

# THE WORD ORDERS OF ENGLISH AND DUTCH\*

## Collective vs. Individual Checking

*Jan Koster*  
University of Groningen

### 1. Introduction

#### 1.1 The problem

In this article, I will show that the vast collection of word order differences between English and Dutch, including the most characteristic and central ones such as the VO/OV distinction, follows from only one small difference in how phrases for feature checking are defined. Practically all differences, in other words, turn out to follow from a simple difference between the two languages with respect to Pied Piping.

Almost since its inception, modern generative grammar, in this respect at least inspired by structuralist ideas of Greenberg (1966) and others, has assumed that the languages of the world can be divided in those with VSO, SVO or SOV as their deep structure order. Occasionally, some other basic orders of Subject, Verb and Object were proposed (like VOS for Malagasy), but gradually a near-consensus developed that Verb and Object (V and O) are adjacent in underlying structure, that Subjects precede Objects and that the basic word orders are therefore SVO and SOV. Among the Germanic languages, English and the Scandinavian languages are considered to represent the SVO type, while Dutch and German are usually analyzed as SOV (Thiersch (1978), Koster (1975)).

Another aspect of the received view was that the distinction between, say, English and Dutch results from a *parameter*, the so-called OV/VO parameter (see for instance Neeleman (1994)).

This view has never been satisfactory, because, corresponding with the OV/VO distinction, there is a substantial number of other facts not at all covered by the simple parameter. For instance, unlike the VO languages, the Germanic OV languages Dutch and German show a great deal of scrambling, which appears to be typical of OV languages in general vis-a-vis VO languages (cf. Corver and Van Riemsdijk (1993)). Up until recently, this correlation between OV and relatively free word order was completely unexplained.

But also the OV/VO parameter itself is hardly more than a stipulation telling us what we already know, namely that there are word order differences between English and German. A real theory would *explain* the correlations in question and even clarify why there is an OV/VO distinction in the first place.

As it stands, not even the simplest differences in word order between closely-related languages like English and Dutch are accounted for. In Dutch, for instance, it is entirely natural to have adverbials between verb and object (1a), while such is completely impossible in English (1b):

---

\* I would like to thank Marcel den Dikken, Pytsje van der Veen and Jan-Wouter Zwart for helpful comments on an earlier version. All remaining errors are mine.

- (1) a. Ik heb haar gisteren gezien  
 I have her yesterday seen  
 "I have her seen yesterday"
- b. \*I have seen yesterday her

Another remarkable difference is that (apart from topicalization) in English only a subclass of the adverbials can appear to the left of the VP (2), while in Dutch, there are no such limitations:

- (2) a. He probably read a book  
 b. \*He yesterday read a book
- (3) a. Hij heeft waarschijnlijk een boek gelezen  
 he has probably a book read  
 "He probably read a book"
- b. Hij heeft gisteren een boek gelezen  
 he has yesterday a book read  
 "He yesterday read a book"

Even such simple facts of the word order of closely-related and much studied languages have never been satisfactorily explained. I will give a more systematic overview of such word order differences in section 2.

### 1.2 Attempts to derive English VO from Dutch-like OV

Note that if the complement-adjacency condition of Chomsky (1986b) is correct --a reasonable assumption--- the Dutch order NP-Adv-V of (1) can only be a derived order.<sup>1</sup> Since English seems to obey complement adjacency (as shown by (1b)), the English SVO order is more likely to reflect the base word order than the Dutch and German SOV order (which not necessarily obeys complement adjacency).<sup>2</sup> As a matter of fact, I will assume that complement adjacency is a very important clue as to underlying order.

On the other hand, the English word order is at first sight a little "suspect" in that the indirect object is closer to the verb than the direct object. That the direct object is more closely related to the verb than the indirect object is most clearly brought out by idioms:

- (4) a. Mary gave him *the sack*  
 b. \*Mary gave *the sack* him

<sup>1</sup> Throughout this article, I will for conventional reasons refer to NPs where I really mean DPs.

<sup>2</sup> In English, adjacency can be violated by Heavy NP Shift. Furthermore, adjacency is strict for (not too heavy) NPs and APs, but less so for PPs and CPs. The latter are exactly the categories which undergo extraposition in Dutch. In Koster (forthcoming), I reanalyze extrapositions as parallel construals, with properties akin to coordination. See Rijkhoek (1998) for some discussion. Deviations from complement adjacency in English can be analyzed in a similar way. See Pesetsky (1994) for examples.

In early generative grammar, it was often assumed (contrary to fact, as it appeared later) that parts of idioms are adjacent in underlying structure, as in Dutch for examples similar to (4a):

- (5) Marie heeft hem *de bons* gegeven  
Mary has him the sack given  
"Mary gave him the sack"

There are some other similarities between Dutch and English word order as well and, based on the wish to derive the word orders of English and Dutch from a common source, it was proposed sometimes to assume an underlying OV order for both English and Dutch and to derive the English VO order by verb movement to the left (Koster (1988) and for related verb movements for different reasons, see Larson (1988)).

The attempt to derive English VO word order from an underlying OV source by verb movement meets insurmountable problems. First of all, verb movement does not do what it set out to do, i.e., to provide a solution for the lack of adjacency of certain elements in English surface structure. Particles are among the elements most tightly connected to verbs and they almost always form an idiomatic combination with the verb. In Dutch, this is reflected by adjacency of particle and verb:

- (6) a. Jan heeft Peter zijn bal *af gepakt*  
John has Peter his ball away taken  
"John picked Peter's ball"  
b. \*Jan heeft Peter *af* zijn bal *gepakt*

With double objects, the particle always follows *both objects* in Dutch. This is not true at all for English, where (with some exceptions, like the particle *back*) the particle typically follows *the first object* (see Emonds (1976, 81-86) for examples):

- (7) a. The secretary *sent* the stockholders *out* a schedule  
b. \*The secretary *sent* the stockholders a schedule *out*  
c. ?The secretary *sent out* the stockholders a schedule

So, if the particle position is a cue about the underlying verb position, this would lead to a base position for the verb *preceding* the object (*a schedule* in (7a)). To the extent that it is not possible to see particle distribution in English as a clue to underlying verb positions (like in Dutch, cf. Koster (1975)), the argument for English OV based on underlying idiom adjacency is undermined.

But there are other serious problems with deriving English VO from underlying OV by verb movement. It is, for instance, far from clear where the verb would have to move to. If the verb moves to the left within the VP, there is no obvious and theoretically permitted landing site available. There is, for instance no obvious other head position for the V to move to.

Moving the V to a position outside of the VP (for instance the T(ense) position) leads to other problems, because the VP is preposed without tense-bearing auxiliaries (cf. Emonds (1976, 115)):

- (8) They said we shouldn't buy gold, but *buy gold* we will

Most important of all, the verb movement hypothesis runs into trouble with respect to the facts of adverb distribution. In Dutch, adverbials like *yesterday* can appear anywhere on the left of the verb:

- (9)
- a. *Gisteren* heeft hij Marie het boek gegeven  
yesterday has he Mary the book given  
"Yesterday, he gave Mary a book"
  - b. Hij heeft *gisteren* Marie het boek gegeven
  - c. Hij heeft Marie *gisteren* het boek gegeven
  - d. Hij heeft Marie het boek *gisteren* gegeven
  - e. Hij heeft Marie het boek gegeven, *gisteren*

English only has the peripheral orders for *yesterday* ((9a) and (9e)) and the internal orders (particularly (9c) and (9d)) form an obstacle for any attempt to derive English VO order from Dutch-like OV with verb movement.

The problem posed by the internal orders (9c) and (9d) is that there is VP-internal material (*Marie* and *het boek*) to the left of non-VP material, namely the adverbial *gisteren* ("yesterday"). In the past, a common way to derive (9c) and (9d) was by scrambling the NPs to the left (to positions outside of the VP). Since Dutch also has a kind of scrambling which yields marked focusing and since the intonation patterns found in (9c) and (9d) are --in contrast-- entirely neutral, it was also proposed sometimes to relax the complement adjacency condition for OV languages (but not for VO languages) (see Neeleman (1994, ch. 3).

The latter solution is arbitrary and *ad hoc* and would wrongly predict a similar relaxation of verb complement adjacency for English after verb movement (applied to the Dutch O-Adv-V). This prediction is not borne out:

- (10) \*He has [<sub>VP</sub> given<sub>i</sub> Mary *yesterday* the book *t<sub>i</sub>* ]

The other solution (NPs scrambled out of the VP) does not work either. This solution would crucially involve the adverbial *yesterday* to the left of the VP, which is entirely impossible in English:

- (11) \*He has *yesterday* [<sub>VP</sub> given Mary the book --]

Nor can the NPs be scrambled out of the VP in English:

- (12)
- a. \*He has *Mary* [<sub>VP</sub> given -- the book]
  - b. \*He has *Mary the book* [<sub>VP</sub> given -- --]

Given these word order facts, I cannot make sense of the hypothesis according to which English VO order is derived from a Dutch-like OV order under verb movement. All in all, then, a different approach seems in order.

### 1.3 A new approach

#### 1.3.1 Theoretical developments

In fact, the elements for a new approach have been available for quite some time now. In retrospect, the earlier attempts to derive the word orders of English and Dutch from a common source were unsuccessful due to the fact that linguistic theory was insufficiently developed during the 1980s. As I will propose, thanks to theoretical developments during the 1990s, it is now possible to formulate a theory in which all major word order differences between English and Dutch --a massive collection of facts, impossible to cover by a simple OV/VO-parameter-- can be traced to the effects of not more than a single parameter.

The theoretical developments which have made this theory possible are Chomsky's Minimalism (Chomsky (1995)) and Kayne's Antisymmetry Theory (Kayne (1994)). Both theories contributed to the overcoming of certain forms of theoretical stagnation as inherited from the earlier EST-framework.<sup>3</sup> Both theories contain elements which make it possible to give a first more comprehensive (but of course tentative) explanation of the word order facts (and differences) of English and Dutch.

Minimalism has many aspects which I will not consider here. For my present purpose, the most important aspect of Minimalism (and the theoretical developments preceding it) is that it broke the spell of semantics as found in earlier forms of generative grammar, particularly during the long reign of EST-style theories. According to such theories the main function of large classes of transformations (as instances of the transformational residue "move alpha") is to feed semantics, i.e., "movement" connected the "thematic" d-structure with the "scopal" s-structure, so that both could contribute their elements to semantic interpretation. The attempt to develop a separate level of Logical Form sprang from the same philosophy (see Koster (1993) for a critique and the role of the idea of structure-preservingness).

It is not generally recognized yet, but current developments inspired by Minimalism have made it possible to break with this idea of the transformational feeding of semantic interpretation by emphasizing an entirely different function of movement transformations ("move alpha"), namely feature checking. Since Chomsky (1986a) and Pollock's work during the same period (Pollock (1989)) there has been a trend towards extending not only the supply of available functional elements but also giving them full projections. There is much debate about the exact extension of this class of functional projections, but no matter how this debate will end, there is little question that the extension of the class of functional projections as such has given the field a desirable creative impulse during the 1990s.

It seems to me that with these new functional projections, we can perhaps say something new about the question why natural languages have movement transformations ("move alpha") in the first place: *the function of movement, it seems, is not to feed semantic interpretation, but to make functional structure visible in an economical way*. This, of course, does not answer why languages have functional structure at all, but given the fact that functional structure exists, movement makes much sense because it often makes two kinds of structure lexically visible with only one set of lexical items.

According to our current theories, a typical projection consists of a Spec, a head and a complement. The complement positions of functional projections are usually occupied, but the Spec and head positions are generally available as landing sites for "movement". Occasionally, languages have special elements to make functional structure visible (like

---

<sup>3</sup> See for instance Chomsky (1977).

English *do*-support), but more often than not functional structure is made visible indirectly by "moving" lexical material from the lexical projections into the available landing sites. In short, thanks to the new perspective created by Minimalism and other recent developments, "move alpha" can be seen as a kind of parasitic lexicalization of functional structure.

A second potential important development of the 1990s has been Richard Kayne's idea (following from his Antisymmetry Theory) that functional projections are only found on the left of lexical projections. This idea has led to a very substantial rethinking of existing analyses and it has revolutionized in my opinion the study of many languages, including the study of Dutch and German (see Zwart (1997) for the most complete account).

### 1.3.2 *New developments in the study of German and Dutch*

The new theoretical perspectives make it possible, as I will show, to derive the word orders of English and Dutch from a common source in a surprisingly simple way. As it stands, the new theory of Universal Grammar not only makes it possible, but in fact requires that not the OV word order of Dutch, but the VO order of English reflects the deepest underlying order. An immediate indication of that is that English preserves verb-complement adjacency (1b), while Dutch does not (1a) (most of the time I will argue on the basis of Dutch, but the facts of German are very similar).

Deriving Dutch OV-order from English-like VO-order can be done by movement of the object to a functional position to the left of the VP, as was first proposed by Vanden Wyngaerd (1989). Vanden Wyngaerd proposed a rule ("object shift") moving the object to the Spec of AgrOP, but the exact label the functional projection is not crucial, as long as we assume that the object is moved to a functional projection to the left of the VP.

Particularly inspired by Jan-Wouter Zwart's work ((1993), (1994) and (1997)), Kaan (1992), Koster (1994) and Den Dikken (1996) and others worked out the obvious consequence of Vanden Wyngaerd's idea, namely that there is no longer a good reason to see Dutch as OV at its deepest level. The arguments for "Dutch as an SOV language" (Koster (1975)) are still valid, but are now seen as being about a derived level, namely the level resulting from a rule like Vanden Wyngaerd's object shift. As before, the SVO order of Dutch main clauses is seen as derived from the (now intermediate) SOV order by Verb Second.

