CONSTRAINTS ON ORDER

Technical Note 364

October 1985

By: Hans Uszkoreit, Computer Scientist
Artificial Intelligence Center
Computer Science and Technology Division
and
Center for the Study of Language and Information

Constraints on Order[*]

Hans Uszkoreit
Artificial Intelligence Center
SRI International and
Center for the Study of Language
and Information at Stanford University

Abstract

Partially free word order[1] as it occurs in German and probably to some extent in all natural languages arises through the interaction of potentially conflicting ordering principles. A modified linear precedence (LP) component of Generalized Phrase Structure Grammar (GPSG) is proposed that accommodates partially free word order. In the revised framework, LP rules are sets of LP clauses. In a case in which these clauses make conflicting ordering predictions, more than one order is grammatical. LP clauses may refer to different types of categorial information such as category features, morphological case, thematic roles, discourse roles, and phonological information. The modified framework is applied to examples from German. It is demonstrated how the new LP component constrains the linear ordering of major nominal categories.

1. Advantages and Shortcomings of the ID/LP mechanism

It is no coincidence that none of the applications of GPSG to German that are described in the literature predate the introduction of the immediate dominance/linear precedence (ID/LP) formalism into the framework. Although Stucky (1981) had already been able to give a GPSG account of a free-word-order language using metarules for the reordering of the constituents, it was actually not before the separation of linear precedence from immediate dominance due to Gazdar and Pullum (1981) that the framework became attractive for linguists working on languages with

1

extensive order variation. For instance, all GPSG grammar fragments for German published so far employ LP rules for determining the position of the verb in the sentence (Johnson, 1983; Nerbonne, 1982, 1984; Uszkoreit, 1982, 1983, 1984).

The ID/LP notation in GPSG removes the information about linear order from phrase structure rules and encodes it in a separate rule set. This move leads to two sets of rules, immediate dominance (ID) and linear precedence (LP) rules. The following example demonstrates the relationship between an ordinary phrase-structure grammar and the corresponding ID/LP grammar.

(1)

traditional PSG	ID/LP grammar	
$A \rightarrow B\,C\,D$	ID rules	LP rules
$A \rightarrow B\,D\,C$		
$B \rightarrow B\,D$	$A \rightarrow B, C, D$	$B < C$
$C \rightarrow A\,B\,C$	$B \rightarrow B, D$	$B < D$
$C \rightarrow B\,A\,C$	$C \rightarrow A, B, C$	
$C \rightarrow B\,C\,A$		

An ID rule of the form $C_0 \rightarrow C_1, ..., C_n$ only states that a constituent of category C_0 may immediately dominate a set of n constituents if there is a one-to-one category assignment mapping from these constituents to the multiset of categories $C_1, ..., C_n$. An LP rule $C_k < C_l$ determines that any constituent of category C_k has to precede any constituent of category C_l whenever C_k and C_l occur as sibling constituents.

A relevant feature of the formalism is that the ordering statements encoded in LP rules are not only notationally detached from ID rules but

that they also apply independently. Every subtree of Depth 1 has to comply with all LP rules, independently from the ID rule that was used to generate it. The empirical claim implicit in this definition is discussed in Gazdar and Pullum (1981) under the term exhaustive constant partial ordering (ECPO).

Versions of the framework differ in the exact place of LP rule application within the modular definition of the grammar[2]. Yet, these differences are not essential for the extension of the LP component to be proposed and no explication or comparison of the approaches will be presented.

Shortly after the ID/LP formalism was introduced into the framework, it was already considered for describing word order variation (Pullum 1982, Stucky 1982, Uszkoreit 1982). Three classes of linear ordering can be handled easily with the formalism that Gazdar and Pullum had suggested: fixed order, ordering variation that depends on a certain syntactic feature, and free order variation. Examples for the three classes can be found in German. In German, just as in English, determiners always precede the nominal which with combine. This can be encoded by an LP rule as in (2):

(2) DET < NOM

The position of the verb in German sentences depends on the status of the clause. In a main clause, the finite verb precedes its siblings[3]; all other verbs in main and subordinate clauses follow their siblings. Two LP rules suffice to state this variation:

(3a)   MC:+ <   X
(3b)   X   < MC:-

A feature cooccurrence restriction ensures that the feature +MC (main clause) is only assigned to finite verbs. X is a variable that ranges over the whole set of categories. As the LP rules under (3a) and (3b) indicate, syntactically determined ordering variation is treated as fixed constituent order. A mnemonic category symbol such as NOM is just an abbreviation for a feature set: N:+, V:-, BAR:1. These features do not have any special status. Thus there is no formal difference between LP rules that refer to category symbols such as (2) and LP rules that refer to features such as the two rules under (3a) and (3b).

The third possibility, i.e., totally free ordering variation, does not require any special rules. The absence of LP rules that impose a linear order on a set of sibling constituents will permit all permutations of these constituents. Let us assume for the moment that instrumental and locative adverbials in German are not ordered with respect to each other. Both (4a) and (4b) are equally acceptable:

(4a) Der Bauer arbeitete mit   einer alten Hacke auf   dem Feld.
      the farmer worked   with an   old   hoe   in   the field

(4b) Der Bauer arbeitete auf dem Feld mit   einer alten Hacke.
      the farmer worked   in   the field with an   old   hoe

No rule is needed to allow for the ordering variation.

The ID/LP formalism provides a general and elegant way for describing syntactic phenomena that exhibit two types of ordering relations: totally fixed order and totally free order. Since the two types of ordering relations are encoded by the presence or absence of LP rules, grammars can be

4

easily written for languages that exhibit both fixed and free word order phenomena. This makes the formalism more suitable for describing languages with partially free word order, such as German, than previous frameworks.

It is very likely that all natural languages possess partially free word order. So-called free-word-order languages such as Warlpiri, a Central Australian language, show at least some ordering requirements such as the second position of the auxiliary clitic. On the other side of the spectrum there are languages like English, which are assumed to have a rather rigid word order. However, even in these languages a certain degree of ordering freedom can be found, just as in the placement of adverbs and adverbials in English. It seems then that the ID/LP version of GPSG provides an ideal tool for capturing the existing degrees on the scale of ordering freedom.

