Grammar Formalisms
and the Description of Word Order Variations

Alexandr Rosen
Institute of Theoretical and Computational Linguistics
Charles University, Prague, Czech Republic
alexandr.rosen@ff.cuni.cz

Abstract

After general remarks on the relationship between word order and its account in linguistic formalisms the report presents an overview of some typical phenomena related to the linear order of words, classified according to several criteria: degree of variability, source of constraints, and extent of syntactic domain. The phenomena are illustrated by examples from German, English and Czech. Next, some existing solutions to the problems of representing discontinuous constituents (crossing dependencies) and constraining word order are presented, mainly within monostratal unification-based frameworks. Merits of the solutions are assessed by comparing their coverage with the typical examples.

Contents

1 Introduction

The order of words, or of strings of words, in a sentence is one of the vehicles for conveying the content of a message, the other such vehicles being the choice of words, morphological and syntactic categories, syntactic structures, intonation and stress, etc. All these ingredients mutually interact, and thus a specific word order of a sentence results from the functioning of a number of factors.

The factors are of various types: phonological, syntactic, semantic and pragmatic, and the relative weight of the types in various languages differs. In configurational languages, such as English or Chinese, where grammatical relations are expressed primarily by fixed positions of complements with respect to their heads, the role of syntactic factors is predominant. In languages where grammatical relations are expressed by affixation, such as Latin or Russian, pragmatic (or discourse-related) factors play a much more important role.

Constraints arising due to the presence of discourse factors are embodied especially in the distinction between that part of an utterance referring to 'old', 'given', and 'already known' facts, and the other part introducing new information, i.e. the distinction between topic and focus. In most languages and in the unmarked case, the topic part of an utterance tends to precede the focus part. Within the topic and focus dichotomy, the ordering of elements is partially conditioned by a finer hierarchy reflecting their status in the discourse.

The effects which the 'syntactic' and 'discourse' factors have on linear order within a sentence are to a certain extent isomorphic. This fact is reflected in some grammar frameworks, where similar mechanisms are used to model the effects of both factors[1]. However, mainly for historical reasons, the word-order phenomena related to discourse factors have received much less attention than the much less variable effects of syntax. Indeed, most grammar frameworks have been built around versions of context-free phrase-structure grammar which implicitly rely on a single basic (syntactically determined) word order. The tension between the rigid constituency-based tree structure and various word-order phenomena present a real challenge especially for frameworks which do not admit several separate levels of representation.

Thus, since the beginning of the 1980s, a number of proposals have been put forward to reconcile the monostratal constituency-based frameworks with linear order phenomena in a more or less principled fashion. Gazdar (1978) proposed to treat unbounded dependencies by slashed categories and Gazdar & Pullum (1981) introduced the rules of linear precedence. Both proposals became major features of the Generalized Phrase-Structure Grammar (Gazdar *et al.* 1985). Kay (1985) and Karttunen & Kay (1985) used regular expressions to impose ordering constraints on constituents in English and Finnish within the formalism of functional unification grammar. In Pollard & Sag (1987), linear precedence (LP) rules were related to the obliqueness hierarchy of complements and thus behaved according to specific subcategorization requirements.

With the LP component, the treatment of word order assumed the position of another dimension of linguistic description. However, in their original shape, LP rules were insufficient for expressing the interaction of the various factors having influence on word order in languages less configurational than English. The proposal of complex LP rules (Uszkoreit 1984) partially remedied this defect. The problem remained of how to apply LP rules to nonsibling nodes when other than flat structure analysis was assumed as necessary and of how to impose constraints whose interactions cannot be expressed by means of a simple hierarchy. Various solutions have been suggested in Reape (1989, 1990), Oliva (1992a,b), Uszkoreit (1991), Engelkamp *et al.* (1992), Uszkoreit *et al.* (1994), and elsewhere.

Since the description of word order regularities within monostratal constituency-based frameworks has begun to assume a proper position, the role of discourse-related factors has been realized, which

[1] The slash mechanism is used both for topicalization and *wh*- questions/relative clauses.

opened the way for attempts to incorporate insights concerning the topic-focus distinction and the theory of communicative dynamism (see, e.g., Sgall *et al.* 1986). It seems, however, that more research is needed in the formal aspects of a possible union as well as more empirical data (e.g. from psycholinguistic experiments) to verify the hypotheses.

In short, the aim of this study is to review various proposals which attempt to solve the problem of reconciling constituent-based descriptions with the variability of word order. Due to the constraints of time, a bias towards HPSG-inspired frameworks will be obvious: very little will be said about polystratal constituency-based approaches, such as various incarnations of GB theory. An attempt to provide a broader perspective will be made by comparison to an approach avoiding constituency. In Section 2, an attempt to characterize and classify various phenomena related to word order is presented, and in Section 3 these phenomena are exemplified. In Section 4, several solutions are presented and their merits evaluated.

2 Preliminaries

There seem to be two sorts of problems which have to be solved by any formalism in order to cope with variable word order. Both are connected with the crucial issue of mapping the surface string of phonemes/characters to an abstract representation:

- Some word order variations may be difficult to derive and represent because they give rise to discontinuous constituents[2] or crossing dependencies (non-projective structures)[3].
- If we assume that a framework allows for some flexibility of word order and that word order is never *completely* free, constraints on word order taking into account relevant factors have to be applied.

More specifically, frameworks based on a traditional (i.e., English-oriented) notion of constituency which assume a single basic order of constituents and employ context-free phrase-structure grammar run into the following difficulties:

- Order variations going beyond constituent boundaries (perhaps bounded in a larger constituent, such as clause) result in discontinuity of such constituents and additional mechanisms have to be employed to relate the 'dislocated' part of the constituent to its 'proper' position predicted by the basic order.
- Order variations among sister constituents lead to a multiplication of rules with the resulting loss of generality. If order among non-sister constituents is not fixed or mutually independent, some other mechanism must be used to constrain it.

Grammar frameworks which deny constituents as the basic syntactic units, but maintain a tree structure (i.e. dependency grammars prohibiting crossing dependencies) are in some respects more flexible in dealing with word order variations than constituency-based approaches:

- Order variations among immediate dependents can be compared to order variations among sister constituents with the important difference that the dependents always include all complements

[2]Frameworks may differ in what they regard as a constituent with the result that what is discontinuous in one framework may not be discontinuous in another.

[3]A tree is non-projective if, in a linearized order, a node follows its head while the head is preceded by a dependent of that node, or if a node precedes its head while the head is followed by a dependent of that node.

and adjuncts of the head, which is not their sister. Problems of dislocating and ordering complements and adjuncts having a common head but scattered over different constituents do not arise[4].

- If the boundary of a subtree is transgressed, a way of relating the dislocated element to its head has to be found as in the case of constituency-based approaches.

Yet other frameworks admit discontinuity or crossing dependencies into the very skeleton of grammar description. This approach makes the problems of *representing* the related linguistic phenomena rather easy, perhaps too easy, and the question remains whether the price paid by such unconstrained formalisms is not too high. In the following discussion I will refer to the more constrained approaches, unless specified otherwise.

Next, I will attempt to classify the types of phenomena related to word order according to three aspects: (1) degree of variability, (2) source of constraints (prevailing type of factors involved), and (3) size of syntactic domain affected[5].

1. **Degree of variability**, i.e., how many different orderings are the result of the functioning of factors responsible for a given phenomenon:

   No variability – position of syntactic units in the surface string is fixed (Wackernagel's position of clitics in some Slavic languages, clause-final position of verb in Japanese).

   Limited variability – position of syntactic units in the surface string is variable, but restricted (position of finite verb in German clauses, the order of subject and verb in English).

   Free variability – under specific conditions, all permutations of syntactic units are possible (order of complements and adjuncts in the German middle field).

2. **Source of constraints**, i.e., what type of factors is responsible for the ordering pattern:[6]

   Phonology – ordering depends on some phonological factors such as intonation, stress, but also length (Heavy-NP shift in English is triggered by the length of the NP string).

   Syntax – ordering is conditioned by (morpho)syntactic factors, such as syntactic function (or morphological case). In highly configurational languages this is the richest source of word order constraints, but there is evidence for a general unmarked ordering pattern also in other language types.

   Pragmatics – the ordering pattern is constrained by discourse-related factors, such as the distinction between items presumed by the speaker (writer) to be already known by the hearer (reader) and items introduced as new facts.

[4]In this respect, a dependency tree is equivalent to a flat constituent structure, although in a dependency tree the position in the linearized order of the head (= governor) relative to its dependents is not obvious. (According to the theory of Functional Generative Description, see Sgall *et al* 1986, main verb, i.e. the head of a clause, is usually positioned at the border between contextually bound and nonbound elements.)

[5]Another distinction can be made: the location of a syntactic unit in the surface string is determined either with respect to other units, or with respect to the beginning or end of the clause (sentence). Most word order phenomena can be stated in terms of the former, i.e. relative to other units. The latter view of the phenomena has been called *topological* and can be applied to the description of the German clause types or Wackernagel's position of clitics.

[6]This is not an exhaustive list, at least two other types of factors could be mentioned: semantic and stylistic. Semantic factors are at work especially in nominal collocations like *Attorney General*. Obsolete or bookish style is often marked by a word order variation.

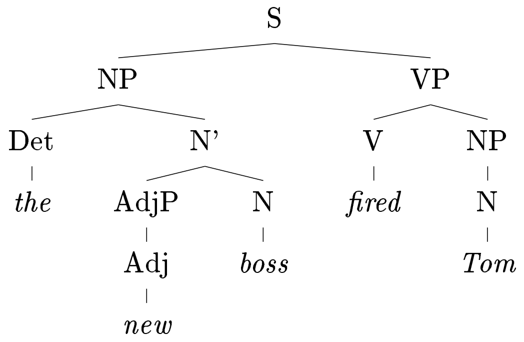


Figure 1: A constituency tree.

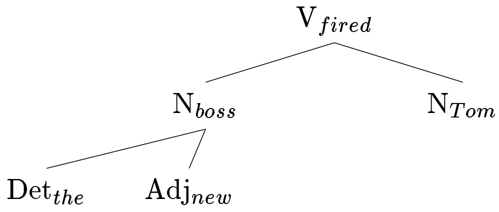


Figure 2: A dependency tree.

3. **Extent of syntactic domain**, i.e. how wide (and deep) is a section of the syntactic structure within which the ordering pattern is manifested, or where are the boundaries for the type of variation[7]:

Local tree is a domain which corresponds to sister constituents in a phrasal approach, or to a head with its dependents in a dependency approach. Thus, the domain is very much theory-dependent, and can range from a head with zero or one modifier[8] in binary branching phrase structures to a head with all its modifiers in flat or dependency structures. Speaking about local trees therefore makes sense only in the context of a specific grammar framework. In Figure 1 the local trees are: NP VP, Det N', AdjP N, V NP; in Figure 2: N_{boss} V N_{Tom}, and Det Adj N_{boss}. Word-order variations in local domains do not result in discontinuous/non-projective structures.

Head domain of a (lexical) head consists of the (lexical) head and all its modifiers (or, strictly speaking, their mother nodes). In a dependency approach, head domain is equal to local tree. In Figure 1 the head domain of the N *boss* is: Det AdjP N_{boss}, and the head domain of the V *fired* is $NP_{thenewboss}$ V NP_{Tom}[9]. In constituency-based approaches word-order variations in head domains can lead to discontinuous constituents. In dependency trees non-projectivity does not occur when elements in a head domain are permuted.

