan speaks with SE  
 ‘Ivan speaks with himself’

These Russian sentences are parallel to the Dutch sentences in (61). They are however grammatical. If one assumes that in Russian as in Dutch *sebjja* does not mark predicate as reflexive, it is unclear why Condition B is not violated in (66). There are several ways out of this problem.

One can assume that *sebjja* is not a *se*-type anaphor, and in fact is a *self*-type anaphor. This would explain the grammatical status of examples in (66), however this solution lacks any predictive power. Russian anaphor behaves in many ways similarly to Dutch: it is monomorphemic, exhibits subject orientation, and can be used in long-distance contexts. None of these properties is characteristic of English *himself*.

Another solution to this problem would be to assume that any verb in Russian can be intrinsically reflexively marked. It is unclear how this can be accounted for formally. Assuming this as a condition in the lexicon again lacks predictive power: it is unclear why lexicons in Dutch and

Russian would differ in this respect.

#### 1.2.4.2 Possessive anaphors

Since Reinhart and Reuland 1993 limit the domain of binding theory to argument positions, their theory cannot account for distribution of possessive anaphors, such as examples in (67) from Russian:

- (67) a. Ivan<sub>i</sub> razgovarivaet so svoej<sub>i</sub> ženoj  
Ivan speaks with self's wife  
'Ivan<sub>i</sub> speaks with his<sub>i</sub> wife'
- b. Ivan<sub>i</sub> rasskazal Borisu<sub>j</sub> o svoej<sub>i/\*j</sub> žene  
Ivan told Boris about self's wife  
'Ivan<sub>i</sub> told Boris<sub>j</sub> about his<sub>i/\*j</sub> wife'
- c. \*Ivan<sub>i</sub> skazal čto svoja<sub>i</sub> žena ušla k drugomu  
Ivan said that self's wife left to other  
'Ivan said that his wife left to someone else'

Under Reinhart and Reuland's 1993 theory, example (67-a) is expected to be grammatical, since there are no reflexive predicates, and none of the predicates is reflexively marked. Therefore, this example shows that *svoj* in Russian can be used logophorically. The next example (67-b) is only grammatical when the reflexive refers to the subject: it exemplifies subject-orientation of possessive anaphors in Russian. It behaves in exact same way as an argument anaphor in Russian. The binding theory of Reinhart and Reuland 1993 cannot explain this behavior, even though one can claim that it is a condition of logophors that they can only be coreferenced with the most prominent argument, i.e. the subject. Finally, the ungrammaticality of (67-c) also cannot be explained: if *svoja* is used logophorically, it is unclear why in this example it cannot be coindexed with the matrix subject.

### 1.2.4.3 Subject anaphors

One further problem with the reflexivity-based account of binding is subject anaphors. As pointed out by Bruening 2006 based on Sung 1990, a number of East Asian languages permit local reflexives in subject positions of embedded clauses (examples below are from Sung 1990):

- (68) a. Zhangsan<sub>i</sub> shuo Xiaoming<sub>j</sub> xiangxin ta-ziji<sub>\*i/j</sub> neng kaoguo. (Chinese)  
Zhangsan say Xiaoming believe himself can pass the exam  
'Zhangsan<sub>i</sub> said that Xiaoming<sub>j</sub> believes that he himself<sub>\*i/j</sub> can pass the exam'
- b. John<sub>i</sub>-un [Mary<sub>j</sub>-ka [caki-casin-i ttokttokha-ta] -ko sayngkakha-n-ta] -ko  
John-TOP [Mary-SUBJ [self-NOM be.smart-DECL] -COMP think-PRES-DECL] -COMP  
malha-ess-ta. (Korean)  
tell-PAST-DECL  
'John told that Mary thinks that she is smart'

The important fact about the examples in (68) is that subject anaphors can only have a local antecedent: a root clause subject is not a possible antecedent. These reflexives are also not logophors for one more reason: they still need a c-commanding antecedent. A reflexivity-based theory cannot account for these cases: an anaphor and its antecedent are not coarguments of any predicate.

### 1.2.4.4 Binding and non-argumental positions

Reinhart and Reuland 1993 claim that only arguments of predicates are subject to the binding theory. Anaphors located within adjuncts are logophoric, and therefore should not obey binding theory principles. It is argued that in these positions anaphors and pronominals are not in complementary distribution, such as in (69).

- (69) a. Max<sub>i</sub> put a gun near him<sub>i</sub>/himself<sub>i</sub>.  
b. Max<sub>i</sub> saw a ghost near him<sub>i</sub>/himself<sub>i</sub>.  
c. Max<sub>i</sub> pulled a cart toward him<sub>i</sub>/himself<sub>i</sub>.

If these anaphors are in fact logophors, one would expect that they need neither c-commanding antecedents nor local antecedents. This is however not the case.

- (70) a. Max<sub>i</sub> said that Mary put a gun near him<sub>i</sub>/\*himself<sub>i</sub>.  
b. Max<sub>i</sub> said that Mary saw a ghost near him<sub>i</sub>/\*himself<sub>i</sub>.  
c. Max<sub>i</sub> said that Mary pulled a cart toward him<sub>i</sub>/\*himself<sub>i</sub>.

In (70) the existence of the intervening subject breaks the relation between the anaphor and *Max*. The ungrammaticality of these sentences is unexpected under the reflexivity-based theory of binding. Similar results can be seen if one tries to construct examples with anaphors within PPs which lack a c-commanding antecedent. The following examples are taken from Hestvik and Philip 2001:

- (71) a. Clinton<sub>i</sub>'s car carried a picture of himself<sub>i</sub> on the roof.  
b. \*Clinton<sub>i</sub>'s car backfired/collapsed/explored behind himself<sub>i</sub>.

The first sentence in (71) demonstrates that in logophoric contexts c-commanding antecedents are not necessary. Now, if anaphors within adjunct PPs are also logophors, it is unclear why the second example in (71) is ungrammatical.

## 1.3 Minimalist theories of binding

### 1.3.1 Is binding compatible with minimalism?

The traditional binding theory of Chomsky 1981/1993 has a filter-like nature: the representation(s) of the sentence is (are) evaluated at a certain moment, and if it does not satisfy certain surface criteria, the sentence is marked as ungrammatical. Filter conditions are hard to formulate in current minimalist theories: first, one has to specify at which levels filters are evaluated; however with the lack of LF/SS/DS in current theory one is only left with a derivational approach, when the filter is evaluated at any point of the derivation, and a (multiple) Spell-Out/phasal approach: the filter

is evaluated when a certain “chunk” of the structure is transferred to the interpretative component, i.e. filter is evaluated at the interface.

More problematic is how to evaluate binding conditions. If one assumes that syntactic derivations are based on Merge/Move, and Agree operations, which value features, it is not immediately clear where binding fits in this picture. A number of attempts to reduce binding to feature-valuation and/or to derive binding conditions from other properties of the grammar have been published in the minimalist literature. Below I will present the most important of these.

### **1.3.2 Co-constituent theories of binding**

A particular group of analyses of anaphoric relations comes from the idea of co-constituency. According to these theories, the anaphor gets its reference by starting as a co-constituent with its antecedent, in fact it is often claimed that it is the only way an anaphor can be conceived as coreferential with its antecedent. In this section I explore two such theories, presented in Kayne 2002 and Zwart 2002 respectively.

#### **1.3.2.1 Kayne 2002**

For Kayne 2002, the only way of deriving coreference is through movement: coreference must be established through movement of an antecedent from a “doubling” constituent, both in case of pronominals and reflexives.

According to Kayne 2002, coreference between two expressions is obtained from a base structure where two DPs — one an R-expression, and one a pronominal — start as a “double” constituent. For instance, the derivation of the sentence in (72-a) is given in (72-b).

- (72) a. John<sub>i</sub> thinks he<sub>i</sub> is smart.  
b. thinks [John he] is smart →  
John<sub>i</sub> thinks [<sub>t</sub><sub>i</sub> he] is smart.

While Kayne does not get into details about the precise nature of this “doubling” constituent, it would be reasonable to assume that its category is DP. This entire DP, headed by *he*, receives its  $\theta$ -role in base position, and *John* moves to a higher position in the sentence to receive its own  $\theta$ -role. Similar considerations work for case assignment. The lower nominative case is assigned to the entire “double” DP, and the higher nominative case is assigned to *John*, after it moved out of the “doubling” constituent.

The problem for Kayne starts when he tries to explain ungrammaticality of basic monoclausal sentences with pronominals, such as (73).

- (73) a. \* $\text{John}_i$  thinks highly of  $\text{him}_i$ .  
b. \* $\text{John}_i$  considers  $\text{him}_i$  intelligent.

In order to prevent grammatical derivations of (73) similar to the derivation in (72-b), he proposes the following principle:

- (74) The pronoun (hence the doubling constituent) must move to a position above the subject theta-position (i.e. outside the thematic part of the structure).

This  $A'$ -position (lower Spec,TP) exists in the case of biclausal sentence, like in (72-a), but is absent in case of monoclausal (73-a), since according to Kayne 2002, “there is no appropriate licensing position for the pronoun within VP or between VP and the subject theta-position.” Similarly, in the case of (73-b), such a position is absent because “the small clause is too “small,” i.e. that there is no available intermediate pronoun position within it.”

Now consider Kayne’s account of reflexive binding. As in the case of pronominals, coreference with reflexives must be derived via movement from the relevant co-constituent. Kayne 2002 is vague in his description of this co-constituent formally, only providing a structure as in (75):

- (75) a.  $\text{John}_i$  thinks highly of  $\text{himself}_i$ .  
b. thinks highly of  $D^0$  [John-he] ( $'s$ ) self

Presumably, “doubling” constituent [*John-he*] serves as a possessor of the nominals head *self*. The major question he has to answer is what allows a derivation in which *John* ends up in the Spec,TP position, i.e. a derivation which was not allowed in case of pronominals, since the proposal in (74) must still hold. Kayne argues that a possessive-like structure has a Spec which can be considered an A'-position sufficient to fulfill the requirement of the “doubling” constituent [*John-he*], and therefore further movement of *John* to a subject position in (75) is allowed.

Several problems are evident with this analysis. First, the nature of a “doubling” constituent is unclear. What motivates its creation in syntax? Is it some selectional features of the pronominal that force its Merge with the antecedent? If that is the case, how can one handle the cases of pronominals which do not have linguistic antecedents within the sentence? It seems to me that the requirement that coreference must be established by Merge should be abandoned, in order to allow inter-sentential antecedents. That however will undermine the theory, necessitating to keep the traditional binding principles.

The second problem is the stipulative nature of the condition in (74). How can this condition be formulated within current minimalist framework? Is there any feature of the constituent which must be checked in an appropriate A'-position? Kayne does not provide an answer to this question.

Apart from the unclear nature of a “doubling” constituent and condition (74) being stipulative in its nature, Kayne acknowledges one further problem of this analysis. Since pronouns can have antecedents located very far, i.e. several clauses away from them, he has to allow virtually unbounded movements out of the “doubling” constituent, for instance to derive a sentence like (76).

(76) John<sub>i</sub> thinks that Mary believes that Sue told everybody that Ivana is spreading rumors about him<sub>i</sub>.

However such unbounded long-distance movement of *John* must be prohibited in the case of reflexives:

(77) \*John<sub>i</sub> thinks that Mary believes that Sue told everybody that Ivana is spreading rumors about himself<sub>i</sub>.

Determining the binding domain of a reflexive, one of the most important problems of the binding theory, cannot be resolved by Kayne's analysis. In fact, he cannot even provide an account why binding is mostly a local phenomenon.

### 1.3.2.2 Zwart 2002

Zwart 2002 proposes another analysis of binding deriving coreference from initial Merge, similar to Kayne 2002. However, one of the major differences between his analysis and the one by Kayne stems from the fact that for Zwart only reflexive pronouns merge with their antecedents, and not pronominals. Anaphoric properties of pronouns are in fact derived from that initial Merge: if a pronoun is merged with its antecedent, it is coreferential with the element it is merged with, it acquires a special phonological form, and gets pronounced as a reflexive counterpart. Therefore, the lexicon does not have a special entry for anaphors: pronominals merged with an R-expression *become* reflexive pronouns.

Similarly to Kayne's approach, the antecedent moves out of the "double" in order to get a  $\theta$ -role and case, and the c-command condition on binding is derived as a consequence of the c-command condition on movement.

The fact that merge results in a pronoun being pronounced as a reflexive and the fact that pronominals are not involved in this process eliminates one of the major problems for Kayne, i.e. locality of reflexive binding. From Zwart's theory it follows that at least at a certain moment of the derivation, an anaphor and its antecedent must be in a very local configuration, i.e. they at least must originate in the same clause. At the same time, one would need to explain how long-distance binding can be reduced to the same process.

Consider for instance the French example in (78):



- (78) On<sub>1</sub> souhaite que les gens ne disent pas du mal de soi<sub>1</sub>. (French)  
one wishes that the people NEG speak NEG of ill of self  
'One always wishes that people do not slander oneself.' (Pica 1991)

According to Zwart's theory, one has to assume that *on* and *soi* initially are merged together, and *on* is further moved out of that constituent to the subject position of the matrix clause, where it acquires its  $\theta$ -role and nominative case. This movement is a long-distance movement for the purposes of checking A-features, something which is not generally accepted in current theory: one needs to make further stipulations to allow such movements in only a limited number of situations.

Another problem for this analysis is how it can deal with reflexive possessives.

- (79) The children<sub>i</sub> like each other<sub>i</sub>'s toys.

Zwart would have to assume that *the children* starts together with *each other*, and then moves out to the subject position. This movement from the left specifier of the object DP, i.e. would violate the left-branch condition even though in English such violations lead to ungrammaticality. One would have to explain why this constraint is not active in these cases.

The last thing to mention about this analysis is the fact that it cannot handle condition B. Any instance of pronominal coreference is accidental, and therefore one has to explain why accidental coreference between a pronoun and an antecedent is impossible in examples like (80):

- (80) John<sub>i</sub> likes him<sub>i</sub>.

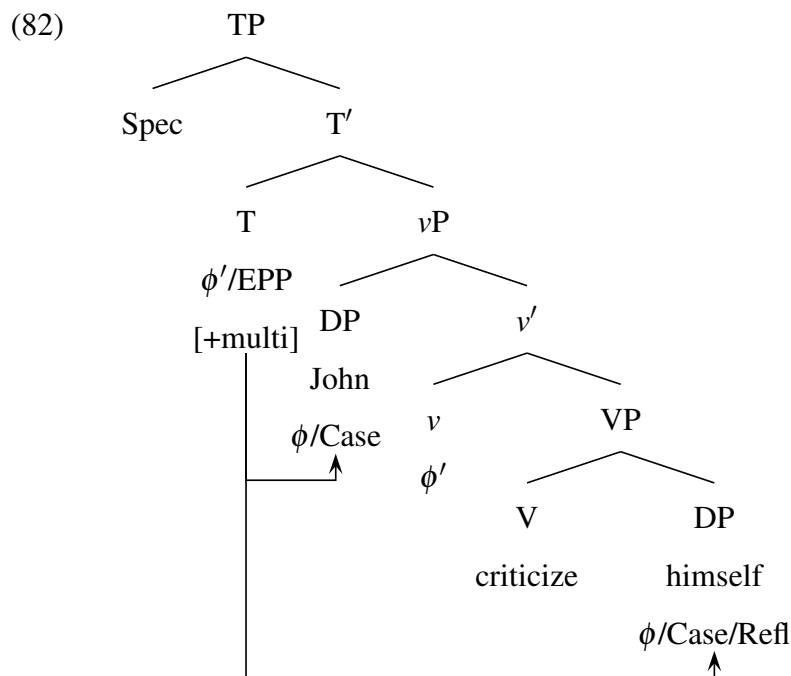
### 1.3.3 Feature theories of binding

#### 1.3.3.1 Hasegawa 2005

The theory described in Hasegawa 2005 tries to derive binding relations from Agree. He claims that the relation between a reflexive and its antecedent can be obtained from the process of Multiple Agreement. Consider for instance his derivation of the sentence (81):

(81) John<sub>i</sub> criticizes himself<sub>i</sub>.

He claims that in a case when the object is reflexives, the head T must be specified for a property of “multiple agreement”, which he formally considers as a [+multi] feature on T-head. T first agrees with the subject *John* in Spec,*v*P and further goes on to probe the object *himself*. It is unclear what triggers this Agree operation, apart from the [+multi] feature. According to Hasegawa, this multiple agreement “induces the binding relation between these two goals.” The corresponding structure is shown below in (82).



Hasegawa 2005 claims that phasehood and the Phase Impenetrability Condition (PIC) are responsible for locality of binding. Since T cannot probe the reflexive if it is too far, i.e. in a phase already sent to interpretation, long-distance binding is impossible.

This approach is very attractive in that it reduces binding to already existing operations, like (Multiple) Agree. It resolves the problem of the binding domain by reducing it to a phase: and this reduction naturally follows from the properties of the Agree relation.

There are however a number of problems with the empirical coverage of this analysis. The first set of problems stem from the mysterious nature of T probing an object and the [+multi] specification. It is unclear which feature triggers an Agree relation between T and an anaphor, and which feature gets checked as a result of this operation.

It is also unclear why exactly T would probe an anaphor in object position, and not some other element. For instance, one would expect that in the case of ditransitive verbs, T could probe any of two objects of the verb, and if the reflexive is located lower in the structure, a binding relation would never be achieved. Further, it is also unclear if T can probe into complex structures: this again would depend on the real reason why T needs to enter a Multiple Agree relation. For instance, it is unclear if it would be able to probe reflexive possessives in sentences such as *John loves self's wife* in languages where such sentences are available (e.g. Russian).

Further, this analysis predicts that anaphors must be subject oriented, since T would first always probe the subject in Spec,vP. While true for SE-type anaphors, it is not the case for English, see Barss and Lasnik 1986, where they show that for ditransitive verbs the surface binding relationship hold, i.e. an anaphor located within an argument PP of the verb *introduce* can for instance be bound by the direct object:

(83) John introduced Mary<sub>i</sub> to herself<sub>i</sub>.

There is no obvious way of deriving this binding relation from the process proposed in Hasegawa 2005. This analysis does not provide an immediate way to differentiate between properties of SE- and SELF-type anaphors.

Last, since Hasegawa crucially relies on the PIC to derive his locality condition, it is unclear how long-distance binding relations and syntactic conditions on it could be derived.

The analysis of binding I propose below in Chapter 3 takes some intuitions from Hasegawa 2005 and proposes a new analysis where these problems are resolved.

### 1.3.3.2 Reuland 2005, 2011

The most recent approach to binding by features is pursued in a series of works by Reuland (2001, 2005, 2011).

His analysis is concerned with the question of how binding relations between SE-anaphors and their antecedents can be established via Agree-based framework. Reuland starts by observing that SE-type anaphors are unspecified for gender and number features. He utilizes Pesetsky and Torrego's 2007 framework in order to explain how coreference is obtained in case of a sentence such as (84):

- (84) Iedere professional voelde [zich aan de kant geschoven]. (Dutch)  
Every professional felt [himself to the side pushed]  
'Every professional felt himself pushed to the side' (Reuland 2005)

Assumptions from Pesetsky and Torrego 2007 used by Reuland are as follows (taken verbatim from Reuland 2005, pg. 510)<sup>1</sup>:

- (85) a. Both interpretable and uninterpretable features come as valued and unvalued;  
b. Agree involves valuation and feature sharing → a feature chain with one valued instance is valued and every feature must end up with at least one interpretable instance;  
c. Structural Nominative Case on the external argument DP (EA) is unvalued T.

Consider the structure in (86):

- (86) [Tns [EA [V\* [V SE ... ]]]]

According to Pesetsky and Torrego 2007, there is a T-feature on T, V\*, and Subject. This dependency is established by the following process:

---

<sup>1</sup>in section 2.5.1 I discuss Pesetsky and Torrego 2007 system in greater details. This short list of assumptions in (85) is sufficient for the purposes of discussing Reuland's account.

- (87) a. T first probes down in search of a match for its T-feature. The subject has an unvalued T-feature (=Nominative case). Agree takes place, and the T-features on Subject and T get shared.
- b. Since the T-feature on T is still unvalued, T probes further down and finds V\* with a valued instance of T-feature. Agree takes place, and all instances of T-feature (on T, on the Subject, on V\*) become valued and now share the same value.

In order to extend this analysis to the valuation of features on SE, Reuland proposes that  $\phi$ -features are now involved in this Agree process, and T-dependency as described in (87) extends to  $\phi$ -feature dependency. The location of  $\phi$ -features important for this process is shown in the structure in (88):

(88) [Tns<sub>u</sub> $\phi$  [SE<sub>u</sub> $\phi$  [EA<sub>val</sub> $\phi$  [V\*<sub>u</sub> $\phi$  [V (SE<sub>u</sub> $\phi$ ) ... ]]]]]

The derivation goes on as described in (89):

- (89) a. V\*'s EPP feature triggers movement of SE to the edge of V\*P;
- b. The dependency between T-features is established similar to the process in (87).
- c. As for  $\phi$ -feature dependency, T first probes SE and their  $\phi$ -features get shared, even though they are still unvalued.
- d. T continues to probe and finds EA as its goal. T-features and  $\phi$ -features on T and EA get shared. That results in valuation of  $\phi$ -features on Subject, T, and SE.
- e. T continues to probe to get a value for its T-feature, and finds V\* with valued T-features as a goal. T-features on T, V\*, and Subject get valued, and  $\phi$ -features on the verb get valued as well.

I find several steps unclear in this derivation. The first step requires SE to move to the edge of V\*P because of the EPP feature of V\*. It is unclear what the nature of this EPP-feature is and why EA does not satisfy it. Even if one assumes that EPP features must be satisfied by Move, and not Merge, it is still unclear why SE is the only thing which can satisfy this feature, for instance in

cases with ditransitive verbs. More needs to be said on when this EPP is operational. It is unclear if Reuland assumes that this EPP feature is present only if a SE-anaphor is an object of the verb. If this is the case, it is unclear how one can enforce it in the current minimalist framework. If this feature is independent on the nature of the object and is universally present on V\*, then one would predict that object shift to the edge of V\*P is always necessary.

Further, this must be a covert movement, since at the Spell-Out SE-anaphor does not need to be present at the V\*P edge, see for instance examples like (90) from Dutch and (91) from Russian:

- (90) a. Max wast zich. (Dutch)  
 Max washes SE  
 ‘Max washes himself’
- b. Max schaamt zich.  
 Max shames SE  
 ‘Max is ashamed.’
- (91) a. Ivan nenavidit sebja (Russian)  
 Ivan hates se  
 ‘Ivan hates himself’
- b. Ivan razgovarivaet s soboj  
 Ivan speaks with se  
 ‘Ivan speaks with himself’

In fact, the assumption of object shift is problematic for Russian: while Russian is a free-word order language which allows alternative word orders, the existence of a process in Russian similar to object shift in Germanic is controversial. Even though some pronominals tend to move to a preverbal position, this is not true of the reflexive *sebja*.

One more problem worth mentioning is the problem of non-argument anaphors. Do they also move to the edge of V\*P? And if they do, at which level, since the overt instances of such movement might violate some other constraints on movement (left-branch extraction in case of movement of possessive anaphors, extraction out of DPs in languages which do not allow it, etc.)

It is also impossible to derive long-distance binding relations: one has to assume (like Reuland

2011 does) that long-distance binding is necessarily logophoric. I have already discussed problems with this approach above.

Lastly, it is unclear how English-type SELF-anaphors get their value in an Agree-based framework. They come from the lexicon with valued  $\phi$ -features, and therefore cannot participate in the process discussed above.

### 1.3.3.3 Hicks 2009

One more approach to binding theory within the minimalist framework is presented in Hicks 2009. He convincingly argues against the approach by Reuland 2005 that using  $\phi$ -features is sufficient to determine binding relations. Hicks claims that while  $\phi$ -features of the anaphor and its antecedent must necessarily match, they by themselves do not encode referential properties. For instance, in the sentence *John loves himself* the  $\phi$ -features of *John* and *himself* match, *himself* could in theory refer to some other male individual, which is not the case. He emphasizes that Reuland 2005 does not have an account of how referential dependencies are established after  $\phi$ -features have been valued and matched between an anaphor and its antecedent.

In order to resolve this problem Hicks 2009 proposes the feature [VAR], i.e. a variable feature, which is unvalued on reflexives but valued on referential expressions. According to him, the sentence *John loves himself* would have the following feature specification<sup>2</sup>:

(92) John<sub>[VAR: x]</sub> loves himself<sub>[VAR: \_]</sub>

Anaphoric binding can therefore be reduced to feature-valuation via an Agree relation between an anaphor and its antecedent.

Reducing binding to feature valuation allows Hicks to limit the binding domain to the domain

---

<sup>2</sup>He also introduces one more feature, an operator feature [OP] to take care of bound variable cases. According to him, OP feature can have values  $\forall$  and  $\exists$  in order to take care of sentences such as

- (i) a. Every boy loves himself.
- b. Every<sub>[Op:  $\forall$ ]</sub> <sub>[VAR: x]</sub> boy loves himself<sub>[VAR: \_]</sub>

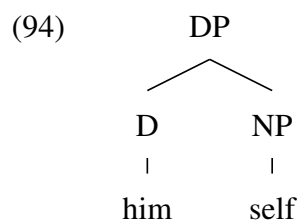
of Agree, i.e. to a phase in Chomsky 2001 “Derivation by Phase” framework. The problem for Hicks 2009 lies in the direction of probing. In the Phase Theory framework it is assumed that unvalued features probe their c-command domain until they find a goal. This relation must be reversed in case one assumes the [VAR]-feature based theory of binding. A reflexive with an unvalued [VAR]-feature is located in the c-command domain of its antecedent with a valued [VAR]-feature. Probing of an antecedent by an anaphor would contradict the definition of probing.

There is also another reason why this probing would be impossible: it is traditionally assumed that only heads, and not phrases, can probe. Therefore DP cannot probe its c-command domain, even if the previous problem with the probing could be resolved.

In order to overcome these issues, Hicks 2009 extends the suggestion by Rezac 2004, “with the search space for a probe extending ‘upwards’ if no goal can be found in its c-command domain.” The most important issue which Hicks 2009 explains is why the following sentence with the subject anaphor is ungrammatical:

(93) Himself<sub>i</sub> loves John<sub>i</sub>.

He proposes that anaphors like *himself* have complex structure, with *him* occupying the D-position within a DP and *self* being an NP-complement of D:



According to this structure, only *self* is in the c-command domain of D-head *him*, and therefore in (93) this D-head will not be able to probe outside of DP and find *John* as its goal. In fact, postulating this structure would *always* prevent an anaphor from probing down: the only possible direction for probing is upwards, as needed for Hicks’s analysis.

Using phase-limited Agree relation and upwards-probing, Hicks 2009 derives binding in simple



sentences as well as in raising, control, and ECM clauses.

He is also able to account for cases when binding interacts with A'-movement, such as (95):

(95) John<sub>i</sub> wondered [<sub>CP</sub> [<sub>DP</sub> which pictures of himself<sub>i/j/k</sub>] Bill<sub>j</sub> claimed [<sub>CP</sub> ~~which pictures of himself~~ Paul<sub>k</sub> had bought ~~which pictures of himself~~ ]]

According to Hicks 2009 probing “upwards” can happen at any moment of the derivation, including the point when the *wh*-phrase is in any of the intermediate CP specifiers. While this approach is consistent with the derivational theory of binding, it is unclear what allows the reflexive to postpone probing of the [VAR] features, instead of valuing them right away as potential antecedent becomes available. For example, why doesn't probing happen immediately after *Paul* has been merged into structure for example (95)?

Even though Hicks proposes solutions to several issues related to this updated definition of probing, usage of this non-standard definition of probing in my opinion is an important problem for adopting Hicks's analysis.

Further problem with this analysis is that it cannot explain long-distance binding dependencies, without *ad hoc* postulating that some domains are not phases or referring to discourse strategies of assigning reference to an anaphor (i.e. claiming that they are in fact logophors). In his discussion of Icelandic long-distance binding facts, Hicks 2009 agrees with the suggestion from Thráinsson 1976 that “the meaning expressed by the NCBR [non-clause bounded reflexive] is related to the meaning reflected in the subjunctives that correlate with it.” It is unclear how this approach can be extended to cases of other languages, like Russian, where binding into subjunctive embedded clauses is prohibited, and it is also unclear how subject orientation of these reflexives can be derived.

## 1.4 Conclusion

Above I outlined the history of binding theory and a number of revision to classical binding theory suggested in the recent literature. I showed that while being on the right track, none of these

analyses covers broad enough territory. Crucially, the problem of binding domains still remains. Reducing it to a phase goes a long way in the right direction from the point of view of minimalism, the notion of *phase* is unfortunately not sufficient to explain cases of long-distance binding without resorting to a theory of logophors. In Chapter 3, I propose my analysis of binding, which combines features of the three minimalist analyses presented above, while in the next Chapter 2 I consider certain problems of phase theory and how they can be resolved in order to prepare for the discussion of binding later in this dissertation.

# Chapter 2

## Deriving Phases

In this chapter I will review Phase Theory in its current formulation based on Chomsky (2000, 2001, 2004, 2008). I argue that phase theory cannot account for a range of facts, for instance long-distance dependencies. I start off by reviewing the major definitions and assumptions of phase theory, and proceed to show the theoretical and empirical gaps. At the conclusion of this chapter I propose how phases can be *derived* based on feature-based syntax. This proposal allows us to take into consideration the featural content of different syntactic categories and derive their phasehood based on it. I claim that the domain becomes closed, i.e. *phasal*, just in case certain features within the domain are either valued or shared with an element outside of this domain.

### 2.1 Overview of phase theory

The arguments that the derivation of the sentence proceeds cyclically, and the need to minimize the computational load gave rise to *Phase Theory*, which was introduced in a series of recent works by Chomsky (2000, 2001, 2004, 2008). Instead of selecting a Lexical Array (LA) for the entire sentence at once, Chomsky (2001) proposes that subarrays  $LA_i$  are cyclically “placed in active memory,” and the computation must exhaust  $LA_i$  before selecting the next subarray. As soon as the subarray  $LA_i$  is exhausted, the syntactic object is formed, and a next subarray  $LA_{i+1}$  is picked

up from the LA. The syntactic objects which are created after exhausting each subarray  $LA_i$  are referred to as *phases*.

While the general idea remained unchanged in the series of works on Phase theory, the motivations for postulating particular syntactic categories as phases has evolved. In Chomsky (2001), it is suggested that phases need to be “propositional,” and possible candidates are either  $v^*P$ , a verbal phrase with a full argument structure<sup>1</sup>, or CP, a phrase with full set of force indicators.

In Chomsky 2004, the Interface Condition (IC) starts playing an important role in defining phases. The following quote from Chomsky 2004, pg. 124 mentions both completeness of the phases at the semantic level as well as the phonological level:

Ideally, phases should have a natural characterization in terms of IC: they should be semantically and phonologically coherent and independent. At SEM,  $vP$  and CP (but not TP) are propositional constructions:  $vP$  has full argument structure, and CP is the minimal construction that includes Tense and event structure and (at the matrix, at least) force. At PHON, these categories are relatively isolable (in clefts, VP-movement, etc.). These properties do not, however, yield exactly the right distinctions:  $vP$  with  $v$  nontransitive is relatively isolated and is a domain for QR, though these cannot be phases for Spell-Out. Call these *weak phases*. Then the strong phases are those that have an an EPP-position as an escape hatch for movement and are, therefore, the smallest constructions that qualify for Spell-Out.

Having defined phases, Chomsky claims that in order to obey cyclicity of the derivation, one needs a condition governing the timing of the Spell-Out of the phasal domains. He postulates the Phase Impenetrability Condition – two slightly different versions of which are cited below in (1) and (2).

---

<sup>1</sup>Notice, that full argument structure for Chomsky (2001) necessarily implies presence of the external argument, as he specifically claims that unaccusative and passive  $vPs$  are not phases.

(1) *Phase Impenetrability Condition (PIC)* (Chomsky 2000, pg. 108)

In phase  $\alpha$  with head H, the domain of H is not accessible to operations outside  $\alpha$ ; only H and its edge are accessible to such operations

(2) *Phase Impenetrability Condition (PIC)* (Chomsky 2001, pg. 14)

The domain of H is not accessible to operations at ZP; only H and its edge are accessible to such operations. (Here, the following configuration is assumed: [<sub>ZP</sub> Z ... [<sub>HP</sub>  $\alpha$  [ H YP ]]])

Notice that while the difference in wording seems to be minimal, these definitions differ on *when* the domain of the phase becomes inaccessible to operations from outside of the phase.

According to definition of the PIC in (1), the domain of the phase becomes closed immediately upon the merge of the next head. In particular, if  $\nu$ P is a phase, that version of PIC would mean that anything within the  $\nu$ -complement becomes inaccessible as soon as T is merged into the structure. For the embedded CP phase, the merge of the next head (V) from the matrix clause would trigger all material below the C-level in the embedded clause to become inaccessible from the matrix clause.

Different predictions are made by the version of PIC in (2). According to that version, a phase becomes closed only upon the merge of the next *phase* head. Therefore, the complement of  $\nu$  in a  $\nu$ P phase becomes inaccessible to further operations as soon as the C head is merged, and the complement of the embedded C becomes inaccessible as soon as the matrix  $\nu$  is merged into structure.

Notice, that if one assumes the split left periphery hypothesis (along the lines of Rizzi 1997), an important question arises as to which exact part of the extended CP is phasal. Chomsky 2008, pg. 143 mentions that “C is shorthand for the region that Rizzi (1997) calls the “left periphery,” possibly involving feature spread from fewer functional heads (maybe only one).” At this moment I will not elaborate on this issue further at the moment; it will be taken up below.

The Spell-Out process is described in Chomsky 2004, pg. 107, and is summarized in (3):

- (3) a. The phonological component  $\Phi$  maps D-NS to PHON.
- b. The semantic component  $\Sigma$  maps D-NS to SEM.
- c. TRANSFER hands D-NS over to  $\Phi$  and to  $\Sigma$

Crucially, Chomsky assumes optimal design, according to which all of the transfers described in (3) occur “cyclically in parallel,” and the units submitted to different components (PHON and SEM) are the same: phases.

In the next sections I will discuss various problems of Phase Theory. They will be grouped into two categories: theoretical problems and empirical problems. Theoretical problems are theory internal problems which I will take on in section 2.2. In section 2.3 I will provide data on a number of phenomena which do not allow a satisfactory account within current phase theory as outlined in Chomsky 2000, 2001, 2004, 2008.

## 2.2 Theoretical gaps

There are a number of problems with the phase theory framework. In this section I will consider theoretical gaps: those include the gaps where the internal theoretical proposals are either inconsistent or have not been sufficiently explored. *Per se* these problems cannot be considered as arguments against phase theory; they do need extra investigation, and most possibly warrant revision of the certain modules of the phase theory.

The following list contains the questions addressed below in this chapter.

- What exactly is a phase? Are all CPs and  $\nu$ Ps phasal or there are factors which determine under which conditions they are?
- What exactly is Spell-Out? How does it work and what can be transferred to the interfaces?
- What exactly are *strong* and *weak* phases? Do weak phases undergo Spell Out? Are they inaccessible to syntactic operations? What determines cross-linguistically whether a phase is weak or strong?

- What about other syntactic categories, such as DPs and PPs? Are they always phasal? Do the languages without articles differ from the languages with articles in terms of phasal properties of DPs? What can be said about PPs?

Some of the questions will not be answered until much later in this thesis; some will not be answered here at all.

### 2.2.1 What exactly is a phase?

As I mentioned earlier, CPs and (transitive) *v*Ps are claimed to be (strong) phases by Chomsky. Now two important questions arise 1) Are *all* CPs phases? 2) Is it true that only *transitive v*Ps are phases? There are several ways one can try to answer that question, which are summarized in (4):

(4) *Which domains are phases?*

	Option 1	Option 2	Option 3	Option 4
<i>v</i> P	all	non-defective (transitive)	all	non-defective
CP	all	non-defective (tensed?)	non-defective	all

Four different options are available here. One can claim that for each major category which can be a phase (*v*P or CP), either it is always a (strong) phase, or it is only a (strong) phase when it is “non-defective” in a certain sense. Notice that here I use a term “non-defective” and specify one of the possibilities what this term might mean when applied to a particular category.

Chomsky 2004 claims that for *v*P this “non-deficiency” means that it has a full argument structure, including an external argument. Therefore, unaccusative and passive *v*Ps are not phases and therefore do not undergo Spell-Out upon the merge of the next available C-head. This assumption is however questionable. For instance, Legate 2003 provides a number of arguments that any

vP is a phase, including passive and unaccusative. Her arguments are based on examination of reconstruction effects, antecedent-contained deletion, and parasitic gaps.<sup>2</sup>

While the works mentioned above and references therein talk about different types of vP with respect to being phasal, not much has been done on determining whether CP is phasal only if certain conditions hold or always. What could “deficiency” of CP mean? It is possible that CP is always phasal, as in Options 1 and 3 from (4) above. Another possibility is to claim that phasehood of CPs depends on the type of T-head in the complement of C. For instance, one might argue that infinitival CPs are not phasal, while indicative CPs are always phasal. Looking at more fine-grained distinctions, one could make a cut at any point on the continuum in (5), arguing that everything to the left of the cut point is phasal, while everything on the right is not.

(5) Indicative → Subjunctive → Control → ECM → Raising

The technical realization of this solution is non-trivial in the early Phase Theory framework, however the framework from Chomsky 2008 which allows for the mechanism of feature transfer from C to T, would be able to handle it.

It is also possible to claim (see Adger 2004) that some types of infinitival clauses (especially at the right edge of the continuum in (5)) lack a CP layer altogether, and therefore some of the infinitival embedded clauses are non-phasal for that very reason.

The distinction given in (5) is a very coarse one. Each of the major elements on that scale may be further subdivided into finer subtypes of clauses. For instance, indicative embedded clauses can be divided into factives and non-factives. De Cuba 2007 (and references therein) discusses various differences between factive and non-factive complements, and some of these difference can be taken to be reflecting difference in phasehood of the various complements.

Further, subjunctive embedded complements can also be of various types. For instance, Serbo-Croatian (Progovac 1993a,c) as well as Bulgarian are usually claimed to have two types of sub-

---

<sup>2</sup>In his reply, Den Dikken 2006 goes against Legate’s arguments, concluding that even though her attempts at establishing phasehood of vPs fail, there is still some good evidence that vPs are phasal.



junctives (Krapova 2001) with different syntactic properties. Type I subjunctives are selected by the epistemic verbs such as ‘hope’, ‘believe’ and volitional/desiderative verbs ‘want’, ‘wish’, while Type II subjunctives are selected by aspectual (‘begin’, ‘continue’) and subject control verbs (‘know’, ‘try’, ‘forget’).

Krapova 2001 argues that Type I subjunctives take *pro* as their subject, while in Type II subjunctives the subject is PRO: the referent of the empty category in Type I subjunctives can be either the matrix subject or any other entity, while the referent of the empty category in Type II subjunctives is necessarily a matrix subject. Further, Type I subjunctives allow for case-marked overt subjects, while Type II subjunctives do not. The relevant facts are exemplified in (6). Similar facts hold for Greek subjunctives.

(6) *Bulgarian Type I and Type II subjunctives (Krapova 2001)*

- a. Ivan<sub>i</sub> iska [brat mu / toj<sub>i/j</sub> / e<sub>i/j</sub> da sledva.]  
 Ivan want.3SG brother his / he / e DA study.3SG  
 ‘Ivan wants his brother/him to go to the college’ / ‘Ivan wants to go to college’
- b. Ivan<sub>i</sub> usnja [\*brat mu / \*toj<sub>i/j</sub> / PRO da spečeli mnogo pari.]  
 Ivan succeeded.3SG brother his / he / PRO DA make.3SG a lot of money  
 ‘Ivan succeeded in making a lot of money’

Further, Spanish and Catalan subjunctives (Piccolo 1985) differ in their syntactic behavior depending on whether they are selected by a volitional verb, such as ‘to want’ or ‘to hope’, or by some other verb requiring subjunctive, such as ‘to doubt’ or a negative form of the verb ‘to believe’. With the former group of verbs, the embedded pronominal subject cannot be coreferential with the subject of the matrix clause (*Subject Obviation Effects*), (7), while with the latter such a coreference is possible, (8).

(7) *Obviative subjunctives in Catalan (Piccolo 1985)*

- a. \*En Jordi<sub>i</sub> espera que [e<sub>i</sub> vingui]  
 DET George hopes that (he) comes<sub>SUBJ</sub>  
 ‘George hopes to come’

- b. \*Tu<sub>i</sub> sents que [ e<sub>i</sub> telefonis al Joan]  
 you regret that (you) call<sub>SUBJ</sub> John  
 ‘You regret that you call John’

(8) *Non-obviative subjunctives in Catalan (Picallo 1985)*

- a. No crec que telefoni  
 not believe.1SG that call<sub>SUBJ</sub>  
 ‘(I) do not believe that (I) call’
- b. Dubtava que telefonés  
 doubt.3SG that call<sub>SUBJ</sub>  
 ‘(He/she)<sub>i</sub> doubted that (he/she)<sub>i</sub> called’

From what we have seen, it is clear that even long-distance phenomena in clauses of the same type, like in subjunctives, cannot be accounted for solely by stipulating the status of all subjunctive complements as being phasal (or non-phasal). Finer distinctions are certainly needed as several types of subjunctives exist. The same can be said for other types of clauses: indicatives can be divided into for example factives/non-factives, infinitival clauses have several types, such as control, ECM, raising, and all of them might be different with respect to phasal properties.

### 2.2.2 Spell-out gaps

(Strong) phases are sent off to interpretation as soon as the head of the next phase is merged into the structure. If all the features within the phase are valued by the time the head of the next phase is merged, sending the phase to interpretation will not pose any problems. The computational system will be able to interpret the phase. However if the embedded clause is dependent on the matrix clause (for example, the tense in the subordinate clause is dependent on the tense in the matrix clause, as in the case of infinitives and subjunctives) the situation is not as clear.

Consider the case the case of embedded subjunctive clauses. Under assumptions outlined in Section 4, the T-feature on the embedded *v* does not get its value until the matrix T is merged into the structure. If one follows standard assumptions that *v*P and CP are phases, there are three phase

boundaries between the embedded verb and the matrix T: embedded  $vP$ , embedded  $CP$ <sup>3</sup>, and matrix  $vP$ . By the time the matrix T is merged into the structure, the embedded  $vP$  phase must have been already sent to interpretative component. Notice, that by the time  $vP$  was sent to interpretation, the T-feature on  $v$  was not valued, and the event described in the embedded  $vP$  does not yet have a temporal anchoring. Sending unvalued features to the interpretative component must lead to a crash at LF; any sentence with subjunctive embedded clause will be found ungrammatical, contrary to the facts.

Similar considerations apply to infinitival embedded clauses. A non-finite verb does not have specification for Tense, and its T-feature gets valued only upon the merge of the matrix T (following the architecture proposed in Pesetsky and Torrego 2001). At least two phase boundaries separate the embedded verb from the matrix T<sup>4</sup>. Following the traditional approach to phases, the embedded  $vP$  phase must be sent off to interpretation necessarily before merge of matrix T, therefore with the unvalued T-feature on the verb. Similarly to the subjunctive case, that would lead to a crash at the LF interface.

### 2.2.3 Strong and weak phases: what is the difference?

Chomsky 2004, pg. 124 talks about *weak* phases. He defines a weak  $vP$  to be a  $vP$  with a non-transitive verb. According to him, while “relatively isolated” and a domain for Quantifier Raising (QR), they still cannot undergo Spell-Out. Further, according to Chomsky 2004, *strong* phases are “the smallest constructions which qualify for Spell-Out”, and need to have an escape hatch for movement. Several questions immediately arise from this discussion.

First, one needs a more specific explanation for how the existence of the escape hatch is connected with the ability to be a domain for Spell-Out.

Second, if weak phases do not have an edge which needs to be used as an escape hatch and are

---

<sup>3</sup>I will address the question of whether the subjunctive CP is a phase in the later parts of this thesis, but as the current discussion is concerned, it is immaterial.

<sup>4</sup>Again, at this point I remain agnostic on whether infinitival embedded clauses have a CP layer and whether they constitute a phase. I will get back to this question in section 2.5.4

not Spelled Out, what exactly is phasal about them? In which sense are they “relatively isolated”?

This mentioning of weak phases is the last one in the series of works by Chomsky. In the more recent papers (Chomsky 2004, 2008), he does not mention weak phases at all. In this regard, the status of the weak/strong distinction must certainly be clarified. It is possible that such a distinction does not exist — as I mentioned above, Legate 2003 argues against passives and unaccusatives being *weak* phases.

#### **2.2.4 Other syntactic categories: NPs, DPs and PPs**

Above I showed some theoretical gaps with respect to whether  $\nu$ P and CP can be counted as phases. The question of whether other syntactic categories have phasal properties can also be considered a gap within the current formulation of phase theory. For instance, the question of whether PPs are phases is not explored in a series of works by Chomsky. At the same time, the functional structure of the PPs has been considered in Abels 2003; Bošković 2004; Kitada 2007; Matsubara 2000; Svenonius 2010. Chomsky (2000, 2001) suggests that DPs are also phases, and a number of recent papers agree with this proposal. For details see Bošković 2005; Den Dikken 2007; Despić 2011; Heck et al. 2008; Svenonius 2004. Bošković (2010) claims that even NPs in article-less languages can also qualify as phases. It is clear that diagnostics for phases do not give uniform results when applied to syntactic categories in various languages.<sup>5</sup>

### **2.3 Empirical problems**

In section 2.2 I outlined some theory internal gaps within the phase theory framework. I did not deal with any specific data, and simply demonstrated internal inconsistencies of the theory and also

---

<sup>5</sup>I will not address issues of whether DP, NP, and PP are phases in this thesis: I will only work with clausal domains such as VP,  $\nu$ P, TP, and CP, since featural content of these domains is less controversial than featural content of NP, DP and PP. The feature-based theory I propose in section 2.5.4 in order to derive phases makes use of featural content of various syntactic elements and categories, and would be able to make predictions about phasal status of other syntactic categories, however I leave this issue for further research.

showed that there are a number of questions which must be answered in order to fully understand the consequences of the theory.

In this section I will outline the empirical problems. I will present data which cannot be easily accounted for under phase theory framework. Most of these phenomena are non-local, and therefore are not (obviously) phase bound. These long-distance processes I will discuss are:

- (9) a. Long-distance binding
- b. Negative Polarity Items
- c. Long-distance agreement
- d. Sequence of tense

### 2.3.1 Long-distance binding

As I showed in Chapter 1, reflexive pronouns and reciprocals must find their antecedent in a certain domain, traditionally a tensed clause (or a minimal domain with a SUBJECT). This generalization is known as Principle A of binding theory:

- (10) **Principle A (Chomsky 1981/1993):** An anaphor must be bound within its governing category.

There have been many attempts in a previous literature to give a precise definition of the governing category. The relevant discussion is given in section 1.1.3.1. As we saw above, phase theory disallows long-distance dependencies: the complement of the phasal head is inaccessible to operations as soon as the head of the next phase is merged. Thus the standard cases when, say, the object anaphor is bound by the subject fall nicely under phase-based binding (i.e. a theory of binding when the domain is a phase, i.e.  $vP$  or  $CP$ ): the complement of  $v$  is still accessible at the point the subject is merged; therefore if binding is a syntactic process, there will be no phase theory violation in establishing a relation between the subject and the object.

(11) John<sub>i</sub> praises himself<sub>i</sub>.

Binding is however not always clause bound. Examples below (some are from Huang 2000) provide sentences from a number of languages where the anaphor can get bound outside of a certain domain.

In example (12) one can see sentences from Russian and Norwegian showing that anaphoric binding is possible into DPs. Notice, that Russian is a language without articles, and Norwegian has articles. The English examples in (12) demonstrate that the internal structure of DP, such as presence of a possessor, can change possibilities of binding: in (12-c) the sentential subject is a possible antecedent, while in (12-d) only the possessor can serve as a binder for an anaphor. Examples in (13) show binding into PPs in English.

(12) *Binding into DP/NP*

- a. Ivan<sub>1</sub> pročital [rasskazy o sebe<sub>1</sub>]. (Russian)  
Ivan read<sub>PST</sub> stories about self  
'Ivan read stories about himself'
- b. Jon<sub>1</sub> likte [din artikkel om seg<sub>1</sub>]. (Norwegian)  
Jon liked your article about self  
'Jon liked your article about self.' (Hellan 1991, pg. 30)
- c. John<sub>1</sub> read [books about himself<sub>1</sub>]
- d. John<sub>1</sub> read [Peter<sub>2</sub>'s books about himself<sub>\*1/2</sub>].

(13) *Anaphor binding into PP (Reinhart and Reuland 1993)*

- a. Max speaks with himself.
- b. Max saw a gun near himself.

The sentences in (14) show the differences between English and Danish with respect to binding into small clauses. As one can see, a reflexive pronoun located in the small clause allows a matrix subject as its antecedent, while in English such coreference is prohibited.

(14) *Anaphor binding into a Small Clause*

- a. Larsen<sub>1</sub> betragter Jorgen some farlig for sig<sub>1</sub>. (Danish)  
Larsen considers Jorgen as dangerous for self  
'Larsen considers Jorgen dangerous for self.' (Pica 1986)
- b. \*Lucie<sub>1</sub> heard Max praise herself<sub>1</sub>.

A similar case is found in (15). Here one can see that Russian is a language which allows long-distance binding into infinitival complement clause. At the same time, not all languages do: the English example shows that coreference between the object of the infinitival complement and the matrix subject is impossible.

(15) *Anaphor binding into an Infinitival Clause*

- a. Professor<sub>1</sub> poprosil assistenta<sub>2</sub> PRO<sub>2</sub> čitat' svoj<sub>1/2</sub> doklad. (Russian)  
Professor asked assistant read self's report  
'The professor asked the assistant to read self's report.' (Progovac 1993c)
- b. \*John<sub>1</sub> told Marie to draw himself<sub>1</sub>

The next possible type of clause which often allows long-distance binding into it is subjunctive, (16). Most Romance languages, such as French and Italian, generally disallow long-distance binding, unless the embedded clause is subjunctive. The same is true for Icelandic. On the other hand as we saw above, Russian allows long-distance binding into infinitival complements, but not into subjunctives.

(16) *Anaphor binding into a Subjunctive Clause*

- a. On<sub>1</sub> souhaite que les gens ne disent pas du mal de soi<sub>1</sub>. (French)  
one wishes that the people NEG speak NEG of ill of self  
'One always wishes that people do not slander oneself.' (Pica 1991)
- b. Jón<sub>1</sub> segir að María elski sig<sub>1</sub>. (Icelandic)  
J. said that M. loves<sub>SUBJ</sub> self  
'John said that Maria loves him' (Sigurðsson 1990)

- c. [Quel dittatore]<sub>1</sub> spera che i notiziari televisivi parlino a lungo delle  
 [That dictator] hopes that the news TV talk<sub>SUBJ</sub> for a long time about  
 proprie<sub>1</sub> gesta. (Italian)  
 self's deed.  
 'That dictator hopes that the TV news will talk about his deeds for a long time.'  
 (Giorgi 1983)
- d. Ivan<sub>1</sub> xočet čtoby Maša<sub>2</sub> pocelovala sebja\*<sub>1/2</sub> (Russian)  
 Ivan wants that<sub>SUBJ</sub> Maš kiss<sub>SUBJ</sub> self  
 'Ivan wants that Maša kisses self'

Finally, a number of languages allow for anaphoric binding into indicative clauses. Relevant examples from Malay and Korean are given in (17).

(17) *Anaphor binding into an Indicative Clause*

- a. Siti<sub>1</sub> mengingatkan Mohamed<sub>2</sub> yang saya tahu dirinya<sub>1/2/3</sub> seorang penjenayah.  
 Siti remind Mohamed that I know self<sub>3SG</sub> one criminal  
 (Malay)  
 'Siti reminded Mohamed that I know he/she is a criminal.' (Cole and Hermon 2005)
- b. Cheolsu<sub>1</sub>-nun Youngshik<sub>2</sub>-i caki<sub>1/2/3</sub>-lul coaha-nun-keot-ul Youngsu<sub>3</sub>-ka  
 Cheolsu-TOP Youngshik-NOM self-ACC like-ASP-COMP-ACC Youngsu-NOM  
 alkoitta-ko saengkakha-n-ta. (Korean)  
 know-COMP think-ASP-DEC  
 'Cheolsu thinks that Youngsu knows that Youngshik likes himself.' (Choi 1997)

One can think of the following hierarchy of the long-distance binding domains:

- (18) DP/PP → Small Clause → Infinitive → Subjunctive → Indicative

Languages which allow binding into larger structures would in general allow binding into a small structure. The table in (19) provides some examples:<sup>6</sup>

---

<sup>6</sup>Notice, that the fact that language allows for long-distance binding into one of the above listed domains does not mean that all instances of such binding are grammatical. For instance, long-distance binding in Italian and Icelandic subjunctives is claimed to be governed by certain discourse properties ("point-of-view", Giorgi and Pianesi 1997).



(19) *Various possibilities for long-distance binding*

	DP/PP	Small Clause	Infinitive	Subjunctive	Indicative
Old Icelandic	✓	✓	✓	✓	✓
Italian	✓	✓	✓	✓	✗
Russian	✓	✓	✓	✗	✗
(no known languages)	✓	✓	✗	✗	✗
English	✓	✗	✗	✗	✗

Unsurprisingly, long-distance binding has always been problematic for locality based theories of syntax. Long-distance reflexives were often assumed to be logophors (Hicks 2009; Huang and Liu 2001; Reinhart and Reuland 1993 a.o.). For instance, Huang and Liu claim about Chinese reflexive pronoun *ziji*:

A fundamental distinction must be drawn between the anaphoric and the logophoric uses of the reflexive, the former subject to syntactic conditions of anaphoricity, and the latter subject to certain pragmatic conditions of logophoricity.

Therefore, they refute a syntactic account of the distribution of *ziji*, and show that a pragmatic analysis fits the data best. While an approach like that is possible, most languages which allow long-distance reflexivization exhibit certain syntactic constraints on the distribution of the reflexive. Such problems with logophor theory are outlined above in section 1.2.3.

### 2.3.2 Negative Polarity Items

The next potentially long-distance phenomenon I consider is licensing of Negative Polarity Items (NPIs). The main difference between NPIs in English on the one side and Korean and Japanese on the other is that English NPIs can be licensed long-distance by a negation present in the superordinate clause, (20), while Japanese and Korean NPIs obey a clausemate condition on negation, (21) and (22).

(20) *NPIs in English*

- a. John didn't eat anything.
- b. Mary didn't say that John ate anything.

(21) *NPIs in Japanese (Nakao and Obata 2007)*

- a. Mary-wa nanimo tabe nakat-ta.  
Mary-TOP anything eat not-PAST  
'Mary did not eat anything.'
- b. \*Mary-wa [John-ga nonimo tabe-ta to] iwa-nakat-ta.  
Mary-TOP John-NOM anything ate C say-not-PAST  
'Mary did not say that John ate anything.'

(22) *NPIs in Korean (Nakao and Obata 2007)*

- a. Mary-ka amuto ani manna-essta.  
Mary-NOM anybody not met  
'Mary did not meet anybody.'
- b. \*Mary-ka [John-i amukesto saessta-ko] mit-ci ani hayessta.  
Mary-NOM John-NOM anything bought-C believe not did  
'Mary did not believe that John bought anything.'

Even in Japanese, however, there are cases when the clause-mate condition does not make the right prediction. The following example (23) involve “subjunctive” complement<sup>7</sup>. While not perfect, the example at the same time presents a contrast with the clearly ungrammatical example from (21).

(23) ??Bill-ga [John-ga Mary-to-sika au yoo(ni)] nozoma-nakat-ta (Japanese)  
Bill-NOM John-NOM Mary-with-NPI meet C hope-NEG-TNS  
'Bill<sub>(Neg)</sub> hoped [John meet [<sub>(NPI)</sub> only Mary]].' (Yamashita 2003)

Further, the example (24) from Yamashita 2003 is also problematic for the Clause-Mate Condition.

---

<sup>7</sup>See Uchibori 2000 for detailed discussion of “subjunctives” in Japanese.

