Chapter 4:[1] Binding in a Local Derivational Approach

Silke Fischer

1 Introduction

In the previous chapter, a derivational analysis of reconstruction has been proposed, which means that optimization takes place in the course of the derivation when the syntactic structure has not yet been completely built up. In contrast, the analysis of binding data in chapter 2 was based on different premises. There it has been assumed that complete sentences are part of the input and thus completely accessible during the optimization procedure. Hence, it can be characterized as a global, representational analysis. However, this kind of approach is not straightforwardly applicable to examples involving reconstruction (cf. chapter 2, section **??**), since the grammaticality status of this type of sentences crucially depends on intermediate derivation steps which might no longer be recoverable once the derivation has been completed.

In this chapter, I will therefore address the question of whether the binding theory outlined in chapter 2 can be integrated into a local derivational syntactic approach. I explore what must be assumed for binding once we restrict ourselves to a derivational framework and discuss the theoretical consequences of such an enterprise.[2] Moreover, I set out to propose a theory

[1] Cf. Fischer, Silke. 2004. *Towards an Optimal Theory of Reflexivization.* Doctoral dissertation, University of Tübingen.

[2] Former derivational approaches to binding include Hornstein (2001), Kayne (2002), and Zwart (2002), which share the underlying assumption that an antecedent and its bindee start out as one constituent and the binding relation is created by movement (cf. chapter 1). In contrast to these proposals, the present approach focuses on crosslinguistic variation and optionality and neither assumes movement into θ-positions (cf. Hornstein (2001)) nor a single phrase containing both the bound element and its antecedent at some stage of the derivation. Moreover, the domain accessible in the course of the derivation will be reduced to a minimum.

that is empirically not inferior to the approach developed in chapter 2 but captures the same amount of data as well as some universal generalizations that can be observed with respect to binding.

This chapter is organized as follows. In section 2, I take a closer look at derivational theories in general and explore their theoretical implications. In section 3, I discuss how much the accessible domain can be restricted in a derivational approach to binding. The conclusion that will be drawn is that the theory with the most restrictive notion of accessibility does not raise more problems than a more liberal theory and is therefore to be preferred from a conceptual point of view. In section 4, I address some technical issues of the new analysis according to which binding corresponds to feature checking between the bound element and its antecedent. Finally, section 5 constitutes the main part of the chapter, because here I develop an optimality-theoretic approach to binding in a derivational framework and show that it captures the same data as the theory presented in chapter 2.

2 Theoretical Considerations on Derivational Theories

It has been argued in the literature that derivational theories are not only superior to global ones from a conceptual point of view, because they induce a reduction of complexity (cf. Chomsky (1995) and subsequent work, Epstein *et al.* (1998), Epstein & Seely (2002)),[3] but that they are furthermore supported by strong empirical evidence (cf., for example, Epstein & Seely (2002), Heck & Müller (2000), Müller (2002, 2003, 2004b)). Let us therefore take a closer look at the underlying architecture of derivational theories.

A derivational theory differs from a representational approach in the following way. In a representational theory, a sentence is not built up stepwise in a derivational manner; instead, it is represented by a static structure that can be compared to the outcome, i.e., the final stage, of the derivation of a sentence in a derivational model. Syntactic principles can therefore only refer to this representation, and derivational notions like 'movement' have to be re-

[3]However, cf. Brody (1995, 2002) for a different point of view.

placed with notions like 'chain'. In a derivational approach, by contrast, it is assumed that sentences are built up step by step using the operations Merge and Move, and consequently we can already start computing the structure in the course of the derivation. As a result, at each point in the derivation, material that has not yet been used is in principle not accessible.[4] This means that there is no possibility of look-ahead with respect to syntactic structures that have not been created yet. Moreover, it is possible that access to earlier parts of the derivation is also restricted, and this is what I refer to as 'local derivational approach'.

In the literature, such a local theory has first been proposed by Chomsky (2000 and subsequent work), who introduces the so-called *Phase Impenetrability Condition* in order to restrict the accessible domain 'downwards' (figuratively spoken if we think of syntactic trees).[5] The first version he comes up with is given in (1), which is based on the definitions below (cf. Chomsky (2000:106; 108), Chomsky (2001a:4f.), Chomsky (2001b:12f.); cf. also Müller (2004b)).

(1) *Phase Impenetrability Condition 1* (PIC_1):
The domain of a head X of a phase XP is not accessible to operations outside XP; only X and its edge are accessible to such operations.

(2) The *domain* of a head corresponds to its c-command domain.

(3) CP and vP are *phases*.

(4) The *edge* of a head X is the residue outside X′; it comprises specifiers and elements adjoined to XP.

As a result, the accessible domain is reduced as soon as a phase is completed; material below the head of a completed phase is no longer accessible (cf. (5), where underlined XPs represent phases and material that is not accessible is crossed out).

[4] Of course, the derivation has access to the remaining numeration, but the crucial point is that the syntactic structure that is going to be built up is not available.

[5] But cf. also van Riemsdijk (1978) and Koster (1987) as far as the general idea is concerned that operations are restricted to some extremely local domain.

(5) *Accessible domain under PIC*$_1$:

a. $[_{\text{YP}} \ldots \text{Y} \; [_{\underline{\text{XP}}} \ldots \text{X} \; \cancel{[_{\text{WP}} \ldots \text{W} \; [_{\underline{\text{UP}}} \ldots \text{U} \ldots]]}]]$

b. $[_{\underline{\text{ZP}}} \ldots \text{Z} \; [_{\text{YP}} \ldots \text{Y} \; [_{\underline{\text{XP}}} \ldots \text{X} \; \cancel{[_{\text{WP}} \ldots \text{W} \; [_{\underline{\text{UP}}} \ldots \text{U} \ldots]]}]]]$

However, Chomsky weakens this version of the *Phase Impenetrability Condition*, because he considers it too restrictive if VP-internal Nominative NPs are taken into account (which occur, for example, in Icelandic): In order to be licensed, they have to establish an Agree relation with T; however, following the PIC as defined in (1), these Nominative objects are no longer accessible when T enters the derivation (cf. (5-a) with Y=T, X=v, W=V). Hence, he proposes the modified version given in (6), which expressly makes reference to the next phase and thus enlargens the accessible domain since material is only rendered inaccessible when the next phase has been completed (cf. (7)). Thus, VP-internal material (for instance, a Nominative object) is still accessible when T is merged into the derivation, because the next phase, CP, has not yet been reached (cf. (7-a) with Y=T, X=v, W=V).

(6) *Phase Impenetrability Condition 2* (PIC$_2$):
The domain of a head X of a phase XP is not accessible to operations outside ZP (the next phase); only X and its edge are accessible to such operations.

(7) *Accessible domain under PIC*$_2$:

a. $[_{\text{YP}} \ldots \text{Y} \; [_{\underline{\text{XP}}} \ldots \text{X} \; [_{\text{WP}} \ldots \text{W} \; [_{\underline{\text{UP}}} \ldots \text{U} \; \cancel{[\ldots]}]]]]$

b. $[_{\underline{\text{ZP}}} \ldots \text{Z} \; [_{\text{YP}} \ldots \text{Y} \; [_{\underline{\text{XP}}} \ldots \text{X} \; \cancel{[_{\text{WP}} \ldots \text{W} \; [_{\underline{\text{UP}}} \ldots \text{U} \ldots]]}]]]$

However, from a conceptual point of view this weakening of the *Phase Impenetrability Condition* undermines the whole enterprise of a local derivational syntactic theory, since it enlarges the "representational residue" (cf. Brody (1995, 2002)), and moreover, the question arises as to whether the integration of further constructions would not require a further weakening of the PIC (for example, the integration of binding phenomena; cf. section 3).

In order to overcome the conceptual objections, Müller (2004b) therefore proposes a strengthened version of the PIC which does not refer to phases but to all kinds phrases and is thus called *Phrase Impenetrability Condition*:[6]

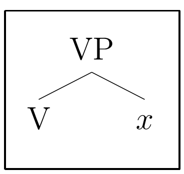
(8) *Phrase Impenetrability Condition* (PIC_3):
The domain of a head X of a phrase XP is not accessible to operations outside XP; only X and its edge are accessible to such operations.
(cf. Müller (2004b:297))

(9) *Accessible domain under* PIC_3:

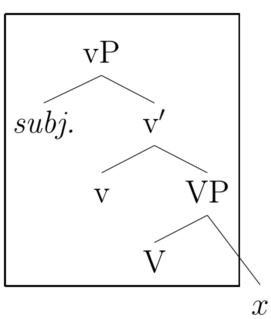
a. $[_{\mathrm{YP}} \ldots \mathrm{Y}\ [_{\underline{\mathrm{XP}}} \ldots \mathrm{X}\ \cancel{[_{\mathrm{WP}} \ldots \mathrm{W}\ [_{\underline{\mathrm{UP}}} \ldots \mathrm{U} \ldots]]}]]$
b. $[_{\underline{\mathrm{ZP}}} \ldots \mathrm{Z}\ [_{\mathrm{YP}} \ldots \mathrm{Y}\ \cancel{[_{\underline{\mathrm{XP}}} \ldots \mathrm{X}\ [_{\mathrm{WP}} \ldots \mathrm{W}\ [_{\underline{\mathrm{UP}}} \ldots \mathrm{U} \ldots]]]}]]$

The effect of the *Phrase Impenetrability Condition* is also exemplified in the following trees, which illustrate how the accessible domain – marked by the frame – shifts when the derivation proceeds. As (10)-(12) show, an item x in the object position will already be unaccessible when vP is completed.

(10) *x still accessible:*

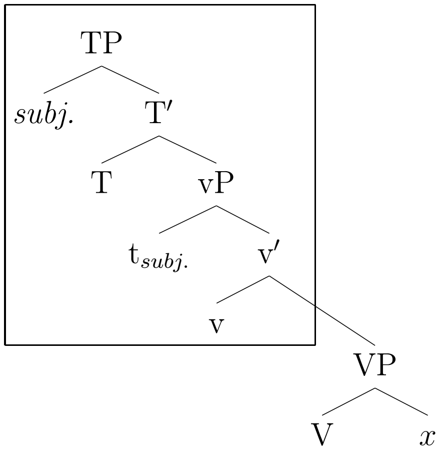


(11) *x no longer accessible:*



[6]Long-distance agreement, as in the case of Nominative objects, would then have to be reinterpreted as involving successive-cyclic feature movement, because the object position and T are obviously not accessible at the same time (cf. (9) with Y=T, X=v, W=V).

(12) *x no longer accessible:*



We will come back to this crucial observation in section 4, but first it will be investigated how much we can restrict the accessible domain if we try to address binding from a derivational perspective.

3 Minimizing the Accessible Domain – Comparing PIC_1, PIC_2, and PIC_3

3.1 General Considerations

If we want to integrate Binding Theory into a derivational framework, we first have to understand how binding principles generally work. What we usually do if we evaluate a binding relation is consider the configuration that holds between the bound element and its antecedent, and based on this information the binding principles allow us to draw conclusions about the grammaticality status of the binding relation. Consider, for instance, the sentences in (13).

(13) a. I know that $[_{\mathrm{TP}}$ Max_1 $[_{\mathrm{vP}}$ t_{Max} hates $\mathrm{himself}_1$/*$\mathrm{him}_1]]$
b. Max_1 knows that $[_{\mathrm{TP}}$ Mary $[_{\mathrm{vP}}$ t_{Mary} likes him_1/*$\mathrm{himself}_1]]$

According to the standard analysis following Chomsky (1981) (cf. the outline in chapter 1 and 2), we have to find out what the binding domain for the bound element is, check whether binding takes place within this domain, and finally Principle A and B of the Binding Theory tell us whether the bound element must be realized as anaphor or pronoun. With respect to (13) we would thus find out that the embedded vP corresponds to the binding domain (– it contains a subject $\neq x$ –) and therefore get the result that the bound element is bound within this domain in (13-a) but not in (13-b), which correctly predicts that we must use the anaphor in the first case and the pronoun in the latter. Similarly, the analysis proposed in chapter 2 presupposes that we know the domain in which binding takes place; only then can we evaluate which realization form the bound element is assigned: In (13-a) we have binding within the θ-domain, hence the anaphor turns out to be the optimal form, in (13-b) the element is only bound in its root domain, therefore it must be realized as pronoun.

In short, in order to be able to draw these conclusions, we must at least know the embedded vP in (13); knowing this part of the derivation, we can then infer that binding takes place within the governing/θ-domain in the case of (13-a), or that binding must take place outside the governing/indicative domain in the case of (13-b), and the binding principles can apply successfully. Thus it seems that we need to be familiar with a certain amount of structure in order to evaluate binding relations. However, the previous section has shown that principles like the PIC restrict access to parts of the derivation. It remains to be seen how this dilemma can be solved.

3.2 Local Binding in English

As described above, there are three different versions of the PIC in the literature, PIC_2 being the most liberal one in the sense that it tolerates a relatively large accessible domain, and PIC_3 being the most restrictive version because it reduces the accessible domain to a minimum. In the following, I will discuss the consequences for binding under the different PIC versions and focus on the question of how much we can restrict the accessible domain if we want to integrate binding into a strictly local derivational theory.

The subsequent derivations are to be read as follows: those parts that are no longer accessible are crossed out; in order to facilitate a direct comparison between the different PIC versions, the examples are ordered in such a way that a_1-z_1 represents the derivation under PIC_1, a_2-z_2 refers to PIC_2, and a_3-z_3 corresponds to the derivation under PIC_3. If the accessible domain is the same under all three PIC versions, the index is omitted. As in chapter 2, the bound element is generally abbreviated as x, and it is assumed that the information as to which items are engaged in a binding relation is indicated by (co-)indexation (with the indices being part of the numeration already).

Let us now consider the derivations of the sentences in (13), starting with (13-a) (repeated in (14)).

(14) I know that Max$_1$ hates himself$_1$/*him$_1$.

a. [$_{\text{VP}}$ hates x_1]

b$_1$. [$_{\text{vP}}$ Max$_1$ hates [$_{\text{VP}}$ t$_{hates}$ x_1]]
c$_1$. [$_{\text{TP}}$ Max$_1$ [$_{\text{vP}}$ t$_{Max}$ hates ~~[$_{\text{VP}}$ t$_{hates}$ x_1]~~]]
d$_1$. [$_{\text{CP}}$ that [$_{\text{TP}}$ Max$_1$ [$_{\text{vP}}$ t$_{Max}$ hates ~~[$_{\text{VP}}$ t$_{hates}$ x_1]~~]]]
e$_1$. [$_{\text{VP}}$ know [$_{\text{CP}}$ that ~~[$_{\text{TP}}$ Max$_1$ [$_{\text{vP}}$ t$_{Max}$ hates [$_{\text{VP}}$ t$_{hates}$ x_1]]]~~]]
f$_1$. [$_{\text{vP}}$ I know [$_{\text{VP}}$ t$_{know}$ [$_{\text{CP}}$ that ~~[$_{\text{TP}}$ Max$_1$ [$_{\text{vP}}$ t$_{Max}$ hates [$_{\text{VP}}$ t$_{hates}$ x_1]]]~~]]]
g$_1$. [$_{\text{TP}}$ I [$_{\text{vP}}$ t$_I$ know ~~[$_{\text{VP}}$ t$_{know}$ [$_{\text{CP}}$ that [$_{\text{TP}}$ Max$_1$ [$_{\text{vP}}$ t$_{Max}$ hates [$_{\text{VP}}$ t$_{hates}$ x_1]]]]]~~]]

b$_2$. [$_{\text{vP}}$ Max$_1$ hates [$_{\text{VP}}$ t$_{hates}$ x_1]]
c$_2$. [$_{\text{TP}}$ Max$_1$ [$_{\text{vP}}$ t$_{Max}$ hates [$_{\text{VP}}$ t$_V$ x_1]]]
d$_2$. [$_{\text{CP}}$ that [$_{\text{TP}}$ Max$_1$ [$_{\text{vP}}$ t$_{Max}$ hates ~~[$_{\text{VP}}$ t$_{hates}$ x_1]~~]]]
e$_2$. [$_{\text{VP}}$ know [$_{\text{CP}}$ that [$_{\text{TP}}$ Max$_1$ [$_{\text{vP}}$ t$_{Max}$ hates ~~[$_{\text{VP}}$ t$_{hates}$ x_1]~~]]]]
f$_2$. [$_{\text{vP}}$ I know [$_{\text{VP}}$ t$_{know}$ [$_{\text{CP}}$ that ~~[$_{\text{TP}}$ Max$_1$ [$_{\text{vP}}$ t$_{Max}$ hates [$_{\text{VP}}$ t$_{hates}$ x_1]]]~~]]]
g$_2$. [$_{\text{TP}}$ I [$_{\text{vP}}$ t$_I$ know [$_{\text{VP}}$ t$_{know}$ [$_{\text{CP}}$ that [~~$_{\text{TP}}$ Max$_1$ [$_{\text{vP}}$ t$_{Max}$ hates [$_{\text{VP}}$ t$_{hates}$ x_1]]]~~]]]]

b$_3$. [$_{\text{vP}}$ Max$_1$ hates [$_{\text{VP}}$ t$_{hates}$ ~~x_1~~]]
c$_3$. [$_{\text{TP}}$ Max$_1$ [$_{\text{vP}}$ t$_{Max}$ hates ~~[$_{\text{VP}}$ t$_{hates}$ x_1]~~]]
d$_3$. [$_{\text{CP}}$ that [$_{\text{TP}}$ Max$_1$ ~~[$_{\text{vP}}$ t$_{Max}$ hates [$_{\text{VP}}$ t$_{hates}$ x_1]]~~]]
e$_3$. [$_{\text{VP}}$ know [$_{\text{CP}}$ that ~~[$_{\text{TP}}$ Max$_1$ [$_{\text{vP}}$ t$_{Max}$ hates [$_{\text{VP}}$ t$_{hates}$ x_1]]]~~]]

f3. [$_{vP}$ I know [$_{VP}$ t$_{know}$ ~~[$_{CP}$ that [$_{TP}$ Max$_1$ [$_{vP}$ t$_{Max}$ hates [$_{VP}$ t$_{hates}$ x_1]]]]]~~]]

g3. [$_{TP}$ I [$_{vP}$ t$_I$ know [~~$_{VP}$ t$_{know}$ [$_{CP}$ that [$_{TP}$ Max$_1$ [$_{vP}$ t$_{Max}$ hates [$_{VP}$ t$_{hates}$ x_1]]]]]]~~]]]

With respect to PIC$_1$ and PIC$_2$, the crucial point in the derivation is represented in (14-b$_1$) and (14-b$_2$), respectively. At this point, the binder is merged into the structure, and the bound element is still accessible. Hence, the binding relation can be evaluated although we do not know yet the complete derivation.[7] With the more restrictive PIC$_3$, it is slightly different; when the binder enters the derivation in (14-b$_3$), the bound element is no longer accessible. Let us therefore go one step back and discuss whether the stage represented in (14-a) allows us to draw conclusions about the binding relation in this example.

Apart from that part of the derivation that has already been built in (14-a), certain subsets of the numeration provide us with some more information. Following Chomsky (2000 and subsequent work), all derivations are based on a so-called lexical array (LA), a set comprising all lexical items that are going to be used in the derivation. In the course of the derivation,

> each phase is determined by a subarray LA$_i$ of LA, placed in "active memory". When the computation exhausts LA$_i$, forming the syntactic object K, L [language] returns to LA, either extending K to K′ or forming an independent structure M to be assimilated later to K or to some extension of K. (Chomsky (2001b:11f.))

This means that at a given point in the derivation we are not only familiar with that part of the already built structure which is in the accessible domain, but we also know the material that is going to be merged into the present phase – however, the syntactic structure that is going to be built up is not known.

[7] The considerations here in section 3 are independent of the version of binding principles we choose. I will therefore not refer to any particular Binding Theory but keep the discussion as general as possible, since the problems binding faces in a derivational approach seem to be valid universally. However, a concrete technical implementation will be proposed in the subsequent sections.

(15) a. LA (lexical array):= set of lexical items used in a derivation;
b. LA_i (subarray):= 'subset' of LA which is selected at that point in the derivation when phase number i begins; it contains the material used for the construction of phase number i; (strictly speaking, LA_i is not necessarily a subset of LA since it can also contain more complex objects composed of elements of LA).
(cf. Chomsky (2000:106, 2001b:11; fn.22))

With respect to the example above, this means that at the stage represented in (14-a) there is only one lexical item left in the current subarray LA_1: *Max*, which is coindexed with x. Hence, *Max* must be merged into a position within vP that c-commands x – there is no other possibility. As a consequence, it can be concluded that x will be bound within the current phase, although binding has not yet taken place, and thus examples like these do not pose a problem for PIC_3.

3.3 Pronominal Binding in English

Let us now turn to the derivation of (13-b), repeated in (16).

(16) Max_1 knows that Mary likes him_1/*himself_1.

a. $[_{\text{VP}}$ likes $x_1]$

b_1. $[_{\text{vP}}$ Mary likes $[_{\text{VP}}\ t_{likes}\ x_1]]$
c_1. $[_{\text{TP}}$ Mary $[_{\text{vP}}\ t_{Mary}$ likes ~~$[_{\text{VP}}\ t_{likes}\ x_1]$~~$]]$
d_1. $[_{\text{CP}}$ that $[_{\text{TP}}$ Mary $[_{\text{vP}}\ t_{Mary}$ likes ~~$[_{\text{VP}}\ t_{likes}\ x_1]$~~$]]]$
e_1. $[_{\text{VP}}$ knows $[_{\text{CP}}$ that ~~$[_{\text{TP}}$ Mary $[_{\text{vP}}\ t_{Mary}$ likes $[_{\text{VP}}\ t_{likes}\ x_1]]]$~~$]]$
f_1. $[_{\text{vP}}\ \text{Max}_1$ knows $[_{\text{VP}}\ t_{knows}\ [_{\text{CP}}$ that ~~$[_{\text{TP}}$ Mary $[_{\text{vP}}\ t_{Mary}$ likes $[_{\text{VP}}\ t_{likes}\ x_1]]]$~~$]]]$
g_1. $[_{\text{TP}}\ \text{Max}_1\ [_{\text{vP}}\ t_{Max}$ knows ~~$[_{\text{VP}}\ t_{knows}\ [_{\text{CP}}$ that $[_{\text{TP}}$ Mary $[_{\text{vP}}\ t_{Mary}$ likes $[_{\text{VP}}\ t_{likes}\ x_1]]]]]$~~$]]$

b_2. $[_{\text{vP}}$ Mary likes $[_{\text{VP}}\ t_{likes}\ x_1]]$
c_2. $[_{\text{TP}}$ Mary $[_{\text{vP}}\ t_{Mary}$ likes $[_{\text{VP}}\ t_{likes}\ x_1]]]$
d_2. $[_{\text{CP}}$ that $[_{\text{TP}}$ Mary $[_{\text{vP}}\ t_{Mary}$ likes ~~$[_{\text{VP}}\ t_{likes}\ x_1]$~~$]]]$
e_2. $[_{\text{VP}}$ knows $[_{\text{CP}}$ that $[_{\text{TP}}$ Mary $[_{\text{vP}}\ t_{Mary}$ likes ~~$[_{\text{VP}}\ t_{likes}\ x_1]$~~$]]]]$

f$_2$. [$_{\text{vP}}$ Max$_1$ knows [$_{\text{VP}}$ t$_{Max}$ [$_{\text{CP}}$ that [~~$_{\text{TP}}$ Mary [$_{\text{vP}}$ t$_{Mary}$ likes [$_{\text{VP}}$ t$_{likes}$ x_1]]]~~]]]

g$_2$. [$_{\text{TP}}$ Max$_1$ [$_{\text{vP}}$ t$_{Max}$ knows [$_{\text{VP}}$ t$_{knows}$ [$_{\text{CP}}$ that ~~[$_{\text{TP}}$ Mary [$_{\text{vP}}$ t$_{Mary}$ likes[$_{\text{VP}}$ t$_{likes}$ x_1]]]~~]]]]

b$_3$. [$_{\text{vP}}$ Mary likes [$_{\text{VP}}$ t$_{likes}$ ~~x_1~~]]

c$_3$. [$_{\text{TP}}$ Mary [$_{\text{vP}}$ t$_{Mary}$ likes ~~[$_{\text{VP}}$ t$_{likes}$ x_1]~~]]

d$_3$. [$_{\text{CP}}$ that [$_{\text{TP}}$ Mary ~~[$_{\text{vP}}$ t$_{Mary}$ likes [$_{\text{VP}}$ t$_{likes}$ x_1]]~~]]

e$_3$. [$_{\text{VP}}$ knows [$_{\text{CP}}$ that ~~[$_{\text{TP}}$ Mary [$_{\text{vP}}$ t$_{Mary}$ likes [$_{\text{VP}}$ t$_{likes}$ x_1]]]~~]]

f$_3$. [$_{\text{vP}}$ Max$_1$ knows [$_{\text{VP}}$ t$_{knows}$ [$_{\text{CP}}$ ~~that [$_{\text{TP}}$ Mary [$_{\text{vP}}$ t$_{Mary}$ likes [$_{\text{VP}}$ t$_{likes}$ x_1]]]]~~]]

g$_3$. [$_{\text{TP}}$ Max$_1$ [$_{\text{vP}}$ t$_{Max}$ knows ~~[$_{\text{VP}}$ t$_{knows}$ [$_{\text{CP}}$ that [$_{\text{TP}}$ Mary [$_{\text{vP}}$ t$_{Mary}$ likes [$_{\text{VP}}$ t$_{likes}$ x_1]]]]]~~]]

The last point in the derivation when the bound element x is still accessible under PIC$_1$ is represented in (16-b$_1$). In contrast to the previous example, the antecedent has not yet been merged into the derivation at this stage. However, one can see that in (16-b$_1$), the current phase has just been completed (LA$_1$=$\{\}$); hence it can be concluded that binding does not take place within this phase, and this information might be sufficient for the binding principles to evaluate this binding relation.[8] Basically the same considerations hold for the derivation under PIC$_2$.

Regarding PIC$_3$, the last point in the derivation when x is still accessible is represented in (16-a). However, we know furthermore that the only element left in LA$_1$ is *Mary*, which is not coindexed with x. Therefore it can be concluded that x will not be bound within the current phase (= embedded vP), which means that the restrictive PIC$_3$ basically leaves us with the same information as the more liberal PIC$_1$ and PIC$_2$.

[8] For instance, it can be concluded that binding takes place outside the subject domain (i.e., the traditional binding domain), thus both the traditional Principle B and the constraints from chapter 2 would predict that x must be a pronoun. Note, however, that in other languages the information "binding takes place outside the subject domain" might not suffice to conclude that the bound element must be realized as pronoun (cf. chapter 2 and the discussion of long distance anaphora below).

In the previous two examples, the subarray LA_1 played a crucial role; however, one might doubt whether it always contains enough information to ensure such an early evaluation of the binding relation. Let us therefore turn to some more complex examples.

3.4 The Complex NP Problem

Since in the analyses above, LA_1 contained at most one element at the crucial stage, let us first examine what happens if more than one element is left in LA_1. In (17), we have the following situation: As far as PIC_1 and PIC_2 are concerned, there is again a point in the derivation when both coindexed elements are accessible (cf. (17-$b_{1,2}$)); hence the example does not offer any new insights. However, under PIC_3 the example differs from the previous ones insofar as at the last point at which x is accessible (i.e., in (17-a)), LA_1 contains more than one lexical item – what we need to complete the first phase is the complex NP *the man whom Max_1 saw*.

(17) The man whom Max_1 saw threatens him_1/*$himself_1$.

a. $[_{\text{VP}}$ threatens $x_1]$

$b_{1,2}$. $[_{\text{vP}}$ the man whom Max_1 saw threatens $[_{\text{VP}}\ t_{threatens}\ x_1]]$

c_1. $[_{\text{TP}}$ the man whom Max_1 saw $[_{\text{vP}}\ t_{subj.}$ threatens $\cancel{[_{\text{VP}}\ t_{threatens}\ x_1]}]]$

c_2. $[_{\text{TP}}$ the man whom Max_1 saw $[_{\text{vP}}\ t_{subj.}$ threatens $[_{\text{VP}}\ t_{threatens}\ x_1]]]$

b_3. $[_{\text{vP}}$ the man whom Max_1 saw threatens $[_{\text{VP}}\ t_{threatens}\ \cancel{x_1}]]$

c_3. $[_{\text{TP}}$ the man whom Max_1 saw $[_{\text{vP}}\ t_{subj.}$ threatens $\cancel{[_{\text{VP}}\ t_{threatens}\ x_1]}]]$

There are now two possibilities. Since the construction of the complex NP proceeds in parallel (cf. Chomsky (2001b:fn.22)), LA_1 might already contain the full structure when VP is completed. On this assumption, we can foresee in (17-a) that although the coindexed element Max_1 will be merged into the current phase, it will not c-command and therefore not bind x – and thus it should be possible to determine the realization of x at this stage.

However, there is a second possibility. Of course, the complex NP must be built before it can be merged into the derivation, and thus it must be part of LA_1 at some stage; but this might as well happen *after* the completion of VP (the last stage in the derivation when x is still accessible under PIC_3).[9] On this assumption, we cannot know at the stage of (17-a) whether x will be bound by *Max* within the embedded vP or not; the material in LA_1 does not allow us to draw any conclusions – for instance, the complex NP might turn out to be *Max_1, whom the man saw*, in which case x would indeed be bound.

Hence, it seems that we are forced to conclude that for sentences like this one PIC_1 or PIC_2 are more suitable and that PIC_3 might be too restrictive. However, further examples will reveal that it is an illusion that the two more liberal PIC variants do not face problems like these.

3.5 German *A.c.I.*-Constructions: Binding Across Two Successive Phases

If we take a closer look at the examples in (14) (*I know that Max_1 hates $himself_1$/*him_1*) and (17) (*The man whom Max_1 saw threatens him_1/*$himself_1$*), we find one crucial similarity. In (14), we have a relatively local binding relation; binding occurs within one phase. In (17), there is no binding relation at all, but the two coindexed elements also enter the derivation within the same phase. Thus, the two coindexed elements are in both examples part of the same phase; and since under both PIC_1 and PIC_2

[9] For instance, the derivation might proceed as follows:

(i) a. $[_{\mathrm{VP}}$ threatens $x_1]$
LA_1={the, man, whom, Max_1, saw}
$\rightarrow$ x still accessible, but complex NP not built yet
b. $[_{\mathrm{vP}}$ $[_{\mathrm{v}}$ threatens$]$ $[_{\mathrm{VP}}$ $\mathrm{t}_{threatens}$ $\cancel{x_1}]]$
LA_1={$[_{\mathrm{NP}}$ the man whom Max_1 saw$]$}
$\rightarrow$ complex NP built, but x no longer accessible
c. $[_{\mathrm{vP}}$ $[_{\mathrm{NP}}$ the man whom Max_1 saw$]$ $[_{\mathrm{v}}$ threatens$]$ $[_{\mathrm{VP}}$ $\mathrm{t}_{threatens}$ $\cancel{x_1}]]$
LA_1={}

the whole phase is accessible at the stage when it is completed, there is a point in the derivation when both elements are accessible. This explains why these examples do not pose a problem for these two PIC versions. However, the question arises as to what happens if the coindexed elements enter the derivation in different phases.

