

## THE POSITION OF THE HEAD AND THE DOMAIN OF SCRAMBLING

Norbert Corver & Henk van Riemsdijk

Grammatical Models Group / CLS / Tilburg University  
CORVER@KUB.NL                      RIEMSDYK@KUB.NL

*Paper presented to the LP'96 Conference, Prague, August 20-22 1996*

### 1. Introduction<sup>1</sup>

The purpose of this article is to present some initial findings of a survey study on scrambling, i.e. the phenomenon of variable word order within a clause. The starting point for this research has been the long held intuition that (the most 'neutral' subtypes of) scrambling (a) do not cross heads and (b) are frequent in head-final languages (SOV) and rare or even absent in head-initial languages (SVO). But any generalization along these lines have never been firmly established, let alone derived. Our program is to investigate the validity of a number of generalizations in this domain and to speculate (at best) about ways to derive them within current syntactic theory.

We will proceed as follows. In section 2 we will discuss some properties of scrambling in Dutch. From this single language investigation we will draw in section 3 some initial tentative conclusions as regards the relation between the position of the head and the domain of scrambling. In section 4 we will check these initial generalizations extrapolated from Dutch against a language sample of 23 languages. After having presented our cross-linguistic data, we will discuss to what extent our initial conclusions as regards the relation between scrambling and headedness derived from Dutch are corroborated or rejected. We will end this paper (section 5) with some speculations on the way to derive the descriptive generalizations from syntactic theory.

Before turning to the empirical domain of this study, i.e. the phenomenon of scrambling, we would like to spend some words on the use and usefulness of survey studies in generative typological research. As is well-known, this research methodology, which is characteristic of the Greenbergian typological research tradition, is not very common in

---

<sup>1</sup> The present material is based on a pilot investigation carried out in the fall of 1995. We would here like to collectively thank our many informants, who are listed by name in the tables in section 4. The first results were presented at the workshop on Final Heads, held at Tilburg University in November 1995. In the spring of 1996 further presentations were given at the University of Trondheim, to the Kansai Association of Theoretical Linguistics at Kwansei Gakuin University, at Tokyo Metropolitan University, and at Tohoku University in Sendai. Thanks are due to our hosts there, Lars Hellan, Masaru Honda, Heizo Nakajima, and Sige-Yuki Kuroda, and to the audiences in all these places for stimulating discussion.

the Chomskyan generative research paradigm. Generative studies generally focus on single languages or very limited numbers of languages. Although there are good reasons for adopting this single language research strategy (see below), we hold the opinion that the Greenbergian survey methodology needs reappraisal within the generative framework. As a matter of fact, there is no intrinsic incompatibility between generative linguistics and the survey-methodology. Why then is this research methodology so unusual in generative linguistics? In what follows, we will give a brief historical sketch of the generative approach towards Universal Grammar and try to make clear why initially the emphasis was on single language studies rather than comparative typological studies, and why we think the survey-method can now be useful to generative research.

In trying to define the notion "possible (grammar of a) human language", the theory of Universal Grammar must meet two conditions. On the one hand, (the structure of) UG must be restrictive enough to account for the fact that the child is able to acquire the grammar of a natural language in a fairly short time and on the basis of fairly limited evidence. At the same time, UG should not be too restrictive in view of the attested diversity among natural languages. Within the framework of generative grammar, this tension between cross-linguistic uniformity and cross-linguistic diversity has led to a conception of UG as a system specifying absolute grammatical properties and parametrized grammatical properties (the so-called principle-and-parameters model of Universal Grammar). The former properties define the common core of all natural language grammars (e.g. the structure dependency of grammatical rules, the c-command requirement on displacement); the latter determine the (limited) range of variation within which grammars may vary from one language to another by defining a restricted set of values that a grammatical property may take (e.g. [+head initial] or [-head-initial] for the so-called headedness parameter).

Although it is obvious that the questions of cross-linguistic uniformity and cross-linguistic diversity are equally important for the proper definition of the structure of UG, it is quite clear from a retrospective view that initially generative linguistics has emphasized the question of uniformity. The question of linguistic diversity and its consequences for the theory of UG became a serious topic of research only around the mid-seventies.

What was the reason for this relatively late development of comparative research within generative linguistics? The answer to this question relates to a methodological choice within the generative research program. The basic approach to the construction of UG was -- and, to a large extent, still is -- the in-depth study of (large fragments of) particular grammars of individual languages.

A direct consequence of this approach to the construction of UG is that the cross-linguistic study of a grammatical pattern X is possible only if in-depth analyses are

available of this pattern for the languages that are compared. Thus, although the question of diversity (i.e. the typological question) is intrinsically related to the question "what is a possible human language?" ("the question about UG"), the former question was hardly addressed in early generative studies because of the lack of detailed generative descriptions of languages other than English.

As a result of this methodological choice and because of the fact that initially most practitioners of early generative grammar were natives of English, focussing through introspection on attested and non-attested patterns of their native language, the empirical basis for systematic cross-linguistic study was absent in early generative linguistics.<sup>2</sup> Consequently, the question of uniformity was the major research question as regards the construction of a theory of UG.

On the basis of single-language studies, one tried to extract abstract grammatical properties which were postulated as universal properties of natural language grammars. This line of research led to precise characterizations of intra-linguistic cross-constructional similarities and the formulation of general universal rule schemas and general linguistic principles governing these rule schemas. To mention two examples: Jackendoff's (1977) elaborate study of the internal make-up of English phrasal structure led to the formulation of some general cross-categorial requirements on phrasal structure; Chomsky's (1977) examination of the presence of locality (i.e. island) effects in different types of English constructions (e.g. clefts, relatives, comparative) led to the formulation of a cross-constructional, putatively universal displacement schema, viz. Move WH.

Around the mid-seventies, the largely mono-linguistic nature of generative research had disappeared. A large number of generative studies appeared on different languages (e.g. French, Italian, Dutch, German, the Scandinavian languages). These studies either took the single-language view and presented detailed analyses of a phenomenon of a particular language (e.g. the Verb second effect in Dutch (Koster (1975)), causative formation in Italian (Rizzi (1978)), cliticization in French (Kayne (1975)) or took the cross-linguistic view and compared some grammatical pattern with its English equivalent. Two exemplifying studies of the latter sort are Emonds's (1978) comparative study of the distribution of verbs in French and English and Rizzi's (1978) comparative study of locality (i.e. Subjacency) effects on displacement phenomena in English and Italian. The latter study was an important landmark in generative research for two reasons: First, it showed that the absence of a grammatical pattern in language A (i.e. extraction across wh-islands in English) and its presence in language B (i.e. Italian) should not immediately

---

<sup>2</sup> For some early generative studies of languages other than English: Cf. Kuroda (1965), Bordelois (1974), De Rijk (1972).

lead to a rejection of a universal constraint (viz. Subjacency) formulated on the basis of a single-language study. Rather, and secondly, it showed that the locality constraint ("Don't cross two bounding nodes") is constant across languages and that the cross-linguistic differences must be attributed to a different selection of the set of bounding nodes (more specifically, "S" for English and "S'" for Italian).

As large portions of the grammar of various languages became available, the crosslinguistic study of grammatical structures became a more serious topic of research within the generative program. This shift from a single-language approach to the construction of UG towards a (single and) cross-linguistic approach to the construction of UG culminated in the so-called Principles-and-Parameters approach towards universal grammar (Chomsky 1981). This approach tries to answer the question of diversity in terms of a theory of parameters (i.e. options in the grammatical system whose (limited) values are specified by UG and fixed by the child on the basis of linguistic environmental input) and faces questions like what are the options within the UG-system?, are there any limits on the number of values a parameter can have?, in what components of the grammar are they to be found?, etc.

Although the question of diversity has obtained a more prominent position within the generative research program, there is a tendency to restrict the empirical basis for typological research to a small number of languages. Furthermore, these languages are often related to each other in that they belong to the same (sub)family. Some indication of this is given by the following arbitrary sample of comparative generative studies.

