Matrix Unloaded:

Binding in a Local Derivational Approach[1]

Abstract:

The central question I address in this paper is a theoretical one: Is it possible to integrate binding into a local derivational syntactic approach, and what would be the theoretical consequences of such an enterprise? The answer to the first question will be positive, and thus I will develop such an approach to binding which takes into account both the crosslinguistic variation we encounter in the field of reflexivity and some universal generalizations that can be observed with respect to binding. The theory I propose basically works as follows: (i) Binding corresponds to feature checking between binder and bindee (=:x). (ii) The concrete realization of x is determined in the course of the derivation in an optimality-theoretic competition. In the beginning, x is equipped with a realization matrix which contains its possible realizations (anaphoric or pronominal specifications). (iii) After the completion of each phrase, optimization takes place and might restrict x's realization matrix depending on the respective language and the domain that has been reached. As a result, anaphoric specifications might be deleted. (iv) When checking takes place, the optimal realization matrix of x is mapped to PF, where the concrete realization of x is finally determined in a post-syntactic process.

1. Introduction

It has been argued in the literature that derivational theories are not only superior to global ones from a conceptual point of view (reduction of complexity; cf.

Chomsky 2000 and subsequent work),[2] but that they are furthermore supported by strong empirical evidence (cf., for example, Epstein et al. 1998, Epstein and Seely 2002, Heck and Müller 2000, Müller 2003, Müller 2004).

In a derivational approach, 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.[3] 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 derivational approach has first been proposed by Chomsky (2000), who introduces the so-called *Phase Impenetrability Condition* (PIC) in order to restrict the accessible domain 'downwards' (figuratively spoken if we think of syntactic trees).[4] The first version he proposes works as follows: the accessible domain is reduced as soon as a phase (= vP or CP) is completed; material below the head of a completed phase is no longer accessible (cf. Chomsky 2000: 108, 2001: 13). Later on he weakens this first version in view of dependencies which extend beyond this domain (cf. Chomsky 2001: 14, where he considers VP-internal Nominative NPs). However, from a conceptual point of view this weakening of the PIC 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 (see below).

In order to overcome the conceptual objections, Müller (2004) 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*. In the following, I will also adopt this version of the PIC.

(1) *Phrase Impenetrability Condition* (PIC):

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.

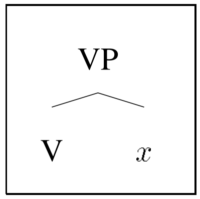
(cf. Müller 2004: 297)

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

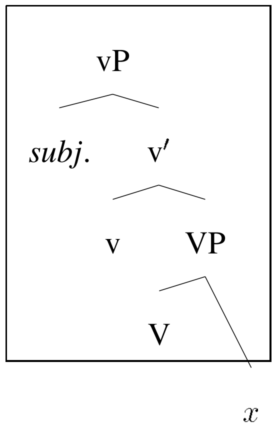
(3) The *edge* of a head X is the residue outside X′; it comprises specifiers and elements adjoined to XP.

The three trees in (4), (5), and (6) illustrate how the accessible domain (which is marked by the frame) shifts when the derivation proceeds: When V and x are merged, the latter would still be accessible if additional material were merged into the SpecV position (cf. (4)). However, as soon as VP is completed, the accessible domain shifts, and when the subject is merged into Specv, x is no longer accessible (cf. (5)). (6) illustrates how the accessible domain is shifted further when the derivation moves on and the next phrase is reached; at this stage, the whole VP (including its edge) has been rendered inaccessible.

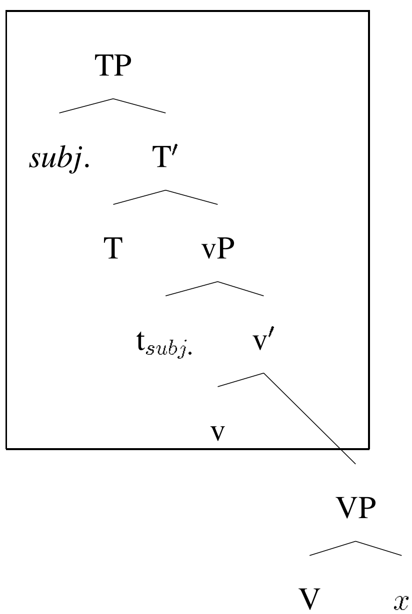
(4) *x still accessible:*



(5) *x no longer accessible:*



(6) *x no longer accessible:*



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.[5]

2. Basic Observations and Former Approaches

There are two reasons why binding seems to pose a problem for a local derivational theory. First, binding is obviously not a strictly local phenomenon, as the following examples show, which illustrate pronominal binding in English and long distance binding in Icelandic respectively.[6]

(7) John$_1$ thinks that Mary likes him$_1$.

(8) Jón$_1$ segir að Pétur raki sig$_1$/hann$_1$/*sjálfan sig$_1$.