Yet, there are two aspects of the formalism that should dampen our enthusiasm. One is a metatheoretical concern based on traditional views of markedness and grammar complexity. Because fixed-order constructions require rules that are not needed for free order constructions, fixed word order appears more highly marked than free word order. This bias by itself does not provide a valid argument against the theory, because no other theory has offered a more plausible view on the markedness of word order. But the mere fact that one major component in the grammar of a free-word-order language is much smaller than the corresponding component in a fixed-word-order language--without any indication of a compensation for the difference in other components--is somewhat disturbing.

The second objection to the formalism is stronger since it can be substantiated by empirical facts. The earlier references to partially free word order were based on the view that the partial ordering freedom results from mixing of totally free and totally fixed order. This view contradicts accepted linguistic observations about many of the relevant languages. In Slavic languages, and in Finnish, Hungarian, German, and many other languages, a different type of partially free order has been observed. Here partially free order is not a mixture of free and fixed order but a degree of freedom that lies between free and fixed order. The constraints on this type of partially free order are determined by the interaction of several ordering principles.

A bona fide case of partially free word order is presented in the next section. It is shown that the original ID/LP framework is not equipped to describe the relevant ordering constraints.

2. The Order of Complements and Adjuncts in German

2.1. The Partiality of Ordering Freedom

The order of the major nominal categories in German provides an excellent example for partial ordering freedom. The phenomenon can be seen most clearly by considering the linear order among the complements and adjuncts of the main verb. In main clauses, these elements follow the finite verb, with the possible exception of a single fronted constituent. They can in turn be followed by a nonfinite main verb and nonfinite auxiliary verbs. Since the string of complements and adjuncts might be surrounded by other material, it is traditionally referred to as the middle field *(Mittelfeld)* of the clause.

At first glance it might seem that the linear order in the middle field is a case of truly free variation. For many sentences, all permutations of the major constituents in the middle field give rise to other grammatical sentences of German. Consider the following permutations of a sentence whose middle field contains a subject and two objects. The subject is marked nominative case, the direct and indirect objects are marked accusative and dative cases, respectively.

(5a) Dann wird der Doktor   demPatienten   die Pille geben.
              NOM   DAT   ACC
      then   will   the doctor   the   patient   the   pill   give

(5b) Dann wird der Doktor die Pille dem Patienten geben.
(5c) Dann wird dem Patienten die Pille der Doktor geben.
(5d) Dann wird dem Patienten der Doktor die Pille geben.
(5e) Dann wird die Pille der Doktor dem Patienten geben.
(5f) Dann wird die Pille dem Patienten der Doktor geben.

All six permutations can be roughly papraphrased as: *the doctor will then give the patient the pill.*

The three permuting constituents in (5) are definite nonpronominal noun phrases. If we consider the full range of phrases that can fill the roles of subject and objects, we immediately find numerous ungrammatical sequences among the permutations of grammatical sentences. The sentence in (6b) differs from the one in (6a) only in the order within the middle field.

(6a) Dann wird es   ihm   einen Ball geben.
                 NOM DAT   ACC
      then   will   it   him   a   ball   give

(6b) *Dann wird einen Ball ihm es geben.

2.2. Ordering Principles for the Middle Field

In this subsection, I present a number of regularities that affect the linear order within the middle field. A small subset of uncontroversial regularities will be selected that translates into five ordering principles. These principles will be used later to demonstrate the proposed extension to the LP component.

Different types of ordering regularities have been observed. First, there are certain regularities that have been attributed to some underlying or unmarked order. Lenerz (1977) contains a comprehensive investigation of the unmarked order among complements and certain types of adverbials. Two tendencies are found that affect the order among subject and objects:

(7a) The subject tends to precede the objects.

(7b) The indirect object tends to precede the direct object.[4]

Whereas Lenerz attributes these tendencies to an unmarked order that depends on grammatical functions or grammatical relations such as subject and direct object, Uszkoreit (1984) argues for an alternative analysis that resembles the viewpoint of functional sentence perspective[5]. According to this view, it is the thematic roles of constituents that determine their unmarked order. Evidence is mainly drawn from classes of verbs with nonagentive subjects and from passive constructions. Uszkoreit (1984) also provides evidence for including grammatical case in the description of unmarked order. However, in German, grammatical case plays a very minor role in determining the order of complements and adjuncts of verbs. For the limited range of examples to be considered here, the predictions of an unmarked order based solely on thematic roles coincide with the predictions of an unmarked order that includes grammatical case. Therefore, the following three ordering principles (OP) that are supported by Lenerz' and Uszkoreit's findings suffice for the current discussion.

(8a) OP1: The agent precedes the theme.
(8b) OP2: The agent precedes the goal.
(8c) OP3: The goal precedes the theme.

In the absence of a generally accepted theory of thematic roles, I have followed common practice and borrowed the names for the roles from case grammar.[6] Nothing that follows hinges on this choice.

Other ordering regularities have to do with the pragmatic status of the constituents or with syntactic features that often bear pragmatic information.

(9a) Known information tends to precede unknown information.
(9b) New information tends to precede old information.
(9c) Pronominal elements tend to precede nonpronominal elements.
(9d) The topic tends to precede other constituents.
(9e) The focus tends to follow other constituents.
(9f) Definite NPs tend to precede indefinite NPs.

There exists very little agreement about the questions that underly pragmatic notions such as topic, focus, new, or unknown information. It is not clear yet which of the principles in (8) overlap and to what degree they can be reduced to a smaller set of regularities. However, it is uncontroversial that the linear order is affected by pragmatic properties of constituents that need not be syntactically marked. Linguists agree that focussed constituents in the middle field tend to follow other constituents. The following ordering principle is added:

(10) OP4: Focussed constituents follow other constituents.

(11) OP5: Personal pronouns precede nonpronominal constituents.

