

SRI International

LINEAR PRECEDENCE IN DISCONTINUOUS CONSTITUENTS: COMPLEX FRONTING IN GERMAN

Technical Note 371

October 1985

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

The research reported in this paper has been made possible by a gift
from the System Development Foundation.



333 Ravenswood Ave. • Menlo Park, CA 94025
(415) 326-6200 • TWX: 910-373-2046 • Telex: 334-486

Linear Precedence in Discontinuous Constituents: Complex Fronting in German*

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

0. Abstract

Syntactic processes that have been identified as sources of discontinuous constituents exhibit radically different properties. They seem to fall into several classes: leftward "extractions," rightward "movements," "scrambling" phenomena, and parenthetical insertions. Current linguistic theories differ as to the formal tools they employ both for describing the participating syntactic phenomena and for encoding the resulting representations.

In this paper, the general problem of determining the linear order in the discontinuous parts of a constituent is discussed. The focus lies on frameworks that use their feature mechanisms for connecting the noncontiguous elements. It is then shown that the current framework of Generalized Phrase Structure Grammar (GPSG) is not suited for describing the interaction of leftward extractions, scrambling, and constraints on linear order. The relevant data come from German fronting. Previous analyses (Johnson 1983; Nerbonne 1984; Uszkoreit 1982; 1984) have neglected certain types of fronting or failed to integrate their account of fronting properly with an analysis of linear precedence. The critical constructions involve the fronting of main verbs together with some complements and adjuncts as in the following example:

- (1) Unbemerkt seine Brieftasche stehlen kannst du ihm nur in der Oper ε.
unnoticed his wallet steal can you him only at the opera
Only at the opera can you steal his wallet unnoticed.

If a widely accepted condition that restricts fronting to a single constituent is not sacrificed, a binary branching clause structure is required. The branching structure, however, prevents LP rules from specifying the linear order of the arguments and adjuncts of the verb.

The proposed modification of the framework redefines the relationship between syntax and lexicon. A subcategorization or valency feature is associated with an uninstantiated entry for a lexical head. Its value is a set of possible types of obligatory and optional complements including constituents that have traditionally been analyzed as adjuncts. When a lexical entry is instantiated, complex LP rules impose a linear order on a selected subset of complements. The order depends on syntactic, thematic, pragmatic, and phonological information. In the syntax, the head picks up its complements one by one, just as in categorial grammars, thus creating the branching structure suggested in previous proposals.

The modified framework admits grammars that preserve the feature analysis for unbounded dependencies, that abolish the metarule analysis for the scrambling of adjuncts and complements, and that treat certain stylistic reorderings such as heavy-NP-shift as simple results of LP rule applications.

1. The Problem

A syntactic unit is called discontinuous if its terminal elements do not form a contiguous substring of the sentence. Every syntactic framework that recognizes discontinuous syntactic units has to provide means for connecting their noncontiguous components. The connection is usually expressed in the syntactic representations. However, it might also be encoded in the grammar without appearing in the representations. The metarules of GPSG have often been used for establishing such connections.

If the connections are to be expressed in the representations, again several strategies can be used. The standard method in transformational grammar is based on sequences of (nontangled) trees as syntactic representations. In such a system, there

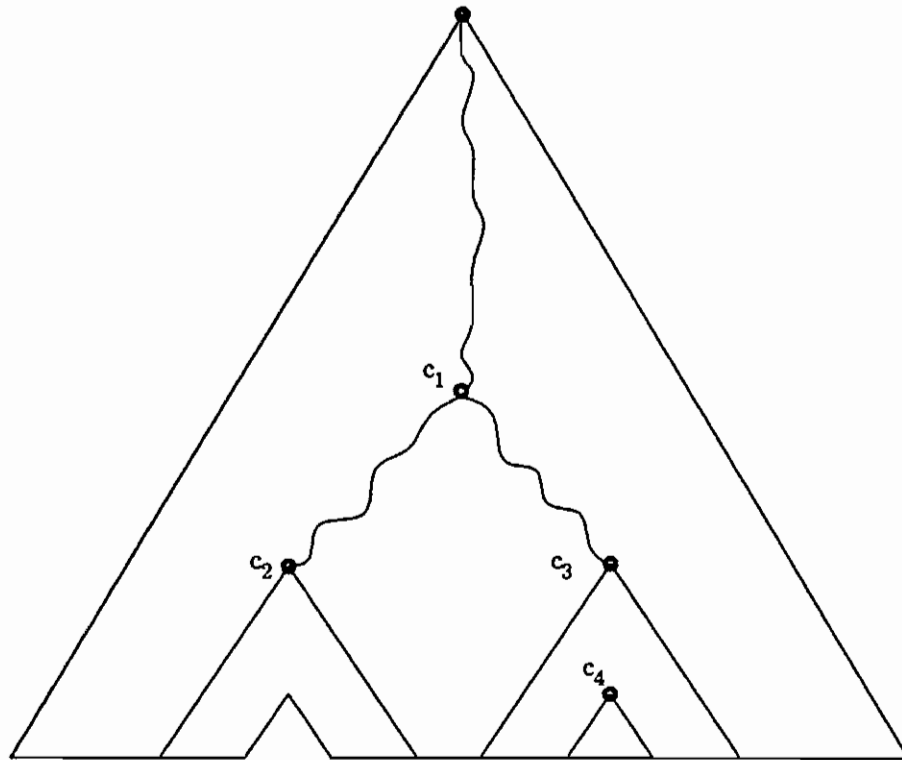
is at least one tree that assigns a single node to the discontinuous constituent. Thus, there is a structure in the representation in which the constituent is contiguous.