John says that Peter shave$_{subj}$ 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 morphologically simple anaphor (henceforth referred to as SE anaphor), as morphologically complex anaphor (henceforth referred to as SELF anaphor), or as pronoun. This is exemplified by the following German sentences, where the bound element becomes less anaphoric the less local the binding relation gets.[7]

(9) *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 more needs to be said.

In fact, there have already been earlier attempts to approach binding within a derivational framework; proposals of this type include Hornstein (2001), Kayne (2002), and Zwart (2002), which share the underlying assumption that an antecedent and its bindee start out as one constituent before the former is moved to a higher position (to be more precise, according to Hornstein's (2001:152) proposal, anaphors *are* "the residues of overt A-movement"; in Kayne's (2002) and Zwart's (2002) theory, antecedent and bindee are merged together in the beginning as separate items).

Apart from theory-internal problems each of the above-mentioned proposals faces (cf. Fischer 2004b for a more detailed discussion), in particular the question of how optionality and crosslinguistic variation can be accounted for remains largely unanswered. Optionality, as in examples like *Max_1 glanced behind himself_1/him_1*, is difficult to capture if the underlying mechanisms for anaphoric/pronominal binding exclude each other. In Hornstein's (2001) analysis, for instance, pronominal binding only emerges as last resort option if movement and thus anaphoric binding is illicit, and Zwart (2002) suggests that all instances of nonaccidental coreference are based on a sisterhood relation between antecedent and bindee, which inevitably yields anaphoric binding, whereas pronominal binding can only be the result of accidental coreference (which does not involve a single constituent in the beginning). Similarly, the status of SE anaphors and consequently also optionality between SE and SELF anaphors remains largely unclarified.

Moreover, the broad range of crosslinguistic variation with respect to binding poses a problem for these proposals. If binding is reduced to movement alone, it is hard to see why, for example, English *Max*$_1$ *glanced behind himself*$_1$/*him*$_1$ translates into German *Max*$_1$ *schaut hinter sich*$_1$/*??sich selbst*$_1$/**ihn*$_1$ and Italian *Max*$_1$ *ha dato un'occhiata dietro di sé*$_1$/**dietro se stesso*$_1$/*?dietro di lui*$_1$ – that is, why languages differ so much with respect to the realization form of their bound element. Furthermore, it is impossible to account for examples which involve binding into islands (as in *Max*$_1$ *will come if he*$_1$ *has to*).[8]

In contrast to the above-mentioned proposals, the approach developed here can account for both crosslinguistic variation (by (restricted) constraint reranking) and optionality (by the use of constraint ties). And although movement will also play a role, it is a special type of movement which might therefore also exhibit different properties (cf. section 3).

3. Binding as Feature Checking

Generally, it can be concluded that in order to be able to evaluate a binding relation, we need to know the exact configuration that holds between the bindee and its antecedent. But since in a local derivational approach, the base position of the bound element might no longer be accessible when the binder is merged into the derivation (cf. (5), for example), x must be dragged along until both elements are accessible at the same time. Thus, the question arises of what triggers movement of the bound element?

If we stick to the assumption that the ultimate goal of all movement operations is feature checking, the solution that presents itself is that binding relations correspond to feature checking relations with the binder as probe which attracts the bound element as goal. And since feature checking must take place in a relatively local configuration, let us assume that the feature which encodes the binding relationship between probe and goal, $[\beta]$, can be checked as soon as the two elements are accessible at the same time. This leaves us with the following configuration for $[\beta]$-feature checking, where XP represents the antecedent and YP the bound element.[9,10]

(10) $[_{\text{ZP}}\ \text{XP}_{[*\beta*]}\ \text{Z}\ [_{\text{WP}}\ \text{YP}_{[\beta]}\ \text{W}]]$

However, before the antecedent eventually enters the derivation, it must be ensured that $x_{[\beta]}$ steadily moves along to the current accessible domain. This means that there are additional movement steps which do not immediately result in feature checking and must therefore be motivated differently. I assume that they are triggered by the following constraint.[11]

(11) *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 2004: 297)

(12) *Potential Availability*:

A feature [F] is potentially available if (i) or (ii) holds:

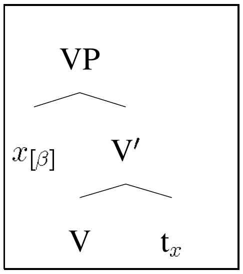
(i) [F] is on X or edgeX of the present root of the derivation.

