The configurational matrix[*]

AD NEELEMAN AND HANS VAN DE KOOT

Abstract

In this paper we explore what primitives of syntax may explain why grammatical relations are obligatory, unique, local and sensitive to c-command. We propose that these properties follow from the way information percolates in syntactic trees (as regulated by compositionality) and the way information in nodes is organized (as regulated by set theory).

1 Introduction

Grammatical relations – binding, movement, predication and the licensing of negative polarity items – display a cluster of properties that Koster (1987) refers to as the configurational matrix. They are obligatory in that a dependent element must have an antecedent. Moreover, the antecedent must be unique, in a c-commanding position and sufficiently close to the dependent.

An explanation for the configurational matrix is in order, as it can hardly be accidental that it holds of the four relations mentioned above, which are prima facia very different in nature. Moreover, as Koster points out, the properties of grammatical relations are by no means necessary – non-syntactic relations such as coreference do not display them. It must be the case, then, that the configurational matrix reflects primitives of syntax.

In this paper, we explore what the primitives in question could be. We argue that a grammatical relation is established if a function introduced by a dependent is copied upward until it directly dominates a node which satisfies it (section 2). This allows us to capture the properties of c-command (section 3) and obligatoriness (section 4). We further argue that nodes are minimally ordered sets of attributes (section 5). From this, it can be derived that the antecedent in a grammatical relation is unique (section 6), but
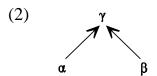
that the dependent need not be (section 7). After a preliminary discussion of A-movement (section 8), the two claims just mentioned are also shown to underlie the final property.

2 Relations as functions
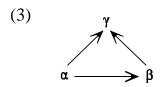
2.1 Against sideways relations

Syntactic structures result from an operation that combines two expressions α and β into a more complex expression γ.[1] It is commonly held that at least the categorial features of γ are copied from either α or β, with the effect that a projection line is created. This implies that syntactic junctures involve upward copying of information:

(2)

The question arises whether in addition to upward copying other elementary syntactic relations must be allowed. It seems that this is the case, given the existence of grammatical dependencies conditioned by specifier-head agreement, m-command or c-command. At first blush, these do not involve mother-daughter relations, but rather relations between sisters (or between a sister and a node dominated by a sister). Thus, one may be inclined to think that the representation in (1) must be enriched with a sideways relation. Holding on to a minimal theory, this sideways relation must presumably be sisterhood, as in (2).[2]

---

[1] Throughout this paper we will talk about syntax as if it is a derivational system. It should be noted from the outset that we only do so for presentational purposes and that nothing of substance depends on it.

[2] Indeed, in many theories sisterhood or a variant thereof is adopted. Government-binding theory and its successors assume that selection and internal Θ-role assignment take place under sisterhood (cf. Chomsky 1981, 1986b). Similar statements are made in LFG (cf. Bresnan 1982), GPSG (cf. Gazdar et al. 1985) and HPSG (cf. Pollard & Sag 1987, 1994). In categorial grammar, combinatory rules apply to adjacent entities only (cf. Steedman 1993). Adjacency and sisterhood are of course different notions, but like sisterhood, adjacency can be seen as a minimal sideways relation.

(3)

Although widely assumed, sisterhood is an unlikely primitive relation. First, two nodes which are in a sideways relation always have a common ancestor. In (2), for example, sisterhood between α and β only holds in virtue of the presence of γ. By definition, α and β are sisters if the node that directly dominates α also directly dominates β. A definition in terms of (direct) domination is unexpected if sisterhood is a truly primitive notion. Put differently, since there are indispensable relations between α and γ and between β and γ, and since these already connect α and β, it seems redundant to assume an independent – sideways – association of α and β. Sisterhood does not adhere to the null hypothesis in (3).

(4) *Accessibility*: Relations between nodes require immediate domination.

A grammatical theory not based on accessibility allows relations that do not have a 'medium' and therefore it in principle allows any node in a tree to be *directly* related to any other node. This implies that in addition to sisterhood other relations can be formulated, many of which are not attested in natural language.[3] The optimal theory makes it impossible to formulate such sideways relations.
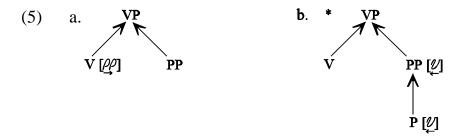
   In addition to the conceptual complication just discussed, theories based on sisterhood give rise to an empirical problem. Any theory must distinguish between internal and external relations. Anaphoric binding and the licensing of negative polarity items involve relations between a maximal projection and an element external to it. By contrast, direct Θ-marking and c-selection hold between a head and elements contained in its projection. As we will now show, the strict locality of internal relations cannot be understood in a theory that assumes sisterhood and upward copying as primitives.

   By definition, the projection of a head consists of those nodes to which its categorial features are copied. This implies that a natural distinction between internal and external relations is available. Internal relations involve properties that can be copied as far as the head's categorial features but no further. External relations involve properties that can be

---

   [3] There are many examples in the literature of direct relations which are weaker than sisterhood. One is m-command. Another example is Chomsky's (1986a) suggestion to define direct Θ-marking in terms of an extended version of sisterhood which ignores functional categories. On that definition the subject would be a sister of VP, even though I' dominates VP but not the subject.

copied independently of categorial features and hence beyond the maximal projection of the head that introduces them.[4]

If, in addition to copying, sisterhood is assumed as a primitive, internal relations cannot be kept strictly internal. This can easily be demonstrated on the basis of c-selection. Although a head may select the category of arguments within its maximal projection, an argument may not select for the category of the head whose projection dominates it. Thus we have verbs that select for a PP complement (cf. 4a) but not PP complements that select for a V (cf. 4b). However, after copying, the selectional properties of a preposition may appear on its maximal projection, where they could be satisfied by a verb under sisterhood.

(5) a. [tree diagram: VP with V [PP] and PP]   b. * [tree diagram: VP with V and PP [V], dominating P [V]]

Of course, there are various ways in which this problem could be circumvented, but these all require additional assumptions. One could stipulate, for example, that c-selectional features can be copied to intermediate but not to maximal projections. Although this would yield the required locality constraint, it amounts to little more than a restatement of the data.

Another option would be to say that c-selectional properties cannot be copied at all. This may seem more principled but it is a stipulation as well. It is clear on empirical grounds that projection involves more than the copying of categorial features. The gender of a noun, for instance, must be present on NP. Indeed, the simplest assumption is that all information in the head is copied upward under projection. Therefore, restricting the copying of c-selectional properties is ad hoc.

Moreover, there is evidence that selectional properties can be copied. A head can select for the category of a non-sister. Depending on one's analysis, one of the variants of (5) presents a case in which the verb *looked* selects a PP separated from it by an adverbial.

---

[4] We will use 'maximal projection' and 'intermediate projection' as convenient shorthands for the highest and all other nonterminal nodes, respectively. However, as will become apparent in section 5.2, we adhere to a theory of bare phrase structure.
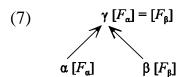
(6)   a.   The director looked at the telegram pensively.
      b.   The director looked pensively at the telegram.

It seems that copying of selectional features must be allowed more generally, given the existence of VP-shell structures like *John met* [*Mary* [t_v *nude*]], where the trace of the verb selects *Mary*. Similarly, in double object constructions like *John gave* [*Mary* [t_v *a book*]] only one of the objects can be a sister of the verb. Selection of the other therefore requires copying.

   In conclusion, sisterhood cannot be a primitive of grammar. On the one hand, it is defined in terms of the more primitive notion of domination; on the other, it leads to complications in keeping internal relations internal.

2.2 Against the copy-only theory

The only way to improve on this situation is to decompose the sideways relation in (2) into the two relations that hold independently between α and γ, and β and γ. An initial hypothesis might be that upward copying is the only primitive relation. If so, sideways relations could be viewed as instances of feature identification, where identity holds if the same feature is copied to a mother from both its daughters. Whenever a node has a feature that must undergo identification, the presence of an identical feature on its sister is forced, thus giving the impression of a sideways relation:
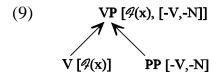
(7) $\gamma \ [F_\alpha] = [F_\beta]$

          $\alpha \ [F_\alpha]$ $\beta \ [F_\beta]$

This proposal has the virtue that it can keep internal relations internal. Assuming that c-selectional requirements cannot be copied beyond the maximal projection of the head that introduces them, it follows that the category of a head cannot be c-selected by (the head of) its complement. Thus, the structure in (4b) is ruled out.

   It cannot be correct, however, that upward copying is the only primitive relation. Although the schema in (6) may seem a natural way to implement agreement relations, extending it to c-selection is problematic. C-selection is not a case of agreement, because the c-selecting head and the c-selected category do not necessarily have the same categorial features. Consequently, percolation of categorial features from the selected

category will in many cases lead to incompatible feature values on the dominating node. This is illustrated below for a verb selecting a prepositional complement.

(8)   *   **VP [+V,-N]/[-V,-N]**

                V [+V,-N]   PP [-V,-N]

Of course, one could assume that the verb has a second set of categorial features which is used to encode c-selection, but this begs the question how the verb can contain conflicting features in the first place.

  The problem at hand is more general. Take the case of anaphoric binding, which in a theory based on copying and identification is most naturally represented as matching of φ-features on a dominating node. Such a proposal would require, for instance, that in a sentence like *I showed John to himself* the number and gender features of *John* and *himself* are copied to VP. However, VP does not have referring properties and can therefore not contain the copied information.

  One can improve on the situation just sketched by invoking function application rather than feature identification. Three assumptions are necessary. (i) A verb which selects a PP contains a function, say $\mathcal{F}(x)$, that looks for an appropriate set of categorial features. The minimal function that expresses this is $\mathcal{F}(x) = x$, where the domain is specified as {PP}. (ii) This function is copied to VP and satisfied if the relevant set of features is copied to VP as well. (iii) After copying, the categorial features inherited from the PP do not clash with those inherited from the verb, as they are now interpreted as the argument of $\mathcal{F}(x)$:

(9)   **VP [$\mathcal{F}(x)$, [-V,-N]]**

          V [$\mathcal{F}(x)$]   PP [-V,-N]

The use of functions to encode c-selection avoids the problem that the verb must have two distinct sets of categorial features if it selects for a PP. However, it is hard to maintain that the proposal depicted in (8) reduces categorial selection to two instances of copying. In fact, an additional rule is adopted, namely that of function application. The categorial features copied from PP are not a property of VP; rather they are interpreted as the argument of $\mathcal{F}(x)$. So, although function application may be dependent on copying, it is

not an instance of this process. This is confirmed by the fact that its structural description requires the presence of a function and the argument of that function in the same node. The rule can be formulated as in (9), where *n* stands for a node and *a* stands for the argument that $\mathcal{F}(x)$ is looking for. In contrast, the structural description of copying requires immediate domination.

(10) *Function application*: $Contain$ (n, $\mathcal{F}(x)$) & $Contain$ (n, a) $\Rightarrow \mathcal{F}(a)$

Let us summarize the argumentation so far. A theory based on copying and sisterhood must be rejected on conceptual grounds: sisterhood is an unlikely primitive of grammar. A theory based on just copying must be rejected on empirical grounds, as grammatical relations do not always involve matching features. Hence, the most promising theory at this point is based on copying plus the interpretational rule in (9). This theory is very close to the one we will ultimately adopt, but it faces a final problem, having to do with the structural description of function application.