A weak point of the theory developed so far is that the parameters distinguishing English from Dutch and German still seem arbitrary. What it comes down to is that sometimes a Spec is overtly filled (like the Spec of AgrOP in Dutch) and sometimes "at LF" (as in English), a distinction made in terms of "strong" and "weak" feature checking in Chomsky (1995, ch. 3). This definitely seemed an advance over the old, very crude OV/VO parameter, because at least the new formulation made the parameter distinguishing English from Dutch very similar to the independently needed parameter distinguishing languages with Wh-movement (like English) from languages with Wh-*in situ* (like Japanese): in both cases the difference boils down to overt filling of a Spec (Wh-movement, Object shift) or covert filling "at LF" (Wh-*in situ*, the "object-*in situ*" of VO languages like English).

In spite of this potential for a greater uniformity of the theory of parameters, the strong-weak distinction has remained entirely arbitrary (in spite of unsuccessful attempts to formulate the distinction morphologically).

In what follows, I will assume a version of Minimalism and Antisymmetry Theory and build further on the work of Vanden Wyngaerd, Kaan, Zwart and others. I will not only show that the (huge) word order differences between English and Dutch follow from one simple

parameter, I will also suggest that a new theory of parametrization is in order. Parametrization is mainly about how the functional structure on the left of lexical projections is made visible ("lexicalized"). I will suggest that with respect to the word orders of English and Dutch the job is done by a parameter based on a generally agreed upon dimension of language variation, namely the size of percolation domains (language-specific differences in Pied Piping structure). I will, in other words, suggest that as for word order, English and Dutch only differ with respect to the size of the checking phrase, where Dutch checks its VP-elements individually, while English checks the same functional heads by percolation of the VP-internal features to a higher projection (Pied Piping).

As I will show, my proposal makes the earlier weak-strong distinction superfluous. If the proposal is correct, it will not only explain the massive word order differences between Dutch and English, it will also explain *why* English is VO and Dutch OV.

A further advantage of the theory to be proposed is that it makes it possible to return to the old idea of a universal base order. In harmony with Kayne (1994), the version of the theory I will propose *requires* such a universal base order, namely VO. The fact that both English and Dutch word order can be derived from such a universal base in a revealing way, strengthens the case of the general theory.

Last but not least, I hope that the theory presented here will cast some light on the old issue of configurational versus non-configurational languages in general. Some of this distinction at least can be found in the languages at issue: English has a very rigid word order, while Dutch and German have so much scrambling that they can be considered free-word-order languages to some degree.<sup>4</sup> So far, the fact that such major distinctions exist among languages (even among very closely related ones) has hardly been explained. From the parameter to be proposed below, it follows that Dutch VP-elements can be distributed over the universal functional structure to the left of the VP, while for English V-complements the VP works as a "cage", thereby rigidly freezing English word order into immobility.

## 2. The facts

### 2.1 Similarities and differences

Although English, German and Dutch are closely related Germanic languages, with some obvious similarities and overlap in structure, the word order differences are bewildering at first sight. Before going into my theoretical background assumptions and before presenting the one parameter from which all differences seem to follow, I will give an overview of the facts to be explained. The main facts, some of them with wide-ranging ramifications, can be summarized in five classes:

- (13)
- 2.2 English is VO, Dutch and German OV
  - 2.3 Neither English, Dutch nor German has rightward scrambling
  - 2.4 Unlike Dutch and German, English has no leftward scrambling
  - 2.5 In Dutch and German *all* AdvS occur to left of VP, in English only a subclass
  - 2.6 English Adv order shows paradoxes of scope

---

<sup>4</sup> Actually, the situation is more complex. English shows somewhat more word order variation VP-internally, while Dutch and German show VP-external word order variation, thanks to scrambling.

I will now briefly discuss these classes of facts one by one.

## 2.2 English is VO, Dutch and German OV

It has been a standard assumption since the early 1970s that English is an SVO language, while Dutch and German are underlyingly SOV (see Thiersch (1978) and Koster (1975)). The difference shows up immediately in subordinate clauses (which including sentences with auxiliaries like (14) in Dutch):

- (14) a. John has read *the book*  
b. Jan heeft *het boek* gelezen  
John has the book read

These word orders are quite rigid, as we cannot have the Dutch order in English (15a) or the English order in Dutch (15b):

- (15) a. \*John has *the book* read  
b. \*Jan heeft gelezen *het boek*

It has been recognized for a long time that there are some exceptional conditions under which the otherwise rigid OV (or SV) pattern can be broken (see, for instance, Neeleman (1994, 85)):

- (16) a. Toen heb ik ontmoet *de Here Jezus*  
then have I met the Lord Jesus  
b. Heden is overleden *onze geliefde grootmoeder*  
today has died our beloved grandmother

In these cases, the postverbal NPs have stress (cf. so-called equatives in Ross (1969)). According to Neeleman (*loc.cit.*), these examples involve extraposition. The object in (16a) is definitely not in the English object position, because with adverbial material following the verb, it is preferably *not* adjacent to the verb:

- (17) a. Gisteren heb ik *ontmoet* tijdens een wandeling *de Here Jezus*  
Yesterday have I met while walking the Lord Jesus  
b. ?\*Gisteren heb ik *ontmoet de Here Jezus* tijdens een wandeling

I reject extraposition analyses for such constructions. Following Kayne (1994), I will assume that Universal Grammar has a general ban on rightward movements. As I have argued elsewhere, most so-called rightward movements do not at all have the properties of "move alpha" and are better analyzed as specifying parallel construals along the lines of coordination. In this article, I will not further discuss extraposition phenomena and refer the reader to the relevant literature.<sup>5</sup>

---

<sup>5</sup> Koster ((1995) and (forthcoming)). See also Rijkhoek (1998) and De Vries (1999).



### 2.3 Neither English, Dutch nor German has rightward scrambling

If scrambling (seen as free adjunction of XPs) would exist, it would come as a big surprise that neither English nor Dutch or German shows any evidence at all for free rightward movement. Apart from a few extraposition phenomena (discussed elsewhere, see note 5), free rightward movement is simply not possible in English:

*English:*

- (18) a. John read *the book* yesterday  
b. \*John read  $t_i$  yesterday *the book*<sub>i</sub>
- (19) a. John gave *Bill* a book  
b. \*John gave  $t_i$  a book *Bill*<sub>i</sub>
- (20) a. Mary made *Sue* happy  
b. \*Mary made  $t_i$  happy *Sue*<sub>i</sub>
- (21) a. Mary was *happy* yesterday  
b. \*Mary was  $t_i$  yesterday *happy*<sub>i</sub>

Stowell (1981) sought to explain the ungrammaticality of *John saw yesterday Bill* in terms of a Case-adjacency condition, but such an approach seems pointless under a theory with free rightward adjunction: in (18b) the relevant Case could easily be transmitted via the trace, just as in leftward movements like Wh-movement (cf. *Whom<sub>i</sub> did you see t<sub>i</sub>*). Moreover, the adjacency phenomenon does not seem to be restricted to Case-bearing NPs, witness the illicit AP-movement in (21b). Apart from extraposition phenomena (discussed elsewhere, see note 5), XPs cannot be moved to the right at all in English. This simple fact makes at once problematic a theory that permits adjunction of XP in the first place.

The facts are exactly the same in Dutch and German. I will limit myself to Dutch:

*Dutch:*

- (22) a. Jan heeft *het boek* gelezen  
John has the book read  
"John has read the book"
- b. \*Jan heeft gelezen *het boek*  
John has read the book
- (23) a. Jan heeft *Mary* een boek gegeven  
John has Mary a book given  
"John has given Mary a book"
- b. \*Jan heeft een boek gegeven *Mary*  
John has a book given Mary

- (24) a. Mary heeft Suus *gelukkig* gemaakt  
 Mary has Sue happy made  
 "Mary has made Sue happy"
- b. \*Mary heeft Suus gemaakt *gelukkig*  
 Mary has Sue made happy
- (25) a. Mary is *gelukkig* geweest  
 Mary is happy been  
 "Mary has been happy"
- b. \*Mary is geweest *gelukkig*  
 Mary is been happy

In other words, on the basis of English and Dutch (and German) alone, there is not the slightest reason to weaken linguistic theory by introducing the very powerful device of free adjunction to the right.

#### 2.4 Unlike Dutch and German, English has no leftward scrambling ("the cage problem")

So far, English and Dutch have appeared to be remarkably similar as to their scrambling possibilities. If we focus on leftward scrambling, however, the parallels all of a sudden break down completely: Dutch and German have ample leftward scrambling while English has none whatsoever. In Dutch, for instance, objects can naturally be moved across adverbials on their left:

- (26) Jan heeft *het boek* gisteren *gelezen*  
 John has the book yesterday read  
 "John read the book yesterday"

This is a very natural word order in Dutch and the lack of verb-complement adjacency (the hallmark of deep structure order) has been taken as evidence for scrambling for a very long time. Originally, leftward scrambling was interpreted as free adjunction to the left and since Vanden Wyngaerd (1989) it has often been seen as obligatory movement to the Spec of AgrOP or the Spec of some other functional projection on the left of VP. This latter interpretation has made it possible to analyze Dutch as underlyingly SVO, which has brought the ideal of a universal base order a little bit closer.

APs are more difficult to scramble than NPs in Dutch, but acceptable examples are not hard to find:

- (27) Jan is *erg gelukkig* in zijn jeugd geweest  
 John has very happy in his childhood been  
 "John has been very happy in his childhood"

All of this is impossible in English. Neither NPs (28) nor APs (29) can be scrambled to the left:

(28) \*John *the book*<sub>i</sub> read *t*<sub>i</sub> yesterday

(29) \* John *very happy*<sub>i</sub> has been *t*<sub>i</sub> during his childhood

If leftward scrambling in Dutch and German is just free leftward adjunction to the VP, it is entirely mysterious why this process, which produces such natural word orders in Dutch and German, is completely impossible in English. One could of course stipulate that complement-verb adjacency is relaxed in OV languages (as in Neeleman (1994, 74)) or that the V-position on the left periphery of the VP blocks scrambling in English somehow, but that would be nothing more than repeating the problematic facts that we want to explain in the first place.

In fact, the English VP behaves like a "cage" for its complement XPs: they can be moved neither to the right (as we saw in the previous section) nor to the left (apart from fronting rules, like Wh-movement).<sup>6</sup> This cage character of the English VP gives English its rigid word order, distinguishing it from Dutch and German and the many other languages of the world with relatively free word order. As it stands, it is highly unsatisfactory that linguistic theory cannot explain why some languages (like English) have very rigid word order, while others (even closely related languages, like Dutch and German) are very permissive as to their word order (but see note 4).

In sum, the least we expect of an explanatory theory of English word order is that it explains the cage character of the English VP.

Note that, although VP-internally English shows some more possibilities than Dutch and German, some apparent rigidity exists within the English VP as well. In German, and even in Dutch in the appropriate contexts, the unmarked order of indirect and direct object can be reversed:

- (30) a. Jan heeft ons *het geld* teruggegeven  
John has us the money back given  
"John has given us the money back"  
b. Jan heeft *het geld* ons teruggegeven  
John has the money us back given

Sentence (30a) has the standard word order IO-DO, but the DO-IO order of (30b) also is possible to many speakers of Dutch (particularly because of the focus bearing particle *terug* ("back"); see further Zwart (1993, 49)). According to more or less traditional analyses, this reversal of order could be interpreted as VP-internal scrambling. This is again impossible in English:<sup>7</sup>

- (31) a. John has given us *the money* back  
b. \*John has given *the money* us back

Actually, the reversal of IO and DO involves scrambling to VP-external positions, which is possible in Dutch and German, but not in English, as I will show. All in all, we see that none

---

<sup>6</sup> I will assume that the English VP serves as a "cage" only for the basic licensing of VP-specific elements. This excludes scrambling, but not Wh-movement or external licensing of the VP-internal subject. Subjects are standard elements of clauses (cf. The EPP) and, unlike objects, not specifically selected by verbs.

<sup>7</sup> Note, however, that English VP-internal word order is also more flexible in a number of respects. In English, the VP and its orders are left intact, while Dutch and German undergo scrambling, which involves fixed positions.

of the Germanic languages under consideration has rightward scrambling (as demonstrated in section 2.3), while English sharply deviates from Dutch and German in that it is also closed to the left.

### 2.5 In Dutch and German all Adv<sub>s</sub> appear to left of VP, in English only a subclass

One of the biggest obstacles in earlier attempts to derive the word orders of English and Dutch from one underlying source was the curious difference in the distribution of adverbial expressions in both languages. In Dutch practically all adverbials can appear to the left of the VP, while English has sharp restrictions in this respect. The English adverbial *probably*, for instance, can appear to the left of the VP in a position that is excluded for numerous other adverbial expressions, such as *yesterday*, *everywhere*, *hard*, etc. (see Jackendoff (1972, ch. 3 for more examples):

- (32) *English:*
- a. He probably [<sub>VP</sub> saw Bill]
  - b. \*He yesterday [<sub>VP</sub> saw Bill]
  - c. \*He everywhere [<sub>VP</sub> saw Mary]
  - d. \*He very hard [<sub>VP</sub> worked]

The adverbs forbidden on the left can all appear on the right:

- (33)
- a. He saw Bill yesterday
  - b. He saw Mary everywhere
  - c. He worked very hard

If adverbs are just freely adjoined to VPs, as has often been assumed, this asymmetry is totally unexpected. Why would it be possible to be adjoined to the right of the VP for *probably* (32a) but not for *yesterday*? If there were a semantic reason for this distinction, adjoining *yesterday* to the right of the VP (as assumed for (33a)) would be just as impossible.

