

# INTRODUCTION

*C. Jan-Wouter Zwart*

This 37th volume of the *Groninger Arbeiten zur germanistischen Linguistik* presents a collection of working papers on Germanic syntax, written within the framework of, inspired by, or addressing critically the stage of generative grammar that has come to be called *the Minimalist Program*, after the title of Chomsky (1992). As the title of this volume betrays, we are taking the minimalist approach to include the generalizations concerning word order typology presented in Kayne (1992, 1993).

For the readers' convenience, I will include here an introduction to the Minimalist Program, which is based on Chomsky (1992) and Kayne (1993).<sup>1</sup>

## 1 Introduction to the Minimalist Program

The *Minimalist Program* (Chomsky 1992) is the latest developmental stage of the theory of (Transformational) Generative Grammar (Chomsky 1957).<sup>2</sup> Its main goal is to derive all conditions on derivations and representations from so-called *bare output conditions*, i.e. from conditions on the representations that constitute the interfaces of the grammatical component with other components of the cognitive system. In this respect, it is properly characterized as a logical consequence of earlier stages of the theory, arrived at by way of sharpening some notions that were relevant in the earlier stages, and by eliminating certain other notions that turned out to be redundant in the process.

As in earlier stages of the theory, the Minimalist Program considers grammar to be a *derivational system*. A sentence is first built up in a basic form, then modified through processes of movement, deletion, and insertion, until it reaches a final form which may serve as input to other components of the cognitive system. However, unlike earlier stages of the theory, the mechanism creating the basic representation and the mechanism performing the other operations (movement, insertion, deletion) are the same (it is the mechanism of *Generalized Transformation*).

As in earlier stages of the theory, movement takes place because elements must be *formally licensed*. Unlike earlier stages of the theory, however, the need for formal licensing is the *only* reason for movement to take place. In addition, it is assumed that elements can *never* be formally licensed in a position they occupy in the initial representation.

As in earlier stages of the theory, movement may take place before or after the point in the derivation at which the instructions to the PF-system (the articulatory-perceptual system) are issued. Unlike earlier stages, however, it is now assumed that movement preferably takes place *after* this particular point in the derivation, so that overt movement is, in a way, the marked option.

<sup>1</sup> The material of this introduction is taken from Chapter 1 of Zwart (1993), with minor adaptations.

<sup>2</sup> Earlier stages that can be distinguished are *the Standard Theory* (Chomsky 1965), *the Extended Standard Theory* (Chomsky 1970; Jackendoff 1972), *the Revised Extended Standard Theory* (Chomsky 1973, 1977; Chomsky and Lasnik 1977; Chomsky 1980), *the Government and Binding Theory or the Principles and Parameters Approach* (Chomsky 1981, 1986a, 1986b).

As before, the amount of overt movement may differ from language to language. But, unlike before, the presence or absence of overt movement is the *only* instance of parametric variation in syntax among languages.

In the next four subsections, the key aspects of the Minimalist Program are briefly sketched. Section 1.5 presents certain aspects of Kayne's typological generalizations concerning word order (Kayne 1993).

### 1.1 Building Up Trees: Generalized Transformation

Representations are built up in a bottom-up fashion by a mechanism called *Generalized Transformation*. A Generalized Transformation combines two phrase markers. Two phrase markers are combined by expanding one of the two phrase markers (the 'target phrase marker') so as to include an empty position. This expansion takes place by adding to the target phrase marker a *projection* of the target phrase marker. This projection is binary branching and has two daughters: the target phrase marker and an empty position. This empty position is substituted for by the other phrase marker. The whole process, illustrated in (1), yields two sister phrase markers connected in a binary branching subtree.<sup>3</sup>

- (1) I. *Two independent phrase markers*
- |                |                 |
|----------------|-----------------|
| V<br> <br>kiss | NP<br> <br>Mary |
|----------------|-----------------|
- II. *Expansion of the target phrase marker*
- |                                   |                 |
|-----------------------------------|-----------------|
| V'<br>/  \<br>V    e<br> <br>kiss | NP<br> <br>Mary |
|-----------------------------------|-----------------|
- III. *Substitution of the empty position in the projection of the target phrase marker*
- |   |
|---|
| V'<br>/  \<br>V    NP<br>      <br>kiss  Mary |
|---|

The projection of the target phrase marker has the same categorial features as the target phrase marker. The phrase level of the projection of the target phrase marker is determined by the rules of X-bar Theory (Chomsky 1986b, going back to Chomsky 1970, Jackendoff 1977).

These rules specify that the ultimate projection of an X (or X<sup>0</sup>, or *head*) will be an XP (or X<sup>n</sup>, or *maximal projection*), and that there is an intermediate projection X' (*X-bar*)

<sup>3</sup> Binary branching is a result of this particular formulation of the Generalized Transformation mechanism. The attractiveness of binary branching has been argued for several times in the literature (Kayne 1984, E. Hoekstra 1991).

which is the immediate projection of X. This is illustrated in the following two rewrite rules:

- (2)        a.    XP → (ZP) X'  
               b.    X' → (YP) X°

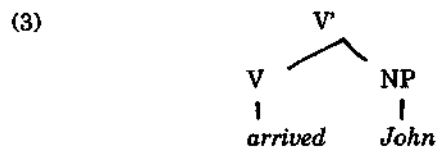
The order of the elements to the right of the arrows in (2) is irrelevant.<sup>4</sup> The sister of X°, YP in (2b), is called *complement*; the sister of X', ZP in (2a), is called *specifier*.<sup>5</sup>

The Generalized Transformation illustrated in (1) combines two independent phrase markers. Therefore, it is called a *binary* operation. Lexical insertion is a typical binary operation.

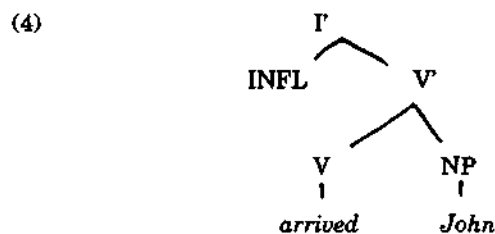
It is also possible that the empty element created by expanding the target phrase marker is substituted for by an element *contained in* the target phrase marker. This would be called a *singular* operation.

Consider a standard case of raising to subject, as in *John arrived*. In this type of construction, *John* is generated as a complement of *arrived*, and moves to the subject position at some point in the derivation (Burzio 1981, Chomsky 1981).