- (24) ?Bill-ga Pam-ni [ Mary-to-sika<sub>i</sub> John-ga t<sub>i</sub> atta to] tutae-nakat-ta  
 Bill-NOM Pam-DAT Mary-with-NPI John-NOM met C tell-NEG-TNS  
 ‘Bill <sub>(Neg)</sub> told Pam [[<sub>(NPI)</sub> only Mary]<sub>i</sub> John met t<sub>i</sub>].’

As claimed in Progovac 1991, Serbo-Croatian has two types of NPIs: NI-NPIs and I-NPIs. In what follows I will specifically concentrate on NI-NPIs. NI-NPIs can be licensed in the following two situations. First, they can be licensed by clausemate negation. The relevant example is given in (25-a). Notice, that in general if the embedded clause is indicative, such as a complement of *kazati* ‘say’, the NI-NPIs cannot be licensed by matrix negation, see example (25-b). Second, NI-NPIs can be licensed in the complements of the volitional verbs such as *želeti* ‘wish’ and *hteti* ‘want’, when the subject of the embedded clause is phonologically empty, see contrast between (25-c) and (25-d) Unsurprisingly, both of these verbs select subjunctive complements. Relevant examples are given in (25).

(25) *NI-NPIs in Serbo Croatian*

- a. Milan ne voli nikoga.  
 Milan not loves no one  
 ‘Milan does not love anyone.’
- b. \*Mira ne kaže da vidi nikoga  
 Mira not says that sees no one  
 ‘Mira didn’t say that she saw anyone’
- c. Mira ne želi da e vidi nikoga.  
 Mira not wishes that sees no one  
 ‘Mira does not want to see anyone’
- d. \*Mira ne želi da Petar vidi nikoga.  
 Mira not wishes that Petar sees no one  
 ‘Mira doesn’t want Petar to see anyone’

The data above is taken partially from Progovac 1991. Some of my informants disagree with the grammaticality of (25-c), and judge this example as marginal at best. They do however exhibit the contrast in example (26) (Ivana Mitrović, p.c.):

- (26) a. \*Mira ne kaže nikoga da vidi  
 Mira not says no one that sees  
 ‘Mira didn’t say that she saw anyone’
- b. Mira ne želi nikoga da vidi  
 Mira not wants no one that sees  
 ‘Mira doesn’t want to see anyone’

Data similar to Serbo-Croatian can also be observed in French. As example (27) shows, NPI *personne* ‘no one’ can not be licensed by matrix negation if it is located in the subordinate indicative clause. On the contrary, if the embedded clause is subjunctive, the extraclausal licensing of NPI leads to only a mild deviance.

- (27) *NPIs in French (Mathieu 2001)*
- a. \*Il n’ a mentionné qu’ il avait vu personne.  
 he NEG has mentioned that he had seen no one  
 ‘He didn’t mention that he had seen anyone.’
- b. ?Je ne demande que tu vois personne.  
 I NEG ask that you see<sub>SUBJ</sub> no one  
 ‘I didn’t ask that you see anyone.’

The data presented above is problematic for phase theory. One can assume that NPIs have unvalued features, which need to be valued by a negation under an Agree relation — this Agree relation can be viewed as NPI licensing. If so, the Phase Impenetrability Condition would require the CP or vP containing an NPI to be sent to Spell Out before the licenser in the matrix clause is merged, and the derivation will crash. Further, as I showed above, the differences between the clause types which allow long-distance NPI licensing is languages such as Japanese, French, and Serbo-Croatian posit the problem which I mentioned above: what exactly counts as a phase? Is it true that in languages where subjunctives are more transparent with respect to NPI licensing than indicatives, they are either not phases at all or “weak” phases?

### 2.3.3 Long-distance agreement

In this section I will demonstrate data involving long-distance agreement. While this phenomenon is well known from Hindi-Urdu (see Bhatt 2005 a.o.), I will mostly concentrate on Daghestanian languages, where long-distance agreement is also extremely wide-spread. In Daghestanian, in simple clauses verbs usually agree in class<sup>8</sup> with the absolutive argument.<sup>9</sup>

If the verbs takes a clausal argument, there are two major possible patterns. In the first one, the matrix verb agrees with the entire complement clause, by taking an unmarked gender. The other, more unusual, *Long-Distance Agreement* (LDA) is a phenomenon when the matrix verb agrees with the absolutive argument located within the complement clause. Examples from Tsez are given in (28). In (28-a) one can see that the matrix verb *-iyxo* ‘know’ agrees with the entire clause, and is marked with the default gender IV. In example (28-b) the matrix verb agrees with the absolutive argument within the embedded clause.<sup>10</sup>

(28) *Long-distance agreement in Tsez, Polinsky and Potsdam 2001*

- a. enir [uža magalu bac’rufi] r-iyxo  
mother [boy bread.III<sub>ABS</sub> ate].IV IV-know  
‘The mother knows the boy ate the bread.’
- b. enir [uža magalu bac’rufi] b-iyxo  
mother [boy bread.III<sub>ABS</sub> ate] III-know  
‘The mother knows the boy ate the bread.’

Haspelmath 1999 demonstrates similar facts in Godoberi, another Caucasian language:

(29) *Long-distance agreement in Godoberi, Haspelmath 1999*

- a. ilu-ŋi quča b-al-i q’<sup>o</sup>araŋ-anta bu-k’-a  
mother-DAT book<sub>ABS</sub> N-read-INF want-CONV.PRS N-be-AOR  
‘Mother wanted to read a book’

---

<sup>8</sup>Here I will use the terminology used in Caucasian linguistics. Noun class can normally be thought of as gender, and Daghestanian languages can have up to 8 different classes. Noun distribution over classes is not always predictable.

<sup>9</sup>The majority of Caucasian languages have ergative-absolutive case systems.

<sup>10</sup>Notice, that this construction though widespread, only occurs with specific matrix predicates.

- b. *ilu-ti quči-be r-al-i q'°araŋ-anta ru-k'-a*  
 mother-DAT book-PL<sub>ABS</sub> PLN-read-INF want-CONV.PRS PLN-be-AOR  
 'Mother wanted to read books'

Long-distance constructions in both Godoberi and Tsez have been claimed to be monoclausal (Haspelmath 1999; Polinsky 2003), and if so, unusual agreement patterns in these languages are not in fact long-distance, and therefore will not present any major problems for a phase theory approach.

Another Daghestanian language which exhibits Long-Distance Agreement is Khwarshi (Khalilova 2009).

(30) *Long-distance agreement in Khwarshi, Khalilova 2009*

- uža-l l/b-iq'-še k<sup>ʃ</sup>aba zihe b-ot'uq'q'-u.*  
 boy<sub>OBL-LAT</sub> IV/III-know-PRS black cow(III) III-come-PST.PTCP  
 'The boy knows that the black cow has come.'

In Khwarshi, the verb *-iq'* 'know' is one verb which allows for long-distance agreement. In example (30) it agrees either with the entire embedded clause (in this case it shows class IV agreement), or with the absolutive argument of the embedded clause (III class agreement). The main differences between Khwarshi and other Daghestanian languages mentioned above is that constructions with the verb 'know' are biclausal. Khalilova 2009 considers Khwarshi data, some of which is based on the behavior of reflexives to demonstrate that. She shows that a complex reflexive pronoun, which is strictly local and subject oriented, cannot take its antecedent in the matrix clause, a result which is surprising if one were to assume a reanalysis of such sentences as monoclausal (a Clause Union analysis).

- (31) *išet'-il<sub>i</sub> l/b-iq'-še kand-ij iŋe.iŋeyoyul<sub>j</sub>/\*<sub>i</sub> kaɣat b-ešut't'-u.*  
 mother-LAT IV/III-know-PRS girl-ERG REFL.APUD.VERS letter<sub>III</sub> III-send-PST.PTCP  
 'The mother knows that the girl has sent herself a letter.'

If the construction in (31) were monoclausal, one would expect that *kandi* 'girl<sub>ERG</sub>' would not be

able to be the antecedent of the reflexive, as it would stop being a subject.<sup>11</sup>

Again, the data considered above is problematic for the Phase Impenetrability Condition. If one adopts Khalilova 2009 analysis that *know*-constructions are biclausal, and thus involve a CP layer in the embedded clause, one would need to take one of the following two positions. The first position would be to claim that complements of the verb ‘know’ are CPs which are either non-phasal or are “weak” phases, and therefore do not undergo Spell Out, and stay accessible to the Agree operation all the way until the merge of the matrix subject. Clearly, such a view is undesirable: what is so special about the verb ‘know’ so that its complements have such special properties? The second position would be to assume that Agree is not governed by phases and can look inside the phase. This position is of course also problematic for the notion of Spell Out proposed by Chomsky. In order to resolve this problem, it would be necessary to revise certain assumptions of phase theory. That’s the position which I will pursue in the following chapters of this dissertation.

### 2.3.4 Sequence of Tense

The term *Sequence of Tense (SOT)* refers to the behavior of embedded tenses usually when the matrix clause has past tense. According to Enç 1987, two distinct possibilities arise in English when the embedded past tense is embedded under the matrix past tense.

(32) *Sequence of Tense in English, Embedded tense = PAST*

- a. Mary found out that John failed the test.
- b. John heard that Mary was pregnant.

Sentence (32-a) exhibit what Enç calls a *shifted reading*. The event of finding out described in the matrix clause is located in the past moment  $t' < t$ , where  $t$  is a time of utterance. Further, the embedded clause is evaluated with respect to that time  $t'$ , and therefore the event described there

---

<sup>11</sup>Khalilova also provides other diagnostics based on the scope of negation and behavior of adverbs, according to which constructions with the verb ‘to know’ are in fact biclausal. I will not replicate them here.

(failing the test by John) is evaluated at some point  $t'' < t'$ . As Enç 1987 puts it, “the complement tense shifts the evaluation time away from the time at which the matrix sentence is evaluated.” Similarly, a shifted reading is available for (32-b).

There is however another possibility for (32-b). On that reading (*simultaneous reading*), John hears at a certain time in the past,  $t' < t$ , that Mary is pregnant at the time of hearing,  $t'$ . As Enç claims, such readings are only possible to stative complements. If one assumes the standard approach which posits the existence of a time-shifting operator and locates the embedded past tense in the past with respect to matrix tense, the existence of the simultaneous reading is unpredictable. Note that not all languages behave like English. For instance, in Russian the embedded tense is evaluated with respect to the matrix tense, (33); in particular, the simultaneous reading is expressed by using a morphological present tense in the embedded clause:

(33) *Absence of SOT in Russian*

- a. Ivan skazal čto Maša žarit kartošku.  
Ivan say<sub>PST</sub> that Mary fry<sub>PRES</sub> potato  
'Ivan said that Mary was frying potatoes'
- b. Ivan skazal čto Maša požarila kartošku.  
Ivan say<sub>PST</sub> that Mary fry<sub>PERF.PST</sub> potato  
'Ivan said that Mary had fried potatoes'

Further, in English the reference point of the embedded past tense can also not be an utterance time. For instance, the following example is semantically awkward in English:

(34) #John said a year ago that Mary was pregnant a month ago.

Incoherence of the example (34) shows that the embedded past tense is not just a regular past tense evaluated with respect to the moment of utterance: there needs to be a relation between the matrix tense and the embedded tense. Clearly, one can postulate that the embedded past tense can be interpreted either as a past or as a present with respect to the matrix tense, but that will be just a stipulation: the precise way both readings are derived would still be unclear, and the nature of this



dependency of the embedded tense on the matrix tense will have to be explained further.

Ogihara 1996 presents Japanese data showing that Japanese is also a non-SOT language, and the choice of tenses in Japanese is similar to Russian:

- (35) a. Taroo-wa [<sub>S</sub> Hanako-ga byooki-da ]-to it-ta.  
Taro-TOP Hanako-NOM be-sick-PRES that say-PAST  
'Taro said that Hanako was sick [at that time]'
- b. Taroo-wa [<sub>S</sub> Hanako-ga byooki-dat-ta ]-to it-ta.  
Taro-TOP Hanako-NOM be-sick-PAST that say-PAST  
'Taro said that Hanako had been sick'

Several accounts of SOT exist. In what follows I will briefly summarize some of them and relate them to the main topic of this chapter: long-distance dependencies.

Enç 1987 makes a parallel between binding theory and tense dependency. For her, tenses are treated as referential arguments (and they bear indices). Enç proposes *Anchoring conditions* in (36), which are formulated similarly to the principles of binding theory:

(36) *Anchoring conditions (Enç 1987)*

- a. Tense is anchored if it is bound in its governing category, or if its local Comp is anchored. Otherwise, it is unanchored.
- b. If Comp has a governing category, it is anchored if and only if it is bound within its governing category.
- c. If Comp does not have a governing category, it is anchored if and only if it denotes the speech time.

If the Tense is bound within its governing category, it is dependent on its binder, and one would get simultaneous readings, as in (32-b). In matrix clauses, the tense is always anchored to the speech time. In complement clauses, there are two options for the embedded past tense to be governed: either by the local Comp, or by the matrix Tense directly. In the former case, the local Comp is bound by the matrix tense, and therefore the embedded past tense is interpreted as shifted.

In the latter case, one gets a simultaneous reading. Further, in order to account for languages without SOT, such as Russian and Japanese, Enç proposes that in these languages tenses have to be anchored by Comp, and cannot be bound by the superordinate tenses, since, according to her, Russian and Japanese Comps are stipulated to always have a temporal index.

A different account for the SOT phenomenon is given in a series of works by Abusch (1988, 1994, 1997). In the last version of her account given in Abusch 1997, she provides a theory where all temporal relations (relating indices to local evaluation time) are transferred by intensional operators such as *believe* or *desire* cumulatively down the tree using a feature-passing mechanism, “so that a tense embedded in an intensional context has access to a set of temporal relation variables.” Further, there are certain constraints on the tenses, for instance, for past tense at least one of the relations passed to it must be the temporal precedence relation.

(37) Mary believed it was raining.

Therefore, in sentence like (37), there are two ways of satisfying the temporal constraint: either both the embedded tense relation and the matrix tense relations are temporal precedence relations satisfying the requirement, resulting in the shifted reading, or only the matrix tense relation is a temporal precedence relation, and the embedded tense is coindexed with it, in which case we obtain a simultaneous reading. However, it is hard to see how Abusch (1997) can account for non-SOT languages, like Russian: if feature-transmission is a universal property of intensional predicates, it is unclear why Russian and Japanese would be different from English.

The main problem of this account lies in the precise nature of this feature-transmission mechanism. This process is obviously counter-cyclic because of the top-down nature of the feature-transfer mechanism. Apart from that, it is also problematic from the phase theory point of view: the condition of the embedded clause depends on the way the tense is transmitted from the matrix clause.

Now consider the contrast between English (38) and Russian (39). If the clause embedded under past tense matrix clause has a future tense in it, like (38-a), the event described in the embedded

clause is evaluated with respect to the moment of utterance  $t$ . Therefore in this example, Mary's winning of the lottery was said to happen at some moment  $t'' > t$ , and therefore the continuation of the example "but she didn't" makes the sentence (38-a) awkward. In order to refer to the future with respect to the event described in matrix clause (at  $t'$ ), one would have to use conditional *would*, (38-b).

(38) *Sequence of Tense in English, Embedded tense = FUT/COND*

- a. #John said that Mary (will win)<sub>FUT</sub> the lottery, but she didn't.
- b. John said that Mary (would win)<sub>COND</sub> the lottery, but she didn't.

(39) *Sequence of Tense in Russian, Embedded tense = FUT*

Ivan skazal čto Maša vyigraet lotereju, no ona proigrala.  
 Ivan said that Mary win<sub>FUT</sub> lottery but she lost  
 'Ivan said that Mary would win the lottery, but she lost'

Now, in Russian the future tense embedded under the past tense allows for such a possibility. Therefore Russian sentence (39) which is analogous to English (38-a) does not exhibit any semantic awkwardness.

To account for SOT in English, Kondrashova 1999 proposes the following rule in (40):

(40) *Sequence of Tense Rule for English (Kondrashova 1999)*

In a sequence/clause [...  $\tau$  ...  $\forall \mu$ ], where  $\tau$  is tense variable,  $\mu$  is morphological tense feature of the verb,  $R$  is reference time,  $t^*$  is speech time, and  $t$  is variable over moments, if  $\|t\| = R$  such that  $\forall t \in R, t < t^*$ , then spell out morphological markers that correspond to the three features of  $\mu$  as follows:

PRES as *-ed*,

PAST as *had -en*,

FUT as *would*.

This rule is however problematic for phase theory. How is a sentence (38-b) derived? In order to Spell Out the embedded tense as *would+V*, a certain process of agreement between the matrix tense and embedded tense must take place. Notice, that matrix verb cannot trigger this process because its T-features are valued — and thus it cannot probe down to the embedded T. However, by the time the matrix T is merged into derivation, the embedded T has already been sent to Spell Out (to both  $\Sigma$  and  $\Phi$ ), and therefore inaccessible to syntactic operations at that moment: Agree between matrix T and embedded T is banned based on phase theory.

In order to resolve this problem in the following parts of this dissertation I will propose that embedded C participates in this tense feature transfer, similar to Russian subjunctives.

## 2.4 Phenomena to consider

In this section I will consider long-distance phenomena which behave differently based on the clause type, such as extraction processes which have different syntactic properties depending on whether the clause the movement originates in is indicative/subjunctive/infinitive etc. Nothing about the existence of such long-distance processes is surprising given the phase theory: the escape hatch on the phase edge allows for long-distance extraction from the phases. At the same time, if the extraction is allowed out of certain types of clauses and disallowed out of other types, an explanation is in order why the escape hatch exists and available in one case and does not exist or unavailable in the other. Potentially, such an explanation can be based on the phasal properties of a particular type of embedded clause, and on finer distinction among different types of phases.

Such processes that I will review in this section are:

- (41) *Processes which behave differently with respect to clause types:*
- a. Wh-movement
  - b. Scrambling
  - c. Clitic-climbing

### 2.4.1 Long distance extraction I: Wh-movement

In French, as claimed by Tsoulas 1995, *wh*-extraction out of subjunctive and infinitival complement clauses is noticeably better than extraction out of indicatives. The relevant examples are given in (42). Example (42-a) shows that long-distance *wh*-extraction is prohibited out of indicative complements while examples (42-b) and (42-c) demonstrate the possibility of long-distance *wh*-extraction out of subjunctive and infinitive complements respectively.

(42) *Long-distance wh-extraction in French*

- a. \*Que te demandes-tu qui a dit que Theo a vu t?  
what you.REF ask.2SG-you who has say.PPL that Theo has see.PPL  
'What do you wonder who said Theo saw?'
- b. Que te demandes-tu qui a exigé que Sophie écrive t?  
what you.REF ask.2SG-you who has require.PPL that Sophie write.SUBJ.3SG  
'What do you wonder who required that Sophie write?'
- c. Que te demandes-tu qui a décidé de voir t?  
what you.REF ask.2SG-you who has decided to see.INF  
'What do you wonder who decided to see?'

In Georgian (Harris 1981), *wh*-extraction out of indicative clauses to form a direct question is prohibited, as illustrated in (43-b), which is an attempt to create an indirect question from the declarative (43-a). In fact, as shown in (43-c), leaving the *wh*-word *in situ* is also ungrammatical.

(43) *Impossibility of long-distance wh-questions in Georgian*

- a. (ase) vpikrob rom Nino moigeb  
thus think.1SG that Nino win.FUT.INDIC  
'I think that Nino will win'
- b. \*vin pikrob (rom) moigeb?  
who think.2SG that win.FUT.INDIC  
'Who do you think will win?'
- c. \*pikrob (rom) vin moigeb?  
think.2SG that who win.FUT.INDIC  
'Who do you think will win?'

However, when the clause is embedded under one of three verbs: *unda* ‘want’, *šeužlia* ‘can, able’, and *sč’irdeba* ‘need’, such extraction is possible. All of these verbs are subcategorized for a subjunctive complement.

(44) *Long-distance wh-questions in Georgian; volitional matrix verbs*

- a. es minda rom Ninom moigos  
this want.1SG that Nino win.FUT.SUBJ  
‘I want Nino to win’
- b. vin ginda rom moigos?  
who want.2SG that win.FUT.SUBJ  
‘Who do you want to win?’

As it stands now, Phase Theory cannot easily make this distinction. Fine-grained distinction between phasal properties of different types of CPs is needed to account for examples above.

## 2.4.2 Long distance extraction II: Scrambling

The next type of movement which I consider in this section is (long-distance) scrambling. Consider example (45) from Progovac 1991.

(45) *Scrambling in Serbo-Croatian (Progovac 1991)*

- a. \*Petra Mira ne tvrdi [ da e vidi τ]. (Indicative)  
Peter<sub>ACC</sub> Mira not claims that sees  
‘Mira didn’t claim that she saw Petar’
- b. Petra Mira ne želi [ da e vidi τ]. (Subjunctive)  
Peter<sub>ACC</sub> Mira not wants that sees  
‘Mira doesn’t want to see Petar’

In Serbo-Croatian, long-distance scrambling is usually prohibited out of indicative embedded clauses. For instance, the embedded object *Petra* in (45-a) cannot be scrambled into the matrix clause. At the same time, if the embedded clause is subjunctive, as in (45-b), such long-distance scrambling is allowed.

Similarly, one needs a finer distinction between different CP types to be a part of Phase Theory to explain these differences.

### 2.4.3 Long distance extraction III: Clitic-climbing

As has been claimed in the literature, in Romance languages clitic-climbing is only allowed out of non-finite clauses. Example (46) demonstrates that clitic climbing is not allowed from subjunctive complements, and it is allowed from infinitival complements.

(46) *Spanish clitic climbing*

- a. \*Juan los quiere que (María) vea.  
Juan 3PL.ACC wants that Maria see<sub>SUBJ.3SG</sub>  
'Juan wants Maria to see them'
- b. Juan los tiene que ver.  
Juan 3PL.ACC has that see  
'Juan has to see them.'

The situation is different in Serbo-Croatian. Clitic climbing is still prohibited out of indicative complements, (47-b), similar to Romance languages. At the same time, subjunctive complements of volitional verbs, allow clitic climbing, (47-d), as opposed to subjunctives in Romance. This is reminiscent of the possibility of long-distance scrambling which in Serbo-Croatian is possible out of subjunctive clauses as opposed to indicatives, as I showed above.

(47) *Serbo-Croatian clitic climbing*

- a. Milan kaže da ga vidi.  
Milan says C him<sub>ACC</sub> sees  
'Milan says that he can see him'
- b. \*Milan ga kaže da vidi.  
Milan him<sub>ACC</sub> says C sees
- c. Milan želi da ga vidi.  
Milan wishes C him<sub>ACC</sub> sees  
'Milan wishes to see him.'

- d. Milan ga želi da vidi.  
Milan him<sub>ACC</sub> wishes C sees

Again, clausal types distinctions need to be a part of phase theory if one wants to derive these properties of clitic-climbing from the properties of phases.

#### **2.4.4 Conclusion**

Above I demonstrated that various syntactic phenomena behave differently with respect to different types of clauses and different types of VP/*v*Ps. If one wants to claim that all syntactic dependencies are phase-based (a desired result in current minimalist framework), phase theory needs to be revised so that various types of clauses and verbal projections can be taken into consideration. For instance, it is not enough to claim that CP is a phase: one needs to make finer distinction between control CP, subjunctive CP, indicative CP, etc. In what follows I propose how domains acquire phasal status based on the properties of the elements which are contained within such domains. While I will not provide an account of how all syntactic phenomena considered above would work within revised phase theory framework, in Sections 3-5 I show how this revision of phase theory can be applied to anaphoric binding.

### **2.5 Deriving phases**

In the rest of this chapter I propose that phases can in fact be *derived*. I show that there is no need to stipulate which category is a phase, rather the syntactic-semantic properties can explain which domains must have phasal properties. Before taking on this goal, I will outline the theoretical framework I am going to assume in this.

#### **2.5.1 Theoretical assumptions about feature-checking**

Throughout this chapter I follow the approach to feature-checking outlined in Pesetsky and Torrego 2007, which I will summarize below. It is based on the possibility of feature sharing, and allows a



feature to have several instances in various locations within the syntactic tree. The crucial operation is the following version of Agree stated in (48).

(48) *Agree: Feature Sharing Version* (from Pesetsky and Torrego 2007)

- a. An unvalued feature  $F$  (a probe) on a head  $H$  at syntactic location  $\alpha$  ( $F_\alpha$ ) scans its  $c$ -command domain for another instance of  $F$  (a goal) at location  $\beta$  ( $F_\beta$ ) with which to agree.
- b. Replace  $F_\alpha$  with  $F_\beta$ , so that the same feature is present in both locations.

This Feature-Sharing version of Agree may create multiple instances of a single feature in various syntactic locations within the structure. After probing by a head with an unvalued feature, the features of a goal and a probe enter into an Agree relation, and both become instances of the same feature.

Another crucial assumption which is needed to maintain feature sharing is the elimination of Chomsky's Valuation/Interpretability Biconditional that allows only uninterpretable and unvalued  $\langle uF -val \rangle$  and interpretable and valued  $\langle iF +val \rangle$  features. In Pesetsky and Torrego 2007 framework, two more types of features are allowed: uninterpretable and valued  $\langle uF +val \rangle$  and interpretable and unvalued  $\langle iF -val \rangle$ . In this framework, unvalued features ( $\langle uF -val \rangle$  or  $\langle iF -val \rangle$ ) act as probes.

For instance, the T-feature on T is interpretable, but unvalued, and that allows it to be a probe. On the contrary, the T-feature on a finite verb is uninterpretable (no semantic interpretation happens within the verb itself), but valued, since verb comes from the lexicon with morphologically specified tense.

## 2.5.2 Feature content of $v/N/T$

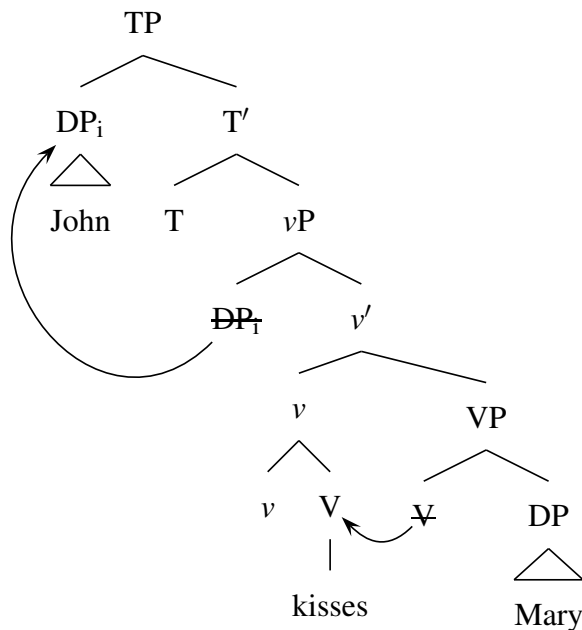
One of the questions I will explore in this section is what the featural content of  $v$  and  $V$  is as they come from the lexicon. At the moment, I will concentrate my attention to the simple case of a

monoclausal transitive sentence with no reflexive pronouns, such as the one in (49):

(49) John kisses Mary.

Traditionally it is assumed that *Mary* is a complement of the verb, undergoing a Merge operation with the V *kisses*, forming a VP. Following that, V universally raises to *v*, and the subject *John* originates at Spec, *v*P position. After the *v*P is complete, T merges into the structure, and probes the *v*+V verbal complex, presumably in order to check its T-features. In English, this probing does not result in verb-raising for certain language-specific reasons, while in French it results in verb raising to T (see, for example, Adger 2004). Further, the existence of the EPP feature on T would force the closest DP, in this case the subject *John* to raise to Spec,TP position. Similar to *v*+V raising, the existence of the EPP feature is language specific: the Subject raises in the majority of SV languages, such as English, French, Russian, etc., and (arguably) stays in Spec,*v*P position in VS languages, such as Irish.

(50) Basic derivation of (49)



Now, let's consider this derivation from the formal standpoint of Pesetsky and Torrego. DPs come from the lexicon endowed with  $\phi$ -features, which are interpretable and valued (at least for R-expressions), i.e.  $\langle i\phi +val \rangle$ . In case of a simple sentence (49), the subject DP *John* comes with  $\langle i\phi +val \rangle$  and is endowed with specification  $\langle masc., 3rd, sg. \rangle$ , and object DP *Mary* comes with  $\langle fem., 3rd, sg. \rangle$  specification of its  $\phi$ -features.

The next question is, with which features does a verb come from the lexicon? As argued in Pesetsky and Torrego 2001, 2007, a verb has a T-feature. They argue that this feature is valued on the verb, but uninterpretable, as the semantic locus of tense interpretation is T, and not V. Pesetsky and Torrego are not distinguishing between  $v$  and V in their derivations, assuming that V and  $v$  form a head-complex at the moment they are probed and enter in Agree relation with T. I will nevertheless assume that valued T-feature must be present on verb itself, and not on  $v$ , and it is specified as  $\langle uT +val \rangle$ .

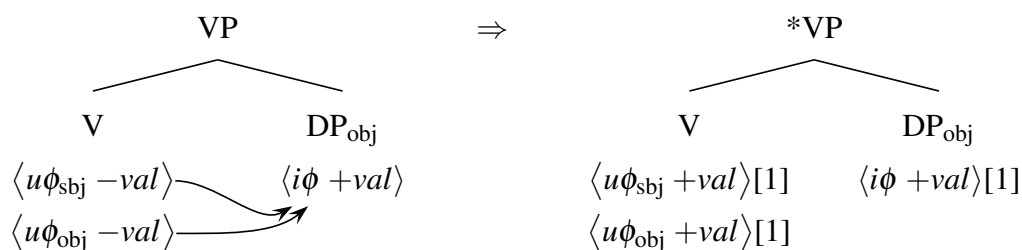
Further, the issue of whether  $\phi$ -features are present on the verb has not received enough consideration in the previous literature. Notice that in order to incorporate languages with object agreement on the verb, two sets of unvalued  $\phi$ -features must be checked with the subject and object respectively. Two general possibilities exist: either both unvalued sets are present on the same head, or they are split between two heads.

Let's consider first if it is possible to have both sets of  $\phi$ -features on the same head (V and  $v$  are the potential heads in question). Assume for the sake of simplicity that V is such a head, and it must have its sets of  $\phi$ -features checked on the object and subject. The potential derivation can proceed as follows. V with two sets of unvalued uninterpretable  $\phi$ -features,  $\langle u\phi_{sbj} -val \rangle$  and  $\langle u\phi_{obj} -val \rangle$ , is merged with an object, which is endowed with a set of valued interpretable  $\phi$ -features  $\langle i\phi +val \rangle$ . At this moment, the verb must probe the object, since it is  $c$ -commanded by the verb. Which of the two sets of  $\phi$ -features on the verb will get valued at this moment? There is no inherent formal property of the object DP that would distinguish it from, say, subject DP, and would force only object-related  $\phi$ -features on the verb to get valued. There is also no inherent formal difference between two sets of  $\phi$ -features on the verb, that would force only object-related

set to probe, and would protect the subject-related  $\phi$ -features to get valued. In fact, in the current minimalist framework, features of the probe get valued as soon as it finds a goal which has the same type of features. Therefore, both sets of  $\phi$ -features on verb will receive their value from the object DP, and neither of them will be left unvalued in order to get a value from the subject, contrary to what is desired. Similar argumentation is possible if one assumes that subject- and object-related  $\phi$ -features are both present on some other head, say  $v$  or T.

Schematically this is illustrated in (51).

(51) *Two sets of  $\phi$ -features on V get valued simultaneously*

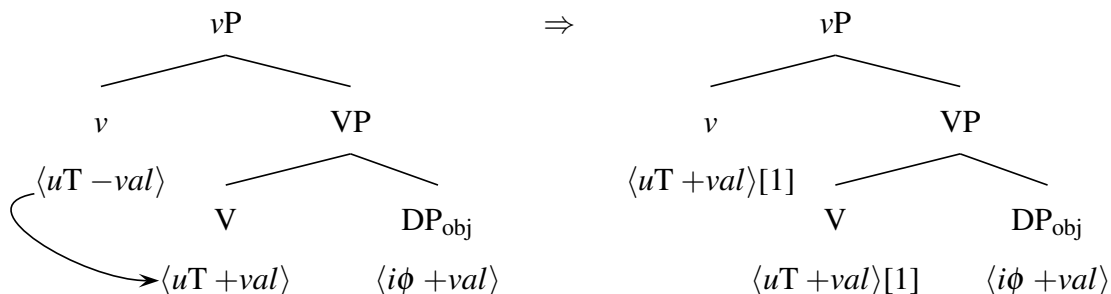


The second option, where subject- and object-related  $\phi$ -features responsible for verb agreement are located on different heads, remains the only possibility. Which heads house subject agreement and object agreement? I argue that  $v$  is the host of object agreement  $\phi$ -features. Theoretically this is a desired assumption since the VP domain is related to internal argument structure of the verb (see for example Kratzer 1996). Below I return to the question of where the subject-agreement  $\phi$ -features are located.

On these assumptions object agreement happens on  $v$  as soon as it probes the direct object. Formally this proposal amounts to the following. Both  $V$  and  $v$  come from the lexicon endowed with T-features.  $V$  is merged with the direct object DP, and no probing for  $\phi$ -features occurs at this moment. Next,  $v$  is merged into structure. I argue that there is a T-feature present on  $v$ . Further, this feature is also present on  $V$ . Following Pesetsky and Torrego, I assume that it is uninterpretable on both  $v$  and  $V$ , valued on  $V$ , and unvalued on  $v$ . Therefore, when  $v$  is merged, it probes down, finds the T-feature on  $V$ , and enters a sharing relation with it. This step is shown below. The arbitrary

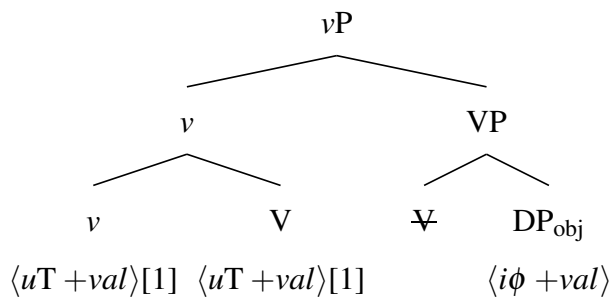
index [1] signifies that T on  $v$  and T on V are instances of the same T-feature.

(52) *v-V T-feature sharing*



Further, V universally raises to  $v$ , resulting in the creation of a  $v+V$  head-complex; this configuration is shown in (53).

(53) *V-to-v raising*



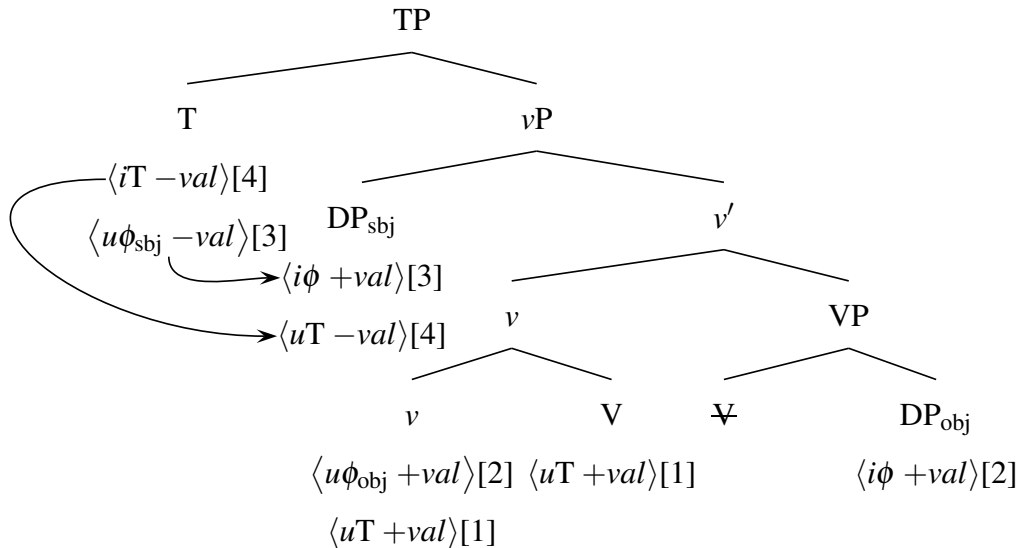
However, the  $\phi$ -features of  $v$  are still not valued. That forces  $v$  to probe the direct object DP, and that probing results in valuation of  $\phi$ -features of the  $v$ . Schematically, this is shown in (54), where the (arbitrary) index [2] signifies feature-sharing in Pesetsky and Torrego framework<sup>12</sup>.

<sup>12</sup>Note that I use the subscripted  $\phi_{obj}$  for illustration purposes, the minimalist computational system cannot make a distinction between subject and object agreement  $\phi$ -features.



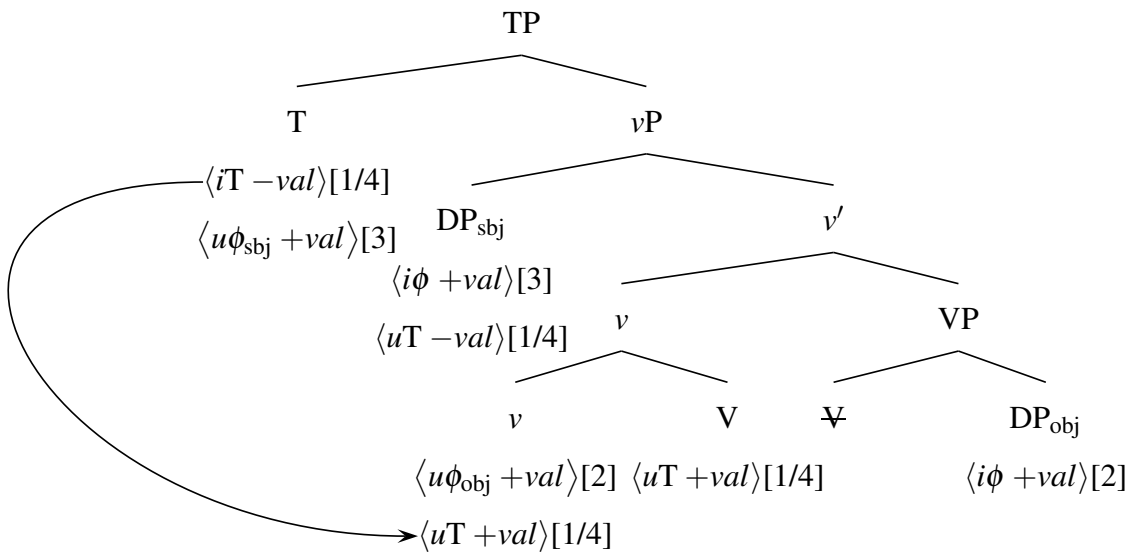
T nor on the subject DP. In (55), indices [3] and [4] indicate feature-sharing that took place after this probing.

(55) *T probes the subject DP*



The next step is crucial and results in the valuing of all unvalued features. T continues to probe down, as its T-features are still unvalued, and finds the head complex  $v+V$  as its goal, as it has valued T-feature. Probing of  $v+V$  results in valuation of T-features on T and consequently on the subject DP, as these T-features are shared. Features marked with [1] and [4] in (55) become instances of the same feature, and acquire the same index, shown in the corresponding trees by [1/4]. This last probing is shown in (56).

(56) *T probes v+V complex*



Following the valuation of all features, the subject DP raises to Spec,TP in order to satisfy the EPP feature of T. The resulting configuration with all features being valued is shown in (57), the derivation succeeds.





closed for further operations: it is impossible to probe into them.

Let us look at this intuition from a feature valuation point of view. The answers to the following questions may provide an idea of what one should be calling a phase:

- Which featural specifications can make a domain a phase?
- When there is no more need to probe/access the features on a certain element, i.e. which domains can be closed in order to satisfy the Phase Impenetrability Condition?
- How must the features on the elements within a domain be specified in order for the domain to be able to undergo Spell Out?

Consider a domain D selected by a head H, with an element inside the domain with a feature F, [ H [<sub>D</sub> ... ⟨F⟩ ... ]]. As argued by Pesetsky and Torrego 2007, there are four possible specifications for this feature: ⟨uF –val⟩, ⟨uF +val⟩, ⟨iF –val⟩, ⟨iF +val⟩. Let us consider these possibilities.

First, imagine that the feature F is valued, i.e. we have a domain D defined as in (58):

- (58) a. [ H [<sub>D</sub> ... ⟨iF +val⟩ ... ]]  
b. [ H [<sub>D</sub> ... ⟨uF +val⟩ ... ]]

In this case such a feature cannot prevent the domain D from being sent off to the interpretative component. The feature F is valued, and therefore will not be problematic at Spell Out: there is no crash at the syntactic level. Therefore, domains in which all features are valued can be closed for future probing, and be spelled out. A domain D in which all features have valued instances can qualify as a phase. However, if at the level of semantic evaluation, the feature F will not have an interpretable instance, the derivation will crash at the semantic level.

Now, let's consider another situation when the feature F is unvalued, i.e. the domain is defined as in (59):

- (59) a. [ H [<sub>D</sub> ... ⟨uF –val⟩ ... ]]  
b. [ H [<sub>D</sub> ... ⟨iF –val⟩ ... ]]

If one tries to ship the complement of the head H (the domain D) to the interpretative component crash is inevitable. The feature F has no value, and therefore D cannot be closed for probing yet. However, in some cases the situation can be saved.

Imagine that an instance of feature F is present on the head H. Further consider a situation when the feature F on H is unvalued, i.e. H serves as a probe for a feature F within the domain D, and the feature F becomes shared between its instance within the domain D and its instance on the head H.

- (60) a. [ H<sub><{u/i}F-val>[1]</sub> [D ... <uF-val>[1] ... ]  
 b. [ H<sub><{u/i}F-val>[1]</sub> [D ... <iF-val>[1] ... ]

Since F is now present on both the head H and within the domain D, there is no need to access the instance of the feature F present within the domain D. In fact, there is no longer a need to value the instance of F within the domain D. As soon as the instance of F on H is valued, so is the instance of F within D.

Thus, I showed that while D might not be interpretable or valued at the interface with the semantic component, there is still no need to access it for the syntactic system from the economy considerations. The domain D can be closed for probing as soon as all its edge features are either valued, or shared with a certain head H, external to the domain. I will call such domains *phasal*.

- (61) **Phasal domain, Definition 1:** A domain D is *phasal* if for every feature F within D,  
 a. F is valued, or  
 b. a head H merged with D has an instance of a feature F that entered sharing a probe-goal relationship with an instance of a feature F within D.

Note that such domains D partially correspond to complements of phase heads in Chomsky 2001.

There is one major problem with this definition. The element embedded deep inside the domain D with an unvalued feature F will prevent not only D from closing, but all domains containing D,

until the moment its feature F is valued. The problem with that is that there will be no more locality in syntax: any unvalued feature will be able to get its value from an probe located much higher in the structure, i.e. situations like (62) will be impossible to prevent:

$$(62) \quad [ H_{\langle iG -val \rangle, \langle iF +val \rangle} \dots [D1 \dots [D2 \dots XP_{\langle uG +val \rangle, \langle uF -val \rangle}$$

In such cases, the presence of the unvalued F-feature on XP will prevent domains D1 and D2 from closing. After that, if a head H is inserted with an unvalued instance of G feature, and a valued instance of F-feature, H will be able to probe XP in order to value its G-feature. This valuation will also result in valuation of F-feature on XP, as a byproduct of G-feature probing.

An example of this situation is *wh*-movement. Assume that H from (62) is an interrogative C with  $\langle iQ +val \rangle$  and  $\langle uwh -val \rangle$ , and XP is a *wh*-phrase with  $\langle uQ -val \rangle$  and  $\langle iwh +val \rangle$ . The existence of an unvalued Q-feature on the *wh*-phrase will prevent any intermediate CP from closing, and therefore the matrix interrogative CP will be able to probe *wh*-phrase directly. Therefore, there will be no successive cyclic movement, contrary to all existing evidence.

Similarly, if one assumes that NPIs are endowed with an unvalued interpretable feature which must enter into a probe-goal-relation with some higher negative head, one would predict that a domain containing NPI will never be closed until the negative probe is merged. Therefore, arbitrary long-distance NPIs are expected to exist, which is not the case, at least in the case of Japanese/Korean (Nakao and Obata 2007).

In fact, from the approach described above it follows that the elements with unvalued features will never have to move to the phase edges, or undergo any successive cyclic movement in general, since it will always be possible to target them directly by a higher head, since they prevent domains from closing.

The proper version of the definition above would be to close domain D only if the features *relevant for this domain* are valued. For instance, TP can be closed if the T-feature of T is either valued, or shared with an above element, say C. This will predict that there cannot be syntactic probing into the finite indicative TP from the matrix clause, no matter what is contained in the

embedded  $vP$ . Similarly, VP can be closed if its arguments have valued features, such as  $\phi$ -features, and its T-feature is either valued or shared with  $v$ . The final version of the definition of a *phasal domain* is given in (63):

- (63) **Phasal domain, Final definition:** A domain  $D$  is *phasal* if for every *D-relevant* feature  $F$  within  $D$ ,
- a.  $F$  is valued, or
  - b. a head  $H$  merged with  $D$  has an instance of a feature  $F$  that entered a sharing probe-goal relationship with an instance of a feature  $F$  within  $D$ .
- (64) **D-relevant features:** Assume that  $D$  is a domain, such as VP,  $vP$ , TP, or CP. Then, *D-relevant features* are the following:
- a. If  $D=VP$ , *VP-relevant* features are:  $\phi$ - and Case-features of the nominals within V-complement and VP specifier.
  - b. If  $D=vP$ , *vP-relevant* features are:  $\phi$ -features of the Spec, $vP$ , i.e.  $\phi$ -features of the subject, interpretation related features on  $v$ , i.e. Definiteness feature.
  - c. If  $D=TP$ , *TP-relevant* features are: T-features of T and  $\phi$ - and T-features of the element in Spec,TP.
  - d. If  $D=CP$ , *CP-relevant* features are: Clause-type feature on C and *wh*- and Q-features of C and Spec,CP.

In the next chapters of this dissertation I show the consequences of this definition for binding. Here I would like to exemplify a few cases of phasal domains.

## 2.5.4 Deriving phases

In this section I show what predictions are made by the approach taken above.

Let's consider the derivation of a simple transitive sentence, such as the one described in the previous section, (49), repeated here as (65):

(65) John kisses Mary.

Which domains for this sentence are phasal, in the sense described in (63)?

As soon as  $\nu$  is merged into structure and probes V, VP becomes phasal:

Following the definition of a *phasal domain*, VP,  $\nu$ P are phasal domains, since they only depend on  $\phi$ - and Case-features of the DPs inside of them, and all of them are valued, since in sentence (65) all DPs are referentially independent (i.e. they are neither NPIs, nor reflexives, nor any other DPs requiring special environments), and not on T-features.

Now, after T is merged with  $\nu$ P, the T-feature on T gets its value after it probes the  $\nu$ +V complex. Therefore by the time the embedded C is inserted, TP becomes a phasal domain. Now, indicative C does not depend on the matrix C, and therefore CP will also be a phasal domain.

Now if we consider the embedded indicative transitive clause, nothing will change: the embedded CP will still be closed, and all features within it are valued; further matrix VP,  $\nu$ P, TP and CP will also be phasal domains, in the same way as their counterparts in the embedded clause.

Therefore, the theory outlined in the previous section predicts that in indicative clauses VP,  $\nu$ P, TP, and CP are all phasal domains. This result is summarized in (66), where ✓ means that the domain is phasal, while ✗ means that it is open.

(66) In indicative embedded clauses:

a.  $\nu$ P and VP, TP, and CP are phasal domains;

	CP <sub>m</sub>	TP <sub>m</sub>	$\nu$ P <sub>m</sub>	VP <sub>m</sub>	CP <sub>e</sub>	TP <sub>e</sub>	$\nu$ P <sub>e</sub>	VP <sub>e</sub>
b. Status:								
✓ — phasal	✓	✓	✓	✓	✓	✓	✓	✓
✗ — non-phasal								

(here and below the index ‘m’ stands for ‘matrix’, ‘e’ — for ‘embedded’)

### 2.5.4.1 Control clauses

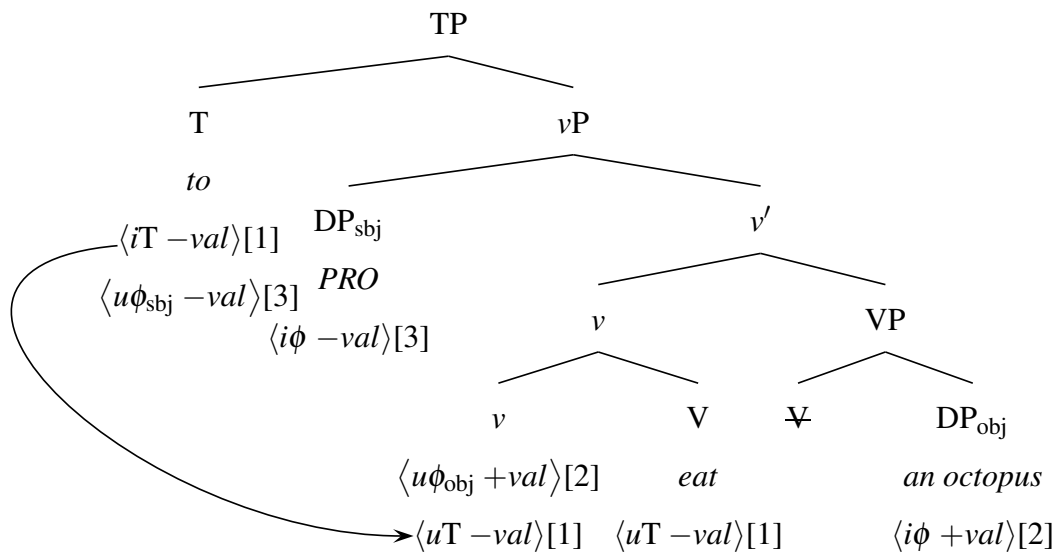
Now, let's consider embedded control clauses. According to Pesetsky and Torrego 2007, infinitival verbs come with an unvalued T-feature, and it is valued through a relation with the matrix verb. Is infinitival ECM VP still phasal? I outline the full derivation of control clauses in Chapter 4, but for now I will show how the proposed system works. At the moment  $v$  is merged, VP is a phasal domain, since the T-feature is not a *relevant feature* for the VP domain. Further, when T is merged,  $vP$  has a PRO subject, which is unspecified for  $\phi$ -features, and therefore control  $vP$  is not a phasal domain. However, when T is merged into structure, PRO will raise to Spec,TP, and by then  $vP$  becomes a phasal domain.

Now, when T probes  $v$  for its T-feature, no valuation occurs, since the T-features on the infinitival  $v+V$  are unvalued. Therefore, the T-feature on T (a *relevant feature* for TP) is unvalued. Further, the existence of the PRO subject with unvalued  $\phi$ -features in Spec,TP position (also a *relevant feature* for TP) will also prevent TP from acquiring phasal status. Thus, the embedded control TP at this moment cannot be closed for syntactic probing.

(67) *Control vP*

a. John tried PRO to eat an octopus.

b.



Now, assume that control C has an instance of unvalued uninterpretable T-feature, and an instance of a clause-type feature (specifying that it is a control clause), which must be valued under a relation with the matrix T/C. The T-features on C will probe the T-features on T, and become shared (however still unvalued). This probing however will not allow TP to become phasal: PRO in Spec,TP would prevent it. Now, CP will also not be able to become closed, because of the clause-type feature. In fact, CP will stay open until the matrix T/C is merged.

Further, unvalued features on C will prevent matrix the VP/vP from closing, since they can only be closed if the arguments of the verb have all their features valued.

Therefore, I conclude the following: in control clauses, embedded vP is a phasal domain, closed for syntactic probing after embedded T is merged, while embedded TP, CP, and matrix VP and vP are open until the matrix T/C is merged. This is summarized in (68).

(68) In sentences with embedded control clauses:

- a. (i) Embedded vP and VP are phasal domains;
- (ii) TP and CP are open domains;
- (iii) Matrix vP and VP are open domains.

	CP <sub>m</sub>	TP <sub>m</sub>	vP <sub>m</sub>	VP <sub>m</sub>	CP <sub>e</sub>	TP <sub>e</sub>	vP <sub>e</sub>	VP <sub>e</sub>
b. Status:								
✓— phasal	✓	✓	✗	✗	✗	✗	✓	✓
✗— non-phasal								

#### 2.5.4.2 ECM Clauses

Now consider the situation with ECM embedded clauses.

Consider the derivation of an infinitival sentence like the one in (69):

(69) John expects Mary to drink milk.

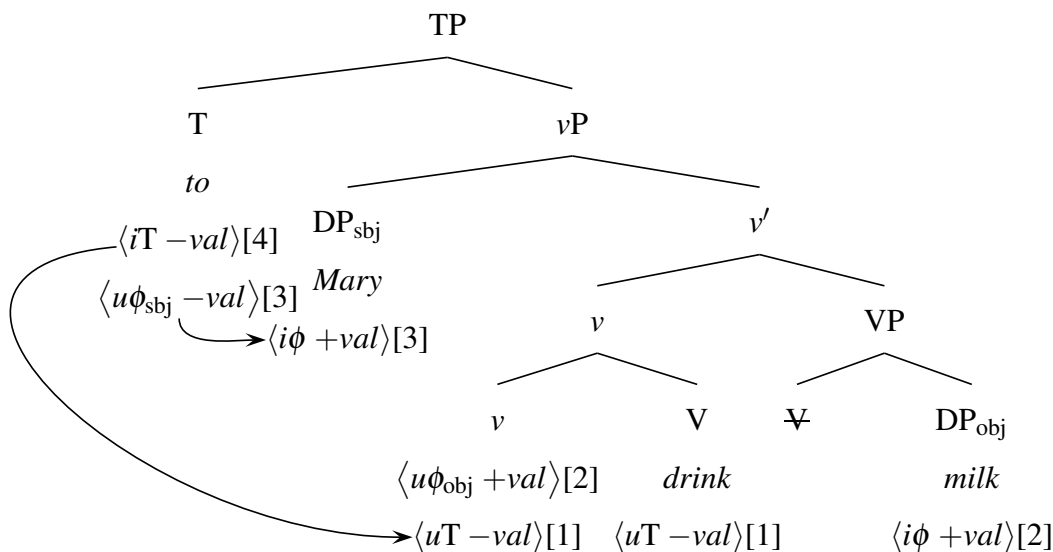
I assume as for Control clauses that the embedded infinitival V has an unvalued uninterpretable



instance of a T-feature. Further, I follow Chomsky in assuming that there is no embedded CP layer in embedded ECM clauses.

The embedded  $vP$  of this sentence with featural specification of its elements is shown in (70) at the moment T probes  $v+V$ .

(70) *ECM vP*



In this case, there are no  $vP$ -relevant unvalued features within  $vP$ , and therefore  $vP$  is a phasal domain. Now, consider the embedded TP. The embedded subject is caseless before the matrix case-assigner is merged. Further, the T-feature on T is unvalued, since there is no value for T on the embedded infinitival verb. However, upon merge of the matrix V, T acquires its value, and the subject in the embedded Spec,TP acquires its (accusative) case. Therefore, the embedded TP in ECM constructions is phasal, according to the definition of *phasal domain*.

Matrix VP and  $vP$  will also be phasal, as the embedded TP will not prevent them from closing, since all  $TP$ -relevant features are valued in the embedded clause.

(71) In embedded ECM clauses:

- a.  $vP$ , VP, and TP are phasal domains;

	CP <sub>m</sub>	TP <sub>m</sub>	vP <sub>m</sub>	VP <sub>m</sub>	TP <sub>e</sub>	vP <sub>e</sub>	VP <sub>e</sub>
b. Status:							
✓— phasal	✓	✓	✓	✓	✓	✓	✓
✗— non-phasal							

### 2.5.4.3 Object shift

Now, what can prevent these domains from becoming phasal, for instance, how can one get object shift if VP is phasal? Recall, that object shift is not a semantically vacuous operation: the shifted object gets a definite interpretation, as shown by the following data from Icelandic (from Alexiadou et al. 2007):

- (72) a. Hann las ekki bækur.  
 he reads not books  
 'He doesn't read any books.'
- b. Hann las bækurnar ekki.  
 he read the books not  
 'He doesn't read the books'

In order to explain the possibility of object shift in case of a definite direct object, let us assume that it is triggered by a certain feature,  $\langle OS \rangle$ . What are the specifications of the OS features and where is it located in the syntax?