Recursive head domain of a (lexical) head includes the (lexical) head and recursive head domains of its modifiers. Thus, it should include only lexical heads. In Figure 1, as well as in Figure 2, the recursive head domain of the verb *fired* includes Det, Adj, N_{boss}, V and N_{Tom}.

In most cases, however, the extent of a domain often exceeds the head domain but does not reach the full depth of the recursive head domain. Therefore, the term will be used for

[7] Here, the term *domain* will cover both bounded and unbounded phenomena.
[8] The term *modifier* will henceforth be used to cover both complements and adjuncts.
[9] Let me assume here that the subject NP and the determiner are complements.

such non-fully recursive head domains in contrast to **fully recursive head domains** and head domains. In fact, the following two aspects can be used to make further distinctions within the class of recursive head domains: (a) depth of recursion, and (b) boundedness.

(a) **Depth of recursion**:

arbitrary – i.e., from the competence point of view (all cases of unbounded dependency)

restricted – head domain is exceeded only slightly and/or there is a limit on the level of embedding (comparison)

(b) **Extent of boundedness** – this distinction can be made within both fully and partially recursive head domains. Boundedness of the ordering pattern in a specific syntactic structure means that no element of the structure may occur outside the structure. Unless specified otherwise, finite clause (without embedded finite clauses) will be assumed as the default syntactic structure for the condition of boundedness, but ordering phenomena bounded within other structures also exist:

noun phrase

non-finite clause

finite clause – unbounded dependencies in German (with some exceptions)

unbounded – English topicalization or *wh*-clauses

In the following section, examples illustrating a fraction of phenomena related to word order will be presented.

3 Some data

Examples will be given in an order which should reflect the increasing complexity of the ordering pattern[10]. The distinction according to the extent of the domain will have precedence over the source and (closely connected) variability distinctions. A binary branching structure will be assumed to support the class of local domain phenomena.

As it is difficult to make the classes clear-cut, the above ordering can only roughly correspond to the grouping of examples into the following subsections:

1. Trivial cases (local domain): syntactic constraints (determiner $\prec$ noun in English), syntactic/pragmatic constraints (adjective – noun in Czech).

2. Head-domain phenomena: largely syntactic constraints (subject – verb in English), pragmatic constraints (order of verb modifiers in Czech).

3. Bounded recursive-head-domain phenomena: phonological constraints (heavy-NP shift in English), syntactic constraints (clitics in Czech), pragmatic constraints (order in the German middle field).

4. Unbounded recursive-head-domain phenomena: syntactic constraints (*wh*-clauses in English), pragmatic constraints (topicalization in English).

[10]This order corresponds to the order in which the classes partitioning the individual aspects have been introduced in the previous section.

3.1 Trivial cases

There is a number of ordering constraints from which the backbone of most constituency-based frameworks is derived. These constraints concern local trees, have no variability and are syntactically determined. The less word-order freedom in a language, the more such constraints. See (1) for some English structures which are impossible in any other order.

(1) a. the picture

b. in the picture

c. the picture of a horse

d. the picture (which) I want

e. to say

f. is said

g. is not said

h. has been said

i. could have been said

j. say that I am sleeping

k. say what you want

Similar examples could be given e.g. for German, where, however, the order of auxiliaries is reversed (in the verb-final sentence type, see (2)).

(2) gesagt worden sein können
said been have could

But also in Czech, a free word-order language, similar examples can be found:

(3) a. v krabici
in box-LOC

b. krabice, kterou chci
box-NOM which-ACC want-1SG

c. řekni, že spím
say-IMP that sleep-1SG

d. řekni, co chceš
say-IMP what want-2SG

e. přestat chtít začít
to stop to want to start
'to stop willing to begin'

f. začít chtít přestat
to start to want to stop
'to start willing to stop'

There are many invariant word-order patterns in Japanese, which is, besides its verb-final and topic-initial requirements, another free word-order language. In (4), no other permutation of syntactic units is possible. The ordering constraints are still confined to local domains.

(4) Yukiko san to wa dēto shita dake desu.
Yukiko Ms. with TOP date did only is
'I just dated Yukiko, that's all.' (MT89)

Some orderings are strongly preferred, see (5) for examples of adjective $\prec$ noun and noun $\prec$ nominal attribute. However, in frozen expressions (5d) or special (obsolete) style (5e) the order can be different.

(5) a. velká krabice
big-NOM box-NOM

b. krabice sušenek
box-NOM biscuits-GEN

c. v té velké krabici
in that-LOC big-LOC box-LOC

d. Spojené státy americké
united-PL-NOM states-PL-NOM american(adj)-PL-NOM
'the United States of America'

e. našeho rodu kořeny
our-SG-GEN family-SG-GEN roots-PL-NOM
'the roots of our family'

3.2 Head-domain phenomena

Similarly invariable as the local ordering patterns illustrated in (1) are some head-domain orderings[11]:

(6) a. buy Mary a present

b. put it on the shelf

c. put on the shelf the book I got yesterday about Max and Moritz

d. put on the shelf the **Bible** and not the other book

If *it* in (6b) is replaced by a stressed or phonologically 'heavy' noun phrase (6c and 6d), the order of complements can be reversed. In fact, most ordering patterns in this subsection exhibit less invariance than those restricted to local domains (such as the subject – verb order in the English example (7d),

[11] Please bear in mind that binary branching constituent structure is assumed to differentiate between local and head domains.

here the usual subject $\prec$ verb order is also possible). Nevertheless, some are considered stable enough to further support the foundations of standard English-oriented frameworks. In English, these cases include the order of (non-subject) complements relative to other complements and to the verb (7a), and the order of subject and verb within a specific clause type (declarative, interrogative) (7b, 7c).

(7) a. buy a present for Mary

b. John said: "I love Mary."

c. "I love Mary," John said.

d. "I love Mary," said John.

The position of **adverbial particle** of English complex verbs is still less restricted:

(8) a. look up the number

b. look the number up

On the other hand, in verb-initial or verb-second German clauses with verbs which require a **separable prefix**, the prefix is obligatorily placed at the end of the clause:

(9) a. Wir bringen ihn morgen gegen 5 Uhr am Tennisplatz um.

b. Er sieht nur selten fern.

A mixture of pragmatic and phonological factors determine the ordering in cases which involve **extraposition**[12]. Extraposition to the end of a clause restricted to head domain of a verb means "movement" of a complement of the verb out of its structurally predicted position over another element, complement or adjunct. A typical example is the heavy-NP (also heavy-PP) shift, i.e., the extraposition of a phonologically long or strongly focussed noun (prepositional) phrase:

(10) I merely wanted to bring back into the discussion the evidently quaint notion that there are ideas at stake in linguistics, which play a major role in hiring decisions. (D. Pesetsky in Linguist List 940317)

Indeed, other categories like prepositional phrases, verb phrases and clauses can also be extraposed under similar conditions[13]:

(11) a. I talked to John for an hour about this problem. (KJ87)

b. They tried for many years to force him to make a decision one way or the other, but to no avail. (St87)

c. I told John yesterday that I wanted pizza. (KJ87)

[12]Traditionally, only cases where the extraposed item was replaced by a pronoun were labelled extraposition. Here, the term will be used rather in the sense of dislocation. Of course, the term implies a canonical order predicted by commonly assumed rules of English grammar.

[13]Interestingly, in (11a) and (11b) the time duration adjunct seems to be strongly focussed, at least in one reading. This supports the claim that there is an unmarked 'systemic' ordering, in the examples corresponding to the placement of the time adjunct after other elements, which, when violated, results in focussing the first element violating the unmarked order.

Extraposition to the left occurs less frequently:

(12) a. Happy I am that Bill came. (St87)

b. Believe I do that Sam is a liar. (St87)

Due to the German clause structure, cases of extraposition in German often go beyond head domains. Whereas the ordering in (13a) can be described within the head domain of *versuchte*, in (13b) it is the recursive head domain of *hat* which has to be ordered[14].

(13) a. ... daß der Mann versuchte, das Buch zu lesen (Rp90)

b. ... daß der Mann versucht hat, das Buch zu lesen (Rp90)

Finally, the **placement of parentheticals** seems to be rather free and confined to the head domain:

(14) a. Of course, the good man talked about his sheep.

b. The good man talked, of course, about his sheep.

c. * The good, of course, man talked about his sheep.

3.3 Bounded recursive-head-domain phenomena

Let me start with some syntax-driven, rather invariable and seemingly highly bounded phenomena. The adjunct *so* and its **degree complement** in (15a) are both in the (recursive) domain of *happy* (and bounded within the adjectival phrase). The degree complement can also be adjacent to its head, as in (15b). However, in (15c), the recursive domain is the domain of the finite verb (*enough* and its complement are bounded within the finite clause).

(15) a. The professors were so happy that tears ran from their eyes.

b. The professors were happy enough to forget about the new systemization scheme.

c. Enough people have responded for us to go ahead with the decision. (St87)

Comparison is a very similar phenomenon. In (16a) the recursive domain is headed by *liquor* (the comparison structure is bounded within the noun phrase). But in (16b) the recursive domain is headed by the finite verb (the structure is bounded within the finite clause). Parallel Czech examples are in (16c) and (16d).

(16) a. Tom serves more expensive liquor than Dick (does).

b. Tom serves more expensive liquor to guests than Dick (does). (Mc87)

c. Panenku Barbie dostávají i menší děti než (je) Mařenka.
doll-ACC Barbie get-3SG even smaller children-NOM than (is) Mary

d. Rodiče dávají i horší věci svým dětem než je panenka Barbie.
parents give even worse things their children-DAT than is doll Barbie

[14]Under the assumption that the main verb is headed by the auxiliary.

The following phenomena can reach much deeper domains. A subclass of structures called 'unbounded dependencies' should also figure in this subsection, namely all such structures where the long-distance dependency does not cross a finite-clause boundary. All types of 'unbounded dependencies' can occur within the finite-clause boundary and some have to. There are two such types: **purpose infinitive** (17a) and **tough constructions** (17b). For the other types, refer to the next subsection.

(17) a. I bought it for Sandy to eat. (Po91)

b. Sandy is hard to love. (Po91)

Clitics in Czech are usually placed at Wackernagel's position, following the first (syntactic) unit of the clause (19a)[15]. Clitics belonging to a non-finite clause may climb to the second position of the main clause (19b). Two identical clitics are usually impossible in a single cluster of clitics (19c) – (19e). There are complex restrictions on the order of clitics in the clitics cluster, which are different from constraints on the order of non-clitics. The first unit need not be a traditional constituent or subtree (see 19f and 19g).[16] Note that (19h), where the first element consists of a verb with its indirect (rather than direct) object is considerably worse[17].

(19) a. Na cestě do školy jsem ji včera nepotkal.
On way to school AUX(cl.) her(cl.) yesterday meet-PAST-NEG
'I didn't meet her on the way to school yesterday.'

b. Boris se odmítá oholit.
Boris REFL(cl.) refuses shave.
'Boris refuses to shave himself.'

c. Boris se rozhodl oholit se.
Boris REFL(cl.) decided shave REFL(cl.).
'Boris has decided to shave himself.'

d. Boris se rozhodl oholit.
Boris REFL(cl.) decided shave.
'Boris has decided to shave himself.'

e. * Boris se se rozhodl oholit.
Boris REFL(cl.) REFL(cl.) decided shave.

f. Loni v létě na Bahamách se mi zdálo všecko tak jednoduché.
Last year in summer on Bahamas REFL(cl.) I-DAT(cl.) seemed everything so easy
'Last summer on the Bahamas everything seemed to me so easy.'

