

Immediate Precedence in GPSG*

Arnold M. Zwicky
Ohio State University and Stanford University

and Joel A. Nevis
University of Michigan

1. Precedence and immediate precedence. The ID/LP (immediate dominance/linear precedence) format of generalized phrase structure grammar (GPSG) separates the principles expressing conditions on branching from those expressing conditions on the relative order of sister constituents. An LP principle is of the form $X < Y$, where X and Y are (possibly complex) category names; such a principle requires that category X precede category Y whenever the two are sisters.¹

Precedence in this sense need not be immediate. Thus the LP condition in (1), while excluding the orderings of sister constituents V, NP, and PP in (2), permits all of the orderings in (3)--including V PP NP, in which V and NP are not immediately adjacent to one another.

- (1) $V < NP$
- (2) *NP V PP, *NP PP V, *PP NP V
- (3) V NP PP, V PP NP, PP V NP

However, in a language with considerable hierarchical (rather than flat) constituent structure, mere precedence (symbolized by the simple sign '<') and immediate precedence (which we will symbolize by the double symbol '<<') will often amount to the same thing. Thus the LP principle (4) holds in English, where it has the effect of requiring (5), given that Det and Nom have no further sisters under NP. In this particular case in English, a simple precedence condition would suffice.

- (4) $\text{Det} < \text{Nom}$
- (5) $\text{Det} << \text{Nom}$

2. Positioning adverbs in Finnish. Now consider a language with flatter constituent structures and freer word order than English. Consider, for example, Finnish, where word order within a clause is free (except that NPs have continuous, strictly ordered parts) and also where the words (again, except those in NPs) are immediately dominated by S. Finnish permits all six orderings of the constituents in a sentence composed of a subject, a direct object, and a finite V, as in (6):

- (6) a. Juha lyö Heikkiä. 'Juha hits Heikki'
- b. Juha Heikkiä lyö.
- c. Lyö Juha Heikkiä.
- d. Lyö Heikkiä Juha.

- e. Heikkiä lyö Juha.
- f. Heikkiä Juha lyö.

Despite this general freedom of word order, Finnish has a number of adverbs whose location is quite rigidly determined with respect to other words; one such item is the sentential adverb *myös* 'also, too'.²

We must observe first that *myös* belongs to two distinct adverbial classes, that is, that it has two distinct uses: first, in what we will call its 'local' use it is attached to (and normally precedes) the constituent it modifies, as in (7).

- (7) a. Eilen Juha antoi kirjan *myös* Marjalle.
'Yesterday Juha gave a book also to MARJA'
- b. Eilen Juha antoi Marjalle *myös* kirjan.
'Yesterday Juha gave Marja also a BOOK'
- c. Eilen Juha *myös* antoi Marjalle kirjan.
'Yesterday Juha also GAVE Marja a book'

In its 'sentential' use, which is the one of interest to us here, *myös* has scope over the whole sentence, and it must immediately follow the finite V--wherever this V happens to be located in its S. We illustrate this constraint in (8); (8a) has V in second position, (8b) in initial position.

- (8) a. Juha antoi *myös* kirjan Marjalle.
'Also, Juha gave a book to Marja'
- b. Antoi *myös* Juha kirjan Marjalle.

The sentences in (8a) and (8b) are not, of course, pragmatically equivalent; but they are both grammatical, they are semantically equivalent, and their semantics differs from the semantics of the sentences in (7).

The traditional approach to finite verb + *myös* units assumes that the two form a subconstituent in the sentence; we will label this subconstituent V'. Immediate precedence would fall out from immediate dominance under this treatment, just as it does for English in Det and Nom as above, and ordinary precedence would suffice. Now we know of no evidence that actually favors this approach, and there are considerations that speak against it, having to do with the generally flat constituent structure of Finnish Ss.