I argue that the valued instance of  $\langle OS \rangle$  is on the direct object DP, but it is uninterpretable at that location,  $\langle uOS +val \rangle$ . The counterpart of that feature is on  $v$ , where it is interpretable but unvalued,  $\langle iOS -val \rangle$ . Now, let's consider the moment when  $v$  probes V.



VP from becoming phasal. Since VP is closed, no higher head will be able to probe the direct object, and the OS-feature on the object DP won't have any interpretable instance. That would lead to a derivation crash at the semantic interface.

#### 2.5.4.4 Ditransitive $v$ Ps

Now, let us consider double object constructions in English, where the goal DP occupies a position before the theme, such as the sentences in (75).

(75) John showed Mary the world.

I assume the structure of double object construction as described in Larson 1988 based on the thematic hierarchy (76) from Baker 1988; Carrier-Duncan 1985. According to Larson 1988, goals are located lower than themes on the thematic hierarchy, and therefore occupy a lower position in the constituent structure, following the principle P2 in (77).

(76) *Thematic hierarchy*

AGENT > THEME > GOAL > OBLIQUES (manner, location, time, ...)

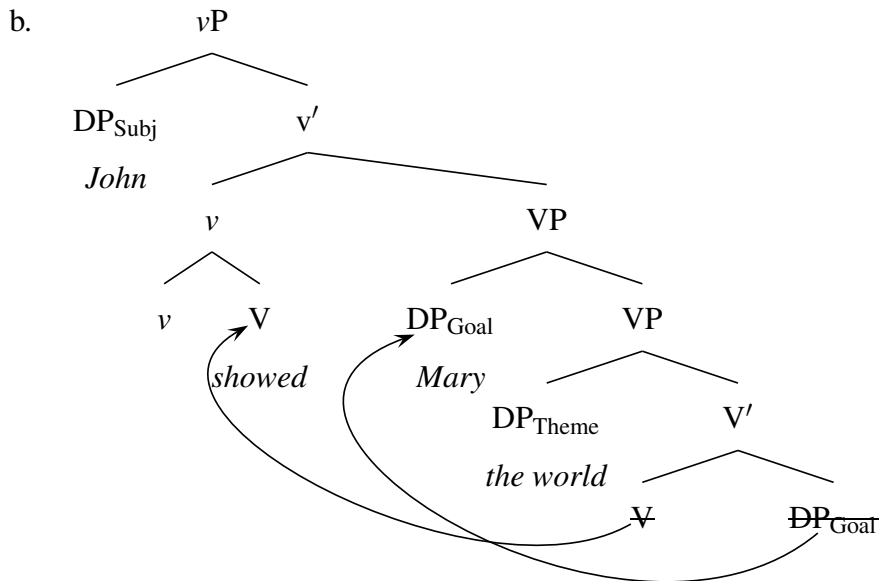
(77) *P2 (Larson 1988, pg. 382)*

If a verb  $v$  determines  $\theta$ -roles  $\theta_1, \theta_2, \dots, \theta_n$ , then the lowest role on the Thematic Hierarchy is assigned to the lowest argument in constituent structure, the next lowest role to the next lowest argument, and so on.

Following recent work by Larson, I assume that the goal argument is merged as a complement to the verb, and the direct object is merged in Spec,VP position. The movement of the goal DP happens in order for the case of the goal DP to be checked. Schematically this is illustrated in (78).

(78) *The structure of the double object construction:*

a. John showed Mary the world.

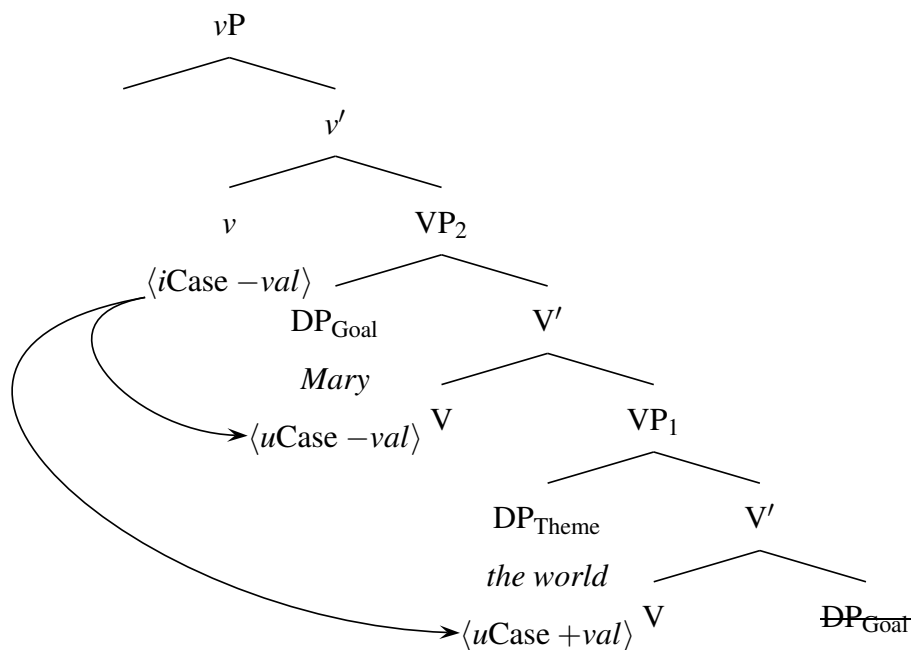


In the recent approach by Larson (see for example Harada and Larson 2009, the movement of the goal DP is needed in order for it to get case, in this case accusative. Working within Pesetsky and Torrego 2007 framework, Larson assumes that the goal DP comes with an uninterpretable unvalued case feature,  $\langle u\text{Case} -val \rangle$ , while the direct object comes with the valued accusative case feature,  $\langle u\text{Case} +val \rangle$ . He also claims that the matching case feature is present on  $v$ , where it is interpretable, but initially unvalued,  $\langle i\text{Case} -val \rangle$ . Therefore,  $v$  must eventually probe the direct object in order to get a value for its Case feature. Movement of the goal DP to the edge of VP ensures that the case feature on the goal DP gets valued. In order to derive the double object construction, Harada and Larson propose the following derivation.

First, the goal DP moves to the edge of VP. Then,  $v$  probes the goal DP since both of them have a Case-feature, and it is unvalued on  $v$ . This probing results in Case-feature sharing, however no valuation occurs: both goal DP and  $v$  have their Case-feature unvalued. After that step,  $v$  continues to probe and finds its goal on direct object, as it also comes with a Case-feature, this time valued. This probing results in sharing a Case feature between  $v$ , the goal DP, and the theme DP, and in

valuation of it on both  $v$  and the goal DP. Schematically it is illustrated in (79).

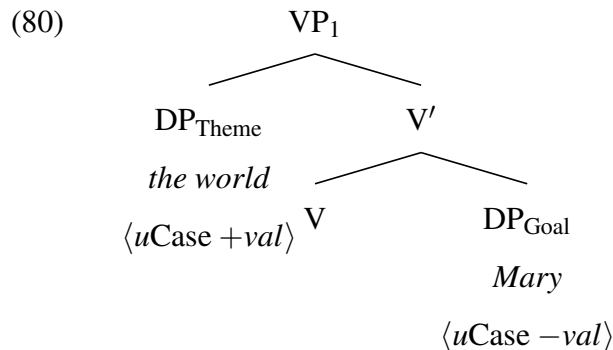
(79) *Case valuation in double object constructions.*



This derivation leads Harada and Larson 2009 to assume that case assignment on the goal DP is “concordial”, i.e it is dependent on the (accusative) case of the theme DP: according to Harada and Larson 2009, case on the goal DP is also accusative. Now, let’s consider how this derivation is relevant to the theory of phases proposed above. Notice, that Harada and Larson 2009 do not have a separate projection where the goal argument moves – he assumes that it moves to the extra specifier position in VP. In the tree (79), I assume the existence of an extra VP projection. To my understanding, nothing in his arguments crucially hinges on the absence of this extra projection, therefore from now on I will assume that the extra VP projection exists.

Which projections here are phasal? Notice, that according to the definition of a *phasal domain*, Case is a *VP-relevant* feature.

Consider first the lower  $VP_1$ . Its structure at the moment of completion is given in (80):



Notice that  $VP_1$  contains an unvalued feature, i.e. the Case-feature on the indirect object *John*. At the moment shown in (80) this is a feature which prevents VP from being closed for syntactic probing: an unvalued feature has not yet been shared with any higher head. Therefore, this  $VP_1$  is not phasal according to the definition in (63).

Now consider  $VP_2$ . After the Goal DP moves to its edge, it is probed by  $v$  in order to have its Case-feature valued, as shown in (79). Further,  $v$  probes the Theme DP, and that results in valuation of the Case features on both direct and indirect object DPs. At that moment, there is no unvalued features left in  $VP_2$ . Therefore,  $VP_2$  is a phasal domain, unlike  $VP_1$ .

#### 2.5.4.5 Conclusion

In (63) I proposed a definition of a *phasal domain*. I claim that a domain D is closed for probing just in case all the features relevant for this domain are either valued or shared with some element above this domain. I proposed which features are considered *relevant* for a given domain D in (64). With the exception of VP, for which the relevant features are the features of V and its complement, the domain-relevant features are in general the features of the edge of the domain, i.e. the head of the domain and the features of the specifier of the domain. As soon as all features of the edge of D are either valued or shared with an element outside of D, the domain D becomes closed. If D is a phasal domain, any remaining unvalued or uninterpretable feature within D becomes inaccessible for further probing. If such features exist, they will never be valued or will never acquire an interpretable instance, and the derivation will crash.

The cases considered above show that the definition of a phasal domain given in (63) does not contradict the standard view of Phase Theory. The fact that the (highest) VP is a phasal domain is consistent with the definition of  $\nu$ P as a phase. According to the standard version of the PIC, the complement of the  $\nu$ P phase, i.e. the highest VP, is closed for further syntactic evaluation. Similarly, in the approach taken in this chapter, VP is closed for syntactic evaluation since it is a phasal domain by the definition in (63), as I showed above.

In what follows I apply revised phase theory to the theory of anaphoric binding and demonstrate how reflexive binding domains can be derived from the revised definition of a phase.



# Chapter 3

## Binding in a feature-based framework

In this chapter I propose a feature-based theory of anaphoric binding. I demonstrate how anaphors acquire their antecedents by introducing a formal feature-checking mechanism within the Pesetsky and Torrego 2007 framework. I also show that the revised phase theory allows to reduce binding domains to phases. This approach to binding allows us to account for the following properties of anaphors:

- (1)
  - a. c-command requirement
  - b. locality of binding (in simple cases)
  - c. subject orientation of monomorphemic anaphors and lack of subject orientation for complex anaphors
  - d. Barss-Lasnik effects (binding in ditransitive constructions)
  - e. backwards binding possibility with psych-verbs

I introduce a formal reflexive feature  $\rho$ , interpretable and unvalued on some head ( $v/V$  or  $T$ ), but uninterpretable and valued on a reflexive. I argue that binding relations between reflexives and their antecedents are obtained when an unvalued instance of the  $\rho$ -feature (on the head) probes a valued instance of the  $\rho$ -feature (on the reflexive). This allows me to reduce the formulation of binding principle A to an Agree relation between two instances of the  $\rho$ -feature. Further, it allows

to independently derive the binding domain as a domain of an Agree operation. I argue that phasal domains (as described in Chapter 2) are in fact binding domains. This allows me to derive locality of binding in simple cases and Barss-Lasnik effects.

I further show how the presence of valued  $\phi$ -features on the reflexive affects the position where an interpretable instance of the  $\rho$ -feature can be placed. I show that in order to value  $\phi$ -features on the monomorphemic reflexive, the  $\rho$ -feature must be placed on the T-head to enforce probing of the reflexive by the T-head, which would result in  $\phi$ -feature valuation. This derives the subject-orientation of the monomorphemic reflexives. In the case of a complex reflexive, its  $\phi$ -features are valued in the lexicon, and therefore the interpretable instance of the  $\rho$ -feature can be placed on  $v$  or  $V$ , which would result in binding by the subject or the object, respectively.

## 3.1 Introducing relevant features

### 3.1.1 The reflexive feature $\rho$

In this section I introduce a formal reflexive feature, which I will denote as the  $\rho$ -feature (“rho”). However, before introducing it I will briefly review the proposal by Reinhart and Reuland 1993.

Reinhart and Reuland 1993 argue that Principle A of the binding theory is in fact a condition of marking predicates reflexive. They propose the following definition of a reflexive predicate:

- (2) a. A predicate is *i-reflexive* iff (at least) two of its arguments are *i*-coindexed (that is, are indexed *i*).
- b. A predicate (formed of P) is *i-reflexive-marked* iff either P is lexically reflexive with respect to an *i*-indexed argument, or one of P’s *i*-indexed arguments is a SELF anaphor.

Having introduced reflexive predicates, Reinhart and Reuland go on to propose the following reformulations of Condition A and B of the binding theory:

- (3) a. *Condition A*: An *i*-reflexive-marked predicate is *i*-reflexive.

- b. *Condition B*: An *i*-reflexive predicate is *i*-reflexive-marked.

These conditions claim that if two arguments of a predicate are coindexed, one of them should be a reflexive pronoun, such as SELF, and it cannot be any other expression, and vice versa, if one of the arguments of the predicate is SELF, the predicate is reflexive, i.e. its arguments must be coindexed, meaning SELF must have the same reference as the other argument of the verb, and cannot refer to another entity.

While attractive conceptually, this approach only explains a subset of binding effects in world languages. As we saw in Chapter 1, any binding into adjuncts, such as prepositional phrases, etc. does not fall under the binding theory according to Reinhart and Reuland. Reflexive pronouns within adjuncts are therefore all analyzed as *logophors*, and require an additional module of grammar which would govern their interpretation. Further, SE-elements in Germanic languages, according to Reinhart and Reuland 1993, also do not fall under Condition A of the binding theory. Reference to logophors is also made during consideration of long-distance binding phenomena, as it is impossible to derive them using the notion of a reflexive predicate. I have provided a detailed sketch of Reinhart and Reuland's 1993 theory in section 1.2.

Here, I build on Reinhart and Reuland's intuition about reflexive predicates, but reduce it to a formal syntactic process by introducing the  $\rho$ -feature. This feature is needed to capture the intuition behind reflexive predicates. More formally, the presence of a  $\rho$ -feature on a head is needed for coreference between a variable contained in the domain of the head (the anaphor) with an expression in the specifier position of the head (the antecedent). For instance, the presence of this feature on a transitive verb will require the complement of a transitive verb to be coreferent with its specifier, in sentences such as (4-a). Further, as I get away from the *purely semantic* notion of a reflexive predicate, the same feature would require any variable within the domain of the head to be bound by an expression in the specifier of the head, (4-b).

- (4) a. John loves himself.  
b. Students love [each other's teachers].

Transferring the realization of the  $\rho$ -feature to the syntax from the semantics allows us to capture a wider range of binding phenomena, as compared to Reinhart and Reuland 1993, without referring to logophors. Notice, that no notion of “reflexive predicate” is necessary anymore: all that is needed is a formal  $c$ -command relation between the  $\rho$ -feature on the head and a variable within its domain; there is no need to limit reflexivity to coarguments of a predicate. This solves a major problem of Reinhart and Reuland 1993, that of non-argument anaphors. In what follows I outline the formal side of the story.

As I mentioned above, the  $\rho$ -feature occurs on a head and must  $c$ -command a feature in its domain. Formally, in the Pesetsky and Torrego’s 2007 framework, this feature on the head must be interpretable, as it works at the interface by establishing binding relations. I take valuation of the  $\rho$ -feature in the current framework to be reminiscent of marking a predicate reflexive in Reinhart and Reuland’s 1993 framework. Therefore, unless the verb is inherently reflexive, or has reflexivity marked on it in a morphological way, I assume that  $\rho$  is unvalued on  $v/V$ . Thus, in most Germanic / Romance / East Asian languages when there is no reflexivity marking on the verb,  $V$  comes from the lexicon with an unvalued  $\rho$ -feature. On the other hand, when the verb is marked as reflexive, for instance in case of Salishan / Iroquoian / South Indian, and certain Slavic languages,  $V$  comes from the lexicon with a valued  $\rho$ -feature. Therefore, whether the verb comes with a valued or unvalued  $\rho$ -feature depends on the verb’s morphology, and partially on the semantics. Notice, that in some languages, e.g. in English, verbs never come endowed with a valued  $\rho$ -feature. Therefore, in a feature-sharing framework, the specification of this feature on the verb is either  $\langle i\rho -val \rangle$  or  $\langle i\rho +val \rangle$ , depending on the language and the semantic properties of the verb.

For Reinhart and Reuland 1993, predicates become reflexive by means of having a reflexive pronoun as one of their arguments. In the framework proposed here, it therefore makes sense to postulate that reflexives also come with a valued  $\rho$ -feature. Further, following the logic of Pesetsky and Torrego 2007, this feature is uninterpretable on reflexives, as reflexivity in a proposed framework is a property of a head. Therefore, this feature on reflexive pronouns and reciprocals is specified as  $\langle u\rho +val \rangle$ . Notice, that as opposed to reflexive pronouns, R-expressions do not have a

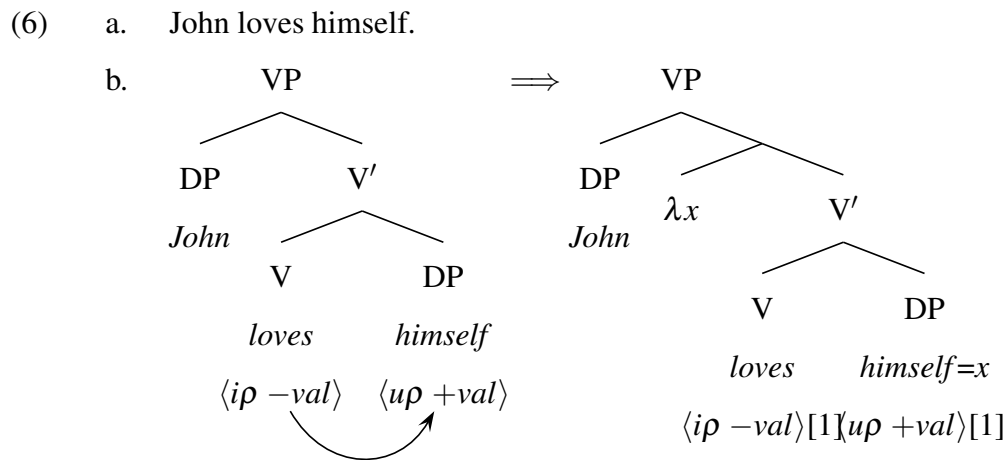
$\rho$ -feature, and therefore, if there is no reflexive with a  $\rho$ -feature in the domain of  $v/V$ ,  $\langle \rho \rangle$  will not be able to get valued, and the derivation will crash; this statement can be taken as a reformulation of Condition A by Reinhart and Reuland 1993.

(5) summarizes the specifications of  $\rho$ -feature on various expressions.

- (5) Distribution of  $\langle \rho \rangle$ :
- on V: *interpretable*  $\langle i\rho \pm val \rangle$
  - on reflexives: *valued*  $\langle u\rho + val \rangle$
  - on R-expressions: *absent*

Now, the question is what the role of the  $\rho$ -feature is, and how it is interpreted at the interface. I argue that an interpretable instance of the  $\rho$ -feature introduces a  $\lambda$ -operator immediately above the position where the  $\rho$ -feature is present. This  $\lambda$ -operator binds the variable located at the position of the uninterpretable  $\rho$ -feature. Below I outline this proposal in detail.

Assume an example in which the  $\rho$ -feature is located on a V-head, and the corresponding valued  $\rho$ -feature is located on the direct object, (6).



If valuation of the  $\rho$ -feature on V occurs after V probed a reflexive, and the  $\rho$ -feature on V and the reflexive became shared, the  $\lambda$ -operator can be introduced at the V position, as shown above. At the moment VP is interpreted,  $\lambda$ -conversion can apply, and the expression located in the Spec,VP

will be substituted for a variable.

- (7) a.  $\exists e \text{ Agent}(e, \text{John}) \text{ John } \lambda x \text{ love}(e) \ \& \ \text{Theme}(e, x)$   
b.  $\exists e \text{ Agent}(e, \text{John}) \text{ love}(e) \ \& \ \text{Theme}(e, \text{John})$

Principle A can now be stated in terms of the  $\rho$ -feature, and is now a requirement that the  $\rho$ -feature must be valued, and there must be an interpretable instance of the  $\rho$ -feature. It is now just a standard requirement on features in general within Pesetsky and Torrego's 2007 framework. Notice that this proposal is similar to the proposal presented in Johnson 2007 about Principle A. However, my proposal is more syntactic in nature; I also take it further and generalize it to the case of monomorphemic anaphors in order to account for their subject-orientation. In Chapter 4 I extend this proposal even further by showing that long-distance binding can also be derived from this principle, if one assumes the theory of phasal domains outlined in Chapter 2. I claim that binding relations can only be evaluated at the moment the domain becomes *phasal* as defined in (63). The necessity of this condition will be seen in my consideration of the double object constructions in languages with complex vs. monomorphemic anaphors. The fact that the VP is not a phasal domain if one of the objects of a sentence is a monomorphemic reflexive prevents binding from applying at the level of VP, therefore preventing the monomorphemic reflexive from being bound by the object. In this case, the first moment the binding can apply is when the  $vP$  is completed and T is merged, and therefore the only possible antecedent of the object monomorphemic reflexive is the subject.

It is crucial to state that the  $\rho$ -feature introduced above is not a direct syntactic counterpart of the notion of *reflexivity* of Reinhart and Reuland 1993. For Reinhart and Reuland, the predicate is reflexive if its arguments are coindexed. No similar requirement is made here. In fact, if one allows Agree relation to be non-clause-bound (later in Chapter 4 I show that this is possible in case of certain defective domains, such as infinitives and subjunctives), the process of coindexation/ $\lambda$ -conversion obtained by the introduction of the  $\rho$ -feature can target elements which are located in different clauses, resulting in the possibility of long-distance binding. The notion of reflexivity

of Reinhart and Reuland 1993 is only a special case of the proposed framework involving the  $\rho$ -feature. While the presence of the  $\rho$ -feature may result in coindexation of the arguments of a predicate, it does not have to.

Further in this chapter I will demonstrate how introduction of the  $\rho$ -feature can account for various binding phenomena in language, such as binding in ditransitive constructions and subject orientation of monomorphemic reflexives, and in Chapter 4 I will extend the theory of binding to capture long-distance binding dependencies, all areas of the binding theory where Reinhart and Reuland's 1993's account fails to provide an explanation, and phase based theories of binding cannot handle, as we saw in Chapter 1.

## 3.2 Deriving binding relations: Principle A

### 3.2.1 Complex anaphors

This section deals with establishing binding relations between a complex anaphor, of the type observed in English (*himself*), and its antecedent. The main difference between complex and monomorphemic anaphors lies in their featural content. Monomorphemic reflexives are unspecified for  $\phi$ -features, while complex ones, such as English *himself*, have valued  $\phi$ -features,  $\langle i\phi +val \rangle$ . In this section I will show the derivation of a simple English transitive sentence with a complex reflexive pronoun. I will also demonstrate that as opposed to monomorphemic reflexives, the final structure predicts that the  $\rho$ -feature can be interpreted lower than  $\nu$ -position. In the next section I will show how this fact can be used to account for the differences between potential antecedents of the reflexive pronoun. It is well-known that complex anaphors are not subject-oriented, and can be bound not only by the subject of the sentence, but also by a direct or indirect object; I show the relevant data from English below (but see Barss and Lasnik 1986 for initial observations), such as examples in (8):

- (8) a. John<sub>i</sub> told Bill<sub>j</sub> about himself<sub>i/j</sub>.

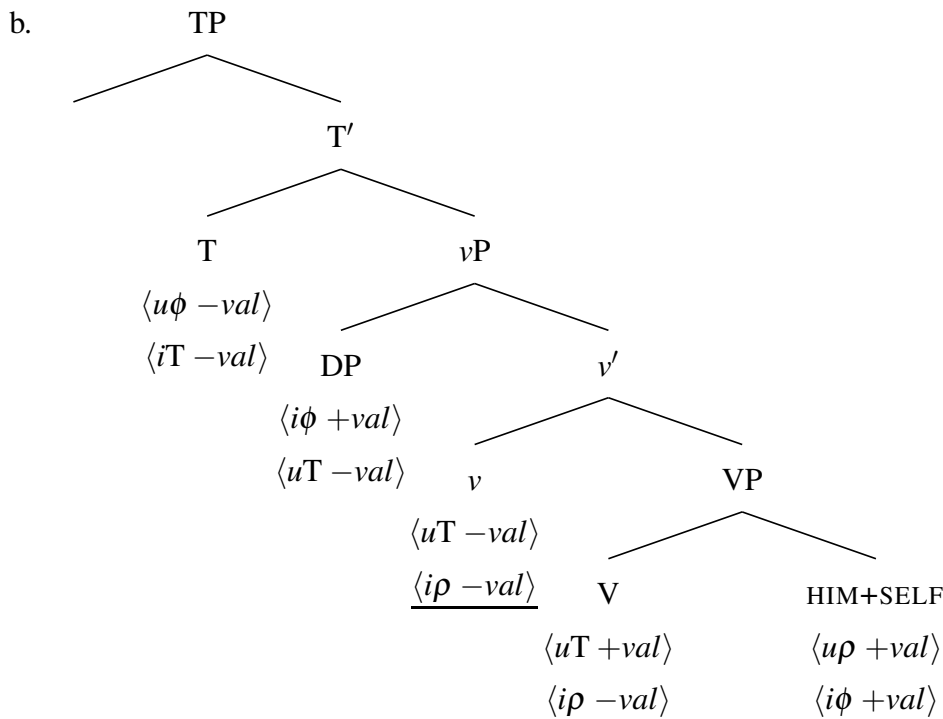
- b. John<sub>i</sub> showed Bill<sub>j</sub> to himself<sub>i/j</sub>

### 3.2.1.1 Transitive sentences

In this section I show how the derivation of a transitive sentence proceeds in a language like English, where reflexive pronouns such as *himself*, *herself* are endowed with a full set of  $\phi$ -features. The basic assumptions about featural content of the major elements are given below in (9). Notice that there are two possibilities for placement of the reflexive  $\rho$ -feature: either on V, or on  $v$ ; they are underlined in (9): only one of them will actually be present when the derivation happens.

(9) *Basic structure and feature content for a transitive sentence with a complex reflexive.*

- a. John loves himself.

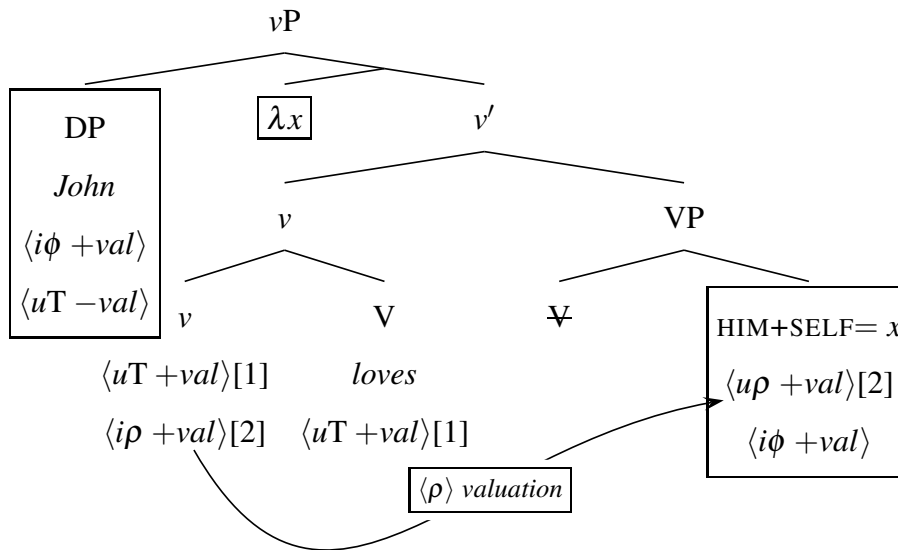


Now, let's consider the full derivation in detail, assuming first that the  $\rho$ -feature is placed on  $v$ . First, the verb is merged with a direct object DP, and after that  $v$  is merged into the structure.  $v$  has an unvalued T-feature, and therefore it probes V, finding its valued T-feature, and enters into



an Agree relation, resulting in T-feature sharing, which is indicated in (10) by the index [1]. As  $v$  also has an unvalued  $\rho$ -feature, it continues to probe, and this time finds its goal in the reflexive pronoun, which has a valued  $\rho$ -feature. Again, this results in sharing of a  $\rho$ -feature between  $v$  and the reflexive HIM+SELF; this is indicated by [2] in (10). Further, the verb (universally) raises to  $v$ . The subject DP is further merged into Spec, $v$ P. The structure after this step is shown in (10). Notice, that after  $v$  probes V, VP acquires phasal status, as all (relevant) features within it are valued, or shared with some element outside VP. Therefore, it can be converted into its semantic form, and shipped to interpretation. The corresponding  $\lambda$ -expression is given in (11). As the Spec, $v$ P is empty (or non-existent),  $\lambda$ -conversion is unavailable until the subject DP is merged.

(10)  $v$ P structure before merge of T.



(11)  $\lambda x \exists e : V(e) \ \& \ \text{Theme}(e, x)$

Now, T is merged with an unvalued T-feature and an unvalued  $\phi$ -feature. It probes the subject DP: T's  $\phi$ -features get valued, and the T-feature on T and the T-feature on the subject get shared. They are still however unvalued. T continues to probe, and finds its target in the  $v+V$  complex. This probing results in simultaneous valuation of the T-feature on T and on the subject DP, as they are

shared and are in fact instances of the same feature (Pesetsky and Torrego 2007). Now,  $\nu$ P gets phasal status. The external argument is introduced, and  $\lambda$ -conversion results in reflexive reading, where both  $\text{Agent}(e)$  and  $\text{Theme}(e)$  of the event described by the verb are the same:

- (12) a.  $\exists e : \text{Agent}(e, \text{Subj}) \ \& \ \text{Subj} \ \lambda x \ V(e) \ \& \ \text{Theme}(e, x)$   
 b.  $\exists e : \text{Agent}(e, \text{Subj}) \ \& \ V(e) \ \& \ \text{Theme}(e, \text{Subj})$

Note that this derivation will only minimally differ if one chooses to place the  $\rho$ -feature on V. In fact,  $\lambda$ -conversion will lead to the same result, where the only possible antecedent of the reflexive is the subject DP.

### 3.2.1.2 Ditransitives and Barss-Lasnik effects

In this section I will demonstrate how binding relations are achieved for ditransitive verbs, when the reflexive pronoun is a part of either the direct or indirect object. The original description of the binding relations in ditransitive constructions is presented in Barss and Lasnik 1986, based on the facts in (13).

- (13) a. I showed John himself.  
 b. \*I showed himself John.

They observe that while Condition A is not violated in (13-a) where the reflexive pronoun follows the R-expression, and the example is grammatical, a similar example (13-b), where the reflexive precedes the R-expression, is ungrammatical.<sup>1</sup> That leads Barss and Lasnik to propose that at

---

<sup>1</sup>While this can be attributed to a problem with Condition C, as *himself* *c*-commands *John*, they show that if the anaphor is embedded inside the DP, similar judgements hold. They provide examples such as (i).

- (i) a. I showed the professors [clones of each other].  
 b. \*I showed [clones of each other] the professors.

Similarly, example (i-a) is grammatical; however, it is clear that since the reciprocal *each other* is embedded within the indirect object DP, the ungrammaticality of (i-b) is clearly not due to a Condition C violation, but due to the fact that the reciprocal is not *c*-commanded by its antecedent, i.e. due to violation of Condition A.

surface structure, the second DP in double object construction is invariably asymmetrically *c*-commanded by the first DP.

Note, that similar facts obtain when one considers a ditransitive verb with a PP argument:

- (14) a. John<sub>i</sub> showed Bill<sub>j</sub> [to himself<sub>i/j</sub>].  
 b. John<sub>i</sub> showed himself<sub>i/\*j</sub> [to Bill<sub>j</sub>].

The first argument *c*-commands the second, and therefore the examples above differ. Example (14-a) is ambiguous: the reflexive *himself* can be bound by either subject or direct object, giving two possible readings. On the other hand, example (14-b) is not ambiguous: the only possible antecedent of *himself* is *John*.

Let us first demonstrate how featural binding can account for the two possible readings of (14-a). I claim that when the  $\rho$ -feature is placed on V, the direct object will serve as a binder for the reflexive. On the contrary, if the  $\rho$ -feature is present on *v*, then the subject DP will serve as the antecedent. Let us consider these derivations in detail.

First, let's assume that the  $\rho$ -feature is present on V. At the initial stage, the derivation proceeds as usual. The verb is merged with the PP *to himself*. Since it has an unvalued  $\rho$ -feature, it probes the reflexive with a valued  $\rho$ -feature: this results in  $\rho$ -feature sharing and its valuation on the verb; this is indicated by [1] in the structure below. Further, the direct object is merged into Spec,VP. The resulting structure is shown below in (15).<sup>2</sup>

---

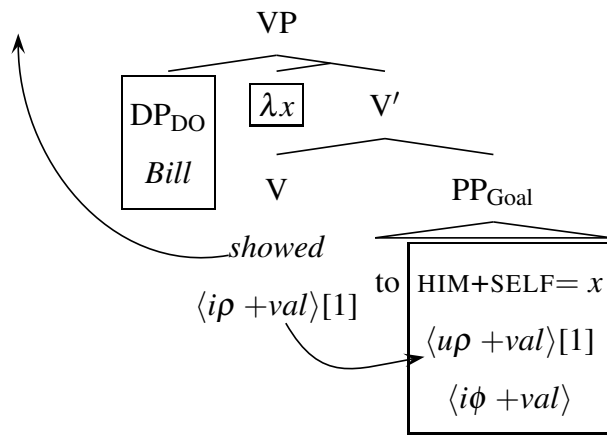
<sup>2</sup>Note that there are apparent violations of a *c*-command condition on binding in examples like (i):

- (i) John talked to Mary<sub>i</sub> about herself<sub>i</sub>.

The approach to binding proposed in this chapter is able to account for this phenomenon by assuming that *to* is semantically pleonastic and the phrase *to Mary* is interpreted simply as *Mary*. The corresponding logical form of (i) before  $\lambda$ -conversion is given in (ii):

- (ii)  $\exists e : \text{Goal}(e, \text{Mary}) \ \& \ \text{Mary} \ \lambda x \text{ talking}(e) \ \& \ \text{Theme}(e, x)$

(15) VP structure of (14-a), object antecedent.

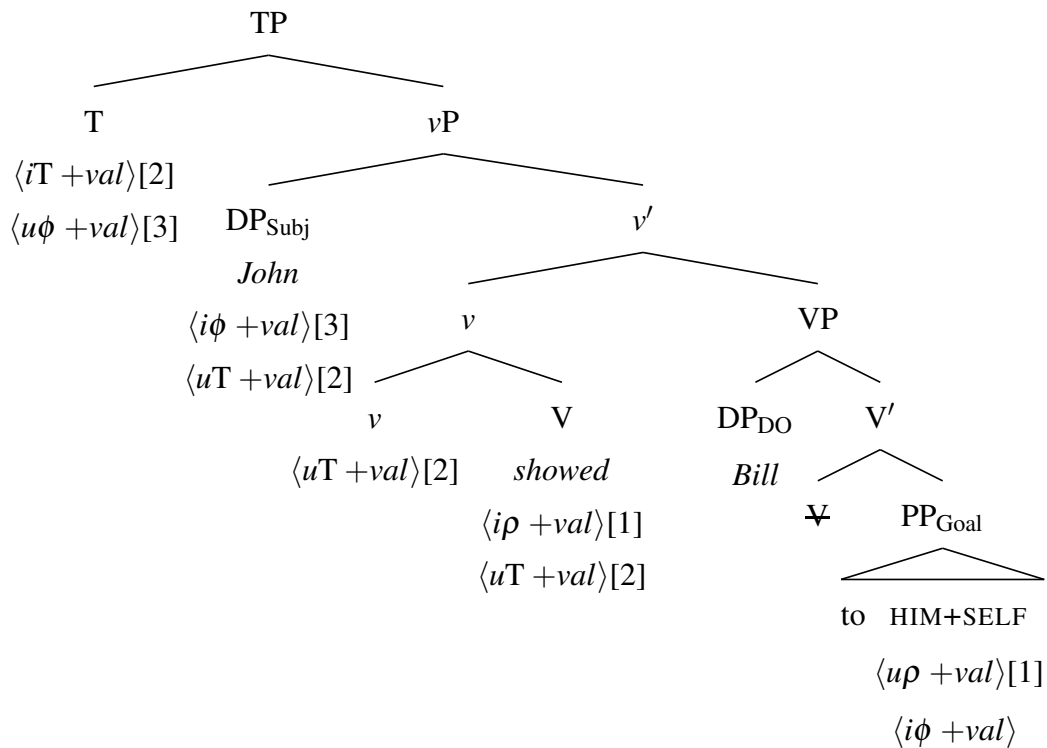


Now, let's consider what happens when  $v$  is merged with VP. As before, it comes with an unvalued uninterpretable T-feature, and therefore probes V, which results in T-feature sharing. Similarly to the previous derivation, VP acquires phasal status. Further, since the  $\rho$ -feature is present on V, the  $\lambda$ -operator is inserted at V position, and the corresponding  $\lambda$ -conversion will result in a reading when the reflexive within the PP is bound by the direct object.

- (16) a.  $\exists e : \text{Theme}(e, \text{DO}) \ \& \ \text{DO} \ \lambda x \ V(e) \ \& \ \text{Goal}(e, x)$   
 b.  $\exists e : \text{Theme}(e, \text{DO}) \ \& \ V(e) \ \& \ \text{Goal}(e, \text{DO})$

After this step, the derivation proceeds in the usual way. V is raised to  $v$ , the subject is merged into Spec, $v$ P. T is further merged with the  $v$ P, and probes first subject DP, and second,  $v$ . As usual, these two probing relations establish T-feature sharing between T, subject DP, and  $v+V$ , and result in valuation of a T-feature on T and the subject DP. The structure of  $v$ P is shown in (17). Following this step,  $v$ P acquires phrasal status, and the subject DP is moved to Spec,TP position. The final logical form (where the reflexive within the PP is bound by the direct object) is provided in (18).

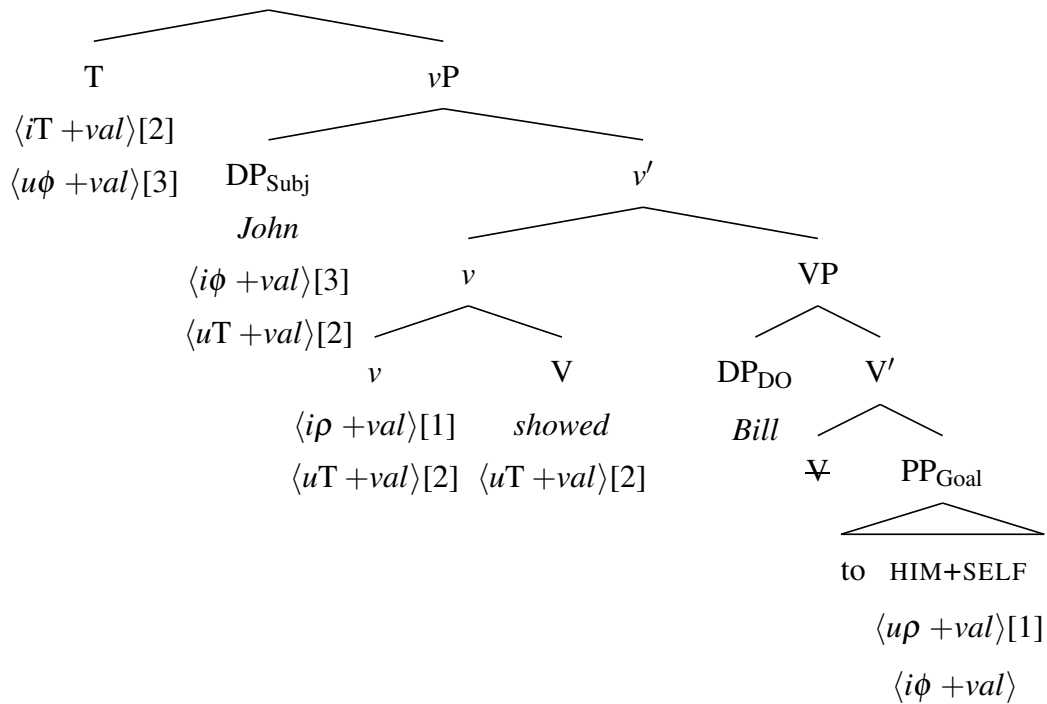
(17) *vP structure of (14-a) after probing by T.*



(18)  $\exists e : \text{Agent}(e, \text{Subj}) \ \& \ \text{Theme}(e, \text{DO}) \ \& \ V(e) \ \& \ \text{Goal}(e, \text{DO})$

Now let us consider a second possibility: placing the  $\rho$ -feature on  $v$ . In this case, VP is constructed in the same way as before, with one difference: the verb does not probe the reflexive within the goal PP, as it does not have an unvalued  $\rho$ -feature. Nevertheless, VP still acquires phrasal status after  $v$  is merged. When  $v$  is first merged with VP, it first probes V in order to value its T-feature, and further it probes the reflexive within the goal PP, in order to value its  $\rho$ -feature. At this stage, VP acquires phrasal status and can be closed. While the resulting structure does not differ much from the one above, the initial placement of the  $\rho$ -feature on  $v$  as opposed to V will result in a different semantic interpretation. The insertion of a  $\lambda$ -operator will not take place until after the  $vP$  phase is closed, i.e. after the merge of T. As a result,  $\lambda$ -conversion will result in the reflexive being bound by the subject DP in Spec, $vP$ , and not by the direct object in Spec,VP position.

(19) *vP structure of (14-a) after probing by T.*



- (20) a.  $\exists e : \text{Agent}(e, \text{Subj}) \ \& \ \text{Subj} \ \lambda x \ \text{Theme}(e, \text{DO}) \ \& \ V(e) \ \& \ \text{Goal}(e, x)$   
 b.  $\exists e : \text{Agent}(e, \text{Subj}) \ \& \ \lambda x \ \text{Theme}(e, \text{DO}) \ \& \ V(e) \ \& \ \text{Goal}(e, \text{Subj})$

Therefore, the location of the  $\rho$ -feature derives each of two possible readings on a sentence with a PP goal and a direct object. If the  $\rho$ -feature is placed on the  $v$ -head, the reflexive within the PP goal will be bound by the subject, while if the  $\rho$ -feature is placed on the V-head, the direct object will serve as the antecedent.

Another consequence of this analysis is that if the theme argument is reflexive, the only potential binder for it is the subject:

- (21) a.  $\text{John}_i \text{ showed himself}_i \text{ to Mary}$   
 b.  $\text{John}_i \text{ showed himself}_{i/*j} \text{ to Bill}_j$

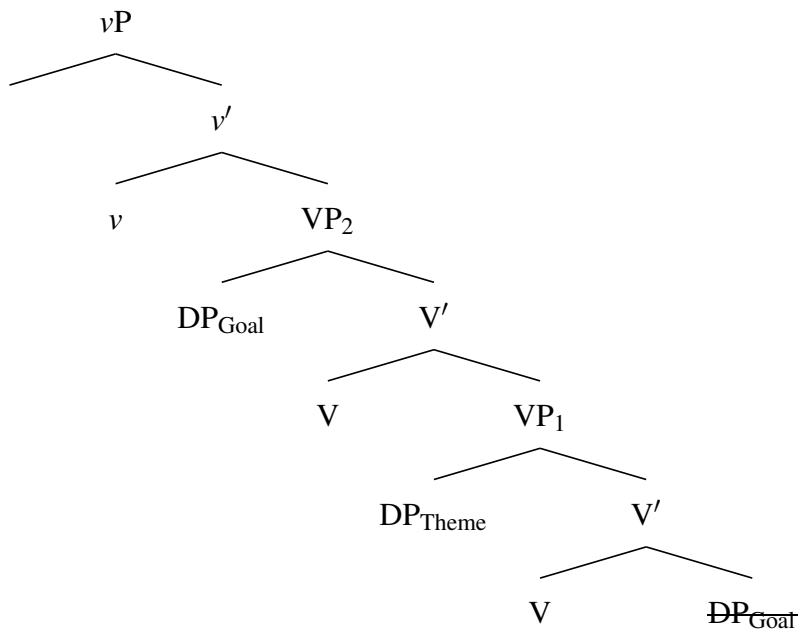
Placement of the  $\rho$ -feature on  $v$  will trigger  $\lambda$ -conversion at the  $vP$  level, and therefore the only possible antecedent for the reflexive is the subject DP.

Now, let us consider double object constructions in English, where the goal DP occupies a position before the theme, such as the sentences in (22).

(22) I showed John Mary.

In section 2.5.4.4 I argued for the following structure of the double object constructions, and demonstrated that in a simple ditransitive sentence, only the higher  $VP_2$  is a phasal domain, while the lower  $VP_1$  is not.

(23) *The structure of double-object constructions*



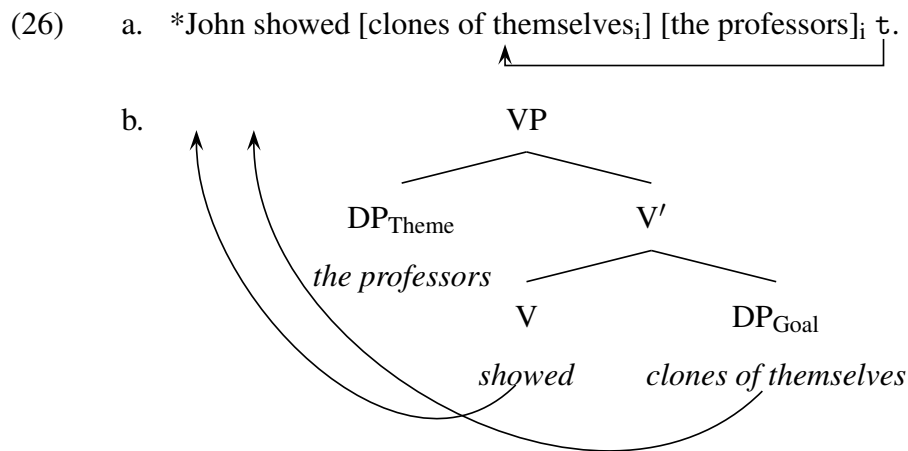
One of the major problem for the derivational theory of binding (discussed in section 1.1.5 based on Belletti and Rizzi 1988) comes from consideration of the sentences like the ones given in (24) and (25):<sup>3</sup>

---

<sup>3</sup>If the subject is a 1st person singular, it cannot serve as a possible antecedent of the reflexive because of  $\phi$ -feature mismatch, therefore in what follows I will concern myself with the possibility of a direct object being bound by a goal

- (24) a. I showed John himself.  
 b. \*I showed himself John.
- (25) a. I showed the professors [clones of themselves].  
 b. \*I showed [clones of themselves] the professors.

Consider how the derivation of the sentence in (25-b) starts. Assuming the hypothesis about low generation of goals according to the thematic hierarchy in (76) (Baker 1988; Carrier-Duncan 1985; Larson 1988), as outlined above, one can see that in this sentence the goal containing *themselves* is originated as a complement of the verb, and the theme, *the professors*, in Spec,VP. Therefore, under a strong derivational binding theory, which assumes that binding can occur at any point during the derivation, the binding relation between *the professors* and *clones of themselves* can be established at the level of VP, (26), before further movement of the goal happens.



Further dislocations of the goal should not make the sentence ungrammatical. However, it is not so: sentence (25-b) is ungrammatical. In what follows, I show that the binding theory as defined above makes correct predictions about the status of sentences in (25).

Let us consider the derivation of the sentence in (25-b) in the framework of this dissertation. Consider a case when the  $\rho$ -feature is present on V. At the beginning, V is merged with the goal

---

DP.



DP, which has an unvalued Case feature,  $\langle u\text{Case} -val \rangle$ , and a valued uninterpretable  $\rho$ -feature. V probes the goal DP in order to value its  $\rho$ -feature. The theme DP with a valued case feature,  $\langle u\text{Case} +val \rangle$  is then merged into Spec,VP position. At this moment, VP does not yet have phasal status: first,  $v$  is not merged into structure yet, and further the case feature on the goal DP is still not valued. Therefore, binding does not apply at this stage, and at this moment binding of the goal DP by the theme DP is impossible. This is a crucial point in the derivation where the current predictions diverge from those of strong derivational binding (Belletti and Rizzi 1988 a.o.). While my proposed theory is also derivational in nature, it limits the points in the derivation where binding can apply, while no such limitations are present in the strong derivational binding theory. In order to resolve this issue, I propose the following principle:

- (27) **Application of binding theory:** Binding relations (i.e.  $\lambda$ -conversion) are applied only at the moment the category acquires its phasal status (as defined in (63)), and not at any moment in the derivation.

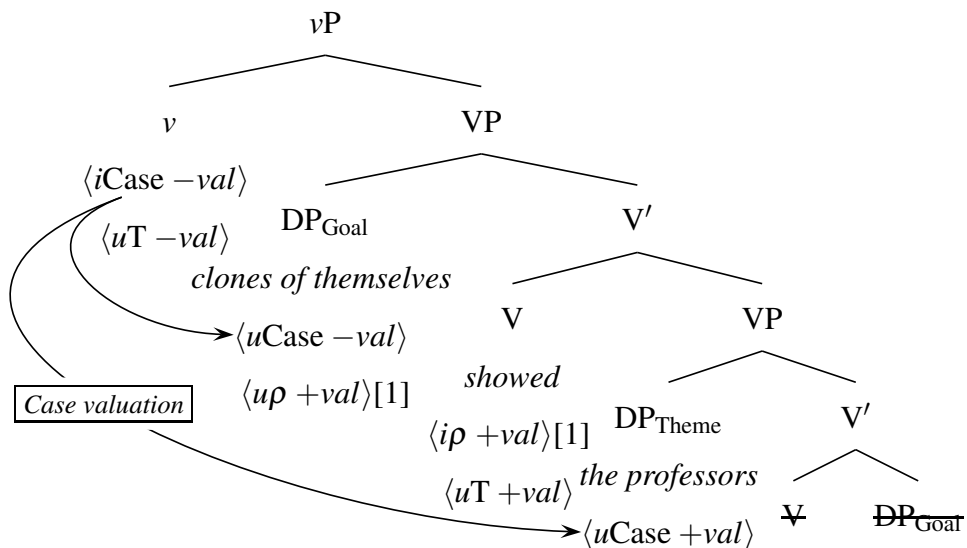
The motivation for the principle in (27) is the following. While feature-valuation happens at the syntactic level grammar,  $\lambda$ -conversion (which as I claimed above is responsible for establishing the binding relations between the anaphor and its antecedent) applies at the interface with semantics. Under assumption that only when the domain acquires phasal status is it sent to the interpretative component, one can see that binding can only apply at the moment the domain is closed, i.e. acquired phasal status.

Consider the consequences of this principle to double-object constructions. Lack of case on the goal DP forces its movement to the edge of VP, and prevents VP from acquiring phasal status early in the course of the derivation.

According to the view on double-object constructions presented above, the goal DP must move to the edge of the higher VP – if it doesn't, it won't be able to have its Case-feature valued. Furthermore, V moves from its lower position to the higher position. After that,  $v$  is merged into the structure. The resulting tree after merge of  $v$  is shown in (28). Notice, that at this time,  $\rho$ -feature

on the reflexive and on V are already shared.

(28) The structure of the *vP* of (25-b).



Now, *v* probes down, and first finds the goal DP; the Case features on *v* and the goal get shared, even though they are not valued yet. The next probing involves *v* as a probe and V as its goal, and results in valuation of T-features. The last probing involves establishing a relation between *v* and the theme DP with a valued Case feature. That probing results in valuation of Case features on both *v* and the goal DP – the goal DP acquires “concordial” accusative case from the theme DP. After the probing happens, the binding theory can finally apply. The only potential position for interpreting the  $\rho$ -feature on V as a  $\lambda$ -operator is above the reflexive variable, i.e. right above the higher VP. The resulting logical form is shown below:

(29)  $\lambda x \exists e \text{ Goal}(e, \text{clones of } x) \ \& \ \text{Theme}(e, \text{the professors}) \ \& \ \text{show}(e).$

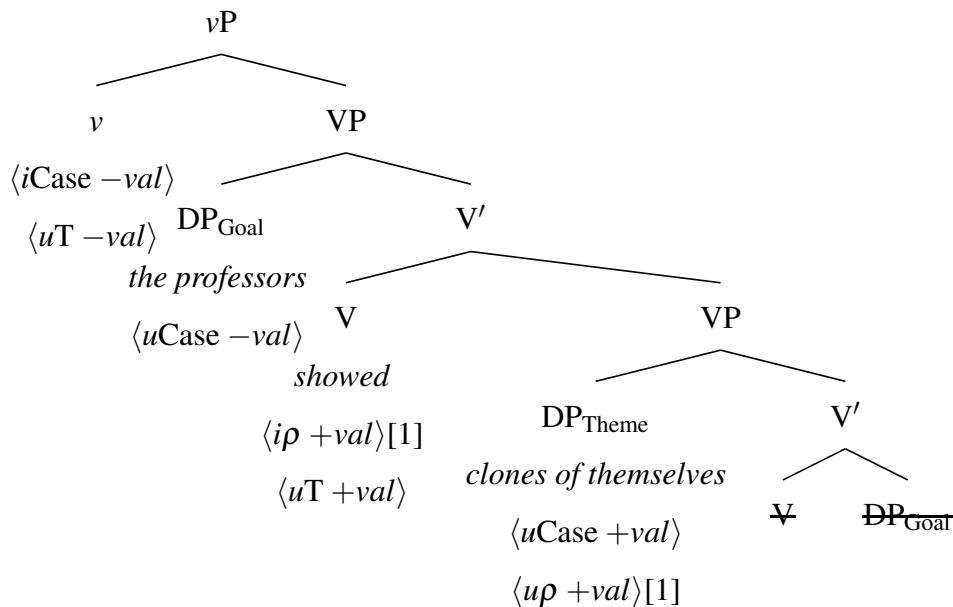
$\lambda$ -conversion cannot apply at this point as the Spec,*vP* position is empty, and nothing can apply to the  $\lambda$ -operator. However, after the subject is inserted in Spec,*vP*, and after T probes first the subject DP, and second the *v*+V complex (to get T-features on T and on the subject DP valued),  $\lambda$ -conversion can apply, and therefore the only potential binder of the goal reflexive is the subject.

The mismatch of  $\phi$ -features of the subject and the goal will make example (25-b) ungrammatical, while a similar example in (30) will have only one interpretation: the reflexive can only be bound by the subject.

(30) [The children]<sub>i</sub> showed [clones of themselves<sub>i/\*j</sub>][the professors]<sub>j</sub>.

Now, the derivation is similar when the theme is reflexive, see example in (25-a). After the higher VP is completed, and  $v$  has finished probing, the higher VP acquires phasal status. The binding theory can apply at this stage, and since the variable is not the highest in the structure,  $\lambda$ -conversion can apply, with the  $\lambda$ -operator inserted in the position directly above the higher V with its  $\rho$ -feature. The goal argument located in the higher Spec,VP will trigger  $\lambda$ -conversion as in (32), and therefore will result in binding of the theme reflexive by the goal argument. The structure is shown in (31).

(31) *The structure of the vP of (25-a).*



- (32) a.  $\exists e \text{ Goal}(e, \text{the professors}) \ \& \ \text{the professors} \ \lambda x \text{ show}(e) \ \& \ \text{Theme}(e, \text{clones of } x)$   
 b.  $\exists e \text{ Goal}(e, \text{the professors}) \ \& \ \text{show}(e) \ \& \ \text{Theme}(e, \text{clones of the professors})$

### 3.2.2 Monomorphemic anaphors

In this section, I will consider how the derivation of a simple transitive sentence with a monomorphemic reflexive is done. For the purposes of simplicity, I will denote the reflexive anaphor as SE. The situation described in this section is representative of languages such as Russian (and other Slavic languages), and a majority of Germanic languages. The major difference between this case and the case described in the previous section is *the lack of valued  $\phi$ -features on the reflexive*: thus, the main questions which have to be answered are how these features are valued, and how this affects the binding possibilities? We will see that the system correctly derives the subject-orientation of monomorphemic anaphors.