A. **cross-dialectal comparison**

- . Kayne (1989): cross-dialectal study of patterns of past participle agreement in Romance.

B. **cross-linguistic comparison within a language-family**

- . Emonds (1978), Pollock (1989): comparative study of the placement of verbs in English and French.
- . Holmberg (1986): a comparative study of word order phenomena in Scandinavian languages.

C. **cross-linguistic comparison across language-families**

- . Rudin (1988): a comparative study of the patterning of clause-initial wh-phrases in Slavic and Balkan languages.
- . Koopman (1984): a comparative study of verb placement in Dutch, Vata and Gbadi.

Thus, the set of languages compared is generally rather small; often for good reasons, e.g. the complexity of the grammatical data that are compared.

The generative approach towards the cross-linguistic study of linguistic patterns differs crucially from the one adopted in the Greenbergian typological research paradigm. This approach tries to determine the bounds of linguistic diversity on the basis of survey studies of large, balanced (i.e. areally and genetically distributed) samples of languages.

It should be clear that there is no intrinsic incompatibility between generative linguistics and the survey-method. The fact that large survey studies are scarcely found within earlier stages of generative research relates to the non-availability of native linguist informants and in-depth analyses of a great variety of languages.

Fortunately, the situation has much improved during the last fifteen years. The amount of languages described within the generative framework has increased enormously. Both single-language studies and small scale cross-linguistic studies have yielded large amounts of information about the grammar of languages such as Basque, Turkish, Hindi, Japanese, Chinese, Warlpiri, Hungarian, Quecha, Hebrew, Rumanian, etc.

Given this improvement as regards the availability of grammatical resources, there does not seem to be any barrier to using the survey method in generative research. In fact, there is much to gain from such an approach. First and foremost, the more languages are studied and compared, the stronger the generalizations with regard to the universal constraints and the more refined a picture is obtained of the bounds on cross-linguistic variation. The survey method further permits a more rapid synthesis of the distribution of some grammatical property and provides useful hints as regards the correctness of intuitive descriptive generalizations extrapolated from the study of one or a limited number of languages.

## **2. Some properties of Scrambling in Dutch**

In this section we will establish a number of properties of scrambling as found in Dutch. Some of these properties will then form the basis for our extrapolation to a number of potential typological generalizations.

To begin with, let us adopt an operational definition of what we mean by scrambling. We take as point of departure the assumption, formalized in X-bar theory, that complements are closer to the head than specifiers and modifiers (adjuncts). What this means is that all structures in which some complement is separated from its head by some modifying element must be derived. Standard cases are found with the most well-known movement processes such as wh-movement and extraposition. But there is a residue of such cases in which the landing site is somewhere within what has been called *Mittelfeld* ('middle field') in the grammatical tradition of the Continental West Germanic languages.

A remark on the notion of ‘movement’ is in order here. We use this term in a pretheoretical fashion to identify structures that deviate from the basic order imposed by X-bar theory as stated above. We do not mean to imply any kind of preconception as to whether these structures should actually be derived by means of a movement operation, or whether they are base-generated as such, where the latter option would imply a reformulation of the relevant notions of X-bar theory. For expository purposes, we nevertheless use the terminology of movement as well as the notation of movement in the examples: italics for the displaced constituent and a trace ([e]) for the canonical position in which we would expect it to occur.

The properties to be established are the following:

- A. There is a contrast between neutral scrambling (N-SCR) and contrastive scrambling (C-SCR);
- B. Scrambling is only possible in head final structures, but not in head initial structures;
- C. Long Scrambling (L-SCR) is only possible out of AP and PP;
- D. Only the verbal domain can be the landing site for L-SCR.<sup>3</sup>

### 2.1. N-SCR must be distinguished from C-SCR

In English, scrambling does not exist, but embedded topicalization does:

- (1) a. \*John will *the car* probably return [e] tomorrow
- b. They say that *the tacos* Judge Bean won’t go for [e] (Ross (1973))

Sentences like b are only acceptable provided the displaced NP receives special intonational prominence and is interpreted in a contrastive way.

In Dutch, both types coexist. The neutral type is characterized by perfectly normal intonation contours, while the contrastive type has an intonation pattern similar to that found in (1)b. The former type, N-SCR, is used for D-linked (discourse-linked) definite NPs (cf. 6) and specific indefinite NPs (cf. 6), while the latter, C-SCR, is typically found with focus particles (cf. 6).

---

<sup>3</sup> This statement may have to be modified depending on the treatment of constructions in which scrambling (or some scrambling like process) is limited to a class of designated elements rather than to the whole class of D-linked noun phrases. In Dutch, for example, L-SCR appears to be possible into the AP-domain, but only if the ‘scrambled’ element is an R-pronoun. In the present article, this complication will be left aside.

- (2) Jan, ik ben *de etter* gelukkig nooit meer [e] tegengekomen (N-SCR)  
John, I have the jerk fortunately never again met
- (3) Je kunt drie van die aanbevelingsbrieven maar beter wat [e] afzwakken  
You had three of those letters of recommendation better somewhat weaken
- (4) Annie Hall, ik heb *OOK DIE FILM* eigenlijk nooit [e] grappig gevonden  
Annie Hall, I have also that movie in fact never amusing considered (C-SCR)

For more discussion of this distinction as well as a proposal for the differential treatment of N-SCR vs. C-SCR, see Neeleman (1994). In the remainder of this article, we will exclusively be concerned with N-SCR.

## 2.2. N-SCR is limited to head final contexts.

Dutch is a mixed-headed language. In short, the underlying structure in the verbal domain is head final, i.e. COMPLEMENT-V, while the fundamental order in the nominal domain is head initial, i.e. N-COMPLEMENT. For prepositions and adjectives, the situation is even more complex: both are sometimes initial and sometimes final.

There are two basic arguments to be derived from this mixed headedness in connection with scrambling. First, Dutch has scrambling in the verbal domain but not in the nominal domain. Second, Dutch has scrambling in head final APs and PPs, but not in head initial ones. This is illustrated in the following examples.

- (5) Ik heb *jouw spullen* gisteren eindelijk [e] weg gegooid  
I have your things yesterday finally away thrown
- (6) [Die twee oude argumenten *tegen haar theorie*] heeft zij onlangs ontkracht  
those two old arguments against her theory has she recently refuted
- (7) a. \*[Die twee oude *tegen haar theorie* argumenten [e] ]  
b. \*[Die twee *tegen haar theorie* oude argumenten [e] ]  
c. \*[Die *tegen haar theorie* twee oude argumenten [e] ]  
d. \*[*Tegen haar theorie* die twee oude argumenten [e] ]

(5) shows a standard case of scrambling in the verbal domain, while (7) demonstrates that scrambling (of a PP this time) within the nominal domain is impossible.

Consider now the case of adjectives:

- (8) a. [Tamelijk zwaar verslaafd *aan vodka*] is hij pas sinds kort  
rather heavily addicted to vodka is he only recently
- b. [Tamelijk zwaar *aan vodka* verslaafd] is hij pas sinds kort
- c. [*Aan vodka* tamelijk zwaar [e] verslaafd] is hij pas sinds kort

As (8)a/b shows, a PP-complement to an adjective may occur to the left or to the right of the adjective. (8)c shows that scrambling is possible. Unfortunately, we cannot know whether the origin of the scrambled PP is to the left or to the right of the adjective. But the facts are fully consistent with the claim that scrambling is limited to head final contexts. This is further confirmed by the fact that adjectives with noun phrase complements, which can only occur preadjectivally, also allow scrambling:

- (9) a. ?[Geheel *het Frans* machtig] was ik pas op mijn 15e  
completely the French in-command-of was I only in my 15th (year)
- b. [*Het Frans* geheel [e] machtig] was ik pas op mijn 15e

The situation with PPs is more straightforward in that only a small number of prepositions can occur postpositionally, in particular a subset of directional prepositions. And it is precisely these that allow scrambling, as shown in (11) as opposed to (10).