First, other sentential adverbs, VP adverbs, and verbal adverbs are clearly, like V and its NP arguments, daughters of constituents larger than V'; in Nevis (1985) it is argued that rules generate some of these as daughters of VP which then appear as daughters of S by virtue of a 'flattening' metarule. Not only would *myös* have to belong to a special subclass generated as a daughter of V' rather than VP or S, but it would also have to be exempted from the general flattening effect.

Second, there are sentential adverbs having multiple positioning within the sentence, including the slot immediately after the finite V. One such adverb is sentential *vain* 'only, just', which occurs in sentence-second position as well as immediately after V. If an ID rule generates *myös* as a daughter of V', then the same rule generates *vain* in this configuration--in which case *vain* requires not only two separate LP rules (as will any adequate analysis), but also two separate ID rules.

We conclude that ID rules generate *myös* as a sister of the finite V. The analytic problem we are then addressing is how to state an ordering constraint on the daughters of S.

3. **Eliminating << in favor of <.** Suppose the class of adverbs like *myös* in its sentential use is labeled [Class:29]. Can the obvious and elegant principle (9) be eliminated in favor of principles using only mere precedence, < ?

$$(9) \quad [+V, -N, Bar:0, +Finite] << [Adv, Class:29]$$

Yes, but the cost is greater than the prize. In general, the precise effect of the principle (10) can be achieved by the conjunction of the two principles (11a) and (11b).

$$(10) \quad X << Y$$

$$(11) \quad \begin{array}{l} a. \quad X < Y \\ b. \quad (\exists Z)(X < Z \quad \& \quad Z < Y) \end{array}$$

Principle (11a) is of course innocuous, but (11b), with its quantification over categories, is quite suspect; countenancing such conditions extends the range of expressible generalizations about linear precedence into new territory, so that conditions like the one in (11'), which allows phrasal but not lexical categories to intervene between X and Y, would be permitted. And in any case there is no intrinsic connection expressed between the content of (11a) and (11b); we would have no reason to expect that (11b) is vastly more likely to cooccur with (11a) than the condition in (11") is.

$$(11') \quad (\exists Z)(X < [Z, Bar:0] \quad \& \quad [Z, Bar:0] < Y)$$

$$(11'') \quad (\exists Z)(X < Z \quad \& \quad Y < Z)$$

(The condition (11'') by itself requires that when X and Y are sisters, one of them must be the last constituent in its construct. In conjunction with (11a) it requires that when X and Y are sisters, Y must be the last constituent in its construct--though no such condition would hold when Y combined with sisters other than X.)

It might seem that a simpler solution would be to replace (10) with the conjunction of two principles (11a) and (11c), where ' $\sim X$ ' stands for the complement of the features mentioned in X.

$$(11) \quad \begin{array}{l} a. \quad X < Y \\ c. \quad Y < \sim X \end{array}$$

Again, no intrinsic connection is expressed between the content of (lla) and (llc), and there are technical infelicities associated with specifying the complement of complex feature descriptions; the immediate-precedence statement in (9), for instance, would have to be replaced by a whole set of LP conditions referring to <, as in (12).

- (12) [Adv, Class:29] < [-V]
 [Adv, Class:29] < [+N]
 [Adv, Class:29] < [+V, -N, Bar:n] (for n>0)
 [Adv, Class:29] < [+V, -N, Bar:0, -Finite]

But there is a much stronger criticism of (lla) & (llc): together, they require that whenever X and Y are sisters, X must be the first constituent in its construct--and this requirement is usually too strong. Finnish *myös*'s must immediately follow V, but V does not have to be clause-initial, as we have already illustrated in example (8a).

Consider even the very configurational language English, and the condition in (13) (the requirement that nothing can intervene between a verb and its direct object, modulo Heavy NP Shift), as illustrated in (14). This condition cannot be handled by the combination of conditions in (15), because (15) would put V first in VP whenever there was an object NP.

- (13) V << NP

(14) *set very quickly the pot on the stove, *gave yesterday a
 marvelous lecture, *ate in the kitchen cheese

(15) a. V < NP
 b. NP < XP (where X = V, P, A, Adv)

But AdvP can come first in VP as well as later in this constituent, as we illustrate in (16a, b).