Let us therefore consider German *A.c.I.*-constructions. In sentences like (18), the bound element is realized as an anaphor; but in comparison to example (14), which also involved anaphoric binding, binding is not as local in this case: x and its antecedent occur in different phases.

This becomes evident if we contrast (18-c) with (18-e); x_1 is part of the embedded vP, whereas Max_1 is merged into the next phase, another vP.[10]

(18) *Ger.* Sarah glaubt, dass Max$_1$ Peter für sich$_1$ arbeiten lässt.
Sarah believes that Max Peter for SE work let
'Sarah believes that Max$_1$ makes Peter work for him$_1$.'

a. [$_{\text{PP}}$ für x_1]

b$_1$. [$_{\text{VP}}$ [$_{\text{PP}}$ für x_1] arbeiten]
c$_1$. [$_{\text{vP}}$ Peter [$_{\text{VP}}$ [$_{\text{PP}}$ für x_1] t$_{arbeiten}$] arbeiten]
d$_1$. [$_{\text{VP}}$ [$_{\text{vP}}$ Peter ~~[$_{\text{VP}}$ [$_{\text{PP}}$ für x_1] t$_{arbeiten}$]~~ arbeiten] lässt]
e$_1$. [$_{\text{vP}}$ Max$_1$ [$_{\text{VP}}$ [$_{\text{vP}}$ Peter ~~[$_{\text{VP}}$ [$_{\text{PP}}$ für x_1] t$_{arbeiten}$]~~ arbeiten] t$_{lässt}$] lässt]

b$_2$. [$_{\text{VP}}$ [$_{\text{PP}}$ für x_1] arbeiten]
c$_2$. [$_{\text{vP}}$ Peter [$_{\text{VP}}$ [$_{\text{PP}}$ für x_1] t$_{arbeiten}$] arbeiten]
d$_2$. [$_{\text{VP}}$ [$_{\text{vP}}$ Peter [$_{\text{VP}}$ [$_{\text{PP}}$ für x_1] t$_{arbeiten}$] arbeiten] lässt]
e$_2$. [$_{\text{vP}}$ Max$_1$ [$_{\text{VP}}$ [$_{\text{vP}}$ Peter ~~[$_{\text{VP}}$ [$_{\text{PP}}$ für x_1] t$_{arbeiten}$]~~ arbeiten] t$_{lässt}$] lässt]

b$_3$. [$_{\text{VP}}$ [$_{\text{PP}}$ für ~~x_1~~] arbeiten]
c$_3$. [$_{\text{vP}}$ Peter [$_{\text{VP}}$ ~~[$_{\text{PP}}$ für x_1]~~ t$_{arbeiten}$] arbeiten]
d$_3$. [$_{\text{VP}}$ [$_{\text{vP}}$ Peter ~~[$_{\text{VP}}$ [$_{\text{PP}}$ für x_1] t$_{arbeiten}$]~~ arbeiten] lässt]
e$_3$. [$_{\text{vP}}$ Max$_1$ [$_{\text{VP}}$ [~~$_{\text{vP}}$ Peter [$_{\text{VP}}$ [$_{\text{PP}}$ für x_1] t$_{arbeiten}$] arbeiten]~~ t$_{lässt}$]

[10] In the following, not the complete but only the relevant parts of the derivations are illustrated.

lässt]

What are the consequences for the different PIC versions? Starting with PIC_2, the most liberal variant, we can observe that the last point in the derivation at which x is accessible is given in (18-d_2), and at this stage the antecedent has not yet entered the derivation. However, the second phase is already being built, and in LA_2 there is only one element left, namely Max_1. Hence, we can infer that x_1 will be bound in the current phase.

This is reminiscent of the analyses of (14) (*I know that* Max_1 *hates* $himself_1$/*him_1) and (16) (Max_1 *knows that Mary likes* him_1/*$himself_1$) under PIC_3; and similarly, we face the same problem if Max_1 is replaced with a more complex NP.

(19) Ich will nicht, dass [der Mann, der gesagt hat, dass Max_1 ein
I want not that the man who said has that Max a
guter Mitarbeiter sei], mich für ihn_1 arbeiten lässt.
good employee be_{sub} me for him work let
'I don't want the man who said that Max_1 was a good employee make me work for him_1.'

As in example (17) above, the subarray LA_2 no longer contains only one element at the last stage in the derivation when x is accessible – LA_2 rather contains all the material needed to build the complex NP (on the assumption that this has not yet been done), and hence we cannot foresee at this point whether the coindexed item Max_1 will finally c-command and thus bind x_1.

Now what about PIC_1 and PIC_3? As to PIC_3, the last point in the derivation at which x is accessible is given in (18-a). At this stage not even the first phase has been completed, so we are left with LA_1={Peter, arbeiten} while the coindexed item Max_1 is in the remaining LA. Therefore we only know that x will not be bound within the current phase; any further predictions are not possible.

For PIC_1, the last point in the derivation at which x is accessible is represented in (18-c_1). This means that the first phase has just been completed, i.e., $\text{LA}_1=\{\}$. On the basis of this information, it is not possible either to make any predictions about the final binding configuration, which means that PIC_1 faces exactly the same problem as PIC_3 in view of sentences like (18).

To sum up, in examples in which the two coindexed elements are no longer part of the same phase but of two successive phases, both PIC_1 and PIC_3 cannot say anything about the binding configuration. By contrast, the more liberal version PIC_2 provides us with the same information as PIC_3 did in the previous examples, when the coindexed items were merged into the same phase; it allows a prediction under certain circumstances, namely if the subarray to which the second element belongs does not contain "too much" material such that we can foresee its designated structural position when x is still accessible.

But what if a binding relation extends over more than two successive phases?

3.6 Long Distance Binding in Icelandic – Binding Across More than Two Phases

In languages like English, this is not that problematic, because here we know that if an element is not bound at least within its subject domain, it cannot be realized as anaphor anyway (cf. example (16) (*Max*$_1$ *knows that Mary likes him*$_1$*/*himself*$_1$)). But in languages with long distance binding, the situation is different.[11]

In the following Icelandic example, even the most liberal PIC version, PIC_2, does not provide enough information to evaluate the binding relation in the course of the derivation. The last point in the derivation at which x is accessible under PIC_2 is represented in (20-c_2), where the second phase has already begun and LA_2={að}; hence we know that x will not be bound within the second phase either, but this information is not enough to draw any conclusions about possible realizations of x. And if this is true for PIC_2, it must definitely be true for the more restrictive PIC_1 and PIC_3.

[11] In fact, if we assume that the competition does not only choose between anaphoric and pronominal binding but also decides whether x can be realized as pronoun or R-expression (cf. chapter 2, section **??**), we face the same problems in English-type languages; in this case, we need to know whether x will be bound in its root domain or not.

(20) *Ice.* Jón$_1$ segir að Pétur raki sig$_1$/*sjálfan sig$_1$/hann$_1$.
John says that Peter shave$_{sub}$ SE/himself/him
'John$_1$ says that Peter would shave him$_1$.'

a. $[_{\text{VP}}\ \text{raki}\ x_1]$

b$_1$. $[_{\text{vP}}\ \text{Pétur raki}\ [_{\text{VP}}\ \text{t}_{raki}\ x_1]]$
c$_1$. $[_{\text{TP}}\ \text{Pétur raki}\ [_{\text{vP}}\ \text{t}_{Pétur}\ \text{t}_{raki}\ \cancel{[_{\text{VP}}\ \text{t}_{raki}\ x_1]}]]$
d$_1$. $[_{\text{CP}}\ \text{að}\ [_{\text{TP}}\ \text{Pétur raki}\ [_{\text{vP}}\ \text{t}_{Pétur}\ \text{t}_{raki}\ \cancel{[_{\text{VP}}\ \text{t}_{raki}\ x_1]}]]]$
e$_1$. $[_{\text{VP}}\ \text{segir}\ [_{\text{CP}}\ \text{að}\ \cancel{[_{\text{TP}}\ \text{Pétur raki}\ [_{\text{vP}}\ \text{t}_{Pétur}\ \text{t}_{raki}\ [_{\text{VP}}\ \text{t}_{raki}\ x_1]]]}]]$
f$_1$. $[_{\text{vP}}\ \text{Jón}_1\ \text{segir}\ [_{\text{VP}}\ \text{t}_{segir}\ [_{\text{CP}}\ \text{að}\ \cancel{[_{\text{TP}}\ \text{Pétur raki}\ [_{\text{vP}}\ \text{t}_{Pétur}\ \text{t}_{raki}\ [_{\text{VP}}\ \text{t}_{raki}\ x_1]]]}]]]$

b$_2$. $[_{\text{vP}}\ \text{Pétur raki}\ [_{\text{VP}}\ \text{t}_{raki}\ x_1]]$
c$_2$. $[_{\text{TP}}\ \text{Pétur raki}\ [_{\text{vP}}\ \text{t}_{Pétur}\ \text{t}_{raki}\ [_{\text{VP}}\ \text{t}_{raki}\ x_1]]]$
d$_2$. $[_{\text{CP}}\ \text{að}\ [_{\text{TP}}\ \text{Pétur raki}\ [_{\text{vP}}\ \text{t}_{Pétur}\ \text{t}_{raki}\ \cancel{[_{\text{VP}}\ \text{t}_{raki}\ x_1]}]]]$
e$_2$. $[_{\text{VP}}\ \text{segir}\ [_{\text{CP}}\ \text{að}\ [_{\text{TP}}\ \text{Pétur raki}\ [_{\text{vP}}\ \text{t}_{Pétur}\ \text{t}_{raki}\ \cancel{[_{\text{VP}}\ \text{t}_{raki}\ x_1]}]]]]$
f$_2$. $[_{\text{vP}}\ \text{Jón}_1\ \text{segir}\ [_{\text{VP}}\ \text{t}_{segir}\ [_{\text{CP}}\ \text{að}\ \cancel{[_{\text{TP}}\ \text{Pétur raki}\ [_{\text{vP}}\ \text{t}_{Pétur}\ \text{t}_{raki}\ [_{\text{VP}}\ \text{t}_{raki}\ x_1]]]}]]]$

b$_3$. $[_{\text{vP}}\ \text{Pétur raki}\ [_{\text{VP}}\ \text{t}_{raki}\ \cancel{x_1}]]$
c$_3$. $[_{\text{TP}}\ \text{Pétur raki}\ [_{\text{vP}}\ \text{t}_{Pétur}\ \text{t}_{raki}\ \cancel{[_{\text{VP}}\ \text{t}_{raki}\ x_1]}]]$
d$_3$. $[_{\text{CP}}\ \text{að}\ [_{\text{TP}}\ \text{Pétur raki}\ \cancel{[_{\text{vP}}\ \text{t}_{Pétur}\ \text{t}_{raki}\ [_{\text{VP}}\ \text{t}_{raki}\ x_1]]}]]$
e$_3$. $[_{\text{VP}}\ \text{segir}\ [_{\text{CP}}\ \text{að}\ \cancel{[_{\text{TP}}\ \text{Pétur raki}\ [_{\text{vP}}\ \text{t}_{Pétur}\ \text{t}_{raki}\ [_{\text{VP}}\ \text{t}_{raki}\ x_1]]]}]]$
f$_3$. $[_{\text{vP}}\ \text{Jón}_1\ \text{segir}\ [_{\text{VP}}\ \text{t}_{segir}\ [_{\text{CP}}\ \cancel{\text{að}\ [_{\text{TP}}\ \text{Pétur raki}\ [_{\text{vP}}\ \text{t}_{Pétur}\ \text{t}_{raki}\ [_{\text{VP}}\ \text{t}_{raki}\ x_1]]]]}]]$

However, the argumentation above only holds if it is assumed that LA as a whole is not accessible during the derivation of phase i (but only LA$_i$). If we consider the remaining elements in LA in (20-b$_1$), (20-c$_2$), and (20-a), we can see that it only contains *Jón*$_1$*, segir*, and, depending on the PIC version, *að* – all the other lexical items have either already been merged into the derivation or are part of the current subarray LA$_i$. And this information would suffice to conclude that the conditions for Icelandic long distance anaphors will be met in the end: *Jón*$_1$ will c-command x_1, and only a subjunctive complement will intervene between the bound element and its antecedent.

But again, we could modify the example in such a way that even this information would no longer be available. Consider the following two Icelandic sentences, which have identical underlying numerations; the only difference is that in (21-a) *Jón*$_1$ is merged into the derivation earlier than in (21-b) (in the third phase instead of the fifth phase), and as a result, only the first one allows anaphoric binding.

(21) a. Max veit$_{ind}$ að Jón$_1$ segir$_{ind}$ að Pétur raki$_{sub}$ sig$_1$.
Max knows that Jón says that Pétur shave$_{sub}$ SE
'Max knows that John$_1$ says that Peter would shave him$_1$.'

b. *Jón$_1$ veit$_{ind}$ að Max segir$_{ind}$ að Pétur raki$_{sub}$ sig$_1$.
Jón knows that Max says that Pétur shave$_{sub}$ SE
'John$_1$ knows that Max says that Peter would shave him$_1$.'

If we now turn to the derivation of (21-a) (illustrated in (22)), we can make the following observation. As in example (20), *Jón*$_1$ is still in LA at the last point in the derivation when x is accessible, independent of the PIC version we choose. However, this time even the remaining items in LA do not allow us to draw any conclusions about the binding configuration. Under PIC$_2$, for instance, LA={Jón$_1$, Max, veit$_{ind}$, segir$_{ind}$, að} (cf. (22-c$_2$)), and thus we cannot yet decide whether long distance binding will be possible or not, because at this stage the derivations of sentence (21-a) and (21-b) are still completely identical. This means that not until LA$_3$ is selected can we decide whether (21-a) or (21-b) is derived and hence evaluate the binding relation. However, when this selection takes place, x is no longer accessible, even under the liberal PIC$_2$.

(22) Max veit$_{ind}$ að Jón$_1$ segir$_{ind}$ að Pétur raki$_{sub}$ sig$_1$.

a. [$_{VP}$ raki x_1]

b$_1$. [$_{vP}$ Pétur raki [$_{VP}$ t$_{raki}$ x_1]]

c$_1$. [$_{TP}$ Pétur raki [$_{vP}$ t$_{Pétur}$ t$_{raki}$ ~~[$_{VP}$ t$_{raki}$ x_1]~~]]

d$_1$. [$_{CP}$ að [$_{TP}$ Pétur raki [$_{vP}$ t$_{Pétur}$ t$_{raki}$ ~~[$_{VP}$ t$_{raki}$ x_1]~~]]]

e$_1$. [$_{VP}$ segir [$_{CP}$ að ~~[$_{TP}$ Pétur raki [$_{vP}$ t$_{Pétur}$ t$_{raki}$ [$_{VP}$ t$_{raki}$ x_1]]]~~]]

f$_1$. [$_{vP}$ Jón$_1$ segir [$_{VP}$ t$_{segir}$ [$_{CP}$ að [~~$_{TP}$ Pétur raki [$_{vP}$ t$_{Pétur}$ t$_{raki}$ [$_{VP}$ t$_{raki}$ x_1]]]~~]]]]

g$_1$. [$_{\text{TP}}$ Jón$_1$ segir [$_{\text{vP}}$ t$_{\textit{Jón}_1}$ t$_{\textit{segir}}$ [~~$_{\text{VP}}$ t$_{\textit{segir}}$ [$_{\text{CP}}$ að [$_{\text{TP}}$ Pétur raki [$_{\text{vP}}$ t$_{\textit{Pétur}}$ t$_{\textit{raki}}$ [$_{\text{VP}}$ t$_{\textit{raki}}$ x_1]]]]]~~]]

h$_1$. [$_{\text{CP}}$ að [$_{\text{TP}}$ Jón$_1$ segir [$_{\text{vP}}$ t$_{\textit{Jón}_1}$ t$_{\textit{segir}}$ ~~[$_{\text{VP}}$ t$_{\textit{segir}}$ [$_{\text{CP}}$ að [$_{\text{TP}}$ Pétur raki [$_{\text{vP}}$ t$_{\textit{Pétur}}$ t$_{\textit{raki}}$ [$_{\text{VP}}$ t$_{\textit{raki}}$ x_1]]]]]~~]]]

i$_1$. [$_{\text{VP}}$ veit [$_{\text{CP}}$ að [~~$_{\text{TP}}$ Jón$_1$ segir [$_{\text{vP}}$ t$_{\textit{Jón}_1}$ t$_{\textit{segir}}$ [$_{\text{VP}}$ t$_{\textit{segir}}$ [$_{\text{CP}}$ að [$_{\text{TP}}$ Pétur raki [$_{\text{vP}}$ t$_{\textit{Pétur}}$ t$_{\textit{raki}}$ [$_{\text{VP}}$ t$_{\textit{raki}}$ x_1]]]]]]]~~]]

j$_1$. [$_{\text{vP}}$ Max veit [$_{\text{VP}}$ t$_{\textit{veit}}$ [$_{\text{CP}}$ að [~~$_{\text{TP}}$ Jón$_1$ segir [$_{\text{vP}}$ t$_{\textit{Jón}_1}$ t$_{\textit{segir}}$ [$_{\text{VP}}$ t$_{\textit{segir}}$ [$_{\text{CP}}$ að [$_{\text{TP}}$ Pétur raki [$_{\text{vP}}$ t$_{\textit{Pétur}}$ t$_{\textit{raki}}$ [$_{\text{VP}}$ t$_{\textit{raki}}$ x_1]]]]]]]~~]]]

b$_2$. [$_{\text{vP}}$ Pétur raki [$_{\text{VP}}$ t$_{\textit{raki}}$ x_1]]

c$_2$. [$_{\text{TP}}$ Pétur raki [$_{\text{vP}}$ t$_{\textit{Pétur}}$ t$_{\textit{raki}}$ [$_{\text{VP}}$ t$_{\textit{raki}}$ x_1]]]

d$_2$. [$_{\text{CP}}$ að [$_{\text{TP}}$ Pétur raki [$_{\text{vP}}$ t$_{\textit{Pétur}}$ t$_{\textit{raki}}$ ~~[$_{\text{VP}}$ t$_{\textit{raki}}$ x_1]~~]]]

e$_2$. [$_{\text{VP}}$ segir [$_{\text{CP}}$ að [$_{\text{TP}}$ Pétur raki [$_{\text{vP}}$ t$_{\textit{Pétur}}$ t$_{\textit{raki}}$ ~~[$_{\text{VP}}$ t$_{\textit{raki}}$ x_1]~~]]]]

f$_2$. [$_{\text{vP}}$ Jón$_1$ segir [$_{\text{VP}}$ t$_{\textit{segir}}$ [$_{\text{CP}}$ að [~~$_{\text{TP}}$ Pétur raki [$_{\text{vP}}$ t$_{\textit{Pétur}}$ t$_{\textit{raki}}$ [$_{\text{VP}}$ t$_{\textit{raki}}$ x_1]]]~~]]]

g$_2$. [$_{\text{TP}}$ Jón$_1$ segir [$_{\text{vP}}$ t$_{\textit{Jón}_1}$ t$_{\textit{segir}}$ [$_{\text{VP}}$ t$_{\textit{segir}}$ [$_{\text{CP}}$ að [~~$_{\text{TP}}$ Pétur raki [$_{\text{vP}}$ t$_{\textit{Pétur}}$ t$_{\textit{raki}}$ [$_{\text{VP}}$ t$_{\textit{raki}}$ x_1]]]~~]]]]]

h$_2$. [$_{\text{CP}}$ að [$_{\text{TP}}$ Jón$_1$ segir [$_{\text{vP}}$ t$_{\textit{Jón}_1}$ t$_{\textit{segir}}$ ~~[$_{\text{VP}}$ t$_{\textit{segir}}$ [$_{\text{CP}}$ að [$_{\text{TP}}$ Pétur raki [$_{\text{vP}}$ t$_{\textit{Pétur}}$ t$_{\textit{raki}}$ [$_{\text{VP}}$ t$_{\textit{raki}}$ x_1]]]]]~~]]]

i$_2$. [$_{\text{VP}}$ veit [$_{\text{CP}}$ að [$_{\text{TP}}$ Jón$_1$ segir [$_{\text{vP}}$ t$_{\textit{Jón}_1}$ t$_{\textit{segir}}$ [~~$_{\text{VP}}$ t$_{\textit{segir}}$ [$_{\text{CP}}$ að [$_{\text{TP}}$ Pétur raki [$_{\text{vP}}$ t$_{\textit{Pétur}}$ t$_{\textit{raki}}$ [$_{\text{VP}}$ t$_{\textit{raki}}$ x_1]]]]]~~]]]]

j$_2$. [$_{\text{vP}}$ Max veit [$_{\text{VP}}$ t$_V$ [$_{\text{CP}}$ að ~~[$_{\text{TP}}$ Jón$_1$ segir [$_{\text{vP}}$ t$_{\textit{Jón}_1}$ t$_{\textit{segir}}$ [$_{\text{VP}}$ t$_{\textit{segir}}$ [$_{\text{CP}}$ að [$_{\text{TP}}$ Pétur raki [$_{\text{vP}}$ t$_{\textit{Pétur}}$ t$_{\textit{raki}}$ [$_{\text{VP}}$ t$_{\textit{raki}}$ x_1]]]]]]]~~]]]

b$_3$. [$_{\text{vP}}$ Pétur raki [$_{\text{VP}}$ t$_{\textit{raki}}$ ~~x_1~~]]

c$_3$. [$_{\text{TP}}$ Pétur raki [$_{\text{vP}}$ t$_{\textit{Pétur}}$ t$_{\textit{raki}}$ ~~[$_{\text{VP}}$ t$_{\textit{raki}}$ x_1]~~]]

d$_3$. [$_{\text{CP}}$ að [$_{\text{TP}}$ Pétur raki ~~[$_{\text{vP}}$ t$_{\textit{Pétur}}$ t$_{\textit{raki}}$ [$_{\text{VP}}$ t$_{\textit{raki}}$ x_1]]~~]]

e$_3$. [$_{\text{VP}}$ segir [$_{\text{CP}}$ að ~~[$_{\text{TP}}$ Pétur raki [$_{\text{vP}}$ t$_{\textit{Pétur}}$ t$_{\textit{raki}}$ [$_{\text{VP}}$ t$_{\textit{raki}}$ x_1]]]~~]]

f$_3$. [$_{\text{vP}}$ Jón$_1$ segir [$_{\text{VP}}$ t$_{\textit{segir}}$ [~~$_{\text{CP}}$ að [$_{\text{TP}}$ Pétur raki [$_{\text{vP}}$ t$_{\textit{Pétur}}$ t$_{\textit{raki}}$ [$_{\text{VP}}$ t$_{\textit{raki}}$ x_1]]]]~~]]

g$_3$. [$_{\text{TP}}$ Jón$_1$ segir [$_{\text{vP}}$ t$_{\textit{Jón}_1}$ t$_{\textit{segir}}$ [~~$_{\text{VP}}$ t$_{\textit{segir}}$ [$_{\text{CP}}$ að [$_{\text{TP}}$ Pétur raki [$_{\text{vP}}$ t$_{\textit{Pétur}}$ t$_{\textit{raki}}$ [$_{\text{VP}}$ t$_{\textit{raki}}$ x_1]]]]]~~]]

h$_3$. [$_{\text{CP}}$ að [$_{\text{TP}}$ Jón$_1$ seg~~ir [$_{\text{vP}}$ t$_{\textit{Jón}_1}$ t$_{\textit{segir}}$ [$_{\text{VP}}$ t$_{\textit{segir}}$ [$_{\text{CP}}$ að [$_{\text{TP}}$ Pétur raki [$_{\text{vP}}$ t$_{\textit{Pétur}}$ t$_{\textit{raki}}$ [$_{\text{VP}}$ t$_{\textit{raki}}$ x_1]]]]]]~~]]

i$_3$. [$_{\text{VP}}$ veit [$_{\text{CP}}$ að [~~$_{\text{TP}}$ Jón$_1$ segir [$_{\text{vP}}$ t$_{\textit{Jón}_1}$ t$_{\textit{segir}}$ [$_{\text{VP}}$ t$_{\textit{segir}}$ [$_{\text{CP}}$ að~~

~~[$_{\text{TP}}$ Pétur raki [$_{\text{vP}}$ t$_{Pétur}$ t$_{raki}$ [$_{\text{VP}}$ t$_{raki}$ x_1]]]]]]]]~~]]

j3. [$_{\text{vP}}$ Max veit [$_{\text{VP}}$ t$_V$ ~~[$_{\text{CP}}$ að [$_{\text{TP}}$ Jón$_1$ segir [$_{\text{vP}}$ t$_{Jón_1}$ t$_{segir}$ [$_{\text{VP}}$ t$_{segir}$ [$_{\text{CP}}$ að [$_{\text{TP}}$ Pétur raki [$_{\text{vP}}$ t$_{Pétur}$ t$_{raki}$ [$_{\text{VP}}$ t$_{raki}$ x_1]]]]]]]]]~~]]

3.7 Conclusion

Let us now come back to the question of how much we can restrict the accessible domain if we want to integrate binding into a strictly local derivational approach.

As the discussion above has shown, all three PIC variants eventually face the same problem. As it stands, they do not seem to provide enough information to evaluate binding relations. This means that we have to find a way how the relevant information can be transferred into the accessible domain. As this move is inevitable independent of the PIC variant we assume, we are free to choose the version that is most attractive from a conceptual point of view, and this is the most restrictive version; henceforth, I will therefore assume that the accessible domain is defined by PIC$_3$, the *Phrase Impenetrability Condition*.

4 Binding as Feature Checking

4.1 Introduction

Against this background, the question arises of whether it is possible to capture an *a priori* non-local phenomenon like binding in such a theory at all. There are two reasons why binding seems to pose a problem for a local derivational approach. First, binding is obviously not a strictly local phenomenon, as the following well-known examples show, which illustrate pronominal binding in English and long distance binding in Icelandic, respectively.[12]

(23) John$_1$ knows that Mary told Sally that Max hit him$_1$.

[12] In fact, even if we consider a relatively local binding relation as in *John$_1$ hates himself$_1$*, the anaphor in the object position is no longer accessible when the subject enters the derivation; cf. the illustration in (i), repeated from section 2

(24) Jón$_1$ segir að Pétur raki sig$_1$/hann$_1$/*sjálfan sig$_1$.
John says that Peter shave$_{sub}$ SE/him/himself
'John$_1$ says that Peter would shave him$_1$.'

Moreover, the locality degree of the binding relation determines the shape of the bound element, which might surface as SE anaphor, as SELF anaphor, or as pronoun. This is exemplified by the following German sentences (repeated from chapter 2), where the bound element becomes less anaphoric the less local the binding relation gets.

(25) *German:*

a. Max$_1$ hasst sich selbst$_1$/sich$_1$/*ihn$_1$.
Max hates himself/SE/him
'Max$_1$ hates himself$_1$.'

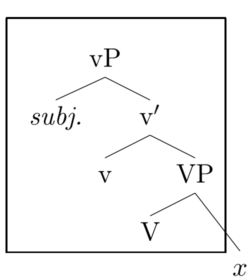
b. Max$_1$ hört sich selbst$_1$/sich$_1$/*ihn$_1$ singen.
Max hears himself/SE/him sing
'Max$_1$ hears himself$_1$ sing.'

c. Max$_1$ schaut hinter sich$_1$/??sich selbst$_1$/*ihn$_1$.
Max glanced behind SE/himself/him
'Max$_1$ glanced behind him$_1$/himself$_1$.'

d. Max$_1$ weiß, dass Maria ihn$_1$/*sich$_1$/*sich selbst$_1$ mag.
Max knows that Mary him/SE/himself likes
'Max$_1$ knows that Mary likes him$_1$.'

What these examples show is that the solution to the locality problem cannot just be to split up the non-local relation into several local ones, as it is done, for example, in the case of *wh*-movement. With respect to binding, something

(i) *x no longer accessible:*



more needs to be said.

4.2 Phrase Balance and Feature Checking

Generally, it can be concluded that in order to evaluate a binding relation, it is necessary that all information concerning this relation is accessible at the same time at some point in the derivation. In short, it is necessary that there is a point in the derivation when both the bound element and its antecedent are accessible. Since the bound element is merged into the derivation first, such a configuration can only arise after the binder has also entered the derivation. However, as this might happen at a stage when the base position of the bound element is no longer accessible (independent of the PIC version we choose, as the previous section showed), it seems to be necessary that the bound element is "dragged along" until it reaches a position which is still accessible when the binder comes in. Thus, the question arises of what triggers movement of the bound element?

In general, movement can be characterized as follows. The ultimate goal of all movement operations is feature checking; thus we are led to conclude that bound elements bear particular features which have to be checked in the course of the derivation. As far as the target of movement is concerned, it is always a position which stands in a very local relation to the element bearing the attracting features, and it is this local relation which licenses feature checking. Regarding the case of binding, we have said that the bound element must move to a position which is still accessible when its antecedent is merged into the derivation. When this happens, we reach a stage in the derivation where the binding relation can be evaluated, which means that afterwards the bound element no longer needs to be moved along (unless it serves as a goal for some higher probe). The position the bound element will have reached at this point can be precisely specified: Assuming that the accessible domain is restricted by the *Phrase Impenetrability Condition* (PIC_3) (cf. the conclusions drawn in the previous section), its target position must be one specifier position below its antecedent – for example, if the binder is a subject, which is merged in Specv, the bound element must be raised at least to SpecV in order to be accessible at the same time.

This means that the relation between the bound element and its antecedent is very similar to that of other probes and goals: goals are generally attracted to a position sufficiently close to the probe for feature checking, and unless the goal bears further features that are attracted by some other higher probe, it stops moving at this point. Let us briefly consider what "sufficiently close for feature checking" actually means. Following Chomsky (1995), the standard situation looks as follows. The probe is a head and the goal is an XP which is attracted to the probe's specifier position such that feature checking takes place in a spec-head relation (cf., for example, feature checking involving *wh*-features, EPP features, Case features, scrambling features etc.).[13] But does this mean that spec-head relations are the only configurations under which feature checking can take place? Against the background of the *Phrase Impenetrability Condition* (PIC_3), which imposes severe locality restrictions on all operations, it seems redundant to introduce a further locality constraint and assume that feature checking is restricted to spec-head relations; instead, it is more attractive to subsume the locality conditions for feature checking under the PIC_3. As a consequence, not only the specifier of the probe serves as potential target for attracted XPs, but also the specifier of the next lower maximal projection.[14] Moreover (and more importantly), if feature checking is not dependent on a spec-head relation, in principle nothing prevents the probe from being a maximal projection; i.e., feature checking needs no longer involve heads.[15,16]

[13]In contrast, Chomsky (2000, 2001b) assumes that movement of the goal to the probe's specifier position is not necessary unless EPP features are involved; the relation Agree (under which feature checking takes place in this approach) does not presuppose a spec-head relation. Similarly to the assumptions developed here it is sufficient that probe and goal are in a c-command relation which is "local enough", the latter being restricted by the *Phase Impenetrability Condition* (PIC_2) and the MLC.