A further class of observed ordering constraints involves phonological properties of constituents. Longer, heavier constituents of the middle field show the tendency to follow shorter, lighter sibling constituents.[7] This tendency can be found in many languages, including English. There is

10

also an influence of the rhythmic pattern that occasionally effects the judgements on marginally acceptable permutations. Although I believe that phonological constraints on word order, especially the heaviness constraint, cannot be considered extragrammatical and therefore neglected in the grammatical description of a language, I exclude them here from the description of partially free word order in German. The problems connected with the specifics of encoding the relevant phonological properties of constituents are not essential for the current discussion.

The five ordering principles derived from the wealth of constraints observed by linguists should suffice to demonstrate the source for partial ordering freedom. They represent the strongest among the factors that determine the linear order of the middle field. For the reader's convenience, they are repeated below.

(12a) OP1:   The agent precedes the theme.
(12b) OP2:   The agent precedes the goal.
(12c) OP3:   The goal precedes the theme.
(12d) OP4:   Focussed constituents follow other constituents.
(12e) OP5:   Personal pronouns precede nonpronominal constituents.

These ordering principles bear a striking resemblance to LP rules. If we assume for German the "flat structure" approach adopted by Uszkoreit (1982) and by most subsequent GPSG analyses of German (Johnson, 1983; Nerbonne, 1982, 1984; Uszkoreit 1983, 1984) , an encoding of the ordering principles as LP rules is supported by the constituent structure. According to this approach, the main verb and all constituents in the middle field are siblings. Yet, it needs to be shown, how such LP rules can be expressed in the ID/LP framework.

2.3. Ordering Principles as LP Rules

For the formalization of the ordering principles, we need to assume grammatical features that express the properties that the principles refer to. I save the discussion of the motivations behind introducing features for thematic roles and discourse roles for later. For the time being I will use the features TR (thematic role), FOCUS, and PPRN (personal pronoun)[8] with the following value assignments (feature:value).[9]

(13a)   TR:AGENT, TR:GOAL, TR:THEME
(13b)   FOCUS:+, FOCUS:-
(13c)   PPRN:+, PPRN:-

It is now easy to formulate LP rules corresponding to the five ordering principles.

(14a)   TR:AGENT   <   TR:THEME
(14b)   TR:AGENT   <   TR:GOAL
(14c)   TR:GOAL   <   TR:THEME
(14d)   FOCUS:-   <   FOCUS:+
(14e)   PPRN:+   <   PPRN:-

If these LP rules were added to the grammar of German, the middle field would be strictly ordered. Moreover, many combinations of constituents would be ruled out. In a sentence containing an agent and a theme, the theme NP can never be a personal pronoun if the agent phrase is a nonpronominal NP.

(15a)   Dann wird   der Doktor   ihn   untersuchen.
                     TR:AGENT   TR:THEME
                     PPRN:-   PPRN:+
        then   will   the   doctor   him   examine

(15b) Dann wird   ihn   der Doktor untersuchen.
                   TR:THEME TR:AGENT
                   PPRN:+   PPRN:-
       then will him   the doktor   examine

Both (15a) and (15b) violate a LP rule, thus there is no order that satisfies the LP component. However, both variants are grammatical. The problem is that LP rules apply conjunctively, i.e., a sequence has to satisfy **all** LP rules at once. This is not how partially free word order works. One ordering principle can be overridden by another ordering principle. Lenerz (1977) finds this type of interaction between the unmarked order and the tendency of the focus to follow nonfocussed constituents. All the tests he designs for determining the unmarked order of complements and adjuncts rely on the observation that one of the two principles can be violated if the other one legitimizes the violation.

There are no tools for describing this mode of interaction in the original ID/LP framework. The form and functioning of LP rules need to be redefined to allow encoding potentially conflicting ordering principles as they are found in German.

3. The Revision of the LP Component

3.1. Definition of LP Rules

LP rules were originally defined as ordered pairs of categories (Gazdar and Pullum, 1981). The LP component of the grammar was defined as a strict partial order, i.e., as an asymmetric, transitive relation in the set of categories. LP rules $\langle \alpha, \beta \rangle$ are written as $\alpha < \beta$ and interpreted as follows:

(16)   A LP rule $\alpha < \beta$ states: if two constituents labelled α and β appear as sibling constituents, then α precedes β.

This can be reformulated: if a LP rule $\alpha < \beta$ applies to a pair of constituents, that is if the constituents are labelled α and β, then the order of those constituents has to comply with the order prescribed by the rule.

The LP component has also occasionally been interpreted as a nonreflexive, antisymmetric, transitive relation. LP rules of the form $\alpha <$ α were suggested to rule out multiple occurrences of α among sibling constituents, because any order of two constituents α would violate the rule.[10] The definition of LP rules in GKPS reflects this change; LP is defined as an antisymmetric transitive relation. The accompanying interpretation of LP rules is slightly reworded but logically equivalent to (16):

(17)   A LP rule $\alpha < \beta$ states: a constituent labelled β cannot appear to the left of a sibling constituent labelled α.

14

We will see that the equivalence does not carry over to certain other definitions of LP rules.

The productive use of complex categories makes things somewhat more complicated. All grammatical symbols that can be found in actually proposed LP rules do not denote fully instantiated categories. Just like the symbols used in ID rules, they stand for feature combinations that abbreviate large sets of actual categories. It is useful to redefine the LP component as a relation in the power set of features (Uszkoreit, 1983) or in the set of partially defined categories (GKPS).

Although the two alternatives collapse for versions of the framework that define categories as sets of feature-value pairs such as in GKPS, the latter definition is more general. It does not assume a certain type of internal structure for categories. As long as we have a weak partial order on the set of complex symbols that tells us for each pair of symbols whether the occurrence of the first in a rule applies to occurrences of the second in the structures on which the rule operates, it does not matter whether categories are feature sets, trees, or other structured entities.

Depending on the assumed definition of category, the ordering relation might be the *subset* relation (for sets), but it might also be *extension* in the sense it is used in GKPS (for trees), or *subsumption* as it is used in connection with unification (for directed graphs). I use the symbol $\subseteq$ for this relation: $\kappa \subseteq \chi$ shall mean κ is subsumed by χ.