The facts become even more mysterious if we realize that *all* adverbials excluded from the left periphery of VPs in English are possible in Dutch in this very position:<sup>8</sup>

- (34) *Dutch:*
- a. Hij heeft waarschijnlijk [<sub>VP</sub> Wim gezien]  
 he has probably Bill seen  
 "He has probably seen Bill"
  - b. Hij heeft gisteren [<sub>VP</sub> Wim gezien]  
 he has yesterday Bill seen  
 "He has seen Bill yesterday"

---

<sup>8</sup> The labelled bracketings in (34) actually involve AgrOP (or AccP) instead of VP. The VP labels are used here for ease of exposition.

- c. Hij heeft overal [VP Mary gezien]  
he has everywhere Mary seen  
"He has seen mary everywhere"
- d. Hij heeft erg hard [VP gewerkt]  
he has very hard worked  
"He has worked very hard"

In Dutch, in contrast, *waarschijnlijk* and *gisteren* are only possible to the right of the VP with a comma intonation, while *overal* and *hard* are excluded altogether in this position:

- (35) a. \*Hij heeft [VP Mary gezien] overal  
he has Mary seen everywhere
- b. \*Hij heeft [VP gewerkt] erg hard  
he has worked very hard

Within the traditional adjunction analyses of adverbials, these facts cannot be excluded by a general ban on rightward adjunctions at first sight, because adverbial PPs appear freely on the right of the VP:

- (36) Hij heeft [VP Mary gezien] in elke stad  
he has Mary seen in each town  
"He saw Mary in each town"

From a comparative syntax perspective these facts are most intriguing and totally unaccounted for so far. Any adequate theory of word order must explain why only a subclass of available adverbials can appear to the left of the VP in English, while in a closely related language like Dutch *all* adverbials can appear to the left of the VP.

## 2.6 English Adv order shows paradoxes of scope

It has been realized almost since the beginning of generative grammar that there is a relationship between "command" and scope (see Klima (1964)). Usually an element *A* which has scope over *B* also c-commands *B*. Interestingly, there also is a relation with precedence: very often the element with the wider scope, *A*, also precedes *B*. However, there are some apparent discrepancies (a subclass of which will be discussed in a minute) from which it has often been concluded that linear order is irrelevant and that the hierarchical ordering (as expressed by c-command) is all important.<sup>9</sup>

The Antisymmetry Theory proposed by Kayne (1994) in principle restores the correlation between hierarchical ordering and linear ordering. Under this theory we expect that if *A* has scope over *B*, *A* also precedes *B*.

---

<sup>9</sup> In fact, the various discrepancies between "precedence" and scope order were used as major evidence in favor of a level of LF, derived with a rule of Quantifier Raising (cf. May (1977)). In my opinion, most discrepancies can be accounted for in some other way (see Koster 1987, ch. 2) and the evidence below gives further support to the idea that "precedence" by and large corresponds with scopal order.

If we look at the actual data of English as conventionally analyzed, we run into a number of problems: sometimes scope does seem to correspond with linear order and sometimes it does not. A preposed Wh-phrase (like *which man* in (37)) both c-commands and precedes the clause over which it has scope:

(37) [Which man [did you see]]?

Similarly, an adverbial like *probably* both precedes and c-commands the VP over which it has scope (*probably* has wider scope than *very hard*):

(38) He [*probably* [worked *very hard*]]

However, the correspondence between scope dependence on c-command and precedence breaks down when *probably* follows the sentence (possible according to Jackendoff (1972, 50) "if separated from the rest of the sentence by a pause and accompanied with a drop in pitch"):

(39) He worked very hard, *probably*

*Probably* has scope over the VP as before, but it follows rather than precedes the material in its scope. If a theory like Kayne's is correct, (39) cannot reflect the underlying order but instead some reordering ("movement") must have taken place. What I hope to clarify in this article is the nature of this reordering.

As a matter of terminology, I will make a distinction between linear scope and anti-linear scope. Scope is *linear* if "A has scope over B" means both that A c-commands B and that A precedes B (as in (37) and (38)). Scope is said to be *anti-linear* if "A has scope over B" means that A c-commands B but does not precede it (as claimed about (39)). A language is said to show *mixed scope* if it has both linear and anti-linear scope. What I will claim in this article is that anti-linear and mixed scope are only apparent and do not happen to exist if the phenomena in question are properly analyzed. In other words, I assume that the following universally holds (as a corollary of Kayne's Antisymmetry Theory):

(40) **Scope generalization**

All scope in natural language is linear

*A priori* this entails a more restricted linguistic theory than a theory that allows mixed, i.e., both linear and anti-linear scope.

We have already seen in the case of (37-39) that conventional analyses of English are based on the less than optimal theory that mixed scope is possible in natural language. This is particularly true for all theories that analyze adverbial extensions of the VP as free adjunction to the right, as in:

(41) She [[played a sonata] today and yesterday]

This sentence can mean that she played two sonatas, which shows that *today and yesterday* has wider scope than *a sonata*. However, as indicated by the conventional bracketing of (41),

this analysis would entail that English has --in part-- anti-linear scope, which is excluded by (40). This means that in all likelihood the conventional analysis of (41) is wrong.

Also the facts discussed in 2.3 cast serious doubt on the correctness of the anti-linear analysis of (41). In 2.3, we saw that in all other cases free adjunction on the right of the VP is impossible in English. Why should we make an exception for adverbials? Linguistic theory is considerably strengthened if we not only exclude all anti-linear scope but rightward adjunction in general.

Interestingly, scope on the right of the VP is not straightforwardly anti-linear anyway. Paradoxically, we also find linear scope:

- (42) a. She played two days only once  
 b. She played only once two days

Both sentences have an interpretation in which the first adjunct has wider scope than the second. Thus, in (42a) *two days* can have scope over *only once* and in (42b) scope can be reversed, so that in both cases scope corresponds with linear order. However, the two adverbials have both scope over the preceding VP, so that --paradoxically-- (42) show both linear and anti-linear scope. There is no way to express this state of affairs in a traditional theory based on free rightward adjunction of adverbials:

- (43) She [[[played] two days ] only once]

The bracketing would correctly express --in accordance with the c-command theory of scope-- that the adjuncts have wider scope than the VP, but the mutual scope of the two adjuncts is just the opposite from what the bracketing (based on rightward adjunction) suggests. This paradox cannot be resolved under the standard rightward adjunction theory, suggesting that it is seriously in error.

It is rewarding to consider the adverbial scope facts of Dutch for a moment. In Dutch, recall, all adverbials can occur to the left of the VP, or at least to the left of the final V in subordinate structures. Interestingly, adverbial scope is always linear on the left of the V:

- (44) Hij heeft [gisteren [hard [gewerkt]]]  
 he has yesterday hard worked  
 "He worked hard yesterday"

The scope structure for this example is linear and therefore entirely regular from the point of view of the most desirable theory of universal grammar: the mutual scope of the adjuncts and the VP shows complete harmony of wide scope, c-command and precedence. The scope paradox observed for the examples in (42) would never arise in the Dutch orders to the left of the V. The equivalents of (42) are completely regular:

- (45) a. Zij heeft [twee dagen [één keer [gespeeld]]]  
 she has two days one time played  
 "She played two days once"  
 b. Zij heeft [één keer [twee dagen [gespeeld]]]  
 she has one time two days played  
 "She played once two days"

In both cases, scope is entirely linear, suggesting that the Dutch examples (in (45)) correspond to the underlying order, while the English examples in (42) reflect a derived order somehow.

A fact of further interest is that Dutch subordinate clauses also show the anti-linear scope order of adjuncts, *but only on the right of the V*. As observed in Koster (1974), the Dutch verb serves as a "mirror center" in this respect (see also Barbiers (1995)):<sup>10</sup>

- (46) a. Hij heeft tijdens de pauze aan zijn vader *gedacht*  
 he has during the break of his father thought  
 "He thought of his father during the break"
- b. \*Hij heeft aan zijn vader tijdens de pauze *gedacht*  
 he has of his father during the break thought
- (47) a. Hij heeft *gedacht* aan zijn vader tijdens de pauze  
 he has thought of his father during the break  
 "He thought of his father during the break"
- b. \*Hij heeft *gedacht* tijdens de pauze aan zijn vader

It is also possible in Dutch to have one PP on each side of the verb:

- (48) a. Hij heeft aan zijn vader *gedacht* tijdens de pauze  
 b. Hij heeft tijdens de pauze *gedacht* aan zijn vader

If you move the V out of this context, both PP orders are derived in Dutch, which was the strongest argument for verb movement ("Verb Second") in Dutch (see Koster (1975)):

- (49) a. Hij dacht aan zijn vader tijdens de pauze  
 he thought of his father during the break
- b. Hij dacht tijdens de pauze aan zijn vader  
 he thought during the break of his father

In current terminology, this means that Dutch main clauses can show mixed scope: anti-linear in (49a) and linear in (49b). This is in sharp contrast with English, in which only the anti-linear order is possible:

- (50) a. He thought of his father during the break  
 b. \*He thought during the break of his father

All in all, as for adverbial scope, the traditional analyses of English make it a language with both linear and anti-linear scope. Mixed scope is undesirable and implausible from a theoretical point of view. The Dutch data suggest that mixed scope patterns are only apparent and that they can be derived from underlying linear scope patterns by two operations: mirror

---

<sup>10</sup> Sentence (46b) is not really ungrammatical, thanks to scrambling possibilities on the left of the VP. The star in (46b) only indicates highly marked word order.



imaging around V and movement of V, or rather, as I will show for English, a constituent containing V.<sup>11</sup>

For an example like (41) (repeated here as (51a), this would mean that it must be re-analyzed as (51b):

- (51) a. She [[played a sonata] today and yesterday]  
b. She [[played a sonata]<sub>i</sub> [today and yesterday [  $t_i$  ] ] ]

This is exactly what I will propose. In the underlying structure, the VP is both c-commanded and preceded by the adjunct *today and yesterday*, so that scope appears to be entirely linear in the real structure, just as is desirable from the point of view of a restricted theory of Universal Grammar. By such a theory I mean a theory with no rightward adjunction whatsoever.

### 3. The theory

#### 3.1 *Introductory remarks*

Before I can show that the remarkable word order differences between English and Dutch can be traced to the different setting of one single parameter, I have to make explicit my theoretical assumptions. By and large, I am assuming a combination of a version of Chomsky's Minimalism and Kayne's Antisymmetry Theory. I also deviate from these standard theories in a number of respects, as will be indicated below. I will first discuss my assumptions about Universal Grammar (3.2), followed by a discussion of assumptions specific to English and Dutch.

#### 3.2 *Universal Grammar*

My assumptions about Universal Grammar can be summarized as follows:

- (52) 3.2.1 The Configurational Matrix (includes Kayne's Antisymmetry)  
3.2.2 Minimalism: movement makes heads visible (feature checking)  
3.2.3 Lexical nuclei are embedded in a universal shell of functional projections  
3.2.4 There is no adjunction (apart from processes involving heads)  
3.2.5 Parametrization: lexicalization and the "size" of checking phrases

In what follows, I will discuss these assumptions one by one.

##### 3.2.1 *The Configurational Matrix (including Kayne's Antisymmetry)*

In Koster (1987), an attempt was made to give a general characterization of all (local) grammatical relations. According to the hypothesis in question, grammatical relations consist

---

<sup>11</sup> The scope paradoxes of English formed part of my original motivation to derive English surface order from the underlying "Dutch" OV order (Koster (1988)). In retrospect, the intuition was correct, but the execution wrong (V-movement instead of VP-movement, as I will show). See also Pesetsky (1994) for a different attempt.

of exactly one dependent element  $\delta$ , which "borrows" its properties from exactly one antecedent  $\alpha$  in a local domain  $\beta$ . It assumes that the essence of grammar is a property sharing scheme, sometimes referred to as the Configurational Matrix:

(53) **The Configurational Matrix**

All (local) grammatical relations have the following form:

$$[\beta \dots \alpha \dots \delta \dots ]$$

where:

- a.  $\alpha$  is an antecedent
- b.  $\delta$  is an element depending on  $\alpha$
- c.  $\beta$  is some minimal domain
- d.  $\alpha$  and  $\delta$  freely share their properties

I will briefly discuss some further properties of this scheme in a moment. The Configurational Matrix is a hypothesis about grammatical representations and should *not* be seen as a commitment to a representational view of grammar instead of a derivational view. It only is about the properties of representations, not about the way they are generated. Also a derivational theory of grammar eventually leads to representations of which we can study the properties. The Configurational Matrix is about a level of abstraction just one step beyond the issue representationalism vs. derivationalism.

The latter issue should not be confused with the empirical issue whether we need "move alpha", or, in current terminology, whether we need both "merge" and "move". One can be committed to derivationalism in general and to derivational "merge" in particular, without being committed to a separate process known as "move".