(ii) [F] is in the workspace of the derivation. (cf. Müller 2004: 298)

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

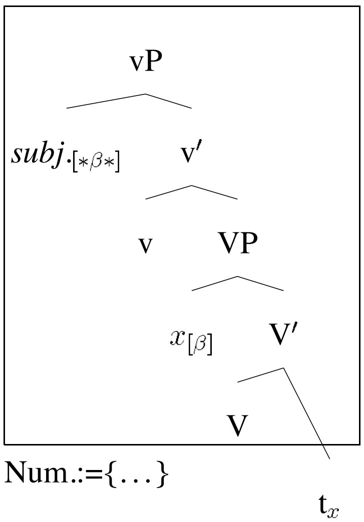
In short, *Phrase Balance* triggers movement of $x_{[\beta]}$ to the edge of the current phrase as long as its antecedent (with the feature [$*\beta*$]) is still in the numeration and thus makes sure that $x_{[\beta]}$ remains accessible. This is illustrated in the following trees. Since *Phrase Balance* forces $x_{[\beta]}$ to move to the edge of VP in (14), $x_{[\beta]}$ is still in the accessible domain at the next derivational stage (cf. (15) and (16)). 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 (15)), $x_{[\beta]}$ stays in its position and feature checking takes place; if the probe remains in the numeration (as in (16)), *Phrase Balance* triggers again movement of $x_{[\beta]}$ to the edge of vP.[12]

(14) *Phrase Balance:*



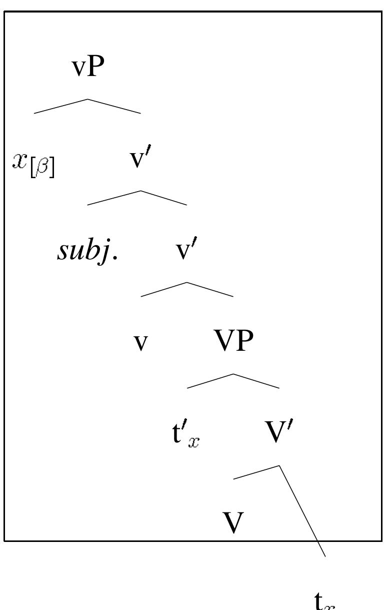
Num.:={subj.$_{[*\beta*]}$, . . . }

(15) *[β]-feature checking:*



Num.:={. . . }

(16) *Phrase Balance:*



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

As alluded to before, the movement x undergoes is special insofar as it is not sensitive to islands. In (17), for instance, the pronoun occurs in an *if*-clause, and its binder *Max* is outside this island, which implies that x has to leave the island in order to be accessible at the same time.

(17) Max_1 will come if he_1 has to.

In fact, the general question arises of how island sensitivity is encoded in a local derivational approach. I assume that it depends on the type of features involved in the movement operation and that feature cooccurrence restrictions might prevent a constituent from raising to the edge of a given phrase since its features might not be allowed in this position. As a result, the constituent in question is rendered inaccessible and cannot leave the phrase (due to the PIC). Thus, the special status of the x-movement has its roots in the features involved – the [β]-features – which are not affected by feature cooccurrence restrictions. However, as indicated in the previous section, examples like (17) pose a severe problem to movement-based approaches which do not differentiate between the type of movement involved in binding relations (here: movement triggered by [β]-features) and other types of movement which are island-sensitive.[13]

4. How to Determine the Optimal Shape of a Bound Element

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.

Following my proposal in Fischer (2003, 2004a, 2004b), I assume that the optimal realization of x is determined in an optimality-theoretic competition in the course of the derivation. In the numeration it is only encoded that there will be a binding relation between the designated antecedent and x. (This is reflected by the features [β]/[$*\beta*$].) 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. The maximal realization matrix of x is thus [SELF, SE, pron].

Following again Fischer (2003, 2004a, 2004b), let us assume moreover that binding is sensitive to domains of different size (cf. also, among others, Manzini and Wexler 1987, Dalrymple 1993). Then we face the following situation: 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. Assume therefore that each time when x reaches one of the domains to which binding is sensitive and x remains unbound, its realization matrix might be reduced in such a way that 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 (cf. the next section as to the technical details).

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.[14] Before Late Insertion takes place (cf. Halle and 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*, 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; that is, syntax itself can be evaluated strictly locally (in contrast to PF – but at this level access to earlier parts of the derivation must be possible anyway since the more deeply embedded material has to be spelled out as well).

Furthermore, 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 only thing that happens is that the lower chain members are specified more precisely in accordance with the predispositions they have already had before.[15] 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."[16]

5. Domains, Constraints, and Candidates

Let us now turn to the technical implementation of the analysis. The theory I propose relies on serial optimization, which means that optimization applies more than once (cf. Müller 2003), namely after the completion of each phrase. The optimal output of a competition serves as input for the next optimization process (cf. Heck and Müller 2000 and subsequent work on derivational OT syntax); the initial input corresponds to the numeration.

The competing candidates differ from each other only with respect to the realization matrix of x. In the beginning, we start with the maximal realization matrix, i.e., in the first competition we have three candidates, O_1, O_2, and O_3, which contain $x_{[SELF,SE,pron]}$, $x_{[SE,pron]}$, and $x_{[pron]}$, respectively. At each step, it is possible to reduce the matrix by deleting anaphoric specifications; however, once the matrix has been reduced, it is not possible to extend it again, since this would constitute a violation of the Inclusiveness Condition.