A binary operation of the Generalized Transformation will first combine *arrived* and *John*, as in (3).



Next, another binary operation will combine the phrase marker in (3) with a functional head in which the tense and agreement features are represented (called *INFL*, for the time being).

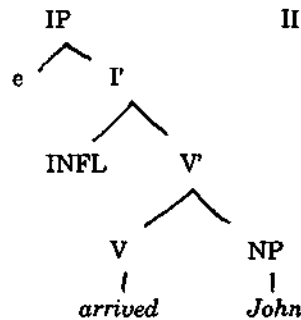


For reasons that do not concern us here, *John* has to move out of the projection of V to a position in the domain of INFL. To this end, I' is expanded in such a way that there will be an empty element in the position of sister of I', to be substituted for immediately by *John*.

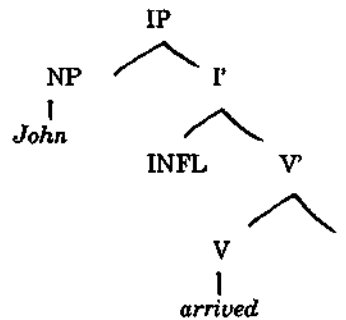
<sup>4</sup> But see section 1.5.

<sup>5</sup> Stuurman (1985) and E. Hoekstra (1991) propose modified versions of X-bar theory, in which the intermediate bar level category disappears and is replaced by a maximal projection. See Zwart (1992b, 1993) for incorporation of the one-level X-bar theory into the minimalist program.

(5) I.



II.



In (5), the target phrase marker is expanded by adding a former subpart of the target phrase marker. No new phrase marker is added to the construction. Therefore this is called a singular operation. All movement operations that were subsumed under the term *Move  $\alpha$*  in the Government and Binding framework are now redefined as singular operations of the Generalized Transformation.<sup>6</sup>

Chomsky (1992) notes that the expansion of a target phrase marker, the introduction of an empty element, and the substitution of that empty element by a second phrase marker, are all part of one indivisible process. The intermediate stages, represented separately above for expository reasons, are never open to inspection as phenomena of language.

Crucially, the Generalized Transformation always adds material *external* to existing phrase markers. It is not possible, Chomsky (1992) suggests, to insert material *inside* a phrase marker.<sup>7</sup>

## 1.2 Licensing Elements: Morphological Feature Checking

A classic distinction exists in linguistic theory between *contentful elements* and *functional elements*. Word stems are contentful elements, whereas inflectional morphemes are functional elements. Functional elements express *agreement relations* between constituents.

In the Minimalist Program, it is assumed that agreement relations are highly local. A maximal projection  $\alpha$  agrees with a head  $\beta$  only if  $\alpha$  is a specifier of  $\beta$ . A head  $\alpha$  agrees with a head  $\beta$  only if  $\alpha$  is adjoined to  $\beta$ .<sup>8</sup> Moreover,  $\beta$  must be a *functional* head.

<sup>6</sup> A note on terminology is in order here. In the earliest stages of Transformational Grammar, a distinction was made between *singular transformations* and *generalized transformations*. The former operate on a single phrase marker, are ordered, and do not introduce meaning-bearing elements; the latter embed a constituent phrase marker into a matrix phrase marker, are unordered, and do introduce meaning-bearing elements (Katz and Postal 1964, Palmatier 1972, and references cited there). In Chomsky (1992, 30f), singular transformations are a subcase of generalized transformations. The two operations work in the same way, the only difference being the origin of the phrase marker substituting for the empty position (the formal identity of generalized transformations and singular transformations was already pointed out in Chomsky 1966:52, cf. also Chomsky 1961:134 note 35). Generalized transformations, especially those governing sentence embedding, have been replaced by the rewrite rules of the base component (Chomsky 1966:65, 1965 chapter 3). Singular transformations gradually developed into *Move  $\alpha$*  (Chomsky 1981).

<sup>7</sup> This, then, is the modern version of the *Strict Cycle Condition* (Chomsky 1973), also called *the extension condition*. Chomsky (1992:33) tentatively restricts the extension condition to substitution in overt syntax.

<sup>8</sup> The locality requirements are further restricted in Zwart (1992b), where it is argued that an element  $\alpha$  agrees with an element  $\beta$  only if  $\alpha$  adjoins to  $\beta$ . This implies that in a specifier-head agreement relation, the specifier does not agree with the head, but with the immediate projection of a head.

In the Government and Binding framework, the distinction between contentful (or *lexical*) elements and functional elements gradually took the following shape.<sup>9</sup> Functional elements are generated as heads of independent phrasal projections. These functional projections are situated outside and on top of the lexical projections. Thus, the inflectional morphemes for tense, person, number, etc., are generated separately from the lexical stems. The stems have to be united with the inflectional morphemes through a process of movement and adjunction.

This yields a sentence structure as illustrated in Figure 1:

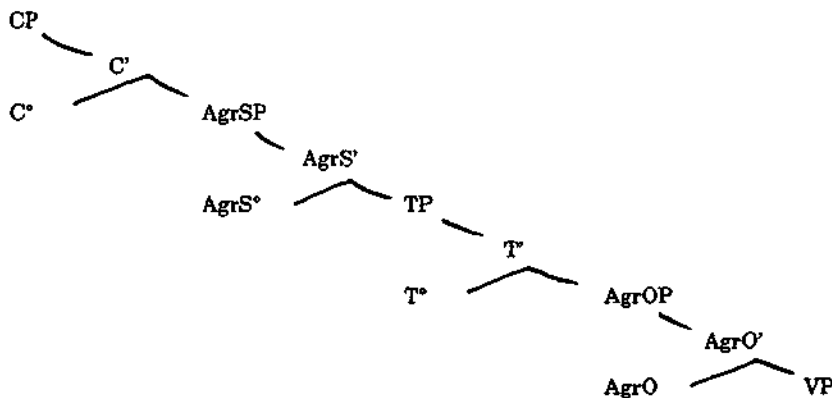


FIGURE 1

In figure 1, *C* stands for the complementizer position, *T* for tense, and *AgrS* and *AgrO* for subject and object agreement morphology, respectively. These functional heads project phrases in accordance with the rules of X-bar Theory given in (2) of section 2.1. *AgrOP*, *TP*, *AgrSP* and *CP* together constitute the functional domain of a syntactic structure, *VP* constitutes the lexical domain.