  The problem becomes apparent when we consider the case of heads that c-select a category of the same type. Such heads contain the features that satisfy the c-selectional function they introduce. Given the rule in (9), the conditions for function application are met and therefore the function should count as satisfied. Yet, there are heads that c-select for a category of the same type; determiners selecting an NP are an example (cf. Hudson 1987 and Abney 1987). To work around this problem, the origin of features in a node must be taken into account: the features that form the argument of the function must be copied from a node that does not contain the function itself. This is expressed by the reformulation of the rule for function application in (10).

(11) *Function application* (revised): $Contain$ (n$_1$, $\mathcal{F}(x)$) & $Contain$ (n$_1$, a) & $Dominate$ (n$_1$, n$_2$)
      & $Contain$ (n$_2$, a) & $\neg$ $Contain$ (n$_2$, $\mathcal{F}(x)$) $\Rightarrow \mathcal{F}(a)$
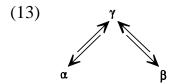
With (10), containment has become a suspicious structural description. Since a function must be satisfied by information that originates in a node different from the one that hosts the function, the question arises why the relevant information should be copied in the first place.
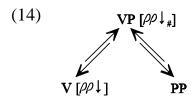
2.3 Upward and downward relations

This brings us to our own proposal. The conclusion seems inescapable that in addition to upward copying, a second primitive relation exists, presumably one of function

application. The conditions under which this rule applies are implicit in what we have said so far. Function application cannot be a direct (horizontal) relation between the complement and the selecting head. It can also not be conditioned by containment. This leaves just one option: like copying, the rule of function application must be conditioned by domination. However, whereas copying is an upward relation (a function is copied to a dominating node), we propose that function application is a downward relation (a function is satisfied if it directly dominates its argument). That is, instead of the rule in (9), we adopt the one in (11). Note that (11) is derived from (10) by omission of the suspicious clause "$Contain$ (n_1, a)".

(12) *Function application* (revised): $Contain$ (n_1, $\mathcal{F}(x)$) & $Dominate$ (n_1, n_2) & $Contain$ (n_2, a) & ¬ $Contain$ (n_2, $\mathcal{F}(x)$) $\Rightarrow$ $\mathcal{F}(a)$

To repeat, we call copying upward, as a node containing an unsatisfied function enters into a copy relation with its mother in order for the function to be satisfied. We call application downward, as a node containing an unsatisfied function enters into an application relation with its daughter in order for the function to be satisfied (as we will see, this downward relation is not anti-compositional). Thus, in addition to the upward relation standardly assumed, we make full use the relationships allowed by accessibility and adopt an additional, downward, relation. We will henceforth indicate mother-daughter relations as in (12):
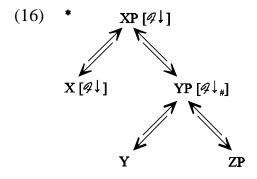
(13)

Selection now receives a straightforward analysis. Suppose, as before, that selectional requirements are functions satisfied by an appropriate set of categorial features. Through the process of upward copying, such a function is transferred from the head that introduces it to the node directly dominating the complement, after which the downward relation between mother and daughter allows for satisfaction. In other words, a c-selectional function is satisfied if it dominates a node of a particular syntactic category. In (13) the function in question is represented as $\rho\rho\downarrow$. For reasons of readability, we use '#' to indicate that the rule in (11) has applied.

(14)

$$\text{VP } [\rho\rho\downarrow_{\#}]$$

$$\text{V } [\rho\rho\downarrow] \qquad \text{PP}$$

There is a crucial asymmetry between the upward and downward relations in (13). The upward relation is one by which information is repeated on a higher node. The downward relation does not involve repetition of information, but establishes the status of functions in the mother node on the basis of the context in which it appears. This, we believe, is not just a property of selection but reflects the way compositionality, as formulated in (14), restricts the flow of information in syntactic structures. As we will now show, this principle forces upward relations to be based on copying and downward relations to be instances of function application.

(15) *Compositionality*: The properties of a non-terminal node are fully recoverable from its daughters; the properties of a terminal node are fully recoverable from the lexicon.
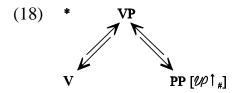
Let us begin by considering why downward copying violates compositionality. Suppose that a function originating in a head X is copied upwards onto X's maximal projection and then copied downwards onto a maximal projection YP, as in (15):

(16) *

$$\text{XP } [\mathcal{G}\downarrow]$$

$$\text{X } [\mathcal{G}\downarrow] \qquad \text{YP } [\mathcal{G}\downarrow_{\#}]$$

$$\text{Y} \qquad \text{ZP}$$

The resulting configuration is ruled out, as the node labelled YP in (15) has a function $\mathcal{G}\downarrow$ which is not present on either of its daughters. Consequently YP is partially non-compositional. A similar line of argumentation applies to downward copying to terminals. It follows that downward relations can never involve copying:

(17) *Theorem I*: Downward relations are instances of function application.

Next, consider upward function application. What such an operation would amount to is that a node imposes a requirement on its mother. It is evident that this kind of selection does not exist and UG should therefore not allow it. There is no need to add a stipulation to this effect, as upward function application violates compositionality. Let us briefly return to a structure discussed earlier, namely one in which a PP-complement selects for a verb. If upward function application were allowed, such selection could be encoded through a function $\mathcal{VP}\!\uparrow$, which is present on the PP and satisfied by the dominating VP. Note that (17) forms a minimal pair with (13):

(18)   *   VP

V   PP [$\mathcal{VP}\!\uparrow_{\#}$]

Compositionality requires that the properties of a node be fully recoverable from its daughters. This is not true of the PP in (17). The function on this node is satisfied, but this in not due to any of its daughters. As explained, $\mathcal{VP}\!\uparrow$ depends on the dominating VP for its interpretation. This implies, then, that upward function application violates compositionality, just like downward copying.

(19) *Theorem II*: Upward relations are copying relations.

The proposal just outlined shares a desirable property with the one depicted in (6): it captures the observation that selection is phrase-bound. Assuming as before that c-selectional functions introduced by a head X cannot percolate beyond XP, it follows that such functions can only be satisfied by elements contained in XP.

   A theory based on upward copying and downward function application echoes earlier work by Brody (1998a), who argues that structures are built up by two operations, namely project and insert. Project involves upward copying of information, whereas insert involves the satisfaction of grammatical functions. Insert is not a sideways relation, but one between the inserted node and the node that has been projected. It could perhaps be argued that this view is implicit in other theories based on generalized transformations as well (cf. Chomsky 1955, 1995b:189-190). Roughly, such theories assume that structures are built by an operation that creates an insertion site $\triangle$ (yielding [$_\gamma$ $\triangle$ α] from α) plus an operation which substitutes a second expression β for $\triangle$ (yielding [$_\alpha$ β α]).

Of course, not everything copied upward is a function: categorial features, for example, are not. However, if a grammatical relation is to be established between two nodes α and β, then α must introduce a function that can be satisfied by properties of β. If the information copied from α does not include a function satisfied by these properties, no relation between α and β can obtain. The only alternative would be to transfer a property from α to β, but this would necessarily involve a step of downward copying, in violation of theorem I. We therefore derive a third theorem:

(20) *Theorem III*: Grammatical relations are mediated by functions.

The ramifications of theorems I-III will become apparent in the following sections.[5]

3. C-command

3.1 Deriving c-command

A recurrent question in the recent literature is whether c-command can be derived from more fundamental properties of grammar. Perhaps the best-known proposal is by Epstein (1995), who derives c-command from a derivational view of syntax. A detailed critique of this analysis is offered in Brody 1998b. In this section we present an alternative to Epstein's approach, one which is neutral with respect to a derivational or representational view of syntax. Our alternative has some affinity with Brody's own proposal, although it differs considerably in execution. What we will argue is that c-command reduces to compositionality and the asymmetry it forces between upward and downward relations.

It is trivial why phenomena involving semantic scope are sensitive to c-command. Consider the case of two adverbials. In *he* [$_α$ *deliberately* [$_β$ *twice knocked on the door*]], *deliberately* takes scope over *twice*, whereas in *he* [$_α$ *twice* [$_β$ *deliberately knocked on the door*]], scopal relations are reversed. Thus, an adverbial attached higher typically takes wider scope. By compositionality, the interpretation of β in these examples is constructed from the meanings of VP and the lowest adverbial; the interpretation of α is constructed from that of the higher adverbial and β. Since information is only copied upwards, the

---

[5] Of course we are not the first to propose that grammatical relations are mediated by functions. This claim is particularly prominent in categorial grammar and HPSG. The main difference between these theories and the one proposed here is the environment in which function application takes place: adjacency in categorial grammar, sisterhood in HPSG and domination in the present theory.
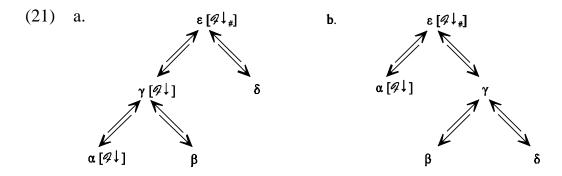
semantics of higher adverbial cannot be part of the semantics of β or indeed of the semantics of the VP. Hence each adverbial takes scope over its sister, as required.[6]

So much is of course uncontroversial. However, c-command is also a condition on syntactic relations, and this is less easy to understand. In general, syntactic dependencies require that the dependent element be c-commanded by its antecedent. A trace, for instance, must be c-commanded by its copy, an anaphor by its binder and a predicate by its subject. In standard analyses, these dependencies are 'direct': dependent and antecedent are related without the transmission of information through intervening nodes. Why a direct relation should require c-command is unclear.

That c-command restricts syntactic dependencies follows immediately in a theory based on accessibility as defined in (3). As stated in theorem III, a grammatical relation must involve a function introduced by a dependent. This function relies on copying and application for satisfaction. Compositionality forces copying to be upward (by theorem II) and application to be downward (by theorem I). C-command now follows from the fact that copying can apply recursively, while function application is restricted to immediate domination as it does not involve transfer of information. A function copied to a dominating node can subsequently be copied further from that node to its mother, and so on. Indeed, unless there is a bounding node for copying (such as a head's maximal projection in the case of c-selection), there is no inherent upper limit to the sphere of influence of a function. The downward relation of function application is strictly bounded, however. A function can only be satisfied by a daughter of the node that hosts it. Satisfaction by a node lower in the tree would violate accessibility.

The effects for grammatical relations are as follows. The element satisfying the function on which the relation is based does not have to be a sister of the element introducing it. The function $\mathscr{A}{\downarrow}$ in (20a) can be satisfied by β, or if β is not an argument of the right type, the function is copied onto ε and applied to δ. The latter possibility is indicated by the '#' on the function in ε. The situation is different in (20b). The function $\mathscr{A}{\downarrow}$, once copied onto ε, can be satisfied by γ. Given accessibility, it cannot be satisfied by either β or δ, however. This is the desired result.