I am agnostic about the issue derivationalism vs. representationalism, but I see no evidence for "move", since all its properties follow from the Configurational Matrix, which also characterizes "merge" (base rules), gapping, anaphora and other grammatical construction types. In order to illustrate this, I will first briefly discuss the essential properties of the Configurational Matrix:

(54) **Properties of the Configurational Matrix:**

- (i)  $\alpha$  precedes  $\delta$  (Kayne 1994)
- (ii) bi-uniqueness: one  $\alpha$  for  $\delta$  and one  $\delta$  for  $\alpha$  (Koster 1987, 1998)<sup>12</sup>
- (iii) bilocality (replaces c-command and locality) (Koster 1987)
- (iv) recursion: both  $\alpha$  and  $\delta$  can be a  $\beta$

There is much evidence that in grammatical relations the antecedent precedes the dependent element. I will therefore assume the strongest possible hypothesis, namely that it is universally the case that  $\alpha$  precedes  $\delta$  in underlying structures. At this point, this only is an empirical generalization awaiting deeper understanding (see Kayne (1994) and Chomsky (1995, ch. 4) for some discussion). There are, of course, many superficial deviations from this pattern, but if the antecedent-precedence hypothesis is correct, all these deviant orders are caused by reordering ("move alpha").

Precedence of the antecedent over the dependent element can be found in a great number of constructions. In local ellipsis (like gapping), for instance, the antecedent always precedes the dependent element (the gap):

- (55) a. John *reads* papers and Mary -- books  
b. \*John -- papers and Mary *reads* books

This is not an accidental property of English but a property holding for other languages as well. So, let us assume that it is universal. Interestingly, there also is backward ellipsis (e.g., Right Node Raising) but that involves a non-local relation and as such falls outside the scope of the Configurational Matrix. Antecedent-precedence only holds for *local* grammatical relations.

Similarly, the major "movement" classes in English (Wh-movement and NP-movement) involve movement to the left rather than to the right:

- (56) a. *Who*<sub>i</sub> did you say *t*<sub>i</sub> told you that you were happy  
b. \*You said *t*<sub>i</sub> told you that you were happy *who*<sub>i</sub>

This pattern seems to be universal as well, because it has been known for a long time that the overwhelming majority of languages with Wh-movement moves Wh-elements exclusively to the left.

Extraposition rules were supposed to move material to the right. Such rules are not only excluded by the Configurational Matrix, I have also shown elsewhere that extrapositions do not have the properties of "move alpha" at all and are better analyzed as parallel construals along the lines of what we find in coordination (see note 5 for references). All of this confirms the claim of Kayne (1994) that "movement" is exclusively to the left.

In the present context it is extremely important that the Configurational Matrix also determines the properties of what used to be called base or X-bar structures. In a head-

---

<sup>12</sup> In Koster (1987), this property is stated as simple "uniqueness" instead of "bi-uniqueness", on the basis of problematic facts (one antecedent with two anaphors) like *John talked with himself about himself*. However, I do no longer consider this a counterexample to "bi-uniqueness", since I think the example can be reanalyzed as involving two links rather than one. For this and other reasons I prefer the much stronger "bi-uniqueness" (one-one relation between antecedents and dependent elements) over simple uniqueness (one-many relation between antecedent and dependent elements).

complement configuration, for instance, the head can be seen as the antecedent  $\alpha$  and the complement as the dependent element  $\delta$ . Since the Configurational Matrix stipulates that  $\alpha$  precedes  $\delta$  in *all* local grammatical relations, it follows that the only possible base order of natural languages is VO (cf. Kayne (1994)). All other orders, such as the OV of German and Dutch, must involve chain formation (displacement by "move alpha"). In other words, head precedence is an instance of the general antecedent precedence we also find in movement and gapping constructions.

I have emphasized the antecedent precedence property of the Configurational Matrix, because it has a direct bearing on the word order issues (OV, VO) we are concerned with. I will only briefly outline the far-ranging bi-uniqueness and bilocality properties.

Bilocality includes Bounding Theory. I will not be further concerned with it here and only mention that a typical local domain is defined by exactly one node  $\beta$ , instead of the two nodes stipulated by classical Subjacency (see Koster (1987) for discussion). One-node locality makes it possible to make a general statement for the local domains of many different construction types. There is, for instance, almost no difference between the island properties of "movement" constructions and gapping (cf. Koster (1978) and (1998)). Classical c-command --an anomaly with its reference to the notion "branching node"-- appears to be nothing else than locality defined on the antecedent rather than on the the dependent element. The properties of bi-locality are as follows:

(57) **Bilocality:**

- (i)  $\delta$  must have an  $\alpha$  in  $\beta$  (  $[\beta \dots \delta..]$  )
- (ii)  $\alpha$  must have a  $\delta$  in  $\beta$  (  $[\beta \dots \alpha..]$  )
- (iii)  $\beta$  has the same value in (i) and (ii) in a given relation
- (iv)  $\beta$  is slightly parametrized (perhaps in predictable ways) for different relations<sup>13</sup>

I will leave it at that and only mention that this formulation makes it possible to have a full unification between the locality properties (formerly c-command and Subjacency) of "movement" and Gapping (Koster (1998)).<sup>14</sup>

Bi-uniqueness means that in any local relation, there is always exactly one  $\alpha$  for  $\delta$  and one  $\delta$  for  $\alpha$ . This is a somewhat neglected property of grammatical relations, responsible for, for instance, the  $\theta$ -criterion and the fact that two verbs cannot be  $\theta$ -related to one subject. In

---

<sup>13</sup> I assume the same value for  $\beta$  for "movement" and "gapping", namely XP+ (i.e. the minimal maximal projection of type XP *plus* its functional extensions. Thus, under this interpretation, the bounding domain is not the VP but its maximal functional extension CP. For X-bar structure, the value of  $\beta$  is just XP. For anaphors, it is the minimal XP containing Tense or a Subject (or some similar domain). In all cases, specification of one node  $\beta$  is enough and reference to more nodes than one (as in classical Subjacency) appears to be based on the use of irrelevant contexts as evidence, such as the end of the VP, where even the two nodes of Subjacency are not sufficient (see Koster (1987, ch. 4) for discussion). The bounding node  $\beta$  appears under the heading "blocking category" in Chomsky (1986a). The barriers theory is less than optimal because it fails to recognize that intermediate projections (like VP and IP, which are not XP+) are irrelevant for bounding theory. Note the systematic exceptional character of IP in the barriers theory as well as the complicated machinery necessary to escape from the VP.

<sup>14</sup> Note that bilocality allows more than c-command, but that the difference is accounted for by independent factors ("movement" only to the left, to a Spec or head position). Bilocality prevents lowering, which appears to be sufficient. This relaxation of c-command makes it possible to explain why the antecedent of gapping cannot be "too low". Apart from the effect of independent factors, there is no difference in locality properties between "gapping" and "movement" (see Koster (1998)).

order to prevent the latter violation of bi-uniqueness, each verb with a subject  $\theta$ -role always has its own subject, requiring PRO in certain cases.

For present purposes, it is important that the following properties of phase structure follow from the Configurational Matrix:

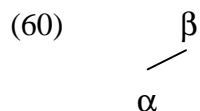
- (58)
- a. binary branching (Kayne (1984))
  - b. the single-complement condition (one  $\delta$  for a head  $\alpha$ ) (Chomsky (1986b))
  - c. the single-Spec condition (one  $\alpha$  for a head  $\delta$ )

Consider for a moment how phrase structure is determined by the Configurational Matrix, in conjunction with what I consider a natural extrapolation from the traditional idea that projections emerge from heads:

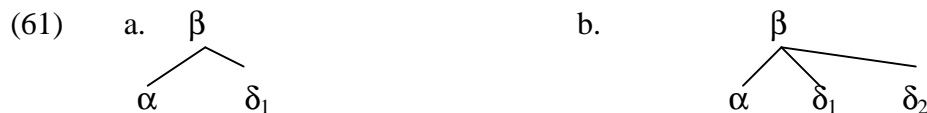
(59) **Head orientation of projection**

- a. only heads project
- b. all non-heads are head-related

Suppose that head projection means that if a head  $\alpha$  projects, it has the property of being dominated by a node  $\beta$ :



The head's property of being dominated by another node can be shared with a non-projecting node  $\delta_1$ , in principle a non-head:<sup>15</sup>



Only the binary branching pattern (61a) is compatible with the Configurational Matrix, while the ternary pattern (61b) would be a violation of bi-uniqueness:  $\alpha$ 's property of being dominated by  $\beta$  can only be shared with exactly one  $\delta$ .<sup>16</sup>

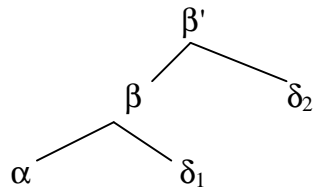
The requirement of bilocality is trivially fulfilled in (61a), since the minimal domain  $\beta$  containing  $\delta_1$  also contains  $\alpha$  and vice versa.

Further adjunction to the right is excluded in (partially) the same way:

<sup>15</sup> I would like to exclude head adjunction to heads entirely (outside of lexically oriented contexts).

<sup>16</sup> I assume that all branching is binary and that intransitive verbs can be reduced to transitive verbs in underlying structure (see Hale and Keyser (1993)).

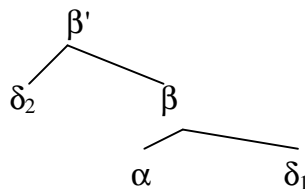
(62)



According to (59b), both  $\delta_1$  and  $\delta_2$  must be head-related. In this structure, too,  $\delta_1$  and  $\delta_2$  would be both head-related to the only available head, namely  $\alpha$ . This is the same type of head-relation, namely the shared property of "being dominated by  $\beta'$ ". This is excluded by bi-uniqueness, which allows only one-one relations of a given type. In general, rightward adjunction is excluded by the conjunction of bi-uniqueness and (59b). This explains the fact --if it is a fact-- that Specs do not occur on the right of a head.

Specs on the left, in contrast, appear to be possible:

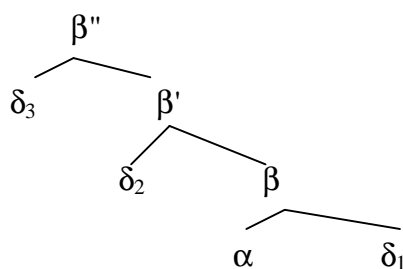
(63)



This structure is possible thanks to antecedent precedence. According to antecedent precedence, the head  $\alpha$  can only be the antecedent of the following  $\delta_1$ , not of the preceding  $\delta_2$ . However, there is another way for  $\delta_2$  to be head-related, namely by making  $\delta_2$  itself the antecedent ( $\alpha$ ) and the head the dependent element ( $\delta$ ). This is the well-known Spec-head relation.

It further follows from bi-uniqueness and (59b) that there is exactly one Spec-head relation per projection. Further adjunction to the left is excluded:

(64)



A structure like this, with two Specs on the left of the head, is excluded because both  $\delta_2$  and  $\delta_3$  would be head-related to  $\alpha$  in the same way (namely in their being dominated by  $\beta''$ ).

All in all, it appears that the Configurational Matrix, in conjunction with a traditional assumption like (59), entails that there is only binary branching and that the only permissible building blocks of phrase structure are head-complement and Spec-head, in the order Spec-head-complement. Both leftward and rightward adjunction (of non-heads) is excluded, which makes the theory more restrictive than most previous theories of phrase structure.

Most important of all, of course, is the fact that this restrictive shape of phrase structure is determined by the same Configurational Matrix which also determines the shape of chain links ("movement"), gapping configurations, and all other local grammatical relations.

### 3.2.2 *Minimalism: movement makes heads visible*

I will not be concerned in this article with the general merits of Minimalism as discussed by Chomsky (1995). Instead, I will focus on one aspect of the current minimalist framework which is crucial for my current purposes and which, as mentioned before, I see as a big step forward with respect to earlier theories of generative grammar. Characteristic of current theories is the idea that "movement" serves purposes of feature checking rather than purposes of semantic interpretation. Furthermore, we see a sharp decline of the interest of levels of representation such as s-structure, d-structure and LF. Personally, I believe that "move alpha" is part of the same older inventory of theoretical apparatus.

As noted in the introduction, linguistic theorizing was for a long time dominated by EST-style theories, which were originally designed with an eye on semantic interpretation ("Interpretive Semantics" as opposed to "Generative Semantics"). Such theories made a distinction between d-structure and s-structure, two levels of representation connected by movement transformations (eventually reduced to "move alpha"). As mentioned above, the main purpose of this design was to combine the thematic meaning contributions of d-structure with the linear, scopal meaning contributions of s-structure, in order to derive a sufficient input structure for semantic interpretation. Later on, LF was added as yet another level of representation, on the basis of an even more questionable form of "movement", namely covert or LF movement.

At least since Katz and Postal (1964), however, there was a parallel development (interestingly, also as part of EST theorizing) which undermined the design in question. Katz and Postal tried to develop a theory in which all semantic interpretation was based on deep structure. In order to do so, they had to code future landing sites --standing for the surface structure aspects of meaning-- into their deep structures. In retrospect, this was a step that has gradually but substantially changed linguistic theory.

This path of structure-preserving movement was further followed in the dissertation of Joseph Emonds (1970), which made it clear that a substantial part of structure derived by "movement" is not different from what is generated by base structures. Implicitly, this undermined the whole project of level-oriented syntax with movement transformations.

The development of trace theory did for surface structure what the Katz/Postal extensions of phrase structure did for deep structure: it coded essential aspects of deep structure into surface structure. In Chomsky (1975), for instance, it is explicitly said that all semantic interpretation can be based on s-structure thanks to the presence of traces.

All in all, it is clear that if most aspects of one level can be coded into another level, the whole notion of levels is undermined. The development of the last ten years is just a logical outcome of the dynamism in question: enrichment of structure is not a supplement to level theory cum "movement" but an alternative to it.