In the Minimalist Program, this analysis is maintained in a simplified form. The major difference concerns the content of the lexical and functional heads. In the Minimalist Program, lexical heads are occupied by fully inflected forms (stems plus inflectional affixes). These forms carry a *feature* associated with the inflectional affix. The functional heads are likewise occupied by features associated with inflectional morphology (instead of by the inflectional morphology itself).<sup>10</sup>

<sup>9</sup> The developments in the Government and Binding period are marked by Stowell (1981) and Pesetsky (1982), on the structure of IP, Chomsky (1986b) on the structure of CP, Kayne (1987), Pollock (1989) and Chomsky (1991) on the division of IP into AgrP and TP, and Abney (1987) on the functional domain of noun phrases, DP. See Fukui and Speas (1986) on the relevance of functional projections for parametrization.

<sup>10</sup> The assumption that abstract features associated with inflectional morphology are of greater syntactic significance than the overt morphology itself is already a crucial part of the Case Theory module of the Government and Binding framework. This Case Theory refers to *abstract Case features* which are associated with nouns and noun phrases regardless the morphological manifestation of Case on these nouns and noun phrases (Vergnaud 1979, Chomsky 1981). This theory of abstract Case is subsumed under the Minimalist Program. As a result, the inflectional features associated with Case are assumed to be present on lexical categories, even if there is no overt morphological manifestation of Case on these categories.

The features associated with the inflectional morphology of lexical categories have to *match* the features represented in the functional heads. Matching is checked under the same strict locality requirements as agreement (in fact, agreement is a subcase of feature matching). Therefore, the requirement that morphological features match triggers movement of lexical elements to positions in the functional domain. *Licensing* inflected elements consists in moving the inflected elements to positions in the functional domain, and checking whether the features associated with the inflection match the features represented in the functional heads.

Recall that movement is an application of the Generalized Transformation mechanism. The structure in figure 1, therefore, is completely built up in the process of moving elements from the lexical domain to positions in which their features can be checked (which yields the functional domain). There is no top-down rule system to ensure that syntactic structures are always like figure 1. The structure in figure 1 is the result of the fact that inflected elements have to be licensed outside of the lexical domain.

The inflectional features relevant to the phenomena of verb movement and noun phrase movement are tense, agreement, and Case.<sup>11</sup> It is very well possible that other features exist, but these three appear to be indispensable features of sentence structure.<sup>12</sup>

The features represented in the functional heads trigger both head movement (to the functional heads) and XP-movement (to the specifier positions of the functional heads). For this reason, Chomsky (1992) distinguishes two types of features represented in the functional heads: *N-features* and *V-features*. N-features are relevant for checking features of XPs (maximal projections), V-features are relevant for checking features of heads.

The derivation of a sentence consists in these two processes only: insertion of elements from the Lexicon (by a binary operation), and movement of elements to the functional domain (by a singular operation).

### 1.3 Restrictions: Economy, Procrastination, Greed

The derivation of a sentence is subject to general conditions of *economy*. The derivation should take as few steps as possible (*economy of derivation*), and the resulting representations should have as few symbols as possible (*economy of representation*) (Chomsky 1991).

One consequence of economy of derivation is that movement always takes the shortest route.<sup>13</sup> Another consequence is that any movement that is not triggered by a

<sup>11</sup> The exact difference between Case and agreement is not very clear in this system. It is assumed that the specifiers of AgrS and AgrO are the positions for checking Nominative and Accusative Case features, respectively. This suggests that Case and agreement are identical concepts. However, Chomsky (1992:42) suggests that, while Nominative and Accusative Case features are checked in the specifier positions of AgrS and AgrO, respectively, the features relevant for checking Case do not reside in AgrS and AgrO, but in T and V, respectively.

<sup>12</sup> Iatridou (1990) contends that the approach to inflectional morphology sketched here leads to an explosion of functional categories, assuming that every functional category discovered in studying the languages of the world should be present in the grammar of every single language of the world. This does not appear to be sound argumentation, since we cannot conclude, in biology for instance, that every aspect discovered in the study of biological systems should be present in every single species of the world. Yet some biological functions appear to be indispensable in any biological system. Likewise, we may assume that a small number of inflectional features are present in all languages of the world, whereas a larger number may be relevant to specific languages only. What is universal, however, is the way inflectional features determine word order.

<sup>13</sup> *Shortness* can be interpreted in two ways, viz. as involving the smallest number of steps and as involving the shortest steps. These two interpretations appear to be contradictory (cf. Chomsky 1992:21).

well-established requirement of morphological feature checking is excluded.<sup>14</sup> Thus, elements, once licensed, are doomed to inertness.

Economy of representation excludes the presence of irrelevant material at any given level of representation. One instantiation of economy of representation is the principle of *Full Interpretation*, which excludes the presence of uninterpretable material at the interface representations.<sup>15</sup>

The derivation of a sentence is a finite process. At a certain point, the process yields a representation that will function as the output of the grammatical system. This representation will serve as the input to other parts of the cognitive system, for instance those having to do with interpretation. The principle of Full Interpretation requires that every element of an output representation should provide a meaningful input to the relevant other parts of the cognitive system. Only these elements are considered to be *legitimate objects* at the interface level.

The features associated with inflectional morphology are considered to be relevant for syntax only. They play a crucial part in the licensing of inflected elements. However, these features are of no direct relevance to components of the cognitive system external to the grammatical component. In other words, the features associated with inflectional morphology are not legitimate objects at the interface level: they cannot be a part of the final representation that is to serve as input to other components of the cognitive system.

For this reason, these features have to be *eliminated* during the derivation. It is assumed that matching features are eliminated as soon as they are checked.

Therefore, a minimal number of derivational steps is required to achieve a minimal representation at the interface of the grammatical component and other components of the cognitive system.

Two other principles are directly derived from economy of derivation.

First, picture the derivation as a step-wise procedure. At each step, economy of derivation will allow only a minimum of activity. Eventually, movements will have to take place, but economy of derivation dictates that these activities take place as late in the derivation as possible. This can be formulated as a separate principle, *Procrastinate* (Chomsky 1992:43).

Second, movement is triggered by the need to license inflected elements (more exactly, by the need to check off the abstract features associated with inflected elements). Elements that are already licensed, or that do not need licensing, are neither forced nor allowed to move. It follows that such elements can never be forced to move in order to assist in the licensing of another element. The trigger for movement always works directly on the element to be licensed. The principle that movement only to help out other elements is disallowed is called *Greed* (Chomsky 1992:47).