---

[6] Where scope relations differ from what is to be expected on the basis of the overt syntactic structure, it is usually assumed that covert movement operations apply (cf. May 1985 and subsequent work). Such operations reconcile the observed interpretation with compositionality.
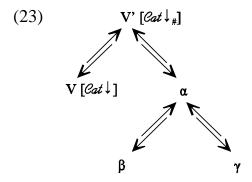
(21)   a.

$$\varepsilon\,[\mathcal{Q}\!\downarrow_{\#}]$$

$$\gamma\,[\mathcal{Q}\!\downarrow]\qquad\qquad\delta$$

$$\alpha\,[\mathcal{Q}\!\downarrow]\qquad\qquad\beta$$

b.

$$\varepsilon\,[\mathcal{Q}\!\downarrow_{\#}]$$

$$\alpha\,[\mathcal{Q}\!\downarrow]\qquad\qquad\gamma$$

$$\beta\qquad\qquad\delta$$

In conclusion, compositionality explains why semantic scope as well as grammatical relations are conditioned by c-command:
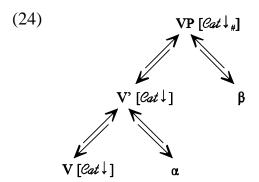
(22) *Theorem IV*: If α is grammatically dependent on β, β c-commands α.
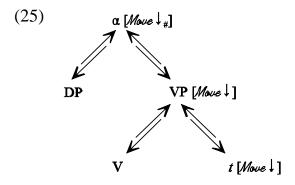
3.2 Effects for grammatical relations

Let us briefly illustrate the effects of theorem IV for c-selection. Here, the dependent category is the verb: it introduces a function $\mathcal{Cat}\!\downarrow$ which is satisfied by the selected category (the function $\rho\rho\!\downarrow$ that we encountered earlier is a particular instantiation). The selected category can thus be seen as the antecedent: it caters to an unsaturated property of the verb. Hence, the selected category must c-command the verb. In the tree below, $\mathcal{Cat}\!\downarrow$ can be satisfied by α, but not by β or γ (compare (20b)). This is the correct result. To give an example, a verb cannot c-select the specifier of its complement.

(23)

$$\text{V' } [\mathcal{Cat}\!\downarrow_{\#}]$$

$$\text{V } [\mathcal{Cat}\!\downarrow]\qquad\qquad\alpha$$

$$\beta\qquad\qquad\gamma$$

On the other hand, the verb need not c-command the c-selected element. In (23), $\mathcal{Cat}\!\downarrow$ can be satisfied by α or, if this is not the case and the function is copied onto VP, by β (compare (20)). Again, this is the correct result, as illustrated by (5) above.

(24)

$$\text{VP } [\mathit{Cat}\!\downarrow_{\#}]$$

V' $[\mathit{Cat}\!\downarrow]$ β

V $[\mathit{Cat}\!\downarrow]$ α

Of course, we expect this pattern to extend to other grammatical relations, given that such relations must also be encoded by functions (by theorem III). For movement, this implies that the trace, which is the dependent, introduces a function. For now let us refer to this function as $\mathit{Move}\!\downarrow$; we will see later on that movement relations are in fact mediated by three different functions.

(25)

$$\alpha \ [\mathit{Move}\!\downarrow_{\#}]$$

DP   VP $[\mathit{Move}\!\downarrow]$

V $t\ [\mathit{Move}\!\downarrow]$

Like the functions involved in categorial selection, these functions can formally be written as f(x) = x, but their domain specification is different (see sections 8 and 9). We assume that all grammatical functions have this format and that they are only differentiated in terms of the argument they take. Thus, functions establish a grammatical relation, but they do not provide an interpretation for it. In the case at hand, the trace and the moved DP are associated in syntax – that the latter must be interpreted as an argument of the verb is something that is established by reading off the syntactic representation at the interface with semantics. In fact, if we take the autonomy of syntax seriously, it is impossible for a grammatical function to yield a semantic object.

If movement relies on function application, as it must, it is correctly predicted that there can be no downward or sideways movement. This restriction is illustrated by the contrasts in (25a,a') and (25b,b') respectively.

(26) a. John said [that [to young children] one should never tell such stories t_{PP}].
     a'. *John said t_{PP} [that [$_{PP}$ to Mary] safe is better than sorry].
     b. [Deciding [which movie to see t_{WH} next]] makes John very happy
     b'. *[Deciding [who to see that new movie next]] makes t_{WH} very happy

Anaphors, like traces, are dependent elements. They must therefore introduce a function, which we may call *Self*↓. Let us provisionally assume that this function is satisfied if it finds a suitable referring expression. We return to the nature of *Self*↓ in section 9; for the moment we conclude that its sensitivity to c-command follows:

(27) a. I expect [John [to defend himself]].
     b. *I expect [himself [to defend John]].

Negative polarity items form another class of dependent elements (cf. Progovac 1994), and they must therefore introduce a function, say *Neg*↓, which looks for a negative operator. The expected c-command effects can be observed:
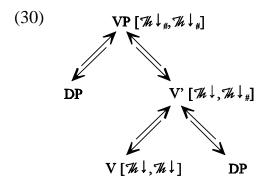
(28) a. [No book [that John wanted to buy]] was on offer in any shop.
     b. *[The book [that no one wanted to buy]] was on offer in any shop.

Predicates form a final class of dependent elements. They introduce a thematic function *Th*↓, which is satisfied by their subject.[7] It should therefore come as no surprise that predication, as already noted by Williams (1980), requires c-command. The behaviour of depictives in Dutch illustrates this fact. The depictive in (28a) can be associated with either the subject or the object, since both DPs c-command it. By contrast, (28b) and (28c) only have a subject-oriented reading, as the object does not c-command the depictive. Finally, (28d) is ungrammatical because the depictive cannot find a c-commanding antecedent.

---

[7] If Chomsky's (1986b) visibility condition is correct, thematic functions look for case bearing constituents. Formally, f(x) = x, where the domain is {NOM, ACC, DAT, ...}.

(29) a. dat [Jan$_i$ [Marie$_j$ [naakt$_{i/j}$ ontmoette]]].
        that John Mary nude met
   b. dat [Jan$_i$ [naakt$_{i/*j}$ [Marie$_j$ ontmoette]]].
   c. dat [Jan$_i$ [[naakt$_{i/*j}$ sprak] [met Marie$_j$]]].
   d. dat [naakt$_{*i/*j}$ [Jan$_i$ [Marie$_j$ ontmoette]]].

On this view of Θ-theory, thematic roles involved in complementation must be functions as well. This implies that in a double object construction two thematic functions are copied upward from the verb. The first of these is satisfied by the verb's sister; the second by the sister of V'. Thus, the following Larsonian representation holds of double object constructions in English:

(30)

$$\text{VP}\ [\mathcal{Th}{\downarrow}_{\#},\mathcal{Th}{\downarrow}_{\#}]$$

$$\text{DP} \qquad \text{V'}\ [\mathcal{Th}{\downarrow},\mathcal{Th}{\downarrow}_{\#}]$$

$$\text{V}\ [\mathcal{Th}{\downarrow},\mathcal{Th}{\downarrow}] \qquad \text{DP}$$

It is to be expected, then, that internal Θ-role assignment requires c-command: the argument must c-command the verb. Indeed, there are no constructions in which a verb assigns a Θ-role to an argument contained in its complement. This is illustrated by the fact that *John saw the bomb explode* does not imply that John saw the bomb; it only implies that he saw the explosion. It is also predicted that the verb need not c-command its arguments, given that copying of thematic functions is recursive. In view of the existence of double object constructions, Θ-role assignment to a non-sister must indeed be allowed. So, internal Θ-role assignment patterns with other grammatical relations in being governed by the fundamental asymmetry between (recursive) upward copying and (strictly local) downward function application.

   The fact that grammatical relations as different as c-selection, movement, binding, negative polarity, predication and complementation adhere to the same requirement must be a consequence of the way the syntax operates, as proposed here. The standard theory can of course capture the data, but at the cost of stipulating, for each grammatical relation, that c-command is a conditioning factor.

3.3 Copying of features versus copying of functions

A final remark is in order. The feature that satisfies a function can itself have been copied. The simplest such case is presented by c-selection, as the categorial features of a selected complement will in many cases have been copied upward in the process of projection. Similarly, the negation feature that licenses a negative polarity item may originate in the specifier of the antecedent.

  This observation in no way threatens the explanation of c-command offered here. Copied features become properties of the node they are copied to. This simple fact restricts copying of features to those nodes with which they are compatible. Categorial features, for instance, cannot be copied to a node that has a different set of categorial features. In the same vein, negation features, if copied, must yield a negative expression.

  Of course, copied functions also become properties of the node they are copied to. However, whereas features indicate what a node *is*, functions indicate what a node *asks for*. In many cases, a node cannot be two things at the same time (say a verb and a preposition). But is can be one thing (say a verb) and ask for another (say a PP). Thus, copying of functions is less restricted than copying of features.

4. Obligatoriness

Besides c-command, Koster's (1987) configurational matrix includes obligatoriness. He views the configurational matrix as a meta-function R, which establishes grammatical relations in a way that need not concern us here. Since functions, on Koster's view, are undefined unless they acquire a value, their satisfaction is obligatory. Thus, he claims, obligatoriness does not have to be stipulated as a substantive property of R.

  Although we believe that the basic insight of Koster's proposal is correct, it is not immediately obvious why an unsatisfied function should be undefined. What Koster intends to express is that a function, once introduced, must have an effect. This in turn requires that the function is satisfied. This condition on the introduction of functions follows straightforwardly from representational economy, as formulated in (30).

(31) *Representational Economy*: Information in nodes is minimized.
A function is introduced in order to establish a grammatical relation. Therefore, if that relation is not established, the function should not have been introduced. Thus, we derive theorem V:

(32) *Theorem V*: Dependent elements obligatorily take an antecedent.

Indeed, it is easily demonstrated that grammatical relations must be established once a dependent is introduced into the tree. The examples in (32) are all ungrammatical. The c-selecting verb in (32a) is combined with a DP rather than a PP; the trace in (32b) does not find a moved element; the anaphor in (32c) does not find a suitable antecedent; the negative polarity item in (32d) is not in the scope of a negative operator; and, finally, the predicate *naakt* in (32e) takes neither *Jan* nor *Marie* as its subject.

(33) a. *John looked Mary
    b. *John bought *t*.
    c. *I expect [you [to defend himself]].
    d. *Anyone better leave town.
    e. *dat [Jan$_i$ [Marie$_j$ [naakt$_k$ ontmoette]]].
       that John Mary nude met

As in the case of c-command, the standard theory can capture these data by stipulating for each of the grammatical relations involved that it is obligatory. However, such a theory does not explain the pervasive nature of this requirement.

5. Nodes as Sets

5.1 The internal structure of nodes

Implicit in the preceding discussion is a particular view of nodes. Since nodes generally have multiple characteristics, they can be considered sets of attributes. As already assumed in the previous section, nodes do not contain more information than necessary (*representational economy*). This implies that the number of attributes in a node is minimized and furthermore that ordering is not introduced into a node unless this is required.

  Let us consider under what circumstances ordering must be imposed on a node. By definition, two unordered sets S_1 and S_2 are identical if each member of S_1 is contained in S_2 and vice versa. This implies that it is meaningless for an unordered set to contain the same member twice, because a set with this property is equivalent to the set that contains the member in question only once. So in the absence of ordering, no node can contain the same attribute twice (at least not in a meaningful way). The node in (33a), in which the function $\mathscr{D}{\downarrow}$ is duplicated, is therefore identical to the one in (33b).

(34) a. $\{\mathcal{A}\!\downarrow, \mathcal{A}\!\downarrow\}$ b. $\{\mathcal{A}\!\downarrow\}$

This raises the question how a node can ever truly contain more than one function of the same type. For instance, how can a verbal projection host two (or more) distinct Θ-roles, as must for instance be assumed for projections of triadic verbs? Of course, one could argue that the semantic label of thematic roles (AGENT, THEME, etc.) is visible in syntax, with the effect that thematic functions are not identical. But apart from the conceptual drawback of introducing interpretative notions into representations manipulated by the syntax, a proposal along these lines lacks empirical confirmation. There is no evidence to suggest that different thematic roles must be distinguished syntactically.