[14]Cf. also Heck (2004) as regards the assumption that feature checking only requires some "sufficiently local" configuration.

[15]In principle, it would be possible to assume an even more local configuration for feature checking involving two XPs: a spec-spec relation between multiple specifiers of the same maximal projection. But as outlined above, from a conceptual point of view it seems to be more reasonable to link feature checking to the PIC_3, under which the next lower specifier position is local enough.

[16]I adopt Sternefeld's (2004) notation according to which features on probes are starred.

(26) *Feature Checking:*
The pair of features [$*$F$*$]/[F] stands in a feature checking relation iff
(i) the element bearing the feature [$*$F$*$] (= *probe*) c-commands the element bearing the feature [F] (= *goal*) and
(ii) both probe and goal are accessible.

(27) *Possible configuration in which feature checking can take place:*
$[_{\text{ZP}}\ \text{WP}_{[*F*]}\ \text{Z}\ [_{\text{YP}}\ \text{XP}_{[F]}\ \text{Y}]]$

Against this background it seems to be a natural assumption that binding can be encoded as feature checking, with the antecedent as probe and the bound element as goal.[17] Let us therefore assume that items that function as bound elements in a derivation bear a feature [β], and their designated antecedents are equipped with a corresponding feature [$*\beta*$]. Because of the PIC_3 the element bearing the [β]-feature will be forced to move successive-cyclically via all intermediate specifier positions to its checking position, which is the first specifier position below the element bearing the [$*\beta*$]-feature. Following Müller (2004b), I assume that the intermediate movement steps are triggered by the constraint *Phrase Balance*.[18]

(28) *Phrase Balance* (PB):
Every XP has to be balanced: For every feature [$*$F$*$] in the numeration there must be a potentially available feature [F] at the XP level.
(cf. Müller (2004b:297))

[17]Note in particular that the c-command requirement of binding is thus encoded in the more general definition of *Feature Checking*; cf. (26).

[18]Since the underlying idea is to restrict look-ahead to the numeration (which does not divulge syntactic structure), the concept of subarrays (LA_i) is abandoned.

Note furthermore that *Phrase Balance* refers to completed XPs. This means that it applies at the point when there is no further material left in the numeration that is merged into XP. Hence, even if a head is merged with its complement and the result is considered to be a maximal projection at this stage, *Phrase Balance* does not yet apply if there is a specifier left in the numeration.

(29) *Potential availability*:
A feature [F] is potentially available if (i) or (ii) holds:
(i) [F] is on X or edgeX of the present root of the derivation.
(ii) [F] is in the workspace of the derivation. (cf. Müller (2004b:298))

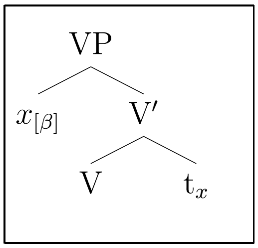
(30) The *workspace* of a derivation D comprises the numeration N and material in the trees that have been created earlier (with material from N) and have not yet been used in D. (cf. Müller (2004b:298))

The following abstract derivation serves as an illustration. *Phrase Balance* forces the bound element to move to SpecU (cf. (31-b)) and SpecY (cf. (31-c)), which turns out to be a position in which the element can enter into a checking relation with the binder, because SpecY is still accessible when the binder is merged into the derivation (cf. (31-d)).

(31) a. workspace: $\{$U, $x_{[\beta]}$ (=bound element), Y, Z, binder$_{[*\beta*]}\}$
b. $[_{\text{UP}}$ $x_{[\beta]}$ U $\text{t}_x]$; workspace: $\{$Y, Z, binder$_{[*\beta*]}\}$
c. $[_{\text{YP}}$ $x_{[\beta]}$ Y $[_{\text{UP}}$ $\text{t}_x{}'$ U $\cancel{\text{t}_x}]]$; workspace: $\{$Z, binder$_{[*\beta*]}\}$
d. $[_{\text{ZP}}$ binder$_{[*\beta*]}$ Z $[_{\text{YP}}$ $x_{[\beta]}$ Y $\cancel{[_{\text{UP}}\ \text{t}_x{}'\ \text{U}\ \text{t}_x]}]]$

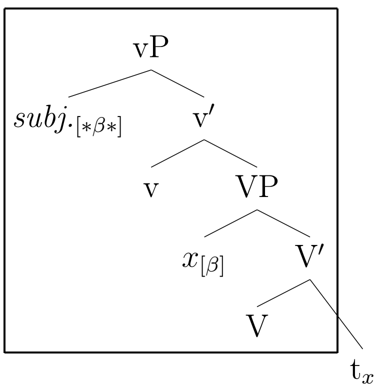
In short, *Phrase Balance* triggers movement of $x_{[\beta]}$ to the edge of the current phrase as long as its antecedent (with the feature $[*\beta*]$) is still in the numeration and thus makes sure that $x_{[\beta]}$ remains accessible. This is illustrated in the following trees. Since *Phrase Balance* forces $x_{[\beta]}$ to move to the edge of VP in (32), $x_{[\beta]}$ is still in the accessible domain at the next derivational stage (cf. (33) and (34)). When vP is built, it depends on the probe as to whether $x_{[\beta]}$ moves on or not: If the probe is merged into the derivation (as in (33)), $x_{[\beta]}$ stays in its position and feature checking takes place; if the probe remains in the numeration (as in (34)), *Phrase Balance* triggers again movement of $x_{[\beta]}$ to the edge of vP.

(32) *Phrase Balance:*



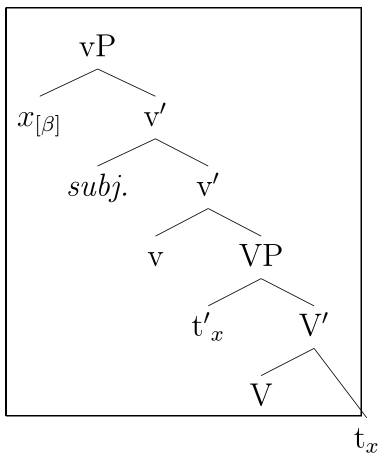
Num={subj.$_{[*\beta*]}$, ...}

(33) *[β]-feature checking:*



Num={...}

(34) *Phrase Balance:*



Num={subj.$_{[*\beta*]}$, ...}

In the next section I will address the question of how a concrete implementation of such a binding theory might look like.

5 Optimal Binding in a Derivational Approach

5.1 An Outline

In the previous section, it has been explained how x gets into the accessible domain; in this section, the issue will be addressed of how the concrete form of x is determined.

As argued in chapter 2, there are good reasons to assume that the concrete realization of bound elements is determined in an optimality-theoretic competition. I will therefore investigate how the approach outlined there can be integrated into a strictly local derivational theory.

The underlying idea is that the numeration of a sentence in which a binding relation is established does not contain the concrete lexical item which will later function as bound element; instead, it is only encoded that there will be a binding relation between a designated antecedent (identifiable by the [$*\beta*$]-feature) and a bindee x bearing the corresponding [β]-feature.[19] However, even if we do not know the concrete form of x at this stage, we know its possible realizations: Depending on the locality degree of the binding relation, x will be realized as SELF anaphor, as SE anaphor, or as a pronoun. Hence, I propose that in the beginning, x is equipped with a realization matrix, i.e., a list which contains all possible realizations of x. I will refer to it with the following notation: [SELF, SE, pron].[20] In the course of the derivation, x's concrete realization will then be determined as follows.

First of all recall that a basic insight in chapter 2 was that binding is sensitive to domains of different size (cf. also, among others, Manzini & Wexler (1987), Dalrymple (1993)). In essence, the following generalization holds: The smaller the domain is in which binding takes place, the more likely it is that the bound element is realized as an anaphor, or, to put it differently, the more anaphoric x is.[21]

Let us now come back to the strictly local derivational approach. In which domain x will eventually be bound can in principle only be inferred when the binder is merged into the derivation and the checking relation is established. However, even if we do not know in the course of the derivation in which domain x will eventually be bound, we do know earlier in which domains x is *not* bound. Hence, if a domain relevant for binding is reached in the course of the derivation and x is still free, we can conclude that it becomes more and more unlikely that x will be realized as an anaphor. On the assumption that in the beginning x is equipped with the complete realization matrix, this observation has the following consequence: Each time when x reaches one of these domains to which binding is sensitive and x remains unbound, its realization matrix might be reduced insofar as the most anaphoric specification is deleted and henceforth no longer available. Whether deletion takes place or not hinges on the respective domain and the language under consideration. Note that due to the introduction of realization matrices, the *Inclusiveness*

[19] Regarding unbound pronouns and inherently reflexive predicates, some slight modifications will be in order; I will come back to these issues in section 5.11 and 5.12

[20] Copies of R-expressions and $\emptyset$ might also be included; cf. section 5.10 and 5.11

[21] The following hierarchy illustrates how anaphoricity decreases, with A $>$ B indicating that A is more anaphoric than B: SELF anaphor $>$ SE anaphor $>$ pronoun $>$ R-expression (cf. also chapter 2).

Condition can thus be respected, which requires that "no new objects are added in the course of the computation" (Chomsky (1995:228)). Although the concrete form of x is determined in the course of the derivation, all possible realizations underly the derivation, and those which must be excluded are gradually deleted.

As alluded to before, x finally stops moving when it can establish a checking relation with its antecedent. Again, the realization matrix is optimized (which means that certain specifications might be deleted), and the result is mapped to PF.[22] Before Late Insertion takes place (cf. Halle & Marantz (1993) and subsequent work on Distributed Morphology), the concrete realization of x can finally be determined, which must match one of the remaining forms in the realization matrix. If there is only one element left in the matrix, the choice is clear, otherwise the remaining form that is most anaphoric is selected.

Once the realization of x is known, the whole chain it heads can be aligned and x can then be spelled out in the appropriate position. This constitutes a minimal violation of the *Phrase Impenetrability Condition* and the *Strict Cycle Condition*,[23] but apparently this is what we have to accept if we want to integrate such a non-local phenomenon as binding into a local derivational approach.

Note, however, that this violation of the locality requirements is restricted to PF and does not occur in narrow syntax. Moreover, it is clear that this step is and must not be abused to carry arbitrary information back to those parts of the structure that are no longer in the accessible domain, or to use this information to change parts of the derivation in a way that could not be foreseen at the point when that part was being built, and thereby undermine the derivational approach as such. Instead, this kind of reference to earlier parts of the derivation is strictly restricted to items that have some connection to the current stage of the derivation via chain formation, and the

[22] As will become clearer in the subsequent sections, once x is bound, the matrix will not change anymore. (This follows from the formulation of the constraints.) Hence, it can immediately be mapped to PF.

[23] *Strict Cycle Condition* (SCC):
Within the current XP α, a syntactic operation may not target a position that is included within another XP β that is dominated by α.

only thing that happens is that the lower chain members are specified more precisely in accordance with the predispositions they already had before.

Thus, chains are like wormholes in physics – they are "hypothetical "tube[s]" [...] connecting widely separated positions", "allowing an object that passes through it to appear instantaneously in some other part of the Universe – not just in a different place, but also in a different time", so to speak.[24] As a result, through this tube lower chain members can be aligned with their head, but other parts of the already built structure are not affected at all.

5.2 Domains, Constraints, and Candidates

Let us now turn to the technical implementation of the analysis. In chapter 2, six different domains have been distinguished which were shown to be relevant for binding: the θ-domain, the Case domain, the subject domain, the finite domain, the indicative domain, and the root domain.[25] Two remarks concerning these domains are in order. First, the definitions of the domains have to be slightly modified, since we have to take into account that the analysis in chapter 2 was a global one, and hence the domain definitions need to be adjusted to the derivational model.

Second, note that according to the considerations in the previous section, it is no longer relevant in which domain binding takes place, instead, we are now interested in the last domain in which x is *not yet* bound. The consequences are twofold. On the one hand, we can dispense with the notion of root domain, because if the smallest domain in which binding takes place is the root domain, it suffices to know that x is not yet bound in the indicative domain, the next smaller domain; that x will be eventually bound can then be inferred from the unchecked [β]-feature on x and the [$*\beta*$]-feature in the numeration. On the other hand, for the case that the binding relation is established in the smallest domain from chapter 2, the θ-domain, we

[24]These are quotations from http://scienceworld.wolfram.com/physics/Wormhole.html and http://www.lcsd.gov.hk/CE/Museum/Space/FAQ/universe/e_faq_universe_1.htm (November 23, 2003).

[25]Recall that this order reflects their increasing size.

have to introduce a new constraint that refers to the situation before the θ-domain is reached, because languages also differ with respect to their binding possibilities in this small domain.

Let me now introduce the relevant definitions, before we can then turn to the analysis of concrete examples.[26]

(35) XP is the *θ-domain* of x if it contains x and the head that θ-marks x plus its external argument (if there is one).

(36) XP is the *Case domain* of x if it contains x and the head that bears the Case features against which x checks Case.

(37) XP is the *subject domain* of x if it contains x and either
(i) a subject distinct from x which does not contain x, or
(ii) the T with which x a checks its (Nominative) Case features.

(38) XP is the *finite domain* of x if it contains x, a finite verb, and a subject.

(39) XP is the *indicative domain* of x if it contains x, an indicative verb, and a subject.

The main difference between the old domain definitions and the ones introduced above is that we often find the additional requirement that the domain contain a subject. This is generally necessary in order to guarantee that the respective domain is not VP, where verbs are usually base-generated, but vP, where a potential binder can be merged into the derivation. Due to obligatory V-to-v movement, this was not necessary in a global approach, where the verb could be considered in its landing site.[27]

[26] Note that in section 5.12, the definitions of finite and indicative domain will have to be slightly revised.

[27] As an alternative, one could try to define the domains via concrete categories (for example, vP= finite domain). However, this does not really simplify things, because there are on the one hand XPs that correspond to different domains (vP can qualify as all kinds of domains), and on the other hand the θ-domain, for example, can be a vP in one sentence and a PP in another one. Hence, at least further specifications such as $\mathrm{vP}_{[finite]}$= finite domain would be needed, and it is unclear to me how a categorial definition of the

Moreover, it does no longer make sense to refer to the *smallest* XP with a particular property; after all we cannot know at a given stage in the derivation whether an XP in the inaccessible domain has before qualified as one of the relevant domains. As a consequence, there are no longer unambiguously defined sets of nodes that constitute the different domains; instead, it is possible that several XPs in the derivation qualify, for instance, as θ-domain, the only requirement being that the accessible domain contains at the same time x, its θ-marker and the corresponding external argument. However, the underlying idea that the domains are ordered in a subset relation can still be maintained if it is understood in such a way that for two domains D_1 and D_2 the relation $D_1 \subseteq D_2$ holds iff the smallest XP that qualifies as D_2 contains the smallest YP that qualifies as D_1.

Let us now take a closer look at the optimization procedure. The theory I propose relies on serial optimization, which means that optimization applies more than once (cf. Müller (2003)); and since we adopt the *Phrase Impenetrability Condition*, it is almost self-evident that optimization takes place after the completion of each phrase (in the sense alluded to before in footnote 18). The optimal output of each competition (plus additional material from the remaining numeration) serves as input for the next optimization process (cf. Heck & Müller (2000) and subsequent work on derivational OT syntax and recall also the analysis proposed in chapter 3, section **??**). As far as the initial input is concerned, it basically consists of the numeration, which contains in particular the designated binder with the [$*\beta*$]-feature and the bound element x, which bears the feature [β] and has the realization matrix [SELF, SE, pron]. Let us now focus on the latter. In the beginning, we start with the maximal realization matrix. Assume then that when the first optimization takes place, we have the option to reduce the matrix by deleting one or two of the most anaphoric elements such that we get the following three candidates: O_1: $\ldots x_{[SELF,SE,pron]} \ldots$; O_2: $\ldots x_{[SE,pron]} \ldots$; O_3: $\ldots x_{[pron]} \ldots$. If O_2 wins the competition, only $x_{[SE,pron]}$ and $x_{[pron]}$ compete when the next optimization takes place, since the realization matrix cannot be extended in the course of the derivation (otherwise it would violate the *Inclusiveness Condition*); it can only be further reduced.

θ-domain would have to look like.

As far as the constraints are concerned, the reflexivity constraints from chapter 2 are now replaced with a universal constraint subhierarchy that does not punish binding of non-maximally anaphoric elements in a given domain; instead, as argued above, the new constraints require x to be minimally anaphoric if binding has not yet taken place. To put it another way, anaphors are punished if they occur unbound in a more or less local domain – and this sounds rather familiar, because it is very similar to the traditional Principle A, which requires that anaphors be locally bound. Hence, the new constraints that replace the reflexivity constraints from chapter 2 (which could be regarded as a version of Principle B) can be considered to be a version of Principle A. Thus, they will be called PRINCIPLE $\mathcal{A}$-constraints.

They have the form outlined in (40), with XD $\in$ {θ-domain, Case domain, subject domain, finite domain, indicative domain}, and are universally ordered in such a way that constraints referring to bigger domains outrank those referring to smaller ones.

(40) PRINCIPLE $\mathcal{A}_{XD}$ (Pr.$\mathcal{A}_{XD}$):
If $x_{[\beta]}$ remains unchecked in its XD, x must be minimally anaphoric.

These constraints work as follows: If the derivation reaches one of the relevant domains and no binding relation is established, i.e., if one can infer from the material in the accessible domain that this is the case, they apply non-vacuously and are violated twice by the candidate with the realization matrix [SELF, SE, pron] and once by O_2, with the matrix [SE, pron]. Hence, the effect of these constraints is that they reduce anaphoric realization possibilities.

As it stands, (40) does not yet suffice to account for languages that have different binding options if binding takes place within the θ-domain, the smallest domain relevant for binding (cf. the remark at the beginning of this section). In order to distinguish between those languages, we need a constraint that applies *before* the θ-domain is reached; thus, the PRINCIPLE $\mathcal{A}$-constraint subhierarchy is extended by the following constraint, which applies non-vacuously in all optimization processes as long as x remains unbound since it refers to maximal projections in general. Hence, it can already apply before the θ-domain is reached and punish unbound anaphors even in such a local domain; informally spoken, it can thus be characterized as an extremely local PRINCIPLE $\mathcal{A}$-constraint.[28]

(41) PRINCIPLE $\mathcal{A}_{XP}$ (Pr.$\mathcal{A}_{XP}$):
If $x_{[\beta]}$ remains unchecked in XP, x must be minimally anaphoric.

As far as the ranking of the PRINCIPLE $\mathcal{A}$-constraints is concerned, (41) is outranked by the rest of the subhierarchy because it refers to the most local domain. The complete universal hierarchy is given in (42), with those constraints referring to bigger domains dominating the constraints referring to smaller domains.[29] This reflects that it is worse if anaphoric x reaches a relatively big domain and is still free.

(42) *Universal subhierarchy 1:*
Pr.$\mathcal{A}_{ID} \gg$ Pr.$\mathcal{A}_{FD} \gg$ Pr.$\mathcal{A}_{SD} \gg$ Pr.$\mathcal{A}_{CD} \gg$ Pr.$\mathcal{A}_{ThD} \gg$ Pr.$\mathcal{A}_{XP}$

There are two cases in which the PRINCIPLE $\mathcal{A}$-constraints apply vacuously – either if the binder is merged into the accessible domain and $x_{[\beta]}$ is checked, or if the accessible domain does not contain any material that corresponds to one of the relevant domains.

In these cases, a second group of constraints decides the competition (cf. (43)). They also form a universal constraint subhierarchy, which is the counterpart of the *SELF-hierarchy from chapter 2. They punish candidates involving a realization matrix for x that does not contain a particular specification. However, while the *SELF-hierarchy preferred non-maximally anaphoric elements, the new constraints are ordererd in such a way that they basically favour anaphoric specifications. This is achieved by the ranking in (44), since it favours realization matrices that have not been reduced.[30]

[28]Note in particular that if the generic label PRINCIPLE $\mathcal{A}_{XD}$ is used, it henceforth also subsumes PRINCIPLE $\mathcal{A}_{XP}$.

[29]As mentioned before, the domain definitions are no longer unique in a derivational approach in the sense that an unambiguously specified set of nodes constitutes a particular domain of x in a given sentence; this is in particular true for the domain referred to in (41). The notions "bigger/smaller domains" are therefore to be understood in such a way that a domain X is smaller than a domain Y if the derivation first reaches a maximal projection that qualifies as X before a maximal projection is reached that qualifies as Y.

[30]Maximal realization matrices do not violate any of the FAITH-constraints, but the

Thus, these constraints function as a counterbalance to the PRINCIPLE $\mathcal{A}$-constraints.

(43) a. FAITH$_{SELF}$ (F$_{SELF}$):
The realization matrix for x must contain [SELF].
b. FAITH$_{SE}$ (F$_{SE}$):
The realization matrix for x must contain [SE].
c. FAITH$_{pron}$ (F$_{pron}$):
The realization matrix for x must contain [pron].

(44) *Universal subhierarchy 2:*
FAITH$_{pron}$ $\gg$ FAITH$_{SE}$ $\gg$ FAITH$_{SELF}$

None of the constraints introduced so far says anything about the concrete realization of x; they only help to determine an optimal realization matrix. Hence, we need an additional rule which applies at PF and determines the final form on the basis of the optimal matrix. Assume that this task is fulfilled by the following principle.

(45) MAXIMALLY ANAPHORIC BINDING (MAB):
Checked $x_{[\beta]}$ must be realized maximally anaphorically.

So let us now apply the theory outlined above and turn to some concrete examples.

5.3 Derivational Binding in German

In this and the following three sections, I provide analyses of the German, English, Dutch, and Italian binding data introduced in chapter 2 to illustrate how the theory works in detail. Let us first turn to the German sentences in (46), repeated from (25). As we saw in chapter 2, these four sentences involve binding relations of different locality degree. In (46-a), the binding relation is already established when the smallest XP that qualifies as θ-domain (i.e., the minimal θ-domain) is reached, namely vP. In (46-b), the antecedent is not

more specifications are deleted, the more (higher-ranked) FAITH-constraints are violated. (Recall that first the SELF specification and then the SE specification is deleted.)

contained in the minimal θ-domain (= embedded vP); it enters the derivation in the matrix vP, which qualifies as Case domain. In (46-c), the minimal θ- and Case domain coincide (= PP), but the binder is not part of it; the binding relation is only established when the minimal subject domain (= vP) is reached. Finally, in (46-d), where the embedded vP corresponds to the minimal θ-, Case, subject, finite, and indicative domain, the binding relation is least local, since the binder only enters the derivation in the matrix vP.

(46) *German:*

a. Max$_1$ hasst sich selbst$_1$/sich$_1$/*ihn$_1$.
   Max hates himself/SE/him
   'Max$_1$ hates himself$_1$.'

b. Max$_1$ hört sich selbst$_1$/sich$_1$/*ihn$_1$ singen.
   Max hears himself/SE/him sing
   'Max$_1$ hears himself$_1$ sing.'

c. Max$_1$ schaut hinter sich$_1$/??sich selbst$_1$/*ihn$_1$.
   Max glanced behind SE/himself/him
   'Max$_1$ glanced behind him$_1$/himself$_1$.'

d. Max$_1$ weiß, dass Maria ihn$_1$/*sich$_1$/*sich selbst$_1$ mag.
   Max knows that Mary him/SE/himself likes
   'Max$_1$ knows that Mary likes him$_1$.'

Let us now consider the derivation of each of these sentences, starting with (46-a) (repeated in (47)).

First, the verb and its direct object, $x_{[\beta]}$, merge (cf. (47-a)) and form VP. However, this phrase is not yet balanced, because there is a starred feature in the remaining numeration, namely $[*\beta*]$, for which there is no corresponding feature $[\beta]$ potentially available (cf. (29)). The only feature that could satisfy this requirement is the $[\beta]$-feature on x, but x is neither in V nor in edgeV nor in the workspace of the derivation. Hence, *Phrase Balance* (cf. (28)) triggers movement of x to the edge of VP. This is indicated in (47-b).

(47) Max$_1$ hasst sich selbst$_1$/sich$_1$/*ihn$_1$.

a. $[_{\text{VP}}\ x_{[\beta]}$ hasst$]$; workspace: $\{\text{Max}_{[*\beta*]}, \ldots\}$

b. $[_{\text{VP}}\ x_{[\beta]}\ [_{\text{V}'}\ \text{t}_x$ hasst$]]$

At this stage, the first optimization takes place (cf. T_1).[31] Since $x_{[\beta]}$ is still unchecked and a maximal projection is completed, PRINCIPLE $\mathcal{A}_{XP}$ applies non-vacuously; further domains to which binding is sensitive have not yet been reached.[32] Moreover, the FAITH-constraints are relevant in the first competition.

As far as the candidates are concerned, the question arises as to whether x keeps the full realization matrix [SELF, SE, pron], with which it is equipped in the beginning, or whether it is reduced to [SE, pron] or [pron].

As to the ranking of the constraints, the universal hierarchy FAITH$_{pron} \gg$ FAITH$_{SE} \gg$ FAITH$_{SELF}$ must be respected; and since in the end both types of anaphors must be optimal in German sentences of this kind, both O_1 and O_2 must win this competition. This is achieved if FAITH$_{SELF}$ and PRINCIPLE $\mathcal{A}_{XP}$ are tied.[33,34]

T_1: VP optimization
(XP reached - $x_{[\beta]}$ unchecked)

Input: $[_{\text{VP}}\ x_{[\beta]/[SELF,SE,pron]}\ [_{\text{V}'}\ t_x$ hasst]]	F$_{pron}$	F$_{SE}$	F$_{SELF}$	PR.$\mathcal{A}_{XP}$
⇒ O_1: $[_{\text{VP}}\ x_{[\beta]/[SELF,SE,pron]}\ [_{\text{V}'}\ t_x$ hasst]]				**(!)
⇒ O_2: $[_{\text{VP}}\ x_{[\beta]/[SE,pron]}\ [_{\text{V}'}\ t_x$ hasst]]			*(!)	*
O_3: $[_{\text{VP}}\ x_{[\beta]/[pron]}\ [_{\text{V}'}\ t_x$ hasst]]		*!	*	

T_1 yields two optimal outputs; this means that there are two possibilities as to how the derivation can proceed (= two optimal derivations). However, since they only differ with respect to the realization matrix of x, the continuation

[31] In the subsequent analyses I ignore the maximal projection(s) that makes up x itself, because at this early stage nothing of interest happens. Moreover, the candidates will be abbreviated and only the different realization matrices will be represented in the subsequent tableaux.

[32] Those PRINCIPLE $\mathcal{A}$-constraints that apply vacuously are generally neglected in the tableaux.

[33] Again, all ties in this analysis are global ties.

[34] Strictly speaking, it cannot yet be excluded that the crucial ranking is FAITH$_{SELF}$ $\gg$ PRINCIPLE $\mathcal{A}_{XP}$; in this case only O_1 would win, which still comprises all possible realizations. However, it would also turn out to be the only optimal candidate in the next optimization, in which case MAB would wrongly predict that only the complex anaphor is licit in sentences like these.

of both variants basically looks as follows.

(48) c. $[_{\text{vP}}\ \text{Max}_{[*\beta*]}\ [_{\text{VP}}\ x_{[\beta]}\ [_{\text{V}'}\ \cancel{t_x}\ t_{hasst}]]\ \text{hasst}]$

At this stage, the θ-domain of x is reached, but since at the same time x's binder enters the derivation, all PRINCIPLE $\mathcal{A}$-constraints apply vacuously when vP is optimized. This optimization is illustrated in $T_{1.1}$ with O_1 from T_1 as input (notation in the tableaux: O_1/T_1), and in $T_{1.2}$ with O_2 from T_1 as input.[35]

Hence, $T_{1.1}$ involves again three candidates, whereas in $T_{1.2}$, only two candidates compete. In $T_{1.1}$, the [SELF, SE, pron] candidate wins; in $T_{1.2}$, this matrix is no longer available and the matrix [SE, pron] is predicted to be optimal.

$T_{1.1}$: vP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ *applies vacuously)*

	Input: O_1/T_1	F_{pron}	F_{SE}	F_{SELF}
⇒	O_{11}: [SELF, SE, pron]			
	O_{12}: [SE, pron]			∗!
	O_{13}: [pron]		∗!	∗

$T_{1.2}$: vP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ *applies vacuously)*

	Input: O_2/T_1	F_{pron}	F_{SE}	F_{SELF}
⇒	O_{21}: [SE, pron]			∗
	O_{22}: [pron]		∗!	∗

Moreover, now that $x_{[\beta]}$ has been checked, MAB can determine the concrete realization of x. According to the derivation in which O_{11}=[SELF, SE, pron] is optimal, MAB selects the complex anaphor; in the derivation where [SE, pron] is optimal (cf. O_{21}), the SE anaphor is chosen as realization of x. Hence,

[35]The derivational history of the candidates is reflected by their indices. Thus a candidate O_{xy} is the y-th candidate in the second optimization process based on the winner O_x from the first competition; O_{xyz} would then be the z-th candidate in the third competition based on the previous winner O_{xy}, and so on.

the analysis makes correct predictions.

Let us now consider the derivation of sentence (46-b), which is repeated in (49). (49-a) represents the derivation when the first phrase is completed and the first optimization takes place.

(49) Max$_1$ hört sich selbst$_1$/sich$_1$/*ihn$_1$ singen.

a. [$_{\text{vP}}$ $x_{[\beta]}$ singen]

At this stage, the θ-domain of x is reached, and since x remains unchecked, both PRINCIPLE $\mathcal{A}_{XP}$ and PRINCIPLE $\mathcal{A}_{ThD}$ apply non-vacuously. As in T$_1$, both O$_1$ and O$_2$ should turn out to be optimal, because both types of anaphors are licit in sentences like these. Hence, PRINCIPLE $\mathcal{A}_{ThD}$ cannot be ranked above FAITH$_{SELF}$; but since the latter is tied with PRINCIPLE $\mathcal{A}_{XP}$ (cf. T$_1$) and PRINCIPLE $\mathcal{A}_{ThD}$ must be universally higher ranked than PRINCIPLE $\mathcal{A}_{XP}$, it must be assumed that PRINCIPLE $\mathcal{A}_{ThD}$ and FAITH$_{SELF}$ are also tied.[36] Thus, we get the following partial ranking for German:

(50) FAITH$_{pron}$ $\gg$ FAITH$_{SE}$ $\gg$ (PR.$\mathcal{A}_{ThD}$ $\gg$ PR.$\mathcal{A}_{XP}$) $\circ$ FAITH$_{SELF}$

T$_2$: vP optimization
(XP/ThD reached – $x_{[\beta]}$ unchecked)

	Candidates	F$_{pron}$	F$_{SE}$	PR.$\mathcal{A}_{ThD}$	F$_{SELF}$	PR.$\mathcal{A}_{XP}$
⇒	O$_1$: [SELF, SE, pron]			$**(!)$		$**$
⇒	O$_2$: [SE, pron]			$*$	$*(!)$	$*$
	O$_3$: [pron]		$*!$		$*$	

[36]Recall from chapter 2, section **??** that I assume that ties are not transitive (cf. Fischer (2001)). The brackets in the ranking in (50) indicate that although both PRINCIPLE $\mathcal{A}$-constraints are tied with FAITH$_{SELF}$, the dominance relation between them is not given up. Thus, (50) is an abbreviation for the following three constraint orders:

(i) FAITH$_{pron}$ $\gg$ FAITH$_{SE}$ $\gg$ PR.$\mathcal{A}_{ThD}$ $\gg$ PR.$\mathcal{A}_{XP}$ $\gg$ FAITH$_{SELF}$ ($\rightarrow$ winner in T$_2$: O$_2$)
(ii) FAITH$_{pron}$ $\gg$ FAITH$_{SE}$ $\gg$ PR.$\mathcal{A}_{ThD}$ $\gg$ FAITH$_{SELF}$ $\gg$ PR.$\mathcal{A}_{XP}$ ($\rightarrow$ winner in T$_2$: O$_2$)
(iii) FAITH$_{pron}$ $\gg$ FAITH$_{SE}$ $\gg$ FAITH$_{SELF}$ $\gg$ PR.$\mathcal{A}_{ThD}$ $\gg$ PR.$\mathcal{A}_{XP}$ ($\rightarrow$ winner in T$_2$: O$_1$)

Since we have again two optimal outputs, there are two competitions when the next phrase is completed.[37]

(51) b. $[_{\text{VP}}\ x_{[\beta]}\ [_{\text{vP}}\ \text{t}_x\ \text{singen}]\ \text{hört}]$

At this point, no new domain is reached, but note that unlike in the global approach in chapter 2, this maximal projection still counts as θ-domain since all defining criteria are met. Thus, the same constraints as in T_2 remain relevant. As a result, we get the realization matrices [SELF, SE, pron] and [SE, pron] as optimal output candidates in $\text{T}_{2.1}$, and [SE, pron] in $\text{T}_{2.2}$.