Let me reformulate the interpretation of LP rules in a way that accomodates the use of partial categories and that at the same time

15

reflects the underlying linguistic concept behind the rule type. To this end, I first define some intuitive notions connected with the application of LP rules. To begin with the standard concept of precedence, I define the relation *precedes* on strings.

(18)   If Φ is a string $\phi_1 \dots \phi_n$, then *precedes*$_\Phi$ is a strict total order on the elements of Φ such that every pair $\langle \phi_i, \phi_j \rangle$, $1 \le i < j \le n$, is in *precedes*$_\Phi$.

Then I state what it means for a string to be admitted by a LP rule.

(19)   A LP rule R *admits* ϕ if and only if R *admits* all members of *precedes*$_\Phi$.

I need to specify next which ordered pairs are admitted by a LP rule. At this point, it is useful to remember what the linguistic basis for LP rules is. Every LP rule is a statement about the admissible order of sibling constituents that belong to the categories mentioned in the rule. This statement does not apply to the order of any other types of constituents. Let me define a relation *applies to* that holds between LP rules and ordered pairs of constituents. The ordered pair $\langle \gamma, \delta \rangle$ shall refer to an ordered pair of constituents labelled γ and δ.

(20)   A LP rule $R = \alpha < \beta$ *applies to* a pair of constituents $\pi = \langle \gamma, \delta \rangle$ iff either $\beta \subseteq \delta$ and $\alpha \subseteq \gamma$ or $\beta \subseteq \gamma$ and $\alpha \subseteq \delta$.

Obviously, a pair has to comply with a LP rule only if the rule applies to it.

(21)   A LP rule R *admits* a pair of constituents n if and only if the
       following holds: if R *applies to* n, then n *complies with* R.

According to (21), a LP rule admits a pair of constituents in two cases:
either the rule does not apply to the pair, or if it does apply, the pair
complies with the rule. Things could not be simpler. All that needs to be
done now is to explain what it means for a pair of constituents to comply
with a LP rule. To this end, I will define a relation *complies with* that
holds between ordered pairs of constituents and a LP rule.   Two
alternative definitions will be considered. I start with the negative
definition that underlies the most current version of GPSG.

(22)   An ordered pair of constituents $\langle \gamma, \delta \rangle$ *complies with* a LP rule
       $\alpha < \beta$ iff it is not the case that $\beta \subseteq \gamma$ and $\alpha \subseteq \delta$.

I will then contrast this definition with the following positive
interpretation.

(23)   An ordered pair of constituents $\langle \gamma, \delta \rangle$ *complies with* a LP rule
       $\alpha < \beta$ iff $\alpha \subseteq \gamma$ and $\beta \subseteq \delta$.

There are clearly formal differences between the two definitions. The
definition under (22) is a filter that does not affect any pair to which the
LP rule does not apply.   This property of the definition permits the
omission of the relation *applies to* in current definitions of the LP
component. The definition under (23), on the other hand, implies that a
LP rule $\alpha < \beta$ applies to every ordered pair of constituents $\langle \gamma, \delta \rangle$ that
complies with it.

17

The two definitions also differ slightly in their linguistic predictions. The only cases in which the alternative definitions give conflicting opinions about the order of two constituents occur when the categories of both constituents in $\langle \gamma, \delta \rangle$ fit both (partial) categories in the LP rule $\alpha < \beta$, i.e., when $\beta \subseteq \delta$, $\alpha \subseteq \gamma$, $\beta \subseteq \gamma$, and $\alpha \subseteq \delta$. For this to happen, the LP rule need not be be of the form $\alpha < \alpha$. Imagine the following fictitious LP rule.

(24) $\{N:+\} < \{V:+\}$

Imagine furthermore that we had to use this LP rule to order pairs of constituents belonging to the following set of categories: NP, VP, and AP. The version of X-BAR Theory that is embodied in GPSG assumes the following feature assignements:

(25)

	N	V
NP	+	-
VP	-	+
AP	+	+

Under the different interpretations (22) and (23), the fictitious LP rule (24) would accept almost the same set of pairs.

(26)

pair	(22)	(23)
NP NP	y	y
NP VP	y	y
NP AP	y	y
VP NP	n	n
VP VP	y	y
VP AP	n	n
AP NP	n	n
AP VP	y	y
AP AP	**n**	**y**

It is only in the sequence AP AP that the two definitions yield contradicting judgements. Under definition (23), it is not possible to use LP rules to prohibit the cooccurrence of certain categories among sibling nodes by creating ordering paradoxes. I am not aware of any empirical evidence that supports either definition in the area of strict word order regularities. The sole purpose of the foregoing comparison was to set the stage for the introduction of a new type of LP rules that can account for partial ordering freedom.

3.2. Complex LP Rules

For the encoding of potentially conflicting ordering principles I redefine LP rules as sets of LP clauses, in which LP clauses have the form of traditional LP rules.

(27) A LP rule is a set of ordered pairs of partial categories
$\{\langle \alpha_1, \beta_1 \rangle, \langle \alpha_2, \beta_2 \rangle, ..., \langle \alpha_n, \beta_n \rangle \}$, written as

$$\left\{ \begin{array}{ccc} \alpha_1 & < & \beta_1 \\ \alpha_2 & < & \beta_2 \\ & \vdots & \\ \alpha_n & < & \beta_n \end{array} \right\}$$

Now I will define how these complex LP rules are interpreted. Two minor extensions to the definitions of *applies to* and *complies with* will suffice.

(28) A LP rule C *applies to* a pair of constituents π iff at least one LP clause in C *applies to* π in the sense of definition (20).

(29) A pair of constituents π *complies with* a LP rule C iff π *complies with* at least one LP clause in C in the sense of definition (23).

The definition of the relation *admits* remains unchanged. It is easy to paraphrase the interpretation of complex LP rules in plain English. A complex LP rule admits all sequences of constituents if and only if it admits all ordered pairs within these sequences. The rule admits an ordered pair if none of the clauses in the rule apply to the pair. If one or more clauses apply to the pair, then the order of the pair has to follow at least one of the ordering statements in the rule.