   Set theory offers a simple and attractive solution. Although an unordered set cannot contain a member twice, an ordered set can. This implies that the information in a node must be structured if it is to contain the same attribute more than once. In what follows, we will represent partial ordering by linking the attributes to a tier that defines positions in a node. (30) implies that this linking will be minimized. Thus, if a node contains a function more than once, then all but one of these functions must be linked to the ordering tier. A transitive verb, for example, can be represented as in (34a). Similarly, a triadic verb has the structure in (34b).[8]

(35)   a. $\{..., \mathcal{Th}\!\downarrow, \mathcal{Th}\!\downarrow, ...\}$ b. $\{..., \mathcal{Th}\!\downarrow, \mathcal{Th}\!\downarrow, \mathcal{Th}\!\downarrow, ...\}$

                            | | |

                           1   1   2

In summary, we have derived the following theorem:

(36) *Theorem VI*: Identical attributes in a node must be ordered.

It can easily be demonstrated that ordering in a node can only originate in a terminal node. Suppose that two identical functions are present on a non-terminal node α. As just explained, these functions can only coexist on α (in a meaningful way) if at least one of them is linked to α's ordering tier. If this is not the case, a set results that is equivalent to one that contains the function in question just once. At first sight, one might think that the set that constitutes α can be structured on an ad hoc basis. This, however, would violate

---

[8] In set theory, the ordering of sets is achieved by indexing. An indexed set is generated by a function that takes as its domain two sets and whose range is a set of pairs consisting of a member of each set. In the examples at hand, the function creates a partial pairing of thematic functions with natural numbers.
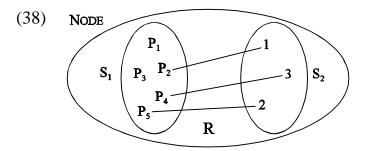
compositionality (cf. (14)); a property would be introduced into α which cannot be recovered from its daughters. In other words, whenever functions in a non-terminal node are ordered, then this ordering must have been inherited from a daughter, which implies that it must ultimately originate in a head. We thus derive the following theorem:

(37) *Theorem VII*: Ordering of functions must be introduced by terminal nodes.

Theorems VI and VII shape parts of Θ-theory. The former explains why Θ-roles must be part of a thematic grid: a predicate can only have multiple Θ-roles if these are distinguished through linking to an ordering tier (cf. Grimshaw 1990 and Jackendoff 1990). The latter explains why Θ-grids are introduced by heads. Introduction by a non-terminal node would violate compositionality.

5.2 Two types of copying

Now that we have seen how nodes are organized, we can address the question what kinds of information may be copied out of a node. It will be clear that functions can be copied individually. However, if a function is linked to the ordering tier prior to copying, that link will not be present in the node to which the function is copied. Formally, linking is a relation (say R) between the set of attributes in a node (S_1) and its tier (S_2, a subset of the natural numbers):

(38)

This implies that a link (that is, a partial order) can only be maintained under copying if R and the sets it relates are copied along. If either of the sets fails to be copied, R is undefined; if R is not copied, no linking obtains. We may conclude, then, that order can only be maintained if the entire node is copied. This is expressed by the following theorem:

(39) *Theorem VIII*: Ordering in a set is only preserved under en bloc copying.

To summarize, the theory developed so far allows two instantiations of the copy operation. Either individual functions are copied, or copying affects the whole node. These properties of the copy operation imply a particular view of projection. It follows from representational economy that copying only takes place if necessary. If information is copied unnecessarily, the receiving node will not be minimal, contra (30).

Since functions must be satisfied, copying of unsatisfied functions is licensed. Categorial features, however, are not functions and may therefore not be copied in isolation. The consequence is that copying of categorial features is parasitic on en bloc copying. Only if copying targets an entire node can the categorial features of that node be transferred to its mother. In other words, en bloc copying must be identified with what is commonly known as projection (that is, copying of categorial features). A projection line ends if copying targets one or more individual functions; vice versa, a projection line is extended whenever en bloc copying takes place.

The result just obtained is in line with the radical view of projection according to which the head is present in every node in its projection line. One proponent of this view is Williams (1994), who claims that the nouns *table* and *chair* project a *table*P and *chair*P, respectively. A similar line is taken in Chomsky's (1995a) theory of bare phrase structure. In the same vein, Brody (1998b) suggests that the head and the maximal projection are two interpretations of the same node.

We are now in a position to sharpen theorem VII above. Copying of structure to a dominating node implies en bloc copying, which in turn implies projection. In conjunction with compositionality, this yields (39).

(40) *Theorem VII* (revised): Ordering of functions in a non-terminal node α is introduced by the head of α.
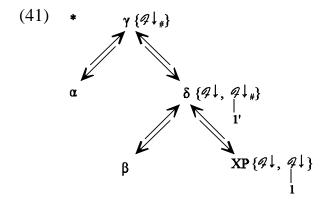
The theory of nodes developed in this section enables us to derive the third property of the configurational matrix: uniqueness.

6. Uniqueness

6.1 Deriving uniqueness

Although external grammatical relations are unique, a head can entertain multiple internal relations. The possibility of multiple internal relations follows from the fact that a head can impose a particular order on the set of functions that it introduces, and that this order

is maintained under projection (or en bloc copying). Let us now consider what happens if a function is satisfied externally to the projection in which it originates.[9]

Crucially, an external function must be copied individually. If en bloc copying takes place, the projection is extended and hence the function is not externalized. From this it follows that no maximal projection can enter into the same external relation twice. If a projection, say XP in (40), were to do so, two identical functions would have to be copied to the dominating node δ. Of course, these two external functions must be ordered on δ (by theorem VI). But as argued in the previous section, order is only maintained under en bloc copying. Hence, the two external functions on XP must copied individually, with subsequent introduction of order in δ. This ad hoc ordering of functions, indicated by a link to 1' rather than 1, violates compositionality (cf. theorem VII). Thus, (40), where XP enters into two external relations, is excluded.

(41)

We thus derive (41).[10]

(42) *Theorem IX*: A dependent can enter into the same external relation only once.

---

[9] Many theories assume a distinction between internal and external relations. One theory that does not is categorial grammar. This implies that in this theory it is hard to state the generalisation that c-selectional requirements can be imposed on internal but not external arguments. It is for instance possible to combine the verb with the subject before it is combined with the object. A further consequence is that the first property of the configurational matrix, c-command, cannot be accounted for in the syntax (cf. Steedman 1993).

[10] Theories that cannot express the distinction between internal and external relations, such as categorial grammar, cannot readily provide a syntactic account of the second property of the configurational matrix, uniqueness of the antecedent.

The following data show that this theorem is correct. The trace in (42a) cannot license two antecedents. Similarly, neither the reflexive in (42b) nor the depictive in (42c) can be associated with two arguments.[11]
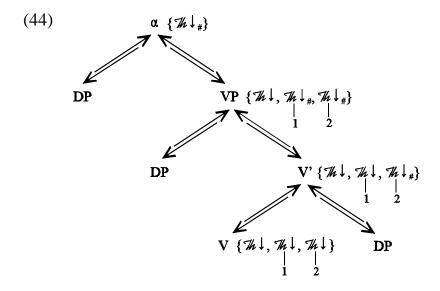
(43) a. *What did John wonder what he bought *t*.
　　b. *John confronted Mary with each other.
　　c. *dat [Jan$_i$ [Marie$_j$ [naakt$_{i+j}$ ontmoette]]]
　　　 that John Mary nude met

As was the case with c-command and obligatoriness, the standard theory can capture these data by stipulating that each of the grammatical relations in question requires a unique antecedent. However, the fact that this is true of all grammatical relations remains unexplained.

　The uniqueness of the subject in predicational relations, demonstrated for secondary predicates in (42c), also holds for primary predicates. As remarked before, a verb may select multiple complements, but it cannot take more than one subject. Hence, a triadic verb such as *give* projects the structure in (43). The grid of the verb in this structure contains three Θ-roles which are ordered through subordination. The assignment of the internal Θ-roles requires that en bloc copying takes place at least twice, after which the external Θ-role is copied in isolation.

---

[11] The point cannot be illustrated for negative polarity items, as there is no way of telling which negative polarity item is related to which negative operator.

(44)

$$\alpha\ \{\mathcal{T}\!h\!\downarrow_{\#}\}$$

Tree (44): α {Th↓#} branches to DP and VP {Th↓, Th↓#, Th↓#} (indices 1, 2); VP branches to DP and V' {Th↓, Th↓, Th↓#} (indices 1, 2); V' branches to V {Th↓, Th↓, Th↓} (indices 1, 2) and DP.

6.2 The lexical encoding of externality

The question now arises how a Θ-role, or more generally a function, is marked as external in a lexical entry. Of course, this can be done by stipulation, but it is preferable to make use of a concept introduced above, namely that of ordering. We have already established that certain functions must be linked to an ordering tier. Perhaps, then, we could assimilate the lexical representation of internal versus external functions to the distinction between linked and unlinked functions. This is in fact necessary if we take representational economy, as given in (30), seriously.

   If information in nodes is minimized, a lexical entry cannot contain a link to the ordering tier if that link remains without consequences in syntax proper or at the interfaces. Recall that the ordering of a set cannot be copied beyond the maximal projection of the head in which it originates. The implication is that if an external function were to be linked to the ordering tier, this information would be lost in the course of the derivation. As a result, linking of an external function does not have consequences, in violation of representational economy. Hence, the syntax interprets linked functions as internal, and functions that are not linked as external. This yields the following theorem (cf. Grimshaw 1990 and Williams 1994 for discussion):

(45) *Theorem X*: External functions cannot be linked to an ordering tier.

So, in (43) the leftmost Θ-role must be the external one. Similarly, the single Θ-role of an unergative verb must be represented as in (45a), not (45b). If (30) is not to be violated,

the Θ-role in (45b) must be internal. Therefore, (45b) is the proper representation of an unaccusative verb.

(46) a. {..., *Th*↓, ...}   b. {..., *Th*↓, ...}
                                                       |
                                                       1

What we have shown so far is that set theory in conjunction with the concept of projection explains why a head can have multiple internal functions of the same type. It also explains why external functions are unique. In addition, we have shown how the distinction between internal and external functions can be assimilated to the distinction between linked and non-linked ones, given representational economy.
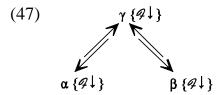
7. Function identification

7.1 The non-uniqueness of dependents

Although a constituent can only have one external function (of a particular type) and hence only one antecedent (of that type), it does not follow from anything we have said so far that antecedents are associated with a unique dependent. On the contrary, syntax imposes no limit on the number of dependents, as we will now demonstrate.
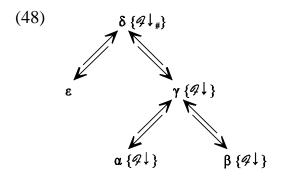
   Recall that nodes are sets. Hence they can only have the same function twice if one of them is linked to the ordering tier. Consider in this light the node γ, created by merger of α and β, which both introduce the function *A*↓. Suppose this function is copied from both α and β. Any ordering of these functions on γ must be ad hoc, given that they originate in different nodes. As we have seen, ad hoc ordering violates compositionality and is hence excluded. The fact that the two functions cannot be ordered on γ implies that they cannot be distinguished and therefore the resulting node is identical to a node that contains the relevant function only once. For all practical purposes the two functions may be said to collapse into one on γ (cf. (46)). We will refer to this as function identification.[12]

------

[12] Function identification also comes for free in unification-based theories such as LFG, GPSG and its descendants. In categorial grammar it must be stated in the form of separate rules. This is, for instance, apparent in the categorial treatment of coordination structures.