#### 3.2.2.1 Transitive sentences

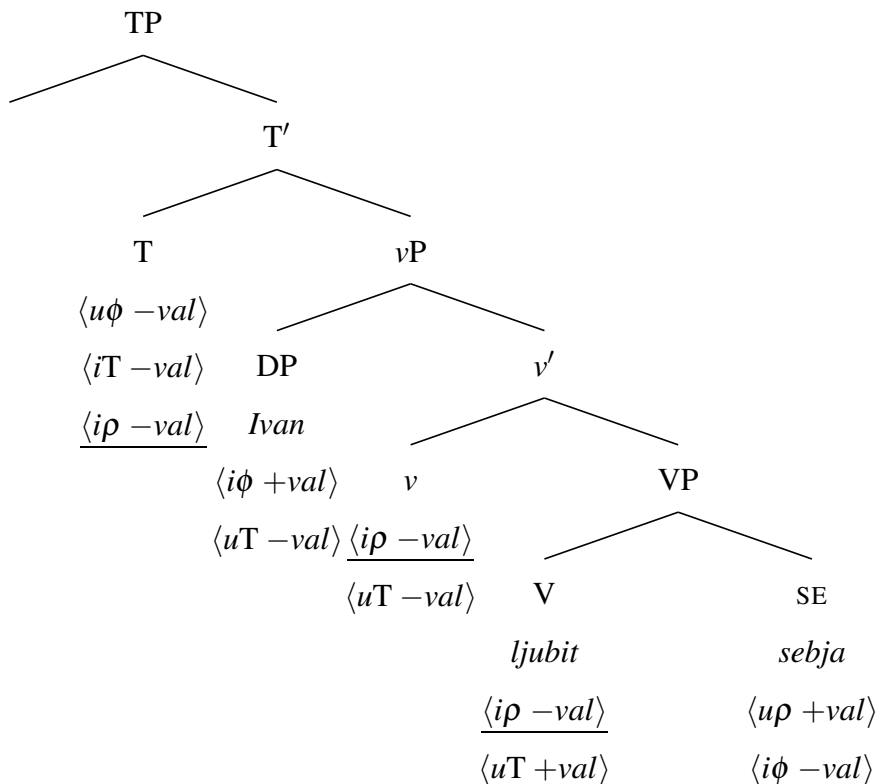
In what follows in this section, I will explore the derivations of sentences in Dutch and Russian with monomorphemic anaphors, which regardless of the language I will be glossing as SE. Dutch is known to have a monomorphemic anaphor *zich*, which is known to be subject-oriented, and is not specified for gender/number. Similarly, Russian monomorphemic reflexive *sebjja* is also unspecified for gender/number, and shares the subject-orientation property with its Dutch counterpart. The examples for basic transitive sentences in Russian and Dutch are given below in (33).

- (33) a. Ringo<sub>i</sub> scheert zich<sub>i</sub> (Dutch)  
Ringo<sub>i</sub> shaves SE<sub>i</sub>  
'Ringo shaves himself.' (Lidz 2001)
- b. Ivan<sub>i</sub> ljubit sebjja<sub>i</sub> (Russian)  
Ivan<sub>i</sub> loves SE<sub>i</sub>  
'Ivan loves himself'

Since these reflexives are not specified for gender/number, I assume that  $\phi$ -features are not valued on the reflexive, but have to be interpreted. Formally that means that they come from the numeration with  $\langle i\phi -val \rangle$  feature specification. Further, being reflexives, these elements must come with a  $\rho$ -feature, which as I argued above is universally uninterpretable on DPs/NPs. Notice,

that while uninterpretable on reflexives themselves, this feature is interpretable on the verb, as being reflexive is a property of a predicate. Further, I assume that this feature is valued on the reflexive. Therefore, the formal specification for this feature is  $\langle u\rho +val \rangle$ .

(34) *The initial structure of a simple transitive sentence with a monomorphemic reflexive direct object, (33-b).*

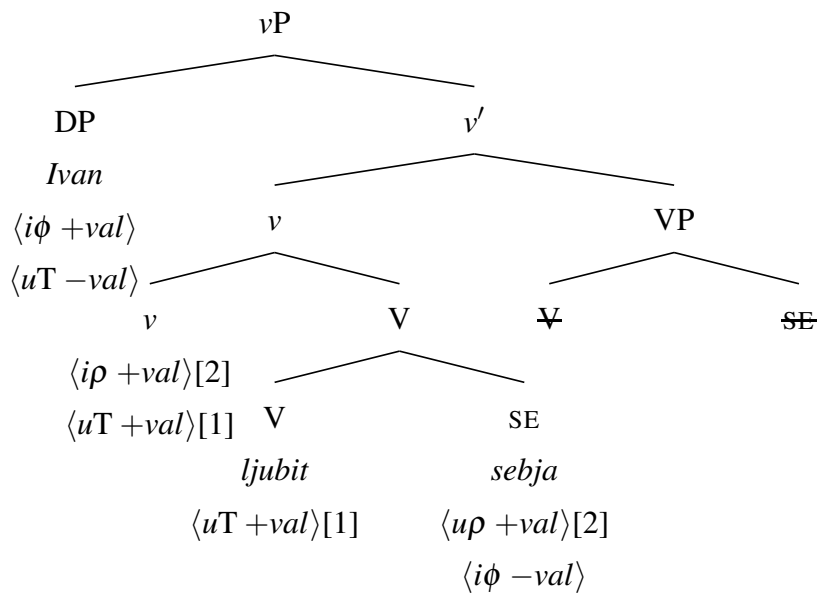


The basic feature specifications on all elements in the structure is given in (34). Notice, that as before, the  $\rho$ -feature can be placed on either V,  $v$ , or T, the corresponding instances of this feature are underlined. For the sake of simplicity, in the further derivation I assume as before that it is placed on  $v$ .

Next, I will show how the derivation proceeds, and how one gets anaphora bound only by the subject.

First, the verb merges with the reflexive. No probing occurs at this moment as there are no common feature between V and SE. Further,  $v$  is merged, and since its T-feature is unvalued, it probes down to V; this probing results in valuation of the T-feature on  $v$ , [1]. At the same time,  $v$ 's unvalued  $\rho$ -feature finds its match in the  $\rho$ -feature of the anaphor, and gets valued, which is indicated by [2] in (35). Following this probing, the V+SE complex is raised to  $v$ . The subject is further merged in the Spec, $v$ P position. Notice that after probing by  $v$ , VP does not acquire phasal status, as opposed to the derivation for English. This is because the presence of an unvalued  $\phi$ -feature on SE which has neither been shared nor valued by any other element prevents VP from being closed.

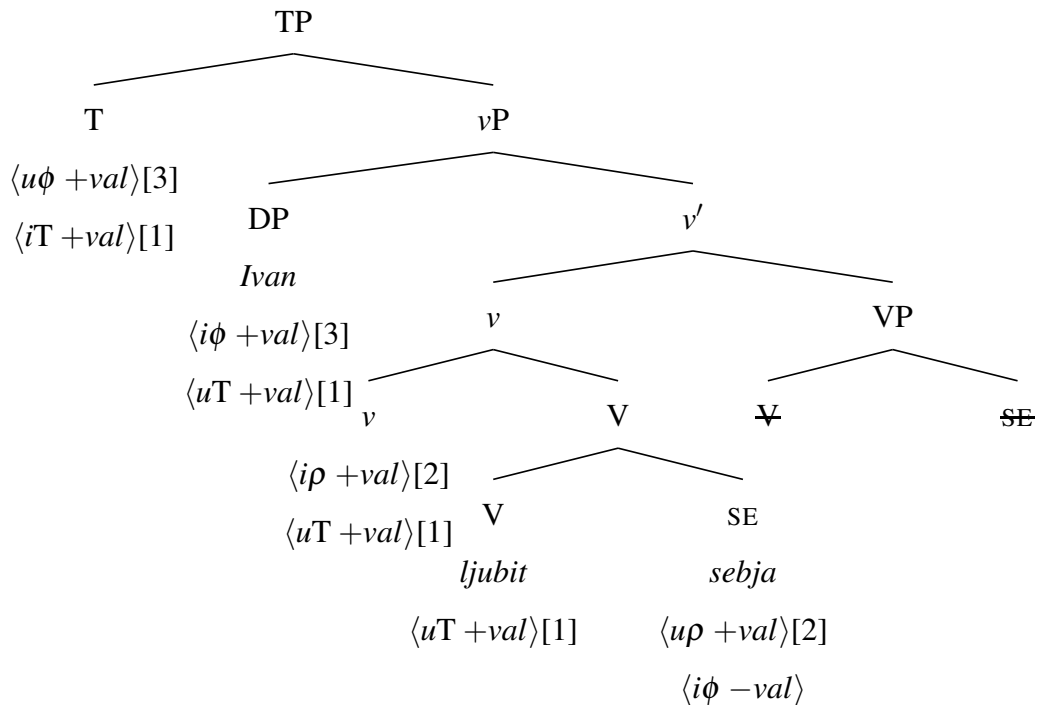
(35)  $v$ P structure of (33-b) before merge of T.



Now, T is merged with the following features:  $\langle u\phi -val \rangle$  and  $\langle iT -val \rangle$ . Both of its features are unvalued, and T probes first finding a subject DP in Spec, $v$ P position. Its  $\phi$ -features get valued, and its T-feature gets shared with the subject. As it is still unvalued, T continues to probe, and finds the  $v$ +V+SE complex, which has a valued T-feature. At this moment, a valuation of T-feature occurs not only on T, but also on subject DP, as its T-feature was shared with the T-feature on T.

That is indicated by index [1] on T's and the subject's T-features. Notice that the  $\phi$ -feature on the reflexive cannot get valued through this probing. T only probes  $v$ , and not SE, even though it is located within the same head-complex.

(36) *Final structure of (33-b) after probing by T.*



The structure above shows the moment after T has probed  $v$ . After this probing the subject DP raises to Spec,TP. Now, notice that all edge features of TP (*TP-relevant features*) are valued: the  $\phi$ -feature on T got its value from the subject, the T-feature on the subject and T are valued from the verb. That means that TP acquires phasal status and will be closed for further probing; no higher probe will be able to supply a value to the reflexive. The derivation will crash.

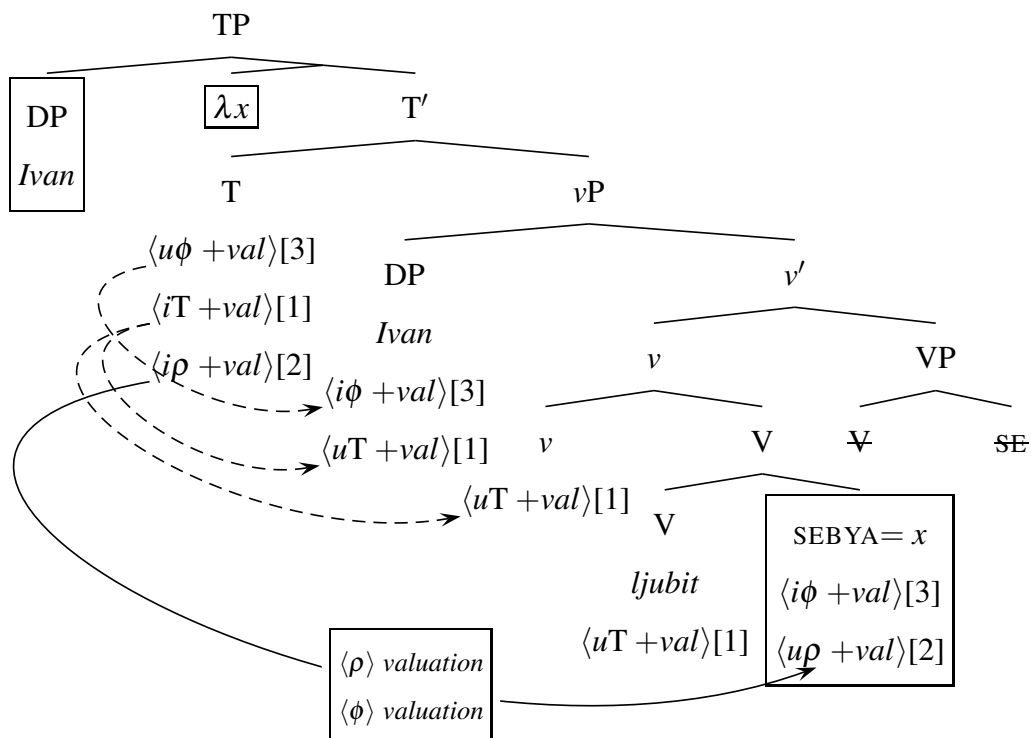
How can the  $\phi$ -features on the reflexive get valued? Let us reconsider the derivation above. It is clear that the  $\rho$ -feature cannot be placed on either  $v$  or V: that leads to derivation crash, as I showed above. The next possibility is to place the  $\rho$ -feature on T. Consider what happens in this case. Up to the moment T probes  $v$  the derivation is parallel to the one discussed above (excluding the  $\rho$ -

feature on  $v$  probing SE — now we assume that  $\rho$ -feature is placed on T). However, after T probes  $v$ , one more probing will happen. T still has an unvalued interpretable  $\rho$ -feature which will enter in a probe-goal relationship with the reflexive: the  $\rho$ -feature on T will acquire its value. However, at the same time, as a byproduct of this probing, the  $\phi$ -feature on T will enter into sharing relation with the  $\phi$ -feature on SE. As a result, the  $\phi$ -feature on SE will get its value from the  $\phi$ -feature on T, i.e. the  $\phi$ -feature of the reflexive will have the same value as the  $\phi$ -feature of the subject. The structure is given in (37). The subject DP is further raised into Spec,TP.

At this moment, all edge features within TP are valued, and TP will be closed for further probing. However in this case the derivation will not crash: the  $\phi$ -features on the reflexive are valued, and therefore TP can be sent to the semantic component. At the interface, a  $\lambda$ -operator is introduced at the position of T, and since the reflexive is in its domain, after  $\lambda$ -conversion, it will end up being bound by the subject DP, as it is the only nominal expression directly preceding T. Similar to English transitive constructions, the logical form of TP before and after  $\lambda$ -conversion takes place is given in (38).



(37) Final structure of (33-b) after probing by T.



- (38) a.  $\exists e : \text{Agent}(e, \text{Subj}) \ \& \ \text{Subj} \ \lambda x \ V(e) \ \& \ \text{Theme}(e, x)$   
 b.  $\exists e : \text{Agent}(e, \text{Subj}) \ \& \ V(e) \ \& \ \text{Theme}(e, \text{Subj})$

Clearly, transitive sentences are non-ambiguous, with the subject being the only possible binder for a reflexive within a direct object, and there is no difference between the semantic interpretation of transitive sentences with complex or monomorphemic reflexives. Below, I examine ditransitive constructions in languages with monomorphemic reflexives, and show that in that case the interpretation is crucially different, and only subject orientation is possible.

### 3.2.2.2 Ditransitive constructions and subject orientation

As I mentioned above, monomorphemic reflexives are subject-oriented. Therefore, in double-object constructions or in constructions with a PP goal, a reflexive located within a lower DP/PP still has only one possible binder — the subject DP. This is illustrated in (39) below.

(39) *Double-Object constructions*

- a. Marina<sub>i</sub> otdala Petra<sub>j</sub> sebe<sub>i/\*j</sub> (Russian)  
Marina gave Peter<sub>ACC</sub> SE<sub>DAT</sub>  
'Marina gave herself Peter'
- b. Ivan<sub>i</sub> rasskazal Petru<sub>j</sub> o sebe<sub>i/\*j</sub>  
Ivan told Peter<sub>DAT</sub> about SE  
'Ivan told Peter about himself (=Ivan)'
- c. Marina<sub>i</sub> otdala Petru<sub>j</sub> sebja<sub>i/\*j</sub>  
Marina gave Peter<sub>DAT</sub> SE<sub>ACC</sub>  
'Marina gave Peter herself'
- d. Jan<sub>i</sub> raadde Peter<sub>j</sub> [ de vrouw naast zich<sub>i/\*j</sub> ] aan (Dutch)  
Jan recommended Peter the woman near SE PRT  
'Jan<sub>i</sub> recommended the woman near him<sub>i/\*j</sub> to Peter<sub>j</sub>'
- e. Bill gave John himself.

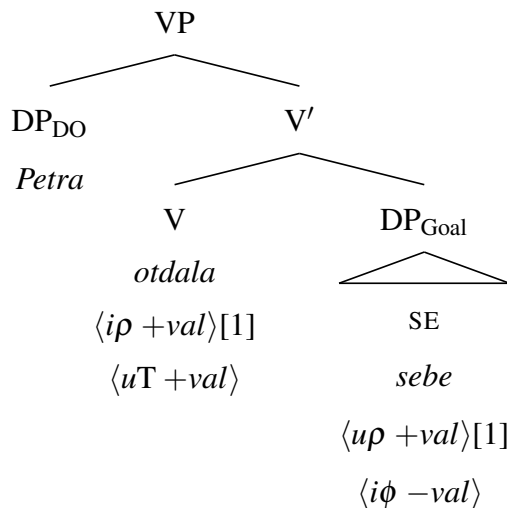
For instance, in the Dutch sentence in (39-d), the SE-element *zich* can only refer to *Jan*, and never to *Peter*. Similarly, in the Russian examples, SE-anaphor *sebe/sebja* can only be bound by the subject DP, *Ivan* in (39-b) and *Marina* in (39-a) and in (39-c). A similar sentence from English, (39-e), is ambiguous: *himself* can be bound by either *Bill* or *John*. Below I demonstrate how subject orientation of monomorphemic anaphors can be derived in the theory of binding proposed above.

Recall that in order to get a reflexive bound by a non-subject argument in case of English, I proposed the positioning of the  $\rho$ -feature on V. VP acquires phasal status after merge of  $\nu$ , and therefore binding theory could apply at this level, resulting in VP internal binding. Now, let us examine how the derivation proceeds in Russian. Following recent work by Bailyn (2010, 2011), I assume that goal argument starts low, as an argument of the verb. I will first consider the derivation of (39-a), as it has basic word order with a low goal. At the beginning of the derivation, the verb is merged with its goal complement *sebe*. The verb has an uninterpretable valued T-feature  $\langle uT +val \rangle$ , and as I assumed before, an interpretable unvalued  $\rho$ -feature  $\langle i\rho -val \rangle$ ; the SE-element comes with unvalued  $\phi$ -features  $\langle i\phi -val \rangle$  and an uninterpretable valued  $\rho$ -feature,  $\langle u\rho +val \rangle$ .

The verb probes the reflexive in order to value its  $\rho$ -features, which become shared, as indicated by [1] in (40). The reflexive is raised to V, and further, the theme DP is merged in Spec,VP.

$v$  is now merged with a VP. It probes in order to value its T-features, and finds its goal in V. The T-features of  $v$  and V get shared, and  $v$  acquires a value for its T-feature. Recall that in the corresponding English derivation with a PP goal, VP acquired phasal status at this level: the  $\lambda$ -operator was inserted in V-position, and  $\lambda$ -conversion resulted in a reflexive variable being bound by an element within Spec,VP. However for languages with a monomorphemic reflexive, this is not the case. The presence of an unvalued  $\phi$ -feature on the reflexive prevents VP from becoming a closed domain, and therefore prevents the application of binding/ $\lambda$ -operator insertion. Therefore, binding is still impossible at this level, and the direct object cannot serve as a binder for the goal DP at this stage of the derivation:

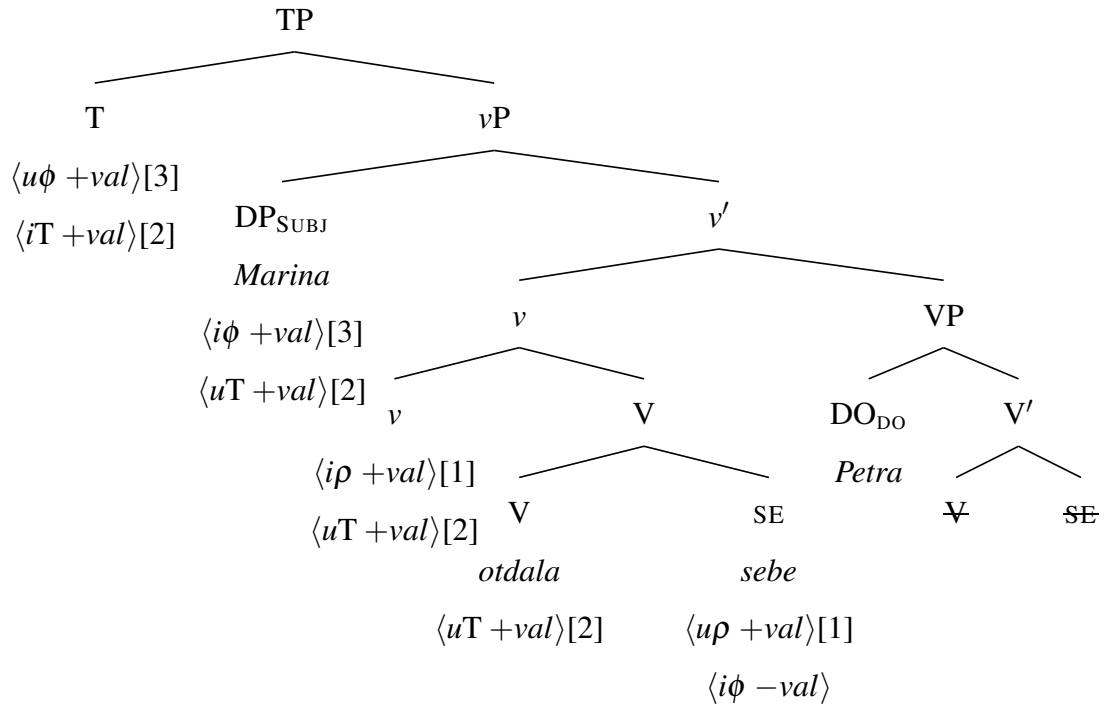
(40) VP structure of (39-a)



Next, V+SE is raised to  $v$ , and the subject is inserted in Spec, $v$ P position. T is further merged with  $v$ P: it has  $\langle u\phi -val \rangle$  and  $\langle iT -val \rangle$  features. As before, it first probes the subject DP, values its  $\phi$ -features, and also shares its T-features. On the second probing, it finds the  $v$ +V+SE complex. This probing results in T-feature valuation on T and the subject DP. Since the  $\phi$ -features of T are now valued, T will not probe the  $\phi$ -features of the reflexive, and they will remain unvalued. The TP

will close after the subject raises to Spec,TP. The presence of unvalued  $\phi$ -features on the reflexive will lead the derivation to crash.

(41) *Final structure of (39-a) after probing by T.*



Therefore, the only possibility in the case of ditransitive constructions is to place the  $\rho$ -feature on T, as in the case of simple transitive sentences. The derivation will be similar to the one described in the previous section, and the first domain which can be shipped to the semantic interface is the TP. As the  $\rho$ -feature on the T-head-complex is valued, a  $\lambda$ -operator is inserted above the T position. The reflexive variable is in its domain, and  $\lambda$ -conversion will result in the subject as binder:

- (42) a.  $\exists e : \text{Agent}(e, \text{Subj}) \ \& \ \text{Subj} \ \lambda x \ V(e) \ \& \ \text{Theme}(e, \text{DP}_{\text{DO}}) \ \& \ \text{Goal}(e, x)$   
 b.  $\exists e : \text{Agent}(e, \text{Subj}) \ \& \ V(e) \ \& \ \text{Theme}(e, \text{DP}_{\text{DO}}) \ \& \ \text{Goal}(e, \text{Subj})$

To summarize, placing the  $\rho$ -feature on V did not help us enforce a theme binder, as it did in the case of English. In English evaluating binding at the level of VP is an option, however, it

is prevented from happening in Russian by the existence of an unvalued  $\phi$ -feature on SE. The derivation would crash.

Now, what happens when the goal is moved higher than the theme, as in (39-b) and (39-c)? While the derivation will proceed in a slightly different way, the moment when the binding theory can be applied does not change. Since the  $\phi$ -features on SE only get their value after T probes SE, TP is still the first phase in the derivation which can be interfaced with the semantic component. At the moment of evaluation, the  $\rho$ -feature will still be present on T, and therefore the only possible antecedent of the reflexive is the subject DP in Spec,TP position.

### 3.2.3 Locality of binding

In the framework of binding proposed above the position of the  $\rho$ -feature is not fixed. It is determined during the derivation, and can be placed on  $v$ , V, or T. However, one question naturally arises: what prevents simple binding from being long-distance? To reformulate the question within the assumptions above, what prevents us from placing a  $\rho$ -feature within a matrix clause in a sentence like (43)?

(43) \*John said that Mary loves himself.

It might seem like placing the  $\rho$ -feature on the matrix verb would prevent embedded  $v$ P and CP from being closed, and therefore the long-distance probing of the reflexive by the matrix verb might be possible. However, this is not the case. I defined phases as domains within which all relevant features are either shared with an element outside of the domain, or valued. If one follows this definition, it becomes clear that in the case of English, VP will still become closed upon merge of  $v$ , exactly as before. The  $\rho$ -feature on reflexives is valued, and therefore does not prevent the domain from acquiring phasal properties. Closure of VP in the case of English will prevent further probing into VP, and if the  $\rho$ -feature is placed on the matrix verb, it won't be able to get its value from the reflexive embedded within the embedded VP, and the derivation will crash.

Similarly, in the case of monomorphemic reflexives, embedded indicative TP will close as soon as the subject is raised into Spec,TP: T-features on T and  $\phi$ -features on the subject are valued. As a result, the  $\rho$ -feature cannot be inserted higher than the embedded T, otherwise the  $\phi$ -features on the embedded reflexives will never get valued, and the derivation will crash.

In the next chapter I will argue that this does not always lead to a crash, particularly in the case of the  $\rho$ -feature being moved out of the VP domain high enough so that the matrix verb can probe it. This is how long-distance binding is derived.

### 3.2.4 Intermediate summary

The following points summarize what was argued above in this chapter:

- Valued  $\phi$ -features on the reflexive (HIMSELF) allow it to be bound early in the course of derivation.
  - VP is a possible binding domain for HIMSELF
  - Binding does not apply at the level of the lowest VP<sub>1</sub> in the case of ditransitive constructions: unvalued case prevents it. In this case, VP<sub>2</sub> is the possible binding domain.
- Unvalued  $\phi$ -features on the monomorphemic reflexive (SE) prevent it from being bound early.
  - Valuation of the  $\phi$ -feature on the reflexive happens after T with a  $\rho$ -feature probes the subject first, and the reflexive second.
  - TP is the only possible binding domain: the subject is the only potential antecedent.

The following table summarizes the binding domains for various types of verbs in the case of different reflexives. Notice that in case of ditransitive verbs by VP<sub>1</sub> I denote the lower VP, and VP<sub>2</sub> is the higher layer of VP.

(44) Binding Theory Domains

	Complex	Monomorphemic
Transitive	vP	TP
[ <sub>VP</sub> THEME GOAL	VP	TP
[ <sub>VP<sub>2</sub></sub> GOAL [ <sub>VP<sub>1</sub></sub> THEME	VP <sub>2</sub>	TP

### 3.2.5 Psych-verbs and binding

#### 3.2.5.1 Introducing psych-verbs

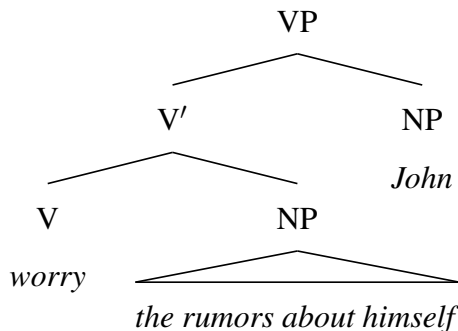
The next issue I will take on in this chapter is the particular status of psych-verbs with respect to binding theory. It has been noted previously (see for instance, Belletti and Rizzi 1988; Bouchard 1992; Landau 2010; Pesetsky 1995) that psych-verbs exhibit so called backward binding, in apparent violation of c-command requirements. The relevant examples from English and Italian are given below in (45)-(46).

- (45) a. [The rumors about himself]<sub>i</sub> worry John<sub>i</sub>.  
b. [Each other]<sub>i</sub>'s supporters worried [Freud and Jung]<sub>i</sub> (Pesetsky 1995, pg. 43)
- (46) a. Questi pettegolezzi su di sè<sub>i</sub> preoccupano Gianni<sub>i</sub> più di ogni altra cosa  
these rumors about SE worry Gianni more than each other thing  
'These rumors about himself worry Gianni more than anything else'  
b. I propri<sub>i</sub> sostenitori preoccupano Gianni<sub>i</sub>  
DET self's supporters worry Gianni  
'His own supporters worry Gianni'

The subject in these examples contains a reflexive, which is not c-commanded by its antecedent at surface structure. This data has long been considered as primary evidence against a surface structure version of Principle A, according to which the reflexive must be c-commanded by its antecedent on the surface. Belletti and Rizzi 1988 proposed that the surface subject originates as a deep structure object, and undergoes raising to the Spec,IP position. According to them, Principle A is satisfied at deep structure, and subsequent raising does not trigger a violation. Based on

this data, they formulate the “anywhere” version of Principle A. According to it, if a reflexive is c-commanded by its antecedent at any time during the derivation, Principle A is satisfied. The structure assumed by Belletti and Rizzi 1988 is given below in (47).

(47) *The structure of the psych-verb VP according to Belletti and Rizzi 1988*



As one can see, the anaphor *himself* within the surface subject is c-commanded by its antecedent *John*, and therefore the “anywhere” version of Principle A is satisfied.

The structure assumed by Belletti and Rizzi 1988 is problematic from the point of view of the current theory. First, right specifiers are dispreferred, and often simply disallowed in the modern framework (cf Kayne 1994). That, however, is the smallest problem with this analysis, as it can be potentially resolved by postulating a  $vP$  projection. Two other problems are much harder to overcome.

One of them is the apparent violation of minimality. Belletti and Rizzi 1988 assume that the NP *the rumors about himself* moves to a Spec,IP (Spec,TP) position. Let us assume that this movement occurs in order to satisfy the EPP feature of the T head. Notice, that the NP *John* is located closer to T than the NP *the rumors about himself*. Therefore, one would expect that the experiencer will be moved in order to satisfy EPP. That, however, is impossible, as the sentence *John worry the rumors about himself* is ungrammatical. The movement of the theme NP therefore violates minimality. Before adopting Belletti and Rizzi’s analysis one has to come up with a theory why no violation is observed in this case.

Another problem is related to case assignment. Compare the following two examples (48):



- (48) a. The pictures of himself worry John.  
b. John worries about the pictures of himself.

One of the major question one has to answer is what the relation is between these two sentences. If they are derived from the same underlying structure, what allows for two alternative derivations? If they are not derived from the same structure, what is the difference between them? No matter what the answer to these questions are, one has to explain how accusative case is assigned to the experiencer in (48-a), and why there is no accusative case nominal phrase in (48-b). In what follows, I propose an analysis of these constructions.

### 3.2.5.2 The causative nature of object experiencer verbs

The first observation is that while these constructions are similar, they in fact might have different truth conditions. Consider the following scenario (Chris Parles, p.c.). Imagine John looking at two pictures of himself, one from a year ago, and one current. He notices that something changed in his face, i.e. a mole on his cheek grew in a year, and looks much bigger on the current picture than on the old one. The only felicitous sentence in this situation will be the one in (48-a). He does not worry about the pictures *per se*, all John worries about is his medical condition. The sentence can be paraphrased as *The pictures of himself cause John to be worried*. Notice that the example in (48-b) does not necessarily have a causative reading, i.e. it can be uttered in a situation when John is worried that someone might steal his pictures. In this case, it is not the pictures that cause him to be worried, but rather the absence of the lock on the door, high crime rate in the neighborhood, etc. I take the causative character of the sentences with object experiencers to be the basis for my analysis.

I propose that there is a Caus(ative) head responsible for the causative reading of the sentences with Object Experiencers, and this head assigns accusative case to the experiencer. This head is present in example (48-a), but absent in example (48-b). Assuming that this head assigns accusative case, one of the puzzles mentioned above, relating to the assignment of accusative case, can be

solved.

### 3.2.5.3 The special status of accusative case in object experiencer constructions

Landau 2010 provides a number of diagnostics on how the accusative case assigned to experiencers is different from accusative case assigned to the objects of simple transitive sentences. I will reproduce some of them in this section.

Consider the case of Russian. Object experiencer constructions are also present in Russian, and have similar properties to those in English, for instance, they also allow backwards binding, which is in general not allowed with non-psych verbs, cf. English example (49) and Russian (50).

- (49) a. Stories about himself<sub>i</sub> worry John<sub>i</sub>  
b. \*Stories about himself<sub>i</sub> describe John<sub>i</sub>

- (50) a. Rasskazy o sebe<sub>i</sub> volnujut Ivana<sub>i</sub>  
stories<sub>NOM</sub> about SE worry Ivan  
b. \*Rasskazy o sebe<sub>i</sub> opisывajut Ivana<sub>i</sub>  
stories<sub>NOM</sub> about SE describe Ivan

Accusative-marked objects of transitive verbs in Russian can undergo the following processes: genitive of negation and *po*-distribution (Landau 2010).

The genitive of negation in Russian is a process that can change an accusative marked element into genitive in negative contexts. The detailed description of this construction is beyond the scope of this dissertation, but cf. Babby 1987; Bailyn 2004 for possible analyses. Examples of the genitive of negation are given below in (51) and (52).

- (51) a. Putin zametil bystro približajuščijsja poezd  
Putin noticed quickly approaching<sub>ACC</sub> train<sub>ACC</sub>  
b. \*Putin zametil bystro približajuščegosja poezda  
Putin noticed quickly approaching<sub>GEN</sub> train<sub>GEN</sub>  
'Putin noticed a quickly approaching train'

- (52) a. Putin ne zametil bystro približajuščijsja poezd  
 Putin NEG notices quickly approaching<sub>ACC</sub> train<sub>ACC</sub>  
 b. Putin ne zametil bystro približajuščegosja poezda  
 Putin NEG noticed quickly approaching<sub>GEN</sub> train<sub>GEN</sub>  
 ‘Putin didn’t notice a quickly approaching train’

If there is no negation in the sentence, the only possibility of case marking *poezd* ‘train’ is by using the accusative case, the corresponding sentence with the genitive case is ungrammatical, see (51). If negation is present, as in examples in (52), the genitive case can be substituted for accusative.

Now let’s consider accusative experiencer objects. If accusative case is assigned by the same head as it is assigned to the object of transitive verbs, we would expect the genitive of negation to be available. This however is not the case. Consider examples in (53).

- (53) a. Fotografii Putina v presse (ne) volnujut ego ženu  
 photographs Putin<sub>GEN</sub> in press (not) worry his wife<sub>ACC</sub>  
 ‘Pictures of Putin in press (do not) worry his wife’  
 b. \*Fotografii Putina v presse (ne) volnujut ego ženy  
 photographs Putin<sub>GEN</sub> in press (not) worry his wife<sub>GEN</sub>  
 ‘Pictures of Putin in press (do not) worry his wife’

If the object experiencer is marked with accusative case, both negative and positive versions of this sentence are grammatical, as expected. At the same time, the object experiencer cannot be marked with genitive case, regardless of whether the sentence is positive or negative. Based on this evidence we conclude that the source of accusative case in object experiencer constructions is different from the source of accusative case for objects of transitive verbs.

The second diagnostic I present here is based on the *po*-distribution test. Accusative objects of Russian transitive verbs can undergo a change into a prepositional phrase headed by a distributive preposition *po*, and that gives the corresponding sentences a distributive reading, see (54).

- (54) a. Ivan vypil butylku vina  
 Ivan drank bottle<sub>ACC</sub> wine  
 ‘Ivan drank a bottle of wine’

- b. Ivan vypivaet po butylke vina každyj den'  
 Ivan drinks PO bottle<sub>DAT</sub> wine every day  
 'Ivan drinks a bottle of wine every day'

Now again let's consider what happens if one tries to apply *po*-distribution to accusative experiencers.

- (55) a. Fotografii Putina v počtovom jaščike volnujut Ivana  
 photographs Putin in mail box worry Ivan<sub>ACC</sub>  
 'Pictures of Putin in a mailbox worry Ivan'
- b. \*Fotografii Putina v počtovom jaščike volnujut po odnomu čeloveku každyj den'  
 photographs Putin in mail box worry PO one person every day  
 'Pictures of Putin in a mailbox worry one person every day'

(55) shows that the object experiencer cannot undergo *po*-distribution, unlike the objects of transitive verbs, and it is not a semantic oddity. In fact, sentences like (55-b) are ungrammatical even if one assumes that Putin's secret police distributes Putin's pictures among mailboxes, one mailbox a day. The conclusion of this test agrees with the conclusion of the previous test: Object Experiencers behave differently from accusative objects of transitive verbs, therefore accusative case in object experiencer constructions is assigned in a different way.

I will take this evidence from Russian to reflect the universal cross-linguistic differences in the nature of accusative case on object experiencers.

### 3.2.5.4 Deriving object experiencer constructions

In this section I propose how object experiencer constructions are derived, based on the existence of the Caus(ative) head. I assume that Caus head selects a *v*P, and projects a CausP projection, right above *v*P.

Consider the sentence (56).

- (56) Pictures of Putin worry John.

I assume that the theme argument, i.e. *pictures of Putin*, is generated as the complement of the verb position, based on UTAH (Experiencer > Theme). I also assume that the Experiencer argument, in this case *John*, is generated in the Spec, $\nu$ P position, as it has the highest thematic role in this sentence, according to UTAH.

Now consider which features are involved in this derivation. Following the framework of Pesetsky and Torrego 2007, I propose that the causative reading of the sentence arises from the causative feature on the Caus head. Following the same logic as before, I assume that the Caus feature is interpretable but unvalued on the Caus head,  $\langle i\text{Caus} -val \rangle$ . Further, I postulate that the counterpart of this Caus feature is present on psych-verbs, this time it is valued but uninterpretable,  $\langle u\text{Caus} +val \rangle$ .

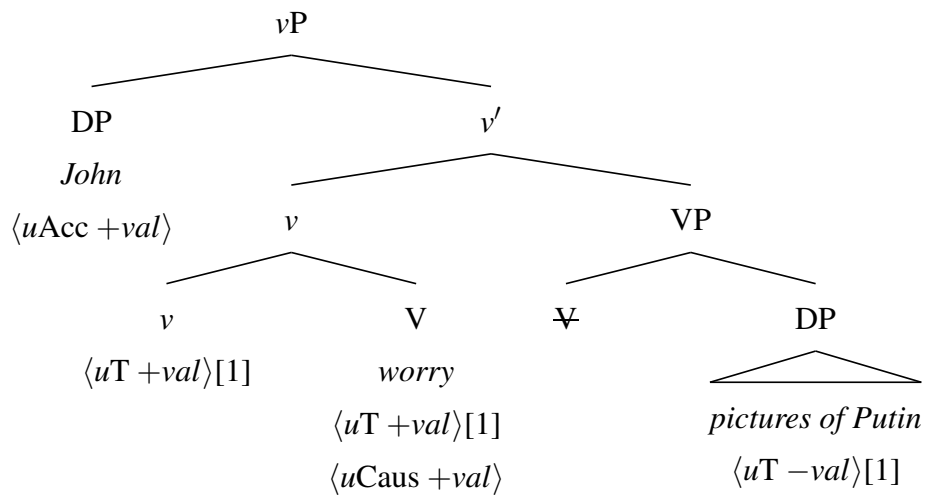
As I showed above, the presence of accusative case on the experiencer gives the sentence a causative reading. Therefore, I propose that the Acc case feature is present on the Caus head, and the Caus head checks accusative case with the experiencer argument.<sup>4</sup>

Now I will show how the derivation proceeds in detail. First, the lexical verb *worry* is merged with the theme DP, *pictures of Putin*. Notice, that since *pictures of Putin* in nominative, it is endowed with an unvalued uninterpretable T-feature,  $\langle u\text{T} -val \rangle$ . Since V is also a bearer of T-features, the T-features on the nominative theme and the verb get shared and valued. After this merge, VP is formed. It is further selected by  $\nu$ , which has uninterpretable unvalued T-features as well, as  $\nu$  probes V in order to check its T-feature. This probing results in sharing of T-features between  $\nu$  and V, and valuation of T-feature on  $\nu$ . The experiencer argument *John* is further merged in the Spec, $\nu$ P position. The resulting structure is shown in (57). Notice that at this time there are no unvalued features within  $\nu$ P as the feature  $\langle u\text{T} -val \rangle$  feature on the theme argument, *pictures of Putin* has been checked earlier, when this argument was merged with V.

---

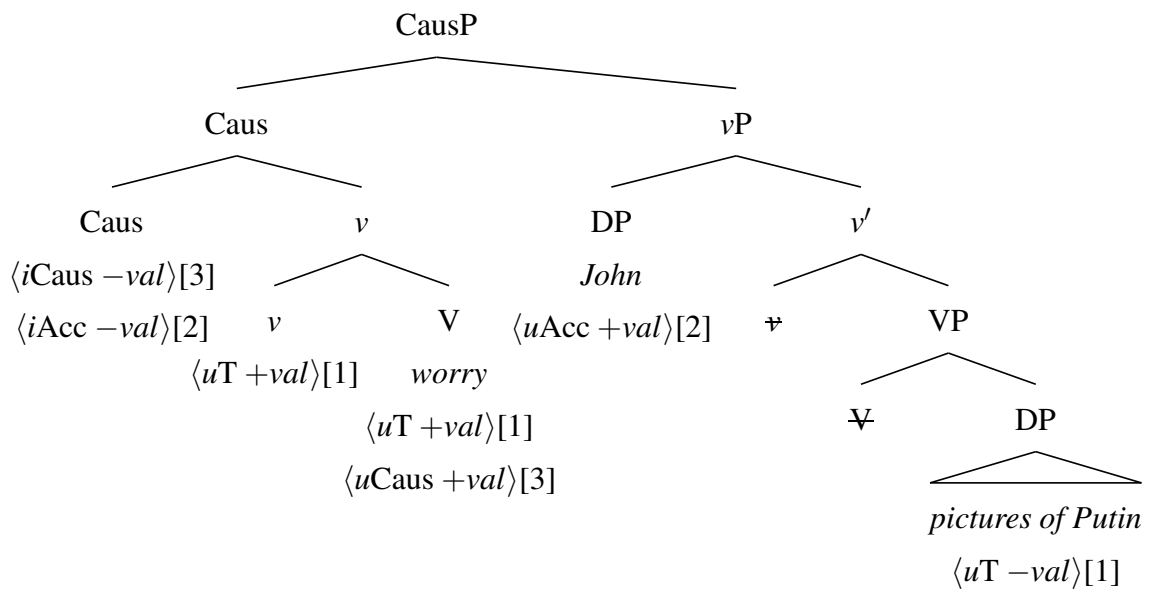
<sup>4</sup>The question is why aren't there two accusative cases in sentences with Psych verbs? Why isn't accusative case assigned by  $\nu$ ? Notice that the in psych-constructions  $\nu$  lacks agentive interpretation. If one assumes that only agentive  $\nu$  assigns accusative case, the problem with the presence of two accusative cases will be solved. For the details of this solution see Bailyn 2011, ch. 4.

(57) The *vP* structure of a psych-construction in (56).



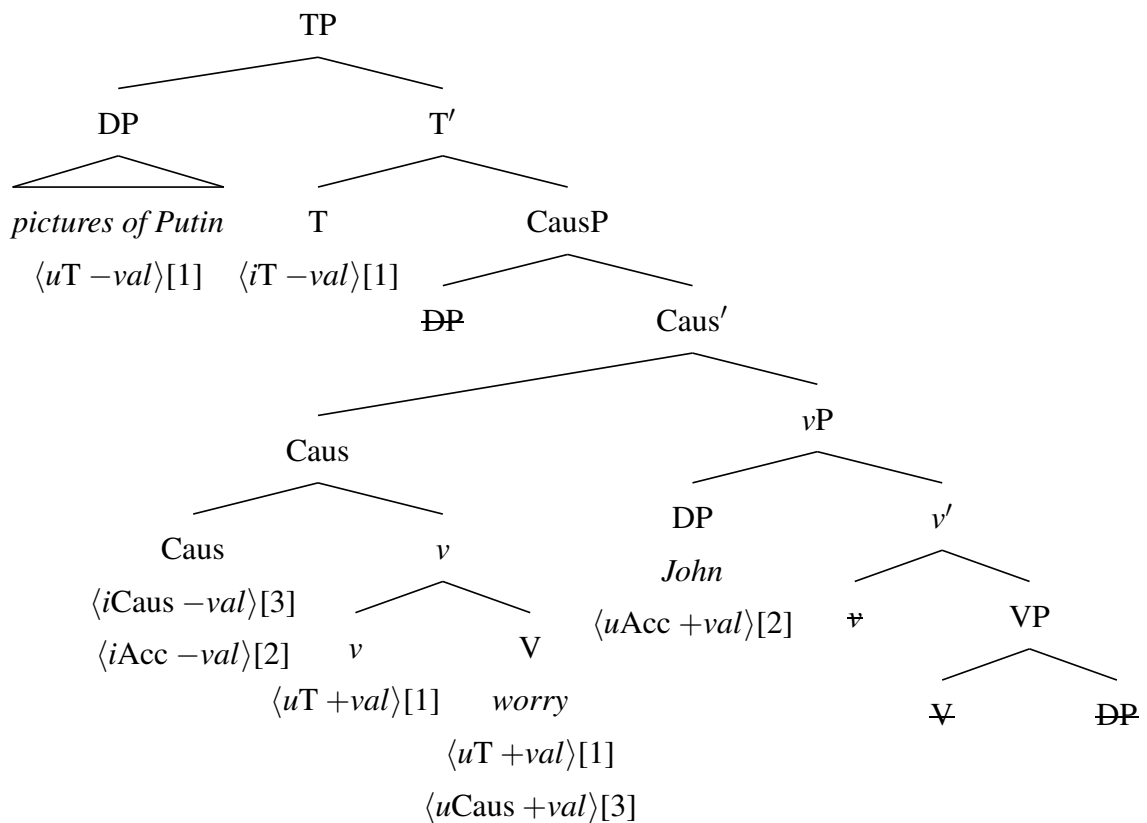
Now, the Caus head with  $\langle u\text{Caus} -val \rangle$  feature is merged with *vP*. Apart from the Caus feature, the Caus-head also comes with the accusative case feature,  $\langle i\text{Acc} -val \rangle$ . It probes the experiencer DP, *John*, in order to value its Case features, its  $\langle \text{Acc} \rangle$  feature becomes valued and shared with the  $\langle \text{Acc} \rangle$  feature of the experiencer DP, which is indicated in the tree in (58) by [2]. It continues to probe in order to value its  $\langle \text{Caus} \rangle$  feature, and finds its goal in *v+V* complex head. The Caus-feature on Caus head gets valued, and the head complex undergoes head-raising to Caus. (58) shows the structure at this moment of the derivation:

(58) Structure of a psych-construction in (56) after probing by Caus.



After that the theme DP *pictures of Putin* raises to the Spec,CausP. T is then merged into the structure, and it probes the nominative theme in order to value its T-features; after this probing, the DP *pictures of Putin* raises to Spec,TP. The final structure is given in (59).

(59) Final structure of a psych-construction in (56).



### 3.2.5.5 Extensions to binding

Now let us consider what happens if the theme DP contains a reflexive, i.e. instead of *pictures of Putin* in (56) we have *pictures of himself*.

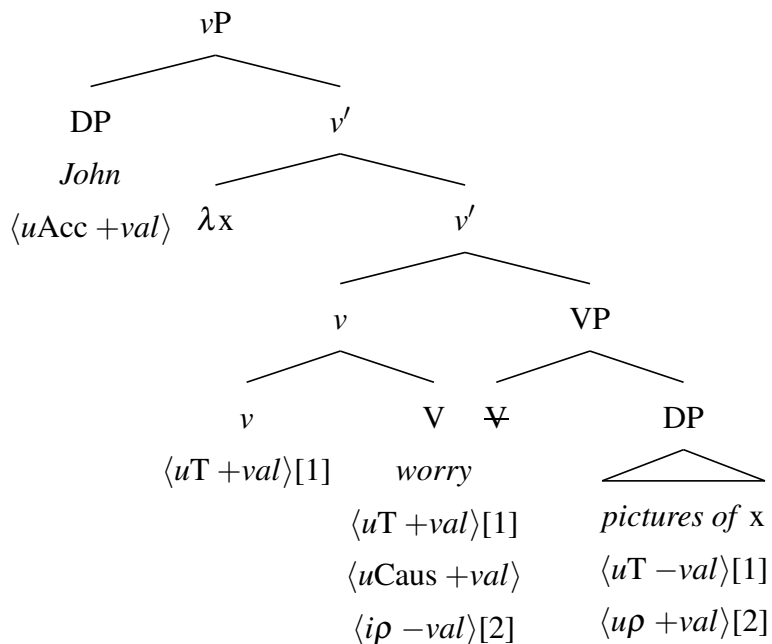
(60) Pictures of himself worry John.

In the English case, the  $\phi$ -features on the reflexive are valued, and therefore will not interfere with the moment binding theory can apply. That means, that the  $\lambda$ -conversion proposed above will apply as early as all features within a phrase are valued. As I demonstrated above, it happens at the moment shown above in (57). At this level, regardless of the initial placement of  $\rho$ -feature, the only possible binder for the reflexive within the theme is the object experiencer.  $\lambda$ -conversion will



apply at this level and produce the desired results.

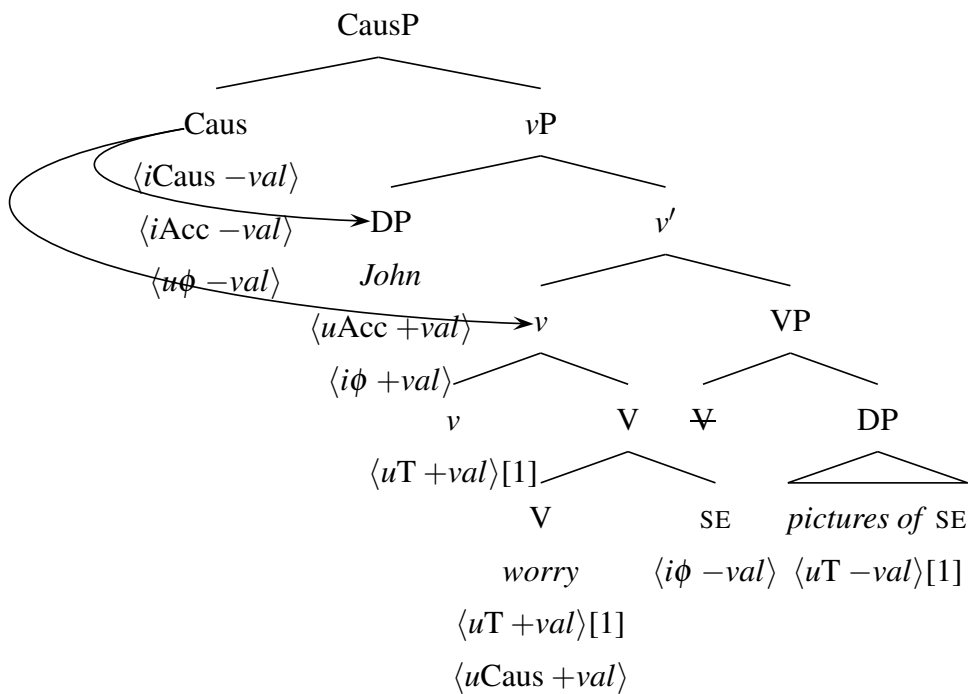
(61) *vP structure of an English psych-construction in (60).*



Now let us turn to psych-construction in languages with SE-reflexives. Their  $\phi$ -features are unvalued, and that prevent binding from applying as early as in English: at the moment *vP* is completed, the  $\phi$ -feature of the reflexive are not yet valued, and therefore the structure cannot be sent to the semantic component. The crucial moment for this derivation is later, when the Caus-head is merged into the structure and probes first the experiencer in Spec,*vP*, and then the *v+V* head complex.

Recall that I assume that monomorphemic reflexives undergo head-raising to V, and further to *v*. Let us assume that Caus head also has  $\phi$ -features, associated with the Acc case it assigns to the experiencer. The tree in (62) shows the structure after the merge of Caus.

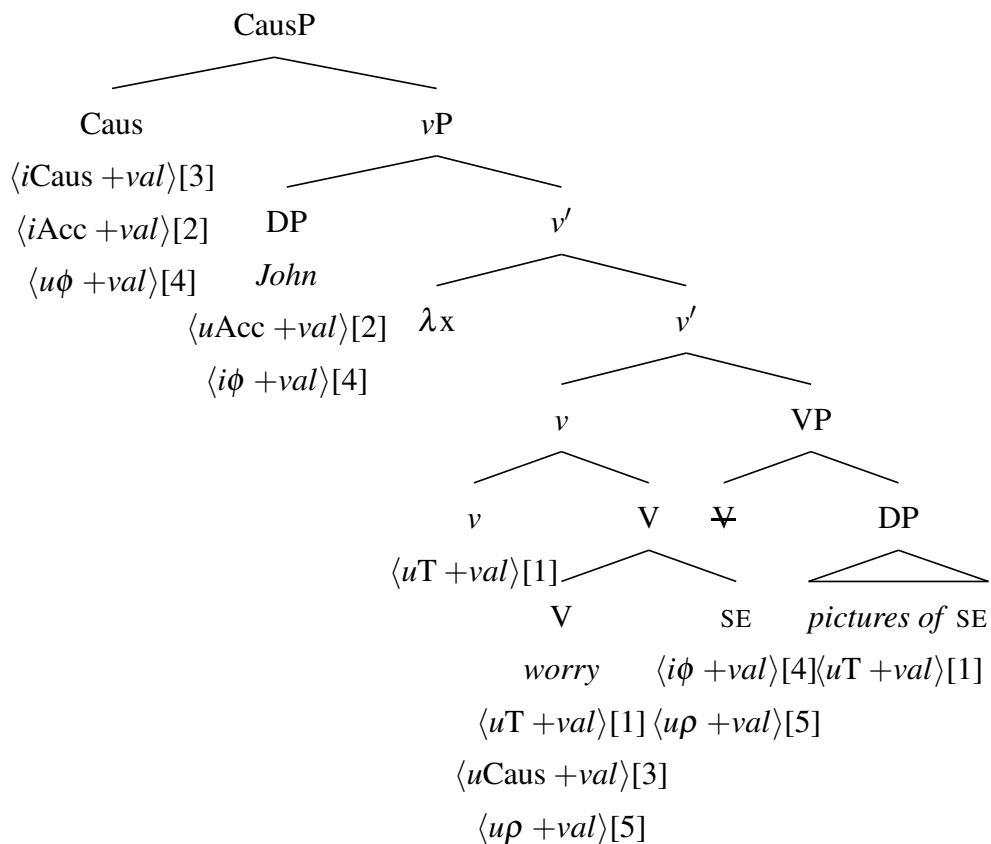
(62) *The structure of an English psych-construction in (60) after the merge of Caus.*



In this case, after Caus probes the experiencer DP in the Spec,vP position, not only does it check its Case feature, but it also gets a value for its  $\phi$ -features from the object experiencer, *John*. The subsequent probing of the  $v+V+SE$  complex results as usual in valuation of the Caus features on the Caus head, but also in valuation of  $\phi$ -features on SE parasitic on this probing. Notice that this is similar to the way SE gets its  $\phi$ -features in simple transitive sentences after being probed by T, with this sole difference that the Caus head does the probing, and the  $\phi$ -features come not from the nominative, but from the accusative argument.

This probing results in all features within vP being valued, and therefore binding can apply at this stage, giving the accusative theme as the only possible binder. The complete structure after probing by the Caus head occurred and  $\lambda$ -conversion can apply is given in (63).

(63) CausP structure of an English psych-construction in (60).



### 3.2.5.6 Psych-verbs: conclusion

As one can see, the proposed theory of binding together with the proposed causative analysis of psych-constructions explains why binding can apply early in derivation at the level of vP, before movement of the thematic argument to the Spec,CauseP (and further to the Spec,TP) position. We again saw that the English case is less problematic: since the reflexive has its  $\phi$ -features valued, it can get bound as soon as other features of other elements are valued.

Languages with SE-type reflexives are more complicated: one needs to explain how  $\phi$ -feature valuation occurs. In both cases demonstrated above, simple transitive and psych-verbs, the mechanism is similar.  $\phi$ -feature valuation of SE is parasitic on the probing by some head H (T or Caus, respectively) of the verbal complex  $v+V+SE$ .