<sup>14</sup> The modification 'well-established' is needed to exclude movements triggered by 'ghost features', whose presence is only motivated in order to account for a specific word order phenomenon.

<sup>15</sup> This principle was first introduced in Chomsky (1986a:98), in the context of a discussion of the relation between Case assignment and theta-role assignment. The idea was that noun phrases must be assigned Case at S-structure, because only then would they be visible for theta-role assignment at LF. Since only noun phrases that carry a theta-role are interpretable at the interface of LF and other components of the cognitive system, the principle of Full Interpretation requires Case assignment at S-structure. This concept has been slightly changed in the Minimalist Program. Case checking eliminates features that cannot be interpreted at the interfaces. Without Case checking, Full Interpretation (and economy of representation) is violated, since unchecked features are uninterpretable.

## 1.4 Parametric Variation: Strength of Features

According to the Minimalist Program, the derivation of a sentence yields interface representations which are subject to the principle of Full Interpretation: they must consist of legitimate objects only. If they do, the derivation is said to *converge*. If not, the derivation is said to *crash*.

The other components of the cognitive system that the grammatical component interacts with are *performance systems*, having to do with, roughly, speech and interpretation. Therefore, there are two types of performance systems: articulatory-perceptual and conceptual-intentional (Chomsky 1992:3). In accordance with this, the grammatical system will yield *two* interface representations, each consisting of instructions for one of the two performance systems. These interface representations are called *PF* (for the articulatory-perceptual performance system) and *LF* (for the conceptual-intentional performance system).

On the assumption that the conceptual-intentional performance system is identical in all humans, the interface representation called LF must be largely identical in all languages. In contrast, the interface representation called PF varies from language to language, as can easily be observed.<sup>16</sup> It follows that the two interface levels PF and LF are not identical.

In the Minimalist Program, it is assumed that the LF interface level is the final stage of a derivation, and that the PF interface level is the reflection of an intermediate stage in the derivation. That is, at a certain point in the derivation, instructions to the articulatory-perceptual system will be issued. This point is called *Spell Out*. The part of the derivation before Spell Out is called *overt syntax*, the part of the derivation after Spell Out is called *covert syntax*.<sup>17</sup>

The problem of comparative linguistics is to find out how and why languages differ in their overt syntax. Recall that the principle of Procrastination dictates that movements take place as late in the derivation as possible. This principle, then, has to be violated to some extent in the grammar of certain, perhaps all, languages. The question is, *Why?*

The only possible answer to this question is that Procrastination must be violated to ensure convergence at the PF interface level. In other words, certain elements that would count as illegitimate objects at PF have to be eliminated in overt syntax. Sticking to the minimalist assumptions made above, it must be the case that certain inflectional features count as illegitimate objects at PF. These features, then, have to be checked and eliminated in overt syntax, through a process of movement of heads and phrases to positions in the functional domain.

The surprising aspect of this mechanism is that *not all* inflectional features count as illegitimate objects at PF. If that were the case, overt syntax would be largely, perhaps completely identical in all languages of the world. As we know, there are very distinct

<sup>16</sup> The point to be made here is actually more subtle. What differs in the PF representation in the various languages is the order of words and phonemes in a string. The way the corresponding instructions are handled by the articulatory-perceptual performance system is just as universal as the way the LF instructions are handled by the conceptual-intentional performance system. The difference between the two interface levels is that word order and/or hierarchical order affects interpretation in the conceptual-intentional system, but not in the articulatory-perceptual system. Therefore, word order must be universal at the LF interface, but not at the PF interface.

<sup>17</sup> There are two significant differences between overt syntax and covert syntax: binary transformations are only allowed in overt syntax (Chomsky 1992:31), and the Strict Cycle Condition does not apply in covert syntax (Chomsky 1992:33).



differences in word order between even so closely related languages as English and French (Pollock 1989).<sup>18</sup>

This, then, appears to be the locus of parametrization between languages: an inflectional feature may or may not be *visible as an illegitimate object* at PF. Those that are visible as illegitimate objects at PF will have to be eliminated in overt syntax. Those that are not visible at PF will *not* be eliminated in overt syntax, by the principle of Procrastination. Features that are *visible* (thus: potentially harmful) at PF are called *strong*; features that are *invisible* (thus: harmless) at PF are called *weak*.<sup>19</sup>

A minimal assumption is that the strong/weak distinction is the *only* instance of parametric variation among languages. This implies that parametric variation is restricted to functional categories (Fukui and Speas 1986). It furthermore implies that there are no directionality parameters, such as directionality of government.<sup>20</sup> The latter implication is supported empirically by Kayne (1993), who argues that movement is always leftward.

### 1.5 Directionality

The structure building process of generalized transformations does not contain a specification of the linear order of head, complement, specifier, and adjunct. Superficial crosslinguistic examination suggests that languages may differ with respect to the linear order of these elements. In the tradition of generative grammar, the attested variation is described in terms of a parametric option: heads may govern to the left or to the right. A head that governs to the left takes its complement to the left in the initial representation, yielding a basic OV structure.

In the minimalist approach, a directionality parameter is no longer available. First, parametric variation must be expressed in terms of the features of functional heads only. A directionality parameter would therefore not suffice to account for the ordering of elements in the lexical domain. Second, government no longer plays a role in the minimalist approach. Therefore, it is unclear whether a directionality parameter could be reduced to properties of an independently established grammatical relation. Third, a directionality parameter would be redundant, since much of the word order variation can be accounted for by the interaction of overt and covert movement.

Kayne (1992) presents empirical evidence showing that movement into the functional domain is invariably leftward. The evidence consists in what we do *not* find, in comparing movement phenomena in the languages of the world. Thus, we can conclude from the general lack of Wh-movement to the right that the specifier position of CP is always to the left. Similarly, there are no known cases where verb movement changes a verb-complement order from VO to OV, which suggests that verb movement to the right does not exist. Hence, the functional projections hosting V-features must all be head initial. Also, the subject precedes the object in almost all languages of the world (Greenberg 1963, Universal 1). Assuming, in connection with this, that AgrSP is hierarchically higher than AgrOP, it also follows that the specifier of AgrSP is situated

<sup>18</sup> English and French are called 'closely related' here not for typological or genetic reasons, but because the same set of functional categories appears to suffice in the description of the two languages.

<sup>19</sup> See Koster 1986, Pollock 1989 for the origin of this terminology.