Another strategy is based on relaxing the nontangling condition on trees. All constituents are represented by a single node, even if their subconstituents are not adjacent. Discontinuities lead to crossing branches. Frameworks that possess a sophisticated feature mechanism can employ feature propagation for encoding the connection in complex category symbols. This is the way in which GPSG handles the fronting in constituent questions and topicalization.

In Phrase-Linking Grammar, certain extracted elements can have two parent nodes. Again, the result is tangled trees. In the syntactic representations of Lexical Functional Grammar (Bresnan, 1982), extracted elements can also be reached through more than one link. However, only one of the links can be part of a (nontangled) constituent tree, the other ones are part of the *functional structure*. The functional structure is a directed graph that allows nodes to be reentrant, i.e., to possess more than one parent node. Reentrant structures are utilized in a similar way in Functional Unification Grammar (Kay, 1985) and in PATR-II (Shieber et al., 1983). Several categorial grammars link the discontinuous elements through functional composition. Many grammatical frameworks use more than one method for dealing with discontinuity.

Most types of syntactic discontinuities can be analyzed as dislocations of a single constituent. In the schematized tree below, this situation is depicted in a quite theory-neutral way.

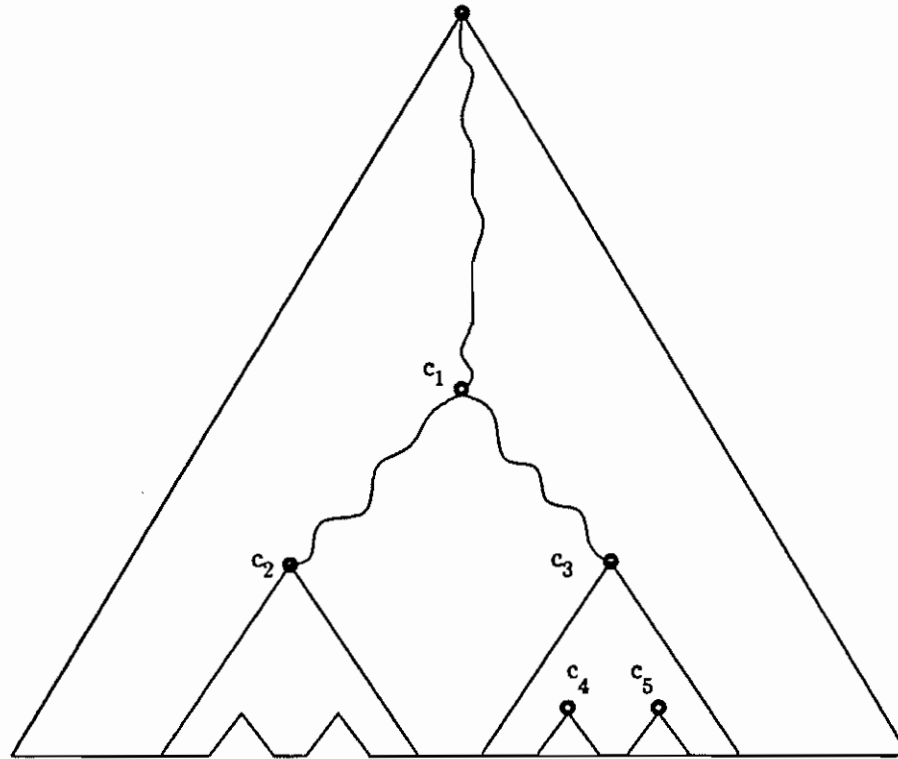
(2)



The constituent c_4 would be inside c_2 as an immediate subconstituent if it had not been dislocated. Depending on the specific type of dislocation and on the framework, the little empty triangle inside c_2 might be analyzed as a trace or gap constituent, c_2 might still immediately dominate c_4 by means of a crossing branch, c_3 might immediately dominate c_4 , or the gap information might be passed through c_1 encoded in features.

The problem to be addressed here arises if more than one subconstituent is dislocated. This case is schematically illustrated below:

(3)



How is the order of c_4 and c_5 determined? The type of discontinuity that is suggested by the general picture is unlikely to be found in natural languages. At least, I am not aware of any dislocation phenomena that involve the extraction of several daughters of a constituent and their insertion inside some other node. However, certain German fronting constructions could be analyzed as special cases of (3).

In a sentence like (1)--repeated below as (4)--several elements of the verb phrase (or sentence) have been fronted.

- (4) Unbemerkt seine Brieftasche stehlen kannst du ihm nur in der Oper ε
unnoticed his wallet steal can you him only at the opera
Only at the opera can you steal his wallet unnoticed.

In this paper, the apparent multiple frontings will be called *complex frontings*. Accordingly, the more common frontings of a single phrase will be referred to as *simple frontings*.

I will show that the order of the fronted elements depends on the order within the nonfronted part of the verb phrase or sentence. Actually, the order of the elements that are preposed in complex fronting constructions is identical to the order of the same elements in sentences in which they are not fronted. A hypothetical unconstrained transformational approach, in which any sequence of symbols in a labelled bracketing can be moved, could describe complex fronting easily by just moving the tail of a verb phrase in its original order. However, the framework of GPSG, which I use as a starting point, does not provide the tools for extracting an ordered sequence of constituents.

It has been suggested in previous GPSG analyses of complex fronting that the preposed string forms a single constituent (Nerbonne, 1982; Johnson, 1983). The appropriate constituents are obtained through assigning binary branching structures to German clauses. This view also forms the basis for my analysis. According to this approach, complex fronting constructions are special cases of (2). The problem that I will discuss arises through the conditions on the order within the single fronted constituent. If its order depends on the order within the remainder of the verb phrase or sentence, this dependency needs to be expressed by the grammar. But the component of GPSG that imposes linear order only applies to sibling constituents. The fronted and nonfronted elements of the verb phrase are not siblings, yet they need to be ordered with respect to each other.