[15] Precise definition of the nature of this unit remains to be specified, see esp. examples (19f) and (19g).

[16] What precisely constitutes the 'first syntactic unit' is rather difficult to define. The very concept of a single syntactic unit is doubtful, as there are clear cases of multiple constituents occurring before a clitic. It seems that, instead of a definition based on syntactic grounds, a definition of the first element in terms of communicative relevance is nearer to real facts. For a proposal along this line, see Avgustinova & Oliva 1995.

[17] In some cases, individual native speaker's judgments may differ and discourse related factors intervene. E.g., some speakers seem to have only weak preference for (19b), as opposed to (18a), or even (18b). (And similarly for (19d)).

(18) a. Boris odmítá se oholit.

b. Boris odmítá oholit se.

g. Pohlídat děti si Novákům troufne jen tetička z Plzně.
baby-sit children-ACC REFL(cl.) Nováks-DAT ventures only auntie from Pilsen
'Only the auntie from Pilsen is brave enough to baby-sit the Nováks' children.'

h. ? Pohlídat Novákům si děti troufne jen tetička z Plzně.
baby-sit Nováks-DAT REFL(cl.) children-ACC ventures only auntie from Pilsen

The example (19g) could as well be classed as a case of **topicalization**. However, topicalization in Czech can be unbounded, so refer to the following section for more examples.

Ordering patterns in German have been dealt with quite extensively in a number of publications (e.g. Uszkoreit 1987, Engdahl & Reape 1990, Reape 1991) and represent an important stimulus for recent developments in the field. Finite-clause bounded phenomena include auxiliary flip, fronting (simple and complex), permutation of syntactic units in the middle field, and extraposition.

Auxiliary flip is a largely syntax-driven phenomenon triggered in verb-final clauses when modal and few other verbs are used in contexts requiring an auxiliary. Although the 'flip' is limited to verb group, recursive head domain is involved (unless the order in the middle field and the verb group is accidently centre-embedding). In (20) two auxiliaries are flipped.

(20) ... daß man ihn hier wird können liegen bleiben lassen. (HN93)

Simple fronting in German (as opposed to complex fronting) involves every major clausal constituent, not necessarily a phrasal one, except finite verb. This includes main verb forms (21a), but not modal verbs (21b vs. 21c and 21d). NP heads, possibly with other parts of NP, can also be fronted (21e, see also 23e), but not always (21f)[18]. Fronting is bounded within a finite clause (constituents from embedded finite clauses cannot be fronted), but is possible from an extraposed infinitival clause (21g).

(21) a. Gegeben hat Peter dem Jungen das Buch heute. (HN89)

b. * Können wird Hans gehen. (KO91)

c. Gehen wird Hans können. (KO91)

d. Gehen können wird Hans. (KO91)

e. (Einen) Wagen hat er sich noch keinen leisten können. (RH90)

f. * Futter wird Cecilia dem Nilpferd dieses *t* morgen geben. (HN93)

g. Über dieses Thema hatte Fritz Peter gebeten, einen Vortrag *t* zu halten. (HN89)

German fronting has also been called 'topicalization', but the fronted element is not necessarily the topic proper (see (22b), where the fronted NP is a contrastive topic), or even a part of the focus (22c).

(22) a. Und was war mit Peter?

b. — Das Fi**nanz**amt hat ihn geschnappt (und nicht die Polizei).

c. — Einem lausigen Schnüffler ist er in die Fänge geraten. (HU87:63-5)

Other examples could suggest that more than one element can be fronted. However, the position that only a single item can be fronted is generally maintained even in such cases of seemingly multiple, or **complex fronting**. In (23a) three adjuncts are fronted[19]. A verb can be fronted together with

[18] This phenomenon has been called **split topicalization**.

[19] See (19f) for a similar pattern in Czech, where the first position in a clause is delimited by clitics.

some of its modifiers, while the order among the fronted elements corresponds to the order they would have if they had not been fronted (23b). A verb cannot be fronted together with its subject, unless the verb is ergative (i.e., the subject does not play the role of Agentive, see (23c)). A part or the whole of the verb group can be also fronted (23d). Similarly as in (21e), a NP complement can be 'split' between the fronted part and the middle field (23e). It is easier to front a verb with its direct object than with its indirect object (23f, 23g)[20].

(23) a. Vom Parkplatz über die Brücke zum Haupteingang läuft man keine 10 Minuten. (HU87)

b. Nachher$_1$ einem Spion$_2$ zustecken$_3$ sollte der Kurier den Brief $\_\__1$ $\_\__2$ $\_\__3$. (HU86)

c. Ein Fehler unterlaufen war ihm noch nie. (HU86)

d. Gefüttert haben (müssen) wird Cecilia das Nilpferd. (HN93)

e. Lösungen für das dritte Theorem gefunden hat er bisher keinen definitiven. (RH90)

f. Das Buch vorgelesen hat er dem Mann.

g. ? Dem Mann vorgelesen hat er das Buch. (HN93)

Ordering in the **German middle field**[21] obeys largely pragmatic factors and exhibits free variability. In (24a), the order of complements corresponds to the unmarked order ('systemic ordering') subject $\prec$ indirect object $\prec$ direct object, but the three complements of *geben* can appear in any order due to the discourse-based source of constraints, which override unmarked order constraints (provided, of course, that an appropriate context precedes). In cases where some complements are pronominalized, non-definite or phonologically 'heavy', even more sources of constraints interact. In (24b) pronominal direct object precedes indirect object and focussed subject follows both. An additional constraint pronominal $\prec$ non-pronominal can be assumed which overrides the unmarked order constraints (i.e., its relative weight is greater, *cf.* the unacceptable example (24c)).

(24) a. Dann wird der Doktor dem Patienten die Pille geben. (HU87:45)

b. Dann wird sie dem Patienten der Doktor geben.

c. *Dann wird dem Patienten sie der Doktor geben.

However, matters are more complex than could be predicted by a simple ordering of constraints expressing precedence relations between single categorial values. In (25a), where both objects are pronominal, the order is direct object $\prec$ indirect object. In (25b–25c) the order of subject and pronominalized object is arbitrary.

(25) a. Dann wird sie ihm der Doktor geben.

b. ... daß es Peter wußte (WoSk)

c. ... daß Peter es wußte (WoSk)

[20] Again, this is paralleled by the Czech examples (19g) and (19h).
[21] Middle field is the part of the finite clause where non-fronted and non-extraposed modifiers are placed.

It seems that it is not possible to subsume all constraints under a single, more general discourse-driven principle. It cannot be assumed that pronouns refer to some already given piece of information and are therefore a part of the topic, in (26b, an answer to 26a) the pronoun is focussed and follows the other complement[22].

(26) a. Und weißt du, wem von den beiden der Schaffner beim Aussteigen geholfen hat?

b. — Beim Aussteigen hat der Stoffel **ihm** geholfen. (HU87:294)

The domain becomes recursive (but bounded within the finite clause) if modifiers of more than one verb are involved. In (27a) the subject of the finite verb is among the non-finite verb complements, in (27b) a structure equivalent to the Dutch cross-serial dependencies (27c) results[23]. However, in (27d), where both verbs can take a dative complement, there is a single (or, at least, preferred) interpretation which does not involve cross-serial dependencies (i.e., 'he promised the man to help the lady').

(27) a. ... daß das Buch Hans dem Mädchen geben wird (KO91)

b. ... daß es ihm jemand zu lesen versprochen hat (Rp90)

c. ... dat Jan Piet de kinderen zag helpen zwemmen (Br82)

d. ... daß er dem Mann der Frau zu helfen versprochen hat (WoSk)

In Czech, there are no 'topological' constraints on the order within such clauses (except for the position of clitics). Other constraints also seem to be much weaker — I find all 120 orderings of (28) at least marginally acceptable. See, however, Oliva (1991), where examples like (29) would be excluded.

(28) Šéf nařídil Honzovi zavřít dveře.
boss-NOM ordered John-DAT to-close door-ACC

(29) Šéf dveře nařídil Honzovi zavřít .
boss-NOM door-ACC ordered John-DAT to-close

Most cases of **extraposition in German** take place within a recursive head domain. Extraposed complements must be *zu*-infinitival VPs and they must not be complements of a raising verb (30).

(30) a. ... daß der Mann versucht hat, das Buch zu lesen (Rp90)

b. ... daß der Mann versucht hat, zu behaupten, das Buch gelesen zu haben (Rp90)

c. ... weil Fritz zu arbeiten versucht/scheint (Rp91)

d. ... weil Fritz versucht/*scheint, zu arbeiten (Rp91)

[22] Another possible candidate for subsumption under the distinction topic/focus might seem to be the definite/non-definite distinction. But again, the two distinctions are orthogonal, as can be seen from (24a), where the focussed direct object can be both definite and indefinite. Another question is what impact on the ordering (relative weight) has the category of definiteness by itself.

[23] Uszkoreit (1987) calls this phenomenon **focus raising**.

e. … weil das Baby sich zu einem Genie zu entwickeln versprach
'… because the baby promised to become a genius'
'… because it was very likely, that the baby would turn into a genius' (Rp91)

f. … weil das Baby versprach, sich zu einem Genie zu entwickeln
'… because the baby promised to become a genius'
'*… because it was very likely, that the baby would turn into a genius' (Rp91)

Cases of **extraposition in English** which go beyond a non-recursive head domain include cases of extraposition from complements and adjuncts. In (31a) and (31b) it is an adjectival complement which is extraposed (a PP and a clause, respectively). The following examples concern extraposition from an NP. In (31c) – (31f) the extraposed item is a relative clause, in (31g) a finite complement clause, in (31h) and (31i) an infinitival clause, in (31j) a PP. In (31k) an adjunct clause is extraposed from a PP.

(31) a. How fond are you of Sally? (KJ87)

b. How certain are you that the Mets will win? (KJ87)

c. Everyone came who I invited. (Ja87)

d. Who came who you invited? (Ja87)

e. I spoke to everyone yesterday who I liked. (Ja87)

f. To whom did you speak who you liked? (Ja87)

g. The assumption was questioned by only a few of those involved that Mary would reject the offer. (St87)

h. Plans have been made to install a new government. (KJ87)

i. The general gave the order to the army to attack. (KJ87)

j. Several books appeared on the shelves with leather bindings. (KJ87)

k. He liked living near (the market) last year where they have a good selection of fish. (St87)

Many constraints have been proposed for the English extraposition and many have been questioned, including the constraint of boundedness within a finite clause (see Stucky, 1987). However, I will not discuss the matter here and continue to assume boundedness.

3.4 Unbounded recursive-head-domain phenomena

The English **unbounded dependency** constructions are typical examples of this class. They include **topicalization** (32a), **wh-questions** (32b), **relative clauses** (32c), **it-clefts** (32d), **pseudoclefts** (32e), purpose infinitives (17a), and *tough* structures (17b)[24]. Examples in (32) are due to Pollard & Sag (1991).

(32) a. Kim, Sandy loves.

[24] The latter two phenomena have been exemplified in the preceding subsection, since they do not go beyond a finite clause boundary.

b. I wonder who Sandy loves.

c. This is the politician (who) Sandy loves.

d. It's Kim (who) Sandy loves.

e. What Kim loves is Sandy.

It is easy to parallel the examples in (32) with more deeply embedded structures, see (33).

(33) a. Kim, everybody believes Tom said Sandy loves.

b. I wonder who John claims Mary thinks Sandy loves.

c. This is the politician (who) you heard people say Sandy loves.

d. It's Kim (who) nobody knows Sandy loves.

e. What I believe Kim loves is Sandy.