# Chapter 4

## Long-distance binding

In this chapter I turn my attention to long-distance anaphoric binding relations. I examine two types of long-distance binding. The first type of long-distance binding considered in this section is the case of *proper* long-distance binding. In this case, reflexives that are present low in the embedded clause, e.g. in the object of the embedded clause, get bound by a matrix element, usually the subject of the matrix clause. In the examples presented in this section I will demonstrate that proper long-distance binding can often be unbounded, i.e. the reflexive can be potentially anteceded by a DP in any higher clause, even though there are certain preferences. I will also take on the issue of tense dependency of binding relations, and propose an analysis of long-distance binding explaining asymmetries between binding into infinitive, subjunctive, and indicative clauses.

Second, I consider cases when the reflexive is located within the subject of the embedded clause, and is bound by an antecedent present in the matrix clause. While this case cannot be considered a case of proper long-distance binding, i.e. the binding domain of the anaphor is not unbounded, it is nevertheless presents a problem for binding theory in general. Under the assumption that the binding domain of a reflexive is a TP (as in the previous literature on the topic), or a phase (as in the currently proposed theory), one would not expect the antecedent of the reflexive to be located in the next higher clause. In this case, the locality condition on binding still holds, since the element which binds the reflexive is the next higher DP in the structure, i.e. it is the closest

potential binder. Cases considered in this section are usually unambiguous.

The interaction between movement and binding is examined in Chapter 5.

## 4.1 Overview of long-distance reflexivization

In this section I explore instances of *proper* long-distance binding: cases when an object reflexive in the embedded clause can be bound by a nominal from the matrix clause. The difference between these cases and cases considered in Chapter 3 is the fact that for an embedded object reflexive there exists a local DP which serves as a potential binder. This option however is not always entertained: in certain contexts described below a binding relationship obtains between a matrix DP and the embedded object anaphor.

The following examples repeated from Chapter 2 demonstrate various possibilities of long-distance reflexivization. The domain in which a reflexive must be bound differs cross-linguistically, and can vary with the tense of the clause. A number of languages do not allow long-distance binding at all: all reflexives must be bound within the minimal clause, containing it. On the other hand, a wide range of languages is more permissive when it comes to possibility of reflexive binding.

The availability of long-distance binding out of infinitival clauses is common in Slavic languages, see the Russian example below. The possibility of long-distance binding into a subjunctive clause is attested in Icelandic, French, and Italian among others. Finally, in a number of languages binding is virtually unconstrained and anaphors within the embedded indicative clause can be bound by the nominals in any higher clause; such languages include Malay, exemplified below, and also a number of East-Asian languages, such as Chinese, Japanese, and Korean (module person constraint in Chinese, see Huang and Liu 2001).

(1) *Binding out of Small Clause*

Larsen<sub>1</sub> betragter Jorgen some farlig for sig<sub>1</sub>. (Danish)  
Larsen considers Jorgen as dangerous for self  
'Larsen considers Jorgen dangerous for self.' (Pica 1986)

(2) *Binding out of an Infinitival Clause*

Professor<sub>1</sub> poprosil assistenta<sub>2</sub> PRO<sub>2</sub> čitat' svoj<sub>1/2</sub> doklad. (Russian)  
Professor asked assistant read self's report  
'The professor asked the assistant to read self's report.' (Progovac 1993b, pg. 755)

(3) *Binding out of Subjunctive Clause*

On<sub>1</sub> souhaite toujours que les gens ne disent pas du mal de soi<sub>1</sub>. (French)  
one wishes always that the people NEG speak NEG of ill of self  
'One always wishes that people do not slander oneself.' (Pica 1991)

(4) *Binding out of Indicative Clause*

Ali<sub>1</sub> berharap Fatimah<sub>2</sub> akan berkahwin dengan dirinya<sub>1/2</sub>. (Malay)  
Ali hope Fatimah will marry with self<sub>3SG</sub>  
'Ali hopes that Fatimah will marry himself/herself.' (Ngoh 1991)

A number of properties of long-distance reflexives have been outlined in the literature. I review them below based on Huang 2000.

The first property concerns what can be a possible antecedent of a long-distance reflexive. It has been noted that the antecedents of a long-distance reflexive tend to be subjects — this is the *subject orientation of a long-distance reflexive*. It is usually true in languages such as Russian, Japanese, Korean, Chinese, Marathi, however a number of counterexamples were observed.

(5) Minine<sub>1</sub> Vinulaa<sub>2</sub> Kalavle ki aapan<sub>1/\*2</sub> turungaat aahot. (Marathi)  
Mini<sub>ERG</sub> Vinu<sub>DAT</sub> informed that self prison<sub>LOC</sub> was  
'Mini informed Vinu that self was in prison.' (Wali 1989, pg. 83)

(6) *Counterexamples to subject orientation*

- a. Die beschrijving van henzelf<sub>1</sub> als communist ergerde de Gaulle<sub>1</sub>. (Dutch)  
 that description of himself as communist annoyed de Gaulle  
 ‘That description of himself as a communist annoyed de Gaulle.’ (Anagnostopoulou  
 and Everaert 1999)
- b. La propria<sub>1</sub> salute preoccupa molto Osvaldo<sub>1</sub>. (Italian)  
 the self health worries very much Osvaldo  
 ‘Self’s health worries Osvaldo a lot.’ (Giorgi 1983)
- c. Taroo-wa Takasi<sub>1</sub> kara Yosiko-ga zibun<sub>1</sub>-o nikundeiru koto-o kiita (Japanese)  
 Taroo-TOP Takasi from Yosiko-NOM self-ACC be-hating COMP-ACC heard  
 ‘Taroo heard from Takasi that Yosiko hated self.’ (Sells 1987)
- d. Vaasu<sub>1</sub> tanu\*<sub>1/2</sub> picci-di ani kamalaa-to<sub>2</sub> ceppeedu. (Telugu)  
 Vaasu self crazy-3F that Kamalaa-to-3F told  
 ‘Vaasu told Kamala that self is crazy.’ (Wali and Subbarao 1991)

While some examples above are not surprising (see 1.1.5) as they involve psych-verbs, examples from Japanese and Telugu demonstrate that non-subjects can be binders for the long-distance reflexives.

The next generalization concerns the preferred reading in the case of long-distance bound reflexives. Possible antecedents of a long-distance reflexive can in principle come from any superordinate clause, but the highest clause subject is in general preferred as an antecedent to subjects of non-root clauses — this is the so-called *the maximality effect*. The following example (7) from Huang 2000 demonstrates it. Interestingly, both the most embedded and the root clauses’ subjects are preferred antecedents compared to the subjects of the intermediate clause, even though all of them are possible.

- (7) Xiaoming<sub>1</sub> yiwei Xiaohua<sub>2</sub> zhidao Xiaolin<sub>3</sub> xihuan ziji<sub>1>3>2</sub> (Chinese)  
 Xiaoming think Xiaohua know Xiaolin like self  
 ‘Xiaoming thinks that Xiaohua knows that Xiaolin likes self’ (Huang 2000)

The third, and one of the most discussed about properties, is the monomorphemicity of long-distance reflexives.

- (8) Jón<sub>1</sub> segir að María<sub>2</sub> elski sig<sub>1/2</sub>/ sjálfan sig\*<sub>1/2</sub> (Icelandic)  
 John says that Mary loves self/ self self  
 ‘John says that Mary loves self/self self.’ (adapted from Sigurðsson 1990)

While this generalization is very robust, and holds within individual languages which have both simplex and complex reflexives, counterexamples still exist:

- (9) *Counterexamples to monomorphemicity*
- a. Xiaoming<sub>1</sub> shuo leisheng ba ta ziji<sub>1</sub> xiao le yi tiao. (Chinese)  
 Xiaoming say thunder BA 3SG self frighten PERF one CL  
 ‘X.-M. said that the loud clash of thunder had given himself a fright.’ (Huang 2000)
- b. Abdullah<sub>1</sub> menolong Aminah mencuci pakaian dirinya<sub>1</sub>. (Malay)  
 Abdullah help Aminah wash clothes self.3SG.POSS  
 ‘Abdullah helps Aminah wash himself’s clothes.’ (Ngoh 1991)
- c. Takasi<sub>1</sub>-wa Hiroshi-ga zibun-zisin<sub>1</sub>-ni kasite kureta kuruma-o kowasite simatta.  
 Takasi-TOP Hiroshi-NOM self-self-DAT lend give car-ACC broken ended up  
 (Japanese)  
 ‘Takasi has broken the car which Hiroshi lent self-self.’ (Kato 1994)
- d. Kaali<sub>1</sub> uqar-p-u-q Pavia<sub>2</sub> immi-nirmi-nit<sub>1</sub> angi-nir-u-sinnaa-nngi-tsu-q (Inuit)  
 Kaali say-IND-3SG Pavia self-OBV-ABL big-CMP-be-can-NEG-PRT-3SG  
 ‘Kaali said that Pavia could not be taller than self.’ (Bittner 1994)

In the following section I will apply the feature-based binding theory from Chapter 3 to the cases of proper long-distance binding.

#### 4.1.1 Feature-based theory of long-distance binding

How can long-distance binding be derived in the feature-based framework? Recall that I proposed that binding is evaluated by phases. The  $\rho$ -feature placed on a head (for instance,  $v/V$ ) introduces a  $\lambda$ -operator at the moment the domain can be sent to the semantic component for evaluation. Upon Merge of a DP into the structure,  $\lambda$ -conversion applies, and the binding relation is established between the nominal and the anaphor.



I argued that the  $\lambda$ -operator is inserted only at the moment when the domain can be sent to evaluation, i.e. when all relevant features within the domain are valued. For instance, I showed that in double object constructions in English the  $\lambda$ -insertion cannot apply before the case feature on the goal DP is valued. Another case included the impossibility of  $\lambda$ -insertion in SE-type languages:  $\phi$ -features are unvalued on the reflexive until T probes the  $v+V$  complex, and therefore, in such languages the first domain at which binding of the SE-anaphor can apply is TP, as opposed to *himself* with valued  $\phi$ -features in English, which can be bound at the level of VP or  $vP$ . This derives the subject-orientation of SE-anaphors. To summarize, I argued that binding applies only at the moment of complete domains.

How can this theory be extended to incorporate possibilities of long-distance binding? In the previous section I showed that the  $\rho$ -feature does not necessarily have to be placed on  $v/V/T$  head: in the case of subject-internal anaphors in English, it must be placed on C in order to allow binding of the subject-internal anaphor by a nominal from the matrix clause. Let us entertain this possibility further.

The first question one has to answer is what happens when one places the  $\rho$ -feature on C in case of object reflexives. For instance, why are the sentences in (10) ungrammatical under coindexation of the anaphoric expressions with the matrix subject?

- (10) a. \*Mary<sub>i</sub> said that Bill loves herself<sub>i</sub>.  
 b. Maria<sub>i</sub> skazala čto Ivan<sub>j</sub> ljubit sebja\*<sub>i/j</sub>. (Russian)  
 Maria said that Ivan loves SE  
 ‘Maria said that Ivan loves himself/\*her’  
 c. Deti<sub>i</sub> skazali čto [ix roditeli]<sub>j</sub> ljubjat drug druga\*<sub>i/j</sub>.  
 children said that their parents love EACH-OTHER  
 ‘Children<sub>i</sub> said that [their parents]<sub>j</sub> love each other\*<sub>i/j</sub>.’

Imagine that the  $\rho$ -feature is placed on the embedded C. I argued above that in English sentences with object reflexives the domains which can be sent to interpretation are VP and  $vP$ , while in Russian the minimal domain which can be sent to interpretation is  $vP$ . At the moment  $vP$  is completed,

T probes the subject DP in Spec, $\nu$ P position and shares  $\phi$ - and T-features with it. Further T probes the  $\nu$ +V complex. The T-features on T, the subject DP, and the  $\nu$ +V complex get shared, however, the  $\phi$ -feature on the reflexive does not get valued, since T does not probe the reflexive SE within the verbal head-complex. Clearly at this level  $\nu$ P cannot be shipped to the semantic component, since the  $\phi$ -features within the head of  $\nu$ P on SE are not valued. The subject DP now raises to Spec,TP.

Next, C is merged with TP, and probes T in order to share its T-features. At this moment, there are no unvalued features at the edge on TP: the T-features and  $\phi$ -features on T are now valued. TP acquires phasal status, and must now be closed for further probing.

Therefore, the semantics receives a TP containing an anaphor with an instance of a valued but uninterpretable  $\rho$ -feature, and with unvalued interpretable  $\phi$ -features. Notice, that since indicative C in cases like (10) does not have  $\phi$ -features, it will no longer be possible to value  $\phi$ -features on the reflexive, since it can only be done by T at the level of TP (first, getting  $\phi$ -features from the subject DP in Spec, $\nu$ P, and further transferring them to the reflexive as a byproduct of  $\rho$ -feature probing.).

The crash of the derivation is predicted by the theory of feature-based binding, and placement of an interpretable instance of  $\rho$ -feature on C does not give rise to long-distance binding in basic cases like (10).

The problem with placing a  $\rho$ -feature on C stems from the fact that TP is a phasal domain, and therefore gets shut down for probing upon completion, preventing further probing which could potentially result in valuation of  $\phi$ -features on SE. The prediction one makes is that if there is something which prevents closure of TP/CP the long-distance binding can potentially be possible. Further, moving the anaphor with its  $\rho$ -feature (or just its  $\rho$ -feature) outside of  $\nu$ P/TP might also affect binding possibilities, allowing  $\rho$ -feature on C (or higher) to probe it. These three possibilities are summarized in (11).

- (11) *Potential environments for long-distance binding availability:*
- a.  $\nu$ P/TP/CP is prevented from closing for an independent reason; or
  - b. Overt movement of the anaphor outside of  $\nu$ P/TP domain; or
  - c. Overt movement of the anaphor's  $\rho$ -feature outside of  $\nu$ P/TP.

Which factors may prevent  $\nu$ P/TP from acquiring a phasal status? One of the major factors affecting closure of  $\nu$ P/TP is Tense. Pesetsky and Torrego 2007 argue that the infinitival form of the verb comes with an unvalued T-feature  $\langle uT -val \rangle$ , and must get its value from the matrix clause under agree relation. Further, as I will argue below, subjunctive complements also rely on the matrix clause in order to be situated in time, and the T-feature on the subjunctive form of the verb is unvalued. Therefore, the prediction is that subjunctive and infinitival complements can (potentially) allow long-distance binding into them.

The next potential environment for long-distance binding is related to anaphoric movement. If the anaphor escapes  $\nu$ P, its  $\rho$ -feature will be able to enter an Agree relation with an interpretable counterpart of the  $\rho$ -feature on C (or even higher), and the derivation will not crash. Scrambling of the anaphor would be an example of this option.

The third possibility is feature movement: if the  $\rho$ -feature raises outside of  $\nu$ P, it can enter in a local relationship with C, and therefore can get probed by it, preventing the crash of the derivation. Recall, that in SE-type languages, reflexive features invariably raise to the  $\nu$ -head, the position where their  $\phi$ -feature valuation occurs. The way to escape  $\nu$ P in this case is equivalent to  $\nu+V$  complex raising out of  $\nu$ P, say to T, and possibly further to C. Notice that neither Russian nor English have V-to-T raising. Based on this fact, one would predict that Russian has less long-distance binding possibilities than, say, Dutch or Icelandic, or Modern Romance languages, which appears to be true.

Below I explore all these possibilities in order, and demonstrate that the predictions made by the theory are borne out.

## 4.2 Interaction of binding and tense

In order to explore the interaction of anaphoric binding with tense properties of the embedded clause I start off by exploring how infinitival and subjunctive embedded clauses can be accounted for within the Pesetsky and Torrego's 2007 framework. First, I describe an analysis of subjunctives based on Russian, and further show that it can be extended to other languages with subjunctives, such as Romance languages and Icelandic. Second, I explore how control clauses are derived.

### 4.2.1 An analysis of the indicative/subjunctive distinction in Russian

#### 4.2.1.1 Preliminary data on Russian subjunctives

Russian subjunctive clauses are introduced by the complementizer *čtoby*. The verb in the subjunctive clause is morphologically in the past tense, and no other verbal forms are allowed, as shown in the example (12):

- (12) a. Ivan xočet čtoby Maša pročitala/čitala LGB.  
Ivan wants that<sub>SUBJ</sub> Maša read<sub>PST.PERF/PST.IMPERF</sub> LGB  
'Ivan wants Masha to read Lectures on Government and Binding'
- b. \*Ivan xočet čtoby Maša čitaet/pročitaet/budet čitat' LGB.  
Ivan wants that<sub>SUBJ</sub> Maš read<sub>PRES/FUT.PERF/FUT.IMPERF</sub> LGB

On the contrary, Russian indicative clauses are introduced by the complementizer *čto*. There is no restriction on the morphology/tense of the verb in indicative clauses, (13):

- (13) a. Ivan skazal čto Maša pročitala/čitala LGB.  
Ivan said that Maša read<sub>PST.PERF/PST.IMPERF</sub> LGB  
'Ivan said that Masha have read/was reading LGB'
- b. Ivan skazal čto Maša čitaet/ budet čitat' LGB.  
Ivan said that Maša read<sub>PRES/FUT.IMPERF</sub> LGB  
'Ivan said that Masha is reading/will be reading LGB'

Despite the fact that the verb in the embedded subjunctive clause is morphologically in the past form, the event denoted by embedded clause is not necessarily situated in the past, either with respect to the event in the matrix clause, or with respect to the speech act. On the contrary, the event described in the embedded clause (a reading of “Lectures on Government and Binding” in (12-a)) is *irrealis* and might happen in the future with respect to the time of the event described in the matrix clause (the volition act in (12-a)). Notice, that in addition to Russian, a number of languages do not have a separate morphological category for subjunctive. For instance, in Serbo-Croatian, verbs in subjunctive clauses are morphologically present, while in Bangla they are marked as imperatives. At the same time, in some Romance languages (Spanish, Italian, and French), the subjunctive is a separate form of the verb, distinct from the past. I propose that the fact that the subjunctive form of the verb is identical to the past tense form in Russian is just an idiosyncrasy.

I assume that even though the verb in the subjunctive clauses is morphologically past, it bears different temporal features: from the data presented and the semantic interpretation of subjunctive sentences I conclude that the subjunctive form of the verb bears an unvalued T feature, unlike verbs in other finite forms (for example, past).

- (14) a. Ivan    xočet čtoby    Maša    pocelovala Borisa                      (Subjunctive)  
       Ivan<sub>NOM</sub> wants that<sub>SUBJ</sub> Maša<sub>NOM</sub> kiss<sub>SUBJ</sub> Boris<sub>ACC</sub>  
       ‘Ivan wants Masha to kiss Boris’
- b. Ivan    skazal čto    Maša    pocelovala Borisa                      (Indicative)  
       Ivan<sub>NOM</sub> said    that Maša<sub>NOM</sub> kiss<sub>SPST.PERF</sub> Boris<sub>ACC</sub>  
       ‘Ivan said that Masha kissed Boris’

That means that in the two sentences in (14) the verb *pročitala* ‘read’ comes from the lexicon embedded with different features (even though those two forms are morphologically indistinguishable): in sentence (14-a) it bears a  $\langle uT -val \rangle$  feature, whereas in the sentence (14-a) it bears a valued instance of the T feature  $\langle uT +val \rangle$ .

This proposal about the subjunctive vs. indicative clauses is summarized below in (15) in what I will call the *Subjunctive parameter*:

(15) *The Subjunctive parameter* (cf. Picallo 1985)

- a. (In Russian,) the subjunctive form of the verb bears a  $\langle uT -val \rangle$  feature;
- b. (In Russian,) finite forms of the verb bear  $\langle uT +val \rangle$  feature.

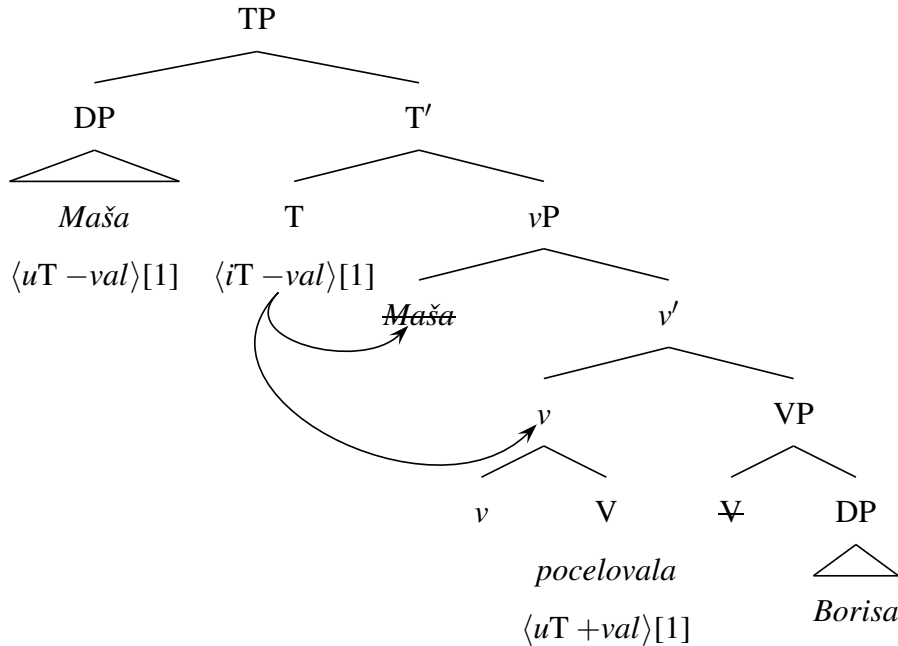
In what follows I will show how the derivation of indicative clauses works in Russian, and then proceed to the subjunctive. I assume the Pesetsky and Torrego 2001, 2007 framework.

#### 4.2.1.2 Indicative clauses

Recall that by the Subjunctive parameter (15), indicative verbs have  $\langle uT +val \rangle$  T-feature. The derivation proceeds in a standard bottom-up way. The verbal projection vP is built in a standard manner with V adjoining to v. If the embedded clause of the sentence is indicative, as in (14-b), after T is merged into the tree structure, its interpretable but unvalued feature  $\langle iT -val \rangle$  probes to find its goal, finding it in the  $\langle uT -val \rangle$  feature on the subject DP (assuming that Nominative case is an instantiation of the T-feature on D, as in Pesetsky and Torrego 2001). After the Agree operation takes place, the features on T and the subject D are linked, and become instances of the same feature. However, since the subject DP's T-feature is unvalued, the shared T feature also remains unvalued. However, because only valued features can be interpreted, T must probe further down in the tree in order to find a value. The second probing finds a goal  $\langle uT +val \rangle$  on the finite verb within the vP projection. After the Agree operation, all three T-features — those on T, the subject DP and V become instances of the same feature, and the valuation of the  $\langle iT -val \rangle$  on T takes place, resulting in the valuation of  $\langle uT -val \rangle$  on subject DP also. After this step, all T-features in the embedded clause are valued. The EPP on T, will be satisfied by further raising of the embedded subject to Spec,TP. Now, there are no unvalued features left in the indicative embedded clause, and its derivation can stop. The resulting structure of the embedded TP before

final valuation takes place is given in (16).

(16) *Feature make-up of the embedded indicative clause in (14-b)*



Notice that there is no movement to the CP domain, as nothing in the CP domain will be able to probe and attract a goal. After merging the complementizer *čto*, the lower CP phase is completed with no elements but the complementizer at its edge.

#### 4.2.1.3 Subjunctive clauses

In the case of Russian subjunctive clauses, applying the analysis proposed above gives surprisingly different results. Following the proposed Subjunctive Parameter (15), the subjunctive verb comes from the lexicon with an unvalued T feature  $\langle uT -val \rangle$ . This contrasts with the verbs in indicative clauses, which enter the numeration with valued T features. Also, I would assume the presence of *čtoby* in the numeration for selectional purposes (I would claim that volitional predicates, such as *xotet* 'to want', select CPs headed by *čtoby*). Therefore, if *čtoby* is not present in the numeration, the derivation will crash.). I assume that *čtoby* also comes from the lexicon endowed with an

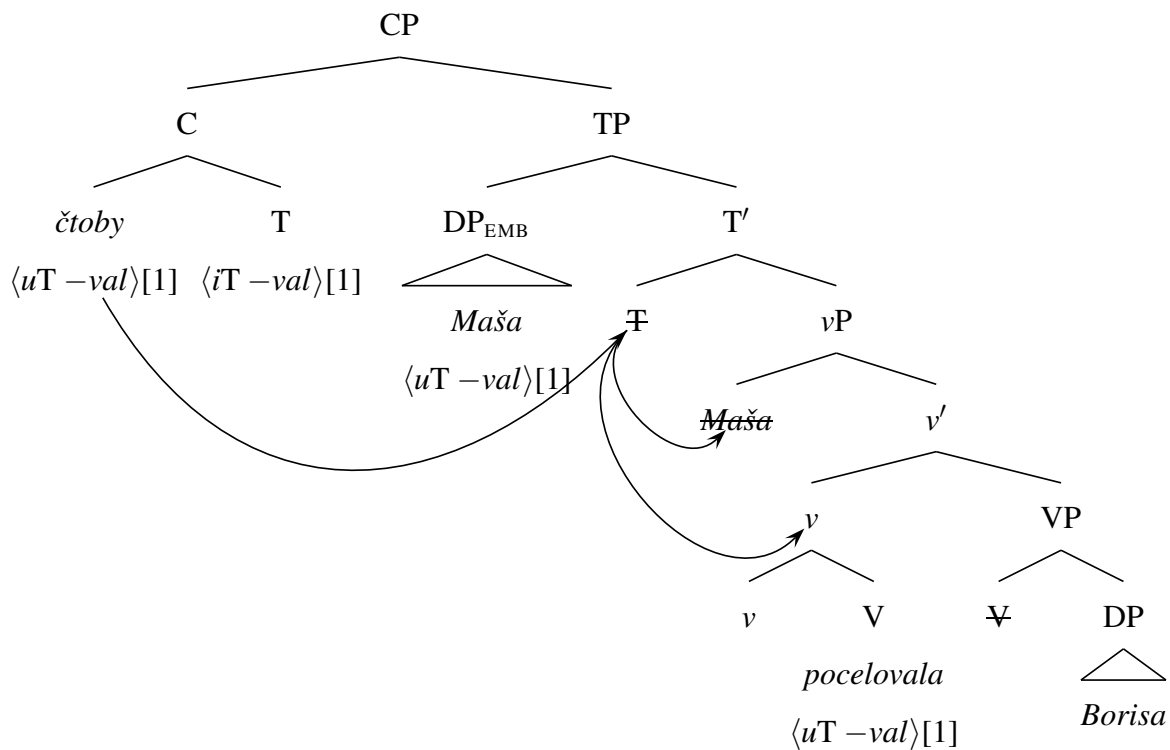
uninterpretable unvalued  $\langle uT -val \rangle$  feature. Now let us consider the derivation of subjunctive clauses.

The embedded  $\nu P$  is built in standard fashion. After that T is merged into the structure. In a similar way to the case of indicative clauses, the embedded T probes and Agrees first with the subject DP, and then with the verb (to be more precise, the  $\nu+V$  complex), resulting in feature sharing among all these elements, making the T-features on T, the subject DP and  $\nu+V$  all instances of the same feature. However, unlike in the case of indicative clauses, no valuation can occur at this point, since the T-feature on the embedded subjunctive verb is not valued. Therefore the derivation proceeds by the merging of *čtoby* in the C-head position.

The T-feature of *čtoby* is unvalued, and therefore must probe down to find its goal. The first goal it finds is a T with an unvalued T-feature. Agree takes place, and the instances of the T-feature on *čtoby*, on T, on the embedded subject, and on the embedded verbal complex become instances of the same feature. Further, the featural bundle created in T adjoins to *čtoby*, and the resulting configuration from the completion of the embedded CP-phase is given in (17), where the index [1] shows which T-features are instances of the same feature, and  $DP_{EMB}$  is the subject of the embedded clause.



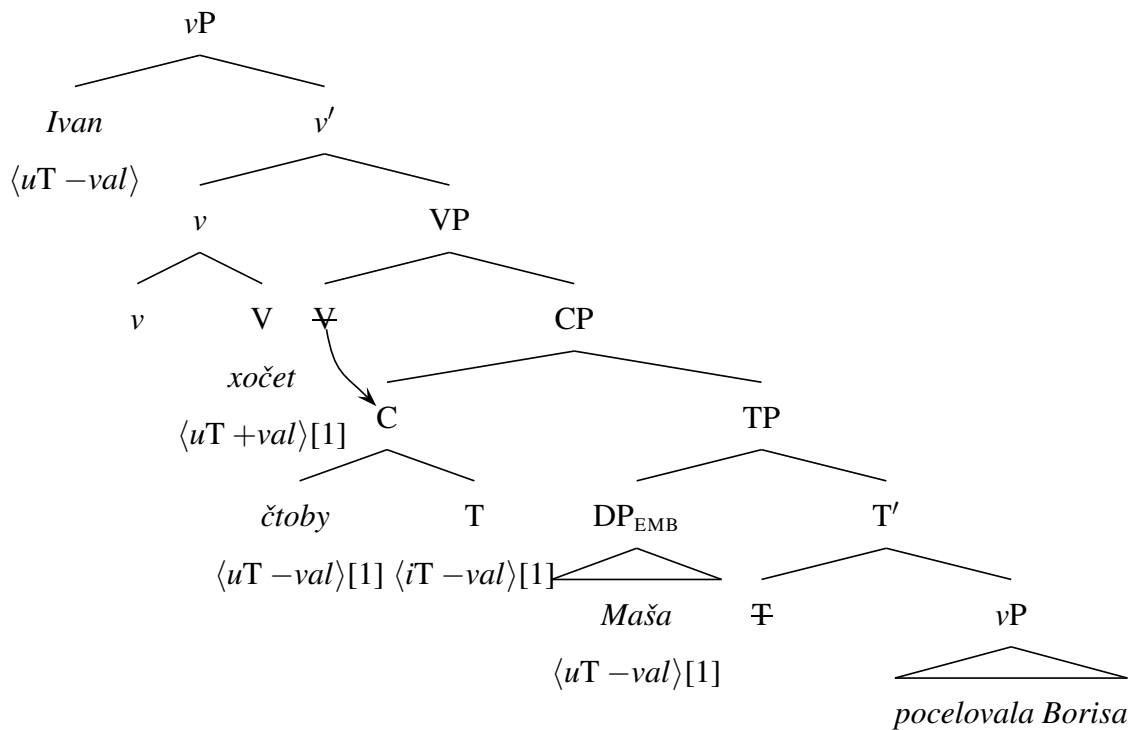
(17) The structure of the embedded subjunctive CP of (14-a)



Crucially, even though there are unvalued features by the end of the derivation of the embedded clause, the derivation does not crash, since the unvalued T-feature was able to move to the edge of CP to the C-head position, and therefore will remain accessible for further Agree relations with the probe from the higher domain.

Next, the elements of the matrix clause are merged in the structure:  $V/v$  with the  $\langle uT +val \rangle$  (since the matrix verb is finite), and matrix subject  $DP_{MATR}$  with the instance of  $\langle uT -val \rangle$ . Recall that V in subjunctive constructions selects a CP headed by *čtoby*. This selectional property would result in an Agree relation between the matrix V and the embedded C. As a byproduct of this selectional relation, the T-features on the matrix verb and on the head-complex in C of the embedded clause will be shared, resulting in valuation of the T-features on the elements of the embedded clause.

(18) *The structure of the matrix vP and the embedded subjunctive CP of (14-a)*

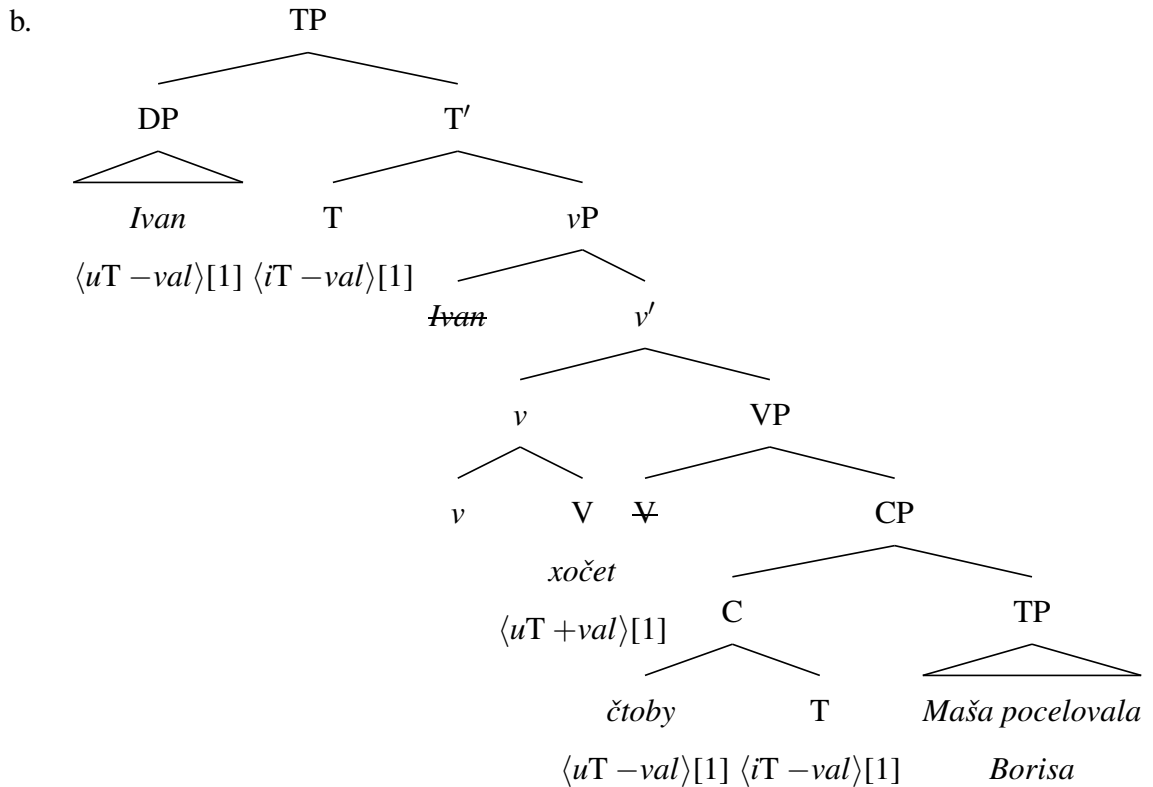


At the next stage, the matrix T, endowed with  $\langle iT -val \rangle$  feature, is merged into the structure. Since it is an interpretable feature, it probes down, finding the T-feature of the matrix subject and Agrees with it, resulting in a shared feature between it and the matrix subject  $DP_{MATR}$ . Further, since the T-feature of the matrix T is still unvalued (as none of the elements with which it has agreed have provided it with a value), it probes down one more time and finds the matrix  $v+V$  complex as a goal. The Agree operation at this stage makes all the T-features on the matrix and embedded Vs, and the T-features in the featural complex located in the matrix  $v$ -head position instances of the same feature, and values them, acquiring the value from the  $\langle uT +val \rangle$  matrix verb.

After this crucial step, all T-features introduced so far in both matrix and embedded clauses are instances of the same T-feature, and all of them become valued. (19) presents the sentence with the subjunctive embedded clause, and gives the tree before the final valuation has taken place:

(19) *The final structure of the sentence with a subjunctive embedded clause.*

- a. Ivan    *xočet*    *čtoby*    Maša    *pocelovala*    Boris  
 Ivan<sub>NOM</sub>    wants<sub>T</sub>    that<sub>SUBJ</sub>    Maša<sub>NOM</sub>    kiss<sub>SUBJ</sub>    Boris<sub>ACC</sub>  
 ‘Ivan wants Masha to kiss Boris’



#### 4.2.2 Deriving infinitival clauses

Pesetsky and Torrego 2007 describe how raising clauses are derived in English. I briefly outline the derivation below, and further propose the possible derivation of control clauses.

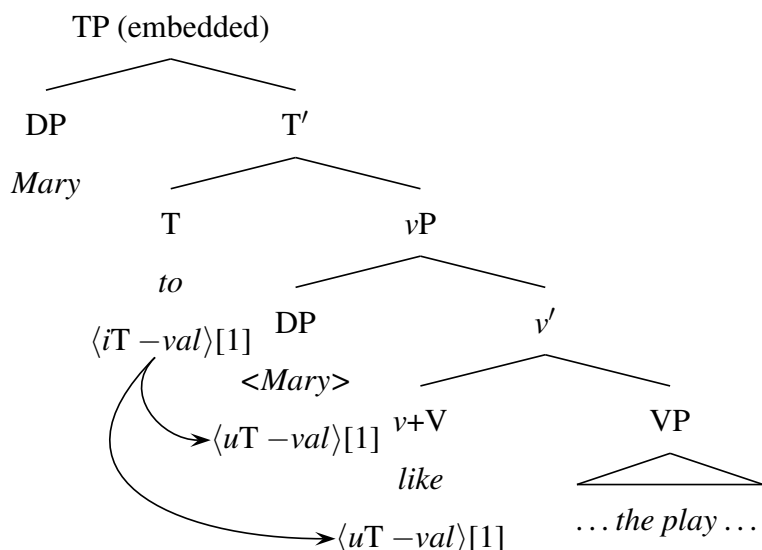
Consider the sentence in (20).

- (20) Mary seemed to like the play. (Pesetsky and Torrego 2007, pg. 275)

Pesetsky and Torrego assume that infinitival V comes from the lexicon with unvalued uninterpretable T-feature,  $\langle uT -val \rangle$ , and therefore Agree relationship between T and V+v in the em-

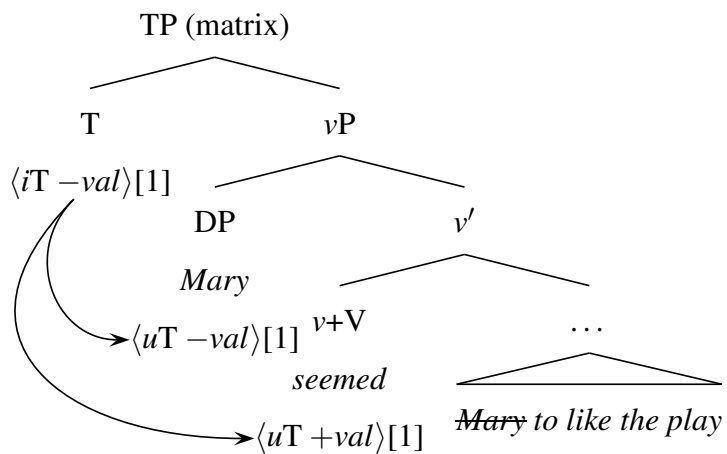
bedded clause would not value the T-feature on T. Further, *Mary* in (20) comes with an unvalued uninterpretable T-feature as well (nominative case). The probing relation is established first between T and the subject, and further between T and *v+V*. While the T-feature is getting shared between all three of these locations, in no location is it valued. Schematically this is shown in (21). The embedded subject further raises to the embedded Spec,TP.

(21) *The structure of raising complements according to Pesetsky and Torrego 2007.*



Since none of the T-features in the embedded clause are currently valued, the valuation must be achieved through some other process. Pesetsky and Torrego propose that raising of the subject DP is exactly what is needed for such a valuation. They assume that the embedded subject raises to the matrix Spec,vP position, and after that it can be probed by the matrix T with its  $\langle iT -val \rangle$  feature. Matrix T probes the raised subject, and enters into a feature-sharing relation with it: the instance of T on the embedded elements and on the matrix T become instances of the same feature. The subsequent probing of the finite *v+V* with  $\langle uT +val \rangle$  by T would result in sharing of all T-features (both in matrix and embedded clauses), and their valuation, the value being acquired from the matrix verb. Schematically this is illustrated in (22).

(22) *The structure of the raising sentence in (20) according to Pesetsky and Torrego 2007.*



At this moment we can compare the derivation of raising constructions by Pesetsky and Torrego 2007 outlined above with the proposed derivation of subjunctive complements from the previous section. The controversy here is at which moment of the derivation the T-features in the embedded clause are valued. According to Pesetsky and Torrego, in raising constructions T-feature valuation is a result of the embedded subject raising to the matrix clause. Its position between matrix T and matrix  $v$  allows it to get probed by the matrix T before the T-feature is valued as a result of subsequent probing of the matrix  $v$ . On the other hand, in the analysis of subjunctives, proposed here, the valuation of T-features in the embedded clause is the result of selection: matrix V selects a proper C-complement, and this selection results in valuation of the T-feature on the embedded C as well as on other embedded elements: T,  $v$ , and the embedded subject:

- (23) a. *Raising constructions*: subject raising leads to T-feature valuation.  
 b. *Subjunctive constructions*: selection by the matrix V leads to T-feature valuation.

The problem arises, which can be looked at in two possible ways. First, one may ask why we don't need raising of the subject in subjunctive constructions. According to Pesetsky and Torrego 2007, the embedded subject must end up between matrix T and  $v$  to get its T-feature valued. However, we may look at this problem in a different way. Why can feature-valuation based on selection

(as I proposed for subjunctive complements) not be applied to raising constructions? The embedded subject in subjunctive complements stays in Spec,TP, and still gets its T-feature valued (gets its nominative case), without any need to raise into the matrix clause. If in raising complements there is a selectional dependency between the matrix verb and embedded TP, why doesn't this dependency result in embedded T-feature valuation, just as in the case of subjunctive complements, without the need for the embedded subject to serve as an intermediary? To summarize the discussion above, in order to reconcile the two analyses proposed above, one has to answer the following two questions:

- (24) a. If in subjunctive complements T-valuation is achieved without long-distance movement, why can't it be achieved in the same way in raising complements?
- b. If in raising complements long-distance subject movement is necessary to achieve T-valuation, why is it not necessary in subjunctive complements?

Keeping these two questions in mind, I will consider how control complements are derived within the same framework (Pesetsky and Torrego 2007). The major difference between control and raising constructions is the presence of PRO in the embedded clause. I explore the featural content of PRO, and propose how the derivation proceeds and how the T-feature on the embedded T/v/V gets valued from the matrix clause  $v+V$  without resorting to subject raising.

The main difference between the raising case and control case lies in fact that no overt nominal starts as the subject of the embedded clause; the subject of the embedded clause is assumed to be PRO. Further, it is traditionally assumed that raising complements lack a CP layer, while CP is present in control complements (Adger 2004).

- (25) a. Raising: [TP ... [vP ... ]]
- b. Control: [CP ... [TP PRO ... [vP ... ]]]

Now, PRO has been claimed to be an element that does not require (nominative) case. In the

adopted framework this translates to the absence of a T-feature on PRO. A further property of PRO is that it needs to be coreferent with a c-commanding element higher in the structure<sup>1</sup>.

The next question one has to answer is what the  $\phi$ -feature specification of PRO is. Applying similar logic to PRO as I did above for SE-type anaphors, I claim that PRO comes from the numeration with unvalued  $\phi$ -features. I claim that they must be valued in syntax upon establishing coreference with a c-commanding element, and this cannot happen before the referent of PRO has been merged into the structure. I leave the question on how this  $\phi$ -feature valuation is established for further research, but the process must be analogous to what I have claimed for reflexives above in Chapter 3.

To summarize, the following three facts distinguish control constructions from raising and from subjunctive constructions:

- (26) a. The absence of a T-feature on the embedded subject (PRO) in contrast with subjunctives and raising constructions.
- b. Unvalued  $\phi$ -features on the embedded subject (PRO) in contrast with subjunctives and raising constructions.
- c. The presence of a CP layer in contrast with raising constructions.

With these facts in mind, we are now in a position to describe how control clauses are derived in English. Consider example (27).

---

<sup>1</sup>I will not touch on the question of so called “arbitrary” PRO, PRO<sub>arb</sub>. The crucial fact here is that it still lacks a T-feature. Notice however that the binding properties of PRO<sub>arb</sub> are different, i.e. *himself* cannot be bound by it:

- (i) a. John<sub>i</sub> likes [ PRO<sub>i</sub> to praise Putin ].
- b. John<sub>i</sub> likes [ PRO<sub>i</sub> to praise himself ].
- (ii) a. [ PRO<sub>arb</sub> to praise Putin ] is inappropriate.
- b. \*[ PRO<sub>arb</sub> to praise himself ] is inappropriate.

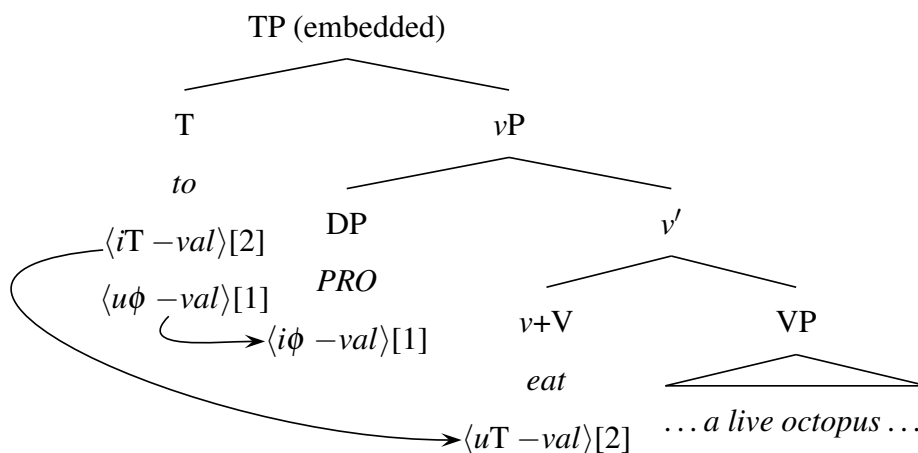
PRO<sub>arb</sub> requires a reflexive unspecified for  $\phi$ -feature, such as *oneself*:

- (iii) [ PRO<sub>arb</sub> to praise oneself ] is inappropriate.

(27) John tried to eat a live octopus.

Embedded  $vP$  is generated in a way similar to raising constructions, PRO starts off at a Spec, $vP$  position, and the infinitival verb is merged with an unvalued uninterpretable T-feature. T further probes PRO in order to establish a sharing relation with its  $\phi$ -features. Unlike in the case of subjunctive complements,  $\phi$ -feature valuation on T does not happen, since PRO lacks them at this moment. It is important to notice that this probing is crucially not triggered by an unvalued T-feature on T. T further probes  $v+V$ , and T-features on T and T-features on  $v+V$  become shared; no valuation occurs at this moment. PRO further raises to Spec,TP.

(28) *The structure of the embedded TP of (27)*



Now, C is merged with TP. Similarly to subjunctive clauses where the tense of the embedded clause depends on the tense of the matrix clause, I propose that C must serve as an intermediary in tense transfer. Formally, I claim that infinitival control C comes with an unvalued uninterpretable T-feature,  $\langle uT -val \rangle$ . C enters into an Agree relation with the infinitival T, and shares its so far unvalued T-features. Assuming that no DP raising takes place in control construction, the remaining part of the analysis will be parallel to the analysis of the subjunctive complements.<sup>2</sup> Matrix  $V/v$  selects for the control C. Similarly to the subjunctive case, this selection results in the

<sup>2</sup>I put aside the possible analysis of control as movement, as in Hornstein 2001. In fact, it is not incompatible with the reconciliation of my analysis of control and raising complements, which I present below.





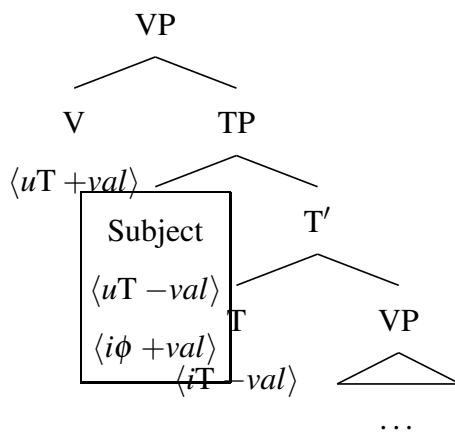
### 4.2.3 Reconciling the analyses

Having shown how the derivation of control complements in English proceeds, and its parallels to Russian subjunctive clauses, I will now try to resolve the problem demonstrated above. The main question was: how to reconcile the analyses of raising and subjunctive clauses? Now, we can add control to the mix:

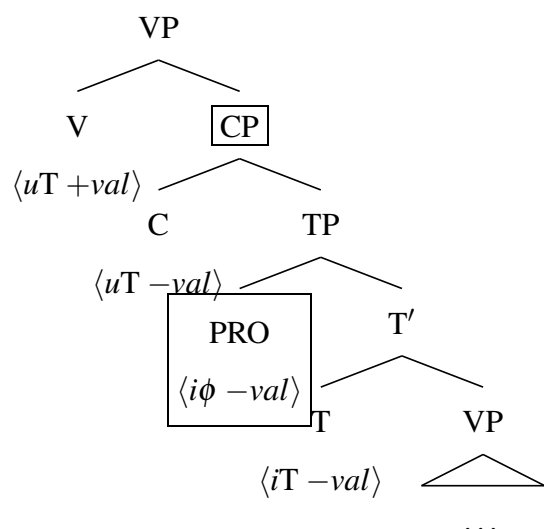
- (30) a. If in subjunctive and control complements T-valuation is achieved without long-distance movement, why can't it be achieved in the same way in raising complements?
- b. If in raising complements long-distance subject movement is necessary to achieve T-valuation, why is it not necessary in subjunctive and control complements, and in case of subjunctive complements even impossible?

Consider the structure of the three constructions mentioned above.

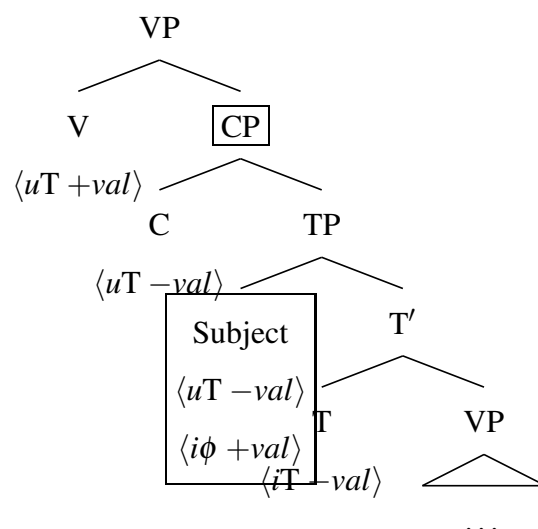
- (31) *Raising constructions*



(32) *Control constructions*



(33) *Subjunctive constructions*



The following facts are true about these constructions, and are important for reconciling the analyses from the previous sections. Raising complements lack a CP layer, but have a T-feature present on their subjects. Subjunctive complements have a CP layer, as well as a T-feature on their subjects. Control complements are the opposite of raising complements: they possess a CP layer, but lack a T-feature on their subjects. The table in (34) summarizes these differences:

(34) Summary of facts<sup>3</sup>:

	Raising	Control	Subjunctives
CP present	✗	✓	✓
T-feature on the subject	✓	✗	✓
φ-feature on the subject	<i>valued</i>	<i>unvalued</i>	<i>valued</i>

Compare these three structures. How does one enforce raising of the subject for Raising complements and disallow it for Subjunctives and Control complements? Notice, that only in Raising constructions is there no intervening projection between the probe, V, and the embedded subject.

<sup>3</sup>The remaining possibility — absence of CP, absence of T-features on the subject — is an ECM construction. However, it is irrelevant for current discussion.

Both subjunctive and control complements have a CP layer headed by C with a T-feature, and this head is closer to V than the embedded subject. On the contrary, in raising construction the subject and embedded T are located in the projection immediately dominated by a matrix VP.

Recall that I assumed that in control and subjunctive complements the valuation of T-feature in the embedded clause is a byproduct of selectional relations. V selects a CP headed by C with a T-feature, and this selectional relation triggers the valuation of the T-feature in the embedded clause, even though in this case the valued instance on the T-feature (on V) is located higher than the unvalued T-features (on the embedded C). While this piggy-backing valuation of the T-feature occurs in a non-standard configuration, it cannot be excluded without a stipulation. In fact, Chomsky 2001 assumes that nominative case valuation on a subject is a byproduct of  $\phi$ -feature checking by T. T's  $\phi$ -features search for a goal, in this case the subject. The subject is "active", since it has an unvalued case feature. The probing happens, and  $\phi$ -features on T get valued, while the Case-feature on the subject is checked piggybacking on this  $\phi$ -feature valuation. Similarly in the case of control/subjunctive complements, matrix V enters in a probe-goal relation with the embedded C (presumably to value its selectional features<sup>4</sup>), and the established relation has T-feature valuation as a side-effect.

Based on the discussion above, it is natural to assume that V-T relationship in raising constructions can also value T-features in the embedded clause, and subject raising is not necessary for that purpose, contra Pesetsky and Torrego 2007. Now, the remaining question is:

- (35) If raising of the embedded subject is not necessary to value T-features within the embedded clause, why does the embedded subject raise in Raising constructions, and stay *in situ* (within the embedded clause) in subjunctives and control complements?

To provide an answer to this question I will argue that raising of the embedded subject to the matrix clause satisfies the EPP property of  $v/V$ . Below, I explore this process in detail.

---

<sup>4</sup>For the purposes of current discussion, I will not discuss how selection works. This is a large topic worthy of a separate research project. However, I assume that selection is a result of probe-goal relation: categorial selectional features on a probe must match the categorial selectional features of a goal.

Assume that  $v/V$  has an EPP subfeature on its T-feature. Consider how this EPP feature can be satisfied in the cases considered above. Consider first the case of control complements. Embedded subject (PRO) lacks a T-feature, and therefore cannot satisfy the EPP subfeature of T on  $V/v$ . The only possible way to satisfy the EPP in this case is to merge a new item from the enumeration, i.e. the matrix subject. No PRO raising can apply in control complements precisely because of the lack of T-feature on PRO.

The second case to consider is the case of subjunctive complements. As opposed to control constructions, the embedded subject of a subjunctive complement has a T-feature. Thus, the structure above presents two potential options to satisfy the EPP of  $v/V$ : by Merge or by Move. If we are to satisfy the EPP subfeature of T by Move, the embedded subject must raise to the  $\text{Spec},vP/\text{Spec},VP$ . Since CP is an intervening category between  $v/V$  and the embedded subject, satisfaction of the EPP will require two steps: first, moving the subject to the  $\text{Spec},CP$ , and second, moving the subject to the desired specifier of the verb phrase. The other option is to satisfy the EPP by Merge. The matrix subject with an unvalued uninterpretable T-feature can be merged into structure. Which option is preferred: one instance of Merge or two instances of Move? Regardless of whether one adopts the Merge-over-Move paradigm (Chomsky 1995), or Move-over-Merge paradigm (Shima 2000), the most economical choice in this case is one instance of Merge. Economy considerations provide us with an explanation of why raising of the embedded subject does not happen in subjunctive complements.

The last case is the case of raising complements. Here, the embedded subjects comes with a T-feature, and therefore has potential to satisfy the EPP of  $v/V$  by Move. In this case, we are considering only one instance of Move, as there is no intervening projection between the source (TP) and the goal. However, there is also an option to satisfy the EPP by Merge, say with an expletive.

Chomsky 1995, pg. 348 argues that Merge is preferred to Move: “we select Merge over Attract/Move if that yields a convergent derivation, irrespective of consequences down the road as long as the derivation converges.” This assumption however is not uncontroversial.

Shima 2000 argues that “Move is selected over Merge.” He goes on to argue that

Merge forms the new object by concatenating two objects that are separate phrase structures; Move forms the new object by concatenating two objects that are in a single phrase structure. In the process of forming phrase structures, it is more economical to look only at an already formed structure than to look at, not only an existing structure, but also lexical items in the numeration, or at an independent syntactic object.

In fact, Shima’s arguments are reminiscent of approach taken in Chomsky 2008. He views Move as an instance of Merge (Chomsky 2008, pg. 140):

Suppose that X and Y are merged (for expository purposes, think of Y as merged to X). Either Y is not part of X (*external Merge*, EM) or Y is part of X (*internal Merge*, IM). In both cases, Merge yields {X, Y}. IM yields two copies of Y in X, Y, one external to X, one within X. IM is the operation *Move* under the “copy theory of movement,” which is the null hypothesis in this framework, required by strict adherence to the NTC [“no-tampering condition”]. Unless there is some stipulation to the contrary, which would require sufficient empirical evidence, both kinds of Merge are available for FL [faculty of language] and IM creates copies.