So, as it stands, we seem to arrive at a theory with one level of representation (apart from the morpho-phonetic extensions) in which the former d-structure is coded by traces and in which the old s-structure is enriched by functional projections. Covert movement and LF have become what they have always been: superfluous artifacts (see Koster (1987) for further

discussion). In our current theories, phrase structure is sufficiently enriched to do everything with theories that allow overt "movement" only.<sup>17</sup>

Actually, we do not need "movement" at all to create chains. What seems sufficient is a theory with lexical and functional projections, the Configurational Matrix (which is much more general in its scope than "movement" theory) and certain assumptions about lexicalization.

I will adopt the idea from standard Minimalism that "movement" serves the purpose of feature checking and reinterpret it as a "movement"-free proposal about lexicalization. In general, lexicalization is necessary to make structure visible. Languages differ somewhat in their system of lexicalization, but all languages avoid redundancy, for instance by the device of ellipsis.

Throughout and in accordance with current practice, I am assuming (as further spelled out in the next section) that lexical projections are embedded in a shell of functional extensions which indicate certain functional roles of the elements contained in the lexical projections. If both functional projections and lexical projections would be completely spelled out, sentences would become extremely long and cumbersome. So, a core problem of language design seems to be how both lexical structure and functional structure can be made visible in an optimally efficient way.

It is clear that there are only very limited possibilities to sacrifice direct lexicalization of lexical projections. Usually, at least the heads of lexical projections must be lexicalized, while languages differ somewhat in the ways in which they lexicalize non-heads (control, pro-drop, discourse-based interpretation of arguments in Chinese, etc.).

Functional structure is largely predictable and only parsimoniously lexicalized by natural languages.<sup>18</sup> Complementizers, for instance, are predictable and can be optionally dropped in many languages. Direct lexicalization of Tense and Agr features is relatively rare in Indo-European, and English *do*-support is one of the few examples that come to mind.

So, how is functional structure made visible in general? It seems to me that "movement" is the optimal answer from the point of view of efficient and parsimonious lexicalization. Since visibility is primarily a matter of heads, all "movement" is aimed at heads. There are three ways to make a functional head visible: 1) by direct lexicalization (as in the case of complementizers and English *do*-support), 2) by head movement (as in Dutch and German Verb Second), and 3) by lexicalizing the Spec with an XP.

Phrase structure (as described in the previous section) is so tightly constrained that each displacement of an XP indicates the presence of an immediately following head. NP-movement in English, for instance, makes AgrS (or some other nominative-related head) visible. Similarly, Wh-movement makes visible a head that bears the <+wh>-feature. In some languages (like Dutch), this can be done by redundantly spelling out the head as well:<sup>19</sup>

---

<sup>17</sup> Although I do not consider "move" as a primitive notion of the theory of grammar, I will continue to use the term "movement" for chain formation.

<sup>18</sup> Some linguists, under an extremely narrow and descriptivist conception of empirical evidence, take the parsimonious lexicalization of functional structure as "lack of evidence" for functional structure. Although all theoretical constructs must be justified, that often takes some time. *A priori* hostility with respect to theoretical constructs (as happened before with trace theory and many other innovations), often points in the direction of an anti-theoretical bias.

<sup>19</sup> Supporting evidence for the view presented here (as pointed out by Jan-Wouter Zwart) is that without a Wh-phrase in the Spec of C, the element *of* is obligatory in Dutch, as in: *Ik vraag me af \*(of) hij komt* ("I wonder if he comes").



- (65) Ik wil weten    *wie*<sub>i</sub> (of)    het *t*<sub>i</sub>    gedaan heeft  
 I want know    who <+wh> it    done    has  
 "I want to know who did it"

The proposed Wh-phrase *wie* ("who") efficiently makes two things visible, namely material from both the lexical and the functional projections involved: 1) the content of the object (to which it is connected by the trace) and 2) the presence of a functional head at the beginning of the clause, which can redundantly be spelled out as *of*. Although there is a lot of dialect variance, most of the time languages avoid the redundancy of spelling out both the Spec and the head of a functional projection. Economically speaking, either lexicalizing the head or the Spec is enough.

Given the tightly constrained theory of phrase structure proposed in the previous section, it would be too narrow a conception of empirical evidence to say that functional projections are only "proven" to exist when the heads are spelled out (as with complementizers or English *do*). In fact exactly the same result can be arrived at by spelling out the Spec (or by movement of a head from the lexical projection). "Movement" is just one of the ways to make a functional head visible.

"Movement" is nothing else than making use of material from the lexical projection to make the functional projection visible. The structures in question can be lexicalized directly, without an extra movement operation. The form of chain links is completely determined by the Configurational Matrix, which also determines the form of all other local relations.

An interesting question is why languages have functional extensions of their lexical projections in the first place. Obviously, there is more to language than just lexical argument structure. Functional structure makes it possible for a language to use the same argument structure (as given, for instance, in a VP) for many different purposes, such as making questions, statements or whatever. So, functional structure contributes much to the expressive potential of natural language. "Movement" makes it possible to do everything in an optimally efficient way by using the structure-identifying power of lexical material more than once.

### 3.2.3 *The lexical nucleus (VP) is embedded in a universal shell of functional projections*

#### 3.2.3.1 *The universal structure*

A next idea that is crucial for my explanation of the word order differences between English and Dutch is that lexical projections are embedded in a shell of functional projections. Since its inception, generative grammar has seen phrase structure as a combination of what are now called lexical projections and functional elements (like the auxiliary position *C* in *Syntactic Structures* (Chomsky (1957)) and the further use of *Aux* in many later versions of Universal Grammar). Particularly since Chomsky (1986a) and Pollock (1989), it has become customary to see such functional elements as heads of full projections with a Spec and a complement, etc. More often than not, the functional elements were thought to be situated on the left of the VP, and with the Configurational Matrix (which incorporates Kayne's antisymmetry idea), there is no choice: universally, functional projections are exclusively on the left of the lexical projections of which they are extensions.

The version of the VP and its extensions I assume is as follows:

(66) **Universal clause structure**

$XP C_{\langle \pm wh \rangle} XP AgrS Adv_1^* XP T XP Dat XP Acc Adv_2^* XP Pred^* [VP]$

Most elements of this structure are familiar from the literature and were not specifically designed to solve the word order problems of English and Dutch which are the topic of this article.<sup>20</sup>

The VP itself consists minimal of a V and has maximally one complement YP in the order V-YP, the only order permitted by the Configurational Matrix. For the word order problems under discussion, this means that all languages, including Dutch and German, are SVO at the deepest level of abstraction.

Among the complements, I assume Small Clauses as familiar from the work of Richard Kayne, Teun Hoekstra and many others.<sup>21</sup> Small Clauses include secondary predications, verb-particle constructions and double object constructions. Thanks to the availability of well-founded Small Clause analyses, we can maintain that a verb has at the most one complement, the only possibility permitted by the Configurational Matrix (thanks to its bi-uniqueness property).

As for [Spec, V] I assume the internal-subject hypothesis of Koopman and Sportiche (1991), i.e., VPs (or rather the  $V^{\max}$  nodes containing the VPs in their analysis) have subjects as their Specs.

Also for the functional projections, I rely on familiar notions. XPs in this scheme stand for Specs of the immediate following head, which has the next XP-head combination as complement. I have omitted this extra structure in (66), but it must be read as:

(67)  $[_{CP} XP C [_{AgrsP} XP AgrS [_{AdvP} XP Adv \text{ etc. etc., .....}]]]$

*C* is the complementizer node, which can be +wh or -wh. *AgrS* stands for Subject Agreement and is usually identified by a Spec (NP or DP) with nominative Case.

I also assume a *Dat* (= Dative) and *Acc* (= Accusative) position, without committing myself to any strict ideas as to the exact nature of these positions. In many earlier theories (such as the one of Vanden Wyngaerd (1989)), *Acc* was seen as the element that determines object agreement (*AgrO*). Since Vanden Wyngaerd used this position as a position where the object (of Dutch) is moved to, one also needs a similar position for indirect objects, hence my distinction between *Dat* and *Acc* as the positions that must be checked (identified) by XPs with dative and accusative features, respectively. The positions *AgrS*, *Dat* and *Acc* correspond with the three non-oblique argument positions which were seen as part of Universal Grammar by the school of Relational Grammar (as promoted by Postal and Perlmutter for some time).<sup>22</sup> At this point, I am not concerned about the exact nature of these positions. My claim is not about labels but about the fact that we need *two* non-subject checking positions (apart from *Pred*, to be discussed next).

So far, my assumptions are conventional. The same can be said about the assumption that there are adverbial positions (the star \* in *Adv* and *Pred* means that the projections in question

<sup>20</sup> The structure (66) is a much simplified version of reality. For its bracketing, see the next example (67). For ease of exposition, I distinguish only two broad classes of Adv<sub>s</sub> in two positions. The elements of Adv<sub>1</sub> often also precede the subject (the XP before *AgrS*). Furthermore, I ignore the finer structure of the VP as discussed, for instance, by Hale and Keyser (1993) and Chomsky (1995, 315).

<sup>21</sup> See for instance Kayne (1984), Hoekstra (1988) and Den Dikken (1995).

<sup>22</sup> See for instance Perlmutter (1983).

(AdvP and PredP) are recursive, to account for the fact that there can be an unlimited number of adverbial expressions in a clause). For my present purposes, I have to assume minimally two classes of adverbial elements: sentence adverbials, like *probably* (*Adv*<sub>1</sub>) and VP-adverbials, like *yesterday* (*Adv*<sub>2</sub>).

More generally, I will analyze adverbials either as *Adv* heads, or as Specs of *Adv* heads (when they are XPs bigger than heads, like PPs with adverbial function).

As in the case of the argument positions, I am only making minimal assumptions about adverbials, which, as such, are not the topic of this article. The minimal picture I present here is no doubt very much idealized and simplified. The overall picture can be refined for instance along the lines of Alexiadou (1997) and Cinque (1997).

### 3.2.3.2 On Pred and Pred Phrases

Crucial for what follows is the existence of a Pred Phrase (PredP) or "enlarged VP" as opposed to the lexical, nuclear VP. The Pred Phrase has a long history in generative grammar (see for instance Chomsky (1965), where it had a somewhat different meaning) and recently, it has been revitalized in the study of Dutch syntax (see Zwart (1993) and (1997) and Koster (1994)). The nuclear VP gives information about the range of V complements, but it does not give a complete picture of functional roles of these complements. In fact, as I assume, one reason why VPs have a functional shell is to provide a range of functional roles for V complements, in particular a range richer than the one entailed by just being a complement.

As it turns out, V complements are either independent arguments or part of a complex predicate. Some V complements, like PPs and APs are always part of a complex predicate, while NPs can fulfill either role. The function of the functional heads *AgrS*, *Dat* and *Acc* is to provide predicate-independent argument positions (their Spec positions) for the NPs of the VP. The function of *Preds* is to indicate that their Specs are part of a complex predicate. *Preds* provide the licensing positions for the non-Case bearing VP complements, particularly APs and PPs.

As I have shown in Koster (1994), there is quite a bit of evidence that in a so-called SOV language like Dutch, there is not just one uniform complement (XP-) position preceding the verb. It makes, on the contrary, a lot of sense to distinguish two kinds of positions: NP positions external to the predicate and general XP positions internal to the predicate. The available empirical evidence includes scrambling. Being to the left of VP-external adverbials is very natural for NPs, while it is much harder for complement PPs and APs:

- (68)
- a. Hij heeft het boek gisteren gelezen  
he has the book yesterday read  
"He read the book yesterday"
  - b. \*Hij is naar huis gisteren gegaan  
he is to home yesterday gone
  - c. \*Hij is ziek gisteren geweest  
he is sick yesterday been

There is some variety in this area in connection with intonation and choice of lexical material, but basically, NPs are in a completely unmarked position on the left of adverbials, while APs

and PPs are not. Idiomatic NPs are more natural to the right of adverbials, just as what we see with APs and PPs (cf. De Hoop (1992)):

- (69) a. ?Hij heeft de clown gisteren uitgehangen  
he has the clown yesterday out hung  
"He acted like a clown yesterday"  
b. Hij heeft gisteren de clown uitgehangen

Another fact described in Koster (1994) is the distribution of stranded prepositions, like *mee* ("with").<sup>23</sup> I assume that stranded prepositions in Dutch are always part of the PredP, and that this is *a fortiori* the case with material to the right of these prepositions. It appears that it is natural for PPs and APs to be to the right of stranded prepositions (70a-b), while it is impossible for most NPs (70c):

- (70) a. Waar heb je *mee* aan je dissertatie gewerkt?  
where have you with at your dissertation worked  
"With what did you work on your dissertation?"  
b. Waar heb je het hek *mee* zwart geverfd?  
where have you the gate with black painted  
"With what did you paint the gate black?"  
c. \*Waar heb je *mee* die dissertatie geschreven?  
where have you with that dissertation written

Interestingly, idiomatic NPs (naturally seen as part of the predicate, pattern like APs and PPs in this respect:

- (71) Waar heb je *mee* de clown uitgehangen?  
where have you with the clown out hung  
"With what did you act like a clown?"

Many languages express the distinction between independent arguments and arguments incorporated in the Pred by two different cases (cf. De Hoop (1992)). In Dutch, the distinction is also made: so-called prepositional objects are never independent arguments but always part of the PredP.