<sup>20</sup> In fact, *government* has no formal status in the Minimalist Program. For example, Case assignment is reduced to feature checking in a specifier-head configuration. The consequences for the Empty Category Principle, which incorporates the notion *proper government*, have not been fully explored.

to the left. Likewise, if the complement of a preposition is extracted, the complement always ends up to the left of the preposition, never to the right of it. Again, this suggests that licensing positions, i.e. specifier positions, under our assumptions, are on the left hand side.

These empirical observations are presented more fully in Kayne (1993). Kayne (1993) in addition argues that the lexical projections in the world's languages are invariably head initial as well. This is an attractive hypothesis, considering the empirical evidence for the universal structure of the projections of the functional domain. However, empirical evidence in support of this hypothesis is infinitely more difficult to obtain, in view of the fact that the observable word order reflects an intermediate state in the derivation of a sentence. In other words, one never knows whether the constituents are in a basic position or not.

Kayne (1993) also presents conceptual argumentation in support of the idea that all phrases are head initial. Kayne proposes that asymmetric *c-command* invariably maps into linear precedence. In order for this mapping to be successful, it must be possible to express the relations between the nodes of a phrase marker that asymmetrically *c-command* each other into a set of ordered pairs  $\langle x,y \rangle$  of the terminal ('lexical') elements dominated by these nodes. The pairing of two terminal elements  $x,y$  thus expresses a relation between  $x$  and  $y$ . Kayne proposes that the set of ordered pairs of these relations must express a *linear ordering*, i.e. a total, transitive, and antisymmetric ordering.

Thus, according to this proposal it must be possible to read the relation of each terminal element to all other terminal elements off of the set of ordered pairs. Crucially, these relations must be antisymmetric, i.e., it is excluded that two terminal elements  $L$  each other, where  $L$  stands for the relevant relation between these two elements.<sup>21</sup>

The axiom that the set of ordered pairs of terminal elements derived from the set of relations between the nodes of a phrase marker that asymmetrically *c-command* each other is a linear ordering of the terminal elements is called the *Linear Correspondence Axiom (LCA)*.

In addition to the LCA, Kayne proposes that the relation expressed by the pairing of terminal elements is a *precedence* relation. I will refer to this hypothesis as the *Extended Linear Correspondence Axiom (ELCA)*.<sup>22</sup>

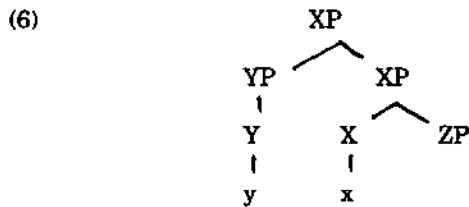
Kayne shows that the adoption of the LCA explains many basic facts of phrase structure, such as binary branching and endocentricity. In this respect, the LCA is compatible with the mechanism of generalized transformations as presented in section 2.1. It follows from the ELCA that adjunction always takes place on the left hand side.

In some respects, however, the LCA appears to be too restrictive, as Kayne notes. In fact, the LCA excludes adjunction of specifiers and adjuncts. Kayne therefore modifies the definitions entering into the notion *c-command* in order to allow adjunction of specifiers. He argues, however, that adjunction of adjuncts (i.e. in addition to adjunction of a specifier) is excluded.

<sup>21</sup> 'Antisymmetry' and 'totality' are two of the three defining properties of linear relations. Kayne assumes that the third defining property, transitivity, also applies to the relations between the terminal elements expressing the relations between the nodes that asymmetrically *c-command* each other.

<sup>22</sup> Kayne (1993:section 5.3) derives the definition of the relation between the terminal elements in a phrase marker as a precedence relation from the hypothesis that every phrase marker contains a root node dominating all other nodes except itself. On the assumption that this root node also dominates an abstract terminal element  $a$ , which, as Kayne argues, has to precede all other terminal elements of the phrase marker, it follows that the linear relation between  $a$  and the other terminal elements is also a precedence relation. Hence, the linear relation between terminal elements must always be a precedence relation. For empirical justification of the Extended LCA, see above.

To see why adjunction of specifiers is difficult, consider the tree structure in (28), where  $y$  and  $x$  represent terminal elements:



Assume the following definition of c-command:

- (7)  $\alpha$  c-commands  $\beta$  iff every  $\gamma$  that dominates  $\alpha$  dominates  $\beta$

C-command is asymmetric where, for  $\alpha$  c-commanding  $\beta$ ,  $\beta$  does not c-command  $\alpha$ .

In (6), YP asymmetrically c-commands X and XP asymmetrically c-commands Y. YP dominates the terminal element  $y$ , and X dominates the terminal element  $x$ . The relation between YP and X therefore can be expressed in the ordered pair of terminal elements  $\langle y, x \rangle$ . But since XP dominates  $x$  and Y dominates  $y$ , the ordered pair of these terminal elements  $\langle x, y \rangle$  is also part of the set of ordered pairs expressing the relations between YP, XP, Y, and X. So now this set contains  $\langle y, x \rangle$  and  $\langle x, y \rangle$ . Hence, the relation between  $x$  and  $y$  (i.e., between a head and its specifier) is not linear, because it is not antisymmetric.

To solve this problem, the pair  $\langle x, y \rangle$  or the pair  $\langle y, x \rangle$  must be excluded. This can be achieved if either YP does not c-command X (kicking out the pair  $\langle y, x \rangle$ ) or XP does not c-command Y (kicking out  $\langle x, y \rangle$ ). Kayne proposes to modify the definition of c-command in such a way that XP no longer c-commands Y. This can be done by excluding segments from the definition of c-command, assuming XP and XP in (6) to be two segments of the same category:<sup>23</sup>

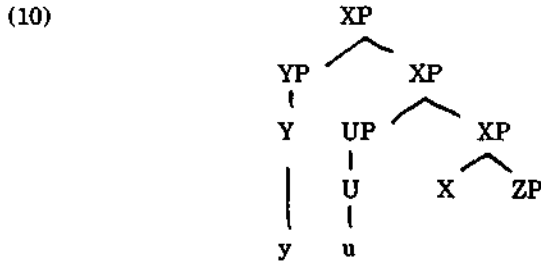
- (8)  $\alpha$  c-commands  $\beta$  iff (i)  $\alpha$  and  $\beta$  are not segments, and  
 (ii)  $\alpha$  excludes  $\beta$ , and  
 (iii) every  $\gamma$  dominating  $\alpha$  dominates  $\beta$
- (9)  $\alpha$  excludes  $\beta$  if no segment of  $\alpha$  dominates  $\beta$

In (6), XP is a segment, hence does not c-command Y by clause (i) of the definition of c-command in (8). This gives the desired result that the relation between  $x$  and  $y$  in (6) is described by  $\langle y, x \rangle$ , hence is a linear relation (hence, following Kayne, a precedence relation).