(47)

Recall that we used exactly the same line of argumentation to explain why external functions are unique, the only difference being that in that case the collapsing functions were copied from the same daughter (cf. (40)). The possibility of function identification has an important consequence: given that the functions copied from α and β in (46) collapse on γ, the antecedent that satisfies the function on γ will also satisfy the functions on α and β. This is represented in (47), where ε is the antecedent satisfying $\mathcal{A}\!\downarrow$.

(48)

We derive the consequence that the dependent in a grammatical relation need not be unique. In (47) ε satisfies only one function in δ, yet it is associated with two dependent categories. This state of affairs is summarized by the following theorem:

(49) *Theorem XI*: An antecedent can enter into the same grammatical relation more than once.

7.2 Effects for grammatical relations

There is a wealth of data corroborating theorem XI. For a start, although a reflexive is never licensed by more than one antecedent (cf. (42)), the example in (49a), taken from Koster (1987), demonstrates that an antecedent can license more than one reflexive. Similarly, a negative operator can license more than one negative polarity item, as in (49b).

(50) a. They talked with each other about each other
     b. John didn't have anything to say about any of the children

In the same vein, it seems that an antecedent can also – at least in principle – bind two traces. Parasitic gaps are a case in point: examples like the one below are comparable in all relevant respects to the examples in (49). Of course, there are many remaining questions surrounding the parasitic gap phenomenon, but it is sufficient here to note that function identification must be part of its analysis.

(51) These are the men that the police arrested t without warning e

A further class of cases confirming the existence of function identification concerns across-the-board phenomena. As is well known, a single antecedent can satisfy multiple functions in coordinated structures. This is illustrated below for movement, internal and external Θ-role assignment, negative polarity items and binding.

(52) a. Which song did [[John sing t] and [Mary dance to t]]
     b. John [bought and read] a book
     c. John [[bought a book] and [read it]]
     d. John didn't like [[any man] or [any woman] he met]
     e. No couple like [[each other] and [each other's parents]]

Whatever the analysis of coordination, it will be clear that each conjunct introduces a function in these examples, and that the coordinate structure as a whole has only one. In other words, coordination involves function identification as made available by the theory developed here.

7.3 A reformulation of the Θ-criterion

The examples in (51b) and (51c) already show that under certain circumstances a single argument can satisfy more than one thematic function. However, the argument based on Θ-theory can be strengthened considerably if we turn to the syntax of secondary predication.

Our starting point is the observation that there is a fundamental bifurcation in the collection of functions introduced so far. Some functions have a licensing capacity with respect to syntactic structure, whereas others do not. An argument, for example, can only be merged with a predicate if there is a thematic role that can be assigned to it. Similarly,

the landing site of movement is licensed by a $\mathcal{M}\!ove\!\downarrow$ function. This contrasts with a function like $\mathcal{N}\!eg\!\downarrow$: negative operators can exist without satisfying this function. The same is true of antecedents in binding relations.
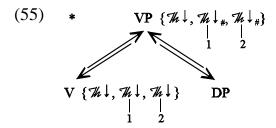
   The distinction between licensing and non-licensing functions allows us to formulate the following condition.

(53) *Licensing Criterion*: The creation of a non-terminal category reduces the number of unsatisfied licensing functions by exactly one.
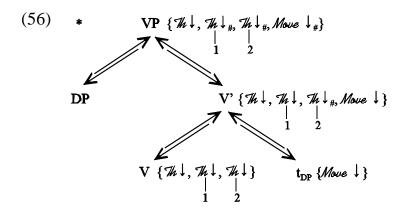
We will see in section 9 that this condition explains why adjuncts are islands. In the present context it is relevant as it subsumes both the Θ-criterion and the ban on movement to a Θ-position. Among the data captured by (52) is the example in (53), which is ungrammatical under the reading that John showed himself to himself in the mirror.

(54) *John showed himself (in the mirror).

This example has two possible representations, both of which violate (52). In the first, the object satisfies two thematic functions in its mother node (the goal and theme roles). This one-step reduction of unsatisfied licensing functions by two gives rise to the ungrammatical tree in (54), a representation usually ruled out by the Θ-criterion.

(55)

Alternatively, (53) could be interpreted as a case of raising to a Θ-position. The theme role would be assigned to a trace bound by the goal DP. This results in the representation in (55), where DP satisfies two licensing functions in its mother. As before, this one-step reduction of unsatisfied licensing functions by two violates (52). Representations like (55) are usually ruled out by the ban on movement to a Θ-position.

(56)   *   VP $\{Th\downarrow, Th\downarrow_{\#}, Th\downarrow_{\#}, Move\downarrow_{\#}\}$
                   1   2

     DP   V' $\{Th\downarrow, Th\downarrow, Th\downarrow_{\#}, Move\downarrow\}$
                              1   2

         V $\{Th\downarrow, Th\downarrow, Th\downarrow\}$ t_{DP} $\{Move\downarrow\}$
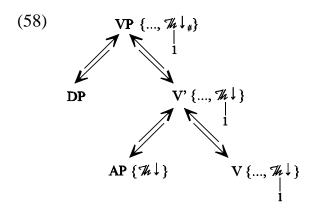   1   2

With the above in mind, let us turn to cases in which an argument is associated with both a primary and a secondary predicate. In (56), for example, the object *Marie* is assigned the verb's internal Θ-role, as well as the external Θ-role of the secondary predicate *naakt* 'nude'.

(57)   dat [Jan [Marie$_j$ [naakt$_j$ ontmoette]]]
      that John Mary nude met

There are several proposals in the literature that complicate the theory of thematic relations in order to accommodate data like these. Essentially, these proposals stipulate that the Θ-criterion holds of primary predication only. Williams (1983) and Chomsky (1986b) do so by amending the Θ-criterion itself. Stowell (1981, 1983) proposes that secondary predication is mediated by a controlled PRO subject, an approach which necessitates otherwise redundant assumptions about control (see Williams 1994 for discussion).

   The complications just mentioned dissolve in the theory presented here as a result of the availability of function identification. The AP's external thematic function in (56) is copied onto a node which, as a result of *en bloc* copying, also contains the verb's Θ-grid. The Θ-role copied from AP cannot be ordered with respect the Θ-roles in the verb's Θ-grid, as that would violate compositionality (cf. (36)). Hence it must collapse with one of the verb's Θ-roles. In (56) it is identified with the verb's internal role, as depicted below:

(58)

VP $\{..., \mathscr{T}\!\!\mathit{h}\!\downarrow_{\#}\}$

$\mid$ 1

DP

V' $\{..., \mathscr{T}\!\!\mathit{h}\!\downarrow\}$

$\mid$ 1

AP $\{\mathscr{T}\!\!\mathit{h}\!\downarrow\}$

V $\{..., \mathscr{T}\!\!\mathit{h}\!\downarrow\}$

$\mid$ 1

Of course, the Θ-role copied from the AP could also be identified with the verb's external Θ-role, giving rise to examples like (58).

(59) dat [Jan$_j$ [Marie [naakt$_j$ ontmoette]]]
      that John Mary nude met

So, on the theory presented here, secondary predication is just another instance of function identification. As a result, the problems with the formulation of the Θ-criterion referred to above dissolve. Crucially, both the representation in (57) and that corresponding to (58) adhere to the condition in (52), according to which a daughter may not satisfy more than one licensing function in its mother. In short, a theory that treats nodes as sets of functions makes available the notion of function identification, which in turn allows a maximally simple statement of the Θ-criterion and which at the same time predicts where apparent violations of this principle occur.

7.4 Prenominal modification

There are two additional pieces of evidence for the analysis of secondary predication outlined above. The first comes from (intersective) prenominal modifiers. Consider the example in (59).
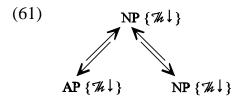
(60) The green door.

Like other APs, *green* is a predicate. It is unclear, however, how it satisfies its external Θ-role, as there does not seem to be a subject to which it can assign it. Given that predicates may not take dominating nodes as subjects, the DP in which *green* is embedded

does not qualify as an antecedent for its thematic function. At first sight, it seems that this problem may be solved by assuming a PRO subject for the AP, as in *the* [PRO *green*] *door* (cf. Van Gestel 1986). However, the interpretation of PRO cannot be determined by a dominating node either, so that the proposed structure recreates the original problem.

Higginbotham (1985) has suggested a mechanism by which examples like *the green door* can be interpreted. The crucial assumption is that NPs are predicates whose external Θ-role is bound by a determiner if the DP is to be used as an argument. This external role is often referred to as the R-role (cf. Williams 1981). What Higginbotham proposes is that the Θ-role of a prenominal modifier and the R-role of the noun are related in syntax by coindexation. The resulting structure is interpreted in such a way that whatever satisfies the R-role will also satisfy the Θ-role of the prenominal modifier. In the case at hand, the R-role is bound by D. *The green door* therefore refers to something that is both green and a door.

Given our proposal, the effects of Higginbotham's analysis come for free: the mechanism of coindexation reduces to function identification in the sense developed above. The structure of prenominal modifiers is as in (60), where both the NP and the AP contain a thematic function that is copied onto the dominating NP node. By compositionality, the AP's external role cannot be ordered with respect to the R-role of the noun and consequently these two roles must collapse. Whatever satisfies the thematic function on the top NP will satisfy the thematic functions of both the lower NP and the AP, as desired.

(61)

NP {𝒯ℎ↓}

AP {𝒯ℎ↓}   NP {𝒯ℎ↓}

We see, then, that a theory which treats nodes as unordered sets of properties provides a unified account of prenominal modifiers and other cases of secondary predication. This is an improvement over existing theories, which need two processes by which Θ-roles can be satisfied: assignment of Θ-roles for secondary predication and coindexation of Θ-roles for prenominal modifiers.[13]

---

[13] Williams (1994) goes some way towards unifying prenominal modification and secondary predication, as he invokes the same mechanism of 'vertical binding' for subject-oriented depictives and prenominal adjectives. Since vertical binding can only involve external properties of the vertical binder, it is not

7.5 Function identification and locality

Predicational relations cannot be established across an intervening subject. This, we will now argue, provides a final argument for an analysis of secondary predication based on function identification.

It is not the case that external thematic roles must be assigned within the m-command domain of the predicate, as witnessed by the example in (61), from Williams 1994. Here, the external Θ-role of the AP *still explosive* is copied out of the projection headed by *while* into the VP, where it collapses with the internal Θ-role of the unaccusative verb *arrive*.

(62) The device arrived [while [still explosive]]

In the light of this possibility, it is somewhat surprising that a similar long-distance predicational relation is not available in the examples in (62).

(63) a. dat Jan$_i$ Marie$_j$ beveelt [PRO$_j$ naakt$_{*i/j}$ te dansen]
          that John Mary orders nude to dance
      b. dat Jan$_i$ Marie$_j$ vraagt [PRO$_i$ naakt$_{i/*j}$ te mogen dansen]
          that John Mary asks nude to be-allowed to-dance

Whatever controls PRO in these examples is also the semantic subject of the secondary predicate *naakt*. This indicates that the external Θ-role of this predicate can only be assigned to PRO. The contrast between (61) and (62) follows straightforwardly from the notion of function identification. When the external thematic function of the secondary predicate is copied onto a node that has a Θ-role independently, it collapses with this Θ-role. Hence, the argument to which this Θ-role is assigned must also be the subject of the secondary predicate. In (62) the external Θ-role of *naakt* collapses with the Θ-role assigned to PRO, as desired.[14] In (61), on other hand, no function identification takes place within the adjunct, as *while* does not introduce thematic functions. Consequently, the depictive's external role will not collapse with another thematic function until it reaches the VP headed by *arrive*.