In general, I will assume that the following is true (with the exception of clausal complements, which are indirectly licensed (see Koster (forthcoming))):

(72) **Licensing**

The XPs of the VP must be functionally licensed, either as independent argument or as part of PredP

Licensing of independent arguments is done through the three Case heads, *AgrS*, *Dat* and *Acc* of (66). Licensing of PredP elements is done via the Spec positions of *Pred* heads. Like in the

---

<sup>23</sup> For stranded prepositions in Dutch, see Van Riemsdijk's classical study on this topic (Van Riemsdijk (1978)).

case of *Adv<sub>s</sub>*, I assume that there can be any number of these heads, which, like all other functional heads must be made visible, for instance by moving something into their Specs. In Dutch, objects must be moved to either the argument positions *Dat* and *Acc* or to the Spec of a *Pred* (as with idiomatic NPs like in (71)). This accounts for the OV character of Dutch:

- (73) Zij heeft [<sub>AccP</sub> *het boek<sub>i</sub>* *Acc* [<sub>VP</sub> *gelezen t<sub>i</sub>* ]]  
 she has the book read  
 "She read the book"

The underlying structure of the VP is head initial (VO) just as in English and the OV order arises by "movement" of the object to the Spec position of the AccP. The all-important question which I will answer in this article is why English fails to take this step, so that its overt order remains VO.

Prepositional objects are "moved" to the Spec of a PredP in Dutch, a position to the right of *Acc* according to (66):

- (74) Hij heeft [<sub>PredP</sub> *aan zijn vader<sub>i</sub>* *Pred* [<sub>VP</sub> *gedacht t<sub>i</sub>* ]]  
 he has of his father thought  
 "He thought of his father"

An example involving both licensing by "movement" to [*Spec, Acc*] and [*Spec, Pred*] is the following (where SC stands for Small Clause):

- (75) Hij heeft [<sub>AccP</sub> *het hek<sub>i</sub>* *Acc* [<sub>PredP</sub> *zwart<sub>j</sub>* *Pred* [<sub>VP</sub> *geverfd* [<sub>SC</sub> *t<sub>i</sub>* [ *t<sub>j</sub>* ]]]]  
 he has the gate black painted  
 "He painted the gate black"

If these "movements" are necessary for licensing, the crucial question arises once more why English fails to show them.

### 3.2.4 *There is no adjunction (apart from processes involving heads)*

So far, all "movements" we have considered serve the purpose of head visibility: a head is made visible by lexicalizing its Spec. As mentioned before, it is also possible to lexicalize a functional head directly, either with an independent lexical element, as in the case of complementizers or English *do*, or by sharing the head of a lexical projection. An example of the latter is the kind of "verb movement" known as Verb Second in the study of Dutch and German. In order to arrive at a restricted theory of grammar, I want to exclude all non-lexical or -morphological adjunction. This means that Verb Second is not a form of head adjunction, but a direct lexicalization of the relevant head, much along the lines of what was formerly called substitution (the original account of Verb Second in Den Besten (1977)). There is no necessity to weaken linguistic theory with a class of non-lexical or -morphological head adjunctions. For so-called verb movements (like Verb Second), ordinary lexicalization of head features (independently needed for the lexicalization of the heads of lexical projections) suffices.

I will assume, then, that all chain formation is nothing else than lexicalization of the universal scheme (66), particularly in a way such that functional projections and lexical

projections share lexical material. None of this involves adjunction, which I would like to limit to lexically oriented processes involving heads, such as incorporations and compounding (such as in verb clusters in Germanic, causative formation in Japanese, etc.).<sup>24</sup>

Since the universal scheme (66) only has functional projections to the left of the VP (the only option permitted by the Configurational Matrix), all chains have their head on the left and their tail on the right. All rightward movements, like traditional extraposition rules, are excluded because there is nothing on the right of the VP to lexicalize. This is confirmed by the strong evidence that extraposed materials are not related to their antecedent according to the properties of "movement" anyway (see Koster (forthcoming)).

Altogether, we arrive a very restrictive theory of chain formation, namely a lexicalization-based theory without adjunctions or lexicalizations on the right of the VP.

### 3.2.5 *Parametrization: lexicalization and the "size" of the checking phrase*

Before discussing which kind of parametrization determines the difference between English and Dutch, something must be said about parametrization in general. Although it is generally agreed upon that differences between languages are due to parametrization, there has not been much of a general theory of parametrization so far. As a result, the notion parameter is often used in a completely *ad hoc* fashion, making the word "parameter" near synonymous with the word "difference", merely describing what has to be explained. Ideally, we want to get rid entirely of *ad hoc* parameters such as OV/VO or weak/strong.

Given my view of chain formation as "economic" lexicalization of the universal scheme (66), it is to be expected that parametrization is by and large about lexicalization. Heads are lexicalized directly (lexical heads, complementizers, English *do*), by head sharing (Verb Second, etc.) or by lexicalizing their Spec ("movement" of XP, adverbial and predicate modifiers). Languages do not do all of this in the same way and this, it seems, is the major area of parametrization.

I will not be further concerned with those forms of parametrization but focus on another form of parametrization abundantly documented for various natural languages, namely the size of the checking phrases found in Specs. This dimension of language variation is uncontroversial and widespread and I will use it to explain the word order differences between English and Dutch, without any recourse to *ad hoc* parameters such as OV/VO or weak/strong.

It has been recognized almost since the beginning of generative grammar that languages differ in the way they define Pied Piping (i.e., in defining phrases larger than strictly necessary for feature checking, see Ross (1967)). Take the following English examples:

- (76)
- a. [<sub>NP</sub> *Who*] did you talk with ?
  - b. [<sub>PP</sub> With [<sub>NP</sub> *whom*] ] did you talk?
  - c. [<sub>PP</sub> With [<sub>NP</sub> the brother [<sub>PP</sub> of [<sub>NP</sub> *which* man]]]] did you talk?

Assuming that Wh-movement makes a <+wh> functional head (C) visible, checking by a minimal Wh-phrase would be sufficient, as in (76a). However, English also allows more inclusive phrases to do the job, such as the PP in (76b). The Wh-phrase can be embedded

---

<sup>24</sup> See for instance Baker (1988) and Zwart (1997) for discussion.

rather deeply, as in (76c). It is usually said in such cases that the <wh>-features of the minimal Wh-phrase percolate to the more inclusive phrases.

Languages differ substantially in the way they define the possible size of their checking phrases. In Dutch, for instance, preposition stranding is impossible in most cases, so that we do not find a counterpart of (76a), but only a pied-piped version:

- (77) a. \*<sub>[NP Wie]</sub> heb je mee gepraat?  
           who have you with talked
- b. <sub>[PP Met wie]</sub> heb je gepraat?  
           with who have you talked

In German, it is even possible to pied-pipe a whole clause, as originally described by Ross (1967) and analyzed in detail by Van Riemsdijk ((1984) and (1994)):

- (78) Der Hund <sub>[CP den zu fangen]<sub>i</sub></sub> ich versucht habe <sub>t<sub>i</sub></sub>  
       the dog which to catch I tried have  
       "The dog which I tried to catch..."

Van Riemsdijk (1984, note 19) cites similar constructions from older forms of Italian (originally from Noordhof (1937)):

- (79) I mei amici <sub>[CP ai quali per scrivere]<sub>i</sub></sub> sono stato a casa <sub>t<sub>i</sub></sub>  
       the my friends to whom for-to write I-am stayed at home  
       "My friends. in order to write to whom, I stayed at home (?\*)"

In sum, Pied Piping shows a lot of variety across languages and, although the observed variety definitely awaits further explanation, the dimension of variation as such is uncontroversial.

Perhaps "large" Pied Piping (involving a whole clause) is what is behind the movement of the whole IP into [Spec, C] in Japanese with the structure (86a) and illustrated in (86b) (example from Kuno (1973)):

- (80) a. [<sub>IP<sub>i</sub></sub> [<sub>C<+wh></sub> <sub>t<sub>i</sub></sub> ]]
- b. [<sub>IP</sub> John ga *dare o* butta]<sub>i</sub> ka <sub>t<sub>i</sub></sub> siranai  
           John nom who acc hit [+wh] know-not  
           "I don't know whom John hit"

If the <wh>-feature of *dare* is percolated to the whole IP, this whole IP can serve as the checking phrase, which leads to movement into [Spec, C] of the whole IP rather than the movement of a minimal phrase, as in English (cf. Kayne (1994)). If this view is correct, Japanese does not really have Wh-*in situ*, but overt Wh-movement like English, the difference being in the size of the checking phrase: a whole IP in Japanese and smaller phrases in English.

The core idea of this article is that a single variation in Pied Piping is the key to an explanation of the vast differences between the word orders of English on the one hand and Dutch and German on the other hand. How it all works will be described next.

### 3.3 The parametric difference between English and Dutch

Recall that the OV order of Dutch was explained in section 3.1.3.2 in terms of feature checking, i.e., the "movement" of the NP to the [Spec, *Acc*] ((73), repeated here as (81)):

- (81)     Zij heeft [<sub>AccP</sub> *het boek*<sub>k<sub>i</sub></sub> *Acc* [<sub>VP</sub> *gelezen* *t<sub>i</sub>* ] ]  
          she has       the book               read  
          "She read the book"

The intriguing question raised by this example is why English does not have a similar "movement" to check *Acc* (to functionally license the object). My proposal is that English does have "movement" to check *Acc*, but that unlike in Dutch and German, not the object but the entire VP is "moved" into the Spec of *Acc*, a form of "massive" Pied Piping as proposed by Koopman and Szabolcsi (1998) in related but somewhat different contexts:

- (82)     She has [<sub>AccP</sub> [<sub>VP</sub> read *the book*]<sub>i</sub> *Acc* [<sub>VP</sub> *t<sub>i</sub>* ] ]

In other words, in English the <acc>-feature of the V and its object is percolated to the whole VP, in a way similar to what was observed for <wh>-features in many languages, including English. If this is correct, the difference between English and Dutch boils down to an extremely simple difference in Pied Piping: "movement" of minimal Case-bearing elements (NPs, as in Dutch and German) versus "movement" of the more inclusive phrase (VP) containing these minimal Case-bearing elements (English).

Exactly the same procedure can be used to license the non-arguments (the predicate elements) of the English VP:

- (83)     He has [<sub>PredP</sub> [<sub>VP</sub> thought of *his father*]<sub>i</sub> *Pred* [<sub>VP</sub> *t<sub>i</sub>* ] ]

In this case, *Pred* is checked via percolation through the VP by the PP *of his father*, licensing the latter not as an independent argument (like *the book* in (82)) but as a part of the predicate.

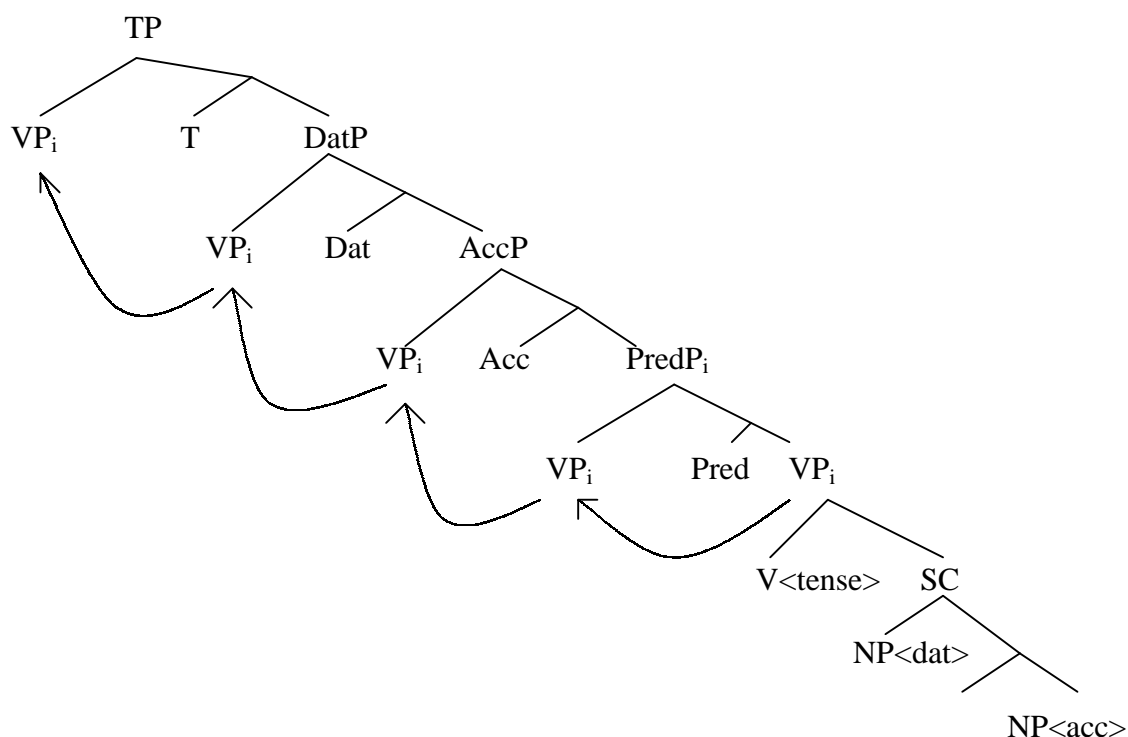
The same can be done with the dative and with the tense-features of the verb. In fact, I will argue that all information related to the verb, as expressed by its features, can be percolated to the VP containing it. I will call the English mode of checking (through Pied Piping) *collective checking* and the Dutch (and German) method *individual checking*.

Since, presumably, English collective checking also includes the tense features of the verb, I will assume that the checking process (which makes the heads necessary for licensing visible) goes as high up in the tree as [Spec, *T*]. Filling [Spec, *T*], however, only serves the purpose of making the head *T* visible. The tensed V itself does not have to be licensed, as expressed by the fact that (72) only mentions VP-internal XPs.