As far as the domains relevant for binding are concerned, I assume that there are six of them: each maximal projection, the θ-domain, Case domain, subject domain, finite domain, and indicative domain.[17] They are defined as follows:

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

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

(20) XP is the *subject domain* (SD) 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.

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

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

As regards the constraints, there are in principle only two groups of constraints which are ordered in two universal subhierarchies. This makes the system relatively restrictive and does not allow for any arbitrary binding system (cf. also section 7). However, different interleaving of the two hierarchies allows for a straightforward account of crosslinguistic variation.

The first type of constraints is sensitive to the above-mentioned domains and favours pronominal specifications (cf. (23)). These constraints require that x be minimally anaphoric if binding has not yet taken place in XD, where XD refers to one of the six relevant domains.[18]

(23) PRINCIPLE $\mathcal{A}_{XD}$ (PR.$\mathcal{A}_{XD}$):

If $x_{[\beta]}$ remains unchecked in its XD, x must be minimally anaphoric.

The constraints work as follows: If the derivation reaches one of the relevant domains and no binding relation is established, the respective constraint applies non-vacuously and is violated twice by candidate O_1 and once by O_2.

The universal subhierarchy of these constraints is ordered in such a way that constraints referring to bigger domains are higher ranked. This reflects that it is worse if anaphoric x reaches a relatively big domain and is still free.

(24) *Universal subhierarchy 1:*

$$\textsc{Pr}.\mathcal{A}_{ID} \gg \textsc{Pr}.\mathcal{A}_{FD} \gg \textsc{Pr}.\mathcal{A}_{SD} \gg \textsc{Pr}.\mathcal{A}_{CD} \gg \textsc{Pr}.\mathcal{A}_{ThD} \gg \textsc{Pr}.\mathcal{A}_{XP}$$

The second group of constraints comprises the following faithfulness constraints.

(25) a. $\textsc{Faith}_{SELF}$ (F_{SELF}):

The realization matrix of x must contain the specification [SELF].

b. $\textsc{Faith}_{SE}$ (F_{SE}):

The realization matrix of x must contain the specification [SE].

c. $\textsc{Faith}_{pron}$ (F_{pron}):

The realization matrix of x must contain the specification [pron].

Since they must function as counterbalance to the $\textsc{Pr}.\mathcal{A}$-constraints, they must be ordered in such a way that anaphoric realizations are preferred. This is achieved by the ranking in (26), since it favours realization matrices that have not been reduced.[19]

(26) *Universal subhierarchy 2:*

$$\textsc{Faith}_{pron} \gg \textsc{Faith}_{SE} \gg \textsc{Faith}_{SELF}$$

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.

(27) *Maximally Anaphoric Binding* (MAB) (not violable):

Checked $x_{[\beta]}$ must be realized maximally anaphorically.

6. Derivational Binding in German

In this section, I provide an analysis of the German data in (28) (repeated from (9)) to illustrate how the theory works in detail.[20] As far as the data is concerned, these four sentences provide examples involving binding relations of different locality degree. In (28-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 (28-b), the antecedent is not contained in the minimal θ-domain (= embedded vP); it enters the derivation in the matrix vP, which qualifies as Case domain. In (28-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 (28-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.

(28) *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 start with the first example (repeated in (29)). (29-b) illustrates the stage after *Phrase Balance* has triggered movement of x to the edge of VP. The first optimization is represented in T_1.[21]

(29) 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 full matrix is available, hence we have three candidates. Moreover, x is still free, but the only domain that has been reached is XP, thus only $\text{PR}.\mathcal{A}_{XP}$ and the FAITH-constraints apply non-vacuously.[22] As to the ranking of the constraints, the universal hierarchy $\text{FAITH}_{pron} \gg \text{FAITH}_{SE} \gg \text{FAITH}_{SELF}$ 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 $\text{PR}.\mathcal{A}_{XP}$ are tied.[23,24]

T_1: VP optimization

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

Input: $[_{\mathrm{VP}}\ x_{[\beta]/[SELF,SE,pron]}\ [_{\mathrm{V'}}\ \mathrm{t}_x$ hasst]]	F_{pron}	F_{SE}	F_{SELF}	PR.$\mathcal{A}_{XP}$
⇒ O_1: $[_{\mathrm{VP}}\ x_{[\beta]/[SELF,SE,pron]}\ [_{\mathrm{V'}}\ \mathrm{t}_x$ hasst]]				$**(!)$
⇒ O_2: $[_{\mathrm{VP}}\ x_{[\beta]/[SE,pron]}\ [_{\mathrm{V'}}\ \mathrm{t}_x$ hasst]]			$*(!)$	$*$
O_3: $[_{\mathrm{VP}}\ x_{[\beta]/[pron]}\ [_{\mathrm{V'}}\ \mathrm{t}_x$ hasst]]		$*!$	$*$	

When the next phrase is built, the antecedent enters the derivation.[25] Hence, x can check its [β]-feature, i.e., it is bound.