---

available for object-oriented predicates. These must therefore be analyzed in a different way, thus blocking complete unification.

[14] The secondary predicate must be attached lower than PRO. See section 9 for an explanation.

In general, then, the radius of application of a function is limited by the fact that it will be identified with any instance of the same function that dominates it. Extrapolating from the result obtained for thematic functions, we derive the following theorem:[15]

(64) *Theorem XII*: The maximal domain in which a function $\mathcal{F}{\downarrow}$ can be satisfied is determined by the first dominating node containing another instance of $\mathcal{F}{\downarrow}$.

This theorem will form the starting point of our discussion of locality in section 9.

8. Derived subjects

8.1 A uniform analysis of derived and non-derived subjects

So far we have implicitly assumed two notions of subject. Some subjects are licensed by thematic functions, whereas others are licensed by the function *Move*$\downarrow$. One can only generalize over these cases by assuming a structural subject position. But given that the notion of subject as external argument is independently motivated (see Williams 1980 and the discussion of secondary predication above), this would constitute an anomaly in the optimal theory.

It is important to realize that this problem originates in the premise that traces uniformly introduce the function *Move*$\downarrow$. At least two sets of data suggest, however, that this is not correct for NP trace. As discussed in Burton & Grimshaw (1992), it is possible to coordinate predicates with VPs which contain the trace of A-movement. Although these authors restrict their data to coordination of transitive and unaccusative VPs, it is also possible to coordinate an AP predicate with a passive VP, as witnessed by the Dutch example below.

(65) Jan sjokt [[$_{AP}$ dronken] en [$_{VP}$ door Marie *t* in de steek gelaten]] door de stad.
John trudges drunk and by Mary in the lurch left through the town

The coordinate structure constraint states that coordinates must be of the same semantic type (although not necessarily of the same syntactic category, cf. Sag et al. 1985). Given that the AP *dronken* is a predicate, this implies that the passive VP *door Marie in de steek*

---

[15] Although locality effects as a result of function identification should also be derivable in unification-based approaches, we are not aware of any specific proposals to this effect.

gelaten must be a predicate as well. In the context of the proposal developed here, the AP and the VP must introduce the same function, namely an external thematic role. After coordination, these functions undergo identification. The resulting thematic function is satisfied by the subject *Jan*.

If it is indeed the case that the passive VP in (64) introduces a thematic role, we must account for its origin. As the passive verb's single thematic function is satisfied by the trace in (64), the function that turns VP into a predicate cannot have been copied from the verb. We are therefore led to the conclusion that it must find its source in the trace. This echoes a similar conclusion drawn by Williams (1986, 1994), according to whom traces of A-movement are a syntactic means of externalizing an internal Θ-role.

Some corroborating evidence for this claim is found in constructions of prenominal modification. As we have already seen, prenominal modifiers are interpreted through function identification. They have an external Θ-role which is identified with the external Θ-role of the noun (see (59)). This implies that passive participles derived from unergatives cannot be used prenominally.[16] If passive morphology absorbs the verb's external (and only) Θ-role, no thematic function remains that can be identified with that of the noun. Indeed, although Dutch freely allows impersonal passives (cf. 65a), impersonal passive participles are barred from occurring prenominally (cf. 65b).

(66) a. Er wordt overal geslapen.
     there is everywhere slept
   b. *de geslapen hond
     the slept dog

By contrast, participles of transitive verbs can occur in this position, as (66) shows.

(67) de door Jan *t* geslagen hond
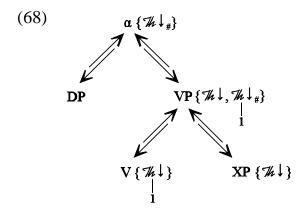   *the by John beaten dog*

Apparently, the prenominal constituent in (66) does have an external Θ-role. Since the passive participle does not have an external Θ-role, and since its single remaining Θ-role is assigned internally, we are again forced to the conclusion that it is the trace which contributes the thematic function identified with that of the noun. This need not surprise us, given that thematic functions and the function introduced by the trace of A-movement

---

[16] For discussion of the thematic properties of participles, see Haider (1984), Hoekstra (1986) and Ackema (1995).

are satisfied by the same kind of object (namely a syntactic argument). The null hypothesis, then, is that the functions in question are identical.
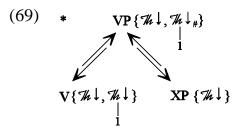
8.2 The anatomy of raising

Let us now consider the syntax of NP raising in more detail. (67) is the structure projected by a monadic unaccusative verb. The internal Θ-role is copied to the VP node, where it is satisfied by the XP in object position, which is either an NP trace (as in *John arrives t*) or a constituent containing such a trace (as in *Surprisingly, John appears [t to have passed the exam]*). XP itself introduces a new Θ-role, which is copied to VP as well. Note that this Θ-role must be external on VP: it cannot collapse with any internal Θ-role (see below) and linking to the ordering tier on an ad hoc basis is ruled out by compositionality. Thus, the presence of a complement introducing a Θ-role implies that an unaccusative verb can still project a predicative VP, and therefore license a subject. As will be clear, the fact that the VP in (67) is a Θ-role assigner explains why it behaves like a predicate in contexts like (64) and (66).

(68)

$$\alpha\,\{\mathcal{T}\!h\!\downarrow_{\#}\}$$

$$\text{DP} \qquad \text{VP}\,\{\mathcal{T}\!h\!\downarrow,\mathcal{T}\!h\!\downarrow_{\#}\}$$
$$_1$$

$$\text{V}\,\{\mathcal{T}\!h\!\downarrow\} \qquad \text{XP}\,\{\mathcal{T}\!h\!\downarrow\}$$
$$_1$$

Although the trace of NP raising is like a secondary predicate in that it contributes a Θ-role, these elements are fundamentally different from the perspective of the licensing criterion (cf. (52)). So far, the cases ruled out by the licensing criterion involved a daughter satisfying two functions in its mother. However, (52) also restricts identification of licensing functions: if two functions collapse, the number of unsatisfied functions is reduced by one.

Crucially, NP trace is an argument: it satisfies a Θ-role in the dominating node. This implies that the Θ-role it contributes cannot be identified with one in its mother, as that would lead to a second reduction in the number of unsatisfied licensing functions. For example, in (68) below the thematic function introduced by XP, which either is or

contains an NP trace, has collapsed with the verb's external Θ-role. This violates (52). Thus the explanation offered earlier for the ban on raising to Θ-positions can be maintained even if certain *Move* ↓ functions are reanalyzed as *Th* ↓ functions.

(69)   *   VP{*Th*↓, *Th*↓#}
                              |
                              1

        V{*Th*↓, *Th*↓}   XP {*Th*↓}
                |
                1

A secondary predicate, on the other hand, does not satisfy a thematic function in the dominating node, so that the function it contributes is free to collapse with a function of the verb (cf. 57).

A consequence of the proposed analysis is that there can be no raising to object position. This would require that the Θ-role introduced by NP trace ends up as an internal Θ-role after one or more steps of copying. But this in turn implies that it is either linked to the ordering tier or identified with a Θ-role already linked to it. As explained, the former option is excluded by compositionality, whereas the latter violates the licensing criterion.

The proposal that NP trace introduces a thematic function allows a unified characterization of subjects. Like base-generated subjects, derived subjects satisfy the external Θ-role of a predicate. The difference between the two cases is the source of the external Θ-role: in one case this is the verb, in the other a trace. Thus, this proposal adheres to the null hypothesis that one notion of subject suffices.[17]

---

[17] On this view of subjects, the question arises what the status is of expletives. For *there*-type expletives the standard theory can be maintained: after (covert) raising of the associate to the expletive, a configuration results in which the trace left behind by movement of the associate introduces a Θ-role satisfied at LF by the associate itself. Chomsky 1981 argues that weather *it* is not an expletive but a pseudo-argument. A similar argument can be made with respect to expletive *it* in examples like *It is generally believed that Greeks like good food.* Bennis (1986) shows that in constructions of this type the verb's Θ-role is satisfied by *it*, which is itself related to an adjunct clause by an interpretational process.

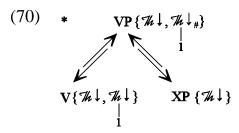## 9. Locality

9.1 Preliminaries

In the preceding sections we have derived three of the four properties which Koster argues constitute the configurational matrix: obligatoriness, uniqueness of the antecedent, c-command of the antecedent and locality. We have seen that grammatical relations are established obligatorily, that the antecedent of an external grammatical relation is unique and that antecedents c-command their dependents. That leaves us with the observation that grammatical relations exhibit locality effects. In our opinion, locality is the most problematic aspect of the configurational matrix, as locality requirements vary for different grammatical relations. Koster deals with this variation by assuming a universal locality condition plus marked, language-specific, domain extensions. Although this approach has its attractions, it raises the problem why other aspects of the configurational matrix cannot be 'extended' in the same way. For example, it does not seem to be the case that the c-command requirement is ever weakened on a language-specific basis. Indeed, such variation is not expected if c-command is part of the invariant core of language. Similarly, if locality is hard-wired, the option of domain extensions is surprising.

It is conceivable that locality is not a unitary phenomenon. For example, it may be that certain nodes do not tolerate functions of a particular type and therefore block copying of such functions. Locality effects that result from this do not seem to be part of the configurational matrix per se. Similarly, contrast between rightward and leftward movement with respect to locality cannot be characterized in terms of the (invariant) configurational matrix. Rather, it seems that such contrast must be attributed to the organisation of the human parser. Incrementality, in conjunction with a filler-driven strategy for the postulation of a trace, strongly disfavours rightward movement. Such movement implies that the parser has to go back into the already analyzed string in order to insert the trace of the moved element, in violation of incrementality in its strictest sense (cf. Rochemont 1992 amongst others).

9.2 Relativized minimality

The question that presents itself at this point is which locality conditions *are* part of the configurational matrix. We have established earlier that one property of syntax, namely the identification of unordered functions, gives rise to locality effects. In particular, a function cannot be copied individually beyond a node that contains the same function. This was expressed by theorem XII, according to which the maximal domain in which a

function $\mathscr{A}\downarrow$ can be satisfied is determined by the first dominating node containing another instance of $\mathscr{A}\downarrow$.

Certain aspects of relativized minimality (cf. Rizzi 1990) can be derived from this theorem. A case in point is the ban on superraising: as is well-known raising to subject is not possible across a subject. Crucially, the thematic function introduced by the trace of NP raising cannot be copied beyond a subject, as it would automatically be identified with the Θ-role assigned to the subject. The only way in which identification could be prevented is by ad hoc ordering of the verb's external Θ-role and that contributed by the trace. Given that this is ruled out by compositionality, (69), where XP either is or contains an NP trace, is the only possible resulting structure.

(70)   *   VP$\{\mathscr{Th}\downarrow, \mathscr{Th}\downarrow_{\#}\}$
               |
               1

V$\{\mathscr{Th}\downarrow, \mathscr{Th}\downarrow\}$ XP $\{\mathscr{Th}\downarrow\}$
     |
     1

Even if (69) were grammatical, it would constitute a case of raising to a Θ-position, rather than NP raising across a subject. But as we have already seen raising to a Θ-position is ruled out by the licensing criterion: the formation of the VP-node in (69), which is identical to (68), involves a reduction of licensing functions by two. The complement satisfies one of the Θ-roles of the verb, while at the same time its own Θ-role is identified with the verb's external one.
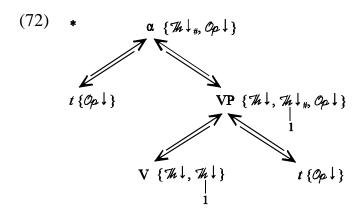
In conclusion the ban on superraising has the same source as that on non-local predication. For the same reason that a predicate must find a local subject (cf. 62), NP movement may not cross a local subject. Both relations must satisfy theorem XII.