In Czech, finite-clause-bounded cases of unbounded dependency (topicalization: 34a – 34e) are rather common:

(34) a. Karla jsme plánovali poslat do Paříže.
Charles-ACC AUX-1PL planned to send to Paris
'Charles we planned to send to Paris.' (MEAN, p.278)

b. Tetě jsme se rozhodli dát růže.
aunt-DAT AUX-1PL REFL decided to give roses
'To Aunt we decided to give roses.' (MEAN, p.278)

c. Nic zlého jsem nemyslil propovědít.
nothing-ACC evil-GEN AUX-1SG think-NEG-PAST to predict
'I have not meant to predict anything evil.' (Šmil)

d. Na takovou glancroli by si moh troufnout ...
for such-ACC important part-ACC AUX-COND REFL could-3SG-MASC venture.
'Such an important part he might be able to act.' (SS94)

e. Slovenský prezident Michal Kováč na páteční tiskové konferenci uvedl,
Slovak president Michal Kováč at Friday press conference stated

že se nepříznivý obraz SR na Západě s příchodem nové vlády
that REFL negative picture SR in West with arrival new government

již daří měnit.
already manages to change
'The Slovak president Michal Kováč stated at the press conference on Friday that since the new government is in power success has been noted in changing the negative picture of the Slovak Republic in the West.' (Caro)

However, examples parallel to (33a) – (33c), where the phenomenon is not bounded within a finite clause, can also be found (36a – 36e)[25]:

(36) a. Karla jsme plánovali, že pošleme do Paříže.
Charles-ACC AUX-1PL planned that send-1PL-FUT to Paris
'Charles we planned to send to Paris.'

b. Tetě jsme se rozhodli, že dáme růže.
aunt-DAT AUX-1PL REFL decided that give-1PL-FUT roses
'To Aunt we decided to give roses.'

c. Tenhle film jsem už nedoufal, že ještě někdy uvidím.
This film-ACC AUX-1SG already hope-NEG-PAST that still ever see-1SG-FUT
'This film I already didn't hope to ever see.'

d. A koho byste řekli, že jsem sehnal na záskok?
and who-ACC would-2PL say-PPLE that AUX-1SG fetch-PAST for substitute
'And whom do you think I fetched at the late notice?' (SS94)

e. Koupil jsem si knihu, kterou jsem netušil,
buy-PAST AUX-1SG REFL book-ACC which-ACC AUX-1SG know-NEG-PAST

že mám doma.
that have-1SG at home
'I bought (for myself) a book which I didn't know I had at home.'

NP heads in Czech can also be topicalized (violating NP-island constraint), as in (37a), or as in the unbounded example (37b). Modifiers of nouns can also be questioned, relativized or topicalized, as in (37c), (37e) and (37g). The examples (37d), (37f) and (37h), at least marginally acceptable, could suggest that the phenomenon is unbounded.

(37) a. Práce to byla galejnická.
work it was slave
'It was slave work.' (MEAN, p.239)

[25] Acceptability judgments may differ in some examples. Furthermore, other factors intervene: (35a) and (36a) are probably better than (35b), and (35c) is even worse. Integration of the "original" embedded verb complement into regular syntactic relationships within the main clause is usually more acceptable, as in (35d):

(35) a. Kdo jste si mysleli, že pojede do Itálie?
who-NOM AUX-2PL REFL think that go-3SG-FUT to Italy
'Who did you think that would/will go to Italy?'

b. Karel jsme si mysleli, že pojede do Itálie.
Carles-NOM AUX-1PL REFL think that go-3SG-FUT to Italy
'Charles we thought that would/will go to Italy?'

c. Karel jsme se rozhodli, že pojede do Itálie.
Carles-NOM AUX-1PL REFL decided that go-3SG-FUT to Italy
'Charles we decided that would/will go to Italy.'

d. O Karlovi jsme se rozhodli, že pojede do Itálie.
about Charles AUX-1PL REFL decided that go-3SG-FUT to Italy
'About Charles we decided that he would/will go to Italy.' (MEAN, p.278)

b. Pivo jsem přece hlásil, že podávají jenom lahvové.
beer AUX EMPH announced-1SG that serve-3PL only bottled
'I did announce that they serve only bottled beer.'

c. Do které chodíš třídy?
to what go-2SG class
'What class are you in?' (MEAN, p.239)

d. Do které sis myslela, že chodím třídy?
to what REFL+AUX thought-2SG-FEM that go-1SG class
'What class did you think I am in?'

e. Zajímavého vám dnes nepovím nic.
interesting-GEN-SG you-DAT-PL today say-1SG-NEG nothing
'Today I have nothing interesting to tell you.'

f. Zajímavého už asi víte, že vám dnes nepovím nic.
interesting-GEN-SG already probably know that you-DAT-PL today say-1SG-NEG nothing
'You probably already know that today I have nothing interesting to tell you.'

g. Zelená není na vojně jenom tráva.
green is not in army only grass
'In the army, it's not just grass which is green.'

h. Zelená mi řekli hned, že na vojně není jenom tráva.
green me told-3PL at once that in army is not only grass
'I was told at once that in the army it's not just grass which is green.'

On the other hand, German unbounded dependencies are usually considered bounded within a finite clause. Counterexamples exist, involving mostly extraction of adverbial phrases, as in (38a). (38b), where a subject is relativized, is acceptable only for some speakers.

(38) a. In welchem Jahr sagte Peter eben, daß er sie zum letzten Mal gesehen hat? (HU87:170)

b. Das ist genau die Lösung, die ich glaube, daß eines Tages gefunden wird. (HU87:171)

The non-existence or restricted acceptability of certain types of unbounded dependency in a language can be viewed as a consequence of the existence of some other means to express the same discourse-related information (such as more word order freedom) or may be simply due to more strict syntactic constraints.

Anyhow, it is clear that while the unbounded dependency constructions have rather limited variability (strong syntactic constraints), the motivation for topicalization and clefts is predominantly pragmatic. The balance is shifted even more towards the syntactic source of constraints in the case of questions and relative clauses.

4 Some solutions

As noted in Section 2, a grammar framework must solve two sorts of problems when dealing with word order variations: (i) the problem of discontinuous or non-projective structures, and (ii) the problem of imposing proper constraints on the linear order given the factors at work.

For the second sort of problems, either an ordering which corresponds to the order of categories in every grammar rule expansion is assumed (any standard CF grammar), or there is a specification of surface linear order, independent from the rule component generating hierarchical structure (binary linear precedence rules, regular expressions constraining surface order).

Three standard strategies for dealing with the first sort of problems are commonly used, very often simultaneously:

1. At some level of description, discontinuity or non-projectivity is admitted in the representation.

2. There are no discontinuities/non-projective structures in the representation. Instead, a multi-level description with different structures is assumed. The connection between discontinuous terminal elements expressed continuously at some higher level is realized by complex relations between the levels. Order in the deep structure can then be defined independently from the surface ordering, e.g., as a basic word order for a given language (GB), or as an order corresponding to degrees of communicative dynamism (Functional Generative Description, see Sgall *et al.* 1986).

3. Different levels of description are collapsed into a single 'monostratal' structure and related by sharing pieces of information. Discontinuous syntactic elements cannot be sisters in a tree (cannot form a constituent), but may be mapped onto a (semantic) representation in which they are interpreted as a unit ('constraint-based formalisms', such as GPSG (Gazdar *at al.* 1985), HPSG (Pollard & Sag 1987)).

In the following, I am going to deal mainly with solutions put forward within the context of the latter (monostratal) approaches.

The toolkit available in monostratal constraint-based formalisms for the treatment of various kinds of discontinuity includes the following items:

1. **Flat structure** – permits permutations and easy implementation of constraints on linear order within head domains. If, however, more complex constituent structure is required, it may be preferable to sacrifice such advantages. In Sag (1987:334) flat structures are proposed for the treatment of extraposition by LP rules. Uszkoreit (1986 or 1987b) discusses the treatment of complex fronting in German.

2. **Binary structure** (or other type of rather *ad hoc* structure) – if the notion of a constituent is defined independently of syntactic relations between heads and their modifiers, constituent structure is free to reflect discontinuity by postulating unorthodox or formal constituents. In Hinrichs & Nakazawa (1989a,b) the verb cluster in German is a constituent. In Oliva (1992a,b) any surface string beginning at a word boundary and spanning the rest of its head domain until its rightmost element is a constituent. This type of structure, in conjunction with raising (see below), allows to treat recursive head domains.

3. **Sequence union**, or separate treatment of surface linear order from the linear order of terminal leaves of the syntactic tree, is perhaps the most principled approach to the treatment of word order variations within finite-clause-bounded recursive head domains. An intuitively plausible constituent structure can be retained, but the cost of ordering operations may be high. This solution was proposed by Reape (1989 and 1990).

4. **Slash mechanism** is used for unbounded dependencies, especially for the cases when a filler or a constituent being referred to stands at the beginning of the clause, while its governor

is embedded somewhere towards the end. The link between the gap (trace) and its filler (or referent) is provided via a common mother node. The first proposal along these lines is due to Harris (1956). Gazdar (1978) suggests that unbounded dependencies can be base-generated when slashed categories (i.e., categories from which a constituent is extracted) are introduced. Since then, 'slash' has become a common tool in monostratal frameworks (Gazdar *et al.* 1985, Pollard & Sag 1991).

5. **Raising** as a means to solve discontinuity is mainly used for finite-clause-bounded recursive-head-domain phenomena, such as the intermixing within the German middle field of complements of various verbs placed in the final verb cluster. All unsaturated subcategorization requirements of embedded verbs are collected in their mother node, which can be a sister to the complements required. See, e.g., Hinrichs & Nakazawa (1989a,b) and Oliva (1991).

Let me now present a sample of actual proposals in some more detail. But first, a word is due on the tools employed to constrain the surface order.

4.1 Constraints on linearization

4.1.1 Binary linear precedence rules (constraints)

The split of context-free grammar into the components of immediate dominance (ID) and linear precedence (LP) was first suggested in Gazdar & Pullum (1981) and used in Gazdar *et al.* (1985) as an important feature of GPSG. LP rules state precedence (not adjacency) relations between pairs of constituents, irrespective of their mother category. In the basic version, these relations stipulate regularities (not preferences) holding among sisters in a local tree. A constituent structure must obey all LP rules. This means that every pair of daughter constituents in a given structure must be checked against the set of LP rules. If a rule is found which is relevant for the pair being checked (i.e., if category specifications in the rule subsume categories in the pair), the pair must occur in the (surface) linear order specified by the rule.

The original version of LP rules had its unmistakable merits in the possibility to express generalization of fixed word order patterns, but failed to account for a number of common constructions even in English. Sag (1987) and Pollard & Sag (1987) propose to relate some LP rules (constraints) to lexical specifications in the head of the structure, namely to the 'obliqueness' hierarchy of complements in subcategorization list. LP constraint 1 requires that lexical heads precede anything else (39a), while LP constraint 2, in its basic version, requires that complements higher in the hierarchy precede those which are lower (39b).

(39) a. LP1: HEAD[LEX $+$] $\prec$ [ ]

b. LP2: COMPLEMENT $\prec\prec$ COMPLEMENT

c. LP2: COMPLEMENT $\prec\prec$ COMPLEMENT[LEX $-$]

d. LP2: COMPLEMENT[MAJ $\neg$V] $\prec\prec$ [LEX $-$]

e. LP3: [MAJ $\neg$N] $\prec$ [FOCUS $+$]

The slightly more sophisticated version of LP2 (39c) requires that less oblique complements precede more oblique complements only in case the latter are not lexical. If adverbial particle is left unspecified for the LEX feature, then both possibilities in (8) are permitted. To solve cases like (40), verbal complements are excluded as the preceding category in the constraint (39d).