- (10) a. [5 uur na *de eerste ontploffing*] volgde de tweede  
5 hours after the first explosion followed the second
- b. \*[5 uur *de eerste ontploffing* na] volgde de tweede
- c. \*[*De eerste ontploffing* 5 uur na [e] ] volgde de tweede
- (11) a. [Nog een stukje verder *de kamer* in] met die spullen!  
Even a bit further the room into with those things
- b. [*De kamer* nog een stukje verder [e] in] met die spullen!

Thus we consider our claim B, that scrambling only occurs in head final structures in Dutch, established.

### 2.3. L-SCR is limited to extraction out of AP and PP

We distinguish long scrambling (L-SCR) from short scrambling (S-SCR). By S-SCR we mean those types of scrambling in which the scrambled element remains within the

domain of its head, whereas by L-SCR we mean cases in which the scrambled element appears outside the domain of its head (or: is extracted from that domain). To be more precise, by ‘domain of the head’ we mean the extended projection of that head in the sense of Grimshaw (1991) and Van Riemsdijk (1990, in prep.). This means that the domain of the verb is IP or CP (i.e. the clause with the complementizer if there is one, without it if there isn’t one). Similarly, the domain of the N is the determiner phrase, DP. And in the case of AP and PP as well, we mean the projection of the lexical adjectival or prepositional head TOGETHER WITH ALL THEIR FUNCTIONAL HEADS.

Since scrambling within the nominal domain is impossible, as shown above, long scrambling out of the nominal domain is excluded *a fortiori*.

- (12) \*Zij heeft *tegen haar theorie* onlangs [die twee oude argumenten [e]] ontkracht  
 she has against her theory recently those two old arguments refuted (cf. (6))

L-SCR out of the verbal domain (i.e. out of the clause) is excluded as well:<sup>4</sup>

- (13) \*Ik geloof dat Martinet *Hjelmslev* altijd zei [dat je beter niet [e] kon lezen]  
 I believe that M. H. always said that you better not should read

With APs, however, extraction out of the AP is possible.

- (14) Jan is *aan dat drankje* toendertijd [tamelijk zwaar [e] verslaafd] geweest  
 Jan has to that drink at-the-time rather heavily addicted been (cf. (8))

Similarly, long scrambling out of a postpositional PP is possible (cf. Van Riemsdijk (1978)).

- (15) Je zou die spullen *de kamer* misschien [iets verder [e] in] moeten brengen  
 you should those things the room perhaps a bit further into have-to bring

We conclude that long scrambling is possible out of (extended) AP or PP, but not out of the extended nominal or verbal domain.

#### 2.4. Only the verbal domain can be the landing site for L-SCR

---

<sup>4</sup> The situation with verb raising complements (a clause union construction) is more complicated. Roughly speaking, neutral long scrambling is excluded here as well, as shown most clearly in causative and perception verb complements. Space precludes more extensive discussion.

As a final property to be established, consider the landing site of the scrambled constituent. In the case of short scrambling, this is trivial: the scrambled constituent remains a dependent of its host projection.<sup>5</sup> But in the case of L-SCR this is an important, though generally overlooked, issue. After all, the constituent out of which a scrambled constituent is extracted could be contained in a noun phrase, a prepositional phrase, etc. In the examples of L-SCR above, the landing site is the verbal domain, the VP that contains the AP or PP out of which extraction has occurred. The following examples show that L-SCR into the nominal domain is not possible.

(16) \*Dit is al [de *aan vodka* derde [tamelijk zwaar [e] verslaafde] jonge man]  
 this is already the to vodka third rather heavily addicted young man

(17) \*[De *het oerwoud* eerste tocht [30 mijl [e] in]] was de moeilijkste  
 the the rain forest first trip 30 miles into was the most-difficult (one)

Examples can also be constructed showing that AP and PP cannot be the landing site for L-SCR, but the demonstration must be deferred to a longer exposé. As we will show below, however, we believe that the fact established for Dutch, viz. that L-SCR is limited to verbal landing sites, may well generalize to many if not all languages.

### 3. Some tentative typological generalizations

In this section, we will extrapolate some tentative cross-linguistic generalizations, directly inspired on the analysis of the rather complex situation found in Dutch.

<b>GENERALIZATION A</b>	<b>Scrambling is limited to head-final contexts</b>
-------------------------	---

This is the impression that gave rise to our investigation in the first place: the typical scrambling languages are head final languages, while the typical head initial languages lack scrambling. Thus, scrambling is found in German and Dutch, but not in English and

---

<sup>5</sup> Technically, it could be adjoined to some intermediate (non-head, non-maximal) node on the main projection line, or it could land in the specifier of some functional head, depending on the exact version of X-bar theory that is adopted. This issue is immaterial in the present context.

French; it is also found in Hindi and Japanese, but not in Arabic or Welsh. The illustration of this correlation can only be a list of suggestive examples.<sup>6</sup>

(18) Ich habe *dieses Buch* gestern glücklicherweise am Flohmarkt gefunden (G)  
I have this book yesterday fortunately at the fleamarket found

(19) \*J'ai *ce livre* heureusement trouvé au marché aux puces hier (F)

(20) a. raam-ne mohan-ko phal de diye (Hindi)  
Ram-SUB Mohan-IO fruits gave  
b. raam-ne *phal* mohan-ko de diye  
c. *phal* raam-ne mohan-ko de diye

(21) a. qara'a muhammad-un *kitaab-an* (Arabic)  
read-PERF-3sg/masc M.-nom book-INDEF-ACC (VSO)  
b. qara'a *kitaab-an* muhammad-un (VOS)  
c. *muhammad-un* qara'a kitaab-an (SVO)  
d. *muhammad-un kitaab-an* qara'a (SOV)  
e. 'al-kitaab-u muhammad-un qara'a-hu (OSV)  
f. *kitaab-an* qara'a muhammad-un (OVS)

The latter examples suggest that scrambling is possible in Arabic. However, it is reported that only the VSO order is unmarked, whereas the other orders contrastively pick out specific discourse topics.

The following examples show that German, like Dutch, limits scrambling in the AP and PP to head final contexts.

(22) a. ??<sub>AP</sub>[*Ganz des Französischen unmächtig*] scheint er nicht zu sein (G)  
entirely of French not-in-command-of seems he not to be  
b. <sub>AP</sub>[*Des Französischen ganz unmächtig*] scheint er nicht zu sein

(23) a. <sub>PP</sub>[*Noch weiter das Dach hinauf*] wird die Schnecke nicht kriechen

---

<sup>6</sup> Baltin (1982) suggests that French does have a limited range of scrambling possibilities, but states himself that this is true only for those speakers of French that allow topicalization, which is known to be largely absent in Standard French, though not, perhaps, in Québécois, the language of one of his two informants. At any rate, the link with topicalization suggests that the examples in question are instances of contrastive rather than neutral scrambling.

- b. even further the roof upward will the snail not creep  
 [pp] *Das Dach noch weiter hinauf*] wird die Schnecke nicht kriechen

<b>GENERALIZATION B</b>	<b>If a language has L-SCR it also has S-SCR</b>
-------------------------	--

We will be brief on this generalization. It follows entirely from the theory of movement processes. For example, languages that have only long distance *wh*-movement and no short *wh*-movement have never been found. The theory, essentially the theory of bounding and successive cyclic movement and its more recent instantiations, predicts this. *Mutatis mutandis*, the same prediction holds for scrambling. If scrambling is not an instance of movement but rather a result of base-generation, the relevant module of the grammar will not be the bounding module but the module regulating the domain within which theta-roles may be deployed by a head. If upward domain extension is included in this module, as it must if it is responsible for scrambling, then the same prediction will ensue.