The framework to be proposed does not apply linearization rules to sibling constituents but to valency lists that include a functor category and its arguments, including so-called free arguments or adjuncts.

The argumentation starts with a brief review of the GPSG analysis of German word order in Uszkoreit (1984). Then the approach of binary branching clause structures

is introduced. The incompatibility of the two analyses leads to the proposal of the new framework. A discussions of some further implications of the proposal concludes the paper.

2. A GPSG Analysis of German Word Order

2.1. Some Basic Concepts of the Framework

The theory of GPSG is built around the hypothesis that phrase-structure grammars are the appropriate model for the description of natural languages. A single level of syntactic representation is assumed. The representations are phrase-structure trees whose nodes are labelled with complex category symbols. Category symbols consist of feature-value pairs. Special rules govern the assignment and propagation of features.

One of the more specific features of the framework needs to be mentioned, since it is essential for the analysis of word order: the separation of immediate dominance from linear precedence. The strategy was introduced by Gazdar and Pullum (1981) under the term ID/LP notation. In this notation, the content of the phrase-structure rule $A \rightarrow BC$ has to be expressed in two rules: the ID rule $A \rightarrow B,C$ and the LP rule $B < C$. The ID rule only states that subtrees are permitted whose mother nodes are labelled A and whose daughter nodes B and C . This rule could also be written $A \rightarrow C,B$. The LP rule $B < C$ states that all constituents labelled B have to precede all sibling constituents labelled C . The order does not depend on the category of the mother, thus the two new rules are stronger than the phrase-structure rule.

The ID/LP formalism permits the concise statement of ordering regularities that, in previous notations, had to be repeated in many rules. Moreover, it makes the empirical prediction that all these regularities hold across phrase-structure rules.

2.2. The Phrase Structure of German

Uszkoreit (1984) presents a grammar for a fragment of German that exemplifies the typical syntactic properties of the language. Most of the phenomena are described in

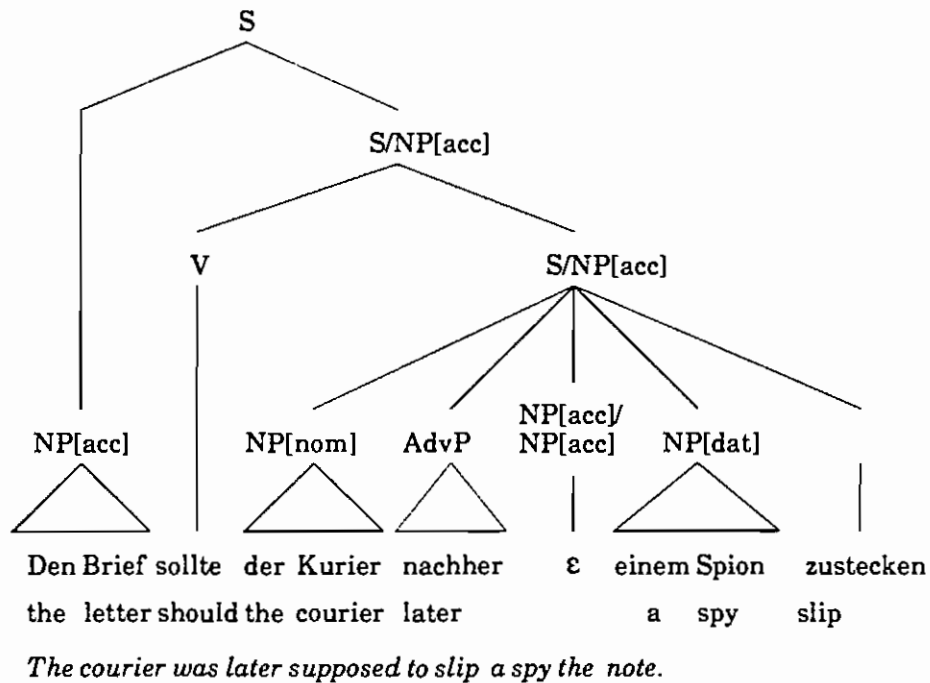
the standard notation of GPSG--sometimes using new strategies of applying the apparatus. The assumed constituent structures are basically those proposed in Uszkoreit (1982). In several points they resemble traditional transformational analyses such as Bach (1970) and Bierwisch (1971).

For German sentences, a structure is suggested in which the main verb and its arguments and adjuncts appear as sibling constituents. The relationship between the "flat" sentences (V^3) and the corresponding verb phrases (V^2) is established through a metarule. Simple fronting in declarative main clauses is treated as long distance dependency. It differs from English topicalization in the restrictions on frontable categories and on the categories that form the path between filler and gap. In accordance with the theoretical premises of GPSG, long distance dependencies are handled by connecting filler and gap through a category-valued gap feature that is assigned by general feature assignment mechanisms to all constituents on the path between gap and filler.

Finite and nonfinite main verbs are introduced by the same set of rules, independently of their position in the sentence. A LP rule ensures that the verb follows its argument and adjuncts if it is not a finite main clause verb. This LP rule is responsible for the final position of the verb in subordinate clauses and in verb phrases, such as in infinitival complements. Another LP rule states that a finite main clause verb precedes its arguments and adjuncts. The latter rule is needed for verb-initial and verb-second main clauses. Both LP rules extend to auxiliary and modal verbs. The second position in declarative sentences and in constituent questions arises through the interaction of the second LP rule and the rules for fronting.