(40) a. Kim appeared to Sandy to be aggressive / to be aggressive to Sandy.

b. Kim appealed to Sandy to be aggressive / * to be aggressive to Sandy.

In the obliqueness hierarchy of the verb *appear* the PP[TO] complement follows the infinitive. LP2 therefore does not apply. On the other hand, the obliqueness hierarchy of PP[TO] and the infinitive for *appeal* is reversed and LP2 applies.

If one tries to extend the set of LP constraints to cover a simple case of heavy-NP shift, such as (6c) or (6d), a simple addition of LP3 (39e) results in violation of LP2. If, however, LP2 and LP3 are treated as disjuncts, together representing a single constraint, the ordering can respond to an interaction between grammatical and discourse-based constraints.

This was the solution proposed by Uszkoreit (1986) to account for the order within the German middle field. Starting from the simple observation that there is a tendency for focussed items to follow other items, for complements to exhibit the order subject $\prec$ indirect object $\prec$ direct object and for personal pronouns to precede other NPs, a complex LP rule for the German middle field was formulated as follows[26]:

(41) +NOM $\prec$ +DAT
+NOM $\prec$ +ACC
+DAT $\prec$ +ACC
−FOCUS $\prec$ +FOCUS
+PRONOUN $\prec$ −PRONOUN

The disjunctive *LP clauses* within the complex rule can carry different weights. In (41) the clauses are ordered by weight, the last clause being the heaviest.

A problem remains, however, how to constrain by LP statements non-siblings and non-adjacent (discontinuous) elements. The main ideas of various proposals to solve it are presented below. In fact, a space for some solution is provided already in Pollard & Sag (1987:188–189), where two kinds of ordering functions constraining the PHONOLOGY attribute of a phrasal sign are assumed. The input to both are the values of the PHONOLOGY attributes of daughter signs. In the case of the first function (called *order-constituents*), the output is a permutation of these values as whole strings. In the case of the second function (called *interleave-constituents*), the output is a permutation of 'pieces of the PHONOLOGY strings of daughter signs'. Whereas the former can constrain local trees only, the latter is able to constrain fully recursive head domains.

4.1.2 Regular expressions

Kay (1985) uses an attribute *pattern* to state ordering of constituents within a *functional description* of FUG. In Karttunen & Kay (1985) patterns like (42a) are used to order nonadjacent constituents in Finnish and patterns like (42b) to place nonfocussed items at a nonfinal position.

(42) a. (... FiniteVerb ... TensedVerb ... MainVerb ...)

b. ((#[New = No]) #1)

The values of *pattern* attributes are actually regular expressions, more powerful than binary LP rules, since they can impose an order such as A B A* B* (see Uszkoreit *et al.* 1994), and are employed in implementations of LFG. A finite-state automaton for the encoding of binary LP constraints is proposed in Uszkoreit (1991).

[26] The constraints could as well be expressed in terms of syntactic functions or thematic roles.

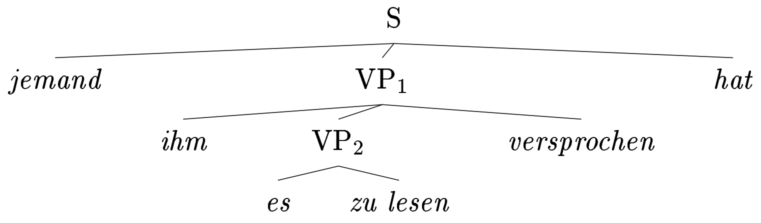


Figure 3: Constituency syntax tree for *es ihm jemand zu lesen versprochen hat.*

4.2 Reordering the sequence of syntactic leaves: sequence union

Reape (1990:25) makes the following observation:

> Nearly all modern grammatical theories derive word order from the terminal yield of phrase structure trees. This includes theories as disparate as GB, LFG, and GPSG[27]. Another way to put this is to say that the word order domain of a constituent is the sequence of the leaves of its constituent structure tree. Therefore, any attempt to explain apparent cases of discontinuous constituency requires that "continuous" surface phrase structure trees be assigned whose terminal yield exhibits the apparent discontinuity.

The notion of *word order domain* corresponds to the concept of fully recursive head domain introduced above. Thus, taking as an example the verb-final German clause in (27b), repeated here as (43), and assuming a syntax tree in Figure 3 with *hat* as the head of the clause, the word order domain of the clause (minus the complementizer) equals the fully recursive head domain of *hat*[28].

(43) ... daß es ihm jemand zu lesen versprochen hat

Reape departs from the identity of the order of syntax tree leaves with the surface order. Surface order is viewed as a separate dimension of grammar and is represented by word order domains. Word order domain of a clause is built monotonically bottom-up from word order domains of the smaller constituents. A word order domain of a daughter may appear non-adjacently within the domain of its mother constituent under the condition that the order of items of the daughter domain is preserved in the mother domain, i.e, all LP constraints imposed on daughter domains are preserved in the merged domain of the mother. The relation between the daughter domains and the mother domain is called *sequence union*[29]. Figure 4 illustrates this successive domain building.

This treatment of word order is then interpreted in a version of HPSG formalism. A few modifications of the standard variant of Pollard & Sag (1987) are necessary:

- A new attribute DOM is added to signs, specifying the relevant word order domain.
- The PHON attribute, according to a modified *Constituent Ordering Principle*, simply concatenates the values of the PHON attributes of the items of the DOM sequence.
- A new *Domain Principle* relates the value of the DOM attribute to the head and non-head daughters. If a (non-head) daughter is specified as UNIONED −, it should not be sequence-unioned (non-verb projections in German) and its domain occurs in the mother domain unchanged.

[27] HPSG is an important exception in that it provides space for non-adjacent surface positioning of items which are siblings in a syntactic tree.
[28] Figures 3 and 4 are adapted from Reape (1991).
[29] This relation is reminiscent of the function *interleave-constituents* in Pollard & Sag (1987:189), mentioned above.

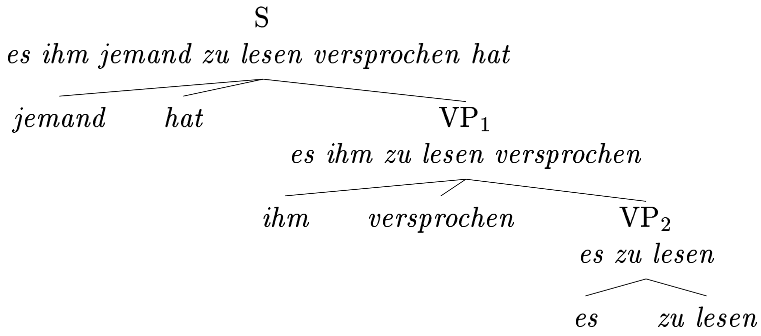


Figure 4: Domain tree for *es ihm jemand zu lesen versprochen hat.*

- LP constraints then operate on the DOM sequences.

The conflict between surface order and the syntactic tree is solved by the introduction of a separate level of description – the word order domain, which is, however, fully integrated into the monostratal framework of HPSG. Ordering constraints are then easily stated as constraints on word order domains. This solution is formally well worked out and (to my knowledge and taste) unrivalled as far as simplicity and elegance are concerned. Yet some doubts have been raised about its efficiency, namely about the necessity to perform the complex operation of sequence union whenever a constituent appears discontinuously in a higher constituent. Also, Reape does not offer a solution to the problem of integrating into the formalism mutually interacting LP constraints (see a complex LP rule in (41)).

4.3 LP constraints in lexical entries

A different strategy is adopted by Engelkamp *et al.* (1992a,b). Rather than separating surface order from the order of leaves in a syntactic tree, they constrain the leaves of a flexible binary-branching syntactic tree by LP constraints compiled into lexical entries. These constraints then operate on recursive head domains as an integral part of a typed feature unification formalism without any additional provisions for linear precedence.

The main idea behind this approach is as follows. An LP constraint A $\prec$ B is compiled into both A and B as a ban on the other category standing somewhere to the right/left. A category projection collects LP-relevant information about all items in its recursive head domain. Whenever a complement or adjunct is combined with its head, its restrictions on the left/right context are checked. In this way, LP constraints can be checked locally in sibling nodes and their violation can result in unification failure.

The integration of this solution into a HPSG framework requires a formidable amount of additions. The structure in (44) specifies the location of new attributes needed for the handling of LP-constraints within an HPSG sign.

(44) $$\left[\text{SYNSEM}\,|\,\text{LOC}\ \left[\begin{array}{ll}\text{HEAD} & \left[\begin{array}{ll}\text{LP-IN} & [\,]\\ \text{LP-OUT} & [\,]\\ \text{LEFT} & [\,]\\ \text{RIGHT} & [\,]\end{array}\right]\\ \text{LP-STORE} & [\,]\end{array}\right]\right]$$

The attributes LEFT and RIGHT specify conditions on the context and are used when the constituent becomes a complement or adjunct. The attribute LP-STORE is used when the constituent becomes the head of a complement or adjunct: it contains information about LP-relevant features from the entire recursive head domain. This information is stored as binary-valued features. If there are five LP-relevant features used to represent the four LP clauses of the complex LP rule in (45) (equivalent to (41) minus the FOCUS clause), the LP-STORE value of a projection like *einem Spion einen Brief zusteckt* is as in (46).

(45) NOM $\prec$ DAT
NOM $\prec$ ACC
DAT $\prec$ ACC
PRO $\prec$ NON-PRO

(46)
$$\left[\text{LP-STORE} \quad \begin{bmatrix} \text{NOM} & - \\ \text{DAT} & + \\ \text{ACC} & + \\ \text{PRO} & - \\ \text{NON-PRO} & + \end{bmatrix}\right]$$

Whenever a complement or adjunct is combined with its head, the head's LP-STORE is modified via the complement/adjunct's features LP-IN and LP-OUT: information about the complement/adjunct's LP-relevant features is added to the LP-STORE.

The solution works for cases like (24a), where no complement is pronominal and the fourth clause of the complex LP rule in (45) does not apply. However, as soon as a pronoun occurs as one of the complements, the proposed solution does not evade the conflict between the 'case' clauses and the 'pro' clause. The LP clauses of the complex LP rule in (45) operate in effect as conjuncts, rather than as disjuncts. In (47a) the RIGHT attribute in the lexical entry of the pronoun *ihm* has the value NOM $-$, which is due to the second LP clause and to the fact that the pronoun's case is dative. But one of the LP-relevant features in the LP-STORE of the head is NOM $+$, due to the subject *der Doktor*. Since the complement's value of the RIGHT attribute and the head's value of the LP-STORE attribute must unify (according to the *Right-Head LP-Principle*), the result is unification failure. The same goes for (47b), where the nominative non-pronominal subject requires that its value of the RIGHT attribute, which is PRO $-$, unifies with the value of the head's LP-STORE, where the value of PRO is $+$.

(47) a. Dann wird ihm der Doktor die Pille geben.

b. Dann wird der Doktor ihm die Pille geben.

In fact, no ordering of complements in the middle field of (47a) is accepted. The addition of another ('focus') clause to account for (24b) would result in additional conflicts between the LP clauses. Thus, although the approach is interesting and innovative, it is not obvious how it could be extended to accommodate complex LP rules.