Notice that the fact that the higher XP is a segment of the lower XP suffices to exclude that the lower XP c-commands Y. Since the higher XP is a segment, the lower XP does not exclude Y, and the c-command relation is barred by clause (ii) of the definition of c-command in (8) (cf. Kayne 1993, note 9).

To see why adjunction is difficult, consider (10):

<sup>23</sup> Following Chomsky 1986b:7.



In (10), not only the specifier UP, but also the adjunct YP is adjoined to XP. YP asymmetrically c-commands U, yielding  $\langle y, u \rangle$ , and UP asymmetrically c-commands Y, yielding  $\langle u, y \rangle$ . Hence the relation between the terminal elements  $y$  and  $u$  is not antisymmetric and therefore (10) is not allowed by the LCA.

Kayne concludes that multiple adjunction (i.e., adjunction of an element in addition to adjunction of a specifier, as in (10)) is universally impossible. It follows that adjunct elements, such as adverbs, can only be present in a structure as specifiers. Thus, for every adjunct there must be a head in the structure creating the required specifier position.

It is important to note that the ELCA does not necessarily follow from the LCA. The ELCA merely interprets linearity in the mathematical sense in one-dimensional terms, yielding precedence. Chomsky (1994) objects to the way Kayne (1993) employs virtual categories (such as intermediate bar nodes) to determine the relation between terminal elements. He therefore rejects the LCA as proposed by Kayne, while accepting the word order generalizations of Kayne in full.

## 2 The Minimalist Approach and Germanic Syntax

At the moment of writing, the linguistic community occupied with Germanic syntax appears to be sharply divided regarding the virtues of the minimalist approach. This is also apparent from some of the papers collected here.

For instance, Wolfgang STERNEFELD, referring to the many leftward movements that Kayne's word order generalizations predict for Germanic, "take[s] it for granted that such a theory is not desirable, perhaps not even feasible, and evidently not appealing as regards simplicity and common sense."

On the other hand, Hartmut CZEPLUCH notes that "there are really no new mechanisms introduced into grammatical theory", since X-bar Theory is maintained (or, in fact, simplified), and verb movement and NP-raising enjoy a long standing tradition in generative grammar.

The point merely seems to be that in the minimalist approach, Germanic has a little bit more of what is already well-known and well-established. Evidently, more of the same does not add to complexity. But even if the minimalist approach were radically different from the government and binding approach, it would be hard to tell which approach would be more appealing as regards common sense. And then, even if the minimalist approach were less satisfactory in this respect, we might contend that a theory that satisfies common sense is probably on the wrong track (see also Koster 1983).

As I have mentioned several times in the above, the minimalist approach is in some sense a natural, though more ambitious, consequence of the government and binding approach. It is more ambitious in that it is an attempt to reach a "bare theory", that is, a theory that does not stipulate any properties of the grammatical system that are not required by conditions on the output of the grammatical component (see Chomsky 1994).

In the course of formulating such a theory, Chomsky proposes a number of descriptive devices, such as N-features and V-features. Given certain assumptions regarding the nature of these features, it follows that movement can only take place in order to eliminate them. This much is new, and determined by the ambition to look at output conditions only. However, the movement needed to eliminate the features is a standard feature of earlier stages of the theory.

More importantly, the idea that movement cannot take place unless it serves some purpose (in other words, unless there is an explanation for the movement) is also a standard feature of the government and binding framework. The Minimalist Program merely restricts the number of purposes for movement.

I have argued in Zwart (1993) that this aspect (that obligatory movement must have an explanation) was somewhat lacking in the standard government and binding approach to verb second in Germanic. The pre-government and binding idea that the verb moves to C(OMP) always was maintained in the government and binding period, and salvaged by the adoption of an ad hoc verb second constraint. The alternative view, in which the verb moves to C in inversion constructions only, and to Infl in all other main clauses, was clearly superior in the government and binding framework, as it could dispense with an ad hoc movement of the subject from Spec,IP to Spec,CP (forced by the SVO word order in subject initial main clauses). The latter, more genuine government and binding approach is quite easily translated into the minimalist framework. And so it seems that the opposition against the minimalist approach to Germanic syntax is often more concerned with pre-government and binding notions and analyses than with the virtues of the government and binding framework itself.<sup>24</sup>

In fact, it is clear that work done on Germanic syntax in the government and binding period has provided one of the corner stones for the Minimalist Program. I am referring to Guido Vanden Wyngaerd's work on object shift (scrambling) in Dutch (Vanden Wyngaerd 1989a, 1989b). Vanden Wyngaerd discovered that scrambling in Dutch has all the properties of A-movement (except one: the ability to license parasitic gaps), and proposed that it involves movement to the specifier position of a functional projection for object licensing. To express the parallel with subject raising to AgrSP, he called this functional projection AgrOP. Independently, AgrOP was proposed by Chomsky (1991) in order to accommodate the short verb movement discovered in Pollock (1989).<sup>25</sup>

In hindsight, the discovery of AgrOP can be seen as a break-through. It made it possible to postulate specifier-head agreement as the only relation of syntactic licensing, and to dispense with government for this purpose. It was also instrumental in shaping the idea of a universal basic SVO order. Since objects in Dutch and German move to a position in the functional domain in overt syntax, it is no longer possible to regard the embedded clause word order as underived. This paved the way to a reanalysis of Dutch and German as SVO-languages, while keeping Koster's (1975) result that the main clause word order is derived from the embedded clause word order (Zwart 1992a).

More generally, the discovery of a licensing position for objects in the functional domain has engendered a view on syntax that is radically different from the traditional view. In the traditional view, it is assumed that 'what you see is what you get', that is: underlying structures are assumed to be maximally visible in some overt representation or other. This assumption governed the debate regarding the basic word order of Dutch and German in the early seventies. The question was, whether the main clause word order or the embedded clause word order *was* the underlying word order. Implicit in this

<sup>24</sup> This is also clear from Gaertner and Steinbach (1994).