The tree under (5) shows the structure of a German declarative sentence:

(5)



The gap appears as ϵ . The representation of the gap feature follows the traditional slash notation: /NP[acc] means that the value of the gap feature is an accusative noun phrase.

2.3. The Order in the Middle Field

Nothing has been said so far about the order in the middle field (Mittelfeld) of the sentence, i.e., the string of nonfronted arguments and adjuncts. The order of the sibling constituents in the middle field is partially free. For some sentences, every permutation of these elements yields a grammatical sentence; for other sentences, certain orders are ungrammatical. Partially free word order, as it probably occurs to some degree in all natural languages, results from the interaction of potentially conflicting ordering principles.

In the case of the middle field in German sentences, these conflicting ordering principles involve a variety of syntactic, pragmatic, and phonological factors. A few of these principles are listed below:¹

- (6) Agent phrases tend to precede theme phrases.
- Agent phrases tend to precede goal phrases.
- Goal phrases tend to precede theme phrases.
- Focus phrases tend to follow nonfocus phrases.
- Personal pronouns tend to precede nonpronominal phrases.

The principles cannot be encoded as the following regular LP rules:

- (7) TR:Agent < TR:Theme
- TR:Agent < TR:Goal
- TR:Goal < TR:Theme
- Focus: - < Focus: +
- PPron: + < PPron: -

Such LP rules would disallow all sentences in which the theme phrase is realized as a personal pronoun and the agent is not. Any sequence of the two phrases would be rejected by one LP rule. Yet, the principles need to be expressed in the grammar. Languages differ in the choice of ordering principles they apply and in the weight that they assign to them. Sentences that violate all of the principles are judged unacceptable by native speakers. The inability to express such ordering principles is a serious shortcoming of the framework.

The original ID/LP mechanism is only suited for totally fixed or totally free ordering phenomena. If there is a LP rule that determines the order of two categories, sibling constituents belonging to the two categories always have to be in this order. If there

is no such rule, nothing else can be used to express ordering conditions between the two categories.

Thus, the formalism can be applied to languages whose word order is partially free in a different sense of *partial*, i.e., languages that mix totally free and totally fixed order. If the order in the middle field of German sentences were totally free, German would certainly belong to this class of languages since the order of the other constituents is fixed. However, neither is the order in the middle field totally free, nor does totally free order occur in any other language I am familiar with.

Uszkoreit (1984) employs a different type of LP rules to adapt the framework to partially free word order. These LP rules contain sets of *LP clauses*. LP clauses in a LP rule may be violated if at least one LP clause in the same rule legitimizes the violation. The effect of the complex LP rules can also be described in a slightly different way: If a LP rule applies to a pair of sibling constituents, the linear order of these constituents must comply with the rule. A LP rule applies to a pair of constituents if at least one LP clause in the rule applies to the pair. A sequence of two constituents complies with a LP rule if it complies with at least one LP clause in the rule.² This definition subsumes traditional LP rules as a special case, as LP rules containing just one LP clause.

The LP clauses in a LP rule can carry different weights. The weights are determined by both grammar and style.

A simplified example of the complex LP rule that applies to the order of arguments and adjuncts in the middle field is presented under (8). The clauses are ordered by weight.

$$(8) \quad \left\{ \begin{array}{l} \text{PPron: +} < \text{PPron: -} \\ \text{TR:Agens} < \text{TR:Thema} \\ \text{TR:Agens} < \text{TR:Ziel} \\ \text{TR:Ziel} < \text{TR:Thema} \\ \text{Fokus: -} < \text{Fokus: +} \end{array} \right\}$$

3. An Alternative Approach

3.1. Complex Fronting

An important difference between topicalization in English and German fronting lies in the selection of frontable categories. The simplified rule system in Uszkoreit (1984) provides, for instance, for the following types of frontings:

- (9a) Den Brief sollte der Kurier nachher einem Spion zustecken.
- (9b) Der Kurier sollte den Brief nachher einem Spion zustecken.
- (9c) Einem Spion sollte der Kurier nachher den Brief zustecken.
- (9d) Nachher sollte der Kurier den Brief einem Spion zustecken.
- (9e) Zustecken sollte der Kurier den Brief nachher einem Spion.

The generalization encoded in the rules could be paraphrased as: any constituent that is immediately dominated by a main-verb clause can be fronted. This includes main verbs as well as their arguments and adjuncts. The extractability of elements from embedded verb phrases and sentences will not be addressed here, because it is not relevant in this context.

Uszkoreit (1984) also discusses types of complex fronting--without including them in the fragment. The following grammatical sentences could not be generated.

- (10a) Den Brief zustecken sollte der Kurier nachher einem Spion.
- (10b) Einem Spion zustecken sollte der Kurier nachher den Brief.
- (10c) Nachher einem Spion zustecken sollte der Kurier den Brief.

In German, the nonfinite main verb of the matrix clause can be fronted together with some--never all--of its arguments and adjuncts. At first glance, the subject seems to be excluded from these complex frontings:

(11a) ?? Der Kurier zustecken sollte den Brief nachher einem Spion.

(11b) ?? Der Kurier den Brief zustecken sollte nachher einem Spion.

But, as Hubert Haider observed (personal communication), there are actually subjects that can be fronted together with a nonfinite main verb:

(12a) Ein wirklicher Fehler unterlaufen war ihm noch nie.
a real error happened was him so far never
He never made a real error.

(12b) Solch eine Frau begegnen kann auch nur dir.
such a woman meet can also only you
Only you can meet such a woman.

Nonagentive subjects of certain verb classes participate in complex fronting. The exact conditions for frontability are not yet known.