4.4 LP constraints in grammar rules

The idea behind the approach presented in Uszkoreit *et al.* (1994) is rather straightforward: instead of specifying LP constraints of every potential complement in the lexicon (values of the LEFT and RIGHT attributes of Engelkamp *et al.* 1992), and testing them by unification with a collection of LP-relevant features of head domain of the head (value of the LP-STORE *ibid.*), lexical and phrasal categories (signs) only provide LP-relevant data about themselves and their head domains while LP constraints are directly expressed in grammar rules which generate binary-branching structure. A complement-head grammar rule stipulates values of LP-relevant attributes for the complement (the sister of the head in the local tree) and also for other complements of the head. Thus, a clause like (48) is licensed by the three grammar rules in (49).

(48) … daß er das Buch dem Mann gibt

(49) a. obj2_head_rule3:
COMPLEMENT: **object2:** pronoun:−
HEAD: **subject:** *nil*, **object:** *nil* OR pronoun:−

b. obj_head_rule2:
COMPLEMENT: **object:** pronoun:− focus:−
HEAD: **subject:** *nil* OR pronoun:− focus:+, **object2:** pronoun:− focus:+

c. subj_head_rule:
COMPLEMENT: **subject:** ___
HEAD: **object:** ___, **object2:** ___

The rule examples is my simplified paraphrase of complex typed feature-structure-based code. Note the instantiation of the feature 'focus' for object and object2 in obj_head_rule2. The rule actually describes the reversed (marked) order of object complements, different from the unmarked order (systemic ordering). The reversed ordering is licensed by the topic/focus distinction (focus −/+). As a result, the rule also says that in the case of conflict between the functional and the pragmatic LP constraints the pragmatic LP constraints win. Thus, in parsing, the indirect object *dem Mann* receives the specification 'focus:+' and the direct object *das Buch* the specification 'focus:−'. Note also that the subject does not receive any focus specification. In fact, the focus feature is instantiated for all complements in a very limited number of orderings and a substantial subset of orderings receives no focus specification at all.

In this way, not only disjunction of LP clauses within a complex LP rule can be incorporated into the grammar, but also the relative weights (priorities) of the individual factors can be taken into account. The cost is obvious: proliferation of rules. In addition, separate rules are needed for head-final (right-branching) and head-initial (left-branching) structures (to account for the three German clause types, the verb-second type is treated as an instance of the verb-initial type with a topicalized constituent). Hence, the section of the grammar describing the three complements and their order, based on the interaction of grammatical function, pronominality, and topic/focus distinction (with an additional factor involving the enclitical pronoun *es*), numbers 18 rules. Ten rules treat head-initial structures (five for subject, three for object and two for object2) and eight rules license head-final structures (one for subject, three for object and four for object2). Six rules are also provided for the ordering of source and goal adjuncts, although only with respect to each other and not to other complements.

The rules actually do more than reflecting relative weights of the individual LP clauses in a complex LP constraint like (41). To do justice to examples like (25b–25c) and (25a), where the hierarchy of the individual factors is violated, the grammar is more explicit and imposes constraints based on combinations of values which go beyond the hierarchy. The coverage of the various interactions of factors is quite respectable. In (50) a list of possible orderings licensed by the grammar rules is given, while (51) lists four negative constraints.

(50) a. SUBJ $\prec$ OBJ2

b. SUBJ $\prec$ OBJ

c. OBJ2 $\prec$ OBJ

d. OBJ PRO $\prec$ OBJ2 PRO

e. OBJ2 PRO −FOC $\prec$ OBJ PRO FOC

f. OBJ −FOC $\prec$ SUBJ FOC

g. OBJ2 −FOC $\prec$ SUBJ FOC

h. OBJ PRO $\prec$ SUBJ

i. OBJ2 PRO $\prec$ SUBJ

j. OBJ PRO $\prec$ OBJ2 −PRO

k. OBJ −FOC $\prec$ OBJ2 FOC

(51) a. OBJ −PRO FOC $\not\prec$ OBJ2 −PRO −FOC

b. OBJ2 PRO $\not\prec$ OBJ CLITIC

c. SUBJ PRO $\not\prec$ OBJ −PRO

d. SUBJ PRO $\not\prec$ OBJ2 −PRO

The ordering constraints (50a – 50c) directly correspond to the first three clauses in the complex LP rule in (41). The constraint (50d) expresses the peculiar fact that if both objects are pronominalized, the order of indirect and direct objects is reversed (see example (25a)). If this is not the case, the direct object should be focussed (50e). This prediction cannot be expressed in terms of a single hierarchy of LP clauses with single values. It is actually a subhierachy involving the class of orderings with both objects pronominalized. The constraints (50f – 50k) express the fact that the constraints involving specifications of pronominality and topic/focus distinction override those involving only syntactic functions. Thus, the rules instantiate the focus feature (at least for some complements) in the following cases of order in the middle field:

(52) a.

(der Doktor/er)	die Pille	dem Patienten
	−FOC	+FOC

b.

(der Doktor/er)	ihm	sie
	−FOC	+FOC

c.

dem Patienten	der Doktor	die Pille
−FOC		+FOC

d.

ihm	der Doktor	sie
−FOC		+FOC

e.

ihm	sie	der Doktor
−FOC	+FOC	

f. die Pille der Doktor dem Patienten
   −FOC +FOC +FOC

g. sie dem Patienten der Doktor
   −FOC +FOC

In (52a) the "focussed indirect object" constraint (50k) overrides the "unmarked order" constraint (50c). In (52b) the "focussed pronominalized object" constraint (50e) overrides the "unmarked pronominalized objects order" constraint (50d). In (52c) the "focussed subject" constraint (50g) overrides the "unmarked order" constraint (50a). In (52d) the +/−FOC specification is due again to (50e) rather than to the violation of the "unmarked order" constraint (50a) (the indirect object is pronominal), therefore the subject is unspecified for the focus feature. In (52e) the constraint (50e) is involved again, and, again, nothing is said about the subject's focus specification (the pronominality of both objects precludes any such commitment by virtue of (50h) and (50i), which override the "focussed subject" constraints (50f) and 50g). In (52f) the direct object precedes both the subject and the indirect object, which is supposed to account for the exhaustive topic/focus specification (due to (50f) and (50k)). In (52g), the constraint (50g) is responsible for the topic/focus specification, while the position of the pronominal object is licensed by (50h) and (50j). In all these cases, none of the negative constraints is violated.

There is one remarkable fact about this grammar: its implementation as a parser is reported to run considerably faster than its other version without word order constraints[30]. The explanation is that a better filtering of ungrammatical substructures compensates for the longer search in a larger corpus of rules.

Clearly, the way from constraints to the head-initial and head-final grammar rules is far from straightforward and an automatic compilation procedure would be badly needed if more factors and more interactions with subregularities are to be reflected. Another extension is needed to account for ordering of complements of embedded and main verbs within the middle field (see examples (27b) and (27d)). Finally, with the breadth of coverage, a more general problem emerges involving the incompability of the symbolic account with the existing continuum between acceptable and unacceptable sentences. While more data from corpus-based and psycholinguistic research is needed, extensive use of numeric preferences might help.

4.5 LP constraints on lists

Another proposal made by Oliva (1992a,b) rests on a special case of binary right-branching syntactic structure, which is actually equivalent to a list (see Figure 5).

Note the empty nodes at the rightmost leave of every subtree and the possibility to accommodate any order within a head domain. The empty rightmost nodes share valency requirements with the lexical head. All nodes along the path from the empty node to the top of the (sub)tree (the maximal projection of the head) are in fact projections of the empty node. The representation resembles flat constituency structures or dependency trees, but it also gives a clear picture about the relation between the surface string and the representation (i.e., the derivation).

As soon as syntactic structures can be treated as lists, different sorts of lists can be distinguished and relations between these sorts specified (as a lattice). Thus, the most general sort (a list consisting of its head and tail) corresponds to a local right-branching binary tree. This sort then subsumes more specific sorts corresponding to (local right-branching binary) trees: the head of the list can be a

[30]A clause *gibt ihm der Mann das Buch* is parsed in 6.750 seconds, while the unrestricted grammar needs 38.433 seconds.

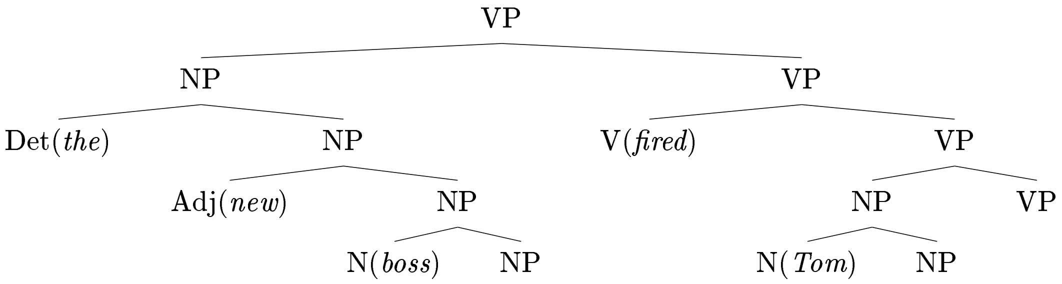


Figure 5: A binary branching syntax tree for the sentence *The new boss fired Tom.*

complement, an adjunct, a minor category, or the head of the phrase, while the tail of the list is the (possibly empty) rest of the phrase. These list sorts are convenient moulds for expressing the principles and grammar schemata commonly assumed in HPSG-inspired frameworks. However, the space of possible lists can be partitioned in many other ways, reflecting, e.g., various factors determining the order of terminal leaves, such as the "topological" constraints of the German clause types. The "ordering" sorts can then interact with the "expansion" sort and the principles to license an appropriate local tree.

Oliva (1992a) gives another example of constraints which can be expressed in this way: constraints based on the syntactic function of modifiers, and also an example of conjuction of partial ordering requirements. In this example, the hypothetical requirements demand that (i) in a clause, the verb always comes second, (ii) complements are ordered as in (50a – 50c), and (iii) pronouns precede nouns. The partial requirements consist of disjunctive constraints on local trees, i.e. on lists. Technically, the constraints are expressed as follows: In (53) an intersection of the most general sort with a parametric "type" (a macro standing for a feature structure) guarantees that any word order constraints are applied to any local tree.

(53) $\begin{bmatrix} \text{first} & [_{First}\ ] \\ \text{rest} & [_{Rest}\ ] \end{bmatrix}$

& ordering($[_{First}\ ]$,$[_{Rest}\ ]$)

The parametric type is itself defined in terms of a conjuction of three other parametric types, representing the three ordering requirements (see 54).

(54) ordering($[_{First}\ ]$,$[_{Rest}\ ]$) ==
n_and_pron_order($[_{First}\ ]$,$[_{Rest}\ ]$) &
compl_order($[_{First}\ ]$,$[_{Rest}\ ]$) &
v_2nd_pos($[_{First}\ ]$,$[_{Rest}\ ]$) .

The definitions of the parametric types in (54) are given as disjunctions of constraints on local trees reflecting the individual ordering factors. Thus, the constraints on complement order are expressed as follows[31]:

(55) a. compl_order($[_{subj}\ ]$,$[_{subj \vee dobj \vee iobj}\ ]$) == $[_{subj}\ ]$.

b. compl_order($[_{iobj}\ ]$,$[_{dobj \vee iobj}\ ]$) == $[_{iobj}\ ]$.