<sup>25</sup> Chomsky assumed that the Agreement projection discovered by Kayne (1987) in past participle agreement constructions in French is the same as the AgrOP needed to describe short verb movement.

view on syntax was the idea that movement from a basic position was always marked in some sense.

However, it has always been clear that A-movement and head movement never have a marked character. Hence, there are processes that involve movement out of a basic position without adding any markedness. Movement of the object to Spec,AgrOP falls into that category, as Vanden Wyngaerd has shown. The important consequence is that the embedded clause word order in Dutch and German, with a shifted object, is unmarked, yet derived. Hence, neither the main clause word order nor the embedded clause word order *is* the underlying word order.

In addition, this leads to the conclusion that objects have two unmarked positions: the theta-position and the Case-position. Similarly for subjects, adopting the VP-internal subject hypothesis. In a given language, only one of the two unmarked positions may be occupied by the relevant noun phrase. The minimalist approach allows no optionality in this respect. However, the approach also predicts languages to vary in exactly this point: some languages realize the object in the Case-position, others do so in the theta-position.

The natural next step would be to hypothesize that *all* elements have two unmarked positions, one for thematic licensing and one for syntactic licensing. Adopting a restrictive theory of phrase structure, as in Hoekstra (1991) or Kayne (1993), this leads to infinite structures with multiple functional projections and multiple movements into the functional domain. However, there is no added complexity, since the structures and the movements are all identical in each case. Moreover, if the new approach to syntax is correct, the proposed structures are not ad hoc postulates to account for petty facts of word order variation, but are the necessary consequence of an encompassing view.

When the minimalist approach was first laid out in class lectures by Chomsky in the Fall of 1991, many had their doubts about the usefulness of the approach for comparative linguistics. I believe it is only fair to say that there is no reason to suspect that the minimalist approach would fare any worse than the government and binding approach. The minimalist program at least presents a framework in which to describe variation. As before, the simple description of a language particular property (whether in terms of strength of features or in some other terminology) does not suffice to explain syntactic patterns. However, a description in terms of strength of features, coupled with a restrictive theory of phrase structure of e.g. Kayne (1993), must be regarded as a very helpful first step in discovering syntactic generalizations, both language internally and crosslinguistically.

### 3 About This Volume

It is with this potential for the advancement of comparative Germanic syntax in mind that Werner Abraham and I started soliciting working papers for publication in this volume. We are most gratified by the results, and wish to thank the contributors for meeting our deadline so speedily.

Of the 15 papers published here, 6 address Kayne's universal SVO hypothesis and related issues (ABRAHAM, CZEPLUCH, DEN DIKKEN, DEN DIKKEN AND HOEKSTRA, STERNEFELD, AND VANDEN WYNGAERD).

A key aspect of the restrictive phrase structure proposed by Kayne is its binary branching character. Binary branching has long been considered a much needed property of phrase structure, for reasons of simplicity and learnability. However, in a recent *Linguistic Inquiry* article Carrier and Randall have argued that resultatives present evidence for ternary branching structures (Carrier and Randall 1992). DEN DIKKEN AND HOEKSTRA take issue with the arguments presented there.

Binary branching is also one of the issues addressed by CZEPLUCH, who critically discusses the evidence for the structure of the English VP that is assumed traditionally, and in "the new MIT approach".

ABRAHAM and STERNEFELD critically address Kayne's (E)LCA (as I have called it above). It was advanced in Zwart (1993) that verb raising phenomena in particular present evidence for the correctness of the hypothesis that Dutch and German are SVO languages (see also Lattewitz 1994). It is this aspect of the syntax of Dutch and German that ABRAHAM looks into in much detail. He concludes with a rhetorical question, suggesting that the Kayne-Zwart-Lattewitz approach is on the wrong track. STERNEFELD similarly rejects Kayne's hypothesis and the movements and structures that would have to be assumed to make it work for German.

DEN DIKKEN and VANDEN WYNGAERD, on the other hand, demonstrate that adopting Kayne's word order hypothesis makes it possible to shed light on long standing problems of verb raising in Dutch. DEN DIKKEN shows that the SVO hypothesis allows one to explain certain puzzling scope phenomena in Verb Projection Raising constructions. VANDEN WYNGAERD presents an account of the *Infinitivus Pro Participio* effect that shows up in verb clusters in Dutch and German that illuminates the structure of both verb clusters and participles.

Of the remaining 9 papers, 8 present particular implementations of the minimalist program (VAN GELDEREN, HAEGEMAN, HOEKSTRA, POSTMA, ROORYCK, TE VELDE, WILDER, AND ZWART). SOLÀ presents an alternative to the minimalist approach, which is nevertheless very much inspired by it.

In Chomsky (1992), morphological features that are relevant in movement and feature checking are considered as purely abstract entities, comparable to abstract Case in the government and binding framework. Consequently, no prediction is made as to the relation between overt morphology and overt movement. SOLÀ, in contrast, argues that overt morphology explains at which position in the structure inflected forms will be spelled out. He illustrates this with an analysis of verb movement in West Germanic.

VAN GELDEREN marries the minimalist approach with the expletive replacement analysis of Chomsky (1986a). Like Solà, she argues for a connection between overt morphology and overt movement, based on an investigation of the history of English.

HAEGEMAN looks into the syntax of scrambling and clitic placement in West Flemish in great detail. She makes the important observation that the A/A' distinction cannot be reduced to the distinction between L-relatedness and nonL-relatedness of Chomsky and Lasnik (1991).

HOEKSTRA discusses a curious agreement phenomenon in the second person singular in Dutch, and proposes that in some cases, heads may check features by moving to a specifier position. Peculiarities of the verbal paradigm in Dutch are also discussed by POSTMA, particularly in the domain of suppletion and analogy.

ROORYCK and ZWART address wh-movement in the context of the generalized transformation mechanism and economy of derivation. ROORYCK analyses free relatives, and argues that they are syntactically equivalent to indirect wh-clauses. ZWART addresses the contradiction between the shortest move and fewest steps requirements of economy of derivation. He argues that the shortest move requirement is a remnant of an older stage of the theory, redundant within the minimalist approach.

In TE VELDE's paper, the syntax of Yiddish is discussed from a minimalist point of view. It is argued that Yiddish, like German and English, utilizes both C and AgrS as positions for the clausal head. Finally, WILDER presents a minimalist approach to coordination.