<b>GENERALIZATION C</b>	<b>L-Scrambling is limited to extraction from AP, PP</b>
-------------------------	--

The following examples show that German, like Dutch, allows L-SCR out of (head final) AP and PP.<sup>7</sup>

- (24) Er ist *des Französischen* nach all diesen Jahren trotzdem immer noch  
 he is of French after all these years nevertheless still

[ [e] unmächtig] geblieben (G)  
 incapable remained

- (25) Die Schnecke wird *das Dach* doch nicht [noch weiter [e] hinauf] kriechen?  
 the snail will the roof presumably not even further upward creep

Turning now to long scrambling out of the verbal (clausal) domain, there is some controversy. Saito (1992) has claimed that long distance scrambling exists in Japanese. However, Ueyama (1994) argues that the relevant examples such as the one cited here are only permissible on a focus reading of the scrambled constituent.

---

<sup>7</sup> By using the term ‘head final’ here we do not wish to imply that AP and PP are taken to be underlyingly head final. What is meant is that in the construction in question, the derived position of the head is final. See below for more extensive discussion.

- (26) *Sono hon-o* John-ga [Mary-ga [e] katta to] omotteiru towa (Jap.)  
 that book-ACC John-NOM Mary-NOM bought COMP think COMP  
 “That book John thinks that Mary bought”

Similarly, Browning & Karimi (1994) have argued that long scrambling in Persian is possible only if the scrambled constituent is contrastive or focussed.

- (27) *Ali ketab-ash-ra* fekr-nemikard ke man be Hasan goftan ke tu bardashti [e]  
 Ali book-his-OM not-think-3sg that I to Hasan say-1sg that you take-2sg  
 “Ali doesn’t think that I told Hasan that you took his book”

**GENERALIZATION D Only the verbal domain can be the target of L-SCR**

There is interesting evidence from Japanese supporting this generalization.<sup>8</sup> First, notice that long scrambling out of AP is possible.

- (28) a. John-ga [sake-ni tsuyoku] natta (Jap.)  
 John-NOM liquor-DAT strong became  
 “John became a heavy drinker”  
 b. *Sake-ni* John-ga [ [e] tsuyoku] natta

Now, consider the following examples.

- (29) a. Taroo-ga [ sono [totemo Sendai-kara tooi] machi]-e itta  
 Taroo-NOM this very Sendai-from far city-ALL went  
 “Taroo went to this city far from Sendai”  
 b. Taroo-ga [ sono [*Sendai-kara* totemo [e] tooi] machi]-e itta  
 c. Taroo-ga [ [*totemo Sendai-kara tooi*] sono [e] machi]-e itta  
 d. Taroo-ga [ [*Sendai-kara* totemo [e] tooi] sono [e] machi]-e itta  
 e. \*Taroo-ga [ [*Sendai-kara*] sono [ [e] totemo tooi] machi]-e itta  
 f. \* [*Totemo Sendai-kara tooi*] Taroo-ga [ [e] sono machi]-e itta  
 g. \* [*Sendai-kara*] Taroo-ga [ sono [ [e] totemo tooi] machi]-e itta  
 h. \* [*Sendai-kara*] Taroo-ga [ [ [e] totemo tooi] sono [e] machi]-e itta

---

<sup>8</sup> Thanks are due to Masayuki Oishi for constructing the argument and to Sige-Yuki Kuroda and Heizo Nakajima for clarifying some issues concerning the examples.

In these examples, we have a noun phrase consisting of a (demonstrative) determiner, followed by a complex AP, followed by the noun (and its allative case marker). The AP consists of a modifier, followed by a nominal (or: PP) complement, followed by the adjective itself. (29)b shows that short scrambling inside the AP is possible, while (29)c shows that the AP as a whole can undergo short scrambling within the noun phrase. In (29)d both short movements are combined, yielding a grammatical result as well. The remaining examples show that long scrambling is excluded here. In (29)g/h the complement of the adjective is scrambled all the way out of the noun phrase. This is excluded since, by hypothesis, scrambling out of the nominal domain is barred (cf. Generalization C above). Similarly, in (29)f, the whole AP is scrambled out of the noun phrase, which is prohibited for the same reason. But consider now (29)e. Here, the complement of the AP is extracted from the AP and lands to the left of the demonstrative determiner as a dependent of the nominal projection. This is excluded as well, supporting Generalization D.

To avoid terminological confusion, let us now define the terms ‘hybrid’ and ‘mixed’ to suit our purposes. We will reserve the term ‘hybrid’ for categories in which the lexical head may be either initial or final within one and the same language. This is what we have found to be the case for Dutch AP and PP. The term ‘mixed’ we will use exclusively for the situation, also found in Dutch, in which certain lexical heads are systematically initial (such as the Dutch noun phrase) while others are systematically final (such as the Dutch verb phrase). We return to the issue of mixed headedness under Generalization F. Here, we turn to hybrid headedness first.

<b>GENERALIZATION E</b>	<b>Only AP and PP can be hybrid</b>
-------------------------	-------------------------------------

Pursuing our strategy of extrapolating from the situation in Dutch (and in most cases German) to other languages, we hypothesize that hybrid headedness is limited to AP and PP. At this point, we have little data to support this claim beyond those reported on in section 4 below. It does appear, however, that hybrid headedness along with dual headedness (as in circumpositional PPs) occurs rather frequently, at least in PPs. Circumpositions are also found in Hungarian (cf. Van Riemsdijk (1990)), while hybridly headed PPs also occur in Kurd as well as in several other languages of the area. At this point we will not pursue the matter any further.

It is useful, however, to clarify the notion of hybrid headedness from an X-bar theoretic point of view. Logically, there are two basic ways in which hybrid headedness could be interpreted: either the complements (sometimes) move around the head, or the head moves across the complements. The former situation is also found, of course, in cases where fixed headedness is assumed. The VP in Dutch, for example, is taken to be

uniformly head final (modulo Verb Second in main clauses), but nevertheless PPs and clauses may be extraposed to a position behind the verb. The situation with AP and PP appears to be fundamentally different, however. Corver (1994,1996) argues that in the AP the lexical head A originates in the initial position and is moved rightward into the final position. Similarly, Van Riemsdijk (1990) has argued that P originates in the prepositional slot and, in specific cases, moves rightward to become a postposition. In both cases, it is argued that the slot that the head moves to is that of a functional head. In view of their relevance to the present discussion, we will present a very succinct summary of the arguments.

The case of PP

In addition to prepositional and postpositional PPs, German also has circumpositional PPs, as in (30). Note that both headlike elements, the initial one and the final one, are adpositional in nature.

- (30) [<sub>PP</sub> eine ganze            Strecke *in* den Wald *hinein*]    (G)  
       a    considerable distance in the wood (into)

Note first that positing a single headed structure with movement of the complement of the adposition will never yield a viable analysis of this case. Note furthermore that it is not surprising that the second head-like element is also adpositional in nature, if we assume that it is a functional head in the extended PP-projection. This categorial identity of the lexical and the functional head is also found in the nominal projection, where determiners, quantifiers, etc. are essentially nominal in nature, and in the verbal projection, where inflection, tense, negation, aspect etc. show up as (defective) verbs when independently lexicalized. In Van Riemsdijk (1990, forthcoming) this observation gave rise to the Categorial Identity Thesis:

- (31) The **Categorial Identity Thesis**: in the unmarked case the lexical head and the corresponding functional head have the same categorial features

The argument, developed in detail in Van Riemsdijk (1990), runs as follows. There are, roughly speaking, four possible analyses.

- (32) Possible analyses:    a.    [<sub>PP</sub> P PP ]  
                                   b.    [<sub>PP</sub> PP P ]  
                                   c.    [<sub>PP</sub> fp° NP P° ]        (fp = functional P).  
                                   d.    [<sub>PP</sub> P° NP fp° ]

(32)a is excluded by four considerations:

A. *P1 determines case, not P2*

- (33) unter der Brücke durch  
under the<sub>NOM</sub> bridge through

*Unter* governs the dative (when it is locative rather than directional), while *durch* always governs the accusative. Assuming, as seems reasonable, that it is the lexical head which determines case, *unter* must be the lexical head governing the noun phrase.