The approach to locality sketched so far can be extended to A'-dependencies: WH movement across a WH operator, for example, is impossible. As in the case of NP raising, this can be understood if we look more closely at the nature of the function introduced by the trace. Relativized minimality effects can only be understood if what satisfies the function involved is shared by all WH traces (and perhaps by traces of other operator movements). If the relevant function was really looking for a copy, we would not expect interference from intervening operators. The copy of *what*, for example, is not identical to the full copy of *who*, so that neither could create a minimality barrier for the other.[18]

---

[18] Some notion of underspecification is required in any theory that employs the notion potential intervening antecedent for a trace, as the potential antecedent may not be a possible antecedent for that

(71) *What do you wonder [who *t* read *t*]?

What we propose instead is that the function introduced by a WH trace – let us call it $\mathcal{O}p\downarrow$ – looks for a WH operator (or perhaps just for any operator). On this view two $\mathcal{O}p\downarrow$ functions that percolate up from WH traces to the same node must be identified. The result is that it is not possible to copy the function introduced by a WH trace past another WH trace. As before, the only way in which identification can be prevented is by an ad hoc ordering of the $\mathcal{O}p\downarrow$ functions contributed by the traces, but this violates compositionality. Hence, (71) is the only possible structure for *t read t* in (70).

(72)   *

$$\alpha\ \{\mathcal{T}h\downarrow_{\#},\mathcal{O}p\downarrow\}$$

$$t\ \{\mathcal{O}p\downarrow\} \qquad \text{VP}\ \{\mathcal{T}h\downarrow,\mathcal{T}h\downarrow_{\#},\mathcal{O}p\downarrow\}$$
$$|$$
$$1$$

$$\text{V}\ \{\mathcal{T}h\downarrow,\mathcal{T}h\downarrow\} \qquad t\ \{\mathcal{O}p\downarrow\}$$
$$|$$
$$1$$

Even if (71) were grammatical, it would not constitute a case of WH movement across a WH operator, as the two $\mathcal{O}p\downarrow$ functions have collapsed. In fact, we have already seen that function identification as in (71) violates the licensing criterion: the formation of the node labelled α in (71) involves a reduction of licensing functions by two. The subject satisfies one of the Θ-roles of the verb, while at the same time its $\mathcal{O}p\downarrow$ function is identified with one introduced by the trace in complement position.

   In conclusion, we have been able to derive some aspects of relativized minimality without invoking additional assumptions. Facts like (72) below, however, suggest that relativized minimality cannot in its entirety be reduced to theorem XII. The ungrammaticality of this example does not follow from function identification, as it involves only a single $\mathcal{O}p\downarrow$ function.

(73) *What do you wonder [whether John read *t*]?

---

trace.

WH islands of this type would be accounted for if a function had to be satisfied at the earliest opportunity. In the example at hand, the $\mathcal{O}p\!\downarrow$ function introduced by the trace looks for a WH operator. As it travels up the tree, the first opportunity for satisfaction presents itself in the form of *whether*. If this opportunity must be taken, it will be impossible for the trace to be associated with *what*. Fortunately, this need not be stipulated.

There is a fundamental difference between the copying of a function and its application. It is clear that copying is an operation that transfers information from one node to another. Function application, on the other hand, does not seem to have this status. A function is read off as satisfied if it occurs in a particular configuration, namely one in which the daughter of the node that hosts it is of the desired type. If we take this distinction seriously, function application will take place whenever the required configuration is created. It follows, then, that a function travelling up the tree is satisfied at the earliest opportunity.[19]

(74) *Theorem XIII*:   Function application is automatic.

At first sight, theorem XIII seems to subsume theorem XII. If so, this is unproblematic, as both are derived from the same set of assumptions. In fact, however, the overlap between theorems XII and XIII is fairly limited. Suppose there were no function identification, so that theorem XII did not hold. Furthermore suppose that a node contained two identical licensing functions. By the licensing criterion, no step of merger may lead to a reduction of such functions by more than one. Hence, a node that satisfies one licensing function is not a potential antecedent for the other. This implies that theorem XIII does not apply to constructions ruled out by theorem XII.

On the basis of theorem XIII two further locality conditions can be understood, namely the head movement constraint and principle A of the binding theory. We now turn to these.

Like the functions introduced by WH trace and NP trace, the one introduced by the trace of head movement must be underspecified. If it were to be satisfied by a copy of the moved head, intervening heads would not have a blocking effect, and the head movement constraint would remain underived (see also note 18). Let us assume, then, that traces of X^0 movement introduce the function $\mathcal{H}ead\!\downarrow$, which is satisfied by any X^0 category.

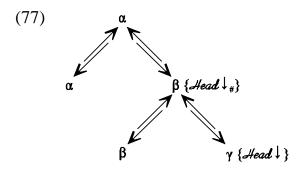[19] The same line of argumentation can be applied to other theories that encode grammatical relations through function application, although we are not aware of specific proposals.

Of course, a function of this type only makes sense if heads can be distinguished from other nodes in the tree. How this can be done is obvious: the head of a node α is that daughter from which it is projected. This characterization implies that the function that encodes head movement is satisfied if below it its finds a copy of the node that hosts it. Thus, in (74) *Head* $\downarrow$ is satisfied by α.

(75)

$\alpha \, \{\mathit{Head} \downarrow_{\#}\}$

$\alpha \qquad\qquad \beta \, \{\mathit{Head} \downarrow\}$

(76)

$\alpha \, \{\mathit{Head} \downarrow\}$

$\alpha \, \{\mathit{Head} \downarrow\} \qquad \beta$

A function cannot be satisfied by the node from which it has been copied (cf. (11)). Consequently, *Head* $\downarrow$ cannot be satisfied by α in (75).

Consider next how the head movement constraint follows from theorem XIII, which states that function application is automatic. In (76) the *Head* $\downarrow$ function is copied to β and satisfied there, since the appropriate configuration obtains. It is hence impossible to relate α to γ using this function.

(77)

α

$\alpha \qquad\qquad \beta \, \{\mathit{Head} \downarrow_{\#}\}$

$\beta \qquad\qquad \gamma \, \{\mathit{Head} \downarrow\}$

This account of the head movement constraint also explains why excorporation from a complex head is impossible for the adjoined category, but allowed for the category to which it is adjoined (cf. Roberts 1990). To see this, consider once again (74) and (75). Suppose that this time these structures represent a complex head rather than a head-complement structure. As indicated in (74), a *Head* $\downarrow$ function introduced by β, the non-head, will be satisfied by α. Consequently, the non-head cannot be excorporated – it cannot be related to a head c-commanding α. By contrast, a *Head* $\downarrow$ function introduced by

α, as in (75), will not be satisfied within the complex head, so that excorporation of α is possible.[20]

9.3 Principle A

Let us now turn to anaphoric binding. At first sight, this relation does not adhere to relativized minimality. For instance, the anaphor *himself* in (77a) is bound by *Bill*, but in the same context it can be bound by *John*, as (77b) shows. Apparently, the function involved in binding can skip a potential antecedent.

(78) a. John showed Bill$_i$ himself$_i$ (in the mirror)
     b. John$_i$ showed Bill himself$_i$ (in the mirror)

On the other hand, it is not the case that this function, $\mathcal{Self}\downarrow$, can skip any potential antecedent. As shown by the examples in (78) it cannot cross a subject.

(79) a. John expected Bill$_i$ to invite himself$_i$
     b. *John$_i$ expected Bill to invite himself$_i$

Several authors have suggested that in one sense or another binding theory is parasitic on Θ-theory. Williams (1980) argues that binding is subject to the predicate opacity condition, according to which an anaphor cannot occur free in a predicate. This provides a natural explanation for the data in (77) and (78). Reinhart & Reuland (1993) propose a theory of reflexivity which requires that an anaphor be bound by a co-argument. Again this explains the contrast between (77) and (78).

---

[20] As in other theories, the question arises to what extent head movement can be an operation of adjunction. In (74), the function $\mathcal{Head}\downarrow$ is satisfied by α. At first sight this implies that it cannot be used to relate a trace of head movement to a category adjoined to α, say γ in [$_\alpha$ γ α]. If this line of argumentation is correct, head to head adjunction is ruled out. There are at least two ways in which this issue can be approached. One could try to argue that the properties of γ are in some way represented on α (cf. Baker 1988 and much subsequent work). Alternatively, one could adopt a more radical approach, according to which head movement is always substitution. Indeed, it has been argued that movement to a functional head is an operation of substitution (cf. Ackema et al. 1992, Nash & Rouveret 1996 and Grimshaw 1997). It has also been argued that structures in which a head is adjoined to another head are not derived by movement (cf. Rosen 1989 and Van Riemsdijk 1998).

In the theory developed here, these data can be understood if we assume that the function $\mathcal{S}elf\downarrow$ looks for an unsatisfied thematic function. This proposal goes back to Williams's (1995) claim that anaphors take Θ-roles as antecedents. In the present terminology, an anaphoric function is satisfied when it is copied to a node whose daughter is predicative in nature. A suggestive piece of evidence for this claim is provided by the Dutch example below.

(80) een [[zichzelf respecterend] mens]
    a himself respecting person

As argued previously, prenominal modifiers are predicates whose external thematic role is identified with the R-role of the noun (cf. Higginbotham 1985). (79) therefore does not contain a subject that can bind the anaphor. Consequently, examples of this type are problematic for standard approaches to binding. They are expected, however, if anaphors take Θ-roles as their antecedent. More specifically, *zichzelf* in (79) relies for its interpretation on the external Θ-role of the prenominal modifier.

   With the above in mind, let us return to the examples in (77). These examples have the structure in (80), in which the verb projects three thematic roles.

(81)

The function $\mathcal{S}elf\downarrow$ is introduced by the direct object and copied onto the dominating node, where it is satisfied by one of the thematic functions on V. There is some optionality as to which thematic function satisfies $\mathcal{S}elf\downarrow$: this can either be the verb's external Θ-role or the Θ-role labelled 1. In the former case, the apparent antecedent is the subject; in the

latter, the indirect object. However, the thematic function labelled 2 is not accessible to $Self\downarrow$ as this would ultimately make the anaphor dependent on itself. What the function $Self\downarrow$ expresses is that the interpretation of an anaphor is linked to a particular thematic role. However, if that thematic role is itself satisfied by the anaphoric expression, a circularity of interpretation results similar to that captured by the *i*-within-*i* condition of Chomsky 1981.[21]

As just mentioned, Williams (1980) accounts for the locality effects observed in (78) by invoking the predicate opacity condition. We can take this proposal a step further, as the opacity of predicates reduces to one more instance of theorem XIII. As soon as $Self\downarrow$ can be satisfied, it must be. Therefore, in the case at hand, the apparent antecedent must be a co-argument of the anaphor.

This account of the locality of anaphoric binding predicts that there are conditions under which the apparent antecedent is not a co-argument. These conditions obtain if the function $Self\downarrow$ is copied to a node whose daughter does not contain an accessible thematic function, as will be the case if $Self\downarrow$ is copied from an exceptionally case-marked subject. The only unsatisfied Θ-role on the sister of an exceptionally case-marked subject will be the one assigned to it, but this role is inaccessible. Hence, $Self\downarrow$ can be copied upward, so that it will ultimately be satisfied by a Θ-role of the matrix predicate. This sequence of operations is illustrated in the tree below:

(82)

Structures of this type are indeed possible. An example in given in (82).