Based on this discussion, I argue that in the case of raising complements moving the embedded subject to satisfy the EPP subfeature of T on  $v/V$  is in fact preferred to merging an expletive. The necessity of subject movement argued for by Pesetsky and Torrego is therefore derived. However, it is not required in order to value embedded T-features; it is required in order to satisfy the EPP-feature of the matrix  $v/V$ , and it is more economical than expletive merge. While Pesetsky and Torrego 2007 are correct in terms of which elements undergo movement, their motivation for these movements is not.

With that in mind, the derivation of the raising complements would be much more similar to the derivation of control and subjunctive complements: the embedded T-features are valued in a same

manner, as a byproduct of selectional relations, while the subject raising satisfies EPP subfeature of T on V.

#### 4.2.4 Tense dependency and long-distance binding

Now with the theory of how subjunctive and infinitival clauses are derived, we can proceed to explaining the binding properties. I will first consider the infinitival cases, i.e. binding of the anaphor within the control complements, and then shift to anaphoric binding into subjunctives.<sup>5</sup>

##### 4.2.4.1 Anaphors within infinitival clauses

The example of anaphoric binding with object control verbs is given below in (36) and (37). The difference between Russian and English is in the availability of binding by the matrix subject. In English, (36), the only possible binder of the embedded object anaphor is PRO. In object control cases it is bound by the matrix object, while in subject control cases, the matrix subject is the only possible coreferent for the reflexive. In Russian, in object control cases, both options are possible: the *se*-reflexive within the embedded clause can be bound by either the matrix object/PRO (giving a reading that Boris has to draw himself), or by the matrix subject (reading when Boris has to draw Ivan). The availability of the latter reading is the crucial distinction between Russian and English, and is a case of “proper” long-distance binding. While the cases of binding by the matrix object in object control cases are unsurprising, as one can claim that the reflexive is actually bound by the PRO, the matrix subject antecedent is problematic for a locality-based theory of binding. At no point in the derivation is reflexive *sebja* in a local configuration with the its antecedent.

- (36) a. \*John<sub>i</sub> told Mary<sub>j</sub> [PRO<sub>j</sub> to draw himself<sub>i</sub>]  
b. John<sub>i</sub> told Mary<sub>j</sub> [PRO<sub>j</sub> to draw herself<sub>j</sub>]  
c. John<sub>i</sub> tried [PRO<sub>i</sub> to draw himself<sub>i</sub>]

---

<sup>5</sup>The analysis proposed in this section relies on feature-movement, and in this it is similar to the analysis of Russian long-distance binding by Rudnitskaya 2000.

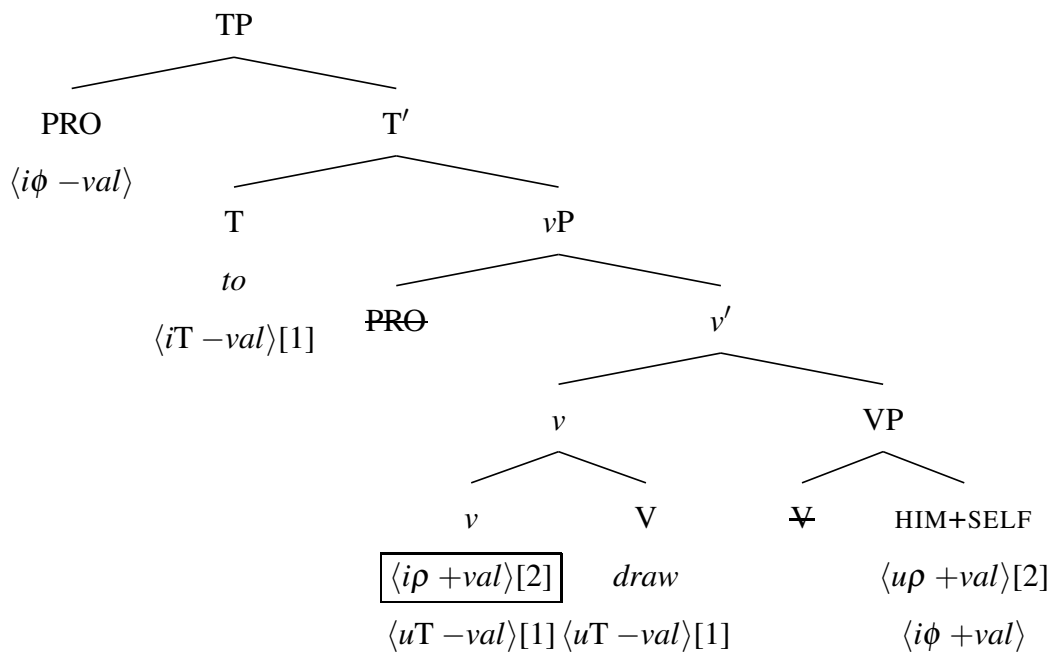
- (37) a. Ivan<sub>i</sub> poprosil Boris<sub>j</sub> [PRO<sub>j</sub> narisovat' sebja<sub>i/j</sub>] (Russian)  
 Ivan asked Boris draw<sub>INF</sub> self  
 'Ivan<sub>i</sub> told Boris<sub>j</sub> to draw himself<sub>i/j</sub>'
- b. Ivan<sub>i</sub> pytaetsja [PRO<sub>i</sub> narisovat' sebja<sub>i</sub>]  
 Ivan tries draw<sub>INF</sub> self  
 'Ivan is trying to draw himself'.

In what follows I show that the analysis of control clauses outlined above together with the featural approach to binding developed in Chapter 3 makes correct predictions in these cases, and the difference between Russian and English follows naturally from the properties of the corresponding reflexives, i.e. the presence or valued vs. unvalued  $\phi$ -features.

Consider first the case of English. Recall, that the English reflexive comes with valued  $\phi$ -features. Further, I argued above that PRO comes with unvalued  $\phi$ -features, which must be valued only after the insertion of a controller. In order to derive the local binding by PRO, no extra assumptions are needed, in fact the derivation will be parallel to that of local binding outlined in Chapter 3. The  $\rho$ -feature can be placed on  $v$ . It probes the reflexive, and their  $\rho$ -features become shared, and the  $\rho$ -feature on  $v$  gets its value from the reflexive. Now, PRO is inserted into Spec, $v$ P position, and at this moment its  $\phi$ -features are unvalued. Further, T is merged with  $v$ P. T probes PRO for  $\phi$ -features, which become shared between T and PRO. Two things are of importance: first, this probing does not result in  $\phi$ -feature valuation, and second, T probes PRO only for  $\phi$ -features, not for T-features, as PRO lacks them. T probes  $v+V$ , and their T-features are shared, though no valuation happens since  $v+V$  does not have a value for its T-feature. PRO is further moved to the Spec,TP position. The resulting configuration is shown in (38).



(38) The structure of the English control TP of (36-c)

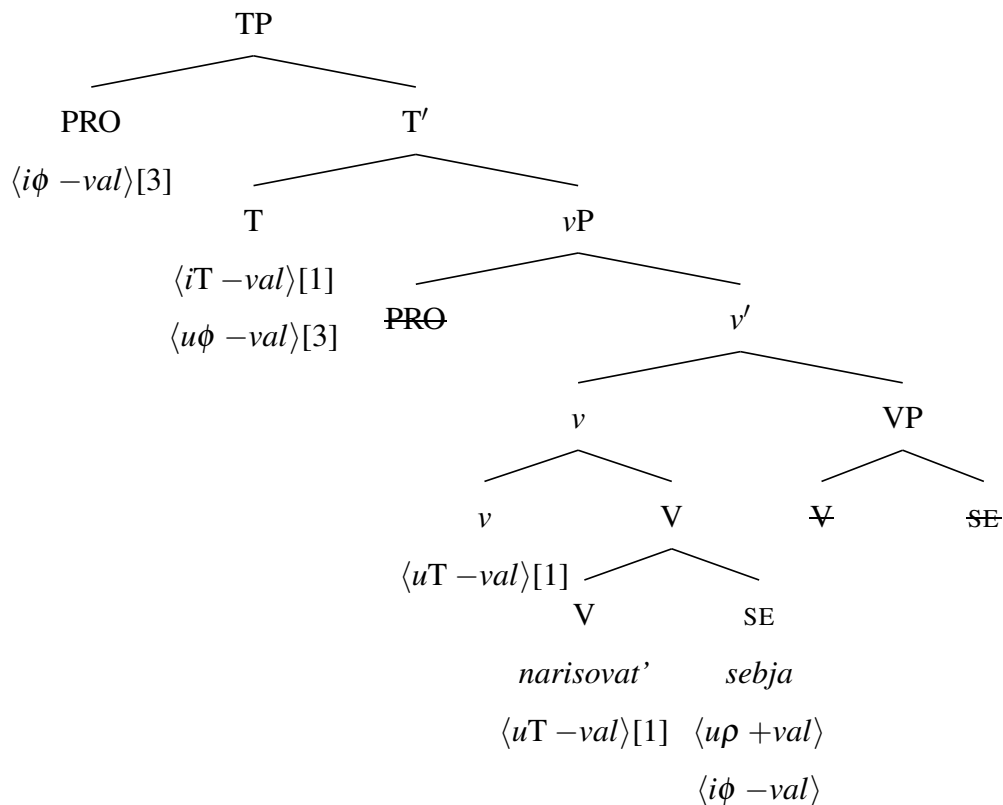


At this level of the derivation, we can consider vP. There are no features left within vP that are either unvalued or not shared with some element outside of vP: the T-feature on v/V is shared with the T-feature on T, the  $\rho$ -features are valued, and the  $\phi$ -feature on the reflexive is also valued; PRO with its unvalued  $\phi$ -feature moved out to Spec,TP. That means that vP can be closed for further probing. A  $\lambda$ -operator will be inserted in place of  $\rho$ -feature, and the result of the subsequent  $\lambda$ -conversion will be binding of the embedded reflexive by PRO.

Now, crucially, if one places an interpretable unvalued instance of a  $\rho$ -feature on some higher head, say C or matrix v, the derivation will crash. The difference between this case and the structure in (38) will be the absence of the  $\rho$ -feature on the embedded v (boxed in (38)). That however will not affect the status of vP: all relevant features within it will still be either valued, or shared with some elements outside of vP. Thus, if one chooses to place a  $\rho$ -feature on a higher head in hopes of achieving long-distance binding, the derivation will crash, since this  $\rho$ -feature will not be able to establish a probe-goal relationship with the valued instance of  $\rho$ -feature on the reflexive. Thus, the impossibility of long-distance binding in English control clauses is accounted for.

The Russian case is different from English because of the absence of valued  $\phi$ -features on the reflexive. Consider the moment of the derivation when  $vP$  is completed and  $T$  is merged into the structure. Similarly to the derivation above,  $T$  probes the  $PRO$  subject, and their  $\phi$ -features become shared, while still being unvalued, since  $PRO$  lacks a value for its  $\phi$ -features. Next,  $T$  probes  $v+V$  complex, which is triggered by a  $T$ -feature on  $T$ . Their  $T$ -features become shared, and as in the previous case, valuation does not occur, since neither of these instances is valued.  $PRO$  is further raised to the embedded  $Spec,TP$ . This moment of the derivation is shown in (39).

(39) *The structure of the Russian control TP for (37-a)*



Compare this case with the English case above. There is one major difference: in the English case, the  $\phi$ -features on the anaphor are valued, while in the Russian case, the  $\phi$ -features of the anaphor are still unvalued. The process by which the  $\phi$ -features on the SE-reflexive are valued in the simple monoclausal case failed to apply in the case of control complement, because of the

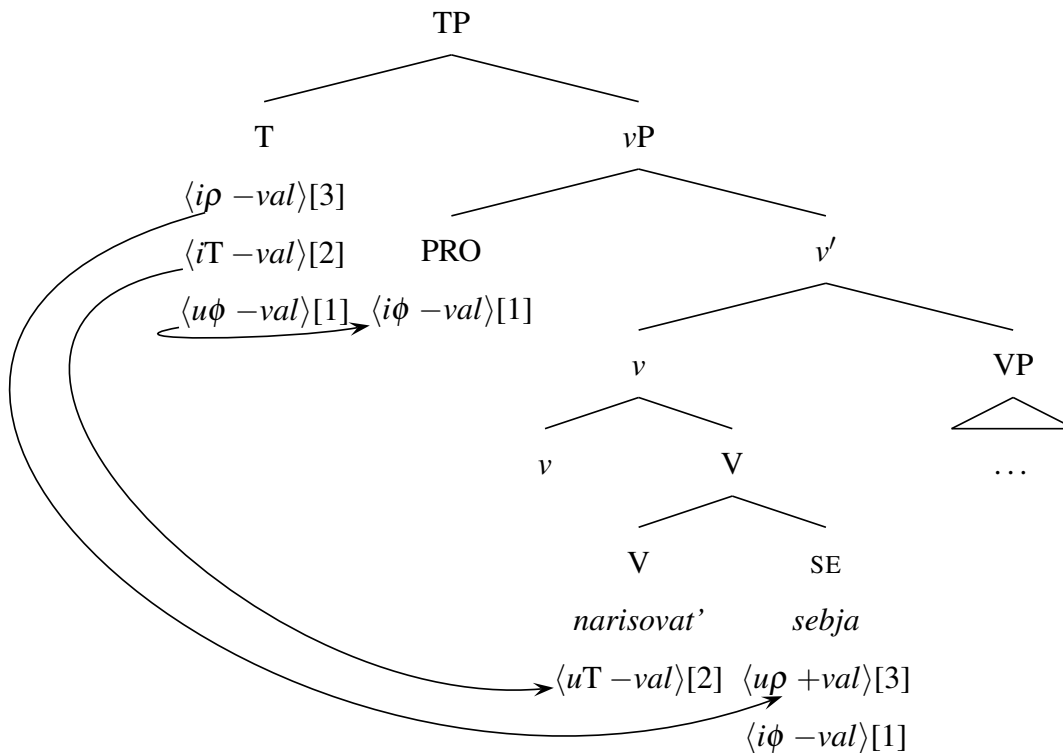
nature of PRO and T. Lack of value of  $\phi$ -features on both T and PRO did not allow subsequent  $\phi$ -features valuation of the SE-reflexive, which in the simple case was piggybacking on probing for  $\rho$ -features.

The presence of unvalued  $\phi$ -features on the reflexive within the  $\nu$ P in the case of SE-type languages will prevent it from closing for further probing, i.e. from acquiring phasal status. Thus, we derived the result that infinitival  $\nu$ P containing a SE-type reflexive in it is not a phase. Potentially, this fact gives us the possibility to bind into  $\nu$ P in the case of control complements in Russian. However, one question which remains to be answered is how  $\phi$ -features on the embedded reflexive get valued at all? Recall that in the feature-based binding framework proposed in Chapter 3, binding cannot apply when there are unvalued features in the structure.

In order to resolve this potential problem, we have to resort to optionality for the position of the  $\rho$ -feature. Since  $\nu$ P in control constructions is not a phase, the  $\rho$ -feature can be placed arbitrarily high, and it will still be able to probe a reflexive within the embedded control  $\nu$ P without violating the PIC. Consider two possibilities.

The first possibility is reminiscent of binding in a simple clause. As I argued in Chapter 3, in SE-languages, the  $\rho$ -feature cannot be placed below T (i.e. on  $\nu$  or V), otherwise there will be no possible derivation which would lead to the valuation of  $\phi$ -features on the reflexive. Assume that an interpretable instance of  $\rho$ -feature is placed on the embedded T. In this case, one more step will be added to the derivation described above: T will probe the reflexive in order to value its  $\rho$ -feature. This probing will result in sharing of the  $\rho$ -feature between T and the reflexive, and will further result in sharing of T's  $\phi$ -feature (already shared with PRO) with the  $\phi$ -feature of the reflexive. The valuation of these  $\phi$ -features will happen immediately after PRO acquires its controller from the matrix clause. By that time, the T-features on the elements within the embedded clause will become valued (value would come from the matrix verb), and binding will be able to apply at this moment. The  $\lambda$ -operator will be inserted above the T-position, and the outcome of this will be local binding of the reflexive by the PRO subject at the level of the embedded TP, similar to the case of a simple clause.

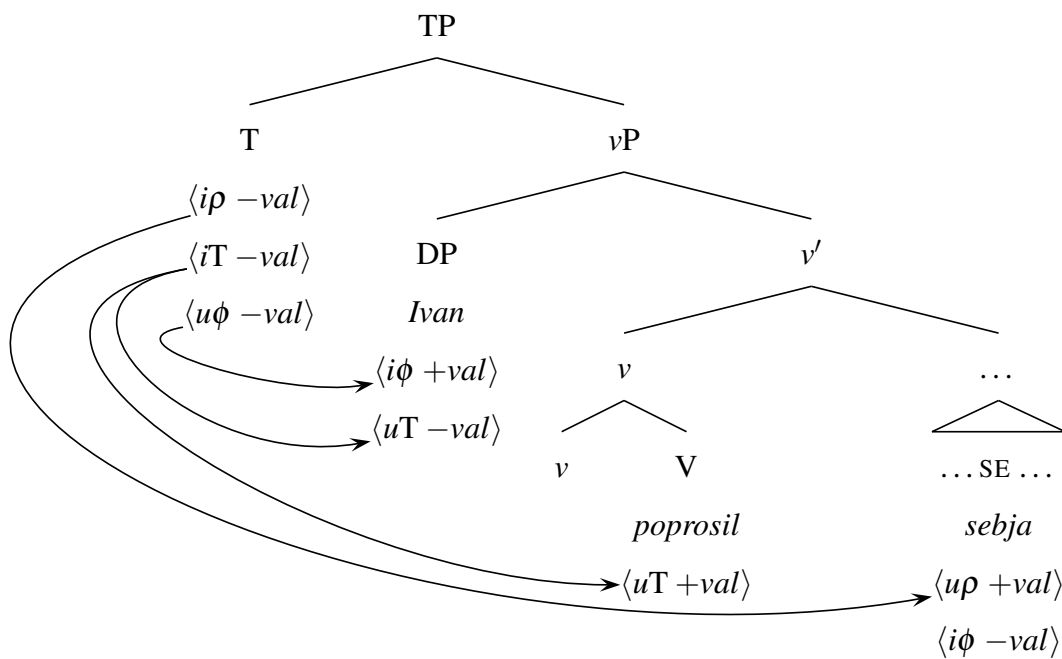
- (40) *Feature valuation in the Russian control TP in (37-a), the interpretable  $\rho$ -feature is on the embedded T.*



Now imagine the second possibility when an interpretable unvalued instance of  $\rho$ -feature is placed on the matrix T. Apart from the  $\rho$ -feature, T comes with  $\langle iT -val \rangle$  and  $\langle u\phi -val \rangle$ . Consider the moment of the derivation when the matrix T is merged. As usual, it will first probe the subject in Spec,vP, and establish a sharing relationship between  $\phi$ -features, and will also share T-features with it. The second step is probing of the matrix  $v+V$ . This step results in valuation of T-features on both T and the subject. Previously, probing by T would stop here, as there were no other features left on T. In the current situation, however, there is one more probing left to do: probing in order to value the  $\rho$ -feature. The only goal for this probing is the reflexive within the embedded vP. Crucially, since vP is not a phase in this case, this probing will not violate the PIC, and can be entertained. As a result of this, the  $\rho$ -feature on the matrix T and the  $\rho$ -feature on the reflexive will be shared, and its instance on the matrix T will become valued. Now, notice that  $\rho$  is not the only

feature shared by T and the reflexive: there is also a  $\phi$ -feature on the reflexive which still has to get valued, and there is a valued instance of the  $\phi$ -feature on the matrix T. Piggybacking on  $\rho$ -probing, the  $\phi$ -feature on the reflexive will get valued at this moment. Now the derivation will converge: all features within both the embedded and matrix clauses are now valued. The matrix subject further raises to Spec,TP. Schematically this process is shown in (41).

(41) *The structure of the Russian matrix TP of (37-a), the interpretable  $\rho$ -feature is on the matrix T.*



The two possibilities for placing an interpretable unvalued instance of  $\rho$ -feature therefore gives us two possible readings in the case of SE-type languages. (42) summarizes binding domains in languages with complex and monomorphemic anaphors in the case of control constructions.

- (42) a. **English:** Binding at the level of the embedded VP or vP by the embedded subject/object.
- b. **Russian:** Binding at the level of the embedded or the matrix TP by the embedded or the matrix subject.

	English	Russian
c. Phasal domains:	embedded VP or $vP$	Matrix TP
Valuation of $\phi$ -feature on the reflexive:	valued in the lexicon	valued through T-probing

In languages with complex reflexives, such as English, the  $\phi$ -features of the reflexive are valued in the lexicon, and therefore  $vP/VP$  will be closed.  $\rho$ -feature cannot be placed any higher than the embedded TP, otherwise it will not be able to probe the reflexive.

In languages with SE-type anaphors, control  $vP/VP$  are prevented from closing since the  $\phi$ -feature of the reflexive is unvalued. In order to value it,  $\rho$ -feature must be placed on the T-head, which would allow T to probe SE, and value its  $\phi$ -features. Further, embedded TP/CP do not close because the T-feature on T in control constructions is unvalued, and that allows placing the  $\rho$ -feature not only on the embedded T, but also on the matrix T, leading to the ambiguity.

#### 4.2.4.2 Anaphors within subjunctives

In this section I show the contrast between long-distance binding in subjunctives which can be observed in non-verb-raising languages with monomorphemic anaphors, such as Russian, and verb-raising languages with monomorphemic reflexives. In discussion below I will concentrate on Icelandic and Italian on the one hand and Russian on the other. Icelandic and Italian are assumed to be verb-raising languages, where  $v+V$  moves to T, while in Russian the  $v+V$  complex stays inside  $vP$ . Consider the data below.

(43) *Icelandic long-distance anaphors*

- a. Jón<sub>i</sub> segir að María elski sig<sub>i</sub> (Subjunctive)  
 John said that Mary loves<sub>SUBJ</sub> self  
 ‘John said that Maria loves him’
- b. \*Jón<sub>i</sub> veit að María elskar sig<sub>i</sub> (Indicative)  
 John knows that M. loves<sub>IND</sub> self  
 ‘John knows that Maria loves him’

(44) *Italian long-distance anaphors* (Giorgi 2006)

- a. \**Quel dittatore<sub>i</sub> ha detto che i notiziari televisivi hanno parlato a lungo*  
that dictator has said that DEM news TV has talked<sub>IND</sub> for a long time  
*delle proprie<sub>i</sub> gesta.* (Indicative)  
about self's deeds  
'That dictator said that the TV news programs talked for a long time about self's  
deeds.'
- b. *Quel dittatore<sub>i</sub> spera che i notiziari televisivi parlino a lungo delle*  
that dictator hopes that DEM news TV talk<sub>SUBJ</sub> for a long time about  
*proprie<sub>i</sub> gesta.* (Subjunctive)  
self's deeds  
'That dictator hopes that TV news programs will talk for a long time about self's  
deeds.'

(45) *Russian long-distance binding*

- a. *Ivan<sub>i</sub> xočet čtoby Boris<sub>j</sub> narisoval sebja<sub>\*i/j</sub>* (Subjunctive)  
Ivan wants that<sub>SUBJ</sub> Boris draw self  
'Ivan<sub>i</sub> wants Boris<sub>j</sub> to draw himself<sub>\*i/j</sub>'
- b. *Ivan<sub>i</sub> skazal čto Boris<sub>j</sub> narisoval sebja<sub>\*i/j</sub>* (Indicative)  
Ivan said that Boris draw self  
'Ivan<sub>i</sub> said that Boris<sub>j</sub> drew himself<sub>\*i/j</sub>'

The data above show an interesting asymmetry. In Italian and Icelandic, both verb-raising languages, long-distance binding is possible when the anaphor is located within the subjunctive embedded clause. For instance, in (43), the monomorphemic reflexive *sig* can be bound by the matrix subject *Jón* in case the embedded clause is subjunctive, (43-a); notice that such long-distance binding is impossible in case of indicative embedded clauses, (43-b). Similarly, in the Italian example (44), *proprie* can refer to *quel dittatore* in case the embedded clause is subjunctive, (44-b), and this coreference is impossible in case of indicative embedded clause, (44-a).

The situation is different in Russian, a non-verb-raising language. Long-distance binding in subjunctives is not allowed, just as in case of the embedded indicative clauses. In both examples in (45), the reflexive *sebja* can only refer to the embedded subject *Boris*, and never to the matrix

subject *Ivan*.

How can this asymmetry be derived in the current framework?<sup>6</sup> The idea of the solution is the following: As I argued above, monomorphemic anaphors raise and adjoin to the  $\nu$ -head. Now, in Russian, there is no further raising of the  $\nu$  head complex, and therefore the anaphor remains buried within the embedded  $\nu$ P. In Icelandic, the  $\nu$  head complex undergoes raising to T, and this move allows the reflexive to escape  $\nu$ P, and further (since I argued that T raises to C in subjunctives), escape the embedded TP. Escaping  $\nu$ P/TP makes it possible for the matrix T to probe it, and can lead to establishment of binding relations between the matrix subject and the embedded reflexive.

Now let me show the details of this solution, and demonstrate how the relevant derivations proceed in the case of Russian and Icelandic.

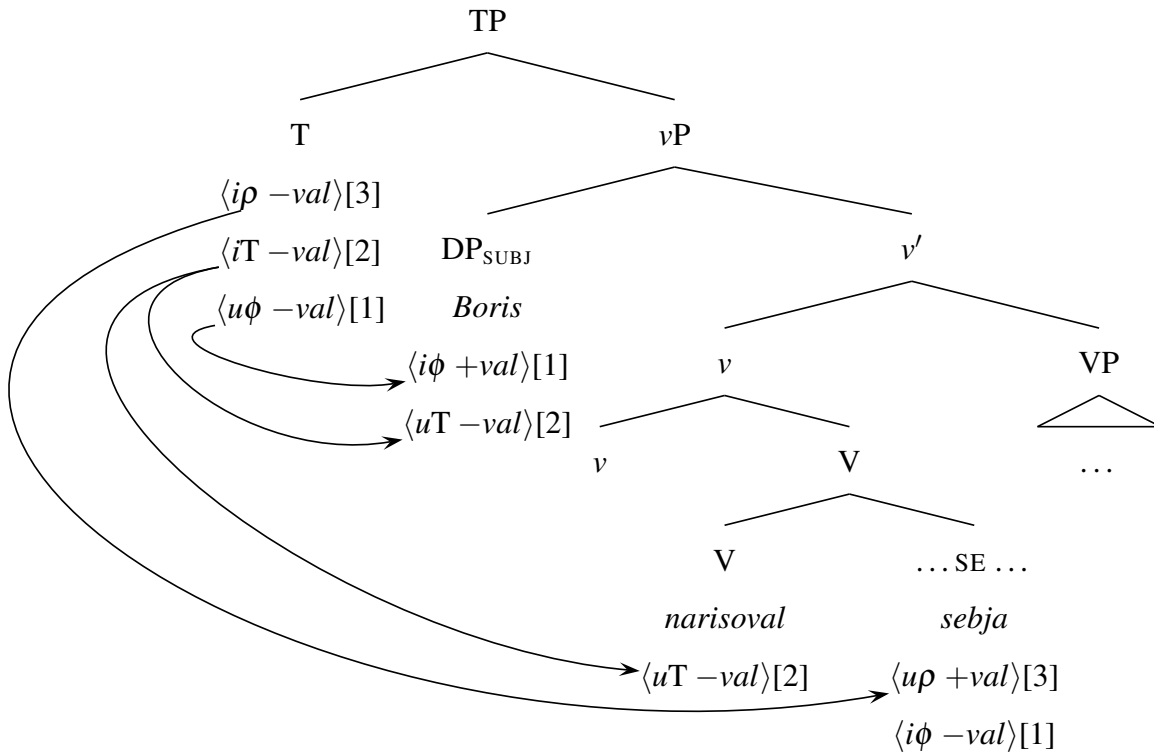
Consider the case of Russian. First notice that local binding can be obtained in a way similar to simple clauses. Assume that the  $\rho$ -feature is present on T.  $\nu$ P is completed in the standard way, and T is merged with it. T probes first the Spec, $\nu$ P in order to get a value for its  $\phi$ -features, and shares the T-feature present on the embedded subject. At the second probing, T finds its goal in the  $\nu$ +SE complex: it has an unvalued uninterpretable T-feature on  $\nu$  (since as I argued above in subjunctives the T-feature on the embedded verb is unvalued), and an unvalued interpretable  $\phi$ -feature on SE, as well as an uninterpretable valued  $\rho$ -feature on SE. As a result of this probing, the T-features on T,  $\nu$ +V, and the embedded subject all become shared, even though they are still unvalued. Further, the  $\rho$ -feature on the reflexive and on T are shared, and valued, and further, the  $\phi$ -features become valued on T as well as on SE (as a byproduct of T probing  $\nu$ +V+SE for  $\phi$ -features). This is shown in (46):

---

<sup>6</sup>Note that it certainly cannot (obviously) be accounted for in any of the other frameworks.



- (46) *The structure of the embedded subjunctive TP in Russian (45-a), the interpretable  $\rho$ -feature is on the embedded T.*

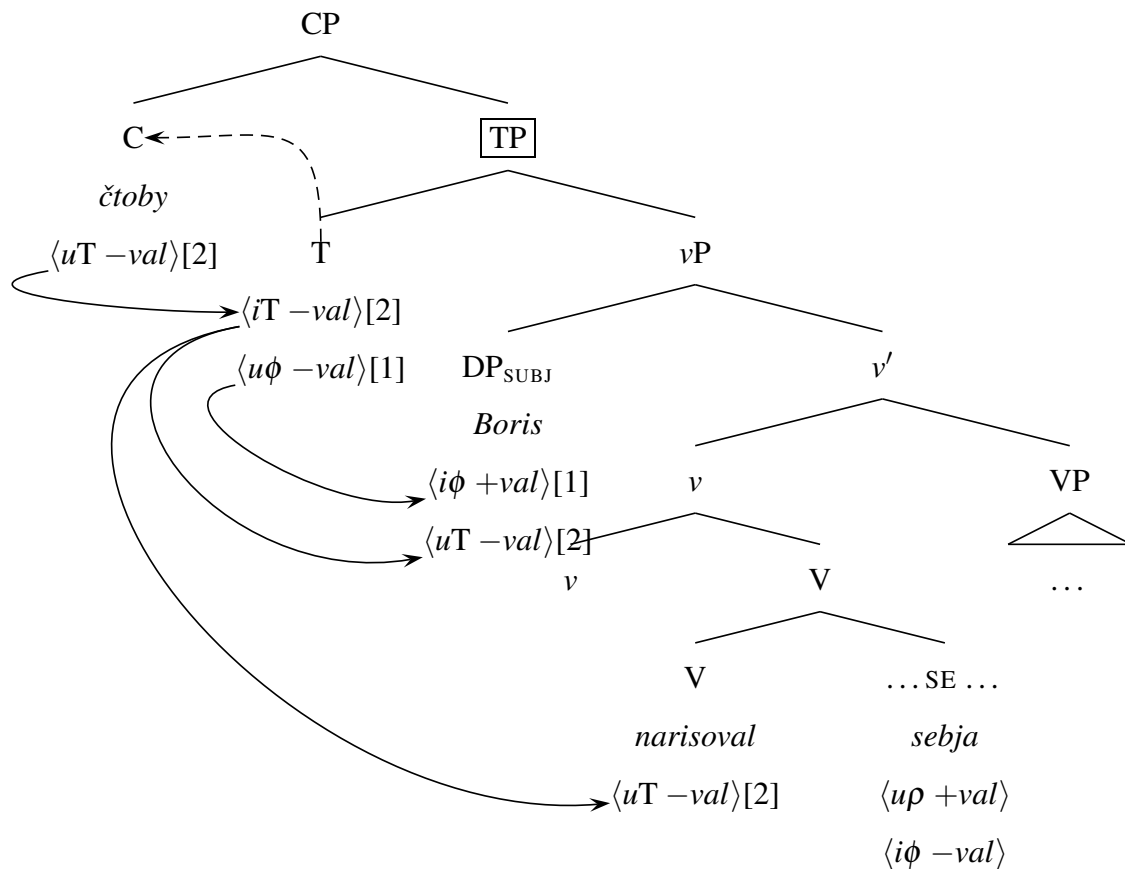


At this moment, all the features within  $vP$  are either valued ( $\phi$ ,  $\rho$ ) or shared with some element outside of  $vP$ .  $vP$  can now be closed, and binding will apply as soon as the embedded subject is raised into Spec,TP position. A  $\lambda$ -operator will be inserted in T-position, and  $\lambda$ -conversion will result in binding by the local embedded subject.

Notice that in the Russian case the derivation would crash if one decides to put the  $\rho$ -feature on the element in the matrix clause. Assume that the  $\rho$ -feature is not placed on the embedded T. In this case, after TP is completed,  $vP$  is still open, as there is an unvalued  $\rho$ -feature on the reflexive which prevents  $vP$  from being closed. Now, C is merged with TP. I argued above that a subjunctive C has an unvalued T-feature, and therefore, C would probe T, and their T-features will become shared. At this moment, TP can be closed: all relevant features on the edge of T are either valued ( $\phi$ -features of the subject), or shared with an element outside of TP (the T-feature on T and

the embedded subject is shared with a T-feature on C). Closing TP would mean that there is no further probing that will be able to target the reflexive within the embedded TP in order to value its  $\phi$ -feature. This moment of the derivation is shown in (47), where the closed TP is marked by a box.

- (47) *The structure of the embedded subjunctive clause in Russian (45-a), no interpretable  $\rho$ -feature in the embedded clause.*

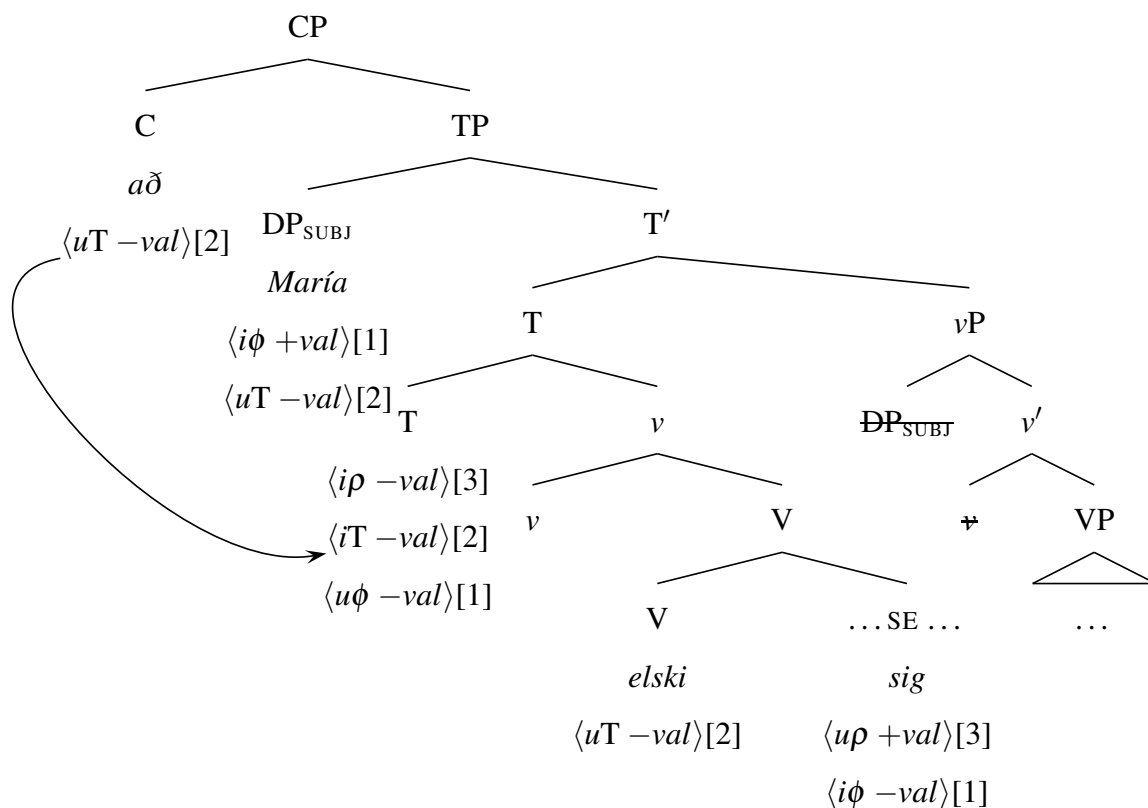


Therefore, it is unsurprising that Russian does not allow long-distance binding in subjunctives. In the remaining part of this section I will show how verb-raising affects binding possibilities. Consider the case of Icelandic.

The first possibility for deriving binding relations in Icelandic mimics the Russian case. After the step shown in (47), the  $v+V+SE$  complex will raise to adjoin to T (presumably since the

T-features on T are strong), the embedded subject will raise to Spec,TP, and C is inserted into the structure with unvalued uninterpretable T-features (see Section 4.2.1.3, where I argued that subjunctive C comes with  $\langle uT -val \rangle$ ). C will probe T in order to share its T-features. At this moment all relevant features within TP are either valued ( $\phi$ ,  $\rho$ ) or shared with an element outside of TP (the T-feature on T, embedded subject, and  $v+V$  is shared with the T-feature on C). TP acquires phasal status, and therefore can be closed for further probing. The binding relations can be obtained in the standard way by inserting a  $\lambda$ -operator above the T position, and the only possible binder of the reflexive in this case is the embedded subject present at Spec,TP.

- (48) *The structure of the embedded subjunctive CP in a verb-raising language, (43-a), the interpretable  $\rho$ -feature is on the embedded T.*



antecedent. The solution will be similar to the way long-distance binding was established in the case of Russian control clauses.

Observe that in order to establish binding by the matrix subject, one needs to place the  $\rho$ -feature on the matrix T. In the case of Russian subjunctives, this was impossible, since the embedded TP was closed for probing because all features within it were either valued or shared with some element outside of it. Now, consider the moment of the derivation when T is merged with  $\nu$ P in Icelandic. In the case described above, the  $\phi$ -features on SE were valued as a byproduct of T with a valued  $\rho$ -feature probing SE for its  $\rho$ -features. Notice that this  $\phi$ -feature valuation can occur only if there is some other feature, such as  $\rho$ , which can trigger probing relation between T and the reflexive. Now, if the  $\rho$ -feature is not placed on the embedded T, but instead on the matrix T,  $\phi$ -feature valuation on the reflexive will not occur in the embedded clause.

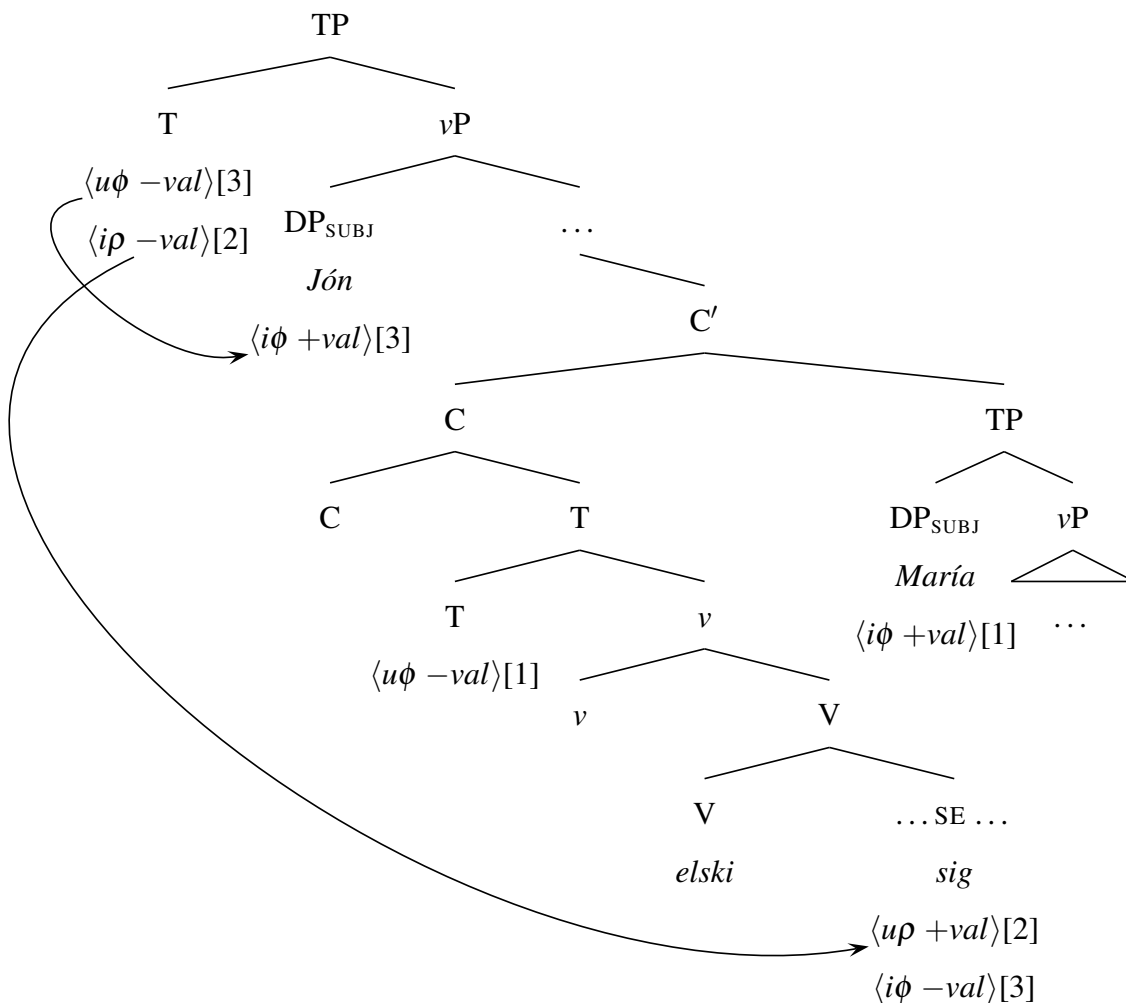
As I showed above, the reflexive “escapes”  $\nu$ P, and is in a T-head complex. After  $\nu$ +V+SE raising to T, the embedded subject is moved to Spec,TP for the EPP. Next, C is merged with TP. Since it is a subjunctive C, it comes with an unvalued uninterpretable T-feature. C will probe T to share its T-features, and after that the T-head complex will raise to C. Similarly to the Russian case, TP will become a closed domain, since all TP-edge features are either valued or shared with an element outside of a TP. TP will be able to close. That however is not a problem: at this moment the SE-reflexive is in the C position, above TP, and therefore is accessible for further probing from the matrix clause.

Further, the matrix  $\nu$ P will be constructed and the matrix T will be merged. Crucially, now it is possible to place a  $\rho$ -feature on the matrix T. The derivation will not crash because the probing between the matrix T and the embedded reflexive is now possible, since the embedded clause is open for probing.

Matrix T will first probe the matrix subject, and get its  $\phi$ -features valued, and further will probe the embedded reflexive in order to value its  $\rho$ -features. This will result in  $\phi$ -feature valuation on the embedded reflexive as a byproduct of  $\rho$ -feature probing. Binding will now be able to apply: a  $\lambda$ -operator will be inserted above the matrix T, and the binder for the embedded reflexive in this

case will be the matrix subject.

(49) The structure of the sentence (43-a), the interpretable  $\rho$ -feature is on the matrix T.



V-to-T-to-C raising in subjunctives allowed me to predict the possibility of long-distance binding in subjunctives in verb-raising languages with monomorphemic reflexives. Notice that this is not a possibility for Russian. If the  $v+V+SE$  complex never raises to T, the reflexive SE will always be c-commanded by T, and therefore, the only way to value its  $\phi$ -features is via probing by an embedded T, and crucially not by matrix T. Therefore in Russian, the embedded subject is the only possible antecedent even in subjunctive embedded clauses. Verb raising is what allows the

reflexive to escape the embedded TP, and that allows an interpretable instance of the  $\rho$ -feature to be inserted in the matrix clause, without leading to a derivation crash.

### 4.3 Anaphors within subjects

The main focus of this section is on reflexives contained within the subjects of the embedded clause and binding possibilities which arise with respect to them. Examples of such a situation are given in (50).

- (50) a. Children<sub>i</sub> know that [ pictures of themselves<sub>i</sub> ] are on sale.  
 b. John<sub>i</sub> told Bill<sub>j</sub> that [ the report about himself<sub>i/?j</sub> ] is on the principal's desk.

First, one can notice that even though proper long-distance binding is prohibited in English, the matrix element may serve as a binder of a reflexive within the embedded clause subject, regardless of the presence of the CP/TP boundary between the antecedent and the reflexive.

Second, as in the case of simple binding, subject orientation does not hold for English subject-internal reflexives. For instance, in (50-b) both *John* and *Bill* can serve as possible antecedents of *himself*.

While English allows such cases, these facts are subject to cross-linguistic variation. Examples in (51) show the Russian case.

- (51) a. \**Deti znajut čto [ svoi fotografii ] budut viset' na stene.* (Russian)  
 Children know that [ SELF'S pictures ] will hang on wall  
 'Children know that pictures of themselves will hang on the wall'  
 b. \**Putin skazal Medvedevu čto [ knigi o sebe ] est' vo vsech bibliotekax.*  
 Putin told Medvedev<sub>DAT</sub> that [ books about SELF ] are in all libraries  
 'Putin told Medvedev that the books about himself are in all libraries.'

Neither of the examples in (51) is grammatical, regardless of the reading. In (51-a) the possessive reflexive *svoi* 'self's' cannot be antecedent by *deti* 'children', while in (51-b) neither the matrix

subject *Putin* nor the matrix object *Medvedev* is a possible antecedent of the reflexive *sebe* ‘self’.

Similar facts hold for other Slavic languages, such as Serbo-Croatian (52-a) and Polish (52-b).

- (52) a. \*Ivan<sub>i</sub> zna da je [čanak o sebi<sub>i</sub>] izašao u novinama  
Ivan knows that AUX [article about SELF] appeared in newspaper  
‘Ivan<sub>i</sub> knows that an article about himself<sub>i</sub> appeared in the newspaper’
- b. \*Jan<sub>i</sub> wie że [artikuł o sobie<sub>i</sub>] ...  
Jan knows that [article about self] ...  
‘Jan<sub>i</sub> knows that an article about himself<sub>i</sub> ...’

Two questions arise with respect to the data above.

- What allows the English reflexive *himself* to be bound by an antecedent outside of its clause?
- What prohibits Russian (and other Slavic) reflexive *sebe/svoi* from being bound using the mechanism used for English?

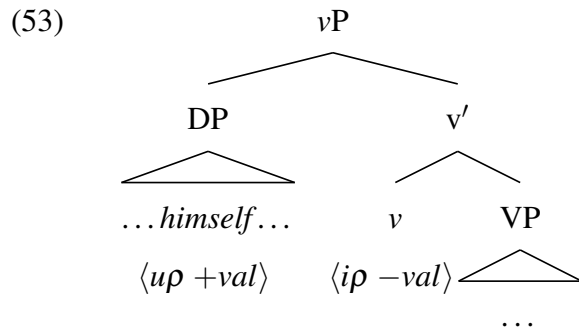
Below I propose an analysis of subject-internal reflexives which answers the questions above. I will show how the framework described in Chapter 3 readily predicts this asymmetry between Russian and English based on the difference in the featural content of reflexives pronouns. First, I demonstrate which mechanism allows English subject-internal anaphors to be bound by an element within the superordinate clause. The required binding relations will be achieved by strategical placement of the  $\rho$ -feature on the embedded C-head. The fact that the  $\phi$ -features on the English reflexive are valued would allow to apply binding at the moment the matrix clause elements are merged into structure.

In Russian placement of the  $\rho$ -feature is also possible on C. The reason the English solution cannot be applied to the Russian case is in the nature of a reflexive. At no point in derivation can the  $\phi$ -features of the anaphor get valued, and therefore the derivation crashes.

### 4.3.1 Subject-internal reflexives in English

In this section I demonstrate how subject-internal reflexives get bound in English. I adopt the binding framework introduced in Chapter 3. I argued above that English reflexives, such as *himself*, *herself*, come from the numeration with two types of features:  $\phi$ -features, which are valued,  $\langle i\phi +val \rangle$ , and an instance of an uninterpretable valued  $\rho$ -feature,  $\langle u\rho +val \rangle$ . I also argued that the uninterpretable  $\rho$ -feature on the reflexive must be probed by the corresponding instance of the interpretable but valued  $\rho$ -feature, which in case of the object anaphor was placed either on  $v$  or  $V$ , resulting in two different binding possibilities: binding by the subject, or binding by the object (direct or indirect).

Now let's consider how the derivation may proceed if the reflexive is contained within the subject position. The subject originated in Spec, $v$ P, and therefore placing the interpretable instance of the  $\rho$ -feature on either  $v$  or  $V$  won't lead to the desired results. Under assumption that only unvalued instances of features serve as probes, neither  $v$  nor  $V$  will be able to probe the valued instance of  $\rho$ -feature if it is located in the Spec, $v$ P position, since this position is not c-commanded by either  $v$  or  $V$ , see (53):



Above I claimed that the  $\rho$ -feature may be strategically placed on an arbitrary head in the clause, in order to achieve the desired binding. While the placement of the interpretable unvalued  $\rho$ -feature is arbitrary, and nothing prevents one from placing it arbitrarily far from the corresponding  $\rho$ -feature on the reflexive, placing it “too far” may lead to a derivation crash, since locality considerations



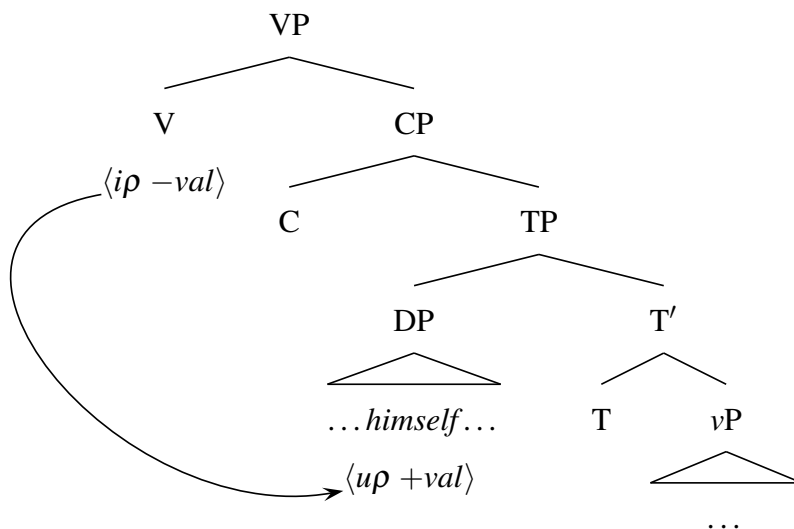
will not allow probing between unvalued and valued instances of a  $\rho$ -feature. For that very reason in the English case of object reflexives the  $\rho$ -feature could not be placed outside of the  $\nu$ P domain: closure of  $\nu$ P/TP would result in the impossibility of  $\rho$ -feature sharing between the potential goal and the potential probe, resulting in a crash of the derivation. However, when the reflexives are present within the subject, we must entertain the possibility of placing an interpretable instance of the  $\rho$ -feature on the matrix V-head. In what follows I demonstrate how the derivation of (50-b), repeated below as (54), proceeds and how the binding is obtained.

(54) John<sub>i</sub> told Bill<sub>j</sub> that [ the report about himself<sub>i/?j</sub> ] is on the principal's desk.

The subject *the report about himself* starts at the Spec, $\nu$ P position, and contains a reflexive *himself*, endowed with a  $\langle u\rho +val \rangle$  feature. T probes first the subject, and shares its T-feature with the T-feature on the subject, and further probes the embedded verb *is* to get a value for a T-feature. The subject is further raised into Spec,TP position (to satisfy the EPP on T). Notice that at this moment probing of the anaphor has not yet been done, and the  $\rho$ -feature of *himself* is still not shared with any interpretable instance of it. Nevertheless, there is no reason for the derivation to crash. Recall that in the case of the object anaphor, it does not escape the  $\nu$ P/TP phase, and therefore if its  $\rho$ -feature is not checked by the time  $\nu$ P/TP is complete, the derivation will crash, as probing into it would be impossible because of the phase impenetrability condition.

The situation is different with a subject-internal anaphor: it is outside the  $\nu$ P/TP phase, and therefore the probing of its  $\rho$ -feature is still possible. Let's assume that the interpretable instance of the  $\rho$ -feature is placed on the matrix V-head. After such a V is merged into structure, it will probe the subject-internal reflexive, and this probing will result in sharing of  $\rho$ -features on V and on the reflexive pronoun. This situation and probing are demonstrated in (55).

(55) *The subject-internal reflexive in English, the interpretable  $\rho$ -feature is on the matrix V.*



Notice that this probing is possible since the reflexive is located at the edge of the TP-phase, and therefore is accessible for probing before CP is closed, i.e. at least until the moment the matrix V is merged into the structure.

At this moment, the  $\rho$ -features on V and on the reflexive are shared; it becomes valued on V, and the  $\rho$ -feature on the reflexive acquires an interpretable instance. That means that all features within the embedded CP are satisfied, and it can be sent to semantic interface. The  $\rho$ -feature in the matrix VP will be converted to a  $\lambda$ -operator, which binds the variable within the subject position:

(56)  $\lambda x \exists e \text{ Theme}(e, \text{pictures of } x) \dots$

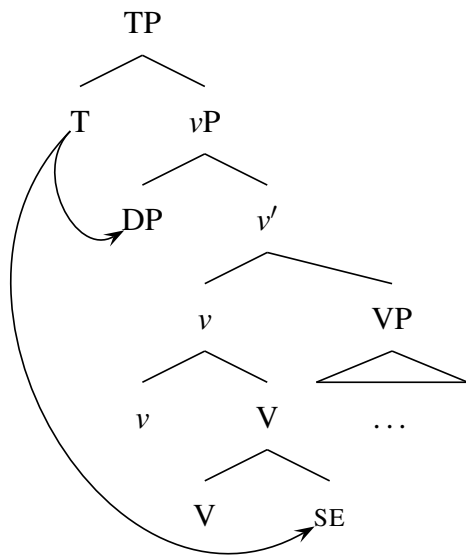
Now,  $\lambda$ -conversion will apply at the moment a higher nominal is merged into structure, in this case the matrix subject *John* or the matrix object *Bill*. This will result in binding of the subject-internal reflexive by a matrix clause element, depending on when lambda-conversion applies. The result will be binding of the reflexive by the matrix object or the matrix subject.

### 4.3.2 Subject-internal reflexives in Russian

In the last section the framework of feature-based anaphoric binding was extended to allow for placement of the  $\rho$ -feature on heads other than  $\nu$  or  $V$ , for instance on  $C$ . In the case of English this resulted in the possibility of subject-internal anaphors in the embedded clause to get bound by a matrix element. However, not all languages allow this: for instance, in Russian binding of subject-internal anaphors by a matrix element is ungrammatical, see (51). The question which remains to be answered is why doesn't a solution similar to English apply to Russian.

As I argued before, the difference between Russian and English reflexives is their specification for  $\phi$ -features. English anaphors come from the lexicon with valued  $\phi$ -features, Russian reflexives do not. As I showed in Chapter 3, English reflexives are allowed to enter in a binding relation in a given domain when all relevant features within this domain have been valued (or shared with elements in the higher domains): the presence of valued  $\phi$ -features on the complex anaphor does not prevent the domain from acquiring phasal status, and does not prevent the application of binding. The situation in Russian is different: the  $\phi$ -features of the reflexives are unvalued in the lexicon, and must be valued through a probe-goal relation, established first between  $T$  and the subject, and second between  $T$  and the reflexive (through  $\nu$ ), shown in (57).

(57) Valuation of the  $\rho$ -feature in Russian through T-probing.



Only after the  $\phi$ -features of the reflexive are valued through this process, can binding relations be established through  $\lambda$ -conversion, as described above.

What goes wrong in case of Russian subject-internal reflexives, (51)? I argue that the problem with binding into embedded subjects in Russian is related to the inability of  $\phi$ -features on a reflexive to get a value. In fact, in order for a reflexive to get its  $\phi$ -features valued, a certain head must enter in a probe-goal relationship with two elements: the reflexive and the element with valued  $\phi$ -features (in any order). In the case of a monoclausal sentence, T with an unvalued interpretable instance of  $\rho$ -feature plays that role. I argue that no such head is available in case the reflexive is located within the embedded subject.