B. *Non-uniform headedness: PrepPs and PostPs*

This solution presupposes that PPs are sometimes prepositional, as in the case of the top PP, and sometimes postpositional, as in the case of the embedded PP. This creates a problem rather than solving one, and it is to be rejected on general grounds.

C. *Distribution of Modifiers: \*P M PP*

On this solution, we would expect the embedded PP to be independently modifiable, which it is not:

- (34) a. hoch auf dem Berg oben      b. \*auf hoch dem Berg oben  
high on the mountain up            on high the mountain up

D. *Extraction of da*

The variant of preposition stranding typical for German -- extraction of *da* with a doubling element *dr* on the preposition -- is blocked in true cases of PPs embedded in PPs, but it is permitted in circumpositional PPs.

- (35) \*Da konnte man den Ton bis drunter hören  
there could one the tone until under hear

- (36) Da hätte er nicht drauf hinauf steigen sollen  
there had he not thereup up climb should

Solution (32)b is excluded by a subset of the above considerations, viz. B and D. Solution (32)c is excluded by consideration A. Solution (32)d is not in conflict with any of the above considerations, hence it constitutes the correct analysis of circumpositional PPs in German. Purely postpositional structures are then straightforwardly analyzed as involving movement of the lexical P to the functional postpositional slot.

### The case of AP

The analysis of Dutch APs presented in Corver (1994, 1996) is essentially as in (37).

(37) [<sub>faP</sub> [<sub>A'</sub> MOD A° PP] fa°] (fa = functional A)

Recall that PP-complements generally occur either to left or to the right of the Adjective

- (38) a. [nauw verwant daaraan] (ADV A PP)  
closely related there-to
- b. [nauw daaraan verwant] (ADV PP A)
- c. [daaraan nauw verwant] (PP ADV A)
- d. ...dat we daaraan toendertijd [nauw verwant] waren  
that we there-to at-the-time closely related were  
(PP scrambled into verbal domain)

A first argument is based on the sensitivity of extraction out of PP to the so-called freezing effect. It is clear that the PP in (38)c/d must have been scrambled into its surface position, regardless of whether it originated pre- or postadjectivally. And as (39)c/d show, stranding of the P in these scrambled positions is blocked. This is what is called the freezing effect. Note now that if we were to assume that (38)b is also derived from (38)a by some form of scrambling, we would expect the freezing effect to be present in the b-example as well, but as (39)b shows, stranding is perfectly possible in this position.

- (39) a. Ik weet waar<sub>i</sub> we toendertijd [nauw verwant [t<sub>i</sub> aan]] waren  
I know where we at-the-time closely related to were
- b. Ik weet waar<sub>i</sub> we toendertijd [nauw [t<sub>i</sub> aan] verwant] waren
- c. \*Ik weet waar<sub>i</sub> we toendertijd [[t<sub>i</sub> aan] nauw verwant] waren
- d. \*Ik weet waar<sub>i</sub> we [t<sub>i</sub> aan]<sub>j</sub> toendertijd [t<sub>j</sub> nauw verwant] waren

The analysis in (37), under which the lexical adjectival head is moved rightward to the functional head position, thereby deriving the b-examples from the a-examples, correctly predicts the contrast between (39)a/b and (39)c/d.

A second argument can be derived from considerations having to do with Across-the-Board (ATB) extraction of a pronominal element from the PP-complement to the adjective. It is well-known that ATB-extraction is only possible if the positions from which extraction takes place are structurally parallel. One of the most significant aspects

of this parallelism is that extraction must either be from two right branches or from two left branches. Consider now the following example.

(40) Dat is het meisje waar<sub>i</sub> ik denk dat Jan  
that is the girl where I think that Jan

[goed [t<sub>i</sub> mee] bevriend] en [financieel afhankelijk [t<sub>i</sub> van]] is  
well with friendly and financially dependent on is

If we assume that the PP in preadjectival position is moved there by some form of scrambling from the postadjectival position, then the pre- and post-adjectival slots are on left and right branches respectively and no ATB-extraction is expected to be permitted. On the alternative hypothesis, however, the PP-complements occupy the same structural position within the conjuncts since only the adjectival head itself is moved. Hence ATB-extraction is permitted, as shown in (40).

On the basis of these two lines of argument, we consider the analysis of hybrid headedness in terms of movement of the lexical head to a functional head position on the opposite side of the projection well established. Note now that, quite apart from its intrinsic interest, this conclusion is also highly relevant for the main issue at hand. For implicit in the claim that (leftward) scrambling is found in head final structures but not in head initial ones is the claim that scrambling may not cross (lexically realized) heads. But such a crossing is exactly what would be required under the alternative analysis, now rejected, under which the right hand side complement of the A or P is scrambled leftward to a position to the left of that head.

Several other equally speculative but just as interesting generalizations come to mind. We will limit ourselves here to one of these with two possible corollary generalizations connected with it.

<b>GENERALIZATION F:</b> <b>Hybrid headedness and mixed headedness are correlated</b>
---

It may not be an accident that Dutch and German are both hybrid and mixed, while languages like English, French, Turkish and Japanese are neither hybrid nor mixed.<sup>9</sup> Some support for such a link can be found in Kurd and Persian. In these languages, the

---

<sup>9</sup> This may well be much too strong. In fact, several cases of potential hybrid headedness in head-final languages have been reported in our sample. Perhaps, if these indications are correct, it would still be possible to preserve the generalization as a one-way implication (mixed implies hybrid, but not the other way round).

verb is final, but the lexical head of noun phrases and adjective phrases is initial. PPs, however, appear to be mixed. Consider first the following examples, taken from Blau (1980) and Windfuhr (1979) respectively, showing that Kurd and Persian are verb final.

- (41) a. qesab gosht de-firosh-ê (Kurd)  
the-butcher meat sells
- b. azad nan de-kiř-ê (Kurd)  
Azad bread buys
- (42) a. az ardařêr bigurêxt (Persian)  
from Ardashir he-fled
- b. řâpur kamar az miyân biguřâd (Persian)  
Shapur the-belt from his-waist opened

The following examples show that the noun is initial within the noun phrase.

- (43) a. were ba sê chuwar shar-êk le shar-ekan-î kurdistan-î  
come!, three four city-INDEF in cities of-Kurdistan
- turkiya-t pê bi-nasên-im (Kurd)  
of Turkey with I-will-make-you-know  
“Come, I will acquaint you with three or four cities of Turkish Kurdistan”
- b. xe|k-î řojhe|at feqîr-e (Kurd)  
peoples-DEF Middle-East miserable  
“The peoples of the Middle East are miserable”
- (44) a. kot-e rang paride (Persian)  
coat-with faded color
- b. kořtan-e řir  
killing-(prep.) lion (= the killing of the lion (obj.))
- (45) a. (hasan) be tehrân raft (Persian)  
Hasan to Tehran went
- b. raftan-(e hasan) be tehrân (Persian)  
the-going (of Hasan) to Tehran

Note now that at least in Kurd there is evidence for hybrid headedness in that circumpositional PPs are found with a prepositional element accompanied within the same phrase by a postpositional element.

- (46) a. *be kam la-da de-ʒo-yn*                      *be ... da* ‘from, through, via’  
 Prep which side-Postp we-go  
 ‘Via which side do we go?’
- b. *lege/ kê-da de-ch-î bo piyase*                      *legel ... da* ‘with’  
 Prep who-Postp will-you-go for walk  
 ‘With whom will you go for a walk?’
- c. *le beyanî-we de-ch-im bo kar*                      *le ... we* ‘from’  
 Prep morning-Postp I-go for work  
 ‘I go to work from early morning’

One might go one step further and speculate about the distribution of headedness in mixed languages. Note, in fact that, by our analysis, Dutch and German have a split between V being final (*pace* Zwart (1993)) on the one hand, and N, A, and P being initial on the other hand. The very same split seems to occur in Kurd and Persian, and, it appears, also Tadjik and some of the Pamir languages<sup>10</sup> (assuming the considerations for an underlying prepositional structure carry over to these languages). So, we very tentatively posit the following implication.