$T_{2.1}$: VP optimization
(XP/ThD reached – $x_{[\beta]}$ unchecked)

	Input: O_1/T_2	F_{pron}	F_{SE}	$\text{Pr.}\mathcal{A}_{ThD}$	F_{SELF}	$\text{Pr.}\mathcal{A}_{XP}$
⇒	O_{11}: [SELF, SE, pron]			**(!)		**
⇒	O_{12}: [SE, pron]			*	*(!)	*
	O_{13}: [pron]		*!		*	

$T_{2.2}$: VP optimization
(XP/ThD reached – $x_{[\beta]}$ unchecked)

	Input: O_2/T_2	F_{pron}	F_{SE}	$\text{Pr.}\mathcal{A}_{ThD}$	F_{SELF}	$\text{Pr.}\mathcal{A}_{XP}$
⇒	O_{21}: [SE, pron]			*	*	*
	O_{22}: [pron]		*!		*	

In the next phrase, the binder is merged into the derivation; hence, the PRINCIPLE $\mathcal{A}$-constraints apply vacuously and again the matrices [SELF, SE, pron] and [SE, pron] win (cf. $\text{T}_{2.1.1}/\text{T}_{2.1.2/2.2.1}$). As a result, MAB determines that x is realized as SELF anaphor if the optimal candidate is O_{111} and as SE anaphor otherwise (cf. $\text{O}_{121/211}$). This prediction is again correct.

(52) c. $[_{\text{vP}}\ \text{Max}_{[*\beta*]}\ [_{\text{VP}}\ x_{[\beta]}\ \cancel{[_{\text{vP}}\ \text{t}_x\ \text{singen}]}\ \text{t}_{h\ddot{o}rt}]\ \text{hört}]$

[37] Recall that *Phrase Balance* generally triggers movement of $x_{[\beta]}$ to the edge until its binder is merged into the derivation.

$T_{2.1.1}$: vP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ *applies vacuously)*

	Input: $O_{11}/T_{2.1}$	F_{pron}	F_{SE}	F_{SELF}
⇒	O_{111}: [SELF, SE, pron]			
	O_{112}: [SE, pron]			$*!$
	O_{113}: [pron]		$*!$	$*$

$T_{2.1.2/2.2.1}$: vP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ *applies vacuously)*

	Input: $O_{12}/T_{2.1}$ or $O_{21}/T_{2.2}$	F_{pron}	F_{SE}	F_{SELF}
⇒	O_{121}/O_{211}: [SE, pron]			$*$
	O_{122}/O_{212}: [pron]		$*!$	$*$

As far as example (46-c) is concerned (repeated in (53)), the first optimization step is illustrated in T_3.

(53) Max$_1$ schaut hinter sich$_1$/*sich selbst$_1$/*ihn$_1$.
  a. [$_{PP}$ $x_{[\beta]}$ hinter t_x]

In sentences like these, where binding takes place in the subject domain, only the SE anaphor is licit in German. As the following tableaux show, this is captured if PRINCIPLE $\mathcal{A}_{CD}$ is ranked below FAITH$_{SE}$ and above FAITH$_{SELF}$. Due to the fact that the PRINCIPLE $\mathcal{A}$-constraints are gradient, O_2 wins in the first competition, and since [SE, pron] remains optimal in the subsequent optimizations, MAB finally selects the SE anaphor as optimal realization for x.

T_3: PP optimization
(XP/ThD/CD reached – $x_{[\beta]}$ unchecked)

	Candidates	F_{pron}	F_{SE}	PR.$\mathcal{A}_{CD}$	PR.$\mathcal{A}_{ThD}$	F_{SELF}	PR.$\mathcal{A}_{XP}$
	O_1: [SELF, SE, pron]			$**!$	$**$		$**$
⇒	O_2: [SE, pron]			$*$	$*$	$*$	$*$
	O_3: [pron]		$*!$			$*$	

(54) b. $[_{\text{VP}}\ x_{[\beta]}\ [_{\text{PP}}\ \text{t}_x$ hinter $\cancel{\text{t}_x}]$ schaut]

$T_{3.1}$: VP optimization
(XP/ThD/CD reached – $x_{[\beta]}$ unchecked)

Input: O_2/T_3	F_{pron}	F_{SE}	PR.$\mathcal{A}_{CD}$	PR.$\mathcal{A}_{ThD}$	F_{SELF}	PR.$\mathcal{A}_{XP}$
⇒ O_{21}: [SE, pron]			$*$	$*$	$*$	$*$
O_{22}: [pron]		$*!$			$*$	

(55) c. $[_{\text{vP}}$ Max$_{[*\beta*]}$ $[_{\text{VP}}\ x_{[\beta]}\ \cancel{[_{\text{PP}}\ \text{t}_x{}'\ \text{hinter}\ \text{t}_x]}\ \text{t}_{schaut}]$ schaut]

$T_{3.1.1}$: vP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ applies vacuously)

Input: $O_{21}/T_{3.1}$	F_{pron}	F_{SE}	F_{SELF}
⇒ O_{211}: [SE, pron]			$*$
O_{212}: [pron]		$*!$	$*$

The analysis of example (46-d) (repeated in (56)) is illustrated in the tableaux T_4-$T_{4.2}$.

(56) Max$_1$ weiß, dass Maria ihn$_1$/*sich$_1$/*sich selbst$_1$ mag.

a. $[_{\text{VP}}\ x_{[\beta]}\ [_{\text{V}'}\ \text{t}_x$ mag]]

When the first optimization procedure takes place, only PRINCIPLE $\mathcal{A}_{XP}$ and the FAITH-constraints apply non-vacuously; and since the former is tied with FAITH$_{SELF}$, both O_1 and O_2 turn out to be optimal in this competition (cf. T_4).

T_4: VP optimization
(XP reached – $x_{[\beta]}$ unchecked)

Candidates	F_{pron}	F_{SE}	F_{SELF}	PR.$\mathcal{A}_{XP}$
⇒ O_1: [SELF, SE, pron]				$**(!)$
⇒ O_2: [SE, pron]			$*(!)$	$*$
O_3: [pron]		$*!$	$*$	

The next phrase that is completed is vP. $x_{[\beta]}$ is still free, but since a subject (*Maria*) enters the derivation, the defining criteria for all domains (θ-, Case, subject, finite and indicative domain) are met at this stage, and therefore *all* PRINCIPLE $\mathcal{A}$ constraints apply non-vacuously.

On the assumption that PRINCIPLE $\mathcal{A}_{ID}$, PRINCIPLE $\mathcal{A}_{FD}$, and PRINCIPLE $\mathcal{A}_{SD}$ (in a word, PRINCIPLE $\mathcal{A}_{ID/FD/SD}$) are ranked above FAITH$_{SE}$, only the candidates with the maximally reduced matrix [pron] win in $T_{4.1}$ and $T_{4.2}$.[38]

(57) b. $[_{\text{vP}}\ x_{[\beta]}$ Maria $[_{\text{VP}}\ \text{t}_x{}'\ [_{\text{V}'}\ \cancel{\text{t}_x}\ \text{t}_{mag}]]$ mag$]$

$T_{4.1}$: vP optimization
(XP/ThD/CD/SD/FD/ID reached – $x_{[\beta]}$ unchecked)

Input: O_1/T_4	F_{pron}	PR.$\mathcal{A}_{ID/FD/SD}$	F_{SE}	PR.$\mathcal{A}_{CD}$	PR.$\mathcal{A}_{ThD}$	F_{SELF}	PR.$\mathcal{A}_{XP}$
O_{11}: [S, S, pr]		*!*		**	**		**
O_{12}: [SE, pr]		*!		*	*	*	*
⇒ O_{13}: [pron]			*			*	

$T_{4.2}$: vP optimization
(XP/ThD/CD/SD/FD/ID reached – $x_{[\beta]}$ unchecked)

Input: O_2/T_4	F_{pron}	PR.$\mathcal{A}_{ID/FD/SD}$	F_{SE}	PR.$\mathcal{A}_{CD}$	PR.$\mathcal{A}_{ThD}$	F_{SELF}	PR.$\mathcal{A}_{XP}$
O_{21}: [S, pr]		*!		*	*	*	*
⇒ O_{22}: [pron]			*			*	

As a result, x will have to be realized as a pronoun in the end – since the realization matrix cannot be further reduced, [pron] remains optimal in the following optimizations until $x_{[\beta]}$ is checked, and the pronominal form must be selected.

[38] For reasons of space, the candidates are abbreviated in some of the subsequent tableaux, and PR.$\mathcal{A}_{ID}$, PR.$\mathcal{A}_{FD}$, and PR.$\mathcal{A}_{SD}$ are represented in one column since they behave alike in these examples and are adjacent on the constraint hierarchy.

5.4 Derivational Binding in English

Let us now turn to English. One particularity we find in English is that English does not have simple anaphors. As a consequence, we can find examples in which both the complex anaphor and the pronoun are licit, as (58-a) shows; this particular type of optionality cannot be found in languages that exhibit a three-way contrast, like German, Dutch, or Italian, where optionality can only arise between SELF and SE anaphors, or SE anaphors and pronouns ((58-b), (58-c), and (58-d) serve as an illustration; cf. also section 5.3, 5.5, and 5.6, respectively).

(58) a. Max_1 glanced behind himself_1/him_1.

b. Max_1 hasst sich selbst_1/sich_1/*ihn_1.
Max hates himself/SE/him
'Max_1 hates himself_1.'

c. Max_1 hoorde zichzelf_1/zich_1/*hem_1 zingen.
Max heard himself/SE/him sing
'Max_1 heard himself_1 sing.'

d. Max_1 ha dato un'occhiata dietro di sé_1/*dietro
Max has given a glance behind SE/behind
se stesso_1/?dietro di lui_1.
himself/behind him
'Max_1 glanced behind him_1/himself_1.'

This restriction is in fact predicted by the present theory, because according to this system only two 'adjacent' candidates can win at the same time. This can be derived as follows: In a single optimization process, only tied constraints can yield two optional candidates (identical constraint profiles cannot arise). However, ties must always involve one Faith-constraint and one Principle $\mathcal{A}$-constraint, because within their group the constraints are universally ordered in dominance relations. Furthermore, the gradience of the Principle $\mathcal{A}$-constraints has the effect that the difference between non-adjacent candidates amounts to "two stars", whereas adjacent candidates differ from each other only with respect to "one star". As a result, depending on whether Faith_{SE} or Faith_{SELF} is involved in the tie, only the matrix pairs [SELF, SE, pron]/[SE, pron] or [SE, pron]/[pron] can win at the same time.[39]

$T_{5.1}$: Optionality – possibility 1

	Candidates	F_{SE}	F_{SELF}	PR.$\mathcal{A}_{XD}$
⇒	O_1: [SELF, SE, pron]			∗∗(!)
⇒	O_2: [SE, pron]		∗(!)	∗
	O_3: [pron]	∗!	∗	

$T_{5.2}$: Optionality – possibility 2

	Candidates	F_{SE}	PR.$\mathcal{A}_{XD}$	F_{SELF}
	O_1: [SELF, SE, pron]		∗∗!	
⇒	O_2: [SE, pron]		∗(!)	∗
⇒	O_3: [pron]	∗(!)		∗

What remains to be investigated is whether optionality between O_1 and O_3 could arise as outcome of two different continuations when [SELF, SE, pron] and [SE, pron] have both been optimal earlier in the derivation (as in $T_{5.1}$). This presupposes that the derivation which is based on [SELF, SE, pron] as input does not reduce the matrix any further until binding takes place, whereas reduction would have to take place in the parallel derivation based on the input [SE, pron]. However, this is not possible, because the two derivations only differ with respect to the input and the resulting candidates; otherwise they are the same. Hence the following conclusion can be drawn: If [SE, pron] is reduced to [pron] at some stage of the derivation, a PRINCIPLE $\mathcal{A}$-constraint that is higher ranked than FAITH$_{SE}$ must have applied non-vacuously. However, this means that this must also be true for the parallel derivation originally based on the input [SELF, SE, pron], and hence the matrix would have to be reduced here as well.

The question therefore arises as to how we can account for languages like English. One possibility would be to assume that in this case the realization matrix lacks the SE form from the beginning. However, if we assume that the matrix does not yet contain the language-specific forms but rather some

[39]Since FAITH$_{pron}$ is not violated by either of the three candidates, it does not play a role here.

universal features that correspond to the SE, SELF, and pronominal form in a more abstract sense, the realization matrices in English would contain a SE form. If [SE, pron] is predicted to be optimal, we then have the following situation: According to this specification and MAB, the ideal realization form would be a simple anaphor; however, since there is no lexical item in English that fits this description, the most anaphoric realization must be chosen that is (i) available in English and (ii) compatible with the optimal matrix. Hence, the pronominal form would have to be selected in English, because the available forms comprise the SELF anaphor and the pronoun, but only the latter is compatible with the matrix [SE, pron]. This means that MAB can only select the most anaphoric form that is *available* in a language. Hence, the selection procedure is reminiscent of principles like the *Subset Principle* as we know it from Distributed Morphology.[40]

Consider now the following English examples. As to their binding behaviour, the first sentence is again an example where binding takes place in the minimal θ-domain; in (59-b), the antecedent enters the derivation when the minimal Case domain is reached; in (59-c), the binding relation is established in the minimal subject domain, and in (59-d), the finite and indicative domain have been reached when binding takes place.

(59) *English:*

a. Max$_1$ hates himself$_1$/*him$_1$.
b. Max$_1$ heard himself$_1$/*him$_1$ sing.
c. Max$_1$ glanced behind himself$_1$/him$_1$.
d. Max$_1$ knows that Mary likes him$_1$/*himself$_1$.

[40] *Subset Principle*:

The phonological exponent of a Vocabulary item is inserted into a morpheme in the terminal string if the item matches all or a subset of the grammatical features specified in the terminal morpheme. Insertion does not take place if the Vocabulary item contains features not present in the morpheme. Where several Vocabulary items meet the conditions for insertion, the item matching the greatest number of features specified in the terminal morpheme must be chosen. (Halle (2000:128))

Starting with the first sentence, the first derivation step yields the structure in (60-a).

(60) Max$_1$ hates himself$_1$/*him$_1$.

a. [$_{\text{VP}}$ $x_{[\beta]}$ hates t$_x$]

If it is assumed that, in contrast to German, FAITH$_{SELF}$ is higher ranked than PRINCIPLE $\mathcal{A}_{XP}$, only O$_1$ is optimal in the first competition (cf. T$_6$), because apart from XP no other domain relevant for binding is reached at this stage.

T$_6$: VP optimization
(XP reached – $x_{[\beta]}$ unchecked)

	Candidates	F$_{pron}$	F$_{SE}$	F$_{SELF}$	PR.$\mathcal{A}_{XP}$
⇒	O$_1$: [SELF, SE, pron]				∗∗
	O$_2$: [SE, pron]			∗!	∗
	O$_3$: [pron]		∗!	∗	

(61) b. [$_{\text{vP}}$ Max$_{[*\beta*]}$ hates [$_{\text{VP}}$ $x_{[\beta]}$ t$_{hates}$ $\cancel{\text{t}_x}$]]

In the next phrase, $x_{[\beta]}$ is already bound, hence the PRINCIPLE $\mathcal{A}_{XD}$ constraints apply vacuously at this stage, and the matrix [SELF, SE, pron] remains optimal (cf. T$_{6.1}$). Thus, MAB selects the complex anaphor as optimal realization, which is the correct prediction.

T$_{6.1}$: vP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ *applies vacuously)*

	Input: O$_1$/T$_6$	F$_{pron}$	F$_{SE}$	F$_{SELF}$
⇒	O$_{11}$: [SELF, SE, pron]			
	O$_{12}$: [SE, pron]			∗!
	O$_{13}$: [pron]		∗!	∗

In example (59-b) (repeated in (62)), XP and the θ-domain have been reached when the first optimization takes place.

(62) Max$_1$ heard himself$_1$/*him$_1$ sing.

a. $[_{\text{vP}}\ x_{[\beta]}$ sing]

Hence, both PRINCIPLE $\mathcal{A}_{XP}$ and PRINCIPLE $\mathcal{A}_{ThD}$ apply non-vacuously at this stage of the derivation; and since only the complex anaphor should win the competition, both constraints must be ranked below the FAITH-constraints, as T_7 illustrates.

T_7: vP optimization
(XP/ThD reached – $x_{[\beta]}$ unchecked)

	Candidates	F_{pron}	F_{SE}	F_{SELF}	PR.$\mathcal{A}_{ThD}$	PR.$\mathcal{A}_{XP}$
⇒	O_1: [SELF, SE, pron]				∗∗	∗∗
	O_2: [SE, pron]			∗!	∗	∗
	O_3: [pron]		∗!	∗		

When VP is completed, $x_{[\beta]}$ is still free, and since its θ-role assigner is still accessible, the accessible domain can still be classified as $x_{[\beta]}$'s θ-domain. Hence, the same constraints apply as in the previous competition, and as a result, the matrix [SELF, SE, pron] remains optimal (cf. $T_{7.1}$).

(63) b. $[_{\text{VP}}\ x_{[\beta]}$ heard $[_{\text{vP}}\ t_x$ sing]]

$T_{7.1}$: VP optimization
(XP/ThD reached – $x_{[\beta]}$ unchecked)

	Input: O_1/T_7	F_{pron}	F_{SE}	F_{SELF}	PR.$\mathcal{A}_{ThD}$	PR.$\mathcal{A}_{XP}$
⇒	O_{11}: [SELF, SE, pron]				∗∗	∗∗
	O_{12}: [SE, pron]			∗!	∗	∗
	O_{13}: [pron]		∗!	∗		

In the next phrase, the binder enters the derivation, and thus only the FAITH-constraints are relevant in the competition illustrated in $T_{7.1.1}$. Consequently, the first candidate wins again, and MAB correctly predicts the complex anaphor to be the optimal realization.

(64) c. $[_{\text{vP}}$ Max$_{[*\beta*]}$ heard $[_{\text{VP}}\ x_{[\beta]}\ t_{heard}\ \cancel{[_{\text{vP}}\ t_x \text{ sing}]}]]$

$T_{7.1.1}$: vP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ *applies vacuously)*

	Input: $O_{11}/T_{7.1}$	F_{pron}	F_{SE}	F_{SELF}
⇒	O_{111}: [SELF, SE, pron]			
	O_{112}: [SE, pron]			∗!
	O_{113}: [pron]		∗!	∗

Let us now turn to (59-c) (repeated in (65)), where optionality between the complex anaphor and the pronominal form arises

(65) Max$_1$ glanced behind himself$_1$/him$_1$.
  a. [$_{\text{PP}}$ $x_{[\beta]}$ behind t_x]

When the prepositional phrase is completed, XP, the θ-domain *and* the Case domain are reached, because the accessible domain does not only contain x's θ-role assigner but also its Case marker (= P). Hence, PRINCIPLE $\mathcal{A}_{XP}$, PRINCIPLE $\mathcal{A}_{ThD}$, and PRINCIPLE $\mathcal{A}_{CD}$ apply non-vacuously; and if we assume that the latter is tied with FAITH$_{SELF}$, both O_1 and O_2 win at this stage of the derivation (cf. T_8) – and this is crucial in order to get the desired optionality in the end.

T_8: PP optimization
(XP/ThD/CD reached – $x_{[\beta]}$ unchecked)

	Candidates	F_{pron}	F_{SE}	PR.$\mathcal{A}_{CD}$	F_{SELF}	PR.$\mathcal{A}_{ThD}$	PR.$\mathcal{A}_{XP}$
⇒	O_1: [SELF, SE, pron]			∗∗(!)		∗∗	∗∗
⇒	O_2: [SE, pron]			∗	∗(!)	∗	∗
	O_3: [pron]		∗!		∗		

As a consequence, there are two competitions when the next optimization takes place. At this stage, $x_{[\beta]}$ is still free and *behind* is still accessible, so the same constraints are relevant as before. As a result, the first two candidates win again in the competition based on the input [SELF, SE, pron] (cf. $T_{8.1}$), whereas in $T_{8.2}$, which represents the second competition, [SE, pron] is predicted to be optimal.

(66) b. [$_{\mathrm{VP}}$ $x_{[\beta]}$ glanced [$_{\mathrm{PP}}$ $t_x{}'$ behind $\cancel{t_x}$]]

$T_{8.1}$: VP optimization
(XP/ThD/CD reached – $x_{[\beta]}$ unchecked)

Input: O_1/T_8	F_{pron}	F_{SE}	PR.$\mathcal{A}_{CD}$	F_{SELF}	PR.$\mathcal{A}_{ThD}$	PR.$\mathcal{A}_{XP}$
⇒ O_{11}: [SELF, SE, pron]			**(!)		**	**
⇒ O_{12}: [SE, pron]			*	*(!)	*	*
O_{13}: [pron]		*!		*		

$T_{8.2}$: VP optimization
(XP/ThD/CD reached – $x_{[\beta]}$ unchecked)

Input: O_2/T_8	F_{pron}	F_{SE}	PR.$\mathcal{A}_{CD}$	F_{SELF}	PR.$\mathcal{A}_{ThD}$	PR.$\mathcal{A}_{XP}$
⇒ O_{21}: [SE, pron]			*	*(!)	*	*
O_{22}: [pron]		*!		*		

In the next phrase, the binder is merged into the derivation, and so the FAITH-constraints determine the outcome of the next optimization procedure. In the competition based on the input [SELF, SE, pron] the maximally specified realization matrix wins again, which means that MAB correctly predicts the complex anaphor to be the optimal realization. In the competition based on the input [SE, pron], a further reduction is also excluded, because this would only be possible if another domain had been reached (– but in this case the constraint would also have reduced the winner in $T_{8.1.1}$). Hence, [SE, pron] is the optimal candidate in $T_{8.1.2/8.2.1}$, and since English does not have a simple anaphor, x is here correctly predicted to be realized pronominally.

(67) c. [$_{\mathrm{vP}}$ $\mathrm{Max}_{[*\beta*]}$ glanced [$_{\mathrm{VP}}$ $x_{[\beta]}$ $t_{glanced}$ ~~[$_{\mathrm{PP}}$ $t_x{}'$ behind t_x]~~]]

$T_{8.1.1}$: vP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ applies vacuously)

	Input: $O_{11}/T_{8.1}$	F_{pron}	F_{SE}	F_{SELF}
⇒	O_{111}: [SELF, SE, pron]			
	O_{112}: [SE, pron]			∗!
	O_{113}: [pron]		∗!	∗

$T_{8.1.2/8.2.1}$: vP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ *applies vacuously)*

	Input: $O_{12}/T_{8.1}$ or $O_{21}/T_{8.2}$	F_{pron}	F_{SE}	F_{SELF}
⇒	$O_{121/211}$: [SE, pron]			∗
	$O_{122/212}$: [pron]		∗!	∗

Let us finally come to the analysis of sentence (59-d), repeated in (68). The first optimization procedure is illustrated in T_9: No other domain than XP is reached and $x_{[\beta]}$ remains unbound, hence only the FAITH-constraints and PRINCIPLE $\mathcal{A}_{XP}$ apply non-vacuously, and as a result O_1 is the winner of the competition.

(68) Max$_1$ knows that Mary likes him$_1$/*himself$_1$.
  a. [$_{\text{VP}}$ $x_{[\beta]}$ likes t_x]

T_9: VP optimization
(XP reached – $x_{[\beta]}$ unchecked)

	Candidates	F_{pron}	F_{SE}	F_{SELF}	PR.$\mathcal{A}_{XP}$
⇒	O_1: [SELF, SE, pron]				∗∗
	O_2: [SE, pron]			∗!	∗
	O_3: [pron]		∗!	∗	

Before the next phrase, vP, is completed, a subject enters the derivation, and since *likes* is also still accessible when optimization takes place, we can conclude that we have reached the θ-domain, Case domain, subject domain, finite domain, and the indicative domain at this stage.

(69) b. [$_{\text{vP}}$ $x_{[\beta]}$ Mary likes [$_{\text{VP}}$ t_x' t_{likes} $\cancel{t_x}$]]

In the analyses of the previous examples, we have already fixed the order of the first three constraints, so we know that they cannot be ranked above the FAITH-constraints. However, if we want the pronoun to be optimal in the end, [pron] must be the optimal realization matrix to which MAB applies.[41] Thus I assume that PRINCIPLE $\mathcal{A}_{SD}$, PRINCIPLE $\mathcal{A}_{FD}$, and PRINCIPLE $\mathcal{A}_{ID}$ are ranked above FAITH$_{SE}$.[42] On this assumption, the reduced realization matrix [pron] becomes optimal (cf. $T_{9.1}$) – and since this matrix cannot be reduced any further, [pron] also remains optimal in the following optimizations until $x_{[\beta]}$ is checked. At this stage, MAB will finally select the pronoun as optimal realization form, which is again the correct prediction.

$T_{9.1}$: vP optimization
(XP/ThD/CD/SD/FD/ID reached – $x_{[\beta]}$ unchecked)

Input: O_1/T_9	F_{pron}	PR.$\mathcal{A}_{ID/FD/SD}$	F_{SE}	F_{SELF}	PR.$\mathcal{A}_{CD}$	PR.$\mathcal{A}_{ThD}$	PR.$\mathcal{A}_{XP}$
O_{11}: [S, S, pr]		*!*			**	**	**
O_{12}: [SE, pr]		*!		*	*	*	*
⇒ O_{13}: [pron]			*	*			

5.5 Derivational Binding in Dutch

Consider now the Dutch data in (70). (In analogy to the previous sections, they represent again examples with the following binding behaviour: (a) binding within the minimal θ-domain; (b) binding within the minimal Case domain; (c) binding within the minimal subject domain; (d) binding within the minimal finite/indicative domain.)

(70) *Dutch:*

[41] As the previous example showed, the pronoun is also the optimal realization form if [SE, pron] wins in the end. However, if binding is so non-local that it takes place even outside the indicative domain, I assume that the pronominal realization is based on the optimal matrix [pron].

[42] At this point it might not yet be evident why all three constraints must be higher ranked than FAITH$_{SE}$; this issue will be addressed in more detail in section 5.8 and 5.9

a. Max$_1$ haat zichzelf$_1$/*zich$_1$/*hem$_1$.
Max hates himself/SE/him
'Max$_1$ hates himself$_1$.'

b. Max$_1$ hoorde zichzelf$_1$/zich$_1$/*hem$_1$ zingen.
Max heard himself/SE/him sing
'Max$_1$ heard himself$_1$ sing.'

c. Max$_1$ keek achter *zichzelf$_1$/zich$_1$/hem$_1$.
Max looked after himself/SE/him
'Max$_1$ glanced behind him$_1$/himself$_1$.'

d. Max$_1$ weet dat Mary *zichzelf$_1$/*zich$_1$/hem$_1$ leuk vindt.
Max knows that Mary himself/SE/him nice finds
'Max$_1$ knows that Mary likes him$_1$.'

As far as example (70-a) is concerned (repeated in (71)), it differs from German insofar as it only allows the complex anaphor as bound element. This is correctly predicted if PRINCIPLE $\mathcal{A}_{XP}$ is ranked below FAITH$_{SELF}$. On this assumption, O$_1$ is the sole winner of the first competition (cf. T$_{10}$), and when the binder is merged into the derivation in the next phrase, [SELF, SE, pron] is predicted to be the optimal realization matrix (cf. T$_{10.1}$). Hence, MAB finally selects the SELF anaphor as optimal realization.

(71) Max$_1$ haat zichzelf$_1$/*zich$_1$/*hem$_1$.

a. [$_{\text{VP}}$ $x_{[\beta]}$ t$_x$ haat]

T$_{10}$: VP optimization
(XP reached – $x_{[\beta]}$ unchecked)

	Candidates	F$_{pron}$	F$_{SE}$	F$_{SELF}$	PR.$\mathcal{A}_{XP}$
⇒	O$_1$: [SELF, SE, pron]				∗∗
	O$_2$: [SE, pron]			∗!	∗
	O$_3$: [pron]		∗!	∗	

(72) b. [$_{\text{vP}}$ Max$_{[*\beta*]}$ [$_{\text{VP}}$ $x_{[\beta]}$ [$_{\text{V}'}$ $\cancel{t_x}$ t$_{haat}$]] haat]

$T_{10.1}$*: vP optimization*
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ *applies vacuously)*

	Input: O_1/T_{10}	F_{pron}	F_{SE}	F_{SELF}
⇒	O_{11}: [SELF, SE, pron]			
	O_{12}: [SE, pron]			$*!$
	O_{13}: [pron]		$*!$	$*$

In example (70-b) (repeated in (73)), both anaphors can function as bound elements. In order to derive this optionality, PRINCIPLE $\mathcal{A}_{ThD}$ must be tied with FAITH$_{SELF}$: As a result, both O_1 and O_2 win in the first competition (cf. T_{11}), because when optimization takes place not only an XP but also the θ-domain of x has been reached.

(73) Max$_1$ hoorde zichzelf$_1$/zich$_1$/*hem$_1$ zingen.

a. [$_{\text{vP}}$ $x_{[\beta]}$ zingen]

T_{11}*: vP optimization*
(XP/ThD reached – $x_{[\beta]}$ *unchecked)*

	Candidates	F_{pron}	F_{SE}	F_{SELF}	PR.$\mathcal{A}_{ThD}$	PR.$\mathcal{A}_{XP}$
⇒	O_1: [SELF, SE, pron]				$**(!)$	$**$
⇒	O_2: [SE, pron]			$*(!)$	$*$	$*$
	O_3: [pron]		$*!$	$*$		

When the next phrase is completed, no new domain relevant for binding has been reached, but x's θ-role assigner (*zingen*) is still accessible, hence, both PRINCIPLE $\mathcal{A}_{XD}$ and PRINCIPLE $\mathcal{A}_{ThD}$ apply again non-vacuously. In the competition based on the input [SELF, SE, pron], the first two candidates are therefore again predicted to be optimal (cf. $T_{11.1}$), and in the second competition, the matrix [SE, pron] wins (cf. $T_{11.2}$).