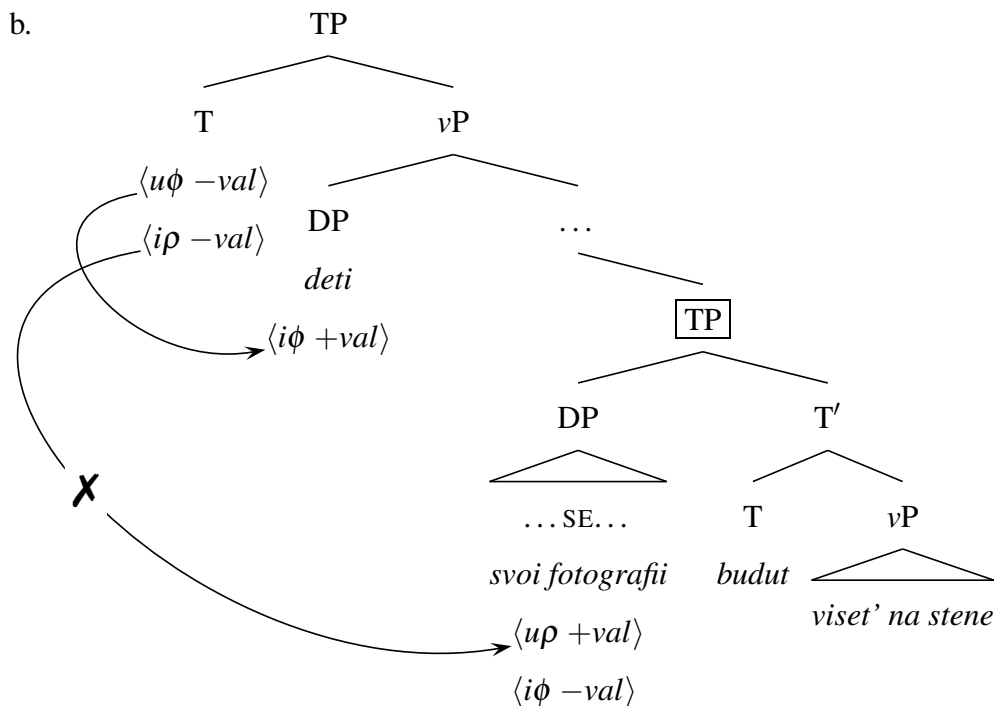
No head in the matrix vP probes an embedded subject or any element which shares features with embedded subject, therefore an analysis similar to English cannot apply to Russian.

The only possibility is therefore matrix T, since it probes the matrix subject. Assume that the  $\rho$ -feature is placed on the matrix subject. Notice that by the time the matrix T is merged into structure, the embedded TP is closed for probing. Its subject is not PRO, and therefore its  $\phi$ -features are valued ( $\phi$ -features of the reflexive *within* the subject are still unvalued, but that is irrelevant for closure of TP). Further, the T-feature on T is either valued if the embedded clause

is indicative, or raised into C if the embedded clause is subjunctive. At the same time, unvalued  $\phi$ -features of the reflexive are located within the subject in Spec,TP. That means that even if the  $\rho$ -feature is placed on the matrix T, it will still not be able to probe a reflexive within the embedded Spec,TP. The derivation will therefore crash.

(58) *The impossibility of the  $\rho$ -feature placement on the matrix T in the case of Russian subject-internal reflexives.*

- a. \**Deti znajut čto [ svoi fotografii ] budut viset' na stene.* (Russian)  
 Children know that [ SELF'S pictures ] will hang on wall  
 'Children know that pictures of themselves will hang on the wall'



Notice, that complex anaphors in Russian, such as reciprocal *drug druga* 'each other', pattern with the English case:

- (59) *Deti<sub>i</sub> dumajut čto fotografii drug druga<sub>i</sub> visjat na stene.*  
 children think that pictures each other<sub>GEN</sub> hang on wall.  
 'Children think that the pictures of each other are hanging on the wall'

This fact can be explained under the assumption that reciprocals in Russian are specified for  $\phi$ -features (presumably, PL). In this case, there is no need to resort to a special mechanism of  $\phi$ -feature valuation on the anaphor, and therefore binding can apply at the same moment as in the English case. Placement of the  $\rho$ -feature on C will result in the embedded clause being converted into a  $\lambda$ -expression upon completion, and introduction of the matrix elements will trigger  $\lambda$ -conversion, resulting in coreference between the reciprocal and the matrix subject or object.

## Chapter 5

# Interaction of binding and long-distance movement

In this chapter I explore the interaction of binding and movement, in particular A'-movement. I consider cases of English and Russian, and show that there are certain asymmetries in case of extracting a reflexive located in a *wh*-phrase in Russian and in English. I will further demonstrate that scrambling in Russian can either give rise to new binding possibilities or not, depending on the type of the embedded clause. I demonstrate how the theory of binding proposed here allows us to account for these facts.

I argue that binding by the matrix clause antecedent is possible if there are features in the embedded clause that prevent the embedded clause from becoming a phasal domain. If the embedded CP is not phasal, it is possible to place an interpretable instance of the  $\rho$ -feature on the matrix element, and it will be able to probe an uninterpretable valued instance of the  $\rho$ -feature on the reflexive. I show which features can prevent the closure of the embedded CP.

First, the Q-feature on the intermediate C in the case of long-distance *wh*-movement allows the *wh*-phrase containing the anaphor to move to the edge of the embedded CP, a position where it can be probed by a matrix element endowed with an interpretable instance of the  $\rho$ -feature.

Second, an unvalued T-feature, as in the case of subjunctives, postpones the closure of TP, and

therefore allows the scrambled element to escape TP, and allows late application of binding. As a result, reflexives scrambled out of subjunctive clauses can have matrix antecedents.

I consider these cases below.

## 5.1 Interaction with long-distance *wh*-movement

In this section I will explore the interaction of binding with long-distance *wh*-movement. The main question is whether A'-movement feeds new binding possibilities when the anaphor is embedded in a *wh*-phrase.

The following data comes from English, and had been widely discussed in the literature on the subject.

- (1) a. John<sub>i</sub> thinks that Bill<sub>j</sub> likes the picture of himself<sub>\*i/j</sub> taken by Mary.
- b. John<sub>i</sub> wonders if Bill<sub>j</sub> likes that picture of himself<sub>\*i/j</sub>.
- (2) a. Which picture of himself<sub>i/j</sub> does John<sub>i</sub> think that Bill<sub>j</sub> likes.
- b. John<sub>i</sub> wonders which picture of himself<sub>i/j</sub> Bill<sub>j</sub> likes.

As I mentioned earlier, long-distance binding is prohibited in English. In either example in (1) the only possible antecedent of the reflexive in the subject of the embedded clause, *Bill*. On the other hand, if the reflexive is embedded in a *wh*-phrase, like *which picture of himself* in example (2), there is another binding possibility. A'-movement of the *wh*-phrase (to the intermediate Spec,CP put the reflexive in a local configuration with the matrix subject, *John*) allows the binding relation between *John* and *himself* to be established. Notice that the *wh*-phrase with the internal reflexive does not have to stay in the embedded Spec,CP as in (2-b). Further A' movement, like in (2-a), will not affect this already established binding relation.

Several questions arise from consideration of these data. The first question is the question of the optionality of binding. Notice that the examples in (2) are ambiguous: both *John* and *Bill* are possible antecedents for the reflexive. The problem with multiple possible antecedents is



again the question of the level of application of the binding theory. In order to get the embedded subject to be the antecedent, binding theory needs to be applied before *wh*-movement of the *wh*-phrase (into intermediate Spec,CP). In order to get the matrix subject as the antecedent, the binding theory should be applied when the reflexive is located in the intermediate Spec,CP. The question is therefore what allows this optionality in binding theory application. To put it in terms of phase theory, the question of optionality can be reduced to the question of why the reference of the reflexive is not fixed at the moment the embedded TP is completed.

This question could not be answered in the framework that postulated fixed domains for binding theory. Assuming that the binding domain is TP (or a domain with an accessible SUBJECT, as was originally claimed by Chomsky 1981/1993), it is unclear what allows extending this domain in case of *wh*-movement.

The second question is the question of obligatory reconstruction of *A'*-movement. Considering that *wh*-movement must obligatory reconstruct at LF, it is unclear why binding by the matrix subject is still possible in (2). Notice, that binding is impossible in example (3).

(3) \*Who does Mary think that [pictures of himself] please.

Here the *wh*-phrase *who* is moved from the embedded object position to the matrix Spec,CP through an intermediate Spec,CP position by *A'*-movement. At the embedded Spec,CP position, *who* c-commands the reflexive within the embedded Spec,TP and therefore is expected to be able to bind it. This coreference however is impossible. The asymmetry one can observe here is the following: *A'*-movement of the reflexive can result in new binding possibilities, while *A'*-movement of the potential antecedent cannot.

In what follows I show that the feature-based analysis of binding can solve these problems. It can account for the optionality, i.e. explain why evaluating of binding relations can be postponed in the case of a reflexive contained in a *wh*-phrase, and can also explain the asymmetries between moving the reflexive vs. moving the antecedent.

### 5.1.1 *Wh*-movement: featural approach

In order to tie in the theory of phases proposed in Chapter 2, the feature-theory of *wh*-movement needs to be outlined in this new light. That is the goal of this section.

In their work, Pesetsky and Torrego argue that in their feature-sharing system, a Q-feature is responsible for formation of questions. They claim that the feature Q is present on both interrogative C and the *wh*-phrase. The instance of the Q-feature on the interrogative C is interpretable and unvalued,  $\langle iQ -val \rangle$ , and it receives its value from the uninterpretable but valued instance of the Q-feature,  $\langle uQ +val \rangle$ , on the *wh*-phrase. In their approach, interrogative CP is formed as shown in (4).

(4) Q-feature valuation on the interrogative CP in Pesetsky and Torrego 2007 framework:

$$\dots C_{\langle iQ -val \rangle} \dots \mathit{what}_{\langle uQ +val \rangle} \dots \longrightarrow C_{\langle iQ -val \rangle [6]} \dots \mathit{what}_{\langle uQ +val \rangle [6]}$$

C enters in an Agree relation with the *wh*-phrase in its c-command domain; this results in Q-feature sharing, which is indicated by (arbitrary) [6] in (4). These two Q-features become instances of the same feature, interpretable in one location, and valued in another.

The successive cyclic *wh*-movement is derived in a similar way, the main difference being the nature of the Q-feature on the embedded (non-*wh*) complementizer. As argued in Pesetsky and Torrego 2007, Q-feature on a head of projection which participates in successive-cyclic *wh*-movement is an unvalued uninterpretable instance of Q,  $\langle uQ -val \rangle$ . They note that categories other than C can host this feature, for instance, if  $\nu P$  is a phase, and successive-cyclic movement proceeds through the edge of  $\nu P$  (see Legate 2003 for evidence that *wh*-movement targets the edge of  $\nu P$ ),  $\nu$  will also have an uninterpretable unvalued instance of Q-feature.

(5) Q-feature valuation on the intermediate CP in Pesetsky and Torrego 2007 framework:

$$\dots C_{\langle uQ -val \rangle} \dots \mathit{what}_{\langle uQ +val \rangle} \dots \longrightarrow C_{\langle uQ -val \rangle [6]} \dots \mathit{what}_{\langle uQ +val \rangle [6]}$$

The Agree-relation which occurs in this situation is shown in (5). The fact that the Q-feature is

uninterpretable on the intermediate C contributes to the fact that this process does not affect the semantics of the *wh*-question.

In what follows I assume that Pesetsky and Torrego's analysis is on the right track, and show how English binding facts can be derived from it, giving rise to ambiguity of the sentences with the reflexive embedded in a *wh*-phrase.

### 5.1.2 Phasal status of embedded CP

In this section I apply the theory developed in Chapter 2 to CP-complements with the Q-feature on C, whether it is interpretable or not. In what follows I consider the definition of a phasal domain proposed earlier, and apply it to CP. In the framework proposed above, the definition of the phasal domain D depends on which features are considered *relevant* for a particular domain D. Having such features unvalued by the time the domain D is completed, or not shared with a head selecting D as its complement, prevents the domain D from having phasal status. On the other hand, if all domain-relevant features are valued at the level of the domain completion, or shared with a higher head, selecting this domain D as its complement, the domain D acquires phasal status, and therefore becomes closed for further probing into it. If any feature is left unvalued within this domain, or if any feature within this domain has no interpretable instance, the derivation would crash, as there would be no way of either valuing a feature, or supplying a feature with an interpretable instance. Crucially, having an interpretable and unvalued feature present in D causes the derivation to crash.

D-relevant features are usually the features traditionally associated with a given domain, such as T-features in case of TP, or features of the complement of V in case of VP. What would be CP-relevant features in this case? Clearly, if CP is associated with a Q-feature, it is desirable to claim Q being a CP-relevant feature. Further, if the Clause-type feature is associated with CP (as for example claimed in Adger 2004), it is reasonable to assume that Clause-typing features are also CP-relevant.

Based on the definition of a phasal domain, if an embedded C is endowed with a Q-feature, CP is not a phasal domain, at least until the moment immediately higher V is merged with it:

there is no way to value a Clause-type feature on the embedded C before a higher element with a matching feature is merged into structure. This situation is similar to the case of subjunctive complements, which also need to have their features valued from a higher element. In case of subjunctive complements, the T-feature on subjunctive T is unvalued, and therefore subjunctive TP and CP are not phasal domains. Similarly in the case of +Q-complements (with either interpretable or uninterpretable Q-features on C), CP does not close immediately upon completion.

Now let's consider the details of this proposal. There are two separate cases which need to be considered. The first is the case of C endowed with an interpretable Q feature,  $\langle iQ -val \rangle$ : in this case CP must necessarily be a complement of a +Q-taking verb, such as *wonder*, *ask*, etc. In this case, the *wh*-phrase stays in the embedded Spec,CP. The second case is the case of uninterpretable Q feature on the embedded C,  $\langle uQ -val \rangle$  — a “purely formal” feature which serves the sole purpose of allowing the embedded *wh*-phrase to establish relation with a higher position in the Spec,CP of the matrix clause where it is interpretable. Let us consider each in turn.

#### 5.1.2.1 $\langle iQ -val \rangle$ CPs

The situation when the Q-feature on the embedded Spec,CP is interpretable corresponds to CPs which are complements of verbs such as *wonder*, *ask*, etc. In this case the *wh*-phrase is interpreted in the embedded Spec,CP. Examples in (6) provide such cases:

- (6)
- a. John wonders which books Bill reads.
  - b. Mary asked her students who bought the books about syntax.
  - c. Peter is interested where he can buy a new phone.

In all of these cases it is a property of the verb which allows it to select the interrogative CP. Notice that with verbs which do not belong to this class, such embedded questions are impossible:

- (7)
- a. \*John thinks which books Bill reads.
  - b. \*Peter is sad where he can buy a new phone.

Further, some verbs from the +Q-group can only take interrogative CPs as their clausal complements; complements headed by *that* are for instance ungrammatical:

- (8) a. \*John wonders that Bill reads “Syntactic Structures.”  
b. \*Mary asked her students that they need to buy more books about syntax.  
c. \*Peter is interested that the new phone will be released next week.

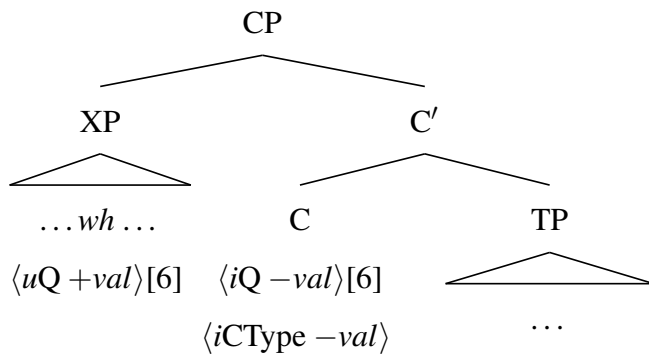
On the other hand, some of the verbs allow both interrogative and non-interrogative CPs as their complements:

- (9) a. Marlin knows that ellipsis is an interesting syntactic phenomenon.  
b. Marlin knows who wrote the best article about right node raising.

These facts show that whether the interrogative CP is grammatical in a given sentence depends on the property of the selecting verb (similar to the case of control complements). In order to formalize this selectional restriction, I postulate that a clause-type feature is responsible for this selectional property (similar to what I claimed in Chapter 2 for control complements). This feature must be valued under an Agree relation with the matrix V. If V has no appropriate clause-typing feature, the sentence will be ungrammatical, since the derivation will crash. Notice that here, as in the case of verbs taking control complements, I argue that verbs selecting +Q-complements come with an appropriate clause-typing feature from the lexicon. Next, consider the derivation from the formal standpoint.

(10) shows the structure of the CP-layer after the *wh*-phrase has moved to its specifier position.

(10) *The structure of the CP-layer after the wh-phrase has moved to its specifier position.*



Here, the Q-feature is not problematic for closing CP: there is a valued instance of it present on the *wh*-phrase, and this valued instance of the Q-feature has been shared with an interpretable instance of the Q-feature on C. There is however a feature which is responsible for selection of the embedded +Q-complement: a Clause-typing feature, denoted in (10) as  $\langle \text{CType} \rangle$ .

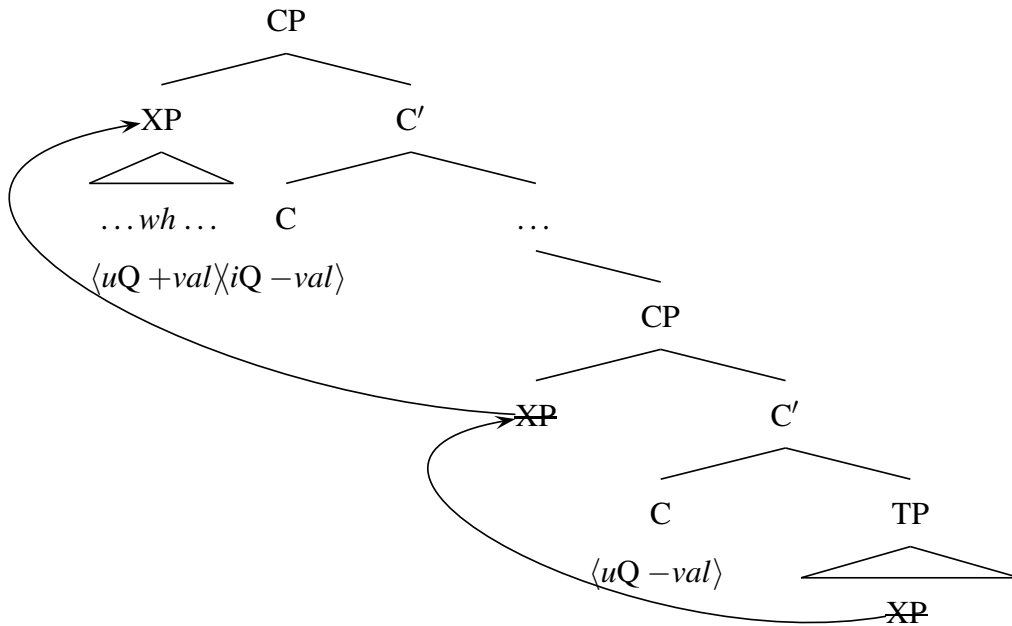
Now, since the matrix verb is responsible for checking the clause-type properties of the embedded clause in this case, it should be able to probe the embedded C. Therefore in this case, when the *wh*-phrase is present in the embedded Spec,CP, CP remains open for further probing by the matrix V and is not a phasal domain. Now, when the matrix V is inserted into the structure, it probes the embedded C. Establishing Probe-Goal relations between the matrix V and the embedded C results in valuation of the Clause-type feature, and results in matrix V closing the embedded CP. Embedded CP acquires phasal status after its head has been probed by the matrix V. In the next section I formally show how this affects binding possibilities.

### 5.1.2.2 $\langle uQ -val \rangle$ CPs

In this section I consider the case of CP with  $\langle uQ -val \rangle$  feature on its C-head. The Q-feature on the C-head is uninterpretable, and therefore does not affect the semantics: it only triggers *wh*-movement to the Spec,CP. Unlike what happens in the case discussed in the previous section, in this case the *wh*-phrase is interpreted in the matrix clause, the complementizer of which is endowed with an interpretable instance of the Q-feature. This is the structure which corresponds to the case

of successive cyclic *wh*-movement. Schematically this case is shown in (11).

(11) *Successive cyclic movement:*



In this case, the only interpretable instance of the Q-feature is located on the matrix C. Once again, I postulate the existence of a Clause-typing feature on the embedded complementizer. The role of that feature in this case would be to signal the fact that the embedded clause hosting a successively-cyclically moved *wh*-phrase in its edge must in turn be embedded under a true interrogative CP. It will be valued only after the matrix C probes the embedded C for a Q-feature, and that would also result in valuing the Clause-typing feature on the embedded C as a byproduct of Q-feature Agree.

Formally, the embedded C is endowed with  $\langle uQ -val \rangle$ , which has in turn been shared with the  $\langle uQ +val \rangle$  on the *wh*-phrase which moved into the embedded Spec,CP. That results in sharing Q-features on the embedded C and Spec,CP. Further, the embedded C is also endowed with  $\langle iCType -val \rangle$ . Upon completion of the embedded CP, it is not a phasal domain since there are CP-relevant features which are still neither valued nor shared with any element outside the domain.

The matrix V is merged with the embedded CP. Absence of interpretation of the Q-feature on the embedded C and absence of value for the Clause-typing feature on the embedded C both

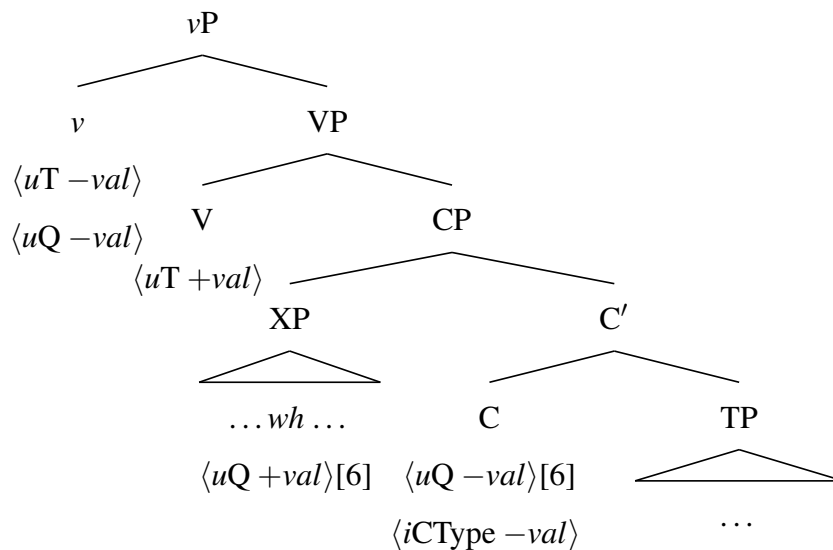
prevent embedded CP from closing. As opposed to the case of +Q-embedded clauses, the matrix V cannot close the embedded CP, as it does not value its Clause-type feature. Now, VP is built, and the matrix  $v$  is merged. As in the case of monoclausal sentences, T-features are shared between V and  $v$ . Recall that in the case of monoclausal sentences, VP will get closed at this moment, after merge of  $v$ , since all VP-*relevant* features are either valued or shared with a higher head. This is not the case now: VP cannot be closed. CP is a complement of V, and features of a V-complement are *relevant* at the VP-level. That prevents VP from becoming a phase upon insertion of  $v$ . Therefore, the first phasal domain in the matrix clause will be  $v$ P, which will get closed as soon as the matrix T is merged. After  $v$ P is closed, the *wh*-phrase will be inaccessible for further probing by a higher +Q-element, since it is located in the embedded Spec,CP position, within  $v$ P. The derivation will crash if no other features are present.

How can one save this derivation of long-distance *wh*-movement? One possibility is to place a Q-feature on  $v$ , which would trigger the movement of the *wh*-phrase to the Spec, $v$ P. The specification of this feature on  $v$  is  $\langle uQ -val \rangle$ . Since the matrix VP and the embedded CP are open for probing at the  $v$ P level, as I argued above,  $v$  will probe the *wh*-phrase in the Spec,CP, in order to value its Q-feature, and assuming that Q-feature on  $v$  is endowed with EPP-subfeature in *wh*-movement languages, *wh*-phrase will move to the Spec, $v$ P. We therefore derived the necessity of successive-cyclic *wh*-movement to the edge of  $v$ P, without resorting to stipulations about the phasal nature of  $v$ P.

Now let us consider how this derivation proceeds in details. Assume that the embedded *wh*-phrase is in the embedded Spec,CP position after it has been probed by the embedded C. Further assume that the matrix V has been merged with CP. As opposed to the case of +Q-complements, embedded CP remains open after merge of V. As before, assume that V comes with a valued uninterpretable instance of T-feature. Next,  $v$  with an uninterpretable unvalued instance of T-feature is merged with the matrix VP. This moment of the derivation is shown in (12).



(12) *The structure of the intermediate CP-domain in the case of -Q-complements.*



In (12),  $v$  probes  $V$  in order to value its  $T$ -feature; they become shared. At this moment  $VP$  cannot be closed: the  $Q$ -feature on the  $V$ -complement ( $CP$ ) does not have interpretation yet, and features of  $V$ -complement are  $VP$ -relevant. That results in both  $VP$  and  $CP$  being open for further probing. As I proposed earlier,  $v$  must also have a  $Q$ -feature, which is unvalued and uninterpretable, otherwise the derivation would crash.  $v$  probes  $XP$  in  $Spec,CP$  position, and the  $Q$ -features on  $XP$ , embedded  $C$ , and  $v$  all become instances of the same feature. While  $Q$ -features within  $CP$  are shared with a probing head ( $v$ ),  $VP$  still cannot be closed as the complement of  $V$  still lacks value on its Clause-typing feature, and  $VP$  still cannot acquire phasal status.

The summary of what I have argued above is given in (13).

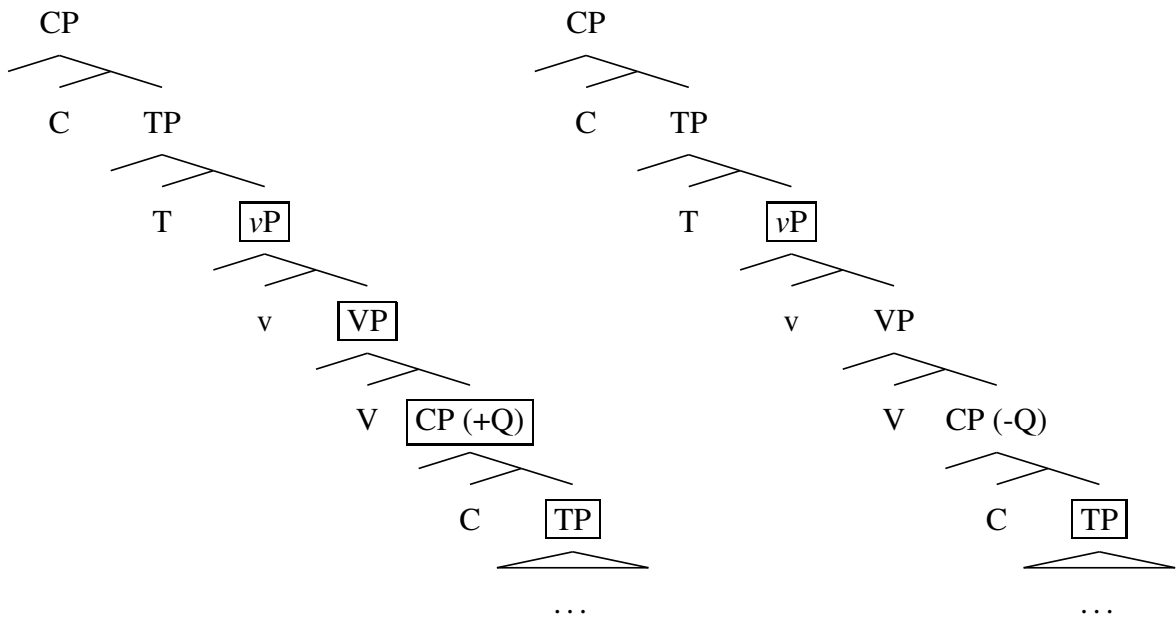
- (13)
- a. Embedded  $C$  which allows for successive-cyclic  $wh$ -movement comes with  $\langle uQ -val \rangle$  feature.
  - b.  $v$  which allows for successive-cyclic  $wh$ -movement comes with  $\langle uQ -val \rangle$  feature. If this feature is absent on  $v$ , the derivation crashes.
  - c. After the  $wh$ -phrase moves to the embedded  $-Q$   $Spec,CP$ ,  $CP$  is not a phasal domain, i.e.  $CP$  does not close upon merge of  $V$ .

- d. After the *wh*-phrase moves to the embedded +Q Spec,CP, CP becomes a phasal domain, i.e. CP closes upon merge of V.
- e. In the case of a -Q CP-complement, after *v* is merged with VP, it is still allowed to probe inside embedded CP. After its Q-feature enters in an Agree relation with the Q-feature on the embedded C/Spec,CP, VP and the embedded CP remain open for syntactic probing.
- f. Matrix T can probe the embedded Spec,CP in case of -Q-CP complements; it cannot probe the embedded Spec,CP in case of +Q-CP complements.

The following diagrams show which domains are phasal in case of +Q CP-complements and in case of -Q CP-complements; they are denoted by boxes.

(14) +Q CP-complements

(15) -Q CP-complements



### 5.1.3 Ambiguities in binding and long-distance *wh*-movement

Recall that English complex anaphors, such as *himself* come from the numeration with valued  $\phi$ -features. In the analysis proposed above it allows them to be bound at the level of VP or  $\nu$ P, without resorting to the mechanism for valuing of  $\phi$ -features: binding can occur as soon as the  $\rho$ -feature on the probe matches the  $\rho$ -feature on the goal, in this case on the reflexive itself.

Now in order to explain at which level binding can occur, we have to answer the following question: what are the possible positions for placing an interpretable instance of the  $\rho$ -feature in case of long-distance *wh*-movement of the reflexive?

Above I showed that CP with a +Q C-head closes upon merge of the matrix V, as all *relevant* features within it are valued and interpretable. This fact allows for probing the anaphor embedded in a *wh*-phrase, which is located in Spec,CP, by an element located in the matrix clause, thus establishing a binding relation between a matrix element and an anaphor. Schematically, this is shown in (16):

$$(16) \quad \dots V_{\langle ip - val \rangle} \dots [CP [ \dots himself_{\langle up + val \rangle} \dots ]_{\langle iQ - val \rangle [6]} C_{\langle uQ + val \rangle [6]} \dots \longrightarrow \\ \dots V_{\langle ip - val \rangle [7]} \dots [CP [ \dots himself_{\langle up + val \rangle [7]} \dots ]_{\langle iQ - val \rangle [6]} C_{\langle uQ + val \rangle [6]} \dots$$

Consider first the case of an interpretable Q-feature on the embedded C. In this case, the binding of the *wh*-phrase internal reflexive by matrix elements is possible, see (17).

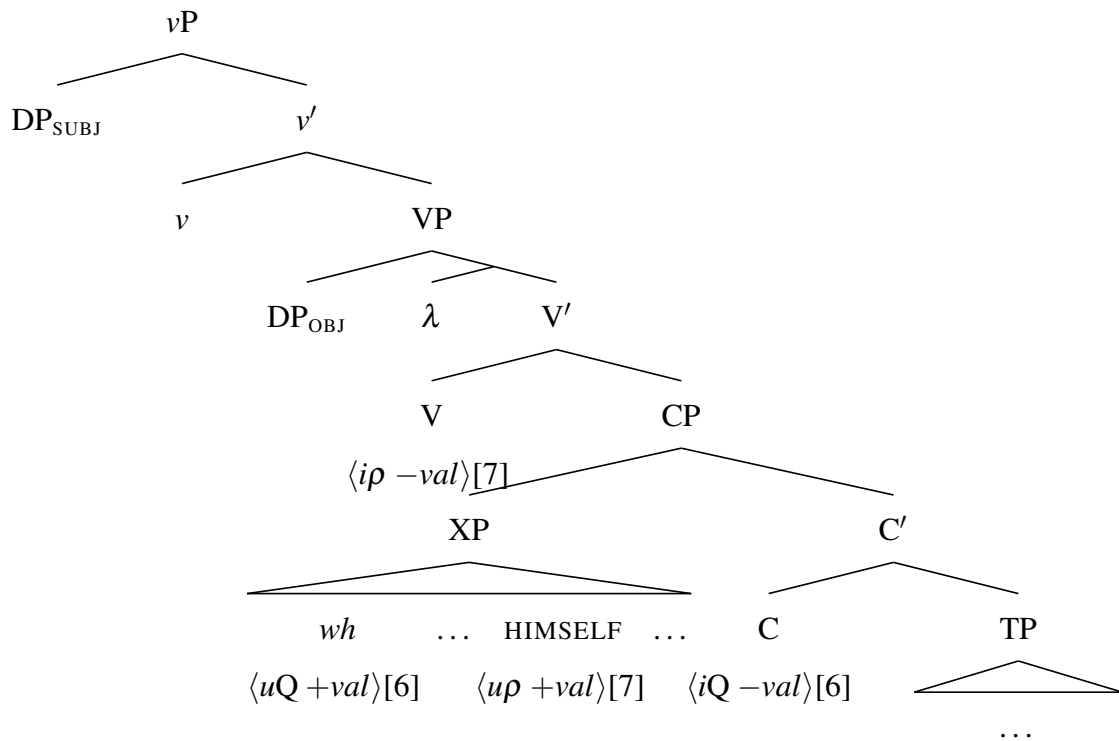
- (17) a. John<sub>i</sub> wonders which pictures of himself<sub>i</sub> Mary published in the newspaper.  
 b. John<sub>i</sub> asked Bill<sub>j</sub> which pictures of himself<sub>i/j</sub> Mary published in the newspaper.

As the examples above show, both the matrix subject and the matrix object can serve as antecedents for the reflexive. Consider how binding by a matrix element can be achieved in a feature-based theory of binding.

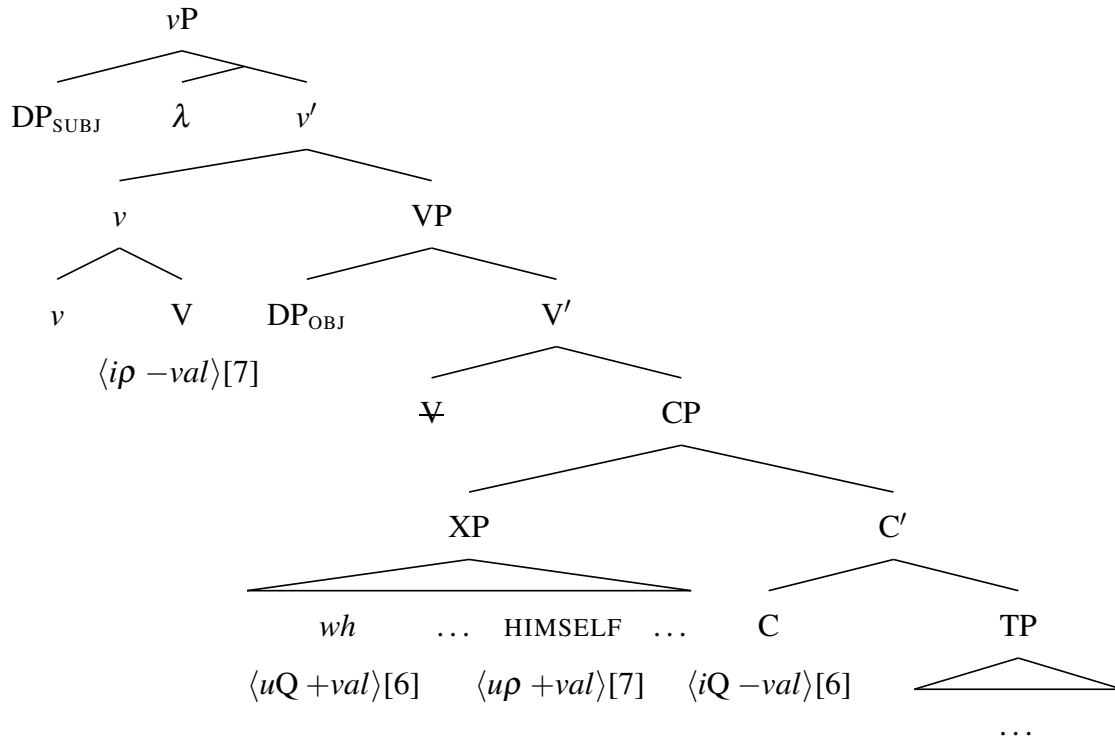
Assume that an interpretable instance of the  $\rho$ -feature,  $\langle ip - val \rangle$ , is placed on the matrix verb. As I argued above, embedded CP is not closed by the time the matrix verb is merged, and the

matrix verb must probe embedded C in order to value its Clause-typing features and establish a selectional dependency. If the matrix verb is endowed with the  $\rho$ -feature, it will probe first the reflexive in the Spec,CP position: their  $\rho$ -features will become shared: they will become instances of the same feature in two locations. Now this case is similar to the case of monoclausal sentences. As soon as the VP-domain acquires phasal status (after  $v$  is merged and probes V), the  $\lambda$ -operator will be inserted immediately above the matrix V, and that would result in binding of the reflexive in the embedded Spec,CP by a matrix element. In case binding relations are evaluated at the level of VP, before  $v$ -to-V raising (which is possible since VP is a closed domain), the binder of the reflexive will be the matrix object. If the binding relations are evaluated at the level of  $v$ P, after merge of T and after the Agree relation took place between T and  $v+V$ , the binder of the reflexive will be the matrix subject. These two situations are shown in (18) and (19). Notice, that only features relevant for binding in the current situation are shown.

(18) *Binding of the wh-internal reflexive by the object in the case of the +Q intermediate CP.*



- (19) *Binding of the wh-phrase internal reflexive by the subject in the case of the +Q intermediate CP.*



Lastly, notice that the binder must be in the clause immediately preceding the clause with the reflexive:

- (20) Steve<sub>i</sub> believes that [John<sub>j</sub> wonders [which pictures of himself<sub>\*i/j</sub>] Mary published in the newspaper].

Coreference of the reflexive with the subject of the root clause, *Steve*, in (20) is impossible. This fact is unsurprising: the intermediate clause will have all of its features valued by the time it is completed. In order to get the root subject to be the antecedent for the reflexive in the most embedded clause's Spec,CP position, the interpretable instance of the  $\rho$ -feature must be placed on the matrix V/v. This will however lead to a derivation crash: matrix V/v will not be able to probe into the intermediate CP, which is a closed phasal domain. Therefore, the interpretable  $\rho$ -feature

on the root clause element and the valued  $\rho$ -feature on the reflexive will not be able to enter in an Agree relation. That would result in two instances of the  $\rho$ -feature present in the structure by the time the derivation is completed: one of them will lack the interpretable instance, and one of them will lack value; the derivation will crash.

Now, the next case to consider is the case of C with an uninterpretable Q-feature,  $\langle uQ -val \rangle$ . Similarly to the previous case, the reflexive embedded in the *wh*-phrase can be bound by matrix elements, either the subject of the object, see (21).

- (21) a. Which pictures of himself<sub>i</sub> does John<sub>i</sub> think Mary likes.  
 b. Which pictures of herself<sub>i</sub> did John tell Mary<sub>i</sub> that Bill likes.

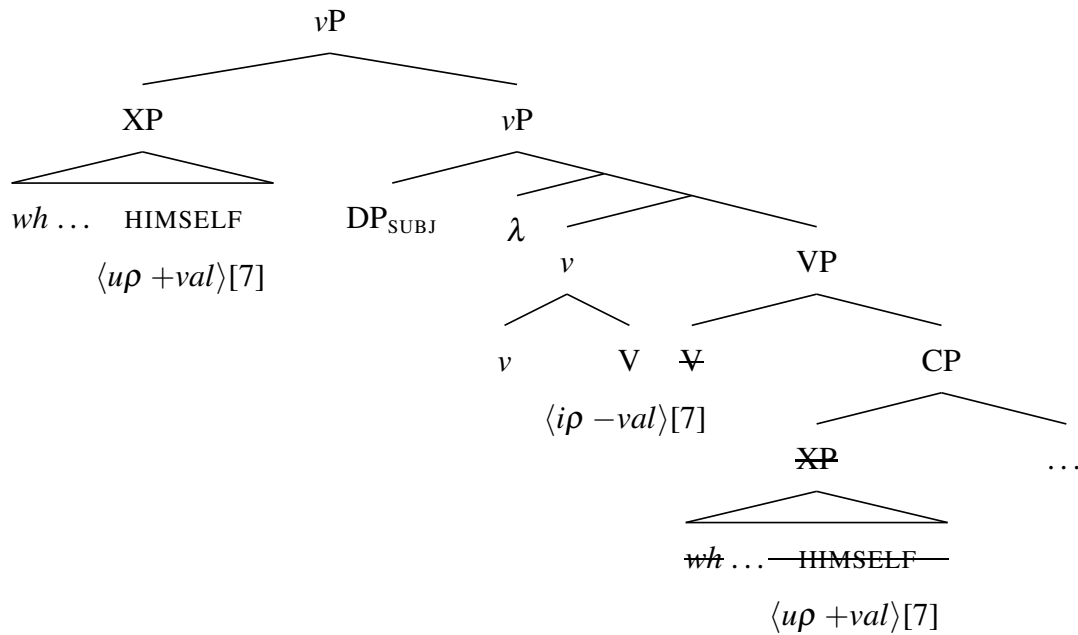
In the first example above, the reflexive *himself* can be bound by the matrix subject.<sup>1</sup> The second example shows that the reflexive *herself* can be bound by the indirect object of the matrix clause; again, any other binding in this example is impossible because of  $\phi$ -feature mismatch.

Above I showed that in this case in order for the derivation to succeed the matrix  $v$  must have a  $\langle uQ -val \rangle$ -feature, and it directly probes the *wh*-phrase in order to value this feature. As I argued, the embedded CP does not close until the matrix  $v$  is merged. Therefore an analysis similar to the previous case is possible. The  $\rho$ -feature can be placed on the matrix V or  $v$ . If it is present on V, then when  $v$  is merged, VP acquires phasal status, and the binding by the matrix object is possible. Otherwise, if the  $\rho$ -feature is present on  $v$ ,  $v$  will probe the reflexive, and the matrix subject will be the binder. Notice, that it does not matter that the *wh*-phrase further moves to the matrix Spec, $v$ P: the  $\rho$ -feature on it has already been probed by  $v$ , and after that keeping the c-command relation between the binder and the bindee is no longer needed: the  $\rho$ -feature on  $v$  will generate a  $\lambda$ -operator, which will bind the variable inside the *wh*-phrase, and  $\lambda$ -conversion will result in binding by the matrix subject:

---

<sup>1</sup>Of course, generally, it can also be bound by the embedded subject — however in this example  $\phi$ -feature mismatch between the reflexive and the embedded subject prevents it from happening.

(22) Binding of the *wh*-internal reflexive by the subject in the case of the *-Q* intermediate CP.



Notice, the same results can be achieved using a different placement of the interpretable instance of the  $\rho$ -feature: on the matrix T. In this case, T will probe the reflexive when the *wh*-phrase is located in the *vP*-adjoined position, and whenever the matrix CP phase will close, the binding will be obtained between the matrix subject in Spec,TP, and the reflexive. As I show next, this is the only possibility of binding the *wh*-phrase internal reflexive in languages with the monomorphemic reflexives, such as Russian.

## 5.1.4 Monomorphemic reflexives in Slavic: interaction of *wh*-movement and binding

### 5.1.4.1 The case of *-Q* complements

In this section I will show what the theory predicts for the case of languages with monomorphemic reflexives. Consider the data from Russian in (23)-(24) where the embedded C has an uninterpretable Q-feature,  $\langle uQ -val \rangle$ .

(23) *Subjunctive wh-movement (from Dyakonova 2009)*

- a. Ivan<sub>i</sub> xočet čtoby Olga<sub>j</sub> nakazala [svoego\*<sub>i/j</sub> syna].  
Ivan wants that<sub>SUBJ</sub> Olga punished self's son  
'Ivan wants that Olga punishes [her/\*his son]'
- b. [Kogo iz svoix detej] Ivan xočet čtoby Olga nakazala?  
which of self's kids Ivan wants that<sub>SUBJ</sub> Olga punished  
'[Which of her/his children] Ivan wants that Olga punishes?'

(24) *Indicative wh-movement*

- a. Ivan<sub>i</sub> skazal čto Olga<sub>j</sub> nakazala [svoego\*<sub>i/j</sub> syna].  
Ivan said that<sub>IND</sub> Olga punished self's son  
'Ivan said that Olga punishes [her/\*his son]'
- b. [Kogo iz svoix detej] Ivan skazal čto Olga nakazala?  
which of self's kids Ivan said that<sub>IND</sub> Olga punished  
'[Which of her/his children] Ivan said that Olga punishes?'

The data are parallel to English: when the *wh*-phrase contains a reflexive, long-distance *wh*-movement feeds new binding possibilities. While long-distance binding by matrix DPs is possible in neither indicative nor subjunctive cases, as soon as a *wh*-phrase containing the reflexive is *wh*-moved, the matrix subject can serve as an antecedent.<sup>2</sup>

The first major difference between Russian and English cases is related to subject orientation: in no case can the matrix object serve as a binder for the reflexive in Russian, see (25)-(26).<sup>3</sup>

- (25) a. Ivan skazal Maše<sub>i</sub> čtoby Olga<sub>j</sub> nakazala [svoix\*<sub>i/j</sub> detej].  
Ivan told Maša that<sub>SUBJ</sub> Olga punish self's children  
'Ivan told Maša<sub>i</sub> that Olga<sub>j</sub> should punish [her\*<sub>i/j</sub> children]'

---

<sup>2</sup>While not all speakers of Russian agree on grammaticality of long-distance *wh*-extraction out of indicative clauses, such as in (24), those who do agree on the fact that binding by the matrix subject is possible.

<sup>3</sup>Notice that both (25) and (26) use the same matrix verb *skazal* 'told', however select for different types of complements: subjunctive in (25) and indicative in (26). The meaning of these examples is however different in a way we expect from subjunctives and indicatives. The subjunctive examples express necessity that something should be done, while being *irrealis*, and indicative examples express that something have happened in the past. In these examples, in (25) Ivan expresses to Mary his desire that Olga punishes the children. In (26), Ivan tells Mary about the fact that Olga have already punished children.



- b. [Kogo iz svoix\*<sub>i/j/k</sub> detej] Ivan<sub>k</sub> skazal Maše<sub>i</sub> čtoby Olga<sub>j</sub> nakazala.  
 which of self's kids Ivan told Masha that<sub>SUBJ</sub> Olga punish  
 '[Which of her\*<sub>i/j</sub>/his<sub>k</sub> children] Ivan<sub>k</sub> told Mary<sub>i</sub> that Olga<sub>j</sub> should punish?'
- (26) a. Ivan skazal Maše<sub>i</sub> čto Olga<sub>j</sub> nakazala [svoix\*<sub>i/j</sub> detej].  
 Ivan told Maša that<sub>IND</sub> Olga punish self's children  
 'Ivan told Maša<sub>i</sub> that Olga<sub>j</sub> have punished [her\*<sub>i/j</sub> children]'
- b. [Kogo iz svoix\*<sub>i/j/k</sub> detej] Ivan<sub>k</sub> skazal Maše<sub>i</sub> čto Olga<sub>j</sub> nakazala.  
 which of self's kids Ivan told Masha that<sub>IND</sub> Olga punish  
 '[Which of her\*<sub>i/j</sub>/his<sub>k</sub> children] Ivan<sub>k</sub> told Mary<sub>i</sub> that Olga<sub>j</sub> have punished?'

Combination of the analysis of monomorphemic reflexive binding and the analysis of possibility of matrix binding of *wh*-moved reflexives provide an account for this data.

Notice that monomorphemic reflexives cannot be bound before they get their  $\phi$ -features valued. I argued that that  $\phi$ -feature valuation is mediated by T with a  $\rho$ -feature, which probes the subject in Spec, $\nu$ P, and further probes a monomorphemic reflexive. I showed that while in languages with complex anaphors, such as English, binding domains are VP and  $\nu$ P, in Russian and other languages with monomorphemic anaphors, binding applies only at the TP level, after the reflexive's  $\phi$ -features have been valued.

The same facts hold in case of *wh*-internal reflexives in Russian. Consider the moment when the *wh*-phrase occupies the embedded Spec,CP. If a  $\rho$ -feature is present on T (as I argued above, this is the only possibility in case of monomorphemic reflexives), T will first probe the subject DP in the Spec, $\nu$ P, and further probe  $\nu$ +V complex. This probing will result in  $\phi$ -feature sharing between T and the matrix subject first, and in T-feature sharing between T and  $\nu$ +V head complex second. Since T has a  $\rho$ -feature, it will also probe the reflexive, and that probing will result in  $\rho$ -feature sharing together with the parasitic valuation of  $\phi$ -features on the reflexive. Next, after the minimal phase containing the matrix TP is closed, binding will be able to apply: a  $\lambda$ -operator will be inserted at the position immediately above T, and after  $\lambda$ -conversion applies, the reflexive will end up being bound by the DP in Spec,TP position, i.e. by the matrix subject.

As before, the impossibility of binding by the matrix object stems from the fact that it is impossible for the reflexive to get its  $\phi$ -features valued through a relation with the matrix object.

#### 5.1.4.2 The case of +Q complements

As I showed above, the Russian facts related to reflexive binding are parallel to the English case: *wh*-internal reflexives interpreted in the Spec,CP of the matrix clause (the case of -Q embedded clauses) can be bound by matrix elements.

However, the situation with embedded +Q complements differs from the English case. Consider the data in (27).

- (27) a. Ivan<sub>i</sub> interesuetsja [kakie rasskazy o sebe\*<sub>i/j</sub>] Boris<sub>j</sub> pročital.  
 Ivan is interested [which stories about SELF] Boris read<sub>PAST</sub>.  
 ‘Ivan<sub>i</sub> is interested which stories about himself\*<sub>i/j</sub> Boris<sub>j</sub> has read.’
- b. Medvedev<sub>i</sub> rasskazal Putinu<sub>k</sub> [kogo iz svoix\*<sub>i/\*k/j</sub> ministrov] Obama<sub>j</sub> obvinil v korrupcii.  
 Medvedev told Putin [which of self’s ministers] Obama accused of corruption.  
 ‘Medvedev<sub>i</sub> told Putin<sub>k</sub> [which of self’s\*<sub>i/\*k/j</sub> ministers] Obama<sub>j</sub> accused of corruption’

The data above show that the reflexive within the *wh*-phrase in the case of indirect questions can only be bound by the embedded subject, and not by the matrix subject. For instance, in (27) only *Boris* and *Obama* can be antecedents of the possessive reflexive *sebe/svoix*.

The feature-based analysis of binding can account for this contrast between Russian and English. As I have shown above, in the case of monomorphemic reflexives, in order to get binding by the matrix subject, the  $\rho$ -features must be placed on the matrix T. Recall that in case of indirect questions, matrix V closes the embedded CP phase, and *v* closes the VP phase, see a tree in (14). As a result of that, matrix T with  $\rho$ -feature will not be able to probe the element located in the embedded Spec,CP. Therefore, the derivation will fail, since there is no way to value  $\phi$ -features of the reflexive. Since English does not rely on this mechanism for  $\phi$ -features valuation, the contrast

between English and Russian is predicted.

This analysis makes a prediction that the Russian complex reciprocal *drug druga* ‘each other’, which as I argued, has valued  $\phi$ -features, will pattern with English complex reflexive *himself/herself*. This is indeed the case as (28) shows:

- (28) a. *Deti interesujutsja [kakie fotografii drug druga] učitel’ pomestil v Internete.*  
Children<sub>i</sub> are interested [which pictures of each other<sub>i</sub>] teacher placed in Internet  
‘Children wonder which pictures of themselves the teacher placed in the Internet.’
- b. [*Putin i Medvedev*]<sub>i</sub> *interesujutsja [kakie sluxi drug o druge<sub>i</sub>] raspuskajut žurnalisty.*  
[Putin and Medvedev] are interested [which rumors about each other] spread  
journalists  
‘Putin and Medvedev wonder which rumors about each other do the journalists spread.’

### 5.1.5 On the impossibility of A'-binding

In this section I take on the issue of the impossibility of reflexives being bound by a DPs located in A' position. The data below show that *wh*-phrases cannot serve as a binder from an A'-position where it have moved.

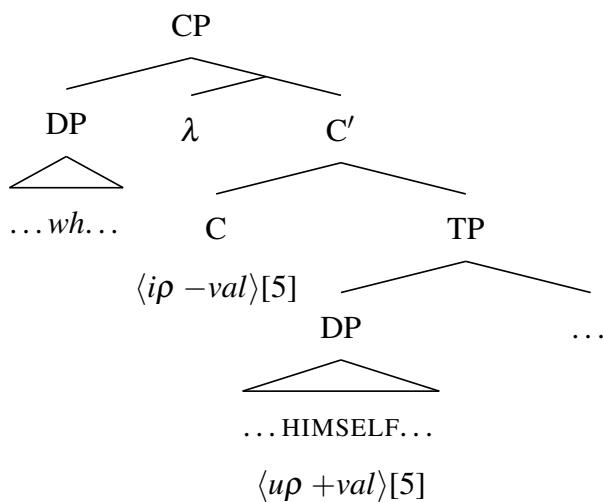
- (29) a. \*John wonders who<sub>i</sub> stories about himself<sub>i</sub> describe.  
b. \*Mary asked which students<sub>i</sub> professors of each other<sub>i</sub> want to kick out of school.
- (30) a. \*Which person<sub>i</sub> do you think (that) children of himself<sub>i</sub> will give flowers to?  
b. \*Which vice-president<sub>i</sub> do you believe (that) employees of himself<sub>i</sub> hate?

In all cases above the *wh*-moved DP either stays in the intermediate Spec,CP, or moves through it, resulting in a c-command relation between it and the reflexive inside the embedded subject. However, the coindexation between the *wh*-phrase and the subject-internal reflexive is impossible.

The theory I proposed above can account for this impossibility without stipulating that A'-binding is impossible.

Consider the case of long-distance *wh*-movement in English. In order to get the reflexive embedded into subject of the embedded clause bound by the *wh*-phrase in the intermediate Spec,CP, one has to postulate that the interpretable instance of the  $\rho$ -feature can probe the subject-internal reflexive, and will generate a  $\lambda$ -operator located between Spec,CP and embedded Spec,TP. This is the only case when the  $\lambda$ -conversion can result in the reflexive being bound by a DP in the intermediate Spec,CP. This hypothetical situation is shown in (31).

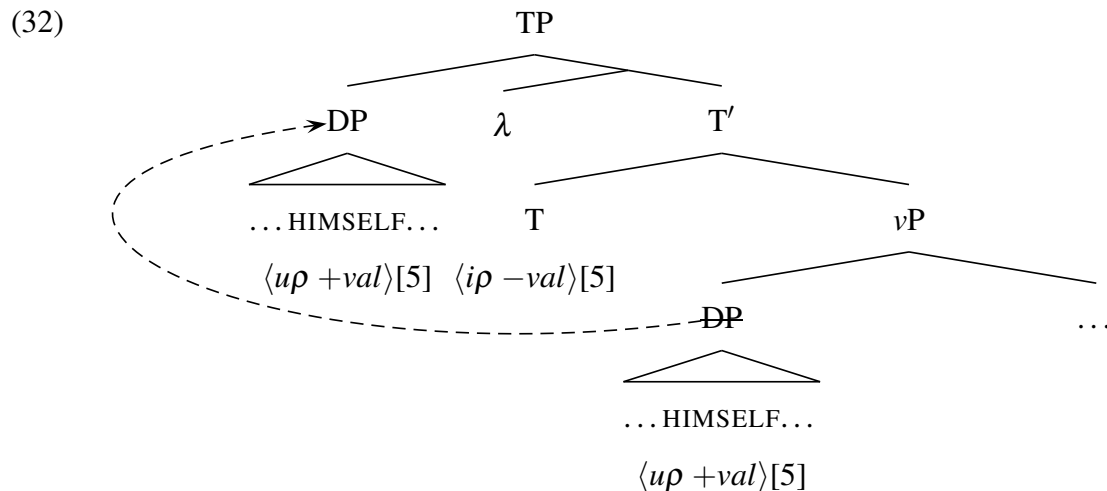
(31) Placement of the  $\rho$ -feature on C.



