

## Principle A and feature valuation

Andrei Antonenko \*

**Abstract.** Traditional binding theory is largely incompatible with minimalist assumptions. In this paper 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. I further demonstrate how this revision of binding theory can derive subject orientation of monomorphemic anaphors, Barss-Lasnik effects, and restrict at which moment of derivation binding theory can apply. In conclusion I show some novel asymmetries observed in *wh*-dislocated reflexives in English vs. Russian indirect questions.

**Keywords.** binding theory; minimalist program; reflexives; principle A

**1. Introduction.** Capturing anaphor distribution within the Minimalist Program faces intriguing challenges. Classical binding relates phrases via indices; agreement in MP relates heads via features. Binding is restricted to local domains; agreement is unbounded. Anaphor binding exhibits “orientation” to different argument positions; agreement has no equivalent concept. For these reasons, all recent approaches to binding in feature-based frameworks (Hicks 2009; Reuland 2011 a.o.) have invoked assumptions non-standard for the MP.

In this paper I demonstrate how binding theory Principle A can be reduced to Agree relation, and show how this approach accounts for the following properties: 1). c-command and locality restriction on binding; 2). Subject orientation of monomorphemic anaphors; 3). Binding in double-object constructions and Barss-Lasnik effects. Finally, I show an asymmetry between binding of *wh*-moved reflexives between Russian and English and show how the proposed theory derives it.

**2. Featural approach to binding.** In this section I adopt the approach to feature-checking outlined in Pesetsky and Torrego 2007, which we 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 (1).

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

---

\*I would like to thank members of The Stony Brook University Linguistics Department and LSA 2018 audience for valuable comments and suggestions. All errors remain my own. Andrei Antonenko, Stony Brook University ([andrei.antonenko@stonybrook.edu](mailto:andrei.antonenko@stonybrook.edu))

- 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 example, 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, but valued, since verb comes from the lexicon with morphologically specified tense.

2.1. THE  $\rho$ -FEATURE. Reinhart and Reuland 1993 argue that Principle A of the binding theory is in fact a condition on 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 reformulation of Principle A of the binding theory:

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

Here, I build on Reinhart and Reuland's 1993 intuition about reflexive predicates, but reduce it to a formal syntactic process by introducing the  $\rho$ -feature. Formally, the presence of an interpretable  $\rho$ -feature on a head is needed for coreference between an element contained in the domain of the head (the anaphor) and an expression in the specifier position of the head (the antecedent). For instance, the presence of this interpretable feature on the  $v/V$ -complex will require the complement of the transitive verb to be coreferential with its specifier, as in (4-a).

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

Further, 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).

The interpretable  $\rho$ -feature occurs on a head and must c-command a  $\rho$ -feature in its domain, in the familiar fashion. Formally, in the Pesetsky and Torrego 2007 framework, this feature on the head must be interpretable, as it works at the interface by establishing binding relations. 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.<sup>1</sup>

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:             | <i>interpretable</i> | $\langle i\rho \pm val \rangle$ |
| on reflexives:    | <i>valued</i>        | $\langle u\rho +val \rangle$    |
| on R-expressions: | <i>absent</i>        |                                 |

Now, the question is what the role of the interpretable  $\rho$ -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.

Consider an example (6) in which the  $\rho$ -feature is located on a V-head, and the corresponding valued  $\rho$ -feature is located on the direct object, (7).

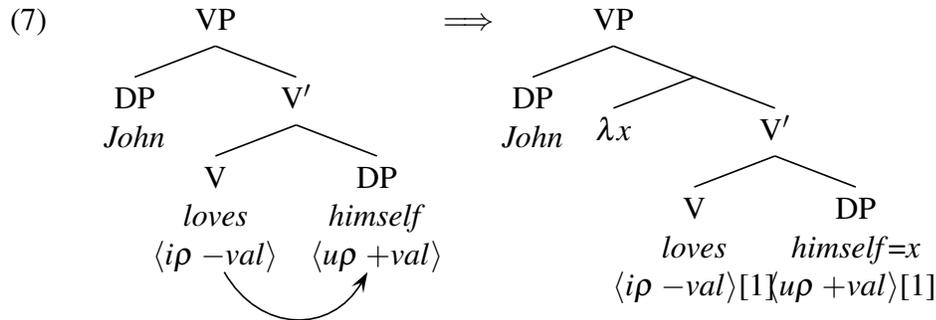
- (6) John loves himself.