LP rules that impose a strict order, i.e., LP rules in the traditional sense, are subsumed under the definition of complex LP rules. In the revised LP

20

component, these rules are singleton rules. Since they contain only one LP clause, this clause has to be complied with whenever it applies.

Before I discuss some of the theoretical consequences of the framework and the underlying theory of linear order, I will use some examples from German for a demonstration of the new mechanism.

4. Test Case German

4.1. A Complex LP rule at Work

In Section 3.3. , some conflicting LP rules (14a-e) were considered for the linear order of the complements and adjuncts of main verbs. They are repeated here as (30a-e).

(30a) TR:AGENT < TR:THEME
(30b) TR:AGENT < TR:GOAL
(30c) TR:GOAL   < TR:THEME
(30d) FOCUS:-   < FOCUS:+
(30e) PPRN:+   < PPRN:-

These rules could not be used for the description of German as separate LP rules. Now they can be combined in a complex LP rule (the numbering on the left is for later reference to the individual clauses):

$$(31) \quad \left\{ \begin{array}{lll} \text{i.} & \text{TR:AGENT} < \text{TR:THEME} \\ \text{ii.} & \text{TR:AGENT} < \text{TR:GOAL} \\ \text{iii.} & \text{TR:GOAL} < \text{TR:THEME} \\ \text{iv.} & \text{FOCUS:} - < \text{FOCUS:} + \\ \text{v.} & \text{PPRN:} + < \text{PPRN:} - \end{array} \right\}$$

In the following examples, I only consider the order among the complements of the main verb, because these are the only constituents that are affected by LP rule (31). The position of the verb with respect to its arguments is determined by the singleton LP rule $\{X < MC:-\}$, which corresponds to (3b).

(32)

Dann wird	der Doktor	demPatienten	die Pille	geben.
	TR:AGENT	TR:GOAL	TR:THEME	
	FOCUS:-	FOCUS:+	FOCUS:-	
	PPRN:-	PPRN:-	PPRN:-	
then will	the doctor	the patient	the pill	give

The only violation of a clause in LP rule (31) is the order FOCUS:+ FOCUS:- in the pair of NPs *dem Patienten* and *die Pille*. However, the phrase *dem Patienten* may precede the phrase *die Pille* because the thematic role of the first is GOAL and the thematic role of the second is THEME. Thus the order of the constituents complies with a clause in the LP rule.

(33)	Dann wird	der Doktor	die Pille	demPatienten	geben.
		TR:AGENT	TR:THEME	TR:GOAL	
		FOCUS:-	FOCUS:-	FOCUS:+	
		PPRN:-	PPRN:-	PPRN:-	
	then will	the doctor	the pill	the patient	give

This time, the order of the object NPs violates the LP clause that makes goal phrases precede theme phrases. Yet, the violation is legitimized by the compliance with the LP clause that is responsible for the postponing of focus phrases.

In the next example the order of the object violates both LP clauses without legitimation. The sentence is ruled out.

(34) ??	Dann wird	der Doktor	die Pille	demPatienten	geben.
		TR:AGENT	TR:THEME	TR:GOAL	
		FOCUS:-	FOCUS:+	FOCUS:-	
		PPRN:-	PPRN:-	PPRN:-	
	then will	the doctor	the pill	the patient	give

The judgement agrees with Lenerz' findings. However, the sentence appears still more acceptable than sentences in which fixed-order

23

constraints are violated.[11] This issue involves the inevitable notion of degrees of grammaticality. Every plausible theory of partially free order needs to provide some account of the range of acceptability that can be observed among ordering variants.

4.2. Degrees of acceptability

The explanation I offer for the distribution of ordering variants on the scale of acceptability is extremely simple. In a full-fledged model of linguistic competence, the LP clauses within LP need to be weighed with respect to each other. In the complex LP rule (31), the last clause will receive the highest weight. Next on the weight scale are the rules that have the agent phrase precede other phrases and by the clause that orders the goal before the theme. The LP clause that refers to the position of the focus phrase comes last on the weight scale, although it would still defeat the heaviness principle and other clauses that are not included in the rule example.

The combined weights of the observed or violated rules in the linearization of a set of sibling nodes determine the acceptability of the order. I do not want to speculate about the nature of the combining function; the considered examples merely indicate that is monotonically increasing.

The resulting model can account for several observations. One has to do with the relative frequency at which LP clauses are violated or observed. One can find longer stretches of text in which the order of the complements is not affected by LP clause iv. in any of the sentences. It is extremely unlikely that longer sequences of sentences could be found in which LP clause v. is never observed. If LP clauses iv. and v. are in

conflict usually v. wins out. The personal pronoun is only postponed if it is stressed as an emphatic or contrastive focus.

(35) Darum hatte der Doktor die Pille IHR gegeben (und nicht ihm).
     therefore had the doctor the pill HER given (and not him)

Therefore the doctor had given the pill to her (and not to him).

A violation of all applicable LP clauses as it could be seen in (34) is marginally acceptable if the violation can be attributed to one of the weaker LP clauses. In the example, the violation would be attributed to LP clause iv. It is compensated by a stronger than usual stress on the focus:

(36)?? Dann wird   der Doktor DIE PILLE dem Patienten geben.
                     TR:AGENT   TR:THEME   TR:GOAL
                     FOCUS:-   FOCUS:+   FOCUS:-
                     PPRN:-   PPRN:-   PPRN:-
        then will   the doctor THE PILL   the patient give

If we assume that the relative weight of an LP clause in a singleton rule is 1 and that the weight of LP clauses in complex rules add up to 1, then we get an explanation for the different degrees of ungrammaticality resulting from the violation of the different rule types. A violation of a singleton rule has more effect on the acceptability of the sentence than the violation of one of the weaker LP clauses in a complex LP rule. Stronger LP clauses of complex rules lie somewhere in between.

It is extremely difficult to discuss sufficiently large sets of examples for the interaction of the weighted LP clauses within a complex LP rule. Even if one restricts the attention to small numbers of permuting elements, feature-value pairs, and LP clauses, the resulting combinatorics is quite

25

impressive. A further complication arises from the distribution of judgements that is typical for marginally acceptable sentences.