The collective checking process of English follows the route as indicated in (84):



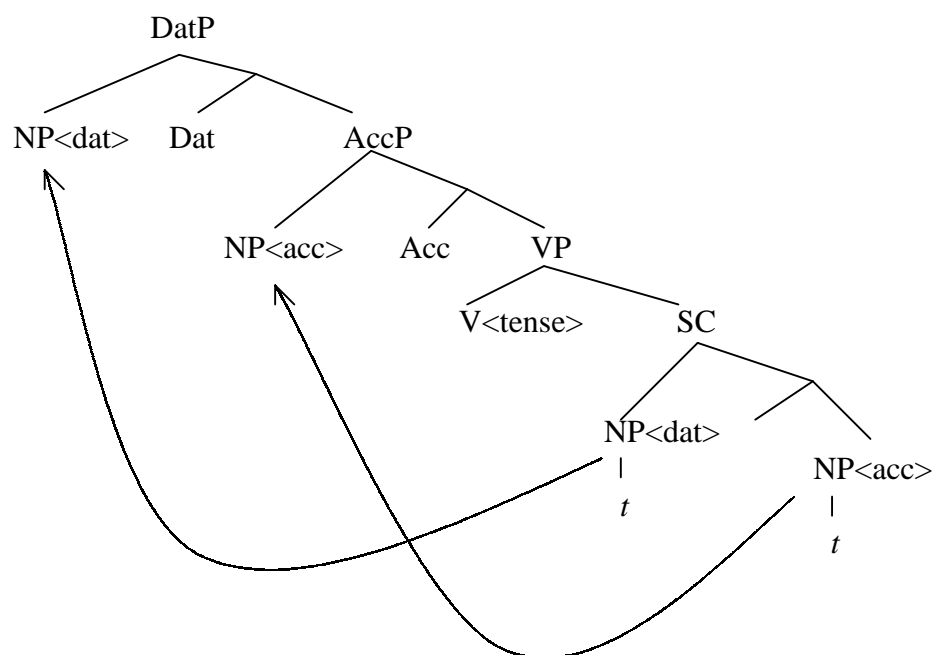
(84) **Collective checking in English (Pied Piping: percolation of V-features to VP)**



I have omitted the subject (as [Spec, V]) here, but I believe that my analysis is compatible with the VP-internal-subject hypothesis. The nominative feature, however, is not an inherent feature of the subject NP and, more importantly, not V-related (like <dat> and <acc>) and therefore not percolated to VP or PredP. It is entirely dependent on VP-external information, particularly on a head determining finiteness. Conventional accounts will do, for instance those which make finiteness dependent on the head *AgrS*. Not much hinges on this matter and I will not discuss it any further here.

In Dutch and German, the features of VP-internal material are not checked collectively (through percolation to VP), but individually. This means that in a structure similar to (84), <dat> and <acc> are not matched with their corresponding heads (*Dat* and *Acc*, respectively) by moving the whole VP, but by moving the NPs in question individually to the Specs of the relevant heads. This absence of Pied Piping yields the familiar OV structures of Dutch and German:

(85) **Individual checking in Dutch and German (without Pied Piping)**



Note that the V is not moved to the head *T* in Dutch, since this head is licensed by moving it to AgrS or C. Without a C, however, the V has to come to the rescue, hence Verb Second in main clauses in Dutch and German (see Zwart (1993) for discussion).

My solution for the word order differences between English and Dutch is entirely based on the generally accepted dimension of language variation of Pied Piping variation and does completely away with the *ad hoc* parameters of earlier theories, such as the meaningless OV/VO parameter or the equally meaningless weak/strong parameter. Both English and Dutch move their checking phrases overtly, the only difference being the size of the moved phrase.

At this point, I would like to claim that all facts that set out to be explained in section 2 are in fact explained by the Pied Piping parameter just discussed. Let us have a look at these facts one by one.

## 4. The facts explained

### 4.1 Overview

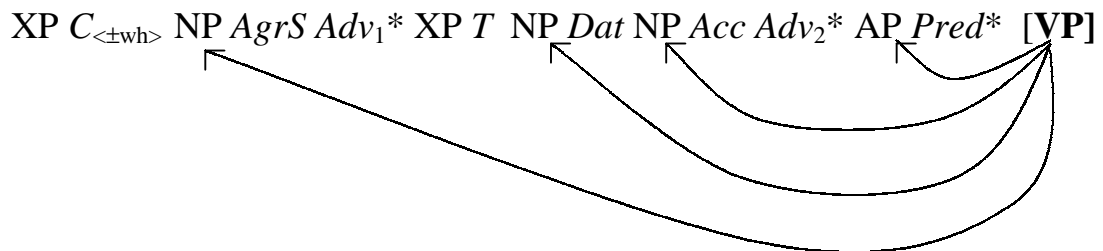
The vast word order differences between English and Dutch (as discussed in section 2) are explained by the Pied Piping parameter in a surprisingly simple way. The parameter can be summarized as follows:

(86) **Pied Piping Parameter**

In Dutch and German the complements of V check their corresponding functional heads *individually*; in English, the complement features are percolated to VP, which checks the functional heads *collectively*

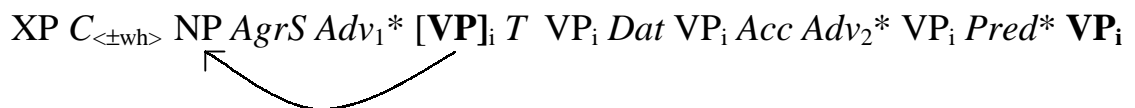
This parameter works against the backdrop of a restrictive theory of Universal Grammar based on the Configurational Matrix (53) and the Universal Clause Structure Scheme (66). This theory excludes non-lexicon-related adjunction and explains all other "movement" of VP material in terms of lexicalization of the functional projections to the left of the VP. If the VP stays where it is (in (66)) and all (or most) of its constituents are used for feature checking, a language shows a great deal of scrambling. This is what we find in Dutch and German:

(87) **Feature checking (and resulting scrambling) in Dutch and German**



In English, the whole VP is used for feature checking. Since it ends up at [Spec, T], (non-Wh) "movement" out of the VP is extremely limited:

(88) **Feature checking (and resulting rigid word order) in English**



After the English VP has moved to its final position in the Spec of T, there are only very few positions available for chain formation ("movement"). The VP-internal subject can still be moved to the Spec of AgrS and that is about it (apart from Wh-movement to the Spec of C which is also possible in Dutch).

In the remainder of this section, I will give a summary of how the individual facts of section 2 are explained.

#### 4.2 English is VO, Dutch and German OV

In the theory just outlined, we can for the first time *explain* why English is VO and German and Dutch OV (at a certain level of abstraction) rather than just stipulating the difference with the arbitrary OV/VO parameter. OV orders are impossible to derive in English, because there are no object positions to the left of the VP anymore when it has arrived at its final position in

[Spec, *T*] (see (88)). Dutch and German do not have Pied Piping, so that not the VP but its constituents do the checking. This leads to OV orders (see (87)). English preserves the original VO order, which, according to the Configurational Matrix (53) is the underlying order of all languages.

#### 4.3 Neither English, Dutch or German has rightward scrambling

Rightward scrambling is excluded by Universal Grammar. Free (non-lexicon-related) adjunction does not exist, only feature checking (lexicalization of functional projections). It follows from the Configurational Matrix (53) that the functional contexts for feature checking are exclusively to the left of the VP (as illustrated by (66), (87) and (88)).

#### 4.4 Unlike Dutch and German, English has no leftward scrambling ("the cage problem")

This fact is also illustrated by (87) and (88). The English checking phrase (VP) is moved up to [Spec, *T*]. Apart from the subject position [Spec, *AgrS*] and the [Spec, *C*] position, there are no further checking positions to the left of [Spec, *T*]. Since English VP-material can move neither to the right (4.3) nor to the left (4.4), the English VP keeps its complements inside (behaves like a "cage" for them). This results in rigid word order.

Scrambling (as in Dutch and German), is the result of individual checking. So, in part at least, the difference between so-called configurational and non-configurational languages follows from collective vs. individual checking.

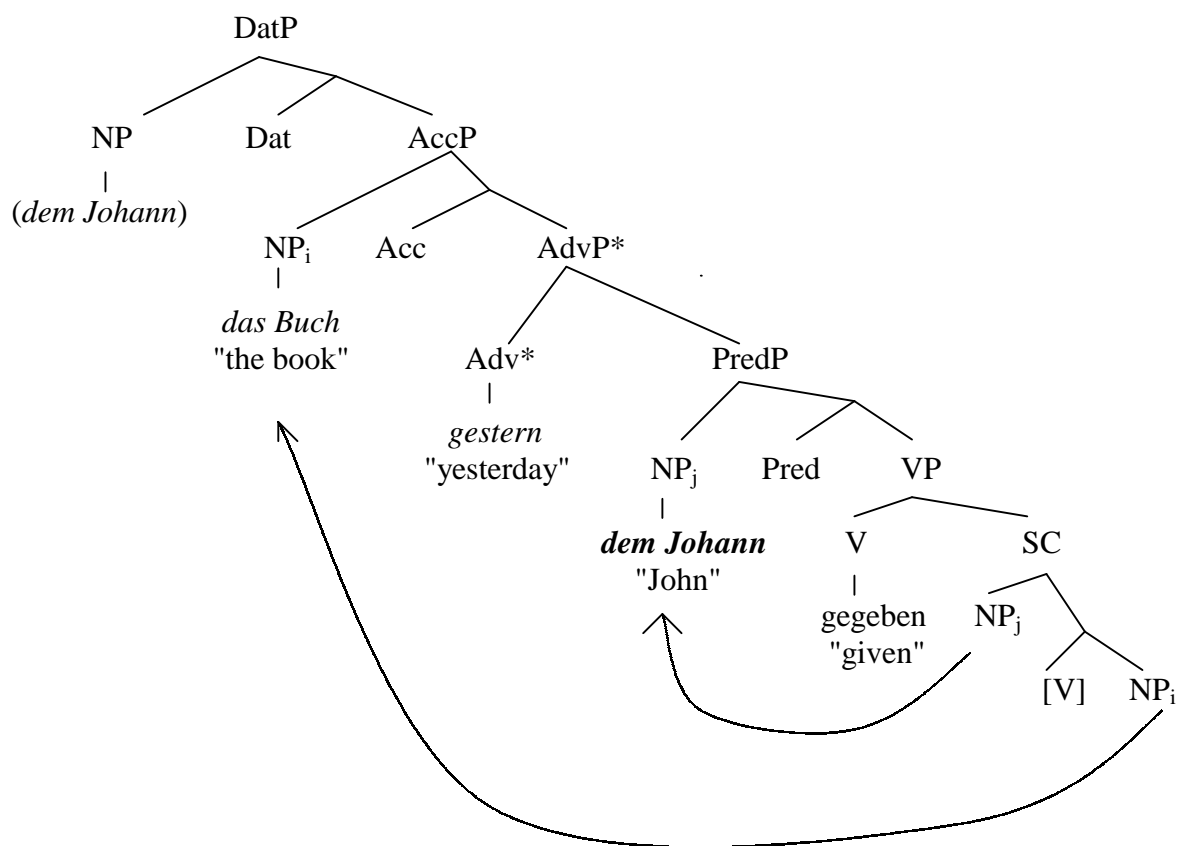
I am assuming that the NP-complements of the VP can be realized either as independent arguments (in the Specs of *Acc* and *Dat*) or as part of the predicate. In the latter case, an NP is moved to a Spec of *Pred*, which is possible thanks to the fact that *Preds* can be made visible by any XP in their Specs. In turn, those Specs are licensed as part of the predicate. This possibility seems to suffice for most scrambling facts in Dutch and German. The individual checking in [Spec, *Dat*] or [Spec, *Acc*] can lead to the unmarked scrambling facts, in which an object is to the left of an adverbial:

(89) Marie heeft *het boek* gisteren gelezen

Unlike what we see in the English VO pattern (\**Mary read yesterday the book*), Dutch and German therefore do not necessarily preserve the verb-object adjacency of the universal underlying structure. As can be seen in (87), the argument positions are to the left of the cluster of adverbials indicated by *Adv*<sub>2</sub>.

However, it is also possible in Dutch and German to reverse the argument order Dat-Acc to Acc-Dat. I will assume that in such cases there always is minimally one NP that is checked as the Spec of a *Pred*. This is illustrated in (90), with an example from German. The dative *dem Johann* appears either in its standard position to the left of *Acc*, or as part of the *PredP*, and therefore to the right of the accusative:

(90) Scrambling in German (and Dutch)



Note that I assume that also in Dutch and German, double objects originate in Small Clauses, along the lines of Kayne (1984). I further assume that different linear orders produce different topic-comment patterns in a dimension independent of what has been discussed in this article.

4.5 In Dutch and German all Advs can appear to left of VP, in English only a subclass

Simplifying for the sake of exposition, we can assume that there are two broad classes of Advs, Adv<sub>1</sub>\* and Adv<sub>2</sub>\* in the Universal Clause Structure (66) (which also shows up in (87) and (88)). The English VP moves to a position, [Spec, T], which only has the class Adv<sub>1</sub>\* to its left. In Dutch and German, the VP remains in its original position, which has both Adv<sub>1</sub>\* and Adv<sub>2</sub>\* to its left. Also the individual checking positions in PredP have both classes of Advs to their left. Thus, in Dutch we find Adv<sub>1</sub>s like *waarschijnlijk* ("probably") and Adv<sub>2</sub>s like *gisteren* ("yesterday") both to the left of the VP (91a), while in English the two types of adverbials end up on different sides of the VP (91b-c):

- (91)
- a. Hij heeft *waarschijnlijk gisteren* gewerkt  
 he has probably yesterday worked  
 "He probably worked, yesterday"
  - b. \*He has *probably yesterday* worked
  - c. He has *probably* worked, *yesterday*

This pattern is entirely as predicted by (87) and (88). In the Dutch and German derived order (87), both *Adv*<sub>1</sub> and *Adv*<sub>2</sub> remain on the left of the [VP], while in the English order (88), the [VP] is "moved" to a position between *Adv*<sub>1</sub> and *Adv*<sub>2</sub>.