In closing, I would like to thank Werner Abraham, who took the initiative for publication of this volume of the *Groninger Arbeiten zur germanistischen Linguistik*, as well as Mr. A. Mollema for editorial assistance.

## References

- Burzio, L. (1981) *Italian Syntax. A Government-Binding Approach*. Reidel, Dordrecht.
- Carrier, J. and J.H. Randall (1992) 'The Argument Structure and Syntactic Structure of Resultatives'. *Linguistic Inquiry* 23, 173-234.
- Chomsky, N. (1957) *Syntactic Structures*. Mouton, The Hague.
- Chomsky, N. (1961) 'On the Notion "Rule of Grammar"'. In J.A. Fodor and J.J. Katz, eds., *The Structure of Language. Readings in the Philosophy of Language*, Prentice Hall, Englewood Cliffs, 1964.
- Chomsky, N. (1965) *Aspects of the Theory of Syntax*. MIT Press, Cambridge, Mass.
- Chomsky, N. (1966) 'The Theory of Transformational Generative Grammar'. In N. Chomsky, *Topics in the Theory of Generative Grammar*, Mouton, The Hague.
- Chomsky, N. (1970) 'Remarks on Nominalization'. In N. Chomsky (1972).
- Chomsky, N. (1972) *Studies on Semantics in Generative Grammar*. Mouton, The Hague.
- Chomsky, N. (1973) 'Conditions on Transformations'. In S.R. Anderson and P. Kiparsky, eds., *A Festschrift for Morris Halle*, Holt, Reinhart, and Winston, New York.
- Chomsky, N. (1977) 'On Wh-Movement'. In P. Culicover, T. Wasow, and A. Akmajian, eds., *Formal Syntax*, Academic Press, New York.
- Chomsky, N. (1980) 'On Binding'. *Linguistic Inquiry* 11, 1-46.
- Chomsky, N. (1981) *Lectures on Government and Binding*. Foris, Dordrecht.
- Chomsky, N. (1986a) *Knowledge of Language. Its Nature, Origin, and Use*. Praeger, New York.
- Chomsky, N. (1986b) *Barriers*. MIT Press, Cambridge Mass.
- Chomsky, N. (1991) 'Some Notes on Economy of Derivation and Representation'. In R. Freidin, ed., *Principles and Parameters in Comparative Grammar*, MIT Press, Cambridge Mass.
- Chomsky, N. (1992) 'A Minimalist Program for Linguistic Theory'. *MIT Occasional Papers in Linguistics* 1.
- Chomsky, N. and H. Lasnik (1977) 'Filters and Control'. *Linguistic Inquiry* 8, 425-504.
- Chomsky, N. and H. Lasnik (1991) 'Principles and Parameters Theory'. In J. Jacobs, A. van Stechow, W. Sternefeld, and T. Vennemann, eds., *Syntax: An International Handbook of Contemporary Research*, Walter de Gruyter, Berlin, 1993.
- Chomsky, N. (1994) "Bare Phrase Structure." *MIT Occasional Papers in Linguistics* 5.
- Fukui, N. and M. Speas (1986) 'Specifiers and Projection.' *MIT Working Papers in Linguistics* 8, 128-172.
- Gärtner, H.-M. and M. Steinbach (1994) "Economy, Verb Second and the SVO-SOV Distinction". To appear in *Working Papers in Scandinavian Syntax*.
- Greenberg, J. (1963) 'Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements'. In J. Greenberg, ed., *Universals of Language*, MIT Press, Cambridge.
- Hoekstra, E. (1991) *Licensing Conditions on Phrase Structure*. Dissertation, University of Groningen. Groningen Dissertations in Linguistics 2.
- Jackendoff, R. (1972) *Semantic Interpretation in Generative Grammar*. MIT Press, Cambridge.
- Jackendoff, R. (1977) *X'-Syntax: A Study of Phrase Structure*. MIT Press, Cambridge.
- Katz J.J. and P.M. Postal (1964) *An Integrated Theory of Linguistic Descriptions*. MIT Press, Cambridge.



- Kayne, R.S. (1984) *Connectedness and Binary Branching*. Foris, Dordrecht.
- Kayne, R.S. (1987) 'Facets of Romance Past Participle Agreement'. In P. Benincà, ed., *Dialect Variation in the Theory of Grammar*, Foris, Dordrecht, 1989.
- Kayne, R.S. (1992) 'Word Order'. Guest lecture at the GLOW Colloquium, Lisbon, April 14.
- Kayne, R.S. (1993) 'The Antisymmetry of Syntax'. Ms., City University of New York.
- Koster, J. (1975) 'Dutch as an SOV Language'. *Linguistic Analysis* 1, 111-136.
- Koster, J. (1983) 'Taalkunde: van common sense naar wetenschap'. *TTT* 3, 198-213.
- Koster, J. (1986) 'The Relation between Pro-drop, Scrambling, and Verb Movements'. Ms., University of Groningen.
- Lattewitz, K. (1994) "A Minimalist View on Verb Raising". Paper presented at the 9th Comparative Germanic Syntax Workshop, Harvard, January 6.
- Neeleman, A. (1994) *Complex Predicates*. Dissertation, University of Utrecht.
- Palmatier, R.A. (1972) *A Glossary for English Transformational Grammar*. Appleton-Century-Crofts, New York.
- Pollock, J.-Y. (1989) 'Verb Movement, Universal Grammar, and the Structure of IP'. *Linguistic Inquiry* 20, 365-424.
- Stuurman, F. (1985) *X-bar and X-plain: A Study of X-bar Theories of the Phrase Structure Component*. Foris, Dordrecht.
- Vanden Wyngaerd, G. (1989a) 'Object Shift as an A-Movement Rule'. *MIT Working Papers in Linguistics* 11, 256-271.
- Vanden Wyngaerd, G. (1989b) 'Raising to Object in Dutch and English'. *Dutch Working Papers in English Language and Linguistics* 14.
- Zwart, C.J.W. (1992a) 'SOV Languages Are Head Initial'. Paper presented at the 8th Workshop on Germanic Comparative Syntax, Tromsø, November 21.
- Zwart, C.J.W. (1992b) 'Matching'. In D. Gilbers and S. Looyenga, eds., *Language and Cognition 2. Yearbook 1992 of the Research Group for Linguistic Theory and Knowledge Representation of the University of Groningen*, 349-361.
- Zwart, C.J.W. (1993) *Dutch Syntax. A Minimalist Approach*. Dissertation, University of Groningen.