(74) b. [$_{\text{VP}}$ $x_{[\beta]}$ [$_{\text{vP}}$ t_x zingen] hoorde]

$T_{11.1}$*: VP optimization*
(XP/ThD reached – $x_{[\beta]}$ *unchecked)*

	Input: O_1/T_{11}	F_{pron}	F_{SE}	F_{SELF}	PR.$\mathcal{A}_{ThD}$	PR.$\mathcal{A}_{XP}$
⇒	O_{11}: [SELF, SE, pron]				∗∗(!)	∗∗
⇒	O_{12}: [SE, pron]			∗(!)	∗	∗
	O_{13}: [pron]		∗!	∗		

$T_{11.2}$: VP optimization
(XP/ThD reached – $x_{[\beta]}$ unchecked)

	Input: O_2/T_{11}	F_{pron}	F_{SE}	F_{SELF}	PR.$\mathcal{A}_{ThD}$	PR.$\mathcal{A}_{XP}$
⇒	O_{21}: [SE, pron]			∗	∗	∗
	O_{22}: [pron]		∗!	∗		

Now the binder enters the derivation, and so the FAITH-constraints alone determine the optimizations at the vP level. In $T_{11.1.1}$, the maximally specified matrix [SELF, SE, pron] wins, and according to $T_{11.1.2}$, [SE, pron] is optimal. Thus, MAB finally correctly predicts that either the SELF or the SE anaphor is the optimal realization of x.

(75) c. $[_{\text{vP}}$ Max$_{[*\beta*]}$ $[_{\text{VP}}$ $x_{[\beta]}$ $\cancel{[_{\text{VP}}\ t_x\ \text{zingen}]}$ $t_{hoorde}]$ hoorde]

$T_{11.1.1}$: vP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ applies vacuously)

	Input: $O_{11}/T_{11.1}$	F_{pron}	F_{SE}	F_{SELF}
⇒	O_{111}: [SELF, SE, pron]			
	O_{112}: [SE, pron]			∗!
	O_{113}: [pron]		∗!	∗

$T_{11.1.2/11.2.1}$: vP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ applies vacuously)

	Input: $O_{12}/T_{11.1}$ or $O_{21}/T_{11.2}$	F_{pron}	F_{SE}	F_{SELF}
⇒	O_{121}/O_{211}: [SE, pron]			∗
	O_{122}/O_{212}: [pron]		∗!	∗

(76) (repeated from (70-c)) is interesting insofar as it is the first example that exhibits optionality between the pronominal and the simple anaphoric form. (Neither did this occur in German nor in English, for obvious reasons.)[43]

(76) Max$_1$ keek achter zich$_1$/*zichzelf$_1$/hem$_1$.

  a. [$_{\text{PP}}$ $x_{[\beta]}$ achter t$_x$]

This type of optionality can be captured if Principle $\mathcal{A}_{CD}$ and Faith$_{SE}$ are tied: When the prepositional phrase is completed, the domains XP, ThD, and CD are reached, which means that in addition to Principle $\mathcal{A}_{XP}$ and Principle $\mathcal{A}_{ThD}$, Principle $\mathcal{A}_{CD}$ is now involved in the competition. On the assumption that the latter is tied with Faith$_{SE}$, optionality between O$_2$ and O$_3$ is predicted (cf. T$_{12}$).

T$_{12}$: PP optimization
(XP/ThD/CD reached – $x_{[\beta]}$ unchecked)

Candidates	F$_{pron}$	F$_{SE}$	Pr.$\mathcal{A}_{CD}$	F$_{SELF}$	Pr.$\mathcal{A}_{ThD}$	Pr.$\mathcal{A}_{XP}$
O$_1$: [SELF, SE, pron]			**!		**	**
⇒ O$_2$: [SE, pron]			*(!)	*	*	*
⇒ O$_3$: [pron]		*(!)		*		

As a result, there are two optimization procedures when the next phrase boundary, VP, is reached.

(77) b. [$_{\text{VP}}$ $x_{[\beta]}$ [$_{\text{PP}}$ t$_x$ achter $\cancel{\text{t}_x}$] keek]

The competition based on the matrix [SE, pron] yields again two optimal outputs (cf. T$_{12.1}$), whereas in the competition based on the input [pron] a further reduction is not possible and this matrix remains optimal (cf. T$_{12.2}$).

T$_{12.1}$: VP optimization
(XP/ThD/CD reached – $x_{[\beta]}$ unchecked)

Input: O$_2$/T$_{12}$	F$_{pron}$	F$_{SE}$	Pr.$\mathcal{A}_{CD}$	F$_{SELF}$	Pr.$\mathcal{A}_{ThD}$	Pr.$\mathcal{A}_{XP}$
⇒ O$_{21}$: [SE, pron]			*(!)	*	*	*
⇒ O$_{22}$: [pron]		*(!)		*		

[43] Recall that some native speakers prefer the weak pronoun instead of *hem* in (76).

$T_{12.2}$: *VP optimization*
(XP/ThD/CD reached – $x_{[\beta]}$ unchecked)

Input: O_3/T_{12}	F_{pron}	F_{SE}	$PR.\mathcal{A}_{CD}$	F_{SELF}	$PR.\mathcal{A}_{ThD}$	$PR.\mathcal{A}_{XP}$
⇒ O_{31}: [pron]		$*$		$*$		

In the next phrase, the binder is merged into the derivation, hence the FAITH-constraints predict that [SE, pron] is optimal in $T_{12.1.1}$, and [pron] wins in $T_{12.1.2/12.2.1}$.

(78) c. $[_{vP}$ Max$_{[*\beta*]}$ $[_{VP}$ $x_{[\beta]}$ ~~$[_{PP}$ $t_x{}'$ achter $t_x]$~~ $t_{keek}]$ keek$]$

According to MAB, the optimal choice is therefore the SE anaphor in the former derivation, and the pronoun in the latter.

$T_{12.1.1}$: *vP optimization*
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ *applies vacuously)*

Input: $O_{21}/T_{12.1}$	F_{pron}	F_{SE}	F_{SELF}
⇒ O_{211}: [SE, pron]			$*$
O_{212}: [pron]		$*!$	$*$

$T_{12.1.2/12.2.1}$: *vP optimization*
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ *applies vacuously)*

Input: $O_{22}/T_{12.1}$ or $O_{31}/T_{12.2}$	F_{pron}	F_{SE}	F_{SELF}
⇒ O_{221}/O_{311}: [pron]		$*$	$*$

(70-d) (repeated in (79)) patterns again like its German and English counterparts: In sentences in which binding takes place outside the Case domain, x must be realized as a pronoun, and this is captured by ranking PRINCIPLE $\mathcal{A}_{SD}$ (and hence also PRINCIPLE $\mathcal{A}_{FD}$ and PRINCIPLE $\mathcal{A}_{ID}$) above FAITH$_{SE}$ (cf. $T_{13.1}$).[44]

(79) Max$_1$ weet dat Mary hem$_1$/*zich$_1$/*zichzelf$_1$ leuk vindt.

a. $[_{VP}$ $x_{[\beta]}$ t_x leuk vindt$]$

[44] I treat the verbal predicate *leuk vindt* like a simple verb and ignore its inherent syntactic structure.

When the first optimization process takes place (cf. T_{13}), only PRINCIPLE $\mathcal{A}_{XP}$ and the FAITH-constraints apply non-vacuously, which means that O_1 serves as input for the next competition.

T_{13}: VP optimization
(XP reached – $x_{[\beta]}$ unchecked)

	Candidates	F_{pron}	F_{SE}	F_{SELF}	PR.$\mathcal{A}_{XP}$
⇒	O_1: [SELF, SE, pron]				$**$
	O_2: [SE, pron]			$*!$	$*$
	O_3: [pron]		$*!$	$*$	

When vP is completed, we reach at once all domains relevant for binding, which means that all PRINCIPLE $\mathcal{A}$-constraints are involved in the next competition. According to the ranking assumed above, [pron] is therefore predicted to be the optimal realization matrix (cf. $T_{13.1}$).

(80) b. $[_{\mathrm{vP}}\ x_{[\beta]}$ Mary $[_{\mathrm{VP}}\ \mathrm{t}_x{}'\ \cancel{\mathrm{t}_x}\ \mathrm{t}_{leuk\ vindt}]$ leuk vindt]

$T_{13.1}$: vP optimization
(XP/ThD/CD/SD/FD/ID reached – $x_{[\beta]}$ unchecked)

Input: O_1/T_{13}	F_{pron}	PR.$\mathcal{A}_{ID/FD/SD}$	F_{SE}	PR.$\mathcal{A}_{CD}$	F_{SELF}	PR.$\mathcal{A}_{ThD}$	PR.$\mathcal{A}_{XP}$
O_{11}: [S, S, pr]		$*!*$		$**$		$**$	$**$
O_{12}: [SE, pr]		$*!$		$*$	$*$	$*$	$*$
⇒ O_{13}: [pron]			$*$		$*$		

Since [pron] serves now as input for the next optimization procedure, it remains the only candidate, because the matrix cannot be further reduced. Hence, [pron] remains optimal in the following optimizations, and when $x_{[\beta]}$ is checked, MAB correctly predicts that x must be realized as a pronoun.

5.6 Derivational Binding in Italian

Last but not least, let us take a look at the corresponding Italian sentences. (Recall that in (81-a) the binding relation is established in the minimal θ-

domain, in (81-b) in the minimal Case domain, in (81-c) in the minimal subject domain, and in (81-d) in the minimal finite/indicative domain.)

(81) *Italian:*

a. Max$_1$ si$_1$ odia/ odia se stesso$_1$/ *lo$_1$ odia.
'Max$_1$ hates himself$_1$.'

b. Max$_1$ ha udito ?se stesso$_1$/ si$_1$ è udito/ *lo$_1$ ha udito cantare alla radio.
'Max$_1$ heard himself$_1$ sing on the radio.'

c. Max$_1$ ha dato un'occhiata dietro di sé$_1$/*dietro se stesso$_1$/?dietro di lui$_1$.
'Max$_1$ glanced behind him$_1$/himself$_1$.'

d. Max$_1$ sa che Maria lo$_1$ ama /*si$_1$ ama/ ama *se stesso$_1$.
'Max$_1$ knows that Mary likes him$_1$.'

As observed before, Italian patterns partly like German and partly like Dutch. In example (81-a) (repeated in (82)), where the binding relation is very local, Italian allows both types of anaphors, like its German counterpart (cf. (46-a)). This result is achieved if FAITH$_{SELF}$ and PRINCIPLE $\mathcal{A}_{XP}$ are tied; on this assumption both O$_1$ and O$_2$ win in the first optimization process (cf. T$_{14}$). Hence, there are two competitions after the completion of the next phrase, one based on the input [SELF, SE, pron] and the other one on the input [SE, pron]. Since at this stage the binder has already been merged into the derivation, the FAITH-constraints determine the outcome of the competitions, which means that no reduction of the matrices takes place, and therefore MAB selects the complex anaphor as optimal realization according to T$_{14.1}$ and the simple anaphor in the case of T$_{14.2}$.

(82) Max$_1$ si$_1$ odia/ odia se stesso$_1$/ *lo$_1$ odia.

a. [$_{\text{VP}}$ $x_{[\beta]}$ odia t$_x$]

T$_{14}$: VP optimization
(XP reached – $x_{[\beta]}$ unchecked)

	Candidates	F_{pron}	F_{SE}	F_{SELF}	$\textsc{Pr}.\mathcal{A}_{XP}$
⇒	O_1: [SELF, SE, pron]				∗∗(!)
⇒	O_2: [SE, pron]			∗(!)	∗
	O_3: [pron]		∗!	∗	

(83) b. $[_{\mathrm{vP}}$ Max$_{[*\beta*]}$ odia $[_{\mathrm{VP}}$ $x_{[\beta]}$ t$_{odia}$ $\cancel{\mathrm{t}_x}]]$

$T_{14.1}$: vP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ *applies vacuously)*

	Input: O_1/T_{14}	F_{pron}	F_{SE}	F_{SELF}
⇒	O_{11}: [SELF, SE, pron]			
	O_{12}: [SE, pron]			∗!
	O_{13}: [pron]		∗!	∗

$T_{14.2}$: vP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ *applies vacuously)*

	Input: O_2/T_{14}	F_{pron}	F_{SE}	F_{SELF}
⇒	O_{21}: [SE, pron]			∗
	O_{22}: [pron]		∗!	∗

(84) (repeated from (81-b)) is an Italian ECM-construction. As in German and Dutch, both the SELF and the SE anaphor is licit in this context, which is correctly predicted if FAITH$_{SELF}$ and PRINCIPLE $\mathcal{A}_{ThD}$ are tied.[45] As a result, both O_1 and O_2 are optimal when the embedded vP is optimized, which corresponds to x's θ-domain (cf. T_{15}).

(84) Max$_1$ ha udito ?se stesso$_1$/ si$_1$ è udito/ *lo$_1$ ha udito cantare (alla radio).

a. $[_{\mathrm{vP}}$ $x_{[\beta]}$ cantare$]$

[45] In fact, one informant of mine preferred the complex anaphor and ruled out the simple anaphor in this example. This is unexpected against the background that *si* is licit in (81-a), where the binding relation is even more local.

T_{15}: vP optimization
(XP/ThD reached – $x_{[\beta]}$ unchecked)

Candidates	F_{pron}	F_{SE}	$\textsc{Pr}.\mathcal{A}_{ThD}$	F_{SELF}	$\textsc{Pr}.\mathcal{A}_{XP}$
⇒ O_1: [SELF, SE, pron]			$**(!)$		$**$
⇒ O_2: [SE, pron]			$*$	$*(!)$	$*$
O_3: [pron]		$*!$		$*$	

When VP is completed, all parts of the derivation are still accessible, but no further domain is reached; hence, the same constraints apply non-vacuously as before, which yields again two optimal outputs in $T_{15.1}$ and [SE, pron] as optimal matrix in $T_{15.2}$.

(85) b. $[_{\text{VP}}\ x_{[\beta]}$ udito $[_{\text{vP}}\ t_x$ cantare]]

$T_{15.1}$: VP optimization
(XP/ThD reached – $x_{[\beta]}$ unchecked)

Input: O_1/T_{15}	F_{pron}	F_{SE}	$\textsc{Pr}.\mathcal{A}_{ThD}$	F_{SELF}	$\textsc{Pr}.\mathcal{A}_{XP}$
⇒ O_{11}: [SELF, SE, pron]			$**(!)$		$**$
⇒ O_{12}: [SE, pron]			$*$	$*(!)$	$*$
O_{13}: [pron]		$*!$		$*$	

$T_{15.2}$: VP optimization
(XP/ThD reached – $x_{[\beta]}$ unchecked)

Input: O_2/T_{15}	F_{pron}	F_{SE}	$\textsc{Pr}.\mathcal{A}_{ThD}$	F_{SELF}	$\textsc{Pr}.\mathcal{A}_{XP}$
⇒ O_{21}: [SE, pron]			$*$	$*$	$*$
O_{22}: [pron]		$*!$		$*$	

In the next phrase, the binder enters the derivation. Thus, the FAITH-constraints determine the competitions at this stage and predict the matrices [SELF, SE, pron] and [SE, pron] to be optimal (cf. $T_{15.1.1}$ and $T_{15.1.2/15.2.1}$ respectively). As a result, MAB selects the two anaphors as optimal realizations.

(86) c. $[_{\text{vP}}$ Max$_{[*\beta*]}$ udito $[_{\text{VP}}\ x_{[\beta]}\ t_{udito}$ $\cancel{[_{\text{vP}}\ t_x \text{ cantare}]}$]]

$T_{15.1.1}$: vP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ applies vacuously)

	Input: $O_{11}/T_{15.1}$	F_{pron}	F_{SE}	F_{SELF}
⇒	O_{111}: [SELF, SE, pron]			
	O_{112}: [SE, pron]			∗!
	O_{113}: [pron]		∗!	∗

$T_{15.1.2/15.2.1}$: vP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ applies vacuously)

	Input: $O_{12}/T_{15.1}$ or $O_{21}/T_{15.2}$	F_{pron}	F_{SE}	F_{SELF}
⇒	O_{121}/O_{211}: [SE, pron]			∗
	O_{122}/O_{212}: [pron]		∗!	∗

In the following example (repeated from (81-c)), Italian patterns like Dutch since it only excludes the complex anaphor in sentences like these. Hence, as has been shown for Dutch in T_{12}, FAITH$_{SE}$ must be tied with PRINCIPLE $\mathcal{A}_{CD}$. On this assumption, O_2 and O_3 are both optimal in the competition illustrated in T_{16}.

(87) Max$_1$ ha dato un'occhiata dietro di sé$_1$/*dietro se stesso$_1$/?dietro di lui$_1$.

a. [$_{\text{PP}}$ $x_{[\beta]}$ dietro di t_x]

T_{16}: PP optimization
(XP/ThD/CD reached – $x_{[\beta]}$ unchecked)

Candidates	F_{pron}	F_{SE}	PR.$\mathcal{A}_{CD}$	PR.$\mathcal{A}_{ThD}$	F_{SELF}	PR.$\mathcal{A}_{XP}$
O_1: [SELF, SE, pron]			∗∗!	∗∗		∗∗
⇒ O_2: [SE, pron]			∗(!)	∗	∗	∗
⇒ O_3: [pron]		∗(!)			∗	

The optimization procedure after the completion of VP yields the same results, since no further domain relevant for binding is reached (cf. $T_{16.1}$ and $T_{16.2}$).

(88) b. $[_{VP}\ x_{[\beta]}$ un'occhiata $[_{V'}$ dato $[_{PP}\ t_x{}'$ dietro di $\cancel{t_x}]]]$

$T_{16.1}$: VP optimization
(XP/ThD/CD reached – $x_{[\beta]}$ unchecked)

Input: O_2/T_{16}	F_{pron}	F_{SE}	$\textsc{Pr}.\mathcal{A}_{CD}$	$\textsc{Pr}.\mathcal{A}_{ThD}$	F_{SELF}	$\textsc{Pr}.\mathcal{A}_{XP}$
⇒ O_{21}: [SE, pron]			∗(!)	∗	∗	∗
⇒ O_{22}: [pron]		∗(!)			∗	

$T_{16.2}$: VP optimization
(XP/ThD/CD reached – $x_{[\beta]}$ unchecked)

Input: O_3/T_{16}	F_{pron}	F_{SE}	$\textsc{Pr}.\mathcal{A}_{CD}$	$\textsc{Pr}.\mathcal{A}_{ThD}$	F_{SELF}	$\textsc{Pr}.\mathcal{A}_{XP}$
⇒ O_{31}: [pron]		∗			∗	

In the next phrase, the binder is merged into the derivation, hence only the FAITH-constraints apply non-vacuously in the subsequent competitions, which means that a further reduction of the input matrices is barred. In $T_{16.1.1}$, this means that [SE, pron] wins and MAB selects the SE anaphor as optimal realization; in $T_{16.1.2/16.2.1}$ it is self-evident that [pron] is optimal because there are no competing candidates, and thus a prononimal realization is also licit.

(89) c. $[_{vP}$ $\text{Max}_{[*\beta*]}$ dato $[_{VP}\ x_{[\beta]}$ un'occhiata t_{dato} $[_{PP}\ t_x{}'$ dietro di $\cancel{t_x}]]$

$T_{16.1.1}$: vP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ *applies vacuously)*

Input: $O_{21}/T_{16.1}$	F_{pron}	F_{SE}	F_{SELF}
⇒ O_{211}: [SE, pron]			∗
O_{212}: [pron]		∗!	∗

$T_{16.1.2/16.2.1}$: vP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ *applies vacuously)*

Input: $O_{22}/T_{16.1}$ or $O_{31}/T_{16.2}$	F_{pron}	F_{SE}	F_{SELF}
⇒ O_{221}/O_{311}: [pron]		∗	∗

The last Italian example (repeated from (81-d)), illustrates binding into a finite embedded clause. Like German, English, and Dutch, Italian exhibits pronominal binding in this case, which is correctly predicted if PRINCIPLE $\mathcal{A}_{SD}$ (and therefore also PRINCIPLE $\mathcal{A}_{FD}$ and PRINCIPLE $\mathcal{A}_{ID}$) are ranked above FAITH$_{SE}$. When the embedded VP is optimized, these constraints are not involved yet, and O_1 and O_2 are predicted to be optimal (cf. T_{17}). However, when the next phrase (= vP) is completed the accessible domain corresponds to the subject, finite, and indicative domain, and all PRINCIPLE $\mathcal{A}$-constraints apply non-vacuously. As a result, [pron] is the winner of all subsequent optimizations (cf., for example, $T_{17.1}$ and $T_{17.2}$) and x will finally have to be realized as pronoun.

(90) Max$_1$ sa che Maria lo$_1$ ama /*si$_1$ ama/ ama *se stesso$_1$.

a. $[_{\text{VP}}\ x_{[\beta]}$ ama $t_x]$

T_{17}: VP optimization
(XP reached – $x_{[\beta]}$ unchecked)

	Candidates	F_{pron}	F_{SE}	F_{SELF}	PR.$\mathcal{A}_{XP}$
⇒	O_1: [SELF, SE, pron]				∗∗(!)
⇒	O_2: [SE, pron]			∗(!)	∗
	O_3: [pron]		∗!	∗	

(91) b. $[_{\text{vP}}\ x_{[\beta]}$ Maria ama $[_{\text{VP}}\ t_x'\ t_{ama}\ \cancel{t_x}]]$

$T_{17.1}$: vP optimization
(XP/ThD/CD/SD/FD/ID reached – $x_{[\beta]}$ unchecked)

Input: O_1/T_{17}	F_{pron}	PR.$\mathcal{A}_{ID/FD/SD}$	F_{SE}	PR.$\mathcal{A}_{CD}$	PR.$\mathcal{A}_{ThD}$	F_{SELF}	PR.$\mathcal{A}_{XP}$
O_{11}: [S, S, pr]		∗!∗		∗∗	∗∗		∗∗
O_{12}: [SE, pr]		∗!		∗	∗	∗	∗
⇒ O_{13}: [pron]			∗			∗	

$T_{17.2}$: vP optimization
(XP/ThD/CD/SD/FD/ID reached – $x_{[\beta]}$ unchecked)

Input: O_2/T_{17}	F_{pron}	PR.$\mathcal{A}_{ID/FD/SD}$	F_{SE}	PR.$\mathcal{A}_{CD}$	PR.$\mathcal{A}_{ThD}$	F_{SELF}	PR.$\mathcal{A}_{XP}$
O_{21}: [S, pr]		*!		*	*	*	*
⇒ O_{22}: [pron]			*			*	

5.7 Summary: Crosslinguistic Variation I

In the previous sections, the four languages German, English, Dutch, and Italian have been analysed in detail. This section provides an overview of their main differences and common patterns.

In a nutshell, the following observations could be made. If binding takes place within the θ-domain (as in sentences of the type *Max*$_1$ *hates* x_1), the bound element can be realized as SELF anaphor in all languages. This is correctly predicted if FAITH$_{SELF}$ is not ranked below PRINCIPLE $\mathcal{A}_{XP}$. However, some languages allow in addition the SE anaphor (cf. German and Italian), while others do not (cf. Dutch). In order to account for the former type of language, the ranking FAITH$_{SELF}$ $\circ$ PRINCIPLE $\mathcal{A}_{XP}$ must be assumed, whereas in languages like Dutch, FAITH$_{SELF}$ must be ranked above PRINCIPLE $\mathcal{A}_{XP}$.

If the binding relation is slightly less local and occurs within the Case domain, the crucial constraint which determines the outcome of the competition is PRINCIPLE $\mathcal{A}_{ThD}$. If it is tied with FAITH$_{SELF}$, both types of anaphors are licit in this context (cf. ECM-constructions in German, Dutch, and Italian). In languages like English where only the complex anaphor is licit (as in *Max*$_1$ *heard* x_1 *sing*), FAITH$_{SELF}$ must be higher ranked than PRINCIPLE $\mathcal{A}_{ThD}$.

In sentences like *Max*$_1$ *glanced behind* x_1, x is bound in its subject domain; hence, the ranking of PRINCIPLE $\mathcal{A}_{CD}$ is decisive. In German, where only the SE anaphor is licit, it must be ranked below FAITH$_{SE}$ and above FAITH$_{SELF}$. If FAITH$_{SE}$ is tied with PRINCIPLE $\mathcal{A}_{CD}$, both anaphors are predicted to be optimal (cf. English, on the assumption that the pronominal realization in examples like these is based on the optimal matrix [SE, pron]). In languages that pattern like Dutch and Italian in allowing a SE anaphor or a pronoun, PRINCIPLE $\mathcal{A}_{CD}$ must be tied with FAITH$_{SE}$.

Finally, none of the languages discussed so far exhibited long-distance anaphora; this behaviour is captured if the three constraints PRINCIPLE $\mathcal{A}_{SD}$, PRINCIPLE $\mathcal{A}_{FD}$, and PRINCIPLE $\mathcal{A}_{ID}$ are ranked above FAITH$_{SE}$. All in all, this yields the following constraint orders for German, English, Dutch, and Italian. Once more, it can be seen immediately that they only differ with respect to different interactions of the two underlying universal constraint subhierarchies, which provide a general frame for possible rankings.

(92) *German ranking:*
FAITH$_{pron}$ $\gg$ PR.$\mathcal{A}_{ID}$ $\gg$ PR.$\mathcal{A}_{FD}$ $\gg$ PR.$\mathcal{A}_{SD}$ $\gg$ FAITH$_{SE}$ $\gg$ PR.$\mathcal{A}_{CD}$ $\gg$ FAITH$_{SELF}$ $\circ$ (PR.$\mathcal{A}_{ThD}$ $\gg$ PR.$\mathcal{A}_{XP}$)

(93) *English ranking:*
FAITH$_{pron}$ $\gg$ PR.$\mathcal{A}_{ID}$ $\gg$ PR.$\mathcal{A}_{FD}$ $\gg$ PR.$\mathcal{A}_{SD}$ $\gg$ FAITH$_{SE}$ $\gg$ PR.$\mathcal{A}_{CD}$ $\circ$ FAITH$_{SELF}$ $\gg$ PR.$\mathcal{A}_{ThD}$ $\gg$ PR.$\mathcal{A}_{XP}$

(94) *Dutch ranking:*
FAITH$_{pron}$ $\gg$ PR.$\mathcal{A}_{ID}$ $\gg$ PR.$\mathcal{A}_{FD}$ $\gg$ PR.$\mathcal{A}_{SD}$ $\gg$ FAITH$_{SE}$ $\circ$ PR.$\mathcal{A}_{CD}$ $\gg$ FAITH$_{SELF}$ $\circ$ PR.$\mathcal{A}_{ThD}$ $\gg$ PR.$\mathcal{A}_{XP}$

(95) *Italian ranking:*
FAITH$_{pron}$ $\gg$ PR.$\mathcal{A}_{ID}$ $\gg$ PR.$\mathcal{A}_{FD}$ $\gg$ PR.$\mathcal{A}_{SD}$ $\gg$ FAITH$_{SE}$ $\circ$ PR.$\mathcal{A}_{CD}$ $\gg$ FAITH$_{SELF}$ $\circ$ (PR.$\mathcal{A}_{ThD}$ $\gg$ PR.$\mathcal{A}_{XP}$)

T_{18}: General predictions[46]

[46] Recall that if binding takes place within domain Y, the crucial PRINCIPLE $\mathcal{A}$-constraint that determines the outcome of the competition is the one which refers to the next smaller domain relevant for binding – hence the notation "XD+1" in T_{18}.

ranking					*optimal realization if binding relation within XD+1*
FAITH$_{SE}$	$\gg$	FAITH$_{SELF}$	$\gg$	**PR.**$\mathcal{A}_{XD}$	SELF anaphor
FAITH$_{SE}$	$\gg$	FAITH$_{SELF}$	$\circ$	**PR.**$\mathcal{A}_{XD}$	SELF/SE anaphor
FAITH$_{SE}$	$\gg$	**PR.**$\mathcal{A}_{XD}$	$\gg$	FAITH$_{SELF}$	SE anaphor
FAITH$_{SE}$	$\circ$	**PR.**$\mathcal{A}_{XD}$	$\gg$	FAITH$_{SELF}$	SE anaphor/pronoun
PR.$\mathcal{A}_{XD}$	$\gg$	FAITH$_{SE}$	$\gg$	FAITH$_{SELF}$	pronoun

Ideally, a theory of binding does not only account for the binding patterns of a particular language but also captures generalizations that seem to hold universally. For example, it can be observed that complex anaphors surface only if the binding relation is relatively local, and the less local the binding relation gets, the more probable it is that first complex anaphors and later also simple anaphors are ruled out, and only pronouns are licit.

These generalizations are captured by the present approach in the following way: If we deal with a local binding relationship, only few, low-ranked PRINCIPLE $\mathcal{A}$-constraints can apply non-vacuously before checking takes place; and since only these constraints favour a reduction of the realization matrix, it is very likely that the candidate with the full specification [SELF, SE, pron] is optimal and the SELF anaphor is finally selected as optimal realization. On the other hand, if the binding relation is less local, more PRINCIPLE $\mathcal{A}$-constraints apply non-vacuously, because x enters bigger and bigger domains unchecked; and since the constraints referring to these domains are higher ranked, it becomes more and more likely that the specification matrix of x is gradually reduced in the course of the derivation and a less anaphoric form is selected as optimal realization. (In the end, only [pron] might be left, and in this case MAB can only choose the pronominal form as optimal form for x.)

Furthermore, it is predicted that if x is realized as SELF/SE anaphor if binding takes place in domain XD, these realizations are also licit if binding is more local, because an anaphoric specification can only win if the corresponding matrix has been in the candidate set – and if it had not won the competitions before, only reduced matrices could have served as competitors.

On the other hand, if x is realized as pronoun, pronominal binding is also possible if binding occurs in a bigger domain, because the reduced matrix [pron] will serve as input for the subsequent competitions, which inevitably yields a pronominal winner.

This shows that the theory developed here is both flexible enough to account for crosslinguistic variation and optionality and restrictive enough to capture universal binding properties and restrict possible binding scenarios (cf. also chapter 2, section **??**).

5.8 Long-Distance Anaphora (LDA) in Icelandic

So far, there has been no need to distinguish between the three highest-ranked PRINCIPLE $\mathcal{A}$-constraints, PRINCIPLE $\mathcal{A}_{SD}$, PRINCIPLE $\mathcal{A}_{FD}$, and PRINCIPLE $\mathcal{A}_{ID}$. However, the ranking of these constraints is crucial if we want to capture the different behaviour of languages that exhibit long-distance anaphora. Let us start once more with the Icelandic examples in (96).