While it looks like this configuration will result in an undesired binding relation, there is a good reason why it is not so. Chomsky 2008 claims that features of C are inherited by T. While he claims that Agree features are definitively inherited, he remains undecided on whether all C features transfer to T. Taking on this suggestion, I assume that it is impossible to have a  $\rho$ -feature on C without having the same instance of a  $\rho$ -feature be present on T.

Now, if the  $\rho$ -feature is in fact present on T, binding of the Spec,TP-internal reflexive by the element in Spec,CP will not happen. In fact, placing a  $\rho$ -feature on T will give a reading where the reflexive within the subject is bound by the subject itself, i.e. *i*-within-*i* reading. Consider in details how this happens.

Consider the moment when  $vP$  is constructed, and the subject is present in  $\text{Spec},vP$  position.  $T$  is merged, and as I assumed above, it has an interpretable instance of the  $\rho$ -feature. The uninterpretable valued instance of the  $\rho$ -feature is currently present inside the nominal occupying the  $\text{Spec},vP$  position. An Agree relation is established between the  $\rho$ -feature on  $T$  and the  $\rho$ -feature on the subject-internal reflexive. Further, the subject undergoes raising from  $\text{Spec},vP$  to  $\text{Spec},TP$ . Upon merge of  $C$ ,  $TP$  will become a phasal domain, since all features within  $TP$  will be either valued or shared with  $C$ . The  $\lambda$ -operator will be inserted immediately above  $T$ . The result of this would be binding of the subject internal variable by the element in  $\text{Spec},TP$ , i.e. subject itself. This situation is shown in (32).



While it seems like this binding relation, when the subject internal reflexive is bound by the subject itself, violates the  $i$ -within- $i$  condition, the status of this condition is not uncontroversial with respect to reflexives. The following sentences are judged at worst as mildly deviant by my informants, while most accept them as grammatical.

- (33)
- a. ?[The man proud of himself<sub>i</sub>]<sub>i</sub> was a famous movie-star.
  - b. ?[The book about itself<sub>i</sub>]<sub>i</sub> was written by Italo Calvino.
  - c. ?[Fans of themselves<sub>i</sub>]<sub>i</sub> are usually pompous jerks.

I will leave the question of why some of the sentences violating *i*-within-*i*-condition are much worse than the ones in (33) unresolved; for further discussion of *i*-within-*i*-condition, one might resort to Hoeksema and Napoli 1990.

## 5.2 Interaction with scrambling: The case of Russian

In this section I will review the interaction of binding and scrambling. Consider the case of Russian. Consider the examples in (34). The object of the embedded clause can be scrambled to a position inside the matrix clause. While there is some disagreement on the relative grammaticality of scrambling out of indicative clauses versus scrambling out of subjunctive clauses, most speakers accept both of them as at most mildly deviant.<sup>4</sup>

(34) *The object Long-Distance Scrambling out of indicative complements*

- a. ?Ivan skazal čto Boris kupil knigu  
Ivan said that Boris bought book  
'Ivan said that Boris bought a book'
- b. ?Ivan knigu skazal čto Boris kupil t  
Ivan book said that Boris bought

(35) *The object Long-Distance Scrambling out of subjunctive complements*

- a. Ivan xočet čtoby Boris kupil knigu  
Ivan said that<sub>SUBJ</sub> Boris bought book  
'Ivan wants Boris to buy a book'
- b. Ivan knigu xočet čtoby Boris kupil t  
Ivan book wants that<sub>SUBJ</sub> Boris bought

For the purposes of further discussion I limit myself to object scrambling, and assume that mild deviance of scrambling in case of indicative complements, (34), is irrelevant for my discussion of the interaction of binding and scrambling.

---

<sup>4</sup>There are further disagreements on whether subject scrambling is grammatical. It is generally assumed that subjects cannot be scrambled out of indicative complements, and their scrambling out of subjunctives leads to deviance. For the purposes of the discussion here, I will not attempt to explain these asymmetries.

The question one has to consider now is what happens if a reflexive (or a DP containing a reflexive) is scrambled from the embedded clause into the matrix clause. The relevant data is given below in (36)-(38).

(36) *Indicative*

- a. Ivan<sub>i</sub> skazal čto Boris<sub>j</sub> narisoval sebja<sub>\*i/j</sub>  
 Ivan said that Boris draw self  
 ‘Ivan<sub>i</sub> said that Boris<sub>j</sub> drew himself<sub>\*i/j</sub>’
- b. Ivan<sub>i</sub> sebja<sub>\*i/j</sub> skazal čto Boris<sub>j</sub> narisoval t  
 Ivan self said that Boris drew  
 ‘Ivan<sub>i</sub> said that Boris<sub>j</sub> drew himself<sub>\*i/j</sub>’

(37) *Subjunctive*

- a. Ivan<sub>i</sub> xočet čtoby Boris<sub>j</sub> narisoval sebja<sub>\*i/j</sub>  
 Ivan wants that<sub>SUBJ</sub> Boris draw self  
 ‘Ivan<sub>i</sub> wants Boris<sub>j</sub> to draw himself<sub>\*i/j</sub>’
- b. Ivan<sub>i</sub> sebja<sub>i/j</sub> xočet čtoby Boris<sub>j</sub> narisoval t  
 Ivan self wants that<sub>SUBJ</sub> Boris draw  
 ‘Ivan wants Boris to draw him/himself’

(38) *Infinitive (object control)*

- a. Ivan<sub>i</sub> skazal Borisu<sub>j</sub> PRO<sub>j</sub> narisovat’ sebja<sub>i/j</sub>  
 Ivan told Boris draw<sub>INF</sub> self  
 ‘Ivan<sub>i</sub> told Boris<sub>j</sub> to draw himself<sub>i/j</sub>’
- b. Ivan<sub>i</sub> sebja<sub>i/j</sub> skazal Borisu<sub>j</sub> PRO<sub>j</sub> narisovat’ t  
 Ivan self told Boris draw<sub>INF</sub>  
 ‘Ivan<sub>i</sub> told Boris<sub>j</sub> to draw himself<sub>i/j</sub>’

Recall the case of *wh*-phrase internal reflexives. No matter whether the *wh*-movement took place out of an indicative complement or a subjunctive complement, it fed new binding possibilities. A *wh*-displaced reflexive could be bound by the matrix subject, resulting in ambiguity of the corresponding interrogative sentence. The examples of scrambling above show an interesting difference from the case of *wh*-movement.

Consider first the case of subjunctive complements in (37). The object reflexive *sebj*a was scrambled from the embedded clause to the *v*P-peripheral position in the matrix clause. In its base position, the only possible binder of the reflexive is the embedded subject, *Boris*. The example with the scrambled reflexive is ambiguous. Both the embedded and the matrix subjects can now serve as antecedents of the reflexive. This is immediately reminiscent of the case of *wh*-movement.

The case of control complements in (38) is unsurprising: the sentence with the reflexive in the base position is already ambiguous. I proposed an analysis of how this ambiguity is achieved in Chapter 4. Clearly, movement of the reflexive should not destroy already established binding relations.

Now consider the case of indicative complements, (36). In the base position the reflexive can only be bound by the embedded subject. Scrambling of the reflexive however does not feed new binding possibilities: regardless of the fact that the reflexive is now in a local configuration with the matrix subject, coreference between them is still impossible. The lack of new binding possibilities is what distinguishes the case of scrambling out of indicative clauses from other cases of long-distance movement I considered above.

The summary of data is given below in (39).

(39) *Summary of data:*

	LD binding	Binding and Scrambling
Indicative	✗	✗
Subjunctive	✗	✓
Infinitive	✓	✓

In the remaining part of this section I show how my analysis predicts the impossibility of new binding relations when a reflexive is scrambled out of an indicative clause. The intuition behind the fact that the new binding relations are impossible when scrambling a reflexive out of the indicative clause is the following.



Recall the case of *wh*-movement. In Chapter 3 I proposed that binding applies at the level of phasal domains. What allows the *wh*-phrase internal reflexives to escape establishing binding relations in the embedded clause is the fact that the uninterpretable Q-feature on an element in the embedded clause prevent projections in the embedded clause to become phasal domains. Now assume that scrambling is not an interpretative phenomenon, i.e. it does not involve feature checking/valuation/interpretation (for further discussion of scrambling see Bailyn 2001, 2003, 2006; Miyagawa 2005; Saito 1992, 2003, 2005 and references therein).<sup>5</sup>

First consider the case of the indicative clause. By the time the embedded TP is completed, there are no unvalued/uninterpretable features left: the tense features are valued and interpretable, there is no Q-feature to prevent TP from closing. TP is a phasal domain, and therefore an interpretable instance of the  $\rho$ -feature *must* be placed on the T-head: placing it higher, on a matrix clause element, will lead to a derivation crash, since it will not be able to get its value from the reflexive. As a result, binding *must* apply mandatorily at the level of the embedded TP, and that would results in reflexive being bound by the embedded subject.

Now, the case of scrambling out of subjunctive embedded clauses is similar to the case of long-distance binding into subjunctive complements in the case of V-raising languages. Recall that the lack of value for the T-feature in the embedded clause allows establishing binding relations to be postponed until the matrix clause T is merged. Here, in Russian, scrambling of the reflexive plays the same role as the V-to-T raising in V-raising languages: it allows the reflexive to escape from  $v$ P-phase. Therefore, if the interpretable instance of the  $\rho$ -feature is placed on the matrix T, it would be able to probe the corresponding instance of the  $\rho$ -feature on the reflexive, which would result in binding relation between the matrix subject and the scrambled reflexive.

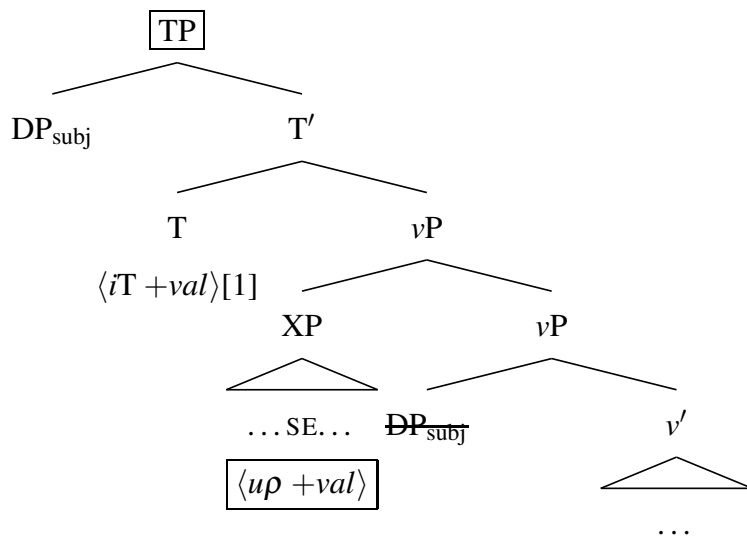
Now I will show how the derivation proceeds in details. First notice that in order to get a reflexive bound by a matrix element, the interpretable instance of  $\rho$ -feature must be present on the matrix T, and not on the embedded T, otherwise one would end up with the local binding. All derivations below proceed under this assumption.

---

<sup>5</sup>Another possibility is to propose that scrambling involves feature-checking, but the corresponding feature is not relevant (see definition (64)) for the respective domains.

Consider the case of indicative complements at the moment the embedded subject is moved to the Spec,TP position. At this moment the phrase which is undergoing scrambling and contains the reflexive is located in the  $vP$ -adjoined position. At this moment, TP has no TP-relevant features which are unvalued/uninterpretable. In fact, the only unvalued interpretable feature within the embedded TP is the valued uninterpretable instance of  $\rho$ -feature on the reflexive. That results in the embedded TP becoming a phasal domain. As a result, matrix T will not be able to probe into the embedded TP in order to get its value and provide the  $\rho$ -feature on the reflexive with an interpretable instance of  $\rho$ -feature. The embedded TP at this moment is shown in (40).

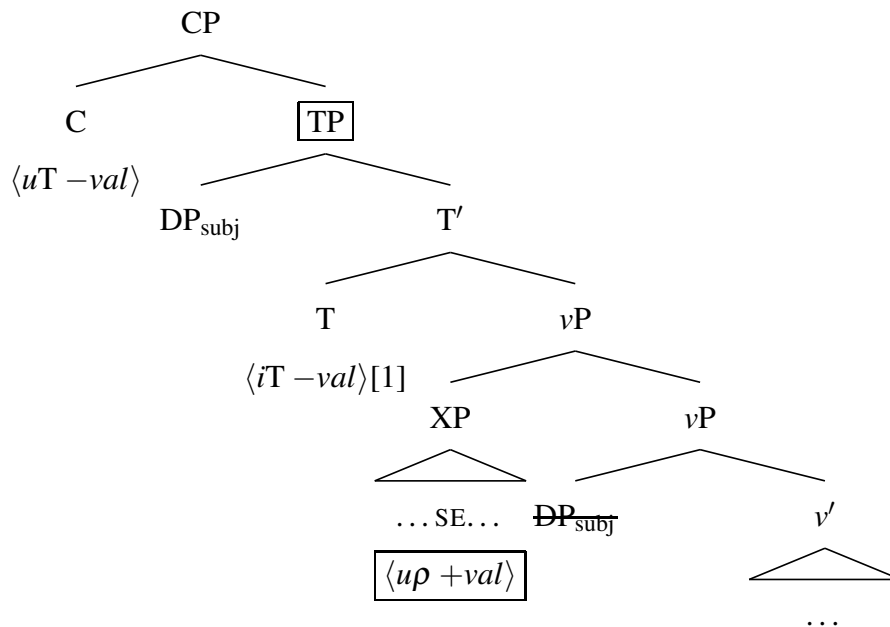
(40) *The embedded indicative TP with an instance of a scrambled reflexive.*



Notice that in this case TP-closure happens before the scrambled phrase is moved out of it. Since scrambling is a semantically vacuous movement, it is not driven by interpretable features, and therefore allowed to happen out of closed phases, in this case, the phrase which undergoes scrambling will be able to move from the  $vP$ -adjoined position to CP-adjoined position after TP-closure.

Now, contrast this case with the case of scrambling out of subjunctive complements. Similar moment in the derivation of the subjunctive complement is shown in (41).

(41) *The embedded subjunctive CP with an instance of a scrambled reflexive.*



In this case the embedded verb does not have a valued T-feature, and therefore after T and  $v/V$  enter in an Agree relation, TP still cannot be closed before T-feature is shared with some higher element, for instance embedded C. The derivation proceeds by merging embedded C endowed with an uninterpretable unvalued T-feature, which probes embedded T's T-feature: the T-feature on T and C become instances of the same feature. As soon as C is merged, Spec,CP becomes an available position for movement of the XP, and it moves there before TP becomes a phasal domain. Now, the  $\rho$ -feature on the reflexive has escaped the phasal domain, and is no longer inaccessible for further probing. This movement of the XP to the edge of CP, triggered not by the features of C, but by the requirement of the scrambled element (see Bošković 2007), allows it to escape TP, and make a reflexive within it accessible for probing by the matrix T.

Recall that in Chapter 4 I demonstrated that an embedded subjunctive CP prevents matrix VP/ $vP$  from closing before the matrix T is merged. That allows matrix T to probe the reflexive within the embedded Spec,CP position in case the matrix T has a matching feature. Therefore, placing  $\rho$ -feature on the matrix T will result in matrix T first probing the matrix subject for  $\phi$ -

features, then matrix  $v/V$  for T-features, as finally the reflexive within the embedded Spec,CP position. This results in valuation of  $\phi$ -features on reflexive as a byproduct of  $\rho$ -probing. The binding will therefore apply at the level of matrix TP, and the reflexive will be bound by the matrix subject.

This results in ambiguity of sentences with the reflexives scrambled out of subjunctive complements.

### 5.3 Conclusion

In this chapter I demonstrated the source of ambiguity in sentences where a reflexive has been dislocated into the matrix clause. I argued that ambiguity arises in cases when there are unvalued features in the embedded clauses which prevent phase closure and therefore allow elements to escape being buried within the closed phase. One such feature is the Q-feature, which while unvalued, allows the element within the reflexive to move to the position accessible by a  $\rho$ -feature within the matrix clause, resulting in binding by the matrix element. Another such feature is an unvalued T-feature, such as T in subjunctives, which postpones the closure of TP, and therefore allows the scrambled element to escape TP, allowing for the late application of binding. This goes along with the general claim of my dissertation that unvalued/uninterpretable features prevent certain domains from being sent to the interpretative component and allow them to stay “open,” and that this domain transparency allows binding relations to be evaluated at a later moment.

### 5.4 Consequences for non-binding phenomena

In this section I explore further consequences of the definition of a phasal domain provided in Chapter 2.5.4. This definition is repeated below in (42):

- (42) **Phasal domain, Final definition:** A domain  $D$  is *phasal* if for every  $D$ -relevant feature  $F$  within  $D$ ,

- a. F is valued, or
  - b. a head H merged with D has an instance of a feature F that entered a sharing probe-goal relationship with an instance of a feature F within D.
- (43) **D-relevant features:** Assume that D is a domain, such as VP,  $\nu$ P, TP, or CP. Then, *D-relevant features* are the following:
- a. If D=VP, *VP-relevant* features are:  $\phi$ - and Case-features of the nominals within V-complement and VP specifier.
  - b. If D= $\nu$ P,  *$\nu$ P-relevant* features are:  $\phi$ -features of the Spec, $\nu$ P, i.e.  $\phi$ -features of the subject, interpretation related features on  $\nu$ , i.e. Definiteness feature.
  - c. If D=TP, *TP-relevant* features are: T-features of T and  $\phi$ - and T-features of the element in Spec,TP.
  - d. If D=CP, *CP-relevant* features are: Clause-type feature on C and *wh*- and Q-features of C and Spec,CP.

According to this definition, for any domain D, if all D-relevant features have either been valued or have been shared with a feature on a higher head outside of the domain, the domain D “closes” for any further syntactic probing into it. If there are D-relevant features which are neither valued nor have found a match on some higher head, this domain D remains open for further probing.

If the phasal domain D contains an unvalued *non-D-relevant* feature which has neither been shared with a higher head nor has been valued, the derivation crashes since no further probing would result in valuation of this feature; this feature remains buried within D.

As one can see, this is a *relativized* definition of a phase: a domain D is phasal if all D-relevant features, i.e. features which are traditionally assumed to be active and checked within this domain, are valued or shared. According to (43), these features are in general defined as features of the head and the specifier of the domain D, with the exception of VP, for which the features of the verb complement are also relevant. Only these features are relevant for the closure of the domain. The presence of an element with unvalued features within a domain D does not prevent D from closing

unless these features are D-relevant. As I showed above, the presence of an anaphor within the domain D in general does not prevent it from closing, as the  $\phi$ -features of the anaphor are not phasal for every possible domain D. For instance, the presence of the object reflexive does not prevent the closure of the  $\nu$ P, TP, etc. If the reflexive, say, in Russian, lacks a value for its  $\phi$ -features, the indicative TP without an interpretable instance of the  $\rho$ -feature on T would still be closed as soon as C is merged and has finished probing, as  $\phi$ -features of the reflexive are not relevant at the TP level. This derives the traditional locality of Principle A. Feature relevance defined in this way prevents domains from staying open indefinitely, and in general prevents long-distance phenomena targeting an unvalued element in a domain with all other relevant feature valued. The absence of long-distance binding into indicatives in languages like Russian and English is the result of this definition.<sup>6</sup>

In next subsections I explore how this definition of a phase can be applied to other syntactic phenomena.

### 5.4.1 Long-distance extraction

It has been traditionally assumed that long-distance extraction, such as *wh*-movement, proceeds successive cyclically through the edges of phases. One of the main arguments for intermediate landing sites of a *wh*-phrase comes from consideration of Irish complementizers, which change depending on whether the *wh*-phrase landed in an intermediate Spec,CP position on its way to a position where it gets its interpretation (McCloskey 1990).

Irish has three complementizers:  $a^L$ ,  $a^N$ , and *go*<sup>7</sup>. The typical example of Irish complementa-

---

<sup>6</sup>Not all languages prohibit long-distance binding into indicatives, Chinese being an example of a language with virtually unconstrained long-distance binding. There are two possible ways to approach this problem. First, while one of the goals of this dissertation was to reduce the reliance on logophors in explaining binding phenomena, logophors do exist in some languages, see my discussion of logophors in section 1.2.3. Chinese *ziji* might be an instance of a logophoric pronoun; only *ta-ziji* might be the “real” anaphor in Chinese.

The second approach is to employ the idea that Chinese lacks tense altogether and must rely on certain other mechanisms in determining semantic tense anchoring of the event, see for instance Lin 2005. This reliance on higher projections might allow binding to be long-distance. I leave this issue for future research.

<sup>7</sup>L triggers ‘lenition’ on following word; N triggers ‘nasalization’ on following word.

tion in given in (44).

- (44) Deir siad [<sub>CP</sub> gur [<sub>TP</sub> ghoid na síogaí í ]].  
say they GO-PAST stole the fairies her  
'They say that the fairies stole her away.'

The complementizer  $a^L$  occurs in structures in which *wh*-movement has occurred, in every position through which *wh*-movement has taken place. In the theory assumed in this dissertation, complementizer  $a^L$  surfaces only when C is endowed with an instance of a Q-feature.

- (45) a. Cé [ $a^L$  bhuaíl tú]?  
who C struck you  
'Who did you hit?'
- b. Cé  $a^L$  mheas tú  $a^L$  chonaic tú?  
who C thought you C saw you  
'Who did you think that you saw?'
- c. Cén t-úrscéal  $a^L$  mheas mé  $a^L$  dúirt sé  $a^L$  thuig sé?  
which novel C thought I C said he C understood he  
'Which novel did I think he said he understood?'
- d. Níl a fhios agam cén t-úrscéal  $a^L$  mheas mé  $a^L$  dúirt sé  $a^L$  thuig sé.  
I don't know which novel C thought I C said he C understood he  
'I don't know which novel I thought he said he understood.'

Further, in the approach by Chomsky 2001,  $\nu$ P's are also phasal domains, and therefore successive cyclic movement must proceed through the edge of  $\nu$ P as well. I will consider the diagnostics for this later in this section, but now let me review the evidence why this is necessary in the revision of phase theory given in this dissertation. This approach was outlined in section 5.1.1. I showed why successive cyclic movement must proceed to the edges of  $\nu$ P and CP, and how it follows from the definition of a phase provided in (42). The crucial point here is that the redefinition of a phase as in (42) makes the same predictions as the traditional phase theory.

The intuition behind this approach was the following. Consider a simple sentence with the object *wh*-phrase, such as the one in (46):

(46) What does John think that Mary bought  $t$ ?

Consider the moment the embedded T is merged. It probes  $v$  for T-features, and after this probing all  $v$ P-relevant features within the  $v$ P are valued. Crucially, the Q-feature on the *wh*-phrase is not  $v$ P-relevant and therefore will not prevent the closure of the embedded  $v$ P. As a result, no higher element will be able to probe the *wh*-element, and the derivation will crash. The only way to prevent this derivation from crashing is to place an uninterpretable unvalued instance of the Q-feature on  $v$ . In this case, the Q-feature becomes relevant for  $v$ P, and as a result  $v$ P will not close. At the same time, placing an instance of the Q-feature on  $v$  would result in movement of the *wh*-phrase to the Spec, $v$ P, assuming that Q-features in English are universally strong (or endowed with the EPP subfeature).

Similar considerations can be applied to the intermediate CP. Failing to place a Q-feature on the intermediate C would result in a crashed derivation. Placing an uninterpretable unvalued instance of the Q-feature on the intermediate C would result in movement of the *wh*-phrase to the intermediate Spec,CP.

Note that there is no such need to place an uninterpretable unvalued instance of the Q-feature on T. After the merge of T,  $v$ P will remain open since Q is a  $v$ P-relevant feature. Now, when C is merged, it will be able to access the edge of  $v$ P since  $v$ P is not phasal. After C is merged and done probing, TP will become a phasal domain. However at this moment the *wh*-phrase will already be moved to the edge of CP, and therefore will be accessible for further probing.

Successive cyclic movement is therefore a natural consequence of the theory of phases proposed in this dissertation. The definition of a phase proposed in (42) requires placement of an uninterpretable unvalued instance of the Q-feature on every intermediate Spec,CP and every intermediate Spec, $v$ P for the derivation to succeed, similar to Chomsky 2001.

One further consequence of the phase theory proposed here is the existence of *wh*-islands, such as the ones shown in (47). In these examples movement of one *wh*-element (marked with an index *i*) to the intermediate Spec,CP blocks further movement of the other *wh*-phrase (marked with an



index  $j$ ) to the matrix Spec,CP.

- (47) a. \*How <sub>$j$</sub>  do you wonder [<sub>CP</sub> what <sub>$i$</sub>  John fixed  $t_i$   $t_j$  ]?  
b. \*What <sub>$j$</sub>  did you wonder [<sub>CP</sub> who <sub>$i$</sub>   $t_i$  saw  $t_j$  ]?

In the examples in (47), the embedded CP has an interpretable instance of the Q-feature,  $\langle iQ - val \rangle$ . As I argued above in 5.1.1, such a CP is a phasal domain: at the moment the *wh*-phrase moved to the Spec,CP, and the matrix V merged into the structure, all CP-relevant features are valued, and therefore CP becomes closed for further probing. The result of that is the impossibility of extraction of the second *wh*-phrase from the closed CP. *Wh*-islands can therefore be derived by using the definition of a phase proposed in (42): the phasal nature of the embedded CP prevents further probing into it.<sup>8</sup>

## 5.4.2 Parasitic gaps

A further diagnostic of phasehood is related to the distribution of parasitic gaps. The idea is presented by Legate 2003 based on Nissenbaum 1998. Nissenbaum claims that in order to license a parasitic gap, a *wh*-phrase must undergo movement to the Spec, $\nu$ P. Legate 2003 uses this as a diagnostics of phases. In traditional phase theory, movement of a *wh*-phrase to the edge of  $\nu$ P is

---

<sup>8</sup>Not all islands can be derived using phases. For instance, subject islands, such as in (i)

- (i) \*Who did [<sub>DP</sub> rumors about  $t$ ] surprise?

are not necessarily phasal phenomena. The extraction out of the exact same DP if it is in the object position is grammatical, (ii):

- (ii) Who does John spread [<sub>DP</sub> rumors about  $t$ ]?

Similarly, adjunct islands (shown in (iii), from Adger 2004) can also be problematic for phase theory approach.

- (iii) a. Hephaestus had run away, before the executioner murdered Hera.  
b. \*Who had Hephaestus run away, before the executioner murdered  $t$ ?

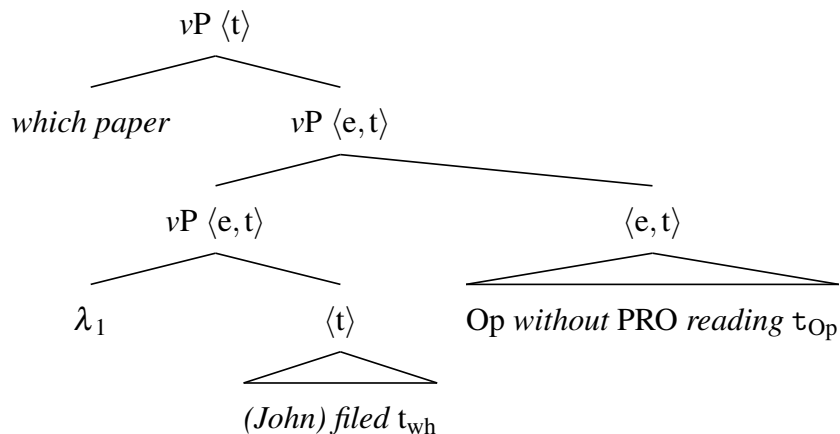
I leave the exploration of these and other types of islands for future research.

mandatory: movement directly to the Spec,CP violates the Phase Impenetrability Condition.

(48) Which paper did John file \_\_\_ without reading \_\_\_.

Legate 2003 claims that “Nissenbaum’s idea is that the structure would be interpretable if (a) a *wh*-phrase from the main *vP* moved to adjoin to *vP*, creating a lambda abstract; and (b) the adjunct clause containing the PG merged countercyclically just below the root.” (Legate 2003, pg. 510). The structure provided by Legate 2003 to demonstrate Nissenbaum’s 1998 approach to parasitic gaps is given in (49):

(49) *The structure of (48) (from Legate 2003).*



In my approach to phases discussed in Section 2.5.4 *vP* can be closed only after *T* is merged into structure. There are however ways to prevent the embedded *vP* from closing. For instance, in Russian embedded control clauses with the monomorphemic reflexive in object position, the embedded *vP* is not phasal, and that allows long-distance binding into Russian control clauses. The prediction one would make is the following: Since the embedded control *vP* with the monomorphemic reflexive is not phasal in Russian, it should fail the parasitic gaps diagnostic for phases, i.e. parasitic gaps should not be licensed within the embedded *vP*. For instance, consider the Russian sentences in (50).

- (50) ?Kogo Putin<sub>i</sub> poprosil Medvedeva predstaviti' \_\_\_ sebe<sub>i</sub> [ ne proverjaja \_\_\_].  
 who<sub>ACC</sub> Putin asked Medvedev introduce<sub>INF</sub> SE<sub>DAT</sub> not checking  
 'Who did Putin<sub>i</sub> asked Medvedev to introduce to self<sub>i</sub> without checking.'

In order to achieve the specified coreference, the embedded *vP* should remain open for probing by the matrix *T*. Since the embedded *vP* is not phasal in this case, the *wh*-phrase does not have to move through its edge, and therefore the parasitic gap should not be licensed. The example in (50) should contrast with the example in (51) with no anaphor.

- (51) ?Kogo Putin poprosil Medvedeva predstaviti' \_\_\_ Obame [ ne proverjaja \_\_\_].  
 who<sub>ACC</sub> Putin asked Medvedev introduce<sub>INF</sub> Obama<sub>DAT</sub> not checking  
 'Who did Putin asked Medvedev to introduce to Obama without checking.'

There are however complications with this prediction. It is unclear that Russian constructions as in (51) and (50) are in fact parasitic gaps. In fact, Russian allows null-subjects in positions which are not licensed by overt *A'*-movement, such as the one in (52):

- (52) Putin vybrosil etu knigu [ ne čitaja \_\_\_].  
 Putin threw away this book not reading  
 'Putin threw away this book without reading it'

There is no overt *A'*-movement in (52), however (52) is grammatical. If it is in fact a parasitic gap, then the conditions on licensing of parasitic gaps must be different in Russian than in English. In fact, Ivlieva 2007 proposes that licensing of parasitic gaps in Russian is conditioned by covert movement to a Topic position, and the example in (52) is in fact a case of parasitic gap. In this case there must be a contrast between the examples in (53). It seems to me that the first sentence with the monomorphemic reflexive is in fact worse than the second sentence without it, but further studies need to be done to show it with certainty.

- (53) a. ??Putin<sub>i</sub> poprosil Medvedeva podarit' sebe<sub>i</sub> etu knigu ne čitaja \_\_\_\_.  
 Putin asked Medvedev gift<sub>INF</sub> self<sub>DAT</sub> this book not reading  
 'Putin asked Medvedev to gift him this book without reading it'

- b. Putin poprosil Medvedeva podarit' Obame etu knigu ne čitaja \_\_\_\_.  
 Putin asked Medvedev gift<sub>INF</sub> Obama<sub>DAT</sub> this book not reading  
 'Putin asked Medvedev to gift Obama this book without reading it'

More research must be conducted in order to apply the definition of the phasal domain in (42) to constructions with parasitic gaps in Russian. One needs to propose a feature-based analysis of parasitic gaps, and examine examples without confounding factors. One could also consider other languages with monomorphemic anaphors which do not allow object drop. I leave these issues for future work.<sup>9</sup>

### 5.4.3 Scope

The next phenomenon which might be affected by the new definition of phase is the phenomenon of inverse scope. The following example in English is ambiguous and allows two readings.

(54) Some boy kissed every girl.

On one reading, there exists one particular boy who kissed every girl. This reading corresponds to the surface scope of (54), when the existential quantifier is interpreted higher than the universal quantifier. The second reading claims that for every girl there is a boy, not necessarily the same one for all girls, that kissed her. In this case the relative scope of the quantifiers is reversed: the universal quantifier scopes over the existential quantifier.

The narrow scope of the quantifier can be obtained when one assumes that the process of Quantifier Raising (QR) targets a  $\nu$ P-adjoined position, while wide scope is obtained when QR targets a higher position. One might expect that if the  $\nu$ P is not a phase, the narrow scope of the object will be impossible.

In order to explore whether this prediction holds, one needs to adopt a theory of scope based on the mechanism of feature-checking.

---

<sup>9</sup>Thanks to Marcel den Dikken (p.c.) for raising this issue.

Further, while the availability of certain readings of (54) might in fact indicate the possibility of movement to the  $\nu$ P-adjoined position, it has to be emphasized that scopal phenomena and QR are much more restricted than extraction. For instance, QR cannot be long-distance. It is more restrictive than binding, which at times can be long-distance. The sentence in (55) is unambiguous as the only possible reading is when the existential quantifier scopes over the existential. It can only mean that there is one particular boy that told that Bill kissed every girl.

(55) Some boy thinks that Bill kissed every girl.

These restrictions on QR cannot be derived directly in the current minimalist framework without extra assumptions. It is unclear why QR cannot proceed successive cyclically, resulting at the end in the inverse scope of the example in (55). Whether these locality restrictions on QR can be derived from the revised theory of phases remains to be seen. One might for instance propose that QR is triggered by a certain feature, say  $\langle \omega \rangle$ . Nothing in the current theory prevents placing an uninterpretable unvalued instance of this feature on  $\nu$  and C, thus resulting in long-distance QR, similar to long-distance *wh*-movement. This however would go against the facts and predict that the inverse scope reading of (55) is possible. To resolve this problem it might be possible to postulate that  $\langle \omega \rangle$  can only surface on  $\nu$ , and cannot be placed on C. This lexical restriction will be able to account for the facts, however it is just a stipulation. A feature-based theory of QR must be proposed in order to evaluate the consequences of the definition of phase in (42) on scope. It is also possible to revise one of the classical checking theories of scope (Beghelli and Stowell 1994, 1997). The interaction between features employed in these analyses and binding remain to be seen in future work.

#### **5.4.4 Future research**

Above I outlined certain predictions that can be made using the revised definition of phase (42). Each of these topics requires reanalysis of a corresponding syntactic phenomenon within the

feature-based system. Introducing relevant features responsible for these phenomena is a topic for future research.

The program I proposed in this dissertation by introducing feature  $\rho$  responsible for establishing binding relations can potentially be applied to all phenomena mentioned above: parasitic gaps, scope, *wh*-movement. One could also take this theory further by introducing a feature responsible for NPI-licensing, say  $\langle v \rangle$ . Using a similar approach to binding, one might analyze NPI as having an uninterpretable but valued instance of this feature, and the licenser of NPI as having an interpretable instance of this feature. As I demonstrated above in section 2.3.2 in a number of languages, such as Japanese, Korean, Serbo-Croatian, French, the availability of long-distance NPI licensing is limited to special types of embedded clauses, for instance subjunctives. In this way NPI licensing is similar to long-distance binding and therefore can be accounted for by using the definition of phase given in (42) and mechanisms similar to the mechanism employed for establishing binding relations.

A final direction for future work is the status of DPs and the nature of DP internal binding. In order to evaluate phasal status of DPs in the revised phase theory framework one needs to subscribe to certain assumptions about feature checking within DP. The debate about the DP structure is still ongoing, and I leave these issues for further research.

I leave the implementation of this program and the analysis of the interaction between the features involved in these various syntactic phenomena for future work.

# Bibliography

- Abels, Klaus. 2003. Successive cyclicity, anti-locality, and adposition stranding. Doctoral Dissertation, University of Connecticut.
- Abusch, Dorit. 1988. Sequence of tense, intensionality and scope. In *Proceedings of the 7th West Coast Conference on Formal Linguistics*, 1–14. Stanford: CSLI.
- Abusch, Dorit. 1994. Sequence of tense revisited: Two semantic accounts of tense in intensional contexts. In *Ellipsis, tense and questions*, ed. Hans Kamp. DYANA-2.
- Abusch, Dorit. 1997. Sequence of tense and temporal *de re*. *Linguistics and Philosophy* 20:1–50.
- Adger, David. 2004. *Core syntax: A minimalist approach*. Oxford University Press.
- Alexiadou, Artemis, Liliane Haegeman, and Melita Stavrou. 2007. *Noun phrase in the generative perspective*. Berlin, New York: Mouton de Gruyter.
- Anagnostopoulou, Elena, and Martin Everaert. 1999. Toward a more complete typology of anaphoric expressions. *Linguistic Inquiry* 30:97–119.
- Babby, Leonard H. 1987. Case, prequantifiers, and discontinuous agreement in Russian. *Natural Language and Linguistic Theory* 5:91–138.
- Bailyn, John. 2001. On scrambling: A reply to Bošković and Takahashi. *Linguistic Inquiry* 32:635–658.
- Bailyn, John. 2003. A (purely) derivational approach to Russian scrambling. In *Formal Approaches to Slavic Linguistics 11: The Amherst meeting 2002*, ed. Wayles Browne, Ji-Yung Kim, Barbara H. Partee, and Robert A. Rothstein, 41–62. Ann Arbor, MI: Michigan Slavic Publications.

- Bailyn, John. 2004. The case of Q. In *Formal Approaches to Slavic Linguistics 12: The Ottawa meeting 2003*, ed. Olga Arnaudova, Wayles Browne, Maria Luisa Rivero, and Danijela Stojanovic. Ann Arbor, MI: Michigan Slavic Publications.
- Bailyn, John. 2006. Against the scrambling anti-movement movement. In *Formal Approaches to Slavic Linguistics 14: The Princeton meeting 2005*, ed. James Lavine, Steven Franks, Mila Tasseva-Kurkchieva, and Hana Filip, 35–49. Ann Arbor, MI: Michigan Slavic Publications.
- Bailyn, John. 2010. What's inside VP? New (and old) evidence from Russian. In *Formal Approaches to Slavic Linguistics 18: The second Cornell meeting 2009*, ed. Wayles Browne, Adam Cooper, Alison Fisher, Esra Kesici, Nikola Predolac, and Draga Zec. Ann Arbor, MI: Michigan Slavic Publications.
- Bailyn, John. 2011. *The syntax of Russian*. Cambridge University Press.
- Baker, Mark. 1988. *Incorporation: a theory of grammatical function changing*. Chicago, IL: University of Chicago Press.
- Barss, Andrew, and Howard Lasnik. 1986. A note on anaphora and double objects. *Linguistic Inquiry* 17:347–354.
- Beghelli, Filippo, and Tim Stowell. 1994. The direction of quantifier movement. *GLOW Newsletter* 32:56–57.
- Beghelli, Filippo, and Tim Stowell. 1997. Distributivity and negation: The syntax of *Each* and *Every*. In *Ways of scope taking*, ed. Anna Szabolcsi, 71–107. Dordrecht: Kluwer.
- Belletti, Adriana, and Luigi Rizzi. 1988. Psych-verbs and theta-theory. *Natural Language and Linguistic Theory* 6:291–352.
- Bhatt, Rajesh. 2005. Long distance agreement in Hindi-Urdu. *Natural Language and Linguistic Theory* 23:757–807.
- Bittner, Maria. 1994. *Case, scope, and binding*. Dordrecht: Kluwer.
- Bouchard, Denis. 1992. Psych constructions and linking to conceptual structures. In *Romance languages and modern linguistic theory*, ed. Paul Hirschbühler and Koerner Konrad, 25–44. Amsterdam: John Benjamins.



- Bošković, Željko. 2004. Object shift and the clause/PP parallelism hypothesis. In *Proceedings of the 23rd West Coast Conference on Formal Linguistics*, ed. Vineeta Chand, Ann Kelleher, Angelo J. Rodríguez, and Benjamin Schmeiser, 101–114.
- Bošković, Željko. 2005. On the locality of the left branch extraction and the structure of NP. *Studia Linguistica* 59:1–45.
- Bošković, Željko. 2007. On the locality and motivation of Move and Agree: An even more minimal theory. *Linguistic Inquiry* 38:589–644.
- Bošković, Željko. 2010. Phases beyond clauses. Downloaded from <http://web2.uconn.edu/boskovic/papers/NPPPhase.Potsdam.pdf> on November 5, 2012.
- Bruening, Benjamin. 2006. What is the right binding theory? Handout from the colloquium talk given in Stony Brook on September 29, 2006.
- Carrier-Duncan, Jill. 1985. Linking of thematic roles in derivational word formation. *Linguistic Inquiry* 16:1–34.
- Choi, Dong-Ik. 1997. Binding principle for long-distance anaphors. *Kansas Working Papers in Linguistics* 22:57–71.
- Chomsky, Noam. 1973. Conditions on transformations. In *A festschrift for Morris Halle*, ed. Stephen R. Anderson and Paul Kiparsky. New York: Holt, Rinehart and Winston.
- Chomsky, Noam. 1980. *Rules and representations*. Oxford: Blackwell.
- Chomsky, Noam. 1981/1993. *Lectures on government and binding: The Pisa lectures*. Mouton de Gruyter.
- Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 89–155. MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. Michael J. Kenstowicz, 1–52. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2004. Beyond explanatory adequacy. In *Structures and beyond. The cartography*

- of syntactic structures*, ed. Adriana Belletti, volume 3, 104–131. Oxford: Oxford University Press.
- Chomsky, Noam. 2008. On phases. In *Foundational issues in linguistic theory. Essays in honor of Jean-Roger Vergnaud.*, ed. Robert Freidin, Carlos P. Otero, and Maria Luisa Zubizarreta, 134–166. Cambridge, MA: MIT Press.
- Clements, George N. 1975. The logophoric pronoun in Ewe: Its role in discourse. *Journal of West African Linguistics* 10:141–177.
- Cole, Peter, and Gabriella Hermon. 2005. The typology of Malay reflexives. *Lingua* 115:627–644.
- De Cuba, Carlos. 2007. On (non)factivity, clausal complementation and the CP-field. Doctoral Dissertation, Stony Brook University.
- Den Dikken, Marcel. 2006. A reappraisal of vP being phasal — a reply to Legate. Downloaded from [http://www.gc.cuny.edu/CUNY\\_GC/media/CUNY-Graduate-Center/PDF/Programs/Linguistics/Dikken/legate\\_reply.pdf](http://www.gc.cuny.edu/CUNY_GC/media/CUNY-Graduate-Center/PDF/Programs/Linguistics/Dikken/legate_reply.pdf) on November 5, 2012.
- Den Dikken, Marcel. 2007. Phase extension. Contours of a theory of the role of head movement in phrasal extraction. *Theoretical Linguistics* 33:1–41.
- Despić, Miloje. 2011. Syntax in the absence of determiner phrase. Doctoral Dissertation, University of Connecticut.
- Dyakonova, Marina. 2009. *A phase-based approach to Russian free word order*. Utrecht: LOT Publishers.
- Enç, Mürvet. 1987. Anchoring conditions for tense. *Linguistic Inquiry* 18:633–657.
- Giorgi, Alessandra. 1983. Toward a theory of long distance anaphors: A GB approach. *The Linguistic Review* 3:307–361.
- Giorgi, Alessandra. 2006. From temporal anchoring to long distance anaphors. *Natural Language and Linguistic Theory* 24:1009–1047.
- Giorgi, Alessandra, and Fabio Pianesi. 1997. *Tense and aspect: From semantics to morphosyntax*. Oxford University Press.
- Harada, Naomi, and Richard Larson. 2009. Datives in Japanese. In *Proceedings of the 5th work-*

- shop on altaic formal linguistics (waf5). mit working papers in linguistics*, ed. R. Shibagaki and R. Vermeulen, volume 54, 3–17. Cambridge, MA: MITWPL.
- Harbert, Wayne. 1995. Binding theory, control, and *pro*. In *Government and binding theory and the minimalist program*, ed. Gert Webelhuth, chapter 4, 177–240. Cambridge, MA: Blackwell.
- Harris, Alice C. 1981. *Georgian syntax: A study in relational grammar*. Cambridge, MA: Cambridge University Press.
- Hasegawa, Hiroshi. 2005. Reflexive binding as agreement and its interactions with the phase system. In *The world of linguistic research: A festschrift for Kinsuke Hasegawa on the occasion of his seventieth Birthday*, ed. Noriko Imanishi, 53–69. Tokyo: Kaitakusha.
- Haspelmath, Martin. 1999. Long distance agreement in Godoberi (Daghestanian) complement clauses. *Folia Linguistica* 33:131–151.
- Heck, Fabian, Gereon Müller, and Jochen Trommer. 2008. A phase-based approach to Scandinavian definiteness marking. In *Proceedings of the 26th West Coast Conference on Formal Linguistics*, ed. Charles B. Chang and Hannah J. Haynie, 226–233. Somerville, MA: Cascadilla Proceedings Projects.
- Hellan, Lars. 1991. Containment and connectedness anaphors. In *Long-distance anaphora*, ed. Jan Koster and Eric Reuland, 27–48. Cambridge, MA: Cambridge University Press.
- Hestvik, Arild, and William Philip. 2001. Syntactic vs. logophoric binding: Evidence from norwegian child language. In *Long-distance reflexives*, ed. Peter Cole, Gabriella Hermon, and C.-T. James Huang, volume 33 of *Syntax and Semantics*, 119–139. San Diego, CA: Academic Press.
- Hicks, Glyn. 2009. *The derivation of anaphoric relations*. John Benjamins.
- Hoeksema, Jack, and Donna Jo Napoli. 1990. A condition on circular chains: A restatement of *i*-within-*i*. *Journal of Linguistics* 26:403–424.
- Hornstein, Norbert. 2001. *Move! A minimalist theory of construal*. Malden, MA: Blackwell.
- Huang, C.-T. James, and C.-S. Luther Liu. 2001. Logophoricity, attitudes and *ziji* at the interface. In *Syntax and Semantics 33: Long-distance reflexives*, ed. Peter Cole, Gabriella Hermon, and C.-T. James Huang, 141–196. Academic Press.

- Huang, C.-T. James, and C.-C. J. Tang. 1991. The local nature of the long-distance reflexive in Chinese. In *Long-distance anaphora*, ed. Jan Koster and Eric Reuland, 263–282. Cambridge, MA: Cambridge University Press.
- Huang, Yan. 2000. *Anaphora: A cross-linguistic approach*. Oxford University Press.
- Ivlieva, Natalia. 2007. Parasitic gaps in Russian. In *Formal Approaches to Slavic Linguistics 15: The Toronto meeting*, ed. Richard Compton, Magdalena Goledzinowska, and Ulyana Savchenko, 132–140. Ann Arbor, MI: Michigan Slavic Publications.
- Johnson, Kyle. 2007. In search of phases. In *Phrasal and clausal architecture*, ed. Simin Karimi, Vida Samiian, and Wendy K. Wilkins, 146–166. John Benjamins.
- Kato, Kumiko. 1994. On reflexives in Japanese: Some syntactic and pragmatic approaches. Master's thesis, University of Reading.
- Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
- Kayne, Richard. 2002. Pronouns and their antecedents. In *Derivation and explanation in the minimalist program*, ed. Samuel D. Epstein and T. Daniel Seely, 133–166. Malden, MA: Blackwell.
- Khalilova, Zaira. 2009. *A grammar of Khwarshi*, volume 234 of *LOT Dissertation Series*. Utrecht: LOT.
- Kitada, Shin-ichi. 2007. Feature-inheritance in prepositional phrases. Downloaded from [http://www.sal.tohoku.ac.jp/eng/pdf/200705\\_kitada.pdf](http://www.sal.tohoku.ac.jp/eng/pdf/200705_kitada.pdf) on November 28, 2010.
- Kondrashova, Natalia. 1999. Barenaked tenses and their morphological outfits. In *Proceedings of NELS 29*, volume 1, 183–195.
- Krapova, Iliyana. 2001. Subjunctives in Bulgarian and Modern Greek. In *Comparative syntax of Balkan languages*, ed. Maria Luisa Rivero and Angela Ralli, 105–126. Oxford University Press.
- Kratzer, Angelika. 1996. Severing the external argument from its verb. In *Phrase structure and the lexicon*, ed. Johan Rooryck and Laurie Zaring, 109–137. Kluwer.
- Landau, Idan. 2010. *The locative syntax of experiencer*. Cambridge, MA: MIT Press.
- Larson, Richard. 1988. On the double object construction. *Linguistic Inquiry* 19:335–391.
- Legate, Julie Anne. 2003. Some interface properties of the phase. *Linguistic Inquiry* 34:506–516.

- Lidz, Jeffrey. 2001. Condition R. *Linguistic Inquiry* 32:123–140.
- Lin, Jo-Wang. 2005. Time is a language without tense: The case of Chinese. *Journal of Semantics* 23:1–53.
- Manzini, M. Rita, and Kenneth Wexler. 1987. Parameters, binding theory and learnability. *Linguistic Inquiry* 18:413–444.
- Mathieu, Eric. 2001. On the nature of French N-words. *UCL Working papers in Linguistics* 13:319–353.
- Matsubara, Fuminori. 2000.  $p^*P$  phases. *Linguistic Analysis* 30:127–161.
- McCloskey, James. 1990. *Transformational syntax and model theoretic semantics: A case study in Modern Irish*. Dordrecht: D. Reidel.
- Miyagawa, Shigeru. 2005. EPP and semantically vacuous scrambling. In *The free word order phenomenon: Its syntactic sources and diversity*, ed. Joachim Sabel and Mamoru Saito, 181–220. Mouton de Gruyter.
- Nakao, Chizuru, and Miki Obata. 2007. Parametric variations in NPI-licensing and the role of LF  $X^0$ -movement. In *Proceedings of the 9th Seoul International Conference on Generative Grammar 2007. Locality and Minimalism*, ed. Doo-Won Lee, 135–152. Seoul, Korea: Hankuk Publishing Co.
- Ngoh, Geok Lun. 1991. A pragmatic analysis of long-distance reflexivization in Malay. University of Cambridge.
- Nissenbaum, Jon. 1998. Derived predicates and the interpretation of parasitic gaps. In *Proceedings of the 17th West Coast Conference on Formal Linguistics*, ed. Kimary Shahin, Susan Blake, and Eun-Sook Kim, 507–521. Cascadilla Proceedings Projects.
- Ogihara, Toshiyuki. 1996. *Tense, attitudes, and scope*. Dordrecht: Kluwer Academic Publishers.
- Pesetsky, David. 1995. *Zero syntax: Experiencers and cascades*. Cambridge, MA: MIT Press.
- Pesetsky, David, and Esther Torrego. 2001. T-to-C movement: Causes and consequences. In *Ken Hale: a life in language*, ed. Michael J. Kenstowicz, 355–426. Cambridge, MA: MIT Press.
- Pesetsky, David, and Esther Torrego. 2007. The syntax of valuation and the interpretability of

- features. In *Phrasal and clausal architecture*, ed. Simin Karimi, Vida Samiian, and Wendy K. Wilkins, 262–294. John Benjamins.
- Pica, Pierre. 1986. Liage et contiguité. In *Recherches sur l'anaphore*, ed. J.-C. Milner, 119–164. Paris: Université de Paris-7.
- Pica, Pierre. 1987. On the nature of the reflexivization cycle. In *Proceedings of NELS 17*, ed. J. McDonough and B. Plunkett, 483–499. UMass Amherst.
- Pica, Pierre. 1991. On the interaction between antecedent-government and binding: The case of long-distance reflexivization. In *Long-distance anaphora*, ed. Jan Koster and Eric Reuland, 119–35. Cambridge, MA: Cambridge University Press.
- Picallo, M. Carme. 1985. Opaque domains. Doctoral Dissertation, CUNY Graduate Center.
- Polinsky, Maria. 2003. Non-canonical agreement is canonical. *Transactions of the Philological Society* 101:279–312.
- Polinsky, Maria, and Eric Potsdam. 2001. Long-distance agreement and topic in Tsez. *Natural Language and Linguistic Theory* 19:583–646.
- Progovac, Ljiljana. 1991. Polarity in Serbo-Croatian: Anaphoric NPIs and pronominal PPIs. *Linguistic Inquiry* 22:567–572.
- Progovac, Ljiljana. 1993a. Locality and subjunctivity-like complements in Serbo-Croatian. *Journal of Slavic Linguistics* 1:116–144.
- Progovac, Ljiljana. 1993b. Relativized SUBJECT: long-distance reflexives without movement. *Linguistic Inquiry* 23:671–680.
- Progovac, Ljiljana. 1993c. Subjunctive: The ‘misbehaviour’ of anaphora and negative polarity. *Linguistic Review* 10:37–59.
- Reinhart, Tanya, and Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24:657–720.
- Reuland, Eric. 2001. Primitives of binding. *Linguistic Inquiry* 32:439–492.
- Reuland, Eric. 2005. Agreeing to bind. In *Organizing grammar: Studies in honor of Henk van Riemsdijk*, ed. Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinhenz, and Jan Koster, 505–513. Mouton de Gruyter.

- Reuland, Eric. 2011. *Anaphora and language design*. MIT Press.
- Rezac, Milan. 2004. Elements of cyclic syntax: Agree and Merge. Doctoral Dissertation, University of Toronto.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar*, ed. Liliane Haegeman, 281–337. Kluwer Academic Publishers.
- Rudnitskaya, Elena. 2000. Long-distance binding in the feature movement framework with reference to Russian. Doctoral Dissertation, CUNY Graduate Center.
- Saito, Mamoru. 1992. Long distance scrambling in Japanese. *Journal of East Asian Linguistics* 1:69–118.
- Saito, Mamoru. 2003. A derivational approach to the interpretation of scrambling chains. *Lingua* 113:481–518.
- Saito, Mamoru. 2005. Further notes on the interpretation of scrambling chains. In *The free word order phenomenon: Its syntactic sources and diversity*, ed. Joachim Sabel and Mamoru Saito, 335–376. Walter de Gruyter.
- Sells, Peter. 1987. Aspects of logophoricity. *Linguistic Inquiry* 18:445–479.
- Shima, Etsuro. 2000. A preference for move over merge. *Linguistic Inquiry* 31:375–385.
- Sigurðsson, Halldór A. 1990. Long-distance reflexives and moods in Icelandic. In *Syntax and Semantics 24: Modern Icelandic syntax*, ed. Joan Maling and Annie Zaenen, 309–346. London: Academic Press.
- Sung, Li-May. 1990. Universals of reflexives. Doctoral Dissertation, University of Illinois at Urbana-Champaign.
- Svenonius, Peter. 2004. On the edge. In *Peripheries: syntactic edges and their effects*, ed. David Adger, Cecile de Cat, and George Tsoulas, 261–287. Dordrecht: Kluwer.
- Svenonius, Peter. 2010. Spatial P in English. In *Mapping spatial PPs*, ed. Guglielmo Cinque and Luigi Rizzi, volume 6 of *The Cartography of Syntactic Structures*. Oxford University Press.
- Thráinsson, Höskuldur. 1976. Reflexives and subjunctives in Icelandic. In *The proceedings of NELS 6*, ed. Alan Ford, John Reighard, and Rajendra Singh, 225–239. UMass Amherst.

- Tsoulas, George. 1995. The nature of the subjunctive and the formal grammar of obviation. In *Grammatical theory and Romance languages*, ed. Karen Zagona, 293–306. Amsterdam: John Benjamins.
- Uchibori, Asako. 2000. The syntax of subjunctive complements: Evidence from Japanese. Doctoral Dissertation, University of Connecticut.
- Wali, Kashi. 1989. *Marathi syntax: A study of reflexives*. Patiala: Institute of Language Studies.
- Wali, Kashi, and K. V. Subbarao. 1991. On pronominal classification: Evidence from Marathi and Telugu. *Linguistics* 29:1093–1110.
- Yamashita, Hideaki. 2003. On the distribution and licensing of negative polarity items in Japanese and the Phase-Impenetrability Condition. In *Proceedings of the Fourth Tokyo conference of Psycholinguistics*, 313–338. Hituzi Publishing Company.
- Zwart, Jan-Wouter. 2002. Issues relating to a derivational theory of binding. In *Derivation and explanation in the minimalist program*, ed. Samuel D. Epstein and T. Daniel Seely, 269–304. Malden, MA: Blackwell.