(30) c. $[_{\mathrm{vP}}\ \mathrm{Max}_{[*\beta*]}\ [_{\mathrm{VP}}\ x_{[\beta]}\ [_{\mathrm{V'}}\ \cancel{\mathrm{t}_x}\ \mathrm{t}_{hasst}]]$ hasst]

As a result, the PR.$\mathcal{A}$-constraints apply vacuously and the FAITH-constraints determine the outcome of the next competition.[26]

Since there have been two winners in the previous competition, there are now two possibilities as to how the derivation can proceed. In the competition based on O_1 from T_1, the matrix [SELF, SE, pron] remains optimal; in the alternative derivation based on O_2 from T_1, there are only two candidates left, since matrix reduction has taken place, and hence [SE, pron] is predicted to be optimal. Thus, in the end MAB correctly selects the SELF anaphor as optimal realization in the former and the SE anaphor as optimal form in the latter case.

$T_{1.1}$: vP optimization

($x_{[\beta]}$ checked: PR.$\mathcal{A}_{XD}$ *apply 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: PR.$\mathcal{A}_{XD}$ *apply vacuously)*

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

Let us now turn to example (28-b) (repeated in (31)).[27]

(31) 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 PR.$\mathcal{A}_{XP}$ and PR.$\mathcal{A}_{ThD}$ 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, PR.$\mathcal{A}_{ThD}$ cannot be ranked above FAITH$_{SELF}$; but since the latter is tied with PR.$\mathcal{A}_{XP}$ (cf. T_1) and PR.$\mathcal{A}_{ThD}$ must be universally higher ranked than PR.$\mathcal{A}_{XP}$, it must be assumed that PR.$\mathcal{A}_{ThD}$ and FAITH$_{SELF}$ are also tied.[28] Thus, we get the following partial ranking for German:

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

T_2: 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]		$*!$		$*$	

Since we have again two optimal outputs, there are two competitions when the next phrase is optimized.

(33) b. $[_{\text{VP}}\ x_{[\beta]}\ [_{\text{vP}}\ t_x$ singen] hört]

At this point, no new domain is reached; 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 $T_{2.1}$, and [SE, pron] in $T_{2.2}$.

$T_{2.1}$: 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}	PR.$\mathcal{A}_{ThD}$	F_{SELF}	PR.$\mathcal{A}_{XP}$
⇒	O_{21}: [SE, pron]			$*$	$*$	$*$
	O_{22}: [pron]		$*!$		$*$	

In the next phrase, the binder is merged into the derivation; hence, the PR.$\mathcal{A}$-constraints apply vacuously and the same candidates win as before. 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. $O_{121/211}$). This prediction is again correct.

(34) c. [$_{\text{vP}}$ Max$_{[*\beta*]}$ [$_{\text{VP}}$ $x_{[\beta]}$ ~~[$_{\text{vP}}$ t_x singen]~~ $t_{h\ddot{o}rt}$] hört]

$T_{2.1.1}$: vP optimization

($x_{[\beta]}$ checked: PR.$\mathcal{A}_{XD}$ apply 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: PR.$\mathcal{A}_{XD}$ apply 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 (28-c) is concerned (repeated in (35)), the first optimization step is illustrated in T_3.

(35) Max$_1$ schaut hinter sich$_1$/*sich selbst$_1$/*ihn$_1$.

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

Here, only the SE anaphor is licit in German. As the following tableaux show, this is captured if PR.$\mathcal{A}_{CD}$ is ranked below FAITH$_{SE}$ and above FAITH$_{SELF}$.[29] Due to the fact that the PR.$\mathcal{A}$-constraints are gradient, O$_2$ wins in the first competition (cf. T$_3$), and since [SE, pron] remains optimal in the subsequent optimizations (cf. T$_{3.1}$/T$_{3.1.1}$), 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]		$*!$			$*$	

(36) b. [$_{VP}$ $x_{[\beta]}$ [$_{PP}$ t$_x$ hinter ~~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]		$*!$			$*$	

(37) c. [$_{\text{vP}}$ Max$_{[*\beta*]}$ [$_{\text{VP}}$ $x_{[\beta]}$ ~~[$_{\text{PP}}$ t$_x$′ hinter t$_x$]~~ t$_{schaut}$] schaut]

$T_{3.1.1}$: vP optimization

($x_{[\beta]}$ checked: PR.$\mathcal{A}_{XD}$ *apply vacuously)*

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

The analysis of example (28-d) (repeated in (38)) is illustrated in the tableaux T_4-$T_{4.2}$.

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

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

When the first optimization takes place, only PR. $\mathcal{A}_{XP}$ 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 are met at this stage, and therefore *all* PR.$\mathcal{A}$-constraints apply non-vacuously.