The frontings in (10a-c) and (12a,b) cannot be accommodated by the analysis, because the fronted strings do not form a constituent. Yet, only a single category is permitted as the value of the gap feature. One conceivable solution that has been suggested at different times involves allowing sets or stacks of categories as gap-feature values.

Disadvantages of this strategy are not only a possible increase in the power of the formalism but also the sacrifice of one of the most reliable regularities in German syntax. Just like English topicalization, German fronting has been used by linguists as a constituent test. The following constructions need to be ruled out:

(13a) ?? Den Brief nachher sollte der Kurier einem Spion zustecken.

(13b) ?? Den Brief einem Spion sollte der Kurier nachher zustecken.

(13c) ?? Nachher einem Spion sollte der Kurier den Brief zustecken.

There are additional reasons for preserving the linguistic generalization that only a single constituent can be fronted in German. Only one of these pertains to the topic of this paper. All the examples presented so far indicate that the order among the fronted constituents is just the order they would have if they had not been fronted. Assuming multiple frontings inevitably leads to the assumption of crossing dependencies:

- (14) Nachher einem Spion zustecken sollte der Kurier den Brief ε ε ε .
-

This contrasts with linguists' findings about other languages in which the nested order is preferred (Fodor, 1978; Maling and Zaenen, 1982; Engdahl, 1985). No instances of genuine multiple fronting, as it occurs in Scandinavian languages, exhibit copying of the order of the traces as seems to be required in German. To encode the nested order, Maling and Zaenen proposed a complex gap feature that can be viewed as a push-down stack. The elements of the stack are the categories of the trace constituents. Since, in a stack, the last stored element has to be removed first, no crossing dependencies can be described. The order gets reversed.

The appropriate storage device for copying linear order is the queue. In a queue, the first element stored is the first element to be removed. However, employing a queue-valued feature would still not suffice to encode the ordering restriction exemplified in (14). In a standard GPSG, the traces of the three fronted constituents in (14) are sibling constituents. Since the gap feature assignment--just like all feature instantiations--operates on ID rules, there exists no means for storing the elements in a queue-valued gap feature in the appropriate order.

The only way to impose the proper order on the set of fronted sibling constituents in German is to analyze them again as sibling constituents in their sentence-initial position. Then the same LP rules can be used for determining their order that are also used for determining the order within regular main-verb clauses. I will not go

into the technical details of such a solution because the discussion of the linear order in the next subsection will show that it does not suffice to determine the linear order among the fronted and the unfronted elements of the main-verb clause independently.

For the time being, I conclude that solutions that salvage the single-constituent condition on fronting are more desirable than those that give it up. The analyses of Johnson (1983) and Nerbonne (1982, 1984) are based on this condition.

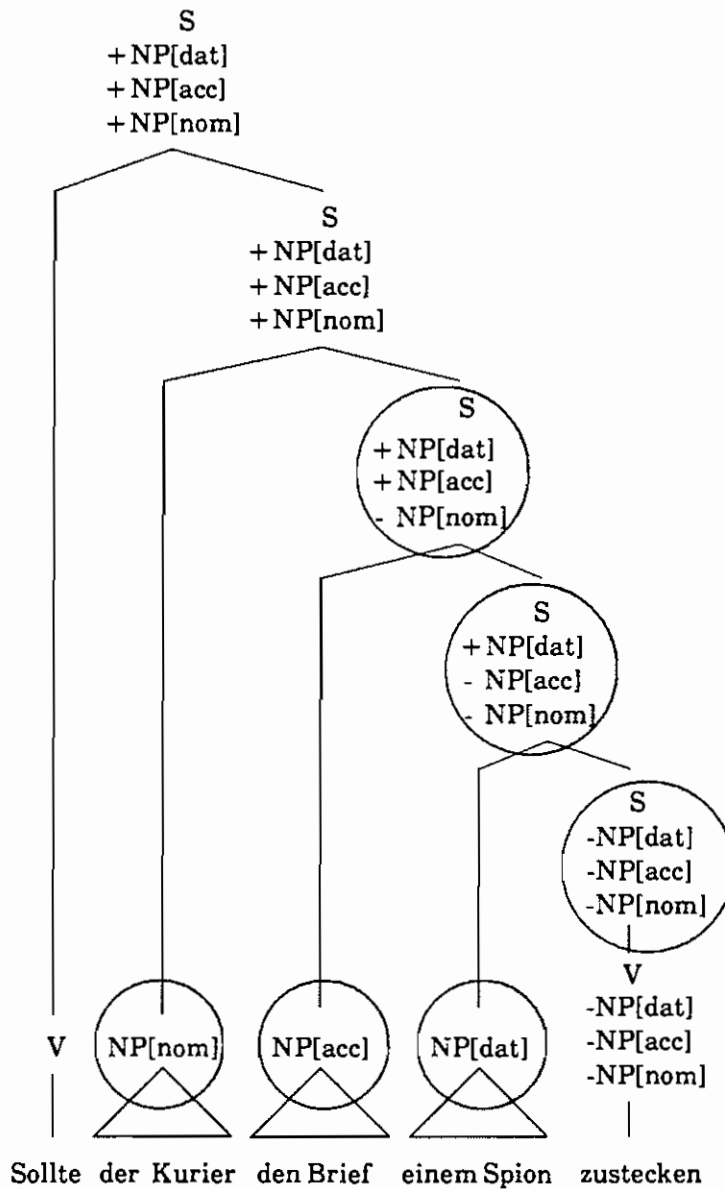
3.2. Branching Clause Structures

Johnson's and Nerbonne's solutions are very similar in their proposed constituent structures. Both provide rules that generate the appropriate constituents for simple and complex fronting. In both approaches, partial verb phrases are created by imposing a binary rightbranching structure on verb-final VPs. Nerbonne's grammar contains these rules in addition to the rules that generate the flat structures. This approach resembles the strategy of *restructuring* or *reanalysis* as it is used in several versions of transformational grammar. Nerbonne relates the rule systems to each other by means of a set of metarules. Johnson's grammar only generates the binary rightbranching structures that are used for complex fronting.