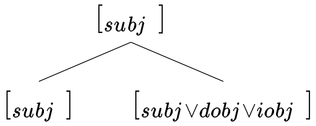
[31]The fact that a complement may not be repeated and that indirect object (direct object) should not appear unless a direct object (subject) is present should be expressed by another component of the grammar.

c. compl_order($[_{dobj}\ ]$,$[_{dobj}\ ]$) == $[_{dobj}\ ]$.

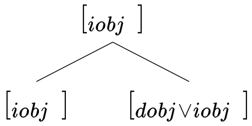
d. compl_order($[_{verb \lor adjunct}\ ]$,$[_{Sort}\ ]$) == $[_{Sort}\ ]$.

The parametric types definitions in (55) constrain the ordering of sorts corresponding to the types of modifiers. The local trees licensed by (55) are as follows:

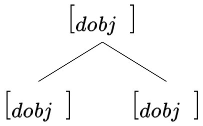
(56) a.



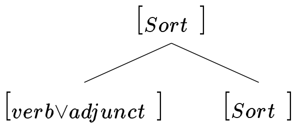
b.



c.



d.



This actually boils down to the effect which the following two schematically expressed LP constraints would have:

(57) $subj \prec dobj \lor iobj$
$iobj \prec dobj$

The other parametric types in (54) are defined in a similar fashion. The whole tree is build from local trees constrained by an intersection of sorts, which are defined in this way.

Now, as things stand, any ordering in which a pronoun violates the complement order, as in (47a), is out. Should the grammar describe a less hypothetical language where such sentences would be acceptable, a rather simple modification of (54) is required – see (58); the new parametric type n_pron_comp_order being defined in (59):

(58) ordering($[_{First}\ ]$,$[_{Rest}\ ]$) ==
n_pron_comp_order($[_{First}\ ]$,$[_{Rest}\ ]$) &
v_2nd_pos($[_{First}\ ]$,$[_{Rest}\ ]$) .

(59) n_pron_comp_order($[_{First}\ ]$,$[_{Rest}\ ]$) == n_and_pron_order($[_{First}\ ]$,$[_{Rest}\ ]$).
n_pron_comp_order($[_{First}\ ]$,$[_{Rest}\ ]$) == compl_order($[_{First}\ ]$,$[_{Rest}\ ]$) .

The effect of (58) and (59) is as follows: while the verb-second constraint applies always, pronominal and complement constraints need not apply both at the same time (in the same local tree). Therefore, if all complements are nouns, or if all complements are pronouns, their order is unconstrained. Similarly, if another constraint on the order of topic and focus is added, it would be effective only for cases when constraints both on complement order and on pronoun – noun order are violated. The interactions between the individual factors in a real language are clearly more complex. For example, in (25a), repeated here as (60), the unmarked order of pronominalized objects is dobj $\prec$ iobj.

(60) Dann wird sie ihm der Doktor geben.

An obvious solution is to provide parametric type definitions with conjunctions of sorts which would parallel constraints listed in (50). To account for (60), neglecting issues of topic and focus, the parametric type **n_pron_comp_order** could be defined as follows[32]:

(61) n_pron_comp_order($[_{subj\wedge noun}]$,$[_{(dobj\vee iobj)\wedge noun}]$) == $[_{subj\wedge noun}]$.
n_pron_comp_order($[_{subj\wedge pron}]$,$[_{dobj\vee iobj}]$) == $[_{subj\wedge pron}]$.
n_pron_comp_order($[_{iobj\wedge noun}]$,$[_{dobj\wedge noun}]$) == $[_{iobj\wedge noun}]$.
n_pron_comp_order($[_{iobj\wedge pron}]$,$[_{(dobj\vee subj)\wedge noun}]$) == $[_{iobj\wedge pron}]$.
n_pron_comp_order($[_{dobj\wedge noun}]$,$[_{dobj\wedge noun}]$) == $[_{\neg(subj\vee iobj\vee dobj)}]$.
n_pron_comp_order($[_{dobj\wedge pron}]$,$[_{(subj\wedge noun)\vee iobj}]$) == $[_{dobj\wedge pron}]$.

The definition (61) can be translated into the following (disjunctive) LP constraints:

(62) $subj \wedge noun \prec (dobj \vee iobj) \wedge noun$
$subj \wedge pron \prec dobj \vee iobj$
$iobj \wedge noun \prec obj \wedge noun$
$iobj \wedge pron \prec (obj \vee subj) \wedge noun$
$dobj \wedge pron \prec (subj \wedge noun) \vee iobj$

The constraints in (62) allow for violation of systemic ordering (unmarked order) only if *iobj* or *dobj* are pronominalized. Constraints in (63) cover some cases of unmarked order violation due to topic/focus articulation.

(63) $dobj \wedge noun \wedge topic \prec (iobj \vee subj) \wedge noun \wedge focus$
$iobj \wedge noun \wedge topic \prec subj \wedge noun \wedge focus$
$iobj \wedge pron \wedge topic \prec dobj \wedge pron \wedge focus$

The extension of the definition of the parametric type n_pron_comp_order in (61) should be obvious. The types (feature structures) unspecified for the sort *topic* or $focus$ can be interpreted as ambiguous in this respect. If items preceding the topic/focus boundary are to be specified as belonging to topic and those following it as belonging to focus, in addition to the extension of (61) by the translation of (63), another conjunct can be added to the definition in (58) yielding (64). The new type is defined in (65).

(64) ordering($[_{First}]$,$[_{Rest}]$) ==
n_pron_comp_order($[_{First}]$,$[_{Rest}]$) &
topic_focus($[_{First}]$,$[_{Rest}]$) &
v_2nd_pos($[_{First}]$,$[_{Rest}]$) .

(65) topic_focus($[_{topic}]$,$[_{topic}]$) == $[_{topic}]$.
topic_focus($[_{topic}]$,$[_{focus}]$) == $[_{topic}]$.
topic_focus($[_{focus}]$,$[_{focus\vee nil}]$) == $[_{focus}]$.

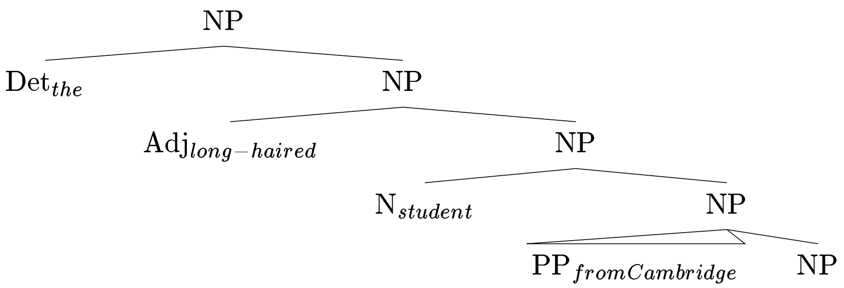
[32]This time, restrictions following from valency requirements are taken into account, and the position of the verb and adjuncts is not.

The definition in (65) has the effect of partitioning the whole sentence into two parts: topic and focus. If no clue is provided by the constraints in (63), the partitioning is ambiguous.

The list structure is certainly a very convenient tool, especially for the description of languages with highly variable word order patterns. If, however, a more expressive structural description is required, the binary right-branching tree becomes more complex. In the example (66), the pronoun *one* stands for something which is not a subtree (constituent) if a structure depicted in (67) is assumed.

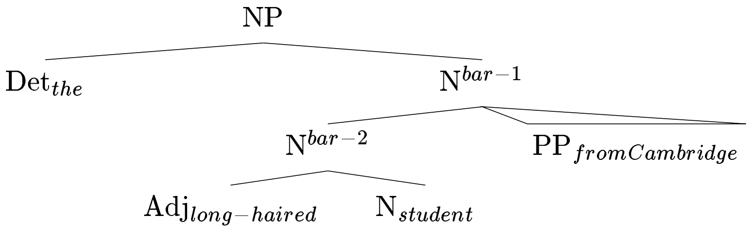
(66) the long-haired student from Cambridge and one from Oxford (MC94)

(67)

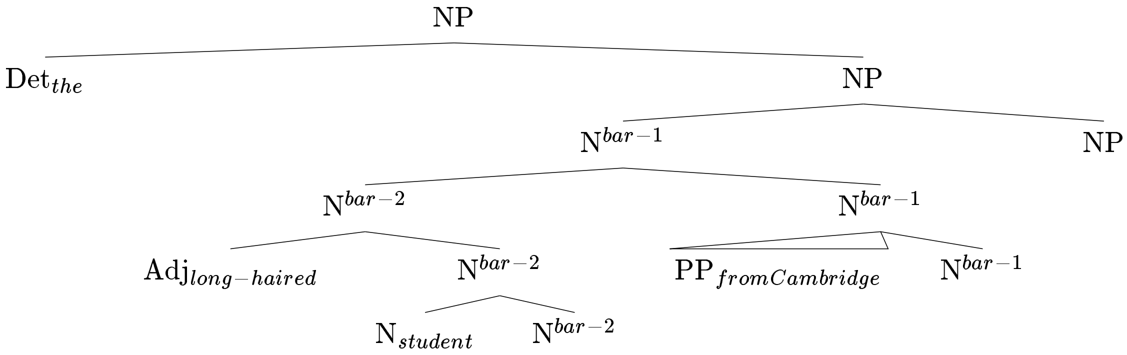


To do justice to the example (66), the common analysis in (68), which represents one of the two readings of the NP *the long-haired student from Cambridge*, should be translated into (69).

(68)



(69)



To conclude, the beauty of the uniform list structure, if it can be maintained, has certain pleasant consequences also in situations when mutually interacting LP constraints are needed, and the possibility to express other theoretical claims concerning word order seems to be open. No extension of the set of rules is required. Complements of embedded verbs occurring among complements of the main verb can be treated by raising and be subject to the same constraints. However, the general problem due to the continuous transition from what is acceptable and what is not remains.

4.6 A different view: deep and surface order

In some approaches rooted in the dependency tradition, the problem of non-adjacent siblings – dependents (crossing dependencies) is solved by allowing structures less restricted than trees. In the monostratal frameworks of Hudson (1990), Covington (1990) and also in the rather practically oriented systems of Sleator & Temperley (1992) and Karlsson (1990) a syntactic unit may be associated with more than one superordinate item. In the polystratal framework of Mel̆čuk, non-projective structures are treated as exceptions to the otherwise general condition of projectivity and handled by transformations (see Mel̆čuk & Pertsov (1987))[33]. However, these frameworks have little to say about other than configurational constraints.

Among dependency-oriented syntactic theories, clearly the one most preoccupied with word order issues is the theory of Functional Generative Description, an heir to the tradition of the Prague School structuralism (see Sgall *et al.* 1986). The theory postulates an underlying level of deep syntax, where sentences are represented as projective dependency trees. The nodes of a tree represent only categories associated with content words and the arcs syntactic relations between them. The nodes are ordered from left to right and the (total) ordering corresponds to the notion of *deep word order*. This order is determined by *communicative dynamism*, which puts less dynamic items before more dynamic (focussed) ones and which is the result of the interaction of the following three factors:

- topic/focus distinction, which partitions the deep word order sequence of entire sentences into two parts (corresponding to given and new information); the same is true also about embedded structures;
- salience hierarchy of elements referred to or implicit in the discourse, in other words, degree of their activation in the relevant context (in its broadest sense); the degree of activation determines the hierarchy of communicative dynamism and the deep word order in the topic (given) part, where 'new topic' precedes 'background';
- systemic (unmarked) ordering of dependency relations – types of modifiers (complements and adjuncts), corresponding roughly to thematic roles; the ordering is specific but invariable for a particular language; systemic ordering determines deep word order in the focus part[34].