**GENERALIZATION F’:**                      **In mixed languages, the split is between V and N/A/P**

Yet another corollary generalization is lurking here. For not only is the split the same in Dutch/German and Kurd/Persian, but in fact the direction is the same: in both cases it is the verb which is final and the other categories which are initial. If this is not an accident, then we might formulate the following generalization.

**GENERALIZATION F’’:                      V final with N/A/P initial exists  
 V initial with N/A/P final does not exist**

It is, of course, not very hard to come up with languages that appear to counter-exemplify these generalizations. For example, it appears that certain Kwa languages including Ewe and Fanti are V-initial but postpositional. However, the possibility cannot

---

<sup>10</sup> Thanks to Yakov Testelec for pointing out that the Pamir languages seem to share this characteristic.

be discounted that the underlying headedness in these languages is different from the most frequent surface word order. In other words, it may well be that these languages are misclassified just as Dutch and German were in Greenberg's original article. Similarly, Hungarian does appear to be underlyingly postpositional, although there is room for argument. In the nominal domain also, the head appears to be final, witness (47)a, although there are cases, described as marginal, in which the noun precedes a PP-dependent.

- (47) a. Péter Mari által való meg-hivasa (Hungarian)  
 P.'s Mary by ('being') (prev.-)invitation
- b. Péter meg-hivasa Mari által  
 P.'s invitation Mary by

Furthermore, the situation in the verbal domain is far from obvious. Depending on the specificity of the object, it either precedes or follows the verb, as illustrated in (48). There is considerable controversy about the underlying order of the verb (cf. É.Kiss (1981), Horváth (1985), so the matter must be left open here.

- (48) a. János levelet irt (OV with non-spec. objects) (Hungarian)  
 J. letter-ACC wrote
- b. János meg-irt-a a levelet (VO with spec. objects)  
 J. (prev.-)wrote-DEF the letter

#### 4. Results from the pilot study

Having extrapolated from Dutch some initial generalizations as regards the relation between the position of the head and the domain of scrambling, we will now check these generalizations against a language sample of 23 languages and see to what extent they are corroborated or rejected.

The data on scrambling were gathered through the use of questionnaires that are responded to by native linguist informants and (nonnative) experts on particular languages. It should further be noted that the sample for this study is rather small by Greenbergian standards (though quite large by generative standards) and not truly balanced. There are two reasons for this. First, this survey study on scrambling is intended as a pilot study meant (i) to provide initial hints towards the correctness of the intuited connection between scrambling and headedness and (ii) to gain experience in

using the survey method in generative typological research. Second, it is quite hard to get a `balanced' sample due to the `imbalanced' availability of native linguistic informants.

The results of this cross-linguistic study are summarized in the following tables:<sup>11</sup>

---

<sup>11</sup> "?" indicates that there is unclarity about the correctness of the YES/NO (e.g. because it is unclear whether the leftward displacement is a true instance of neutral scrambling).

<b>TABLE I</b> <b>V-INITIAL LANGUAGES</b>	HEAD COMPLEMENT ORDER			SCRAMBLING	LONG SCRAMBLING	
	P	N	A	IN VERBAL DOMAIN	OUT OF AP	OUT OF PP
Arabic (Huybregts)	P-X	N-X	A-X	NO	NO	NO
Chinese (Siebesma)	P-X	d.n.a.	X-A?	NO	NO	NO
Greek (Anagnostopoulou)	P-X	N-X	A-X	NO	NO	NO
Hebrew (Shlonsky)	P-X	N-X	A-X	NO	NO	NO
Icelandic (Sigurdsson)	P-X	N-X	A-X	NO	NO	NO
Norwegian (Hellan)	P-X	N-X	A-X	NO	NO	NO
Welsh (Roberts)	P-X	N-X	A-X	NO	NO	NO

<b>TABLE II</b> <b>V-FINAL LANGUAGES</b>	HEAD COMPLEMENT ORDER			SCRAMBLING	LONG SCRAMBLING	
	P	N	A	IN VERBAL DOMAIN	OUT OF AP	OUT OF PP
Basque (Laka)	X-P	X-N	d.n.a.	YES	d.n.a.	NO
Bengali (Bayer)	X-P	X-N	X-A	YES	YES	NO
Hindi (Mahajan)	X-P	X-N	X-A	YES	YES	NO
Japanese (Oishi)	X-P	X-N	X-A	YES	YES?	NO
Korean (Chung/Kim)	X-P	X-N	X-A	YES	YES/NO ?	NO
Turkish (Kural)	X-P	X-N	d.n.a.	YES	d.n.a.	NO

<b>TABLE III</b> <b>MIXED LANGUAGES</b>	HEAD COMPLEMENT ORDER			SCRAMBLING	LONG SCRAMBLING	
	P	N	A	IN VERBAL DOMAIN	OUT OF AP	OUT OF PP
Dutch (HvR/Corver)	P-X	N-X	A-X	YES	YES	SOME
Finnish (Vainikka)	P-X/X-P	N-X	A-X	NO	NO	NO
German (HvR)	P-X	N-X	A-X	YES	YES	SOME
Hungarian (Kiss / Kenesei)	X-P	X-N/N-X	X-A/A-X	YES	YES	YES
Persian (Karimi)	P-X	N-X	A-X	YES	NO	NO

<b>TABLE IV</b> <b>SLAVIC/BALKAN LANGUAGES</b>	HEAD COMPLEMENT ORDER			SCRAMBLING	LONG SCRAMBLING	
	P	N	A	IN VERBAL DOMAIN	OUT OF AP	OUT OF PP
Albanian (Kallulli)	P-X	N-X	?	YES	NO	NO
Bulgarian (V ulchanova)	P-X	N-X	A-X	?	NO	NO
Czech (Starke/Veselovska)	P-X	N-X	A-X	YES?	YES	NO
Macedonian (Tomic)	P-X	N-X	A-X	YES	YES	NO
Polish (Rozwadowska)	P-X	N-X	A-X	NO?	NO	NO
Russian (Trugman)	P-X	N-X	A-X	NO?	YES	NO

With the exception of the fourth table, representing the Slavic/Balkan languages, the subdivision of the languages is based on the headedness parameter. Table I presents data from head initial languages, table II from head final languages, and table III from mixed headed languages. Although the members of the Slavic and Balkan languages presented in table IV are generally considered to be head-initial (hence VO), we have not placed them in the table of V-initial languages for the reason that they appear to blur generalization A, i.e. that scrambling is limited to head final contexts. We will briefly come back to the data of the Balkan/Slavic languages at the end of this section.

To what extent are the generalizations extrapolated from Dutch corroborated or rejected by our sample? Comparison of the third major columns of tables I and II provides striking confirmation of generalization A (scrambling is limited to head final contexts). That is, in VO-languages it is impossible to scramble the complement leftward across the verb; in OV-languages, on the other hand, leftward scrambling of a complement to V within the verbal domain is attested in all the languages of our sample. This contrast is also found in mixed languages. Finnish, which is mixed in the sense of being V-X and X-P (besides P-X), does not permit scrambling, whereas all the other mixed languages, being V-final, do exhibit scrambling within the verbal domain.

The intuition that head-finality of a phrasal category is a necessary (though not sufficient) condition for scrambling within that phrasal domain is also confirmed by short scrambling patterns in domains other than VP, e.g. DP or AP. The relevant outcomes are given in the sixth column of each of the tables. The findings depicted in table I show that short scrambling in N-initial or A-initial structures is blocked. The YES in the cell representing short scrambling in Chinese DPs is in need of some qualification. It is not clear whether there is such a thing as complement to N in Chinese. Things that would normally be regarded as complements selected by the head have exactly the same form as modifiers. Modifiers, even relative clauses, always precede the noun they modify. If these modifiers are base generated in a prenominal position in Chinese, then the permitted permutation of the prenominal modifiers is not in contradiction with generalization A. We should immediately add, however, that further research into these modifier permutations is needed before we can conclude that these patterns of word order variation within the Chinese noun phrase are on a par with the more common scrambling patterns involving complements selected by some lexical head.