(96) *Icelandic:*

a. Jón$_1$ skipaði Pétri$_2$ PRO$_2$ að raka
   John ordered Peter to shave$_{inf}$
   sig$_1$/??sjálfan sig$_1$/hann$_1$ á hverjum degi.
   SE/himself/him on every day
   'John$_1$ ordered Peter to shave him$_1$ every day.'

b. Jón$_1$ segir að Pétur raki sig$_1$/??sjálfan sig$_1$/hann$_1$ á
   John says that Peter shave$_{sub}$ SE/himself/him on
   hverjum degi.
   every day
   'John$_1$ says that Peter shaves him$_1$ every day.'

c. Jón$_1$ veit að Pétur rakar ??sig$_1$/*sjálfan sig$_1$/hann$_1$ á
   John knows that Peter shave$_{ind}$ SE/himself/him on
   hverjum degi.
   every day
   'John$_1$ knows that Peter shaves him$_1$ every day.'

In a sentence like (96-a) (repeated in (97)), where the binding relation is not established unless the finite domain is reached, optimization occurs relatively frequently until x is finally checked. In the following discussion, I ignore these ealier parts of the derivation, since the goal of this section is to investigate what determines long-distance binding; assume therefore that we have already reached the stage when the minimal subject domain (= embedded vP) is reached.

(97) Jón$_1$ skipaði Pétri$_2$ PRO$_2$ að raka$_{inf}$ sig$_1$/??sjálfan sig$_1$/hann$_1$ á hverjum degi.

a. [$_{\text{vP}}$ $x_{[\beta]}$ PRO raka [$_{\text{VP}}$ t$_x$′ t$_{raka}$ $\cancel{\text{t}_x}$]]

(97-a) illustrates the point in the derivation when the embedded vP is optimized. At this stage, the material in the accessible domain allows us to classify this phrase as x's θ-domain, Case domain, and subject domain; and as x remains unchecked, the following four PRINCIPLE $\mathcal{A}$-constraints apply non-vacuously: PRINCIPLE $\mathcal{A}_{XP}$, PRINCIPLE $\mathcal{A}_{ThD}$, PRINCIPLE $\mathcal{A}_{CD}$, and PRINCIPLE $\mathcal{A}_{SD}$. Since we know that the latter outranks the first three constraints (due to the underlying universal subhierarchy), its ranking will determine the optimal realization of x. (Note that the remaining optimizations until the binder enters the derivation in (98-b) can be neglected because no further domain relevant for binding is reached.) If PRINCIPLE $\mathcal{A}_{SD}$ is ranked above FAITH$_{SE}$, [pron] is the optimal matrix, if it is ranked below FAITH$_{SE}$, [SE, pron] is optimal, and if PRINCIPLE $\mathcal{A}_{SD}$ and FAITH$_{SE}$ are tied, both matrices win, and x might therefore be realized as pronoun or SE anaphor. The latter option is chosen in Icelandic (cf. T$_{19}$-T$_{19.2}$).[47,48]

[47]As mentioned in chapter 2, another (possibly older) variant of Icelandic seems to favour the SE anaphor in this context, which is predicted by the ranking FAITH$_{SE}$ $\gg$ PR.$\mathcal{A}_{SD}$.

[48]For reasons of space, I combine the lower-ranked constraints PR.$\mathcal{A}_{CD}$, PR.$\mathcal{A}_{ThD}$, and PR.$\mathcal{A}_{XP}$ in the subsequent tableaux.

Although I do not want to present a detailed analysis of local binding relations in Icelandic, the following Icelandic data provide conclusive information as regards the ranking of the lower-ranked constraints. (The data are again from Gunnar Hrafn Hrafnbjargarson (p.c.).)

(i) a. Max$_1$ hatar sig$_1$/sjálfan sig$_1$/*hann$_1$.
Max hates SE/himself/him
'Max$_1$ hates himself$_1$.'

T_{19}: vP optimization
(XP/ThD/CD/SD reached – $x_{[\beta]}$ unchecked)

Candidates	F_{pron}	F_{SE}	PR.$\mathcal{A}_{SD}$	PR.$\mathcal{A}_{CD/ThD/XP}$	F_{SELF}
O_1: [SELF, SE, pron]			$**!$	$**$	
⇒ O_2: [SE, pron]			$*(!)$	$*$	$*$
⇒ O_3: [pron]		$*(!)$			$*$

(98) b. $[_{\text{vP}}$ Jón$_{[*\beta*]}$ skipaði $[_{\text{VP}}$ $x_{[\beta]}$ Pétri$_2$ t$_{skip.}$ $\cancel{[_{\text{CP}} \ldots]}$]]]

$T_{19.1}$: vP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ *applies vacuously)*

Input: O_1/T_{19}	F_{pron}	F_{SE}	F_{SELF}
⇒ O_{21}: [SE, pron]			$*$
O_{22}: [pron]		$*!$	$*$

$T_{19.2}$: vP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ *applies vacuously)*

---

b. Max$_1$ heyrði sig$_1$/sjálfan sig$_1$/*hann$_1$ syngja.
Max heard SE/himself/him sing
'Max$_1$ heard himself$_1$ sing.'

c. Max$_1$ leit aftur fyrir sig$_1$/sjálfan sig$_1$/*hann$_1$.
Max glanced behind SE/himself/him
'Max$_1$ glanced behind himself$_1$/him$_1$.'

In all three examples, which illustrate binding within the θ-, Case, and subject domain, respectively, both types of anaphors are licit while the pronoun is excluded. (Note that (i-a) patterns like German and Italian, (i-b) like German, Italian, and Dutch, and (i-c) basically like English in allowing the complex anaphor and the next less anaphoric element.) As can be read off T_{18}, this result is predicted if the constraints PR.$\mathcal{A}_{XP}$, PR.$\mathcal{A}_{ThD}$, and PR.$\mathcal{A}_{CD}$ are tied with FAITH$_{SELF}$. Hence, we get the ranking in (ii-a) for Icelandic, which is an abbreviation for the four underlying constraint orders in (ii-b).

(ii) a. (PR.$\mathcal{A}_{CD}$ $\gg$ PR.$\mathcal{A}_{ThD}$ $\gg$ PR.$\mathcal{A}_{XP}$) $\circ$ FAITH$_{SELF}$
b. (i) PR.$\mathcal{A}_{CD}$ $\gg$ PR.$\mathcal{A}_{ThD}$ $\gg$ PR.$\mathcal{A}_{XP}$ $\gg$ FAITH$_{SELF}$
(ii) PR.$\mathcal{A}_{CD}$ $\gg$ PR.$\mathcal{A}_{ThD}$ $\gg$ FAITH$_{SELF}$ $\gg$ PR.$\mathcal{A}_{XP}$
(iii) PR.$\mathcal{A}_{CD}$ $\gg$ FAITH$_{SELF}$ $\gg$ PR.$\mathcal{A}_{ThD}$ $\gg$ PR.$\mathcal{A}_{XP}$
(iv) FAITH$_{SELF}$ $\gg$ PR.$\mathcal{A}_{CD}$ $\gg$ PR.$\mathcal{A}_{ThD}$ $\gg$ PR.$\mathcal{A}_{XP}$

	Input: O_2/T_{19}	F_{pron}	F_{SE}	F_{SELF}
⇒	O_{31}: [pron]		*	*

The next example (repeated from (96-b)) involves binding into a subjunctive complement clause. This means that at the stage when the finite domain is reached (= embedded vP), x is still free (cf. (99-a)). As a result, all PRINCIPLE $\mathcal{A}$-constraints except PRINCIPLE $\mathcal{A}_{ID}$ apply non-vacuously when this phrase is optimized. Hence, the highest PRINCIPLE $\mathcal{A}$-constraint that is involved in this competition is PRINCIPLE $\mathcal{A}_{FD}$, and if it is tied with FAITH$_{SE}$, both O_2 and O_3 are predicted to be optimal (cf. T_{20}).

(99) Jón$_1$ segir að Pétur raki$_{sub}$ sig$_1$/??sjálfan sig$_1$/hann$_1$ á hverjum degi.

a. $[_{\text{vP}}\ x_{[\beta]}$ Pétur raki $[_{\text{VP}}\ \text{t}_x{}'\ \text{t}_{raki}\ \cancel{\text{t}_x}]]$

T_{20}: vP optimization
(XP/ThD/CD/SD/FD reached – $x_{[\beta]}$ unchecked)

	Candidates	F_{pron}	PR.$\mathcal{A}_{FD}$	F_{SE}	PR.$\mathcal{A}_{SD}$	PR.$\mathcal{A}_{CD/ThD/XP}$	F_{SELF}
	O_1: [SELF, SE, pr]		**		**	**	
⇒	O_2: [SE, pron]		*(!)		*	*	*
⇒	O_3: [pron]			*(!)			*

During the next optimization processes (TP, CP, VP), the outcome remains unchanged: Since no new domain is reached, no higher-ranked PRINCIPLE $\mathcal{A}$-constraint gets involved and might force a further reduction of the matrix.[49] Hence, the matrices [SE, pron] and [pron] function as input when the matrix vP is optimized (cf. $T_{20.1}$ and $T_{20.2}$ respectively), and since x is checked at this point in the derivation, [SELF, pron] wins in the former competition and [pron] in the latter. Thus, according to MAB both the SE anaphor and the pronoun turn out to be optimal realizations in this example, which is the desired result.

[49]Note that even if the new accessible domains no longer qualify as θ-, Case, subject, or finite domain (for example, the matrix VP), the result is not blurred, because as long as no new higher-ranked PRINCIPLE $\mathcal{A}$-constraint is activated, the matrices are not reduced any further.

(100) b. $[_{\text{vP}}$ Jón$_{[*\beta*]}$ segir $[_{\text{VP}}\ x_{[\beta]}\ \text{t}_{segir}\ \cancel{[_{\text{CP}} \ldots]}]]$

$T_{20.1}$: vP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ *applies vacuously)*

Input: O_2/T_{20}	F_{pron}	F_{SE}	F_{SELF}
⇒ O_{21}: [SE, pron]			∗
O_{22}: [pron]		∗!	∗

$T_{20.2}$: vP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ *applies vacuously)*

Input: O_3/T_{20}	F_{pron}	F_{SE}	F_{SELF}
⇒ O_{31}: [pron]		∗	∗

(101) (repeated from (96-c)) differs from the previous examples insofar as already the embedded vP qualifies as indicative domain; but since it does not include x's antecedent, all PRINCIPLE $\mathcal{A}$-constraints apply non-vacuously when vP is optimized (cf. T_{21}). What is crucial here is that PRINCIPLE $\mathcal{A}_{ID}$ is ranked above FAITH$_{SE}$: On this assumption, O_3 is the winner of the competition, which leads to the result that [pron] remains the only optimal candidate when x is finally checked (cf. $T_{21.1}$), and MAB finally correctly predicts that x must be realized as a pronoun.

(101) Jón$_1$ veit að Pétur rakar$_{ind}$??sig$_1$/*sjálfan sig$_1$/hann$_1$ á hverjum degi.

a. $[_{\text{vP}}\ x_{[\beta]}$ Pétur rakar $[_{\text{VP}}\ \text{t}_x{}'\ \text{t}_{rakar}\ \cancel{\text{t}_x}]]$

T_{21}: vP optimization
(XP/ThD/CD/SD/FD/ID reached – $x_{[\beta]}$ unchecked)

Candidates	F_{pron}	PR.$\mathcal{A}_{ID}$	PR.$\mathcal{A}_{FD}$	F_{SE}	PR.$\mathcal{A}_{SD}$	PR.$\mathcal{A}_{CD/ThD/XP}$	F_{SELF}
O_1: [S, S, pr]		∗!∗	∗∗		∗∗	∗∗	
O_2: [SE, pr]		∗!	∗		∗	∗	∗
⇒ O_3: [pron]				∗			∗

(102) b. $[_{\text{vP}}$ Jón$_{[*\beta*]}$ veit $[_{\text{VP}}\ x_{[\beta]}\ \text{t}_{veit}\ \cancel{[_{\text{CP}} \ldots]}]]$

$T_{21.1}$: vP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ *applies vacuously)*

Input: O_3/T_{21}	F_{pron}	F_{SE}	F_{SELF}
$\Rightarrow$ O_{31}: [pron]		*	*

5.9 Summary: Crosslinguistic Variation II

Basically, we can distinguish between four different types of languages as regards their behaviour with respect to long distance binding (cf. also chapter 2, section **??**). There are languages which do not allow anaphoric binding in this case (like English and German), some languages only allow long anaphoric binding into infinitive complements (like Russian), type 3 prohibits LDA only if indicative complements intervene (like Icelandic), and the last type even allows intervening indicative complement clauses (like Faroese). This crosslinguistic variation is captured by reranking the constraint subhierarchy (PRINCIPLE $\mathcal{A}_{ID} \gg$ PRINCIPLE $\mathcal{A}_{FD} \gg$ PRINCIPLE $\mathcal{A}_{SD}$) with FAITH$_{SE}$ in different ways; the respective predictions are represented in (103)-(106). The ties in (104-a)-(106-a) predict optionality between anaphoric and pronominal binding; if the pronominal realization is illicit, FAITH$_{SE}$ must be ranked above the respective PRINCIPLE $\mathcal{A}$-constraint(s).[50]

(103) *Languages without LDA:*
PR.$\mathcal{A}_{ID} \gg$ PR.$\mathcal{A}_{FD} \gg$ PR.$\mathcal{A}_{SD} \gg$ FAITH$_{SE}$

[50] As far as (106-b) is concerned, it predicts a language which does not distinguish between simple anaphoric and pronominal forms: Due to this ranking, [SE, pron] always beats [pron], so it is always the same type of element that is inserted in the end – a vocabulary item with the feature specification {SE, pron} if available, or an item with less features otherwise. Hence, English could in principle have this ranking (instead of (103) as assumed in (93)); on this assumption, pronominal binding in English would never be based on the matrix [pron] but always on the realization matrix [SE, pron] instead. However, due to the lack of a corresponding vocabulary item, the less specified form *pron* would always be inserted.

(104) *Languages with intervening infinitive complements only:*

a. *anaphoric or pronominal binding possible:*
$\textsc{Pr}.\mathcal{A}_{ID} \gg \textsc{Pr}.\mathcal{A}_{FD} \gg \textsc{Faith}_{SE} \circ \textsc{Pr}.\mathcal{A}_{SD}$

b. *only anaphoric binding licit:*
$\textsc{Pr}.\mathcal{A}_{ID} \gg \textsc{Pr}.\mathcal{A}_{FD} \gg \textsc{Faith}_{SE} \gg \textsc{Pr}.\mathcal{A}_{SD}$

(105) *Languages with intervening infinitive or subjunctive complements:*

a. *anaphoric or pronominal binding possible:*
$\textsc{Pr}.\mathcal{A}_{ID} \gg \textsc{Faith}_{SE} \circ (\textsc{Pr}.\mathcal{A}_{FD} \gg \textsc{Pr}.\mathcal{A}_{SD})$

b. *only anaphoric binding licit:*
$\textsc{Pr}.\mathcal{A}_{ID} \gg \textsc{Faith}_{SE} \gg \textsc{Pr}.\mathcal{A}_{FD} \gg \textsc{Pr}.\mathcal{A}_{SD}$

(106) *Languages which even allow intervening indicative complements:*

a. *anaphoric or pronominal binding possible:*
$\textsc{Faith}_{SE} \circ (\textsc{Pr}.\mathcal{A}_{ID} \gg \textsc{Pr}.\mathcal{A}_{FD} \gg \textsc{Pr}.\mathcal{A}_{SD})$

b. *only anaphoric binding licit:*
$\textsc{Faith}_{SE} \gg \textsc{Pr}.\mathcal{A}_{ID} \gg \textsc{Pr}.\mathcal{A}_{FD} \gg \textsc{Pr}.\mathcal{A}_{SD}$

5.10 Principle C Derivationally

By now, the distribution of bound anaphors and pronouns has been extensively discussed. The effects of Principle A and B of the standard Binding Theory have been derived by means of two universal, but violable, constraint subhierarchies, which made it possible to integrate the phenomenon of binding into a derivational model and improve at the same time descriptive adequacy. What remains to be shown is how the third traditional binding principle, Principle C, can be integrated into this approach. Since Principle C refers to R-expressions, let us first think about the status of full NPs in this model in general.

We have come across R-expressions in this chapter before, namely the antecedents in the previous examples, and it has tacitly been assumed that these R-expressions are simply part of the numeration. However, if we consider R-expressions that function as potential bindees, something more must be said. If we stick to the assumption that bound elements do not occur in the numeration as concrete items but are represented by means of a realization matrix, bound R-expressions must result from an optimization procedure which is based on a matrix that contains not only the pronominal and two anaphoric forms, but also the R-expression.

However, a realization matrix can only contain different forms whose semantic contribution to the sentence is the same and does not change the underlying meaning of the sentence in any way. Therefore a matrix can only contain an R-expression if its designated antecedent is an R-expression and it can be considered to be a copy of it. For the sake of concreteness, consider the following examples. If we want to say that John likes himself, this meaning is expressed unambiguously with the following form: *John*$_1$ *likes* x_1; whether we have to realize x as *himself*, *him*, or *John* basically depends on the language under consideration and is a question that is answered by the syntactic component in the course of the derivation. However, in a sentence such as *He*$_1$ *likes* x_1, the situation is slightly different. Whether x_1=*him*$_1$ or *himself*$_1$ does not make any difference with respect to semantics, but if x were realized as an R-expression such as *John*, additional information would be added and *John* would not just be an equivalent variant of *him* or *himself* in this case. Thus the R-expression cannot be part of the realization matrix in the latter example.

Hence, we can draw the following conclusion. If the designated antecedent of x is an R-expression, x's realization matrix additionally contains a copy of this R-expression and the maximal realization matrix is then [SELF, SE, pron, R-ex]. Thus, there are in principle two possibilities how R-expressions can emerge. If they do not function as a bound element, they directly form part of the numeration; as bound elements, on the other hand, they are encoded as x in the numeration and can turn out to be optimal if the matrix [R-ex] wins in the end.

Against this background, one type of Principle C effect can be accounted for straightforwardly: The system predicts that R-expressions cannot be bound by pronouns, because in this scenario the realization matrix of x cannot contain an R-expression at all (cf. (107)).

(107) a. *He$_1$ likes John$_1$.

b. *Underlying scenario:*
he$_1$ likes x_1; x=[SELF, SE, pron]
$\rightarrow$ x= R-expression impossible

However, there will have to be another explanation as to why Principle C effects that involve R-expressions being bound by R-expressions must be ruled out. Such a configuration is not prohibited *a priori*, because in this case x's realization matrix contains a copy of the binding R-expression (cf. (108)). Hence, this configuration must be ruled out in the course of the derivation (which is illustrated below).

(108) a. *John$_1$ likes John$_1$.
b. *Underlying scenario:*
John$_1$ likes x_1; x=[SELF, SE, pron, R-ex]
$\rightarrow$ x= R-expression in principle possible

At first sight, it might look unattractive to have different accounts of Principle C effects, but if we think again of those languages where Principle C is violable in certain contexts, this split turns out to be an advantage, because it accounts for the following observation: Although it is possible in languages like Vietnamese that R-expressions are bound by R-expressions, they can never be bound by pronouns (cf. (109); see also chapter 2, section **??**). The former scenario might come about if the constraints are ranked accordingly, but the latter is ruled out in general due to the nature of realization matrices as such.[51]

(109) *Vietnamese:*
John$_1$/*nó$_1$ tin John$_1$ sẽ thắńg.
John/he thinks John will win
'John$_1$ thinks he$_1$ will win.'

Let us now turn to those examples in which R-expressions are bound by R-expressions, as, for instance, in the following German sentences (also repeated from chapter 2).

[51]Example (109) is repeated from chapter 2, section **??** and was quoted from Lasnik (1991).

(110) a. Max$_1$ weiß, dass Maria *Max$_1$/ihn$_1$/*sich$_1$/*sich selbst$_1$
Max knows that Mary Max/him/SE/himself
mag.
likes
'Max$_1$ knows that Mary likes him$_1$.'

b. Max$_1$ mag *Max$_1$/*ihn$_1$/sich$_1$/sich selbst$_1$.
Max likes Max/him/SE/himself
'Max$_1$ likes himself$_1$.'

(111) Peter$_1$ mag seine$_1$/*Peters$_1$ Bücher.
Peter likes his/Peter's books
'Peter$_1$ likes his$_1$ books.'

Since this type of examples involves realization matrices of the sort [SELF, SE, pron, R-ex], the number of candidates in the subsequent tableaux is increased to maximally four different output candidates: O_{n1}=[SELF, SE, pron, R-ex], O_{n2}=[SE, pron, R-ex], O_{n3}=[pron, R-ex], and O_{n4}=[R-ex] (with n=0, 1, 2, ...). Following the assumptions in chapter 2, R-ex counts as the least anaphoric possible realization, hence only the last candidate does not violate the PRINCIPLE $\mathcal{A}$-constraints; the first candidate violates them three times, the second one twice, and the third candidate once. As far as the faithfulness constraints are concerned, the matrix [R-ex] does not only violate FAITH$_{SELF}$ and FAITH$_{SE}$, but also FAITH$_{pron}$. For the sake of completeness, the FAITH-subhierarchy can be complemented along the following lines:

(112) a. FAITH$_{R-ex}$ (F_{R-ex}):
The realization matrix for x must contain [R-ex].

b. FAITH$_{R-ex}$ $\gg$ FAITH$_{pron}$ $\gg$ FAITH$_{SE}$ $\gg$ FAITH$_{SELF}$

Against this background, the derivation of (110-a) (repeated in (113)) proceeds as follows. When the first phrase is completed, only PRINCIPLE $\mathcal{A}_{XP}$ applies non-vacuously, and O_1 and O_2 are predicted to be optimal (cf. T_{22}), which means that there are two competitions after the completion of the next phrase.

(113) Max$_1$ weiß, dass Maria *Max$_1$/ihn$_1$/*sich$_1$/*sich selbst$_1$ mag.

a. $[_{\text{VP}}\ x_{[\beta]}\ [_{\text{V}'}\ \text{t}_x \text{ mag}]]$

T_{22}: VP optimization
(XP reached – $x_{[\beta]}$ unchecked)

	Candidates	F_{R-ex}	F_{pron}	F_{SE}	F_{SELF}	PR.$\mathcal{A}_{XP}$
⇒	O_1: [SELF, SE, pron, R-ex]					$***(!)$
⇒	O_2: [SE, pron, R-ex]				$*(!)$	$**$
	O_3: [pron, R-ex]			$*!$	$*$	$*$
	O_4: [R-ex]		$*!$	$*$	$*$	

When the embedded vP is optimized, x is still free, and since the accessible domain fulfils all domain definitions, all PRINCIPLE $\mathcal{A}$-constraints are involved in this competition.[52] The crucial ranking is now FAITH$_{pron}$ $\gg$ PRINCIPLE $\mathcal{A}_{ID}$; since the matrix [R-ex] violates this FAITH-constraint, it loses against the matrix [pron, R-ex]. Hence, a maximal reduction of the realization matrix is prevented, which would result in the eventual selection of the R-expression as optimal realization, and [pron, R-ex] is predicted to be optimal in both $T_{22.1}$ and $T_{22.2}$.[53]

(114) b. $[_{\text{vP}}\ x_{[\beta]}$ Maria $[_{\text{VP}}\ \text{t}_x{}'\ [_{\text{V}'}\ \cancel{\text{t}_x}\ \text{t}_{mag}]]$ mag$]$

$T_{22.1}$: vP optimization
(XP/ThD/CD/SD/FD/ID reached – $x_{[\beta]}$ unchecked)

Input: O_1/T_{22}	F_{pron}	PR.$\mathcal{A}_{ID/FD/SD}$	F_{SE}	PR.$\mathcal{A}_{CD}$	PR.$\mathcal{A}_{ThD}$	F_{SELF}	PR.$\mathcal{A}_{XP}$
O_{11}: [S, S, pr, R]		$**!*$		$***$	$***$		$***$
O_{12}: [S, pr, R]		$**!$		$**$	$**$	$*$	$**$
⇒ O_{13}: [pr, R]		$*$	$*$	$*$	$*$	$*$	$*$
O_{14}: [R-ex]	$*!$		$*$			$*$	

[52] Since German does not have long distance anaphora, PR.$\mathcal{A}_{ID}$, PR.$\mathcal{A}_{FD}$, and PR.$\mathcal{A}_{SD}$ need not be distinguished in the subsequent tableaux.

[53] Since it is not violated by any candidate anyway, the constraint FAITH$_{R-ex}$ is neglected in these tableaux.

$T_{22.2}$: vP optimization
(XP/ThD/CD/SD/FD/ID reached – $x_{[\beta]}$ unchecked)

Input: O_2/T_{22}	F_{pron}	$\textsc{Pr}.\mathcal{A}_{ID/FD/SD}$	F_{SE}	$\textsc{Pr}.\mathcal{A}_{CD}$	$\textsc{Pr}.\mathcal{A}_{ThD}$	F_{SELF}	$\textsc{Pr}.\mathcal{A}_{XP}$
O_{21}: [S, pr, R]		$**!$		$**$	$**$	$*$	$**$
⇒ O_{23}: [pr, R]		$*$	$*$	$*$	$*$	$*$	$*$
O_{24}: [R-ex]	$*!$		$*$			$*$	

Since there is no further domain that could be reached and could therefore force a further reduction of the matrix, the subsequent optimizations can be neglected. So let us turn to the point in the derivation when the binder is finally merged in, which is illustrated in (115-c). Since the [β]-feature is checked at this stage, only the FAITH-constraints apply non-vacuously in $T_{24.1.1/24.2.1}$, which yields [pron, R-ex] as optimal matrix, and according to the MAB-principle this means that x must be realized as pronoun. Hence, bound R-expressions are excluded in this type of example, because pronouns are the better choice.

(115) c. $[_{\text{vP}}$ Max$_{[*\beta*]}$ $[_{\text{VP}}$ $x_{[\beta]}$ $\cancel{[_{\text{CP}} \ldots]}$ t$_{w.}$] weiß]

$T_{22.1.1/22.2.1}$: vP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ *applies vacuously)*

Input: $O_{13}/T_{22.1}$ or $O_{23}/T_{22.2}$	$F_{R-ex.}$	F_{pron}	F_{SE}	F_{SELF}
⇒ $O_{131/231}$: [pron, R-ex]			$*$	$*$
$O_{132/232}$: [R-ex]		$*!$	$*$	$*$

The next example (repeated from (110-b)) patterns similarly; the only difference is that here the anaphoric forms turn out to be the better alternative.

(116) Max$_1$ mag *Max$_1$/*ihn$_1$/sich$_1$/sich selbst$_1$.
a. $[_{\text{VP}}$ $x_{[\beta]}$ $[_{\text{V}'}$ t$_x$ mag]]

As in the previous example, O_1 and O_2 win when the embedded VP is optimized, as T_{23} shows.

T_{23}: VP optimization

(XP reached – $x_{[\beta]}$ unchecked)

	Candidates	F_{R-ex}	F_{pron}	F_{SE}	F_{SELF}	PR.$\mathcal{A}_{XP}$
⇒	O_1: [S, S, pr, R]					$***(!)$
⇒	O_2: [S, pr, R]				$*(!)$	$**$
	O_3: [pr, R]			$*!$	$*$	$*$
	O_4: [R-ex]		$*!$	$*$	$*$	

Already in the next phrase, the antecedent enters the derivation and $x_{[\beta]}$ is checked; hence, the PRINCIPLE $\mathcal{A}$-constraints apply vacuously, and the matrices [SELF, SE, pron, R-ex]/[SE, pron, R-ex] remain optimal (cf. $T_{23.1}$ and $T_{23.2}$, respectively). As a result, MAB predicts the two anaphoric forms to be the optimal realizations.

(117) b. $[_{\text{vP}}\ \text{Max}_{[*\beta*]}\ [_{\text{VP}}\ x_{[\beta]}\ [_{\text{V}'}\ \cancel{t_x}\ t_{mag}]]\ \text{mag}]$

$T_{23.1}$: vP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ applies vacuously)

	Input: O_1/T_{23}	F_{R-ex}	F_{pron}	F_{SE}	F_{SELF}
⇒	O_{11}: [SELF, SE, pron, R-ex]				
	O_{12}: [SE, pron, R-ex]				$*$
	O_{13}: [pron, R-ex]			$*!$	$*$
	O_{14}: [R-ex]		$*!$	$*$	$*$

$T_{23.2}$: vP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ applies vacuously)

	Input: O_2/T_{23}	$F_{R-ex.}$	F_{pron}	F_{SE}	F_{SELF}
⇒	O_{22}: [SE, pron, R-ex]				$*$
	O_{23}: [pron, R-ex]			$*!$	$*$
	O_{24}: [R-ex]		$*!$	$*$	$*$

The example in (118) (repeated from (111)) can be derived similarly. When the NP in (118-a) is optimized, the matrix [SE, pron, R-ex] is predicted to be optimal (cf. T_{24}), and until the antecedent is merged into the derivation, no further domain relevant for binding is reached. As a result, [SE, pron, R-ex] remains the optimal matrix (cf. $T_{24.1}$), and since German anaphors lack a genitive form, MAB finally selects the most anaphoric form available that is compatible with this matrix – the pronominal form *seine* ('his').

(118) Peter$_1$ mag seine$_1$/*Peters$_1$ Bücher.

a. [$_{\text{NP}}$ $x_{[\beta]}$ Bücher]

T_{24}: NP optimization
(XP/ThD/CD reached – $x_{[\beta]}$ unchecked)

Candidates	F_{pron}	F_{SE}	PR.$\mathcal{A}_{CD}$	PR.$\mathcal{A}_{ThD}$	F_{SELF}	PR.$\mathcal{A}_{XP}$
O_1: [S, S, pr, R]			∗∗∗!	∗∗∗		∗∗∗
⇒ O_2: [S, pr, R]			∗∗	∗∗	∗	∗∗
O_3: [pr, R]		∗!	∗	∗	∗	∗
O_4: [R-ex]	∗!	∗			∗	

(119) b. [$_{\text{vP}}$ Peter$_{[*\beta*]}$ [$_{\text{VP}}$ [$_{\text{NP}}$ $x_{[\beta]}$ Bücher] t_{NP} t_{mag}] mag]

$T_{24.1}$: vP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ *applies vacuously)*

Input: O_2/T_{24}	F_{pron}	F_{SE}	F_{SELF}
⇒ O_{21}: [SE, pron, R-ex]			∗
O_{22}: [pron, R-ex]		∗!	∗
O_{23}: [R-ex]	∗!	∗	∗

As the previous examples showed, it is always the high-ranked constraint FAITH$_{pron}$ which rules out bound R-expressions (with R-expressions as antecedents). However, if this type of Principle C effect is derived by a particular ranking, it should in principle be possible to obviate these effects if FAITH$_{pron}$ is ranked sufficiently low. This is exactly what we find in languages like Vietnamese, where R-expressions may be bound by R-expressions. Hence we can account for the grammaticality of the Vietnamese example in (120) (repeated from (109)) if we assume that (at least) PRINCIPLE $\mathcal{A}_{ID}$ is not ranked below

Faith_{pron} in languages of this type (cf. $T_{25.1}$ and $T_{25.2}$).