The way in which the binary-branching structures are obtained is similar to the basic strategy of categorial grammars. However, the argument positions of the verb are encoded in a set of special features. Johnson's analysis differs from the standard categorial approach in that his "functors," i.e., the verbs, can combine with their arguments in any order.

The following example of binary-branching clause structure differs from Nerbonne's and Johnson's structures only in the assigned node labels:

(15)



The circled constituents can be fronted. The second-highest S node is not frontable. It not only contains an agentive subject, but it also covers all material after the finite verb.

3.3. The Incompatibility of Analysis and Framework

Unfortunately, both Nerbonne's and Johnson's proposals are incompatible with an essential component of the analysis of German in Uszkoreit (1984). Because the arguments and adjuncts of the verb cannot be sibling constituents in the

binary-branching structures, LP rules cannot be used to linearize the middle field. Only the verb itself can still be positioned by LP rules. The LP rules that refer to the verb merely need to be extended to refer to the partially saturated verb-phrase or clause categories. The complex LP rules, however, which, for the first time, permit the formal description of the partially free order among arguments and adjuncts, cannot be applied.

As a possible remedy, one might consider reformulating LP rules as conditions on the application of (ordered) phrase-structure rules. LP rules can be viewed as regular expressions. The application of the rules to strings of symbols is equivalent to forming the intersection of the regular expressions denoted by the LP rule and the string. Since the regular expression of the string denotes a unit set, the intersection can, depending on the success or failure of the rule, either be the unit set containing the string itself or the empty set. Because of this property of the LP rules, it is certainly possible to encode the LP statements in the feature system and use it in strictly unidirectionally branching trees for enforcing the ordering conditions that are embodied in the original LP component.

However, there is a serious problem with this approach. It is an essential property of the ID/LP mechanism that only sibling nodes are ordered with respect to each other. This property represents an important linguistic insight. Without this premise, rules that are as simple as LP rules could not be used to impose the right order on sentences.

To simulate this property in a formalism that reencodes LP rules in the feature system one would have to partition the tree into domains that correspond to the subtrees of Depth 1 in a traditional phrase structure. In such a grammar, the features that contain the LP information need to be flushed between the domains.

A further complication arises through the different positions of main verbs in German clauses. The main verb can precede or follow the string of arguments and adjuncts. This leads to right- and left-branching structures. Accordingly, the LP rules that apply to the order within the main-verb clause have to be encoded twice in the feature system.

It can be concluded then that it is worthwhile to search for a solution that preserves the application of LP rules to the linearization of sibling constituents.

It was mentioned earlier that Nerbonne proposes to provide both flat and branching structures. Further restrictions that prohibit the occurrence of branching structures in the middle field could be imposed to ensure that the LP rules always regulate the order among arguments and adjuncts. Binary branching structures would be confined to fronted constituents. Such a solution would predict that the order among the fronted arguments and adjuncts is not governed by the ordering principles encoded in the complex LP rule (8). This prediction is wrong. The following pairs of sentences demonstrate that the LP rule has to apply to both the fronted and nonfronted strings of arguments and adjuncts. (The question marks indicate that the sentence is less acceptable than the sentences that are not marked.)

- (16a) Der Kurier sollte ihm einen geheimen Brief zustecken.
the courier should him a secret note slip
The courier was supposed to slip him a secret note.
- (16b) ? Der Kurier sollte einen geheimen Brief ihm zustecken.
- (17a) Ihm einen geheimen Brief zustecken sollte der Kurier.
- (17b) ? Einen geheimen Brief ihm zustecken sollte der Kurier.

Even if one used the mighty feature system to enforce flat structures within the fronted constituent, i.e., to restrict the branching structures to exactly the rules that introduce the gap, one would not get the right results. As one might expect, the ordering principles affect the entire string of arguments and adjuncts, even if the string is interrupted by fronting.

- (18a) Darum hatte der Spion ihn dem Kurier aus der Tasche gezogen.
therefore had the spy him the courier out the pocket
Therefore the spy had slipped it out of the courier's pocket.
- (18b) ? Darum hatte der Spion dem Kurier ihn aus der Tasche gezogen.

(19a) Dem Kurier aus der Tasche gezogen hatte der Spion ihn.

(19b) ? Ihn aus der Tasche gezogen hatte der Spion dem Kurier.

At this point, I will stop envisioning more hypothetical analyses in the current model of GPSG and rather proceed to a proposal of a framework that combines the advantages of the categorial-grammar solution with the ones introduced by the enriched LP component.

4. The Proposed Framework

4.1. The Lexicalization of Linear Order

One of the most striking developments in current syntactic theories might be called the *lexicalization of syntax*. An increasing number of regularities that had been described in syntactic rule systems are treated now as lexical phenomena. One of the best examples is the so-called relation-changing rules, i.e., rules that inflict syntactic valency changes.

Thus, the lexicon should not be viewed as a mere list or data base but as a highly complex component of the grammar whose structure represents diachronic, as well as productive synchronic processes. Such a concept of the lexicon provides an attractive model for processes that are neither historically completed nor fully productive, including many processes that exhibit lexical exceptions.