Valuation of the  $\rho$ -feature on V occurs after V probed a reflexive, and after 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

---

<sup>1</sup>Unless the verb is inherently reflexive, or has reflexivity marked on it in a morphological way, we 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.

the Spec,VP will be substituted for a variable.



Principle A is now a standard requirement that the  $\rho$ -feature must be valued, and there must be an interpretable instance of the  $\rho$ -feature.<sup>2</sup>

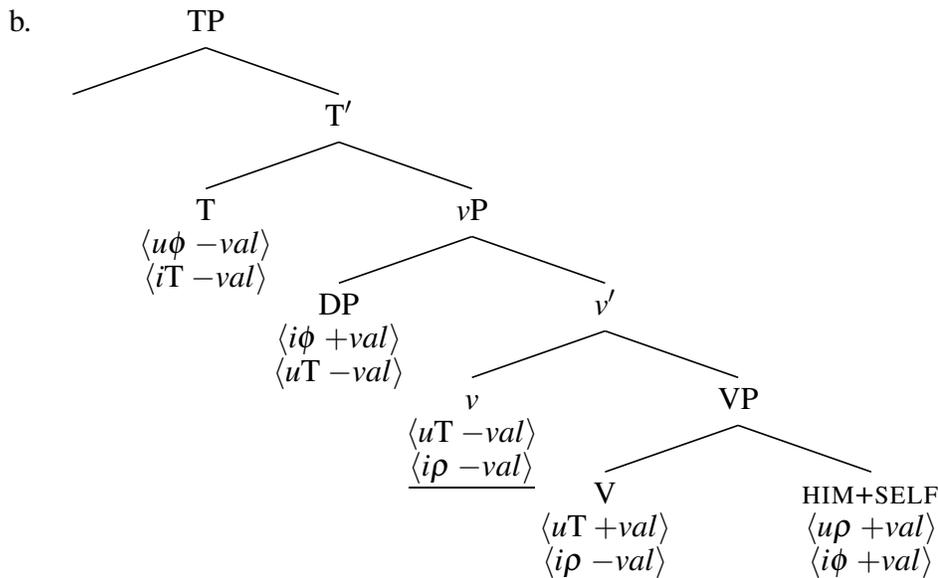
2.2. BASIC CASES OF BINDING IN ENGLISH. This section deals with establishing binding relations between a complex anaphor, of the type observed in English (*himself*), and its antecedent. The main difference between complex and monomorphemic anaphors lies in their featural content. Monomorphemic reflexives 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 in the  $\nu$ P-internal 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>

In a language like English, 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  $\nu$ ; they are underlined in (9): only one of them will actually be present when the derivation happens.

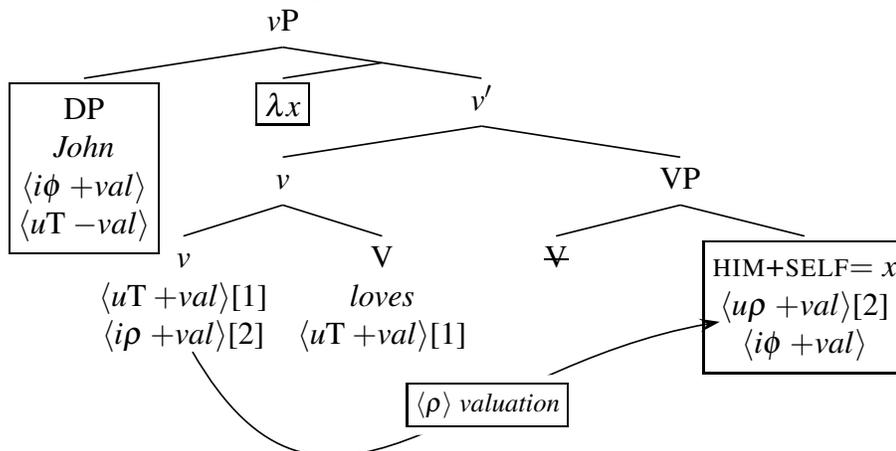
- (9) *Basic structure and feature content for a transitive sentence with a complex reflexive.*  
 a. John loves himself.