Because of these complications I present a very limited set of examples that I have tested with a number of native speakers. The first sentence is the Example (6b), here repeated as (37). I will then stepwise change order and features of the complements. Each new sentence will be more acceptable than its predecessor.

(37) * Dann wird   einen Ball   ihm   es   geben.
                      TR:THEME   TR:GOAL   TR:AGENT
                      FOCUS: +   FOCUS:-   FOCUS:-
                      PPRN:-   PPRN: +   PPRN: +
       then will   a   ball   him   it   give

The order of the complements violates LP clauses i., ii., iii., iv. (twice), and v. (twice). In the next sentence, the subject propnoun will be replaced by a proper name.

(38) ?? Dann wird   einen Ball   ihm   Peter   geben.
                      TR:THEME   TR:GOAL   TR:AGENT
                      FOCUS: +   FOCUS:-   FOCUS:-
                      PPRN:-   PPRN: +   PPRN:-
       then will   a   ball   him   Peter   give

Now LP clause v. is violated only once. Since LP clause v. carries much weight, there is a detectable difference in grammaticality. The number of LP clause violations will be further reduced. In Example (39), the subject becomes the single focus of the sentence:

(39) ? Dann wird   den Ball   ihm   ein Freund geben.
                      TR:THEME   TR:GOAL   TR:AGENT
                      FOCUS:-   FOCUS:-   FOCUS: +
                      PPRN:-   PPRN: +   PPRN:-
       then will   the ball   him   a friend   give

26

LP clauses i., ii., iii., v. are violated. The next sentence, finally, pleases even the most critical informants.

(40) Dann wird ihm den Ball ein Freund geben.
                 TR:GOAL TR:THEME TR:AGENT
                 FOCUS:– FOCUS:– FOCUS:+
                 PPRN:+ PPRN:– PPRN:–
   then will him the ball a friend give

Only two LP clauses are violated: i. and ii. Both violations are legitimized because they permit the application of LP clause iv.

The correlation between the number of violated LP clauses and the degree of ungrammaticality is quite surprising. An even more fine grained model can be derived if stylistic influences are included.

4.3. Grammar and Style

The choice among possible linearizations is heavily influenced by stylistic factors. In the history of transformational grammar this influence became the basis for the assumption that the partially free word order in German belongs to a broad class of stylistic reorderings, or stylistic scrambling. This terminology is misleading. Standard use of the notions *grammar* and *style* would not support a stylistic component that enlarges the language.

I assume that the grammar is a characterization of the sentences of a language and that the stylistic component of a model of language use merely chooses among sentences characterized by the grammar. Whenever there is a choice in producing an utterance among elements that are semantically and pragmatically equivalent with respect to the purpose, content, and context of the utterance, the stylistic component selects on the basis of a variety of criteria.

The framework proposed here for partially free word order often generates grammatical ordering variants that are semantically and pragmatically equivalent with respect to a given utterance. The stylistic component may apply stylistic criteria to select among the variants. As would be expected, the contribution of the grammar remains invariant across utterance situations, whereas the stylistic influence depends on the speaker's stylistic preferences, the medium of communication, the register of speech, and many other factors.

For a realistic model of language production, one would prefer a stylistic component that is integrated with the other parts of the grammar over one that only works as a filter on sets of already generated sentences. The stylistic component can dynamically add to or subtract from the weights assigned to LP clauses. Emphasis on focussed elements, for instance, may lead to a higher weight for LP clause iv. Written texts often show more applications of LP clause iv. than oral utterances because in written texts focus elements cannot be marked by phonological stress.

I summarize the suggested interaction of grammar and style: LP rules rule out every order that is not motivated by one of the general ordering principles encoded in LP clauses. The combined weights of violated LP clauses lead to a preference ranking among the permitted linearizations. The preference ranking can be rearranged by the stylistic component. The application of LP rules can lead to sentences that are of marginal acceptability, if the stylistic component does not compensate for the relative deficiency.

A very convenient factorization of grammar and style could be based on the assumption that only the yes/no decisions resulting from LP rule applications belong in the realm of grammar and that all types of weight

assignment are performed by the stylistic component. This is the position taken in Uszkoreit (1982 and 1984). The view is tempting, yet it seems rather implausible that the invariant preferences of certain ordering principles within a language should be described as stylistic. If in two languages whose ordering principles overlap, different principles are consistently preferred, then these preferences cannot be merely ascribed to style. I will briefly return to this problem in the next section in which I discuss crosslinguistic implications of the framework.

5.   Discussion and Summary

5.1. Cross-linguistic Aspects of the Proposal

In the very first sentence of the paper, I alluded to the possibility that all languages might exhibit some occurrences of partially free word order. It is one of the conjectured universals about free-word-order languages that pragmatic considerations affect the linear order. That just a single ordering principle is at work in these languages is very unlikely. It is much more probable that phonological principles and unmarked order also play a role.

It might still be the case that in some constructions every possible order under every possible feature assignment yields a grammatical sequence. This could easily be described in the proposed framework. Since the stylistic component can operate on the involved LP clauses just as well, it is possible to describe a stylistic influence even for sequences of constituents that seem to be freely permutable.

Languages with totally fixed word order can be handled easily by the framework. However, it is not very likely that such languages exist. A prime example for an alleged fixed-word-order language is English. However, English exhibits partial word order freedom in a number of

29

constructions. The seemingly free order among adverbial phrases underlies ordering principles like those found in German: unmarked order, pragmatically based principles, the heaviness principle, and others. The application of the proposed framework to the particular problem of adverbial placement crucially depends on an analysis that treats these adverbials as siblings. Yet, I do not intend to argue for such an analysis within this paper--although I would certainly like to do so in light of a uniform treatment of partial word order freedom.

Let me instead consider the order among complements which is supposed to be fixed in English. There is a clear difference between the following sentences in their degree of ungrammaticality.