(120) *Vietnamese:*
John$_1$ tin John$_1$ sẽ thắńg.
John thinks John will win
'John$_1$ thinks he$_1$ will win.'

$T_{25.1}$: The emergence of bound R-expressions I

	Candidates	$\text{Pr.}\mathcal{A}_{ID}$	F_{pron}
	O_1: [SELF, SE, pron, R-ex]	∗!∗∗	
	O_2: [SE, pron, R-ex]	∗!∗	
	O_3: [pron, R-ex]	∗!	
⇒	O_4: [R-ex]		∗

$T_{25.2}$: The emergence of bound R-expressions II

	Candidates	$\text{Pr.}\mathcal{A}_{ID}$	F_{pron}
	O_1: [SELF, SE, pron, R-ex]	∗∗!∗	
	O_2: [SE, pron, R-ex]	∗∗!	
⇒	O_3: [pron, R-ex]	∗(!)	
⇒	O_4: [R-ex]		∗(!)

5.11 Inherently Reflexive Predicates Revisited

Let us now come back to those cases where anaphors and pronouns occur without establishing a binding relation, as, for instance, in the following examples involving inherently reflexive predicates.

(121) a. *German:*
Max benimmt sich/*sich selbst/*ihn (wie ein Gentleman).
Max behaves SE/himself/him like a gentleman

b. *Dutch:*
Max gedraagt zich/*zichzelf/*hem.
Max behaves SE/himself/him

c. *Frisian:*
Max hâld him/*himsels.
Max behaves him/himself

(122) *English:*

a. Max behaves like a gentleman.
b. Max behaves himself.

As argued in chapter 2, anaphors that occur together with inherently reflexive predicates do not function as arguments and are not bound by an antecedent. If this is translated into the present derivational approach, it means that no [β]-feature is involved. Let us therefore assume that inherently reflexive predicates are predicates that enter the numeration with an x that does not bear a [β]-feature. As a result, they might occur with an anaphoric or pronominal form, but they do not have to, since x does not stand for an argument. And since no [β] is involved, it follows moreover that the PRINCIPLE $\mathcal{A}$-constraints apply vacuously throughout the derivation.

However, if the universally equally ranked FAITH-constraints were the only constraints relevant for the derivation of the sentences in (121) and (122), we would not expect any crosslinguistic variation and the complex anaphor would be predicted to be optimal in general. Hence, there must be another constraint that can interact with this universal constraint subhierarchy in different ways and which prefers less anaphoric elements. Let us therefore introduce the constraint in (123); on the assumption that anaphoric specification reflexive-marks a predicate, it is violated three times by the matrix [SELF, SE, pron], twice by [SE, pron], once by [pron] and not at all by the fourth candidate, [–], where the realization matrix has been emptied completely.[54]

(123) *REFLMARK$_{inh}$:
Inherently reflexive predicates must be minimally reflexive-marked.

[54]Note that if inherently reflexive predicates are involved, there is no R-expression in the realization matrix; since x does not have an antecedent in these examples, the matrix cannot contain a corresponding copy.

Let us first consider languages like German and Dutch, where inherently reflexive predicates occur with SE anaphors. So let us derive the German example in (124) (repeated from (121-a)). After the verb has been merged with x, VP optimization takes place. (Note that in these examples, x is not moved to the edge of the phrase as it lacks the [β]-feature.) If *REFLMARK$_{inh}$ is now ranked between FAITH$_{SE}$ and FAITH$_{SELF}$, the matrix [SE, pron] is predicted to be optimal (cf. T_{26}) – and on the assumption that the optimal realization is based on the most anaphoric specification that is left in the optimal matrix, the SE anaphor is chosen as optimal realization of x.

(124) Max benimmt sich.

a. [$_{\text{VP}}$ benimmt x]

T_{26}: VP optimization

	Candidates	F_{pron}	F_{SE}	*REFLMARK$_{inh}$	F_{SELF}
	O_1: [SELF, SE, pron]			$*\!*\!*!$	
⇒	O_2: [SE, pron]			$**$	$*$
	O_3: [pron]		$*!$	$*$	$*$
	O_4: [–]	$*!$	$*$		$*$

Note that the last step – from the optimal matrix to the optimal realization – cannot be directly derived from the MAB principle as formulated in (45) (repeated in (125)). In order to make it compatible with derivations involving inherently reflexive predicates, the formulation must be modified in such a way that it does not necessarily presuppose a [β]-feature on x; cf. (126).

(125) MAXIMALLY ANAPHORIC BINDING (MAB) (repeated from (45)):
Checked $x_{[\beta]}$ must be realized maximally anaphorically.

(126) MAXIMALLY ANAPHORIC BINDING (MAB) (revised):
When all [β]-features of x are checked, it is realized maximally anaphorically.

In Frisian, inherently reflexive predicates occur with pronouns (cf. (127), repeated from (121-c)). This is correctly predicted if *REFLMARK$_{inh}$ is higher ranked than $\text{F}\textsc{aith}_{SE}$ but lower ranked than $\text{F}\textsc{aith}_{pron}$ (cf. T_{27}). On this assumption, [pron] is predicted to be the optimal matrix, and MAB selects the pronominal form as optimal realization.

(127) Max hâld him/*himsels.

a. $[_{\text{VP}}$ hâld $x]$

T_{27}: VP optimization

	Candidates	F_{pron}	$*\text{R}\textsc{efl}\text{M}\textsc{ark}_{inh}$	F_{SE}	F_{SELF}
	O_1: [SELF, SE, pron]		∗∗!∗		
	O_2: [SE, pron]		∗∗!		∗
⇒	O_3: [pron]		∗	∗	∗
	O_4: [–]	∗!		∗	∗

The English example in (128) (repeated from (122-a)) lacks any realization of x. This is captured if $*\text{R}\textsc{efl}\text{M}\textsc{ark}_{inh}$ outranks all $\text{F}\textsc{aith}$-constraints, as T_{28} shows.

(128) Max behaves like a gentleman.

a. $[_{\text{VP}}$ behaves $x]$

T_{28}: VP optimization

	Candidates	$*\text{R}\textsc{efl}\text{M}\textsc{ark}_{inh}$	F_{pron}	F_{SE}	F_{SELF}
	O_1: [SELF, SE, pron]	∗!∗∗			
	O_2: [SE, pron]	∗!∗			∗
	O_3: [pron]	∗!		∗	∗
⇒	O_4: [–]		∗	∗	∗

By contrast, if the SELF anaphor occurs with inherently reflexive predicates, as in (129) (repeated from (122-b)), $*\text{R}\textsc{efl}\text{M}\textsc{ark}_{inh}$ must be lower ranked than the $\text{F}\textsc{aith}$-constraints. On this assumption, O_1 wins the competition (cf. T_{29}), and MAB selects the complex anaphor as optimal realization of x.

(129) Max behaves himself.

a. $[_{\text{VP}}$ behaves $x]$

T_{29}: VP optimization

	Candidates	F_{pron}	F_{SE}	F_{SELF}	*ReflMark$_{inh}$
⇒	O_1: [SELF, SE, pron]				$***$
	O_2: [SE, pron]			$*!$	$**$
	O_3: [pron]		$*!$	$*$	$*$
	O_4: [–]	$*!$	$*$	$*$	

To sum up, the crosslinguistic variation we find with respect to inherently reflexive predicates is derived by different interaction between the constraint *ReflMark$_{inh}$ (=*RM$_{inh}$) and the Faith-constraint subhierarchy. The respective predictions are summarized in the following table.

T_{30}: Summary

ranking	*realization of x with inherently reflexive predicates*
$F_{pron} \gg F_{SE} \gg F_{SELF} \gg$ ***RM**$_{inh}$	SELF anaphor
$F_{pron} \gg F_{SE} \gg$ ***RM**$_{inh}$ $\gg F_{SELF}$	SE anaphor
$F_{pron} \gg$ ***RM**$_{inh}$ $\gg F_{SE} \gg F_{SELF}$	pronoun
***RM**$_{inh}$ $\gg F_{pron} \gg F_{SE} \gg F_{SELF}$	$\emptyset$

5.12 Pronouns without Antecedents

Talking about contexts in which anaphoric and pronominal forms seem to occur unbound, let us now pursue the question of how examples of the following type can be derived in this system. (130-a) and (130-b) show that x cannot be realized as anaphor if it lacks an antecedent. (Recall that (130-a) rules out the possibility that the ungrammaticality of anaphors in these examples is connected with the fact that German simply lacks Nominative anaphoric forms.) Moreover, we have to say something about the relation between (130-b) and (130-c).

(130) a. Ihn/*sich friert.
him/SE is cold
'He is cold.'

b. Er schläft.
he sleeps
'He is sleeping.'

c. Peter schläft.
Peter sleeps
'Peter is sleeping.'

Let us start with the latter. (130-c) contains an unbound R-expression; hence we can conclude that the numeration does not contain any x at all, but simply looks as follows: Num$_c$= {Peter (= 'genuine R-expression'), schläft} (ignoring additional functional material). By contrast, if we assume that pronouns generally emerge as the result of a competition between different realization matrices, the underlying numeration in (130-b) corresponds to Num$_b$= $\{x_{[SELF,SE,pron]}$, schläft} – since the sentence does not contain a coindexed R-expression, the matrix lacks a potential copy of it. Hence, the two sentences in (130-b) and (130-c) are based on completely different numerations and do not compete at all. This explains why they are basically interchangeable.

The restriction 'basically' refers to the fact that – although both (130-b) and (130-c) are grammatical – their distribution is dependent on the broader context. If people are talking about Peter anyway, it is more natural to utter (130-b), while (130-c) might sound redundant; however, if Peter has not been mentioned before, it is odd to use the pronominal form. Hence, it can be concluded that although sentence (130-b) does not contain an antecedent, the pronoun must be anchored in discourse, i.e., it must be discourse-bound, and it is therefore not really true that (130-b) and (130-a) contain unbound pronouns. So let us assume that the x in these sentences is also equipped with a [β]-feature and that it is checked by the head in the root phrase if discourse binding is involved. Thus, the numeration of (130-b) contains, *inter alia*, $x_{[\beta]}$ and $C_{[*\beta*]}$. Against this background, let us take a closer look at the derivation of (130-b) (repeated in (131)).

(131) Er schläft.

a. $[_{\mathrm{vP}}\ x_{[\beta]}\ [_{\mathrm{VP}}\ \mathrm{t}_{schl\ddot{a}ft}]$ schläft]; workspace: $\{C_{[*\beta*]}, \ldots\}$

x is merged into the derivation in the second phrase; at this stage, the verb is also accessible, hence we reach x's θ-domain; since x is Case-marked by T, vP does not correspond to its subject and Case domain. However, considering how finite and indicative domain have been defined, it is suggested that vP already fulfils these definitions since it contains a finite/indicative verb and a subject. So if we want to stick to the assumption that the finite and indicative domain are not reached before the subject and Case domain, their definitions have to be slightly modified; therefore, the following revised versions are introduced.

(132) XP is the *finite domain* of x if it contains a finite verb and Case-marked x.

(133) XP is the *indicative domain* of x if it contains an indicative verb and Case-marked x.

According to these definitions, TP is the first XP which qualifies as x's finite/indicative domain in sentence (131), and it is correctly predicted that x must be realized as a pronoun: When TP is optimized, $x_{[\beta]}$ is still unchecked, and since TP not only corresponds to x's θ- and Case domain but also to its subject, finite, and indicative domain, the high-ranked constraints PRINCIPLE $\mathcal{A}_{SD}$, PRINCIPLE $\mathcal{A}_{FD}$, and PRINCIPLE $\mathcal{A}_{ID}$ apply non-vacuously; as a result, [pron] is predicted to be optimal (cf. $T_{31.1}$ and $T_{31.2}$).[55] In the next phrase, $C_{[*\beta*]}$ finally enters the derivation and x is checked. Thus, only the FAITH-constraints apply non-vacuously in $T_{31.1.1/31.2.1}$, but since [pron] is the only candidate anyway, it remains optimal, and MAB selects the pronoun as optimal realization.

[55] Here, we can clearly see that $[*\beta*]$ must be associated with C and not with T if we consider discourse binding. If x functions as subject, only its θ-domain and an XP have been reached before TP is completed – if x were already checked at this stage, this would have the consequence that PR.$\mathcal{A}_{XP}$ and PR.$\mathcal{A}_{ThD}$ would be the only PR.$\mathcal{A}$-constraints that would apply non-vacuously before the realization of x would be determined. However, since these two constraints are relatively low ranked, the matrix would not have been reduced to [pron] and an anaphoric realization would be predicted to be optimal. (Recall from section 5.7, T_{18}, that we can only avoid anaphoric specifications if a PR.$\mathcal{A}$-constraint applies non-vacuously which is higher ranked than FAITH$_{SE}$.)

T_{31}: VP optimization
(XP/ThD reached – $x_{[\beta]}$ unchecked)

	Candidates	F_{pron}	F_{SE}	$\text{Pr.}\mathcal{A}_{ThD}$	F_{SELF}	$\text{Pr.}\mathcal{A}_{XP}$
⇒	O_1: [SELF, SE, pron]			∗∗(!)		∗∗
⇒	O_2: [SE, pron]			∗	∗(!)	∗
	O_3: [pron]		∗!		∗	

(134) b. $[_{\text{TP}}\ x_{[\beta]}\ [_{\text{vP}}\ \text{t}_x\ \cancel{[_{\text{VP}}\ \text{t}_{schlaft}]}\ \text{t}'_{schläft}]\ \text{schläft}]$;
workspace: $\{\text{C}_{[*\beta*]}, \ldots\}$

$T_{31.1}$: TP optimization
(XP/ThD/CD/SD/FD/ID reached – $x_{[\beta]}$ unchecked)

Input: O_1/T_{31}	F_{pron}	$\text{Pr.}\mathcal{A}_{ID/FD/SD}$	F_{SE}	$\text{Pr.}\mathcal{A}_{CD}$	$\text{Pr.}\mathcal{A}_{ThD}$	F_{SELF}	$\text{Pr.}\mathcal{A}_{XP}$
O_{11}: [S, S, pr]		∗!∗		∗∗	∗∗		∗∗
O_{12}: [SE, pr]		∗!		∗	∗	∗	∗
⇒ O_{13}: [pron]			∗			∗	

$T_{31.2}$: TP optimization
(XP/ThD/CD/SD/FD/ID reached – $x_{[\beta]}$ unchecked)

Input: O_2/T_{31}	F_{pron}	$\text{Pr.}\mathcal{A}_{ID/FD/SD}$	F_{SE}	$\text{Pr.}\mathcal{A}_{CD}$	$\text{Pr.}\mathcal{A}_{ThD}$	F_{SELF}	$\text{Pr.}\mathcal{A}_{XP}$
O_{21}: [SE, pr]		∗!		∗	∗	∗	∗
⇒ O_{22}: [pron]			∗			∗	

(135) c. $[_{\text{CP}}\ x_{[\beta]}\ [_{\text{C}_{[*\beta*]}}\ \text{schläft}]\ [_{\text{TP}}\ \text{t}'_x\ \cancel{[_{\text{vP}}\ \text{t}_x\ [_{\text{VP}}\ \text{t}_{schlaft}]\ \text{t}'_{schlaft}]}\ \text{t}''_{schläft}]]$

$T_{31.1.1/31.2.1}$: CP optimization
($x_{[\beta]}$ checked: PRINCIPLE $\mathcal{A}_{XD}$ applies vacuously)

	Input: $O_{13}/T_{31.1}$ or $O_{22}/T_{31.2}$	F_{pron}	F_{SE}	F_{SELF}
⇒	$O_{131/221}$: [pron]		∗	∗

5.13 On the Distribution of the Beta-Features

So far, we have mainly considered examples in which binding relations between two elements have been established (if we abstract away from inherently reflexive predicates and the examples in the previous section). This means that sentences have been excluded which involve elements that are coreferent but do not stand in a c-command relationship (cf. (136)), and sentences in which more than two elements are coreferent (cf. (137) and (138)).

(136) a. Peter$_1$'s sister adores him$_1$/*himself$_1$.
b. His$_1$ sister adores Peter$_1$.

(137) John$_1$ wonders whether he$_1$ should shave himself$_1$/*him$_1$.

(138) John$_1$ only shaves himself$_1$/*him$_1$ in his$_1$ bathroom.

In this section, the question is therefore addressed of how sentences of this type can be derived, and how unwanted derivations resulting from numerations with a different distribution of beta-features can be excluded.

Let us first consider sentence (136-a). What does the underlying numeration look like? As far as the direct object is concerned, we have assumed that pronouns like *him* are encoded as $x_{[\beta]}$ in the beginning, and that its concrete realization form is determined in the course of the derivation. And since in a derivational model we do not know in advance that the coreferent R-expression *Peter* will never c-command x, we might want to try the numeration Num$_1$=$\{$Peter$_{[*\beta*]}$, $x_{[\beta]}$, $\ldots\}$. However, in the course of the derivation it emerges that x is never c-commanded by *Peter*, hence x can never check its [β]-feature (since feature checking requires a c-command relation between probe and goal in the accessible domain (cf. (26) in section 4.2)), and therefore the derivation will eventually crash.

This means that (136-a) must be based on a different numeration. For obvious reasons, the [$*\beta*$]-feature cannot be associated with the NP *Peter's sister* either (after all, it is not coreferent with x) – hence there is only one possibility left: discourse binding. On this assumption, the numeration is Num$_2$=$\{$C$_{[*\beta*]}$, $x_{[\beta]}$, Peter, $\ldots\}$, x eventually checks its [β]-feature against C$_{[*\beta*]}$, and the pronoun is correctly predicted to be the optimal realization form (cf. the previous section). If *Peter* is now coreferent with x or not is not encoded in features but depends on whether *Peter* happens to refer to the same person as "[β]". If it does, we get sentence (136-a) (*Peter*$_1$*'s sister adores him*$_1$), otherwise the following sentence is derived.

(139) Peter$_1$'s sister adores him$_2$.

As to example (136-b), it behaves analogously to (136-a) with *Peter* and x in exchanged positions.

However, the previous examples have also alluded to a first restriction that must be assumed for the distribution of [$*\beta*$]-features. As discussed in section 5.12, discourse binding always implies that the pronominal form is predicted to be the optimal realization. Thus it can be concluded that if the option of discourse binding were generally available, i.e., if [$*\beta*$] could always be associated with matrix C, we would predict that pronominal binding would be a universal option in all binding contexts. Since this prediction is obviously not borne out (cf., for example, *John*$_1$ *likes himself*$_1$*/*him*$_1$), the occurence of $C_{[*\beta*]}$ must be restricted. What would such a restriction look like? Recall that in the analysis of (136-a) and (136-b), the insertion of $C_{[*\beta*]}$ in the numeration was the only possibility to yield a convergent derivation. So let us therefore assume that the following principle holds.

(140) *Restriction on the distribution of [$*\beta*$]-features:*
$C_{[*\beta*]}$ is a last resort option; it is only licit if the association of [$*\beta*$] with a lexical item of the numeration (including underspecified x) does not yield a convergent derivation.

What about sentence (137) (*John*$_1$ *wonders whether he*$_1$ *should shave himself*$_1$*/*him*$_1$)? In principle, we could think of the following six underlying numerations if we take into account all potentially possible distributions of the beta-features.[56]

(141) *Possible distribution of beta-features:*

[56]In the following, I assume that the second coindexed element (linearly speaking) starts out as x and the third one as y.

a. $\text{Num}_1 = \{\text{John}_{[*\beta*]},\ x_{[\beta]},\ y_{[\beta]},\ \ldots\}$
b. $\text{Num}_2 = \{\text{John}_{[*\beta*]},\ x_{[\beta,*\beta*]},\ y_{[\beta]},\ \ldots\}$
c. $\text{Num}_3 = \{\text{John}_{[*\beta*]},\ x_{[\beta]},\ y_{[\beta,*\beta*]},\ \ldots\}$
d. $\text{Num}_4 = \{\text{John}_{[*\beta*,*\beta*]},\ x_{[\beta]},\ y_{[\beta]},\ \ldots\}$
e. $\text{Num}_5 = \{\text{John}_{[*\beta*]},\ x_{[\beta]},\ y_{[\beta]},\ \text{C}_{[*\beta*]}, \ldots\}$
f. $\text{Num}_6 = \{\text{John},\ x_{[\beta]},\ y_{[\beta]},\ \text{C}_{[*\beta*,*\beta*]}\ \ldots\}$

The first numeration, $\text{Num}_1 = \{\text{John}_{[*\beta*]},\ x_{[\beta]},\ y_{[\beta]},\ \ldots\}$, can be ruled out immediately, since it only involves one [$*\beta*$]- but two [β]-features. Hence, one [β]-feature will remain unchecked, and Num_1 must therefore be excluded. (In fact, one of the items with a [β]-feature would not even reach a position in which it could in principle check features against *John*, because there is no need to drag along both x and y to satisfy *Phrase Balance*.)

As to $\text{Num}_2 = \{\text{John}_{[*\beta*]},\ x_{[\beta,*\beta*]},\ y_{[\beta]},\ \ldots\}$, it does not only facilitate a convergent derivation, it also yields the expected results with regard to the predicted realization forms: y is checked by x in its θ-domain, hence y will have to be realized as anaphor, and x is checked by *John* later in the derivation when its realization matrix has already been reduced to [pron] – hence, it is realized as pronoun.[57]

(142) a. *Checking 1:*
[$_{\text{vP}}$ $x_{[\beta,*\beta*]}$ shave [$_{\text{VP}}$ $y_{[\beta]}$ t_{shave} $\cancel{\text{t}_y}$]]
b. *Checking 2:*
[$_{\text{vP}}$ $\text{John}_{[*\beta*]}$ wonders [$_{\text{VP}}$ $x_{[\beta]}$ $\text{t}_{wonders}$ [$_{\text{CP}}$ t''_x whether [$_{\text{TP}}$ t'_x ~~should [$_{\text{vP}}$ t_x shave [$_{\text{VP}}$ y t_{shave} t_y]]]]]]~~

As far as $\text{Num}_3 = \{\text{John}_{[*\beta*]},\ x_{[\beta]},\ y_{[\beta,*\beta*]}\ \ldots\}$ is concerned, there are two possibilities. If the resulting derivation proceeds as indicated in (143), it crashes. Here it is assumed that y remains in edgeV, and x moves on to satisfy *Phrase Balance* and eventually check its [β]-feature against *John*. As a result, the features of y remain unchecked (since self-checking is excluded; cf. footnote 57).

[57]Note that the feature distribution $x_{[\beta,*\beta*]}$ does not facilitate "self-checking" – this is excluded since feature checking requires a c-command relation and the notion of c-command is not reflexive.

(143) *Non-convergent derivation:*

a. [$_{\text{vP}}$ $x_{[\beta]}$ shave [$_{\text{VP}}$ $y_{[\beta,*\beta*]}$ t$_{shave}$ ~~t$_{y}$~~]]; workspace: {John$_{[*\beta*]}$, ...}

b. [$_{\text{vP}}$ John$_{[*\beta*]}$ wonders [$_{\text{VP}}$ $x_{[\beta]}$ t$_{wonders}$ [$_{\text{CP}}$ t$''_{x}$ whether [$_{\text{TP}}$ t$'_{x}$ ~~should [$_{\text{vP}}$ t$_{x}$ shave [$_{\text{VP}}$ $y_{[\beta,*\beta*]}$ t$_{shave}$ t$_{y}$]]]]]~~]]

However, based on Num$_3$ there might be an alternative derivation; if y does not stay in edgeV but moves on to Specv (for instance, because *Phrase Balance* triggers movement to satisfy the needs of another feature [$*$F$*$]), x gets the opportunity to check its [β] feature in an appropriate configuration: under c-command against y's [$*\beta*$]-feature (cf. (144-a) and (144-b)). Afterwards, *Phrase Balance* would force y to move on till it reaches the specifier of the matrix VP, where it could eventually check features with *John* (cf. (144-c)).

(144) *Unwanted derivation:*

a. [$_{\text{vP}}$ $x_{[\beta]}$ shave [$_{\text{VP}}$ $y_{[\beta,*\beta*,F]}$ t$_{shave}$ ~~t$_{y}$~~]]; workspace: {John$_{[*\beta*]}$, X$_{[*F*]}$, ...}

b. *Checking 1:*
[$_{\text{vP}}$ $y_{[\beta,*\beta*,F]}$ $x_{[\beta]}$ shave [$_{\text{VP}}$ t$'_{y}$ t$_{shave}$ ~~t$_{y}$~~]]

c. *Checking 2:*
[$_{\text{vP}}$ John$_{[*\beta*]}$ wonders [$_{\text{VP}}$ $y_{[\beta]}$ t$_{wonders}$ [$_{\text{CP}}$ t$''''_{y}$ whether [$_{\text{TP}}$ t$'''_{y}$ ~~should [$_{\text{vP}}$ t$''_{y}$ x shave [$_{\text{VP}}$ t$'_{y}$ t$_{shave}$ t$_{y}$]]]]]~~]]

Hence, Num$_3$ might yield a convergent derivation – however, it does not yield the correct result. Since in this case x is checked in its base position, it has to be realized as anaphor; and since y is checked when its matrix has been reduced to [pron], we would expect a pronoun in the object position, contrary to the facts. As a result, this derivation has to be excluded. In fact, what seems to go wrong in (144) is that the probe for x is base-generated below the latter and moves across the goal to get into this feature checking configuration. Therefore it must be assumed that the following restriction holds, which finally excludes Num$_3$ as a possible underlying numeration.

(145) [$*F*$] must not move across [F].

And what about Num_4=$\{\mathrm{John}_{[*\beta*,*\beta*]}$, $x_{[\beta]}$, $y_{[\beta]}$, $\ldots\}$? *A priori*, it does not violate any restrictions and yields a convergent derivation. However, it would predict that y has to be realized as a pronoun (since it would be checked when its matrix would have been reduced to [pron]) – and this option must be ruled out. So what might be wrong with the following derivation?

(146) *Unwanted derivation:*
$[_{\mathrm{vP}}$ $\mathrm{John}_{[*\beta*,*\beta*]}$ wonders $[_{\mathrm{VP}}$ $y_{[\beta]}$ $x_{[\beta]}$ $\mathrm{t}_{wonders}$ $[_{\mathrm{CP}}$ t''''_y t''_x whether $[_{\mathrm{TP}}$ t'''_y t'_x should $[_{\mathrm{vP}}$ t''_y t_x shave $[_{\mathrm{VP}}$ t'_y t_{shave} $\mathrm{t}_y]]]]]]]$

What we want to enforce is that y is already checked in its θ-domain, which is only possible if it is checked by x. More generally, if we have more than two coindexed elements in a sentence and the first one (L_1) c-commands all the others, the second one (L_2) all but the first one, the third one (L_3) all but the first and the second one etc., we want to make sure that the beta-features are distributed as follows: $\{\mathrm{L}_{1[*\beta*]}$, $\mathrm{L}_{2[*\beta*,\beta]}$, $\mathrm{L}_{3[*\beta*,\beta]}$, $\ldots \mathrm{L}n_{[\beta]}\}$. This is achieved if we assume that the following restriction holds. According to this rule, the derivation in (145) is already ruled out when the embedded vP is completed.[58]

(147) *Restriction on the cooccurrence of [β]-features:*
Two coreferent (i.e. identical) unchecked [β]-feature must not cooccur in the same accessible domain.

As far as Num_5 ($\{\mathrm{John}_{[*\beta*]}$, $x_{[\beta]}$, $y_{[\beta]}$, $\mathrm{C}_{[*\beta*]}$,$\ldots\}$) and Num_6 ($\{$John, $x_{[\beta]}$, $y_{[\beta]}$, $\mathrm{C}_{[*\beta*,*\beta*]}$ $\ldots\}$) are concerned, they are also ruled out, because they violate (140) – Num_2 already yields a convergent derivation without resorting to discourse binding.

Let us now turn to example (138), repeated in (148).

[58]Note that this principle does not affect configurations as in (i-a); since in this case the [β]-features are not coreferent, they can cooccur in edgev.

(i) Sarah_1 knows that Max_2 thinks she_1 adores him_2.

a. $[_{\mathrm{vP}}$ $y_{[\beta_2]}$ $x_{[\beta_1]}$ adores $[_{\mathrm{VP}}$ t'_y t_{adores} $\cancel{\mathrm{t}_y}]]$; workspace: $\{\mathrm{Sarah}_{[*\beta_1*]}$, $\mathrm{Max}_{[*\beta_2*]}$, $\ldots\}$

(148) John$_1$ (only) shaves himself$_1$/*him$_1$ in his$_1$ bathroom.

Since the sentence involves again three coreferent items, there are in principle again the six potential numerations from (141). The first numeration, Num$_1$=$\{$John$_{[*\beta*]}$, $x_{[\beta]}$, $y_{[\beta]}$, $\ldots\}$ can be excluded along the same lines as before, and we can generally state that each unchecked feature needs a different corresponding starred feature.

The second possibility, Num$_2$=$\{$John$_{[*\beta*]}$, $x_{[\beta,*\beta*]}$, $y_{[\beta]}$, $\ldots\}$, can also be ruled out immediately. Since x can check its $[\beta]$-feature against *John* when it is in SpecV (cf. (149)), it would not have to move any further and would thus never c-command y (which is contained in a vP-adjunct). Hence, it cannot act as a probe for the latter.

(149) $[_{\text{vP}}$ John$_{[*\beta*]}$ shaves $[_{\text{VP}}$ $x_{[\beta]}$ t$_{shaves}$ $\cancel{\text{t}_x}]]$

And what about Num$_3$=$\{$John$_{[*\beta*]}$, $x_{[\beta]}$, $y_{[\beta,*\beta*]}$, $\ldots\}$, where the second $[*\beta*]$-feature is associated with y? In this case, the derivation proceeds as follows: *Phrase Balance* triggers movement of x to the edge of VP. When little vP is built up *John* enters the derivation *before* the PP adjunct is also inserted in Specv. Hence, the first opportunity for x to check its feature involves feature checking with *John* – and not with y.[59] However, this implies that y cannot get rid of its beta-features anymore, and the derivation crashes.

(150) $[_{\text{vP}}$ $y_{[\beta,*\beta*]}$ $[_{\text{v}'}$ $[_{\text{v}'}$ John$_{[*\beta*]}$ shaves $[_{\text{VP}}$ $x_{[\beta]}$ t$_{shaves}$ $\cancel{\text{t}_x}]]$ $[_{\text{PP}}$ t$'_y$ in $\cancel{[_{\text{NP}}\ \text{t}_y\ \text{bathroom}]}]]]$

The fourth possibility is Num$_4$=$\{$John$_{[*\beta*,*\beta*]}$, $x_{[\beta]}$, $y_{[\beta]}$, $\ldots\}$. This attempt is more promising; since both x and y are c-commanded by *John* at some stage in the derivation and there are two $[*\beta*]$-features which trigger movement of the two bound elements to the current accessible domain, the derivation does not crash.[60] Moreover, x is already bound in its θ-domain, hence it is correctly predicted that it must be realized as anaphor, whereas y is only bound when its matrix has been reduced to [pron]. Hence, this numeration yields the correct result. And since Num$_4$ yields a convergent derivation,

[59]Generally, feature checking takes place as soon as possible, i.e., it cannot be delayed.
[60]Note that (145) is respected throughout the derivation; cf. (151).