<sup>2</sup>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.



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).

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



Notice, that after  $v$  probes V, all features within it are valued, or shared with some element outside VP. Therefore, after T probes subject DP and  $v+V$  complex,  $v$ P can be shipped to the inter-

pretative component, and  $\lambda$ -conversion will result in reflexive reading, where both  $\text{Agent}(e)$  and  $\text{Theme}(e)$  of the event described by the verb are the same:

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

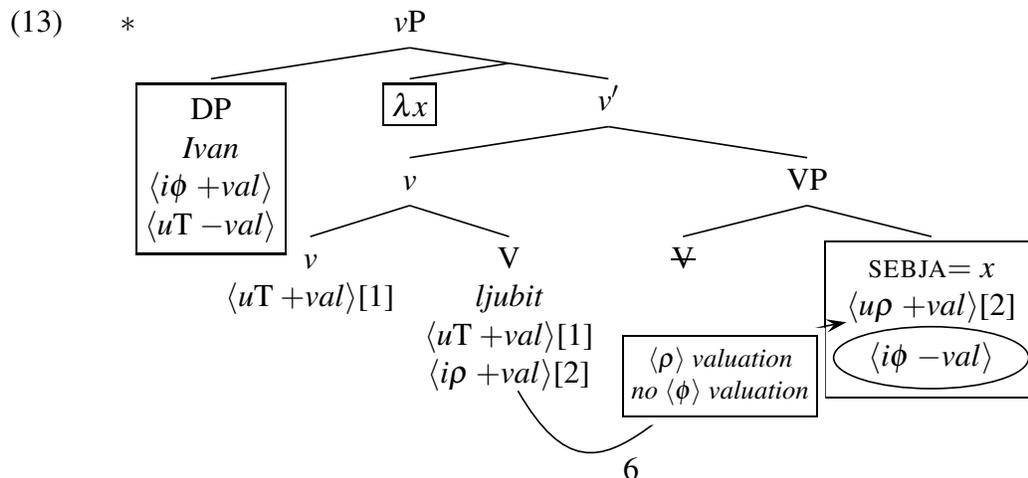
This derivation will only minimally differ if one chooses to place the  $\rho$ -feature on V.  $\lambda$ -conversion will lead to the same result, where the only possible antecedent of the reflexive is the subject DP.

**3. Interaction of  $\rho$  and  $\phi$ : Subject orientation.** In this section, I 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 the presence of unvalued  $\phi$ -features derives subject orientation of this reflexive. A basic transitive sentence in Russian is given in (12).

- (12) Ivan<sub>i</sub> ljubit sebja<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 interpretable. Formally that means that such reflexives 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, but valued,  $\langle u\rho +val \rangle$ .

As before, the unvalued instance of the  $\rho$ -feature can be placed on either V,  $v$ , or T. For the sake of simplicity, in the further derivation I assume as before that it is placed on  $v$ . Now, the derivation will proceed along the same lines as the derivation in the previous section for English. The  $vP$  structure after the merge of the subject DP is shown in (13).