(83) John$_i$ expected himself$_i$ to be the first to leave

A second configuration in which the apparent antecedent of an anaphor is not a co-argument is illustrated in (83a). Here, the anaphor is the object of a preposition that does not take any other arguments. The function introduced by the anaphor can therefore be copied out of the PP and be associated with a Θ-role of a higher predicate. Confirmation for this analysis comes from data like (83b), in which the anaphor is selected by a preposition that also takes a subject (namely *a snake*). Since in this case there is an accessible Θ-role within the PP, the anaphor must take this role as its antecedent. The example in (83b) is hence excluded.

(84) a.   John$_i$ talks a lot [about himself$_i$]
      b.   *John$_i$ saw a snake [near himself$_i$]

In sum, following Williams (1994), we have argued that the antecedent of an anaphor is a Θ-role. It has been shown that the locality condition that restricts binding (the predicate opacity condition) can be reduced to another instance of relativized minimality. If this is indeed the case, the extensions of the notion of syntactic co-argument as proposed by Reinhart & Reuland (1993) fall out naturally.

  Of course, in addition to principle A, the classic binding theory contains principles B and C. These principles do not fall out from the configurational matrix, which means that an explanation for them must be found external to syntax proper. In fact, this is the line taken by Reinhart (1983), Grodzinsky and Reinhart (1993) and Reuland (1997). Although we cannot discuss this in detail, we believe that proposals along these lines can be made compatible with the theory developed here.
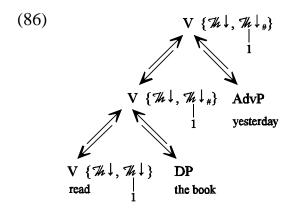
9.4 Adjunct islands

All locality effects discussed so far involve relativized minimality. However, there are other types of locality which presumably involve the barrierhood of non-Θ-marked constituents. In general, non-selected constituents are islands. This restriction follows from the licensing condition if we are more precise about the distinction between categories and segments (see Chomsky 1986a and subsequent work).

  We define category as in (84).

(87)   *   V $\{\mathcal{T}\!h\!\downarrow, \mathcal{T}\!h\!\downarrow\}$
                        |
                        1

V $\{\mathcal{T}\!h\!\downarrow\}$ XP $\{\mathcal{T}\!h\!\downarrow\}$
|
1

Since the two V-nodes in (86) contain a different set of licensing functions, they count as separate categories. Therefore, by the licensing criterion, the creation of the top V-node should have resulted in a reduction in the number of unsatisfied licensing functions. As it has not, the structure is ungrammatical.

The possibility of function identification is illustrated in (87). In this case, the two V-nodes form a multi-segmented category. Since no category has been created, the licensing criterion is satisfied vacuously.

(88)   V $\{\mathcal{T}\!h\!\downarrow\}$
                       |
                       1

V $\{\mathcal{T}\!h\!\downarrow\}$ XP $\{\mathcal{T}\!h\!\downarrow\}$
|
1

In conclusion, the external Θ-role of a predicate merged as an adjunct is identified with a Θ-role of the verb. We demonstrated earlier that this type of identification is *possible* with secondary predicates (cf. (88a)). We have now derived that it *must* take place (cf. (88b)). This captures the long-standing observation that secondary predication is parasitic on primary θ-role assignment (compare, for example, Chomsky's (1986b: 97-98) discussion of the Θ-criterion and Williams's (1994: 47) discussion of relativized heads).

(89) a.   John left the room angry
     b.   *John seemed [that it rained] angry

Note that, as a further consequence, we also derive that secondary predicates cannot be attached higher than the highest argument in the projection that hosts them. In (89), *dronken* 'drunk' must copy its Θ-role to the node labelled α. Since this node does not contain any unsatisfied Θ-roles, function identification cannot take place, in violation of the licensing criterion.

(90) *Jan$_i$ zei dat [$_\alpha$ dronken$_i$ Marie door de stad liep]
      John said that drunk Mary through town walked

The licensing criterion has a wider domain of application than the Θ-criterion: it applies to movement functions as well as thematic roles. Consequently, we make the further prediction that movement out of an adjunct is parasitic on the copying of movement functions in the main projection line. If a movement function which is copied from an adjunct does not collapse with an identical function introduced by the verb, a new category will be created. This category inevitably violates the licensing criterion as the number of unsatisfied licensing functions is not reduced. Hence the data in (90) mirror those in (88).

(91) a. Which book did you file *t* [before you read *e*]
      b. *Which book did you sleep [before you read *e*]

We also derive the anti-c-command condition on the licensing of parasitic gaps (cf. Taraldsen 1981 and subsequent work). If the primary gap c-commands the parasitic gap, function identification is impossible. Copying of the parasitic movement function to α in (91) leads to a violation of the licensing criterion as α does not contain the movement function introduced by the primary gap.

(92) *Who did you explain [*t* [$_\alpha$ met you [before you recognized *e*]]]

Furthermore, since the licensing criterion applies to licensing functions only, we do not expect to find island phenomena with the functions introduced by anaphors or negative polarity items. The following examples confirm this prediction: the anaphor in (92a) and the negative polarity item in (92b) are contained in an island that excludes their respective antecedents.

(93) a. [The scientists]$_i$ made disapproving noises [during each other$_i$'s speeches]
      b. No scientist made disapproving noises [during any of the speeches]

Although we have not attempted to give a comprehensive account of locality, we conclude that relativized minimality and the adjunct island condition belong to the configurational matrix. The joint effect of theorems XII and XIII is that grammatical relations are restricted in their scope by the intervention of potential antecedents (relativized minimality). The licensing criterion restricts copying of functions from non-selected

categories to those environments in which function identification can take place (the adjunct island condition). The relevant data find a parallel in the syntax of secondary predicates. With this final result, we have derived the four properties of the configurational matrix from a minimal theory of syntax.

References

Abney, S. (1987). *The English noun phrase in its sentential aspect*. PhD dissertation, MIT.

Ackema, P. (1995). *Syntax below zero*. PhD dissertation, Utrecht University.

Ackema, P., A. Neeleman & F. Weerman. (1992). Deriving functional projections. *Proceedings of NELS 23, Volume one*, Amherst: Umass.

Baker, M. (1988). *Incorporation*. Chicago: University of Chicago Press.

Bennis, H. (1986). *Gaps and dummies*. Dordrecht: Foris.

Bresnan, J., ed., (1982). *The mental representation of grammatical relations*. Cambridge: MIT Press.

Brody, M. (1998a). Projection and phrase structure. *LI* 29: 367-398.

Brody, M. (1998b). *Mirror theory*. Ms. University College London.

Burton, S. & J. Grimshaw. (1992). Coordination and VP-internal subjects. *Linguistic Inquiry* 23: 305-313.

Chomsky, N. (1955). *The logical structure of linguistic theory*. Ms. MIT. Published by Plenum Press, 1975.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, N. (1986a). *Barriers*. Cambridge: MIT Press.

Chomsky, N. (1986b). *Knowledge of language*. New York: Praeger.

Chomsky, N. (1995a). Bare phrase structure. In G. Webelhuth (ed.) *Government and binding theory and the minimalist program*. Oxford: Blackwell, pp. 383-439.

Chomsky, N. (1995b). *The minimalist program*. Cambridge: MIT press.

Epstein, S. (1995). Un-principled syntax and the derivation of syntactic relations. Ms. Harvard University.

Gazdar, G., A Klein, G. Pullum and I. Sag (1985). *Generalized phrase structure grammar*. Cambridge: Harvard University Press.

Grimshaw, J. (1990). *Argument structure*. Cambridge, Mass.: MIT Press.

Grimshaw, J. (1991). Extended projection. Ms. Brandeis University, Waltham, Mass.

Grimshaw, J. (1997). Projection, Heads and Optimality. *Linguistic Inquiry* 28: 371-422.

Grodzinsky, Y. and T. Reinhart. (1993). The innateness of binding and coreference. *Linguistic Inquiry* 24: 69-101.

Haider, H. (1984). Was zu haben ist und was zu sein hat. *Papiere zur Linguistik* 30: 23-36.

Higginbotham, J. (1985). On semantics. *Linguistic Inquiry* 16: 547-593.

Hudson, D. (1987). Zwicky on heads. *Journal of Linguistics* 23: 109-132.

Hoekstra, T. (1986). Passives and participles. In F. Beukema & A. Hulk (eds.) *Linguistic in the Netherlands 1986*. Dordrecht: Foris, pp. 95-104.

Jackendoff, R. (1990). *Semantic structures*. Cambridge, Mass.: MIT Press.

Koster, J. (1987). *Domains and dynasties*. Dordrecht: Foris.

Nash, L. and A. Rouveret (1997). Proxy categories in phrase structure theory. *Proceedings of NELS* 27.

Pollard, C. and I. Sag (1987). *Information-based syntax and semantics; volume 1*. Stanford: CSLI.

Pollard, C. and I. Sag (1994). *Head-driven phrase structure grammar*. Stanford: CSLI.

Progovac, L. (1994) *Negative and positive polarity; A binding approach*. Cambridge University Press.

Reinhart, T. and E. Reuland (1993). Reflexivity. *Linguistic Inquiry* 24: 657-720.

Reuland, E. (1997). Primitives of binding. Ms Utrecht University.

Rizzi, L. (1990). *Relativized minimality*. Cambridge: MIT Press.

Roberts, I. (1990). Excorporation and minimality. *Linguistic Inquiry* 22: 209-218.

Rochemont, M. (1992). Bounding rightward A'-dependencies. In H. Goodluck and M. Rochemont (eds.) *Island constraints*. Dordrecht: Kluwer, pp. 373-397.

Rosen, S. (1989). Two types of noun incorporation: a lexical analysis. *Language* 65: 294-317.

Sag, I., G. Gazdar, T. Wasow & S. Weisler. (1985). Coordination and how to distinguish categories. *Natural Language and Linguistic Theory* 3: 117-171.

Steedman, M. (1993). Categorial grammar: tutorial overview. *Lingua* 90: 221-258.

Stowell, T. (1981). Origins of phrase structure. PhD dissertation MIT.

Stowell, T. (1983). Subjects across categories. *The Linguistic Review* 2: 285-312.

Taraldsen, K.T. (1981). The theoretical interpretation of a class of marked extractions. In A. Belletti, L. Brandi and L. Rizzi (eds.) *Theory of markedness in generative grammar, Proceedings of the 1979 GLOW conference.* Scuola Normale Superiore, Pisa, pp. 475-516.

Van Gestel, F. (1986). *X-bar grammar.* Dordrecht: Foris.

Van Riemsdijk, H. (1998). Head movement and adjacency. *Natural Language and Linguistic Theory* 16: 633-678.

Williams, E. (1980). Predication. *Linguistic Inquiry* 11: 208-238.

Williams, E. (1981). Argument structure and morphology. *The Linguistic Review* 1: 81-114.

Williams, E. (1983). Against small clauses. *Linguistic Inquiry* 14: 287-308.

Williams, E. (1986). NP trace in theta theory. *Linguistics and Philosophy* 4: 433-447.

Williams, E. (1989). The anaphoric nature of θ-roles. *Linguistic Inquiry* 20: 425-456.

Williams, E. (1994). *Thematic structure in syntax*. Cambridge, Mass.: MIT Press.