#### 4.6 English Adv order shows scope paradoxes (absent from Dutch and German)

In section 2.6, it was observed that the relative scope of adverbials and VP-material is always linear in Dutch (i.e., "more to the left" means wider scope). Anti-linear orders could only be derived by a mirror imaging process (optional addition of Adv heads that have to be made visible) on the right of the verb.

In English, VP-internal material always has narrower scope than adverbials to the right of the VP. This anti-linear scope is excluded by the Configurational Matrix (if scope is thought to correspond with c-command), which is a strong indication that English has a movement rule "disturbing" the linear scope of underlying structure. An example was (51), repeated here as (92):

- (92) a. She [[played a sonata] today and yesterday]  
 b. She [[played a sonata]<sub>i</sub> [today and yesterday [ *t*<sub>i</sub> ] ] ]

In this example, *today and yesterday* has wider scope than *a sonata*. The traditional analysis (adjunction of adverbials to the right of the VP) leads to anti-linear scope and is therefore in conflict with Universal Grammar, which allows linear scope only.

Under the hypothesis proposed in this article, this problem is solved immediately, because the whole VP is "moved" from a position to the right of *Adv*<sub>1</sub>s to a position on their left, as can be seen in (88), as illustrated in (92b) by the movement of the VP [played a sonata].

A further prediction on the basis of (88) is that we also find *linear* scope to the right of the VP, namely in cases with exclusively VP-external material. The mutual order of *Adv*<sub>1</sub>s is not affected by the VP movement proposed for English, so that scope order can remain linear, as in Dutch. That is indeed what we found ((42) repeated here as (93)):

- (93) a. She played two days only once  
 b. She played only once two days

Both examples have a linear scope interpretation, i.e., an interpretation in which the preceding adverbial has scope over the following adverbial.

Of course, English also allows the optional mirror imaging that we assumed for Dutch. This makes it often possible to have two orders:

- (94) a. She read in the yard, yesterday  
 b. She read yesterday, in the yard

A consequence of this analysis is that when anti-linear scope is absolutely obligatory, as in (92) and (95), the elements with narrower scope are part of the VP:

- (95) a. He worked hard yesterday  
b. \*He worked yesterday hard

This is indeed what I have in mind for predicative adjectives like *hard*, which have a VP-internal origin, namely as Small Clause predicates with a Small Clause subject. However, such analyses involve slightly more abstract conceptions of phrase structure which I will not pursue here any further (see Hale and Keyser (1993) for an indication what kind of structures I have in mind). The study of adverbials is still in its infancy and the distribution of these elements involves complications which are beyond the scope of this article (see Alexiadou (1997), Cinque (1998), Sportiche (1994), Rijkhoek (1998) for more elaborate treatments).

## 5. Conclusion

The word orders of English on the one hand and Dutch and German on the other hand show a vast collection of related differences that are not accounted for by the traditional OV/VO parameter. Recent alternatives follow Vanden Wyngaerd (1989) in deriving OV orders by object movement in a way analogous to Wh-movement (feature checking). Although this was a step towards more uniform parametrization, such theories failed to clarify why some languages (like Dutch and German) have these object movements, while others (English, Scandinavian, Romance) fail to show them. Up until now, nothing much was proposed beyond the stipulation that some languages (like Dutch and German) have *strong* features and overt object movement, while other languages (like English) have *weak* features and covert object movement at LF. Parametrization on the basis of weak-strong is of course just as *ad hoc* and unrevealing as a parameter that says nothing else than that there are OV and VO languages (the OV/VO parameter).

The fact that languages differ entails that some arbitrary choice must be made somewhere. However, we are much better off if we can discover parameters which, as a kind, are independently attested and which moreover explain more than one fact.

The Pied Piping parameter (86) proposed in this article is such a desirable device. It is based on alternative choices about the size of checking phrases, an uncontroversial dimension of language variation, which can be illustrated with examples from many languages. By assuming that the English checking phrase for VP-material is the whole VP instead of any of its constituents, we can for the first time explain why there are OV and VO languages in the first place. Moreover, the parameter (86), working together with a restrictive theory of Universal Grammar based on the Configurational Matrix (53) and the Universal Clause Structure Scheme (66), could be utilized not only to predict the OV-VO distinction but also a host of related facts which were never covered at all by the OV/VO parameter.

Another matter is why languages have checking phrases of different size at all. The dimension of differentiation itself is uncontroversial, but a deeper explanation of the variation in question has to be left to future research.

Last but not least, the theory presented in this article suggests a pattern of explanation that makes it perhaps possible to resolve what has been a mystery so far, namely that some languages (like English) have a very rigid word order, while others, even closely related languages like Dutch and German show a remarkable freedom in their word orders.

## References

- Alexiadou, A. (1997) *Adverb Placement: A Case Study in Antisymmetric Syntax*. John Benjamins, Amsterdam
- Baker, M. (1988) *Incorporation: A Theory of Grammatical Function Changing*. University of Chicago Press, Chicago.
- Barbiers, S. (1995) *The Syntax of Interpretation*. PhD Dissertation, University of Leyden, Leiden.
- Besten, H. den (1977), "On the Interaction of Root Transformations and Lexical Deletive Verbs." Ms. University of Amsterdam (also published in Den Besten (1989)).
- Besten, H. den (1989) *Studies in West Germanic Syntax*. PhD Dissertation, Tilburg University, Tilburg.
- Chomsky, N. (1957) *Syntactic Structures*. Mouton, The Hague.
- Chomsky, N. (1965) *Aspects of the Theory of Syntax*. MIT Press, Cambridge, Mass.
- Chomsky, N. (1975) *Reflections on Language*. Pantheon Books, New York.
- Chomsky, N. (1977) *Essays on Form and Interpretation*. North-Holland, New York, etc.
- Chomsky, N. (1986a) *Barriers*. MIT Press, Cambridge, Mass.
- Chomsky, N. (1986b) *Knowledge of Language: Its Nature, Origin and Use*. Praeger, New York.
- Chomsky, N. (1995) *The Minimalist Program*. MIT Press, Cambridge, Mass.
- Cinque, G. (1998) *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford University Press, New York, etc.
- Corver, N. and H. van Riemsdijk (1994) Eds., *Studies on Scrambling: Movement and Non-Movement Approaches to Free Word Order*. Mouton/de Gruyter, Berlin.
- Dikken, M. den (1995) *Particles*. Oxford University Press, New York, etc.
- Dikken, M. den (1996) "The Minimal Links of Verb (Projection) Raising," in, W. Abraham, S. Epstein, H. Thráinsson, and J.-W. Zwart, Eds., *Minimal Ideas*. John Benjamins, Amsterdam.
- Emonds, J. (1976) *A Transformational Approach to English Syntax: Root, Structure-Preserving and Local Transformations*. Academic Press, New York.
- Greenberg, J. (1966) "Some Universals of Grammar with Particular Reference to The Order Meaningful Elements," in J. Greenberg, *Universals of Language*. MIT Press, Cambridge, Mass.
- Hale, K. and S.J. Keyser (1993) "On Argument Structure and Lexical Expression of Syntactic Relations," in K. Hale and S.J. Keyser, *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. MIT Press, Cambridge, Mass.
- Hoekstra, T. (1988) "Small Clause Results." *Lingua* 74, 101-139.
- Hoop, H. de (1992) *Case Configuration and Noun Phrase Interpretation*. PhD Dissertation University of Groningen, Groningen.
- Jackendoff, R. (1972) *Semantic Interpretation in Generative Grammar*. MIT Press, Cambridge, Mass.
- Kaan, E. (1992) *A Minimalist Approach to Extraposition*. MA Thesis University of Groningen, Groningen.
- Katz, J. and P. Postal (1964) *An Integrated Theory of Linguistic Descriptions*. MIT Press, Cambridge, Mass.
- Kayne, R. (1984) *Connectedness and Binary Branching*. Foris, Dordrecht.
- Kayne, R. (1994) *The Antisymmetry of Syntax*. MIT Press, Cambridge, Mass.



- Klima, E. (1964), "Negation in English," in J. Fodor and J. Katz, Eds., *The Structure of Language: Readings in The Philosophy of Language*. Prentice-Hall, Englewood Cliffs, NJ.
- Koopman, H and D. Sportiche (1991) "The Position of Subjects." *Lingua* 85, 211-258.
- Koopman, H. and A. Szabolcsi (1998) "Verbal Complexes." Ms., UCLA, Los Angeles.
- Koster, J. (1974) "Het Werkwoord als Spiegelcentrum." *Spekator* 3, 601-618. [transl.: "The Verb as Mirror Center"]
- Koster, J. (1975) "Dutch as an SOV Language." *Linguistic Analysis* 1, 111-136.
- Koster, J. (1988) "The Residual SOV Structure of English." Ms., University of Groningen, Groningen.
- Koster, J. (1993) "Structure-Preservingness and the End of Transformationalism, in A. de Boer, J.T. de Jong and R.J. Landeweerd, Eds., *Language and Cognition 3. Yearbook 1993 of the Research Group for Theoretical and Experimental Linguistics of the University of Groningen*, Groningen, 1993, 139-148.
- Koster, J. (1994) "Predicate Incorporation and the Word Order of Dutch," in G. Cinque, J. Koster, J.-Y. Pollock, L. Rizzi and R. Zanuttini, Eds., *Paths Towards Universal Grammar: Studies in Honor of Richard S. Kayne*. Washington DC.
- Koster, J. (1995) "Extraposition as Coordination." Presented at Colloquium Max-Planck-Institute and Von Humboldt University, Berlin.
- Koster, J. (1998) "Gapping Moet Blijven." *TABU* 28-3, 99-106.
- Koster, J. (forthcoming) "Extraposition as Parallel Construal". University of Groningen, Groningen.
- Kuno, S. (1973) *The Structure of The Japanese Language*. MIT Press, Cambridge, Mass.
- Larson, R. (1988) "On the Double Object Construction." *Linguistic Inquiry* 19, 335-391.
- May, R. (1977) *The Grammar of Quantification*. PhD Dissertation MIT, Cambridge, Mass.
- Neeleman, A. (1994) *Complex Predicates*. PhD Dissertation Utrecht University, Utrecht.
- Noordhof, H. (1937) *La construction relative en italien*. PhD Dissertation University of Leiden. The Hague.
- Pesetsky, D. (1994). *Zero Syntax*. MIT Press, Cambridge, Mass.
- Perlmutter, D., (1983) Ed., *Studies in Relational Grammar 1*. University of Chicago Press, Chicago.
- Pollock, J.-Y. (1989) "Verb Movement, Universal Grammar and The Structure of the IP." *Linguistic Inquiry* 20, 365-424.
- Riemsdijk, H. van (1978) *A Case Study in Syntactic Markedness*. Foris, Dordrecht.
- Riemsdijk, H. van (1984) "On Pied-Piped Infinitives in German Relative Clauses," in J. Toman, Ed., *Studies in German Grammar*. Foris, Dordrecht.
- Riemsdijk, H. van (1994) "Another Note on Clausal Pied-Piping," in in G. Cinque, J. Koster, J.-Y. Pollock, L. Rizzi and R. Zanuttini, Eds., *Paths Towards Universal Grammar: Studies in Honor of Richard S. Kayne*. Washington DC.
- Rijkhoek, P. (1998) *On Degree Phrases and Result Clauses*. PhD Dissertation, University of Groningen, Groningen.
- Ross, J. (1967) Constraints on Variables in Syntax. PhD Dissertation MIT, Cambridge, Mass.
- Ross (1969) "Adjectives as Noun Phrases," in D. Reibel and S. Schane, Eds., *Modern Studies in English: Readings in Transformational Grammar*. Prentice-Hall, Englewood Cliffs, NJ.
- Sportiche, D. (1994) "Adjuncts and Adjunction." *GLOW Newsletter* 32, 54-55.

- Stowell, T. (1981) *Origins of Phrase Structure*. PhD Dissertation MIT, Cambridge, Mass.
- Thiersch, C. (1978) *Topics in German Syntax*. PhD Dissertation MIT, Cambridge, Mass.
- Vanden Wyngaerd, G. (1989) "Object Shift as an A-movement Rule." *MIT Working Papers in Linguistics* 11, 256-271.
- Vries, M. de (1999) "Extrapolation of Relative Clauses as Specifying Coordination", Ms. University of Amsterdam (to appear in proceedings of Console 7 (1998)).
- Zwart, J.-W. (1993) *Dutch Syntax: A Minimalist Approach*. PhD Dissertation University Groningen, Groningen.
- Zwart, J.-W. (1994) "Dutch is Head Initial". *The Linguistic Review* 11, 377-406.
- Zwart, J.-W. (1997) *Morphosyntax of Verb Movement: A Minimalist Approach to The Syntax of Dutch*. Kluwer, Dordrecht, etc.

*Jan Koster*  
*General Linguistics*  
*University of Groningen*  
*NL-9700 AS Groningen*  
*The Netherlands*

koster@let.rug.nl