When we consider table II, we observe that short scrambling appears to be found in N-final and A-final phrases. However, as indicated by the NO's for Turkish and Bengali, head finality is not a sufficient condition for having scrambling. That is, independent, thus far largely unknown, factors may rule out application of scrambling in these non-verbal domains. As indicated by the question marks accompanying the YES/NO's, there is some vacillation in the judgments of these scrambling patterns. Obviously, more careful study

of the scrambling patterns in these nonverbal domains is definitely required before we can draw any firm conclusions on the basis of these data.

The fourth major columns of Tables I, II and III present data concerning long scrambling. As is clear from Table I, long scrambling out of AP, PP and DP into the verbal domain is impossible. This outcome is not unexpected given generalization A; long scrambling in rigidly head initial languages would cross two heads. Schematically:

(49)  $[_{YP} \mathbf{A} \mathbf{Y} [_{XP} \mathbf{X} \text{--}]]$

The cross-linguistic data on long scrambling in V-final languages (cf. Table II) yield the following picture: With the exception of L-scrambling out of DP in Hindi, long scrambling out of PP and DP is unattested in V-final languages. The only phrasal category which does not block long scrambling is AP.

Table III gives the following information concerning long scrambling in mixed languages: First, long scrambling out of the noun phrase appears to be blocked in all five mixed languages of our sample. This is in line with generalization A: given the N-complement order of these mixed languages, scrambling out of the noun phrase would take place from an N-initial context.<sup>12</sup> Long scrambling out of PP and AP into the verbal domain is blocked in Finnish and Persian, but allowed in Dutch, German and Hungarian. The impossibility of L-scrambling in the former two languages follows again from generalization A: In Persian, scrambling into the dominating head final VP-domain will always take place out of a head initial embedded PP or AP. In other words, the lower phrase forms the head initial context blocking L-SCR. In Finnish, L-scrambling out of APs and PPs having the order P-complement is also prohibited because of generalization A; Both the lower domain (AP/PP) and the dominating verbal domain into which the scrambled element is placed are head initial and, as such, are environments blocking scrambling.

---

<sup>12</sup> Istvan Kenesei and Katalin É. Kiss inform us that if the nominal head N is nom or acc, it can be followed by a complement (cf. (i)). If the nominal head is in another case, there has to be some kind of "attributivization" by means of a semantically empty participle, most often that of the copula (cf. (ii)). Although the latter structure is opaque to scrambling, the former conceivably permits scrambling out of the noun phrase. See also example (1) in the main text.

- (i) Peter Mari altal valo meghivasa  
Peter's Mary by being invitation
- (ii) Peter meghivasa Mari altal  
Peter's invitation Mary by  
'Peter's invitation by Mary'



b. Janek *książkę* kupił [e]

(51) a. Vanja kupil knigu (Russian)  
Johnny bought book

b. Vanja *knigu* kupil

If one takes VO to be the base order of these languages, the OV-order is derived by leftward scrambling of the complement across the verbal head. This appears to contradict our generalization A, which prohibits leftward scrambling across a lexical head. Recall, however, that the typological generalizations formulated in this pilot study only apply to so-called neutral scrambling. As was noted in section 2, neutral scrambling behaves differently in certain respects from focus scrambling (e.g. with respect to long scrambling out of clausal domains). One might hypothesize that a verbal head can only be crossed if the the scrambled elements is (contrastively) focussed. As exemplified in (52) and (53), it is true that the leftward scrambled element can be (contrastively) focussed (focus indicated by capitals):

(52) Vanja *VODU* pil, a Maša *VINO* pila (Russian)  
John water drank, and Mary wine drank  
'John drank water and Mary drank wine' (Harries-Delisle 1978)

(53) A: kogo Janek pocałował, Marię czy Sue? (Polish)  
Who Janek kiss, Mary or Sue  
'Who did Janek kiss, Mary or Sue?'

B: Janek pocałował *MARIE*  
John kissed Mary

Janek *MARIE* pocałował

Although most of our informants agreed that leftward scrambling across the lexical head was permitted under a focus reading of the scrambled constituent, there was much more obscurity about the possibility of having leftward neutral scrambling (see the third major column of Table IV).<sup>13</sup> Recall that the latter type of scrambling typically involves specific indefinite noun phrases and D-linked (i.e. familiar) noun phrases. An illustrative example

---

<sup>13</sup> Dalina Kallulli informs us that in Albanian scrambling of non-focuous accusative objects is only possible in the presence of a doubling clitic.

of a neutrally scrambled object is provided by the following Polish discourse segment (example taken from Lubańska (1996)):

- (54) A: Czy Janek pocałuje Marię? (Polish)  
 Q John will-kiss Mary  
 'Will John kiss Mary?'  
 B: Nie, on *tę idiotkę* kopnie  
 No, he that idiot will-kick  
 'No, he will kick that idiot'

In the sentence uttered by speaker B, the noun phrase *tę idiotkę* refers to *Marię*, an individual familiar from the discourse context. The focussed element in this sentence is the verb *kopnie*. As noted in Lubańska (1996)), there is variation among speakers of Polish in the judgment of this scrambled sentence. Although some speakers accept it, others judge it as incorrect. Importantly, all speakers accept the order in which the familiar noun phrase follows the verbal head (*Nie, on kopnie tę idiotkę*).

In view of the above, it seems premature to conclude that the Slavic/Balkan languages contradict our typological generalization A. Further study of scrambling phenomena in these languages will have to determine whether scrambling can be of the neutral type. For the moment, we will therefore discard these languages from our discussion.

## 5. Extrapolations:

### 5.1. A Left Roof Constraint?

One way of summarizing the results, at least once the Slavic/Balkan languages have been (temporarily) discarded, is given in the following table.

category	head-compl.	scrambling	L-scrambling
V <sup>max</sup>	V-X	NO	NO
	X-V	YES: Dutch, German, Japanese, Turkish	NO
N <sup>max</sup>	N-X	NO	NO
	X-N	YES: Hindi, Japanese	NO / YES: Hindi

$A^{\max}$	A-X	NO	NO
	X-A	YES: Hindi, Japanese	YES: Bengali, Hindi, Japanese
	A-X-fa	YES: Dutch, German	YES: Dutch, German
$P^{\max}$	P-X	NO	NO
	X-P	YES?: exx missing	YES?: exx missing
	P-X-fp	YES: Dutch, German	YES: Dutch, German

All the NOs, needless to say, await further verification. Furthermore, no examples of scrambling within the extended PP have come to light in purely postpositional languages, hence the YES in that shaded cell is, for the time being, purely hypothetical. But note that the hypothesis as such derives from a notion of regularity that is imposed by the table as a whole.

Turning now to the second pair of shaded cells, we have an interesting choice. On the one hand, we have Generalization C above which suggests that the answer here should be NO. In that case we would have to examine more closely the claim that Hindi has L-SCR out of noun phrases with a view to explaining this case away. On the other hand, we might want to revise Generalization C in light of the findings from Hindi. This would be interesting in two ways.

First, it would suggest a close link with Generalizations F' and F'', both of which oppose V and N/A/P. And perhaps it would also lead us to reexamine Generalization E which says that AP/PP can be hybrid, as opposed to (extended) NP/VP and to postulate that hybrid noun phrases can exist as well, as suggested by the data from Turkish, Bengali and Hungarian.