(41) * Peter takes his car the shop to

(42)?? Peter takes to the shop his car

Permutations of complements appear more acceptable than other permutations of siblings. Under certain conditions such as a contrast in heaviness or strong focusing, the permutation of complements even gives rise to acceptable sentences.

(43)   The king considers an enemy any armed foreign warrior who crosses our borders without advanced permission.

(44)   And then he revealed to his audience THE SUPERPLAN.

The interaction between LP rule acceptance and weight assignment permits a simple description of the constraints on these permutations. It should be obvious by now what such an analysis would consist of. I assume complex LP rules for the description of the order of the complements and adjuncts in English. The lesser degree of ordering freedom lies in the weights that are assigned to the LP clauses. This makes it harder for the stylistic componenent to override the standard

30

choice, which depends on either categorial information (NP<PP) or, such as in the case of double objects, on thematic roles (TR:GOAL<TR:THEME). Consequently the conditions for a deviation from the standard order are much more constrained than in German.

This analysis offers a plausible account of well-known ordering differences between German and English. In English, information about the syntactic category plays an important role in ordering the complements and adjuncts of the verb. In German, this role is negligable for most constructions.[12] Whenever the purely categorial information does not stipulate the order between two siblings of a main verb in English, the order follows ordering criteria that also apply to German, such as thematic role, discourse role, and heaviness. Because of the high weight of the ordering clauses that determine the order on the basis of category information, it is much harder for other principles such as the ones based on heaviness or discourse role assignment to override them.

It is one of the main advantages of the framework that it accounts for the cross-linguistic scale of word order freedom along two dimensions of partiality. One dimension is expressed through the number of conflicting LP clauses across LP rules. The other dimension comes in through the assignment of weights to LP clauses.

In Section 2, I mentioned a serious theoretical objection to the original ID/LP framework. LP components differ too radically in their complexity across languages. The revised framework will not assign nearly empty LP components to free word order languages. It is expected that some languages will require more but overall simpler LP rules whereas others might be described in terms of fewer but more complex LP rules.

31

5.2. Metatheoretical Aspects

In this section, I will remark on two metatheoretical implications of the revised framework; the incorporation of new types of features and its contribution to the generative power of GPSG.

LP clauses can refer to traditional categorial information and to pragmatic, thematic, and phonological information. Some integration of syntactic, pragmatic, and thematic concepts is inevitable if we apply a framework to wider classes of languages. A highly constrained surface oriented grammar model with a single level of grammatical representation demands a uniform encoding of properties of constituents.

The encoding of thematic roles in features allows for lexical rules that specify the subcategorization frame for verbs depending on the thematic roles that need to be instantiated (Wunderlich, 1984 and 1985). The incorporation of features for pragmatic properties accomodates topic languages and permits the linking of these properties with the appropriate non-truth-conditional semantic information in a rule-to-rule fashion. For a specific proposal about the syntactic and semantic effects of pragmatic features please refer to Karttunen and Kay (1985). The integration of non-truth-conditional semantics into the ID rules is independently needed for the description of such syntactic construction as clefts, pseudoclefts, and topicalizations. The integration of prosodic information in the grammatical representation of constituents is necessary for describing the various connections between intonation and syntax.

One point needs to be emphasized in connection with the diverse sources of the ordering principles. Providing the LP component with access to the

32

entirety of information that influences the linear order does not mean that any modularity is given up in the organization of the grammar. Pragmatic and phonological factors are determined purely by pragmatic and phonological regularities and mechanisms. Some ordering principles might well be directed by extragrammatical forces such as general principles in the ordering of information. However, if the roles of these principles in linearizing linguistic structures differs accross languages as it can be observed for the principles discussed in this paper, then the grammar needs to provide means for encoding these roles. Complex LP rules together with the weighting mechanism for LP clauses are designed to fulfill this function.

I will finish this section with a remark on the generative capacity of the framework. The design of the GPSG formalism is based on the hypothesis that natural languages are context-free. Original definitions and subsequent revisions ensure that the formalism does not exceed weak context-free power. As Gazdar and Pullum (1981) show, the ID/LP mechanism does not increase its weak generative capacity. Shieber (1984) was able to extend an existing efficient algorithm for the parsing of context-free languages to accomodate the ID/LP framework without a concomitant increase of the upper bound on the parse time as a function of the length of the input string.

It is still a controversial question how to determine the computational complexity of grammar formalisms. Yet, it is worth pointing out that the revised ID/LP framework does not increase the weak generative capacity of GPSG. Assume that we use the LP component in the generation of a fully specified phrase-structure grammar as in earlier versions of GPSG. If the number of ID rules is finite, so is the number of possible

linearizations of these ID rules. But then the set of linearized ID rules that are accepted by the revised LP component is finite, since the LP component is only a filter on the set of possible linearizations of ID rules. The set of nonterminal symbols is not affected by the LP component. Thus, the phrase-structure grammar that contains the rules accepted by the LP component will be context-free. If the LP component is just used for acceptance, it is compatible with Shieber's parsing algorithm. The complexity introduced by weight assignment and stylistic component will hinge on the design of these mechanisms.

5.3. Summary and Outlook

A revision of the ID/LP framework of GPSG has been proposed. It offers an analysis of partially free word order in German and a formalization of a universal linguistic phenomenon that had been a serious problem for previous formal theories of language. However, as usual, the real linguistic work starts after the new formalism has been proposed and demonstrated on a small fragment of natural language.

The new tools need not only be tested as to their applicability and usefulness in the description of wider classes of languages, but they also still require a good deal of shaping. The whole range of ordering principles needs to be investigated and encoded in LP clauses. The actual assignment of weights to LP clauses will require sophisticated empirical work. The weight combination function still needs to be determined. Finally, I have developed the strong feeling, just from having worked through many sets of examples on paper and blackboard, that an empirically sound formal description of partially free word order in any particular language belongs to the class of highly complex linguistic tasks that will not be solved without the extensive use of computers.

34

Footnotes

[1]Throughout this paper, the term *word order* is used in its traditional linguistic meaning, referring to the linear order of constituents. Thus, no distinction is made between *free word order* and *free constituent order*.