- (16) a. very quickly set the pot on the stove
 b. set the pot on the stove very quickly

We conclude that the LP principles of GPSG should be able to refer directly to << as well as to <.

4. **Some observations.** We have two final notes, a long one and a short one. The long one is a remark that immediate precedence is quite commonly called for in the analysis of languages with free constituent order or free word order, in particular in the description of items that must appear in second position or in penultimate position. In the framework we have been sketching, 'second position' translates as 'immediately following an X in first position', and 'penultimate position' as 'immediately preceding an X in last position'. This is not the place to explore how first and last position are to be described; let us simply suppose that there are features First and Last that are associated with the first and last constituents, respectively, in a set of sister

constituents.³ Then the requirement that some item Y appear in second position is described by an LP condition like the one in (17).

- (17) [+First] << Y

Finnish has items restricted in just this way. In particular, the adverbs *muka* 'supposedly' and *toki* 'truly' in their sentential uses must occur immediately after the first daughter constituent in an S, as we illustrate in (18) and (19) with *muka*. The sentences in (19) cannot be understood as paraphrases of those in (18).

- (18) a. (Tässä) vuoressa muka asuu peikko.
this mountain ADV live troll
'In (this) mountain supposedly lives a troll'

b. Peikko muka asuu (tässä) vuoressa.

(19) ≠ (18)

- a. *(Tässä) muka vuoressa asuu peikko.
b. *Muka (tässä) vuoressa asuu peikko.
c. *(Tässä) vuoressa asuu muka peikko.
d. *(Tässä) vuoressa asuu peikko muka.

The ID conditions of Finnish will insure that the class of adverbs to which *muka* in its sentential use belongs--call it [Class:17]--is introduced only as a daughter of S, along with the main verb and its various NP arguments. The LP condition in (20), which requires immediate precedence, then permits such adverbs to occur only in the second slot among the daughters of S.

- (20) [+First] << [Adv, Class:17]

Our short final remark concerns the generative capacity of syntaxes incorporating immediate precedence conditions as well as simple precedence conditions. It is easy to see that each such syntax describes a context-free language (immediate precedence conditions merely eliminate a finite number of ordering possibilities from a finite set), so that no increase in weak generative capacity results from permitting immediate precedence conditions. The gain is entirely in our ability to state generalizations.

Notes

*A version of this paper was delivered at the 1985 annual meeting of the Linguistic Society of America, Seattle. Our thanks to members of this audience, in particular Richard Oehrle, for their comments. This is the version of 16 March 1986.

¹The ID/LP format was set out in Gazdar and Pullum (1981). See Gazdar et al. (1985) for formal development of the proposal.

²Our discussion of the relevant facts about Finnish adverbs draws directly on Nevis (1985).

³These 'position features' are subject to the following conditions:
(a) if a daughter category is specified for a position feature *F*, then its mother category must be specified for *F* as well; and (b) if a mother category is specified for a position feature *F*, then exactly one of its daughters must be specified for *F*. The position features are then a special type of foot feature, subject to the uniqueness requirement in (b) as well as to (universal) LP conditions requiring a [+First] category to precede all its sisters and a [+Last] category to follow all its sisters.

References

- Gazdar, Gerald, Ewan Klein, Geoffrey Pullum, and Ivan Sag. 1985.
Generalized phrase structure grammar. Oxford: Basil Blackwell.
- Gazdar, Gerald and Geoffrey Pullum. 1981. Subcategorization, constituent order and the notion 'head'. In M. Moortgat, et al., 107-23.
- Moortgat, Michael, Harry van der Hulst, and Teun Hoekstra (eds.). 1981.
The scope of lexical rules. Dordrecht: Foris.
- Nevis, Joel A. 1985. Finnish particle clitics and general clitic theory.
Ph.D. dissertation, Ohio State University.