Uszkoreit (1985) proposes to apply the LP rules in the area of syntax that is ascribed to the extended lexicon. It is assumed that the lexicon consists of a data base component that contains uninstantiated or partially instantiated entries. The productive subcomponent of the lexicon contains rule systems that fully instantiate the entries in accordance with syntactic, semantic, pragmatic, and phonological constraints. Three rule sets in this subcomponent are: *feature instantiation rules*, *valency instantiation rules*, and *order instantiation* or (*complex LP*) *rules*.

This is the simplified uninstantiated lexical entry for the German verb *erzählen*:

(20a) {Sem:erzählen', TR:Agens, (TR:Ziel), (TR:Thema), (TR:Time),...}

The entry lists the thematic roles of the verb. For *erzählen*, only the agent phrase is obligatory (in the active voice); for *zustecken*, agent, goal, and theme are obligatory. The diathesis rules of Wunderlich (1984, 1985) or similar rules that determine the valency on the basis of thematic roles and idiosyncratic lexical properties serve as valency instantiation rules.

Feature instantiation rules add the features for the verb and for the arguments and adjuncts. Order instantiation rules linearize the entry. Here is a simplified instantiation for the verb *zustecken*:

(21)

V	NP	NP	NP	
MC: +	Case: acc	Case: nom	Case: dat	
Finit: +	TR:Theme	TR:Agent	TR:Goal	
Person: 3	Person: 3	Person: 3	Person: 3	
Plural: -	Plural: -	Plural: -	Plural: -	
Focus: -	Focus: -	Focus: -	Focus: +	
Lex:erzählt	PPron: +	PPron: -	PPron: -	
Sem:erzählen'				

The ordered and fully instantiated list--at least the sublist containing the arguments and adjuncts--can now be viewed as a subcategorization stack, just as the *subcat* stack in Head-Driven Phrase Structure Grammars (Pollard 1985), the *syncat* list in PATR-II (Shieber et al. 1983), or the sequence of arguments that are encoded in the complex symbols of categorial grammars.³

4.2. Additional Evidence

The categorial approach to subcategorization has a number of advantages. Johnson (1983) notes the position of the finite auxiliary verb in double-infinitive subordinate clauses as additional evidence for the branching structure. In sentences of this type, the finite verb form precedes the other verbs instead of following them as in regular subordinate clauses.

(22a) ..., weil er sie hätte sehen können.
 ..., since he her had see could
 ..., *since he could have seen her.*

(22b) ? ..., weil er sie sehen gekonnt hätte.

Often, the preposed finite verb form also precedes other material, such as functional verb complements (23a), directional adverbials (23b), or, in rare cases, even objects (23c).

(23a) ..., weil er sie hätte aus den Augen verlieren können.
 ..., since he her had out the eyes loose could
 ..., *since he could have lost sight of her.*

(23b) ..., weil er sie hätte nach Hause begleiten dürfen.
 ..., since he her had home accompany may
 ..., *since he could have been allowed to accompany her home.*

(23c) ..., weil er in diesem Buch hätte eine Antwort finden können.
 ..., since he in this book had an answer find could
 ..., *since he could have found an answer in this book.*

In a flat verb phrase or clause, the preposed element has to be inserted between sibling nodes. In the right-branching structure, the preposed element always precedes a whole constituent. An analysis of the latter type that seems compatible with the proposed framework can be found in Johnson (1985).

5. Implications of the Proposal

An obvious advantage of the proposal lies in the reduction of the rule system. In a traditional GPSG, there has to be at least one verb phrase or clause rule for each subcategorization frame. In the proposed model, one general rule schema suffices:

$$(24) \quad S \langle a_1, \dots, a_n \rangle \rightarrow a_0, S \langle a_0, a_1, \dots, a_n \rangle$$

Since the lexical approach to linear order obviates the need for the coexistence of flat and branching clause structures, the multitude of additional ID rules and metarules proposed by Nerbonne (1983) is obsolete.

In the suggested analysis, the fronting of nonfinite main verbs is covered as a subcase of the fronting of unsaturated clauses since the verb, by itself, is a clause that is still missing all of its arguments. The important generalization about fronting in declarative matrix clauses is easily formulated: Frontable constituents need to be immediately dominated by a clause node whose lexical functor is the main verb.

The proposed approach to linear order supports simple solutions to a number of additional phenomena. I will address here some phenomena that involve the permutation of arguments and adjuncts.

Uszkoreit (1984) assumes phrase structures in which arguments and adjuncts are siblings of the verb. These structures, together with the complex LP rules, allow for the grammatical permutations of arguments and adjuncts. Uszkoreit (1984) also discusses several ways to obtain the flat structures. The two predominant

approaches are the derivation of the relevant rules through metarules and the use of the Kleene-star notation in phrase-structure rules.

For the current proposal, the main difference between English and German in the linear order of the arguments and adjuncts lies in the LP component. In both languages, arguments and adjuncts are elements of the valency list. That arguments and adjuncts permute more frequently in German follows from the absence of LP rules that order arguments and adjuncts with respect to the major-category feature such as NP < PP and from the different weighting of LP clauses.

Since such "movement" phenomena as heavy-NP-shift, focus-shift, particle movement, etc. are analyzed as linearizations through LP rules, they cannot be the source of discontinuity. I am certain that this can be extended to new approaches to right-extractions of postnominal modifiers as they are discussed in Kay and Sag (1985), Stucky (1985), and Wittenburg (1985).

I will conclude with some remarks about the generative power of the resulting formalism. The original framework of GPSG only generates context-free languages. Moreover, as utilized in earlier versions of the formalism (Gazdar and Pullum, 1982), every grammar of the framework can be viewed as a definition of a context-free grammar.