Num_5 ($\{\text{John}_{[*\beta*]}, x_{[\beta]}, y_{[\beta]}, \text{C}_{[*\beta*]},\dots\}$) and Num_6 ($\{\text{John}, x_{[\beta]}, y_{[\beta]}, \text{C}_{[*\beta*,*\beta*]} \dots\}$) are immediatedly ruled out by (140).

(151) a. *Checking 1:*
$[_{\text{vP}}\ \text{John}_{[*\beta*,*\beta*]}\ \text{shave}\ [_{\text{VP}}\ x_{[\beta]}\ \text{t}_{shaves}\ \cancel{\text{t}_x}]]$

b. *Checking 2:*
$[_{\text{TP}}\ \text{John}_{[*\beta*]}\ \text{T}\ [_{\text{vP}}\ y_{[\beta]}\ [_{\text{vP}}\ \text{t}_{John}\ \text{shaves}\ \cancel{[_{\text{VP}}\ x\ \text{t}_{shaves}\ \text{t}_x}]]\ [_{\text{PP}}\ \cancel{\text{t}_y\ \text{in}\ [_{\text{NP}}\ \text{t}_y\ \text{bathroom}}]]]$

The findings of this section can thus be summarized as follows. First, we saw that the three types of sentences under discussion can be derived within the current theory, namely on the basis of the numerations indicated in (152-a)-(152-c).[61]

(152) a. *Two coreferent elements involved – no c-command relation:*
Peter$_1$'s sister adores him$_1$/*himself$_1$.
Num=$\{\text{C}_{[*\beta*]}, x_{[\beta]}, \text{Peter}, \dots\}$

b. *Three coreferent elements involved – three c-command relations:*
John$_1$ wonders whether he$_1$ should shave himself$_1$/*him$_1$.
Num=$\{\text{John}_{[*\beta*]}, x_{[\beta,*\beta*]}, y_{[\beta]}, \dots\}$

c. *Three coreferent elements involved – two c-command relations:*
John$_1$ only shaves himself$_1$/*him$_1$ in his$_1$ bathroom.
Num=$\{\text{John}_{[*\beta*,*\beta*]}, x_{[\beta]}, y_{[\beta]}, \dots\}$

However, although it is of course crucial to have a derivation that makes correct predictions, it is also important to rule out alternative derivations that might yield unwanted results. Since in a derivational model look-ahead with respect to syntactic structures must be excluded, we cannot *a priori* asso-

[61]Note that the analysis of (152-b) also extends to sentences of the following type – the only difference being that in this case the first binding relation is an instance of discourse binding:

(i) He$_1$ likes himself$_1$.
Num=$\{\text{C}_{[*\beta*]}, x_{[\beta,*\beta*]}, y_{[\beta]}, \dots\}$

ciate beta-features only with elements that will later establish a c-command relation – this would involve knowledge of syntactic structures that we cannot know at the stage when the features are distributed. Hence, we must in principle permit that beta-features might be associated with all kinds of pairs of coreferent elements, even it they will never occur in a c-command relationship. The task is then to rule out independently those derivations that would formally converge but make wrong empirical predictions. And, as has been shown in the discussion above, this can be achieved if we assume that the three restrictions from (140), (145), and (147), repeated in (153-a), (153-b), and (153-c), respectively, hold.[62]

(153) a. *Restriction 1*
$C_{[*\beta*]}$ is a last resort option; it is only licit if the association of $[*\beta*]$ with a lexical item of the numeration (including an unspecified x) does not yield a convergent derivation.

b. *Restriction 2*
$[*\beta*]$ must not move across $[\beta]$.

c. *Restriction 3*
Two coreferent (i.e. identical) unchecked $[\beta]$-feature must not cooccur in the same accessible domain.

References

Ackema, Peter & Ad Neeleman. 1998. Optimal Questions. *Natural Language and Linguistic Theory* 16: 443-490.

Aissen, Judith. 1999. Markedness and Subject Choice in Optimality Theory. *Natural Language and Linguistic Theory* 17: 673-711. Reprinted in *Optimality-Theoretic Syntax*, eds. Géraldine Legendre, Jane Grimshaw & Sten Vikner, 61-96. Cambridge, MA: MIT Press.

Aissen, Judith. 2000. Differential Object Marking: Iconicity vs. Economy.

[62] Recall moreover that the c-command requirement does not have to be stipulated specifically for binding relations; instead, it follows from the general operation of *Feature Checking* (cf. (26) in section 4.2).

Ms., University of California, Santa Cruz. Appeared in *Natural Language and Linguistic Theory* 21: 435-483, 2003.

Anderson, Stephen. 1986. The Typology of Anaphoric Dependencies: Icelandic (and other) Reflexives. In *Topics in Scandinavian Syntax*, eds. Lars Hellan & Kirsti Koch Christensen, 65-88. Dordrecht: Reidel.

Aoun, Joseph & Elabbas Benmamoun. 1998. Minimality, Reconstruction, and PF Movement. *Linguistic Inquiry* 29: 569-597.

Baker, Carl. 1995. Contrast, Discourse Prominence, and Intensification, with Special Reference to Locally Free Reflexives in British English. *Language* 71: 63-101.

Barnes, Michael. 1986. Reflexivisation in Faroese: A Preliminary Survey. *Arkiv för Nordisk Filologie* 101: 95-126.

Barss, Andrew. 1986. *Chains and Anaphoric Dependence.* Doctoral dissertation, MIT, Cambridge, MA.

Bobaljik, Jonathan & Samuel Brown. 1997. Interarboreal Operations: Head Movement and the Extension Requirement. *Linguistic Inquiry* 28: 345-356.

Bonet, Eulália. 1991. *Morphology after Syntax.* Doctoral dissertation, MIT, Cambridge, MA.

Bresnan, Joan & Judith Aissen. 2002. Optimality and Functionality: Objections and Refutations. *Natural Language and Linguistic Theory* 20: 81-95.

Brody, Michael. 1995. *Lexico-Logical Form: A Radically Minimalist Theory.* Cambridge, MA: MIT Press.

Brody, Michael. 2002. On the Status of Representations and Derivations. In *Derivation and Explanation in the Minimalist Program*, eds. Samuel Epstein & T. Daniel Seely, 19-41. Oxford: Blackwell.

Burzio, Luigi. 1989. On the Non-Existence of Disjoint Reference Principles. *Rivista di Grammatica Generativa* 14: 3-27.

Burzio, Luigi. 1991. The Morphological Basis of Anaphora. *Journal of Linguistics* 27: 81-105.

Burzio, Luigi. 1992. On the Morphology of Reflexives and Impersonals. In *Theoretical Analyses in Romance Linguistics (LSRL XIX)*, eds. Christiane Lauefer & Terrell Morgan, 399-414. Amsterdam: Benjamins.

Burzio, Luigi. 1994. Weak Anaphora. In *Paths towards a Universal Grammar: Studies in Honor of Richard S. Kayne*, eds. Guglielmo Cinque, Jan Koster, Luigi Rizzi & Raffaella Zanuttini, 59-84. Washington, DC: Georgetown University Press.

Burzio, Luigi. 1996. The Role of the Antecedent in Anaphoric Relations. In *Current Issues in Comparative Grammar*, ed. Robert Freidin, 1-45. Dordrecht: Kluwer.

Burzio, Luigi. 1998. Anaphora and Soft Constraints. In *Is the Best Good Enough?*, eds. Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis & David Pesetsky, 93-113. Cambridge, MA: MIT Press.

Chomsky, Noam. 1973. Conditions on Transformations. In *A Festschrift for Morris Halle*, eds. Stephen Anderson & Paul Kiparsky, 232-286. New York: Holt, Rinehart and Winston.

Chomsky, Noam. 1980. On Binding. *Linguistic Inquiry* 11: 1-46.

Chomsky, Noam. 1981. *Lectures on Government and Binding.* Dordrecht: Foris.

Chomsky, Noam. 1986a. *Barriers.* Cambridge, MA: MIT Press.

Chomsky, Noam. 1986b. *Knowledge of Language.* New York: Praeger.

Chomsky, Noam. 1993. A Minimalist Program for Linguistic Theory. In *The View from Building 20*, eds. Kenneth Hale & Samuel Jay Keyser, 1-52. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program.* Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist Inquiries: The Framework. In *Step by step*, eds. Roger Martin, David Michaels & Juan Uriagereka, 89-155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001a. Beyond Explanatory Adequacy. Ms., MIT, Cambridge, MA.

Chomsky, Noam. 2001b. Derivation by Phase. In *Ken Hale: A Life in Language*, ed. Michael Kenstowicz, 1-52. Cambridge, MA: MIT Press.

Chomsky, Noam & Howard Lasnik. 1993. The Theory of Principles and Parameters. In *Syntax, vol. I*, eds. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann, 506-569. Berlin: de Gruyter.

Cinque, Guglielmo. 1990. *Types of A'-Dependencies.* Cambridge, MA: MIT

Press.

Dalrymple, Mary. 1993. *The Syntax of Anaphoric Binding.* Stanford: CSLI Publications.

Dictionary of Contemporary English. [3]1995. Langenscheidt-Longman, München.

Embick, David & Rolf Noyer. 2001. Movement Operations after Syntax. *Linguistic Inquiry* 32: 555-595.

Epstein, Samuel, Erich Groat, Ruriko Kawashima & Hisatsugu Kitahara. 1998. *A Derivational Approach to Syntactic Relations.* Oxford: Oxford University Press.

Epstein, Samuel & T. Daniel Seely. 2002. *Derivation and Explanation in the Minimalist Program.* Oxford: Blackwell.

Everaert, Martin. 1986. *The Syntax of Reflexivization.* Dordrecht: Foris.

Fabricius-Hansen, Cathrine & Arnim von Stechow. 1989. Explikative und implikative Nominalerweiterungen im Deutschen. *Zeitschrift für Sprachwissenschaft* 8: 173-205.

Faltz, Leonard. 1977. *Reflexivization: A Study in Universal Grammar.* Doctoral dissertation, University of California, Berkeley.

Fanselow, Gisbert. 1991. *Minimale Syntax.* Habilitation thesis, University of Passau.

Fanselow, Gisbert & Damir Ćavar. 2000. Remarks on the Economy of Pronunciation. In *Competition in Syntax*, eds. Gereon Müller & Wolfgang Sternefeld, 107-150. Berlin: Mouton de Gruyter.

Featherston, Sam & Wolfgang Sternefeld. 2003. The Interaction of Factors in Judgements of Reflexive Structures: Data from Object Coreference in German. In *Arbeiten zur Reflexivierung*, eds. Lutz Gunkel, Gereon Müller & Gisela Zifonun, 25-50. Tübingen: Niemeyer.

Fischer, Silke. 2001. On the Integration of Cumulative Effects into Optimality Theory. In *Competition in Syntax*, eds. Gereon Müller & Wolfgang Sternefeld, 151-173. Berlin: Mouton de Gruyter.

Fischer, Silke. 2002. Reanalyzing Reconstruction Effects. An Optimality-Theoretic Account of the Relation between Pronouns and R-Expressions. In *Proceedings of Console IX*, eds. Marjo van Koppen, Erica Thrift, Erik Jan van der Torre & Malte Zimmermann, 68-81.

http://athena.leidenuniv.nl/letteren/sole.
Fischer, Silke. 2003. Optimale Reflexivierung. In *Arbeiten zur Reflexivierung*, eds. Lutz Gunkel, Gereon Müller & Gisela Zifonun, 51-73. Tübingen: Niemeyer.
Fischer, Silke. 2004. Optimal Binding. *Natural Language and Linguistic Theory* 22: 481-526.
Fox, Danny. 1999. Reconstruction, Binding Theory, and the Interpretation of Chains. *Linguistic Inquiry* 30: 157-196.
Fox, Danny. 2000. *Economy and Semantic Interpretation.* Cambridge, MA: MIT Press.
Frampton, John. 1990. Parasitic Gaps and the Theory of *Wh*-Chains. *Linguistic Inquiry* 21: 49-77.
Freidin, Robert. 1986. Fundamental Issues in the Theory of Binding. In *Studies in the Acquisition of Anaphora*, ed. Barbara Lust, 151-188. Dordrecht: Reidel.
Freidin, Robert. 1994. Generative Grammar: Principles and Parameters Framework. In *The Encyclopedia of Language and Linguistics, vol. III*, eds. Ronald Asher & J.M.Y. Simpson, 1370-1385. Oxford: Pergamon Press.
Giorgi, Alessandra. 1984. Toward a Theory of Long Distance Anaphors: A GB Approach. *The Linguistic Review* 3: 307-361.
Grewendorf, Günther. 1988. *Aspekte der deutschen Syntax.* Tübingen: Narr.
Grewendorf, Günther. 1989. *Ergativity in German.* Dordrecht: Foris.
Grimshaw, Jane. 1990. *Argument Structure.* Cambridge, MA: MIT Press.
Grimshaw, Jane. 1997. Projection, Heads, and Optimality. *Linguistic Inquiry* 28: 373-422.
Haider, Hubert. 1984. Topic, Focus, and V-Second. *Groninger Arbeiten zur Germanistischen Linguistik* 25: 72-120.
Halle, Morris. 2000. Distributed Morphology: Impoverishment and Fission. In *Research in Afroasiatic Grammar*, eds. Jacqueline Lecarme, Jean Lowenstamm & Ur Shlonsky, 125-151. Amsterdam: Benjamins.
Halle, Morris & Alec Marantz. 1993. Distributed Morphology and the Pieces of Inflection. In *The View from Building 20*, eds. Kenneth Hale & Samuel Jay Keyser, 111-176. Cambridge, MA: MIT Press.

Heck, Fabian. 2004. *A Theory of Pied Piping.* Doctoral dissertation, University of Tübingen.

Heck, Fabian & Gereon Müller. 2000. Repair-Driven Movement and the Local Optimization of Derivations. Ms., University of Stuttgart & IDS Mannheim, http://www.ids-mannheim.de/gra/personal/texte/mu8.pdf.

Heck, Fabian, Gereon Müller, Ralf Vogel, Silke Fischer, Sten Vikner & Tanja Schmid. 2002. On the Nature of the Input in Optimality Theory. *The Linguistic Review* 19: 345-376.

Hestvik, Arild. 1991. Subjectless Binding Domains. *Natural Language and Linguistic Theory* 9: 455-496.

Heycock, Caroline. 1995. Asymmetries in Reconstruction. *Linguistic Inquiry* 26: 547-570.

Hornstein, Norbert. 2001. *Move! A Minimalist Theory of Construal.* Oxford: Blackwell.

Huang, C.-T. James. 1993. Reconstruction and the Structure of VP: Some Theoretical Consequences. *Linguistic Inquiry* 24: 103-138.

Jackendoff, Ray. 1977. *X′-Syntax: A Study of Phrase Structure.* Cambridge, MA: MIT Press.

Jackendoff, Ray. 1992. Mme. Tussaud Meets the Binding Theory. *Natural Language and Linguistic Theory* 10: 1-31.

Kawashima, Ruriko & Hisatsugu Kitahara. 1996. Strict Cyclicity, Linear Ordering, and Derivational C-Command. In *The Proceedings of the Fourteenth West Coast Conference on Formal Linguistics*, eds. José Camacho, Lina Choueiri & Maki Watanabe, 255-269. Stanford: CSLI Publications.

Kayne, Richard. 1972. Subject Inversion in French Interrogative. In *Generative Studies in Romance Languages*, eds. Jean Casagrande & Bohdan Saciuk, 70-126. Rowley, MA: Newbury House.

Kayne, Richard. 1984. *Connectedness and Binary Branching.* Dordrecht: Foris.

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, eds. Samuel Epstein & T. Daniel Seely, 133-166. Oxford: Blackwell.

Kiparsky, Paul. 2002. Disjoint Reference and the Typology of Pronouns. In *More than Words. A Festschrift for Dieter Wunderlich*, eds. Ingrid Kaufmann & Barbara Stiebels, 179-226. Berlin: Akademie Verlag.

Kiss, Tibor. 2001. Anaphora and Exemptness. A Comparative Treatment of Anaphoric Binding in German and English. Ms., University of Bochum, http://www.linguistics.ruhr-uni-bochum.de/∼kiss.

Kitahara, Hisatsugu. 1995. Target α: Deducing Strict Cyclicity from Derivational Economy. *Linguistic Inquiry* 26: 47-77.

König, Ekkehard & Peter Siemund. 2000. Intensifiers and Reflexives: A Typological Perspective. In *Reflexives. Forms and Functions*, eds. Zygmunt Frajzyngier & Traci Curl, 41-74. Amsterdam: Benjamins.

Koppen, Marjo van. 2003. First Conjunct Agreement. Handout, CGSW 2003, University of Durham.

Koster, Jan. 1984. Reflexives in Dutch. In *Grammatical Representation*, eds. Jacqueline Guéron, Hans-Georg Obenauer & Jean-Yves Pollock, 141-167. Dordrecht: Foris.

Koster, Jan. 1987. *Domains and Dynasties.* Dordrecht: Foris.

Kroch, Anthony & Sabine Iatridou. 1992. The Licensing of CP-Recursion and its Relevance to the Germanic Verb-Second Phenomenon. *Working Papers in Scandinavian Syntax* 50: 1-24.

Kuno, Susumu. 1987. *Functional Syntax.* Chicago: The University of Chicago Press.

Kuno, Susumu. 1997. Binding Theory in the Minimalist Program. Ms., Harvard University, Cambridge, MA.

Lasnik, Howard. 1991. On the Necessity of Binding Conditions. In *Principles and Parameters in Comparative Grammar*, ed. Robert Freidin, 7-28. Cambridge, MA: MIT Press.

Lasnik, Howard. 1998. Some Reconstruction Riddles. *University of Pennsylvania Working Papers in Linguistics* 5: 83-98.

Lebeaux, David. 1983. A Distributional Difference between Reciprocals and Reflexives. *Linguistic Inquiry* 14: 723-730.

Lebeaux, David. 1988. *Language Acquisition and the Form of the Grammar.*

Doctoral dissertation, University of Massachusetts, Amherst.

Lebeaux, David. 1991. Relative Clauses, Licensing, and the Nature of the Derivation. In *Syntax and Semantics, 25. Perspectives on Phrase Structure: Heads and Licensing*, ed. Susan Rothstein, 209-239. New York: Academic Press.

Lebeaux, David. 2000. *Language Acquisition and the Form of the Grammar.* Amsterdam: Benjamins.

Legendre, Géraldine. 2001. An Introduction to Optimality Theory in Syntax. In *Optimality-Theoretic Syntax*, eds. Géraldine Legendre, Jane Grimshaw & Sten Vikner, 1-27. Cambridge, MA: MIT Press.

Legendre, Géraldine, Paul Smolensky & Colin Wilson. 1998. When is Less More? Faithfulness and Minimal Links in *Wh*-Chains. In *Is the Best Good Enough?*, eds. Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis & David Pesetsky, 249-289. Cambridge, MA: MIT Press.

Maling, Joan. 1984. Non-Clause-Bounded Reflexives in Modern Icelandic. *Linguistics and Philosophy* 7: 211-241.

Manzini, Rita & Kenneth Wexler. 1987. Parameters, Binding Theory, and Learnability. *Linguistic Inquiry* 18: 413-444.

Marantz, Alec. 1984. *On the Nature of Grammatical Relations.* Cambridge, MA: MIT Press.

Marantz, Alec. 2003. Blocking and Morphology. Handout, University of Stuttgart.

Martin, Roger. 1999. Case, the Extended Projection Principle, and Minimalism. In *Working Minimalism*, eds. Samuel Epstein & Norbert Hornstein, 1-25. Cambridge, MA: MIT Press.

May, Robert. 1977. *The Grammar of Quantification.* Doctoral dissertation, MIT, Cambridge, MA.

May, Robert. 1985. *Logical Form: Its Structure and Derivation.* Cambridge, MA: MIT Press.

Menuzzi, Sergio. 1999. *Binding Theory and Pronominal Anaphora in Brazilian Portuguese.* Holland Academic Graphics, Leiden.

Müller, Gereon. 1995. Anti-Rekonstruktion. Handout, University of Tübingen.

Müller, Gereon. 2000. *Elemente der optimalitätstheoretischen Syntax.*

Tübingen: Stauffenburg.

Müller, Gereon. 2001. Harmonic Alignment and the Hierarchy of Pronouns in German. In *Pronouns – Grammar and Representation*, eds. Horst Simon & Heike Wiese, 205-232. Amsterdam: Benjamins.

Müller, Gereon. 2002. Two Types of Remnant Movement. In *Dimensions of Movement*, eds. Artemis Alexiadou, Elena Anagnostopoulou, Sjef Barbiers & Hans-Martin Gärtner, 209-241. Amsterdam: Benjamins.

Müller, Gereon. 2003. Local vs. Global Optimization in Syntax: A Case Study. In *Proceedings of the Workshop on Variation within Optimality Theory*, eds. Jennifer Spenader, Anders Eriksson & Östen Dahl, 82-91. Stockholm University: Institute of Linguistics.

Müller, Gereon. 2004a. A Distributed Morphology Approach to Syncretism in Russian Noun Inflection. Ms., IDS Mannheim. To appear in *Proceedings of FASL 12*, eds. Olga Arnaudova, Wayles Browne, Maria Luisa Rivero & Danijela Stojanovic.

Müller, Gereon. 2004b. Phrase Impenetrability and Wh-Intervention. In *Minimality Effects in Syntax*, eds. Arthur Stepanov, Gisbert Fanselow & Ralf Vogel, 289-325. Berlin: Mouton de Gruyter.

Müller, Gereon & Wolfgang Sternefeld. 1993. Improper Movement and Unambiguous Binding. *Linguistic Inquiry* 24: 461-507.

Newmeyer, Frederick. 2002a. Optimality and Functionality: A Critique of Functionally-Based Optimality-Theoretic Syntax. *Natural Language and Linguistic Theory* 20: 43-80.

Newmeyer, Frederick. 2002b. A Rejoinder to Bresnan and Aissen. *Natural Language and Linguistic Theory* 20: 97-99.

Newson, Mark. 1997. Pronominalisation, Reflexivity and the Partial Pronunciation of Traces: Binding Goes OT. Ms., Eötvös Loránd University at Budapest. Appeared in *The Even Yearbook 3. ELTE SEAS Working Papers in Linguistics*, 1998, ed. László Varga, 173-222.

Noyer, Rolf. 1998. Impoverishment Theory and Morphosyntactic Markedness. In *Morphology and its Relation to Phonology and Syntax*, eds. Steven G. Lapointe, Diane K. Brentari & Patrick M. Farrell, 264-285. Stanford: CSLI Publications.

Nunes, Jairo. 1995. *The Copy Theory of Movement and Linearization of*

Chains in the Minimalist Program. Doctoral dissertation, University of Maryland, College Park.

Nunes, Jairo. 1999. Linearization of Chains and Phonetic Realization of Chain Links. In *Working Minimalism*, eds. Samuel Epstein & Norbert Hornstein, 217-249. Cambridge, MA: MIT Press.

Nunes, Jairo. 2001. Sideward Movement. *Linguistic Inquiry* 32: 303-344.

Pesetsky, David. 1998. Some Optimality Principles of Sentence Pronunciation. In *Is the Best Good Enough?*, eds. Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis & David Pesetsky, 337-383. Cambridge, MA: MIT Press.

Pesetsky, David & Esther Torrego. 2001. T-to-C Movement: Causes and Consequences. In *Ken Hale – A Life in Language*, ed. Michael Kenstowicz, 355-426. Cambridge, MA: MIT Press.

Petersen, Hjalmar P., Jógvan í Lon Jacobsen, Zakaris Svabo Hansen & Höskuldur Thráinsson. 1998. Faroese: An Overview for Students and Researchers. Ms., University of Iceland & Academy of the Faroes. Appeared as: Thráinsson, Höskuldur *et al.* 2003. *Faroese. An Overview and Reference Grammar.* Tórshavn: Fróðskaparfelag Föroya.

Phillips, Colin. 1998. Linear Order and Constituency. Ms., University of Delaware. Appeared in *Linguistic Inquiry* 34: 37-90, 2003.

Pica, Pierre. 1986. De Quelques Implications Théoriques de l'Étude des Relations à Longue Distance. In *La Grammaire Modulaire*, eds. Mitsou Ronat & Daniel Couquaux, 187-209. Paris: Minuit.

Pollard, Carl & Ivan Sag. 1987. *An Information-Based Syntax and Semantics. Volume I: Fundamentals.* Stanford: CSLI Publications.

Pollard, Carl & Ivan Sag. 1992. Anaphors in English and the Scope of Binding Theory. *Linguistic Inquiry* 23: 261-303.

Prince, Alan & Paul Smolensky. 1993. Optimality Theory: Constraint Interaction in Generative Grammar. Ms., Rutgers University & University of Colorado at Boulder. Appeared as: Prince, Alan & Paul Smolensky. 2004. *Optimality Theory: Constraint Interaction in Generative Grammar.* Oxford: Blackwell.

Rappaport, Gilbert. 1986. On Anaphor Binding in Russian. *Natural Language and Linguistic Theory* 4: 97-120.

Reinhart, Tanya. 1976. *The Syntactic Domain of Anaphora.* Doctoral dissertation, MIT, Cambridge, MA.

Reinhart, Tanya. 1979. The Syntactic Domain for Semantic Rules. In *Formal Semantics and Pragmatics*, eds. Franz Guenther & Siegfried Schmidt, 107-130. Dordrecht: Reidel.

Reinhart, Tanya. 1981. Definite NP Anaphora and C-Command Domains. *Linguistic Inquiry* 12: 605-635.

Reinhart, Tanya & Eric Reuland. 1991. Anaphors and Logophors: An Argument Structure Perspective. In *Long-Distance Anaphora*, eds. Jan Koster & Eric Reuland, 283-321. Cambridge: Cambridge University Press.

Reinhart, Tanya & Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24: 657-720.

Reuland, Eric & Martin Everaert. 2001. Deconstructing Binding. In *The Handbook of Contemporary Syntactic Theory*, eds. Mark Baltin & Chris Collins, 634-669. Oxford: Blackwell.

Reuland, Eric & Tanya Reinhart. 1995. Pronouns, Anaphors and Case. In *Studies in Comparative Germanic Syntax*, eds. Hubert Haider, Susan Olsen & Sten Vikner, 241-268. Dordrecht: Kluwer.

Richards, Norvin. 1997. Competition and Disjoint Reference. *Linguistic Inquiry* 28: 178-187.

Riemsdijk, Henk van. 1978. *A Case Study in Syntactic Markedness: The Binding Nature of Prepositional Phrases.* Dordrecht: Foris.

Riemsdijk, Henk van & Edwin Williams. 1981. NP-Structure. *The Linguistic Review* 1: 171-217.

Rizzi, Luigi. 1986. On Chain Formation. In *Syntax and Semantics, 19. The Syntax of Pronominal Clitics*, ed. Hagit Borer, 65-95. New York: Academic Press.

Rizzi, Luigi. 1990. *Relativized Minimality.* Cambridge, MA: MIT Press.

Roberts, Ian. 1997. *Comparative Syntax.* London: Arnold.

Safir, Ken. 1997. Symmetry and Unity in the Theory of Anaphora. In *Atomism and Binding*, eds. Hans Bennis, Pierre Pica & Johan Rooryck, 341-379. Dordrecht: Foris.

Safir, Ken. 1999. Vehicle Change and Reconstruction in A′-Chains. *Linguis-*

tic Inquiry 30: 587-620.

Sauerland, Uli & Paul Elbourne. 2002. Total Reconstruction, PF Movement, and Derivational Order. *Linguistic Inquiry* 33: 283-319.

Schäfer, Florian. 2002. Relativierte Bindungsminimalität in OT: Einfache vs. Serielle Optimierung. In *Arbeiten zur Reflexivierung*, eds. Lutz Gunkel, Gereon Müller & Gisela Zifonun, 189-217. Tübingen: Niemeyer.

Schütze, Carson. 1995. PP Attachment and Argumenthood. In *Papers on Language Processing and Acquisition, MIT Working Papers in Linguistics 26*, eds. Carson Schütze, Jennifer Ganger & Kevin Broihier, 95-151. MIT, Cambridge, MA.

Sells, Peter. 1987. *Backwards Anaphora and Discourse Structure: Some Considerations.* Report CSLI-87-114. Stanford: CSLI Publications.

Smolensky, Paul. 1995. On the Internal Structure of Con, the Constraint Component of UG. Ms., Johns Hopkins University.

Sternefeld, Wolfgang. 1985. Deutsch ohne grammatische Funktionen: Ein Beitrag zur Rektions- und Bindungstheorie. *Linguistische Berichte* 99: 394-439.

Sternefeld, Wolfgang. 1993. Anaphoric Reference. In *Syntax, vol. I*, eds. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann, 940-966. Berlin: de Gruyter.

Sternefeld, Wolfgang. 2000. *Semantic vs. Syntactic Reconstruction.* SfS-Report-02-00, University of Tübingen.

Sternefeld, Wolfgang. 2004. *Syntax. Eine merkmalbasierte generative Analyse des Deutschen.* Tübingen: Stauffenburg.

Sternefeld, Wolfgang & Sam Featherston. 2003. The German Reciprocal *einander* in Double Object Constructions. In *Arbeiten zur Reflexivierung*, eds. Lutz Gunkel, Gereon Müller & Gisela Zifonun, 239-265. Tübingen: Niemeyer.

Stowell, Tim. 1981. *Origins of Phrase Structure.* Doctoral dissertation, MIT, Cambridge, MA.

Thráinsson, Höskuldur. 1979. *On Complementation in Icelandic.* New York: Garland.

Thráinsson, Höskuldur. 1991. Long-Distance Reflexives and the Typology of NPs. In *Long-Distance Anaphora*, eds. Jan Koster & Eric Reuland, 49-75. Cambridge: Cambridge University Press.

Uriagereka, Juan. 1995. Aspects of the Syntax of Clitic Placement in Western Romance. *Linguistic Inquiry* 26: 79-123.

Vikner, Sten. 1985. Parameters of Binder and of Binding Category in Danish. *Working Papers in Scandinavian Syntax* 23: 1-61.

Vikner, Sten. 2001. *Verb Movement Variation in Germanic and Optimality Theory*. Habilitation thesis, University of Tübingen.

Wilson, Colin. 2001. Bidirectional Optimization and the Theory of Anaphora. In *Optimality-Theoretic Syntax*, eds. Géraldine Legendre, Jane Grimshaw & Sten Vikner, 465-507. Cambridge, MA: MIT Press.

Zwart, Jan-Wouter. 2002. Issues Relating to a Derivational Theory of Binding. In *Derivation and Explanation in the Minimalist Program*, eds. Samuel Epstein & T. Daniel Seely, 269-304. Oxford: Blackwell.