On the assumption that PR.$\mathcal{A}_{ID}$, PR. $\mathcal{A}_{FD}$, and PR.$\mathcal{A}_{SD}$ (in a word, PR.$\mathcal{A}_{ID/FD/SD}$) are ranked above FAITH$_{SE}$, only the candidates with the maximally reduced matrix [pron] win in $T_{4.1}$ and $T_{4.2}$.[30]

(39) 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}	ID/FD/SD	F_{SE}	CD	ThD	F_{SELF}	XP
	O_{11}: [SELF, SE, pron]		$*!*$		$**$	$**$		$**$
	O_{12}: [SE, pron]		$*!$		$*$	$*$	$*$	$*$
⇒	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}	ID/FD/SD	F_{SE}	CD	ThD	F_{SELF}	XP
	O_{21}: [SE, pron]		$*!$		$*$	$*$	$*$	$*$
⇒	O_{22}: [pron]			$*$			$*$	

As a result, x will have to be realized as 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.

7. General Predictions

Ideally, a theory of binding does not only account for the binding patterns of a given 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 then 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 PR.$\mathcal{A}$-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 finally selected as optimal realization. On the other hand, if the binding relation is less local, more PR.$\mathcal{A}$-constraints apply non-vacuously. Hence, it is 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.

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.

Although it would have been beyond the scope of this paper to consider more languages in detail, we can nevertheless see the effects of different interactions between the two universal constraint subhierarchies in general. The different possibilities are represented in table T_5, which illustrates which rankings yield which realization forms. In German, for instance, we thus get the following constraint order (– on the basis of the data from the previous section):

(40) *German ranking:*

$\textsc{Faith}_{pron} \gg \textsc{Pr}.\mathcal{A}_{ID} \gg \textsc{Pr}.\mathcal{A}_{FD} \gg \textsc{Pr}.\mathcal{A}_{SD} \gg \textsc{Faith}_{SE} \gg \textsc{Pr}.\mathcal{A}_{CD}$
$\gg \textsc{Faith}_{SELF} \circ (\textsc{Pr}.\mathcal{A}_{ThD} \gg \textsc{Pr}.\mathcal{A}_{XP})$

T_5: General predictions

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

All this shows that the theory outlined above is both restrictive enough to capture universal binding properties and flexible enough to account for crosslinguistic variation and optionality. Hence, I hope to have shown that binding phenomena do not provide an argument against a local derivational theory of syntax: As the theory developed here shows, a local derivational approach to binding is possible, even if the exact role of post-syntactic realization might require some further discussion.

Notes

[1] For comments and discussion I would like to thank Artemis Alexiadou, Susann Fischer, Fabian Heck, Gunnar Hrafn Hrafnbjargarson, Gereon Müller, Florian Schäfer, Wolfgang Sternefeld, two anonymous *Linguistics* reviewers, and audiences at the Zentrum für Allgemeine Sprachwissenschaft, Berlin (May 2003, OT-SYNTAX^{+} workshop), and at the universities of Tübingen (May 2003), Köln (May 2003, GGS meeting), Durham (September 2003, CGSW 18) and Nijmegen (October 2003, WOTS 7).

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

[3] 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.

[4] As far as the general idea is concerned that operations are restricted to some local domain, cf. also earlier work like van Riemsdijk (1978) and Koster (1987).

[5] Recall that the main goal here is to explore the consequences of such an enterprise in general (namely to integrate binding into a local derivational approach), because if dependencies of this sort could not be captured in such a framework, the local derivational approach as such – though conceptually attractive – would turn out to be problematic from an empirical point of view. It is therefore not the aim of this paper to compare the present approach with non-derivational alternatives which might capture the data as well.

[6] 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 (5).

[7] I assume that SELF anaphors are more anaphoric than SE anaphors, which are in turn more anaphoric than pronouns.

[8] Maybe this is one reason why Zwart (2002) assumes that pronominal binding is completely different from anaphoric binding insofar as only accidental coreference involves pronominal forms. However, even if there are examples that might involve accidental coreference, sentences like (i-a) clearly involve binding: as (i-b) shows, the replacement of the R-expression *Max* with a quantificational phrase is possible, hence we really face an instance of binding.

(i) a. Max$_1$ will come if he$_1$ has to.

b. Every student$_1$ will come if he$_1$ has to.

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

[10] Cf. Chomsky's (2000) notion of Agree and proposals in subsequent work as far as the idea is concerned that feature checking only requires some sufficiently local configuration and not necessarily a spec-head relation.

[11] Chomsky (2000, 2001) assumes that intermediate movement steps of this sort are triggered by optional EPP features which are inserted if "they have an effect on outcome" (Chomsky 2001: 34). However, Müller (2004) argues that this requirement cannot be checked locally and is thus not compatible with a local

derivational approach – in contrast to *Phrase Balance*.