[2]In the traditional two-level model of GPSG, the two types of rules are considered parts of the metagrammar. They are used together with other metagrammatical devices such as metarules and feature instantiation principles to define a context-free phrase-structure grammar of the language, the object-grammar. In newer versions of the theory (Gazdar, Klein, Pullum,and Sag 1985; henceforth GKPS), all rule sets that were formerly considered part of the metagrammar are applied directly as a complex object grammar to admit well-formed trees of the language. Uszkoreit (1985a,b) argues that LP rules need to be redefined as order instantiation rules that operate together with feature and relation instantiation rules on basic lexical entries to derive fully instantiated lexical categories. Sag (1985a,b) proposes a dynamic model of LP rule application that defines LP rules as constraints on nonconcatenative string combination.

[3]The sentence initial constituents in verb-second clauses is not a sibling of the verb.

[4]The unmarked order of two pronominalized objects which differs from the regularly observed unmarked order will be excluded from the current discussion.

[5]Talks by Petr Sgall and Eva Hajicova at CSLI, Stanford 1984.

[6]In the naming of this thematic role, I follow Wunderlich (1984). Among the traditional roles or deep cases that are comprised by *goal* are recipients and benefactives.

[7]The phonological influence cannot be disregarded, even if one assumes that the syntactic and semantic complexity of a constituent also contributes to its *heaviness*.

[8]It is not essential for the current discussion whether these features are individual features or whether they could or should be defined in terms of combinations of other features.

[9]This notation is equivalent to the notation ⟨ feature, value⟩ , which is used in GKPS.

[10]Such rules were employed in rules for English by Gerald Gazdar (personal communication).

[11]In this case, the direct object is more stressed than other focus elements. A discussion of the complex interdependency of certain violations of ordering constraints and marked stress assignment is beyond the scope of this discussion. But see (Lenerz, 1977).

[12]Possible exceptions, such as the influence of syntactic case for the position of free datives are discussed in Uszkoreit (1984).

References

Gazdar, G., E. Klein, G. K. Pullum, and I. A. Sag (1985) *Generalized Phrase Structure Grammar*. Blackwell Publishing, Oxford, England, and Harvard University Press, Cambridge, Massachusetts.

Gazdar, G. and G. K. Pullum. (1981) "Subcategorization, Constituent Order and the Notion 'Head'." In M. Moortgat, H. v.d. Hulst and T. Hoekstra (Eds.), *The Scope of Lexical Rules*, 107-123, Foris, Dordrecht, Holland.

Johnson, M. (1983) "A GPSG Account of VP Fronting in German." University of California at San Diego (unpublished mimeo).

Karttunen, L. and M. Kay (1985) "Parsing in a free word order language." In D.R. Dowty, L. Karttunen, and A.M. Zwicky (Eds.) *Natural Language Parsing*, Cambridge University Press, Cambridge.

Lenerz, J. (1977) *Studien zur deutschen Grammatik*. Volume 5: *Zur Abfolge nominaler Satzglieder im Deutschen*. Narr, Tübingen.

Nerbonne, J. (1984) *German Temporal Semantics: Three-Dimensional Tense Logic and a GPSG Fragment*. PhD thesis, Ohio State University, Columbus, Ohio.

Nerbonne, J. (1982) "Phantoms and German Fronting: Poltergeist Constituents." Paper presented at the Annual Meeting of the LSA, San Diego, California.

Pullum, G. K. (1982) "Free Word Order and Phrase Structure Rules." In J. Pustejovsky and P. Sells (Eds.), *Proceedings of the Twelfth Annual Meeting of the North Eastern Linguistic Society*, Amherst, Massachusetts.

Sag, I. A. (1985a) "Grammatical Hierarchy and Linear Precedence." paper presented at the Conference on Categorial Grammars and Natural Language Structures, May 1985, Tucson, Arizona.

Sag, I. A. (1985b) "Continuous Analysis of Dicontinuous Constituents." paper presented at the Conference on Discontinuous Constituency, July 1985, Chicago, Illinois.

Shieber, S. M. (1984) "Direct Parsing of ID/LP Grammars." *Linguistics and Philosophy*, 7, 135-154.

Stucky, S. (1981) *Word Order Variation in Makua: a Phrase Structure Grammar Analysis*. PhD thesis, University of Illinois, Urbana-Champaign.

Stucky, S. (1982) "Linearization Rules and Typology." In D. Flickinger, M. Macken, and N. Wiegand (Eds.), *Proceedings of the First West Coast Conference on Formal Linguistics*, Stanford University, Stanford, California.

Uszkoreit, H. (1982) "German Word Order in GPSG." In D. Flickinger, M. Macken, and N. Wiegand (Eds.), *Proceedings of the First West Coast Conference on Formal Linguistics*, Stanford University, Stanford, California.

Uszkoreit, H. (1983) "A Framework for Processing Partially Free Word Order." In *Proceedings of the 21st Annual meeting of the ACL*, Cambridge, Massachusetts.

Uszkoreit, H. (1984) *Word Order and Constituent Structure in German*. PhD thesis, University of Texas, Austin, Texas.

Uszkoreit, H. (1985a) "Problematische Konstruktionen für kontextfreie Phrasenstrukturgrammatiken des Deutschen." paper presented at the Annual Meeting of the Deutsche Gesellschaft für

Sprachwissenschaft, Februar 1985, Hamburg (to appear in Klenk, U. (Ed.) *Strukturen und Verfahren in der maschinellen Sprachverarbeitung,* AQ-Verlag, Dudweiler).

Uszkoreit, H. (1985b) "Linear Precedence in Discontinuous Constituents." paper presented at the Conference on Discontinuous Constituency, July 1985, Chicago, Illinois.

Wunderlich, D. (1984) "Über die Argumente des Verbs." Univ. Düsseldorf, (unpublished).

Wunderlich, D. (1985) "'Ergative' Verben sind nicht ergativ!" paper presented at the Annual Meeting of the Deutsche Gesellschaft für Sprachwissenschaft, Februar 1985, Hamburg.