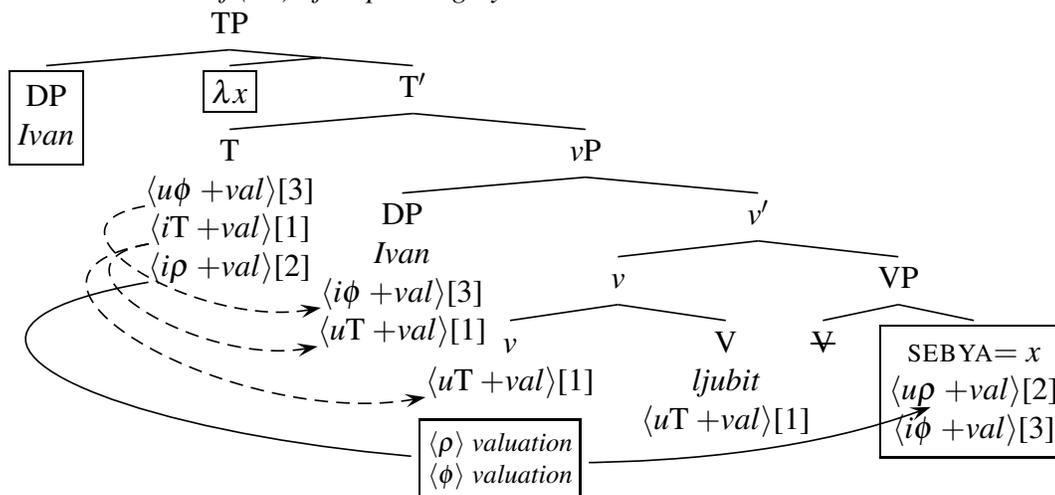
It is similar to the structure in (10), with the exception of the reflexive, which now has unvalued  $\phi$ -features.

After that, 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 will probe first a subject DP in Spec,vP position to value its  $\phi$ -features, and then  $v+V$  complex, to value its T-features. Importantly, T will not probe SE, and as a result  $\phi$ -feature on the reflexive will be left unvalued as no higher probe will be able to supply a value to it. 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 (14). The subject DP is further raised into Spec,TP.

At this moment, all features within the TP are valued (including the  $\phi$ -features on the reflexive) and the derivation will not crash; 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 the element in Spec, TP, i.e. 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 (15).

(14) Final structure of (12) after probing by T.



- (15) 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.

To summarize, our theory predicts that in a language with monomorphemic anaphors,  $\rho$ -feature must be placed on T, and not on  $v$  or V, and that results in binding being possible only from Spec,TP position. We will use this fact below.

#### 4. Interaction of $\rho$ and Case: Barss-Lasnik effects.

4.1. BARSS-LASNIK EFFECTS IN ENGLISH AND DERIVATIONAL THEORY OF BINDING. In this section we 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 (16)-(17).

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

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

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

The first argument  $c$ -commands the second, and therefore the examples above differ. Further note that example (18-a) is ambiguous: the reflexive *himself* can be bound by either subject or direct object, giving two possible readings. On the other hand, example (18-b) is not ambiguous: the only possible antecedent of *himself* is *John*.<sup>3</sup>

The featural binding theory proposed above can account for the two possible readings of (18-a). If the  $\rho$ -feature is placed on  $v$ , the derivation differs minimally from the case of transitive

<sup>3</sup>Note that there are apparent violations of a c-command condition on binding in examples like (i):

- (i) John talked to  $\text{Mary}_i$  about  $\text{herself}_i$ .

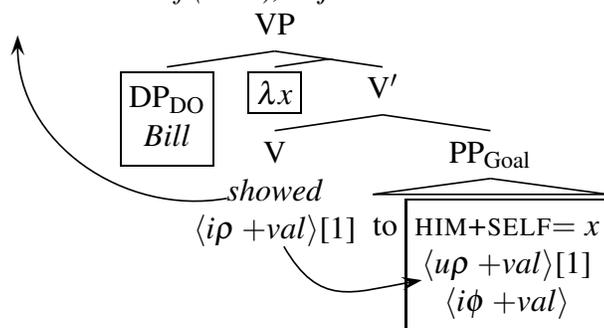
The approach to binding proposed in this chapter is able to account for this phenomenon by assuming that *to* is semantically pleonastic and the phrase *to Mary* is interpreted simply as *Mary*. The corresponding logical form of (i) before  $\lambda$ -conversion is given in (ii):

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

sentences, resulting in binding by the subject.

On the contrary, when the  $\rho$ -feature is placed on V, the direct object will serve as a binder for the reflexive. The structure below shows how binding by the direct object is derived if the  $\rho$ -feature is placed on V. In this case, 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.

(19) VP structure of (18-a), object antecedent.



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

To summarize, 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; 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. John<sub>i</sub> showed himself<sub>i</sub> to Mary  
 b. John<sub>i</sub> showed himself<sub>i/\*j</sub> to Bill<sub>j</sub>

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 higher than the theme, such as the sentences in (22).