[12] Note that theoretical considerations alone have led to the conclusion that movement operations of this type are indispenable. However, keep also in mind the ultimate goal of this paper: The initial question was whether binding can be captured adequately in a local derivational syntactic framework. The answer I will provide is yes, if we adopt the procedures outlined in this and the following sections – whether we want to do this eventually or not is a different question.

[13] Note that this movement does not affect the meaning, i.e. it is not interpreted semantically; instead, x is interpreted in its base position. As a *Linguistics* reviewer suggests, x should therefore have the more complex form $[x]_x$, where x is a variable and the index x is a λ-operator.

[14] 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.

[15] For example, if the syntactic derivation yields the chain CH=($x_{[SE,pron]}$, $x_{[SELF,SE,pron]}$), its tail takes on the specification of the head at PF; hence, the resulting chain is CH=($x_{[SE,pron]}$, $x_{[SE,pron]}$); and since the most anaphoric specification determines the realization form, x would have to be spelled out as SE-anaphor in this example. What I neglect in this paper is the concrete formulation of spell-out rules which determine which member of a chain is finally spelled out. I assume that this hinges again on the features that are involved in the underlying movement operation and thus the creation of the chain; if [β]-features alone are

responsible for the movement operation, the affected constituent is apparently not spelled out in its target but rather in its base position.

[16] These are quotations taken 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).

[17] The differentiation between subject, finite, and indicative domain will not play a role in the examples discussed below, but it is crucial if long distance binding is addressed along the same lines (cf. Fischer 2003, 2004a, 2004b): The definitions in (20)-(22) make it possible to distinguish between binding into an infinitival complement clause (in which case x is still free (i.e. unchecked) in its SD, but checked when FD and ID are reached), binding into a subjunctive complement (x is still free in its SD=FD, but checked when ID is reached), and binding into an indicative complement (in which case all three domains coincide and x is still unchecked when they are reached).

[18] I refer to these constraints as PR.$\mathcal{A}$-constraints, because they are violated if x is still free in a more or less local domain, which is reminiscent of the traditional Principle A of Chomsky's (1981) Binding Theory.

[19] Maximal realization matrices do not violate any of the FAITH-constraints, but the more specifications are deleted, the more (higher-ranked) FAITH-constraints are violated.

[20] Since the focus of this paper is on the theoretical question of whether binding can be integrated into a local derivational framework, I refrain from analysing

further languages in detail for reasons of space; for an analysis of English, Dutch, Italian, and Icelandic along these lines cf. Fischer 2004b.

[21] In the subsequent tableaux, the candidates are abbreviated and only the different realization matrices are represented.

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

[23] All ties in this analysis are global ties. A global tie X $\circ$ Y stands for the two underlying constraint orders X $\gg$ Y and Y $\gg$ X.

[24] Strictly speaking, it cannot yet be excluded that the crucial ranking is FAITH$_{SELF}$ $\gg$ PR.$\mathcal{A}_{XP}$; 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, and MAB would wrongly predict that only the complex anaphor is licit in sentences like these. In fact, this scenario can be observed in Dutch (cf. (i)), which must therefore involve the ranking FAITH$_{SELF}$ $\gg$ PR.$\mathcal{A}_{XP}$.

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

Max hates himself/SE/him
'Max$_1$ hates himself$_1$.'

[25] Material that is no longer accessible is crossed out.

[26] This 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. The derivational history of the can-

didates 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.

[27] In fact, reflexivization in German A.c.I.-constructions is a more complex issue if the bound element does not function as subject of the embedded clause and occurs inside a PP (cf., for instance, Reis 1976, Gunkel 2003, and the references cited there).

As expected, if the embedded subject functions as antecedent, the bound element must be realized as anaphor and cannot be pronominal (cf., for example, (i) with index 2). However, if the binder is the matrix subject, the data are not uniform at all – reflexivization is sometimes possible despite the intervening subject (cf. (i) /index 1; in fact, for me, the anaphor sounds much better), sometimes excluded (cf. (ii)), and sometimes even obligatory (cf. (iii)).

(i) Der König$_1$ ließ die Leute$_2$ für sich$_{1,2}$/ihn$_1$/*sie$_2$ arbeiten.

the king let the people for SE/him/them work
'The king$_1$ made the people work for him$_1$.'

(cf. Gunkel 2003: 126; cf. also Reis 1976: 27)

(ii) Nur mit Unbehagen ließ Fritz$_1$ den Reporter aus ihm$_1$/*sich$_1$ einen

only with uneasiness let Fritz the reporter out of him/SE a

Helden machen.

hero make

'Only with a feeling of uneasiness did Fritz$_1$ allow the reporter to make a hero of him$_1$.'

(cf. Reis 1976: 31)

(iii) Der König$_1$ ließ den Gefangenen vor sich$_1$/?*ihm$_1$ niederknien.

the king let the prisoner in front of SE/him kneel down

'The king$_1$ made the prisoner kneel down in front of him$_1$.'