The formalism proposed here allows for non-context-free grammars. Since there is no upper bound on the length of valency lists, the grammars are not restricted to a finite nonterminal vocabulary. However, the strong generative capacity of the grammars depends on the definition of the grammar rule. The crucial point is the access to the potentially nonfinite component of the categories. Uszkoreit (1985) discusses this question. One of the basic assumptions about the valency lists or stacks of a grammar is that they form a regular language. This follows from the form of the uninstantiated lexical entries and from the power of the lexical instantiation rules. Uninstantiated lexical entries are unordered. They contain a symbol for each thematic role. Thematic roles can either be filled by a unique constituent or by an unbounded number of constituents. There are four types of thematic roles, which are exemplified by the underlined phrases in the example sentences below: the ones that have to be uniquely filled (25a), the ones that may be uniquely filled (25b), the ones

that have to be filled at least once (25c), and the ones that can be filled any number of times (25d).

- (25a) Der Kurier hatte den Brief aus der Tasche gezogen.
the courier had the letter out the pocket slipped
The courier had slipped the letter out of the pocket.
- (25b) Der Spion schrieb einen Brief.
the spy wrote a letter
- (25c) Er wohnt in Frankreich (in Paris in einem Vorort in einer alten Villa...).
he lives in France (in Paris in a suburb in an old villa...)
- (25d) Der Spion traf ihn in Berlin.
the spy met him in Berlin

The three types of instantiation rules together intersect the regular language represented by the uninstantiated entry with another regular language. Therefore, the result is a regular language.

The power of the formalism depends on both the language formed by the valency lists and the way the rules may access these lists since they are the only nonfinite parts of the complex symbols.

The rule schema (24) need only read the top element of the stack. At the same time, this element gets removed from the stack. If this is the only way in which rules may access the stack, especially if rules cannot add to the stack, the generated languages are context-free. It is simple to see why this is the case: for the purpose of the rule application, every stack may be encoded in a single symbol that corresponds to a state in a finite state automaton that recognizes well formed valency lists.

However, it might be necessary to allow rules limited write access to the stacks. Rules might be considered that append stacks in a restricted way to accommodate functional composition. Further research is needed to determine which other types of stack access are needed to describe natural languages and how they influence the generative capacity of the formalism.

Footnotes

*The research reported in this paper has been made possible by a gift from the System Development Foundation.

¹For a discussion of these principles, see Uszkoreit (1984).

²For exact definitions and for a detailed description of the revised LP component and its theoretical implications, please see Uszkoreit (forthcoming).

³Since the verb does not necessarily precede or follow the other elements in all languages, and since there are other functor categories like conjunctions that combine with arguments at both sides, a more sophisticated feature encoding is needed for the subcat stack. One could either attach to each element an additional feature that indicates on which side of the functor the argument is to be found or one could split the linearized entry into the functor itself and into a left- and a right-subcat list. The former strategy goes back to various categorial grammars and was recently adopted for head grammars by Pollard (1984); the second one is employed in categorial grammars that I have been implementing on the PATR-II system.

References

- Bach, E. (1962) "The Order of Elements in a Transformational Grammar of German," *Language* 38, 263-69.
- Bierwisch, M. (1971) *Grammatik des deutschen Verbs*, *Studia Grammatica* 2, Akademie Verlag, Berlin.
- Bresnan, J. (Ed.) (1982) *The Mental Representation of Grammatical Relations*, MIT Press, Cambridge Massachusetts.
- Engdahl, E. (1985) "Interpreting Questions." In: D.R. Dowty, L. Karttunen, and A. M. Zwicky, *Natural Language Parsing*, Cambridge University Press, Cambridge.
- Fodor, J. D. (1978) "Parsing Strategies and Constraints on Transformations." *Linguistic Inquiry* 9.
- 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.
- Gazdar, G. and G. K. Pullum. (1982) "Generalized Phrase Structure Grammar: A Theoretical Synopsis." Indiana Linguistics Club, Bloomington, Indiana.
- Johnson, M. (1983) "A GPSG Account of VP Fronting in German." University of California at San Diego (unpublished mimeo).
- Johnson, M. (1985) "The Double Infinitive Construction in Dutch and German." Stanford University (unpublished mimeo).
- Kay, M. (1985) *Unification Grammar*, Technical Report, Xerox Palo Alto Research Center, Palo Alto, California.
- Kay, M. and I. A. Sag (1985) in this volume [fill in title when known]
- Lenerz, J. (1977) *Studien zur deutschen Grammatik*. Volume 5: *Zur Abfolge nominaler Satzglieder im Deutschen*. Narr, Tübingen.

- Maling, J. and A. Zaenen (1982) "A PSR Treatment of Unbounded Dependencies in Scandinavian Languages." In: P. Jacobson/G. Pullum (eds.) *The Nature of Syntactic Representation*, Reidel, Dordrecht, Holland.
- 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.
- Pollard, C. (1985) "Phrase Structure Grammar Without Metarules." In: Proceedings of the Fourth West Coast Conference on Formal Linguistics, University of Southern California, Los Angeles, 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 Categorical Grammars and Natural Language Structures, May 1985, Tucson, Arizona.
- Sag, I. A. (1985b) "Continuous Analysis of Discontinuous Constituents." in this volume.
- Shieber, S. M., H. Uszkoreit, F.C.N. Pereira, J.J. Robinson, and M. Tyson (1983) "The Formalism and Implementation of PATR-II." In: Research on Interactive Acquisition and Use of Knowledge, Artificial Intelligence Center, SRI International, Menlo Park, California.
- 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.
- Stucky, S. (1985) in this volume [fill in title when known]
- 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. (1985) "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. (forthcoming) "Constraints on Order." To appear in *Linguistics*.
- Wittenburg, K. (1985) in this volume [fill in title when known]
- 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.