- (22) I showed Mary<sub>IO</sub> the world<sub>DO</sub>.

We assume the structure of double object construction as described in Larson 1988 based on the thematic hierarchy (23) 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 (24).

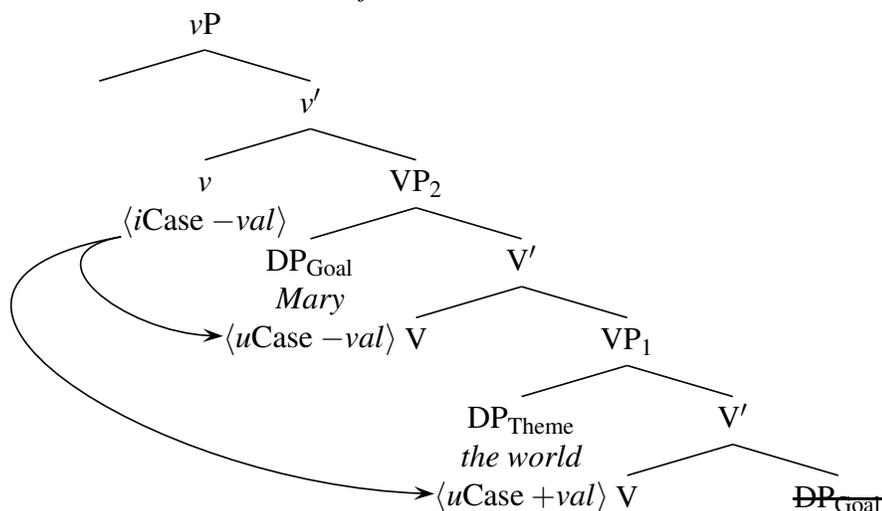
(23) *Thematic hierarchy*  
 AGENT > THEME > GOAL > OBLIQUES (manner, location, time, ...)

(24) *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.

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 (25).

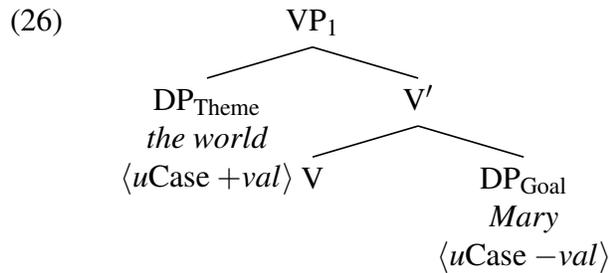
(25) *Case valuation in double object constructions.*



This derivation leads Harada and Larson 2009 to assume that case assignment on the goal DP is “concordial”, i.e. that 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 binding proposed above. Notice, that Harada and Larson 2009 do not have a separate projection where the goal argument moves – they assume that it moves to the extra specifier position in VP. In the tree (25), we 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 we will assume that the extra VP projection exists.

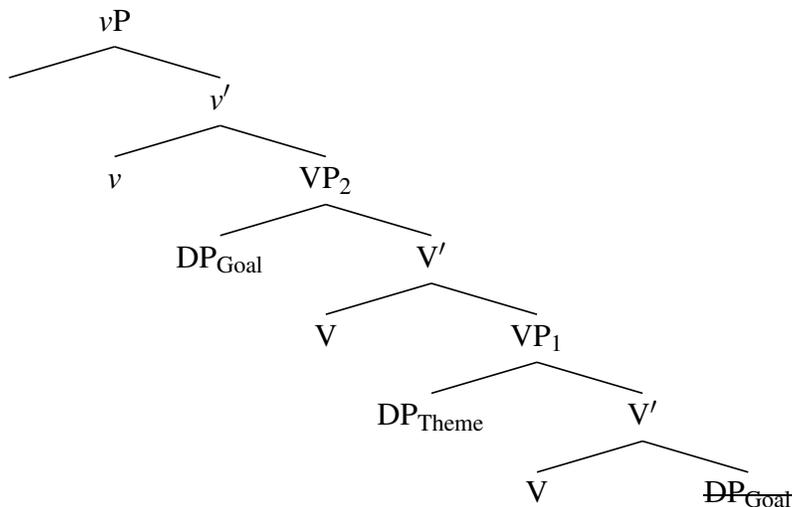
Consider first the lower VP<sub>1</sub>. Its structure at the moment of completion is given in (26):



Notice that VP<sub>1</sub> contains an unvalued feature, i.e. the Case-feature on the indirect object *Mary*. As a result, VP<sub>1</sub> cannot be shipped to semantics, and therefore cannot serve as a binding domain.