The relation between the deep and surface word order is stated in terms of movement rules. The movement rules account for 'syntactic deviations' (where surface order obeys morphosyntactic constraints rather than the hierarchy of communicative dynamism) and 'subjective order' (due to marked sentential stress patterns). In languages with more variable word order the relation between the deep and surface order is more straightforward than in configurational languages.

Thus, similarly as in some approaches described above, the surface word order is determined by several factors, which, however, form a hierarchy which is embodied in the very architecture of framework: some of them determine the deep word order while the rest is at work when transformations between the surface and the deep structure are performed.

A procedure applying the syntactic constraints and principles distinguishing between topic/focus and governing the deep word order in the focus part has been formulated for English in Hajičová *et*

[33] In fact, the exceptions, at least in the surface grammar of English as described in Mel̆čuk & Pertsov (1987), are rather numerous: from the total number of 240 English constructions accounted for, there are 39 which either require or permit non-projectivity.

[34] Systemic ordering can be determined by testing all pairs of types of dependency relations. For Czech, Petkevič (1993) lists 38 dependency relations ordered according to systemic ordering. Less detailed orderings have been suggested also for English and German. According to Hajičová *et al.* (1993) the systemic ordering for English is as in (70). Pfeiffer *et al.* (1994:158) give the systemic ordering for German (see (71)).

(70) Time Actor Addressee Objective Origin Effect Manner $\text{Directional}_{from}$ Means Directional_{to} Locative

(71) Actor Time Cause Purpose Manner Amount Locative Means $\text{Directional}_{from}$ Addressee Objective Origin Effect Directional_{to}

al. (1993) and German in Pfeiffer *et al.* (1994). In both cases, neither complete coverage of data nor full conformance with the theory is claimed. Nevertheless, they constitute an explicit account worth some investigation.

The procedures determine the topic/focus boundary and deep order in the focus part for immediate dependents of the main verb. The input is dependency tree structure analysis of common written sentences taken from technically-oriented texts. The trees are annotated with functional information. In some cases, additional (mainly semantic) information may be needed. Nodes of the tree are ordered according to surface order. The output is the same structure, reordered according to deep order, with a boundary distinguishing among nodes in topic and nodes in focus. Ambiguity in topic/focus assignments results in multiple structures.

For English, the additional input information includes definiteness of nouns, semantic type of verbs (general, specific, intermediate) and adjuncts (specific, such as *on a nice September day, in October seven moths ago*, general, such as *always*, or indexical *here, today, this year*). Similar requirements exist for German. The procedure for German runs as follows[35]:

1. All leftward and rightward dependents of the verb are provisionally marked as t (i.e., in topic) and f (i.e., in focus), respectively.

2. The main verb is marked as t, if it has a very general, lexically weak meaning *sein, vorkommen, werden, tun, verbringen, bekommen.* If it is semantically very specific or clause-final (as a finite verb), the verb gets f. In all other cases it is marked as ambiguous (t/f).

3. If there is a single rightward dependent and a definite noun group at the same time, it is marked t/f.

4. If there are more rightward dependents:

   (a) and a pair of adjacent dependents A B exists for which the systemic ordering is reversed (B A), then A and all other dependents between the verb and A get t, while those following B, except the final one, get t if definite or t/f if indefinite;

   (b) and no such pair exists, then all rightward dependents, except the final one, get t if definite or t/f if indefinite.

5. Two structures are created for every local t/f ambiguity: in one the item is marked as t, in the other as f.

6. A structure without any f item is deleted.

7. If a f item precedes a t item, the order is reversed.

Nodes in the output structure are then separated into the topic and the focus part, while the ordering in the focus part reflects the scale of communicative dynamism.

To compare this procedure ($\mathcal{A}$) with the 'LP constraints in grammar rules' approach ($\mathcal{B}$) sketched above and described in Uszkoreit *et al.* (1994), certain incompabilities have to be realized:

[35] The procedure assumes verb-second or verb-final clauses, with all clause-final verb forms together with the finite verb treated as a single node in the finite verb position. Similarly, all function words are supposed to form complex nodes with respective content words. All dependents (= modifiers) are to be understood as immediate dependents. Indexical adverbs (*hier, gestern*) can be treated as definites, while more general ones (*oft, manchmal*) as indefinites. Pronouns are not treated.

- The current version of $\mathcal{B}$ (unlike $\mathcal{A}$) does not take into account the category of definiteness of nominals.
- $\mathcal{B}$ (unlike $\mathcal{A}$) applies word order constraints only to the middle field.
- $\mathcal{B}$ (unlike $\mathcal{A}$) does not work with the notion of a separate deep order, the discourse-related factors are manifested only through the binary-valued feature +/−FOCUS.
- The current version of $\mathcal{B}$ (unlike $\mathcal{A}$) does not cover adjuncts, with the exception of two types (direction$_{from}$ and direction$_{to}$), which are ordered only with respect to each other.
- The current version of $\mathcal{A}$ (unlike $\mathcal{B}$) does not treat pronominalized complements.
- The current version of $\mathcal{A}$ (unlike $\mathcal{B}$) does not treat verb-final clauses.
- $\mathcal{A}$ gives different results depending on semantic class of the verb.
- $\mathcal{A}$ (unlike $\mathcal{B}$) does not rule out unaccaptable sentences, with the exception of sentences without focus[36].

Hence, only order variations of definite NP complements in the middle field of a verb-second (or verb-first) sentence like (72) are compared.

(72) Dann gibt **Peter dem Mann das Buch**.

The six possible permutations of the three complements are given in (73). Violation of some of the three LP constraints relevant to the example is marked as ×. Or, in other words, the number of × symbols equals the number of times systemic ordering is violated. The symbol ∘ means that the constraint is satisfied. The superscript *t* or *f* means that the procedure $\mathcal{A}$ marks the complement as topic or focus, respectively. The subscript says how the feature FOCUS is instantiated in the grammar $\mathcal{B}$: *t* stands for −FOCUS, *f* for +FOCUS, blank for unspecified value.

					subj ≺ obj2	subj ≺ obj	obj2 ≺ obj
	a.	Peter^t	dem Mann^t	das Buch^f	∘	∘	∘
	b.	Peter^t	das Buch^t_t	dem Mann^f_f	∘	∘	×
(73)	c.	dem Mann^t_t	Peter^f_f	das Buch^f	×	∘	∘
	d.	dem Mann^t_t	das Buch^t_t	Peter^f_f	×	×	∘
	e.	das Buch	dem Mann	Peter	×	×	×
	f.	das Buch^t_t	Peter^f_f	dem Mann^f_f	∘	×	×

With the caveat that $\mathcal{A}$ and $\mathcal{B}$ are not ideal incarnations of the theories behind them, the following observations can me made:

- If $\mathcal{B}$ specifies the feature FOCUS for a constituent, the specification agrees with that predicted by $\mathcal{A}$.
- $\mathcal{A}$ makes stronger claims than $\mathcal{B}$ in the case of unmarked order (as in (a)) as well as in the cases of a single violation of systemic ordering ('functional' LP constraints, as in (b) and (c)). $\mathcal{A}$ and $\mathcal{B}$ give identical results when systemic ordering is violated twice (i.e., when two 'functional' LP constraints are violated).

[36] Indeed, $\mathcal{A}$ is a procedure defined for topic/focus analysis of syntactically parsed input, whereas $\mathcal{B}$ is (an integral component of) a grammar, interpretable in both directions (parser and generator).

- According to $\mathcal{B}$ the grammar has to satisfy at least one constraint, the sentence (e) is therefore out. According to $\mathcal{A}$, the complement *dem Mann* would be marked simultaneously as t and f, which could be interpreted as a failure and the sentence could be excluded as well. However, in a version of $\mathcal{B}$, (d) is also excluded. This exemplifies the problem of the existence of a continuum between the acceptable and unacceptable, rather than of a clear-cut borderline.

Now, whatever position one takes with respect to the theoretical premises of $\mathcal{A}$ and the resulting stronger claims it makes, an integration of the formal tool reformulated in a declarative fashion with a syntactic description would represent an interesting step. Then, the theoretical claims could be tested and compared on a more sophisticated level.

From a more theoretical viewpoint, the following points can be made:

- This (polystratal) approach differs from the monostratal approaches presented here in that it singles out one factor (or a group of factors) determining surface word order as the basic one(s) and relegates all the other factors to the lower strata. The basic factor(s) are responsible for a relatively independent total ordering — the deep order. Clearly, the choice between a hierarchy of word-order constraints applied at separate strata and a hierarchy of such constraints applied simultaneously at a single stratum is conditioned by overall framework of the theory and, from a certain perspective, both approaches can be regarded as equivalent[37]. However, some evidence is needed to decide whether there is an ordering different from the surface one, or whether all ingredients present in a message interact to finally produce a single (surface) order.

- The notion of communicative dynamism hierarchy parallels surface word order in that it imposes total ordering. While linearization of items is a necessity at the surface level, a more complex hierarchy may be the case at the deep level. Again, the choice between total ordering or a more complex hierarchy at the deep level is a matter of conclusions drawn from linguistic evidence.

- The contrast between factors determining deep order in topic (salience hiearchy) and in focus (systemic ordering) gives rise to a rather sharp split. Or is there not such strict "division of labour" between the factors? (Indeed, this point is related to the previous one.)

- While languages (and their syntactic structures) differ in how strong the individual factors determining word order are, the position of discourse-based factors may be challenged by a stronger factor not only in the share of responsibility for the resulting surface order, but also in what position the factor assumes in the overall framework. A parameterized hierarchy of factors may turn out to be an appropriate solution.

- Sentences are represented as projective dependency trees at the deep level. Hence, dependency relations determine pairs of nodes linked by edges and their vertical order. Their horizontal order represents the hierarchy of communicative dynamism. Now if a node representing a less dynamic item A depends on a node representing a more dynamic item B, and there is a node representing an item C, more dynamic than A and less dynamic than C, the condition of projectivity is violated (see Figure 6).

  Examples like (34a) or (34b), cannot be represented projectively without movements if C is the finite verb, B is the embedded infinitive and A is its complement. Separate representation of the two dimensions of the deep tree is a possible solution.

[37]Furthermore, there is no predestination: in a monostratal framework several *levels* of linguistic description may coexist in a single formal structure and nothing can possibly prevent a stratificational description from applying all constraints at once at a single stratum.

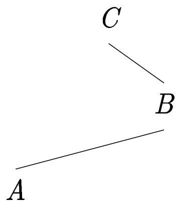


Figure 6: A non-projective tree.

5 Conclusion

The table in Figure 7 compares approaches discussed so far with respect to three aspects: (i) to what domains their word order constraints can be applied, (ii) what solution to the problem of crossing dependency they offer, and (iii) to what extent their word order constraints can interact.

The earlier constituency-based attempts suffer from the prevailing orientation on English: the constraints are largely syntactic and cannot interact with each other nor with discourse-based factors. Furthermore, the syntactic structure cannot easily accommodate word order variations. More recent attempts either go in the direction of dissociating the syntax tree from the surface order of the terminal string, or try to maintain the identity of the order of leaves with the surface order. Both ways have their merits and drawbacks which need to be further investigated, as well as the question of whether there is a basic factor responsible for ordering patterns, manifested via an underlying hierarchy.

Acknowledgments

The core of this report was written during the author's stay at University of Saarland in Saarbrücken, supported by a grant from the Commission of the European Communities in the framework of the Communities' Action for