Second, it would suggest a link with one of the main laws governing rightward movement processes. The main bounding constraint on rightward movement, due to Ross (1967/1986), is the Right Roof Constraint, sometimes also referred to as the Upward Boundedness Condition. The Right Roof Constraint states that EXTRAPOSITION PROCESSES CAN CROSS THE BOUNDARIES OF ANY (EXTENDED) MAXIMAL PROJECTION EXCEPT THE BOUNDARY OF THE EXTENDED VERBAL PROJECTION, I.E, THE CLAUSE. Formulated in slightly more technical terms, extraposition can cross DP, AP and PP, but not IP/CP. Observe now that if we assume the V vs. N/A/P split for L-SCR, as suggested here, we can identify the two constraints because the domain restriction on L-SCR would now amount to what we could call the Left Roof Constraint. Thereby, we would have discovered a degree of symmetry between scrambling and extraposition that had hitherto remained undetected.

## 5.2. The No Crossing Corollary

As stated above, there is an implied corollary in the hypothesis that (leftward) scrambling is limited to head final structures. This is the claim that a scrambled element never crosses a head, at least not a lexically filled one. Any account of scrambling in terms of movement has problems accounting for the no-crossing property. On the other hand, a theory which holds that scrambling (as opposed to C-scrambling (and LC-scrambling)) is base-generation as argued by Neeleman (1994), Bayer & Kornfilt (1994), Kiss (1994) and others, combined with some form of a headedness parameter, say directional theta-identification, yields the no-crossing property automatically. Of course, the property that scrambling only occurs in head-final structures now follows trivially, since scrambling is by definition leftward, where leftward now means ‘limited to leftward theta-identifying structures’. The consequence of this is that theta-identification cannot be limited to strict sister nodes as is commonly assumed. Instead the domain for theta-identification must now, presumably, be defined as something like L(exical)-government, where L-government is taken to mean that a head L-governs elements (on one side) within the lexical part of its (extended) projection. In summary, two assumptions suffice to derive the no-crossing property:

- (S-)scrambling does not involve movement
- theta-identification is limited to directional L-domains

Presumably, the no-crossing property also holds for L-scrambling (as opposed to LC-scrambling). This means that L-scrambling cannot involve movement either, and hence that in these cases theta-identification has to proceed via upward extension of the L-domain. Thus:

- L-scrambling does not involve movement either
- L-domains can be extended upward for AP, PP, and perhaps NP/DP (cf. 5.1.) when governed by V

## 6. References

- Baltin, M. (1982) 'A landing site theory of movement rules' *Linguistic Inquiry* 13, 1-38.
- Beninca, P. (1989) *Dialect variation and the theory of grammar*. Dordrecht: Foris.
- Bordelois, I. (1974) 'The grammar of Spanish causative complements' Ph.D. Dissertation MIT.
- Blau, J. (1980) *Manuel de Kurde (dialecte Sorani)*. Paris: Klincksieck.
- Bayer, J. & J. Kornfilt (1994) 'Against scrambling as an instance of Move-alpha' in N. Corver & H.C. van Riemsdijk (eds.), 17-60.
- Browning, M. & E. Karimi (1994) 'Scrambling to object position in Persian' in N. Corver & H.C. van Riemsdijk (eds.), 61-100.
- Chomsky, N. (1977) 'On *wh*-movement' in P. Culicover, T. Wasow & A. Akmajian (eds.).
- Chomsky, N. (1981) *Lectures on government and binding*. Dordrecht: Foris.
- Corver, N. (1994) 'Phrasal structure and word order within the adjectival system' in *GLOW Newsletter* 32.
- Corver, N. (1996) 'The internal syntax of the Dutch extended adjectival projection,' ms., Tilburg University.
- Corver N. & H.C. van Riemsdijk eds. (1994) *Studies on scrambling*. Berlin: Mouton - de Gruyter.
- Culicover, P., T. Wasow & A. Akmajian eds. (1977) *Formal syntax*. New York: Academic Press.
- Dubuisson, C., D. Lightfoot & Y.C. Morin eds. (1978) *Montreal working papers in Linguistics 11*.
- Emonds, J. (1978) 'The verbal complex V'-V in French' *Linguistic Inquiry* 9, 151-75.

- Greenberg, J.H. ed. (1978) *Universals of human language* Volume 4 Syntax, Stanford University Press, Stanford, California.
- Grimshaw, J. (1991) 'Extended projection' ms., Brandeis University.
- Gross, M., Halle, M. & Schützenberger, M-P, eds (1973) *The formal analysis of natural languages*, Proceedings of the first international conference, The Hague-Paris, Mouton.
- Harries-Delisle, H. (1978) 'Contrastive emphasis and cleft sentences' in J.H. Greenberg (ed.)
- Holmberg, A. (1986) 'Word order and syntactic features in the Scandinavian languages and English' Ph.D. Dissertation Stockholm University.
- Horvath, J. (1985) *Focus in the theory of grammar and the syntax of Hungarian*. Dordrecht: Foris.
- Jackendoff, R. (1977) *X-bar syntax: A study of phrase structure*. Cambridge, MA: MIT Press.
- Kayne, R.S. (1975) *French syntax: The transformational cycle*. Cambridge, MA: MIT Press.
- Kayne, R.S. (1989) 'Facets of Romance past participle agreement' in P. Beninca (ed.), 85-103.
- Keyser, S.J. (1978) *Recent studies in European languages*. Cambridge, MA: MIT Press.
- É. Kiss, K. (1981) 'Structural relations in Hungarian, a "free" word order language' *Linguistic Inquiry* 12, 381-416.
- É. Kiss, K. (1994) 'Scrambling as the base generation of random complement order' in N. Corver & H.C. van Riemsdijk (eds.), 221-256.
- Koopman, H. (1984) *The syntax of verbs: From verb movement rules in the Kru languages to Universal Grammar*. Dordrecht: Foris.
- Koster, J. (1975) 'Dutch as an SOV language' *Linguistic Analysis* 1, 111-36.
- Kuroda, S.Y. (1965) 'Generative grammatical studies in the Japanese language' Ph.D. Dissertation MIT [published as Kuroda (1979), New York: Garland Pub.]
- Lubańska, M. (1996) ms. without title, University of Wrocław.
- Neeleman, A. (1994) 'Scrambling as a D-Structure Phenomenon' in N. Corver & H.C. van Riemsdijk (eds.), 387-429.
- Pollock, J-Y. (1989) 'Verb movement, universal grammar, and the structure of IP' *Linguistic Inquiry* 20, 365-424.
- Riemsdijk, H.C. van (1978) *A case study in syntactic markedness*, Dordrecht: Foris.
- Riemsdijk, H.C. van (1990) 'Functional prepositions' in H. Pinkster & I. Genée (eds.) *Unity in Diversity: Papers Presented to Simon C. Dik on his 50th Birthday*. Dordrecht: Foris.
- Riemsdijk, H.C. van (forthcoming) 'The extension of projections,' to appear in the Tsuru University Working Papers in Linguistics.

- Rijk, R. de (1972) 'Studies in Basque syntax: Relative clauses' Ph.D. Dissertation MIT.
- Rizzi, L. (1978) 'A restructuring rule' in S.J. Keyser (ed.)
- Rizzi, L. (1978) 'Violations of the Wh-island constraint in Italian and the Subjacency condition' in C. Dubuisson, D. Lightfoot & Y.C. Morin (eds.)
- Ross, J.R. (1967) *Constraints on Variables in Syntax*. Ph.D. Dissertation, MIT. [Published as Ross (1986)]
- Ross, J.R. (1973), 'Slifting' in Gross et al. (eds).
- Ross, J.R. (1986) *Infinite Syntax!* Norwood, NJ: Ablex Publishing Company.
- Rudin, C. (1988) 'Multiple questions and multiple WH-fronting' *Natural Language and Linguistic Theory* 6, 445-501.
- Saito, M. (1992) 'Long distance scrambling in Japanese' *Journal of East Asian Linguistics* 1, 69-118.
- Ueyama, A. (1994) 'Against the A/A'-Movement Dichotomy', in N. Corver & H.C. van Riemsdijk (eds.), 459-486.
- Windfuhr, G.L. (1979) *Persian Grammar*. The Hague: Mouton.
- Zwart, C.J.W. (1993) *Dutch Syntax: A Minimalist Approach*. Ph.D. Dissertation, University of Groningen.