Now consider VP<sub>2</sub>. After the Goal DP moves to its edge, it is probed by *v* in order to have its Case-feature valued, as shown in (25). 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<sub>2</sub>. Therefore, VP<sub>2</sub> is the first moment in the derivation of double object construction where binding theory as defined above can be applied.

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



One of the major problem for the derivational theory of binding (see e.g. Belletti and Rizzi

1988) comes from consideration of the sentences like the ones given in (28) and (29):<sup>4</sup>

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

Consider how the derivation of the sentence in (29-b) starts. Assuming the hypothesis about low generation of goals according to the thematic hierarchy in (23) (Carrier-Duncan 1985; Baker 1988; 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, (30), before further movement of the goal happens.

- (30) \*John showed [clones of themselves]<sub>i</sub> [the professors]<sub>i</sub> t.  


Further dislocations of the goal should not make the sentence ungrammatical. However, it is not the case: sentence (29-b) is ungrammatical. In what follows, we show that the binding theory as defined above makes correct predictions about the status of sentences in (29).

Let us consider the derivation of the sentence in (29-b) in the current framework. Consider a case when the  $\rho$ -feature is present on V. At the beginning, V is merged with the goal 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<sub>1</sub> position. As I mentioned above, binding theory cannot apply at this moment and coreference between the goal DP by the theme DP cannot yet be established. 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, we propose the following principle:

- (31) **Application of binding theory:** Binding relations (i.e.  $\lambda$ -conversion) are applied only at the moment the category becomes can be send to semantics, i.e. it has no unvalued or unshared features, and not at any moment in the derivation.

The motivation for the principle in (31) is the following. While feature-valuation happens

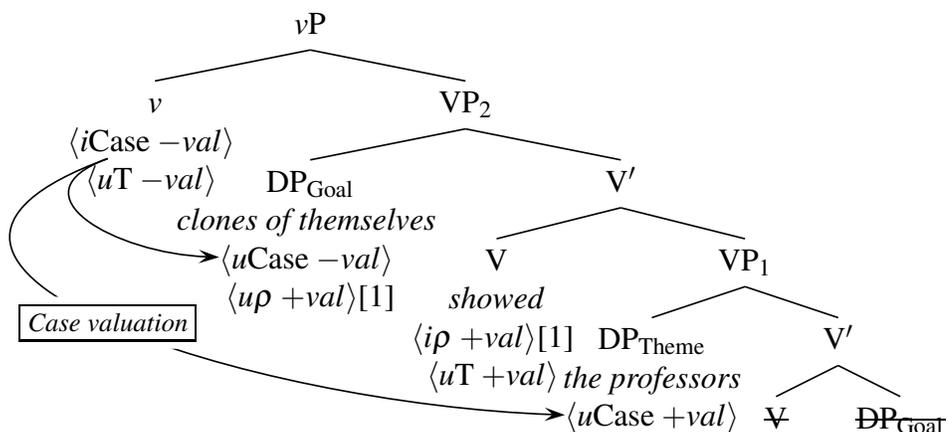
---

<sup>4</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 we will concern ourselves with the possibility of a direct object being bound by a goal DP.

at the syntactic level grammar,  $\lambda$ -conversion (which is responsible for establishing the binding relations between the anaphor and its antecedent) applies at the interface with semantics. For it to be necessary, all features within the domain must be interpretable and valued.

Consider the consequences of this principle to double-object constructions. According to the view on double-object constructions presented above, the goal DP must move to the edge of the higher  $VP_2$  – 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 (32). Notice, that at this time,  $\rho$ -feature on the reflexive and on  $V$  are already shared.