(cf. Reis 1976: 27)

As it stands, the theory outlined above would generally predict a pronominal realization for the binding relations involving the matrix subject. (Due to the intervening embedded subject, the subject domain would be reached before x is checked.) However, since the factors responsible for this variation are not really clear to me, I will neglect these examples.

[28] I assume that ties are not transitive (cf. Fischer 2001). The brackets in the ranking in (32) indicate that although both PR.$\mathcal{A}$-constraints are tied with FAITH$_{SELF}$, the dominance relation between them is not given up. Thus, (32) 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$_{SELF}$

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

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

[29] The optionality in English examples of this type (like *Max*$_1$ *glanced behind himself*$_1$/*him*$_1$) is again attributed to an underlying constraint tie. However, it must

be pointed out that optionality between SELF anaphor and pronoun can only arise since English lacks SE anaphors completely. The crucial ranking in this case is $\text{P}\textsc{r}.\mathcal{A}_{CD} \circ \text{F}\textsc{aith}_{SELF}$, which yields the two optimal matrices [SELF, SE, pron] and [SE, pron]. Thus, MAB selects the complex anaphor as realization form in the former case and would select the simple anaphor in the latter case if such a form were available – however, this is not the case, and hence the pronominal form is chosen since it is the most specified form that is compatible with the matrix [SE, pron]. (This corresponds to an application of the Subset Principle.)

[30] For reasons of space I use the abbreviation XD for the constraints $\text{P}\textsc{r}.\,\mathcal{A}_{XD}$ in the following tableaux.

References

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*, Samuel D. Epstein and T. Daniel Seely (eds.), 19-41. Oxford: Blackwell.

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

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

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

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

Epstein, Samuel D.; Groat, Erich M.; Kawashima, Ruriko; and Kitahara, Hisatsugu (1998). *A Derivational Approach to Syntactic Relations*. Oxford: Oxford University Press.

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

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

Fischer, Silke (2003). Optimale Reflexivierung. In *Arbeiten zur Reflexivierung*, Lutz Gunkel, Gereon Müller, and Gisela Zifonun (eds.), 51-73. Tübingen: Niemeyer.

Fischer, Silke (2004a). Optimal Binding. *Natural Language and Linguistic Theory* 22(3), 481-526.

Fischer, Silke (2004b). Towards an Optimal Theory of Reflexivization. Unpublished doctoral dissertation, University of Tübingen.

Gunkel, Lutz (2003). Reflexivierung in AcI-Konstruktionen. In *Arbeiten zur Reflexivierung*, Lutz Gunkel, Gereon Müller, and Gisela Zifonun (eds.), 115-133. Tübingen: Niemeyer.

Halle, Morris and Marantz, Alec (1993). Distributed Morphology and the Pieces of Inflection. In *The View from Building 20*, Kenneth Hale and Samuel J.

Keyser (eds.), 111-176. Cambridge, MA: MIT Press.

Heck, Fabian and Müller, Gereon (2000). Repair-Driven Movement and the Local Optimization of Derivations. Ms., Universität Stuttgart and IDS Mannheim, http://www.uni-leipzig.de/∼muellerg/mu8.pdf.

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

Kayne, Richard (2002). Pronouns and their Antecedents. In *Derivation and Explanation in the Minimalist Program*, Samuel D. Epstein and T. Daniel Seely (eds.), 133-166. Oxford: Blackwell.

Koster, Jan (1987). *Domains and Dynasties*. Dordrecht: Foris.

Manzini, Rita and Wexler, Kenneth (1987). Parameters, Binding Theory, and Learnability. *Linguistic Inquiry* 18(3), 413-444.

Müller, Gereon (2003). Local vs. Global Optimization in Syntax: A Case Study. In *Proceedings of the Workshop on Variation within Optimality Theory*, Jennifer Spenader, Anders Eriksson, and Östen Dahl (eds.), 82-91. Stockholm University: Institute of Linguistics.

Müller, Gereon (2004). Phrase Impenetrability and Wh-Intervention. In *Minimality Effects in Syntax*, Arthur Stepanov, Gisbert Fanselow, and Ralf Vogel (eds.), 289-325. Berlin: Mouton/de Gruyter.

Reis, Marga (1976). Reflexivierung in deutschen A.c.I.-Konstruktionen. Ein transformationsgrammatisches Dilemma. *Papiere zur Linguistik* 9, 5-82.

Riemsdijk, Henk van (1978). *A Case Study in Syntactic Markedness: The Binding Nature of Prepositional Phrases*. Dordrecht: Foris.

Sternefeld, Wolfgang (2004). *Syntax. Eine merkmalbasierte generative Analyse des Deutschen*. Tübingen: Stauffenburg.

Zwart, Jan-Wouter (2002). Issues Relating to a Derivational Theory of Binding. In *Derivation and Explanation in the Minimalist Program*, Samuel D. Epstein and T. Daniel Seely (eds.), 269-304. Oxford: Blackwell.