(32) The structure of the  $vP$  of (29-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  $VP_3$ . The resulting logical form is shown below:

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

After the subject is merged into Spec, $vP$ ,  $\lambda$ -conversion will apply, and therefore the only potential binder of the goal reflexive is the subject; coreference between *themselves* and *professors* is impossible, and therefore the sentence is ungrammatical: subject binding is impossible since  $\phi$ -feature of the subject and of the reflexives do not match in (29-b).

4.2. DITRANSITIVE CONSTRUCTIONS IN A LANGUAGE WITH MONOMORPHEMIC ANAPHORS AND SUBJECT ORIENTATION. As we 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 (34) below.

(34) *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’

In the Russian examples, SE-anaphor *sebe/sebja* can only be bound by the subject DP, *Ivan* in (34-b) and *Marina* in (34-a) and in (34-c).

Recall that in Russian  $\rho$ -feature must be placed on a  $\phi$ -bearing element, i.e. T, otherwise  $\phi$ -features of the reflexive will not get valued. Hence, binder can only be located in the Spec,TP position, and therefore only subject binding is possible in Russian. The potential binding possibilities for examples in (34) are therefore explained.

Notice that in Russian, dative subjects can also serve as binders:

- (35) Maše<sub>i</sub> nravitsja svoja<sub>i</sub> kvartira.  
 M.<sub>DAT</sub> pleases self’s apartment  
 ‘Maša likes her apartment’

If we assume that dative experiencers are located in Spec,TP (see a.o. Bailyn 2004), such binding possibilities are predicted by the theory.

**5. Interaction of binding and *wh*-movement.** Now consider the following contrast between English and Russian.

- (36) a. John<sub>i</sub> wonders which pictures of himself<sub>i/j</sub> Bill<sub>j</sub> published t.  
 b. Ivan<sub>i</sub> interesuetsja [kakie rasskazy o sebe<sub>\*i/j</sub>] Boris<sub>j</sub> pročital t.  
 Ivan is interested [which stories about SELF] Boris read<sub>PAST</sub>.  
 ‘Ivan<sub>i</sub> is interested with stories about himself<sub>\*i/j</sub> Boris<sub>j</sub> has read.’

In English, when the reflexive is *wh*-moved into the intermediate Spec,CP position, it can be bound by the matrix subject. On the other hand, in Russian, moving reflexive to the intermediate Spec,CP does not lead to any new binding possibilities, and the only possible antecedent in the embedded subject.

Theory outlined in this paper provides an explanation why this is the case. Recall that in

English,  $\rho$ -feature can be placed on either  $\nu$  or  $V$ . As a result, it can probe the uninterpretable instance of the  $\rho$ -feature on the reflexive in Spec,CP, since the embedded CP is not yet a closed phase. Therefore, binding relations will be established between a matrix subject (or object) and the anaphor.

In Russian, in order to achieve binding by the matrix subject, the  $\rho$ -feature must be placed on the matrix T, since placing it on  $\nu$  or  $V$  is impossible in a language with monomorphemic reflexives. However T will not be able to probe into embedded Spec,CP, since by the time T is merged into structure (after completion of the matrix  $\nu$ P), the embedded CP is closed for probing as a phase. As a result, new binding possibilities cannot arise.

## References

- Bailyn, John. 2004. Generalized inversion. *Natural Language and Linguistic Theory* 22:1–49.
- 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.
- Belletti, Adriana, and Luigi Rizzi. 1988. Psych-verbs and theta-theory. *Natural Language and Linguistic Theory* 6:291–352.
- Carrier-Duncan, Jill. 1985. Linking of Thematic Roles in Derivational Word Formation. *Linguistic Inquiry* 16:1–34.
- Harada, Naomi, and Richard Larson. 2009. Datives in Japanese. In *Proceedings of the 5th Workshop on Altaic Formal Linguistics (WAFL5)*. MIT Working Papers in Linguistics, ed. R. Shibagaki & R. Vermeulen, volume 54, 3–17. Cambridge, MA: MITWPL.
- Hicks, Glyn. 2009. *The derivation of anaphoric relations*. John Benjamins.
- 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.
- Larson, Richard. 1988. On the double object construction. *Linguistic Inquiry* 19:335–391.
- 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.
- Reinhart, Tanya, and Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24:657–720.
- Reuland, Eric. 2011. *Anaphora and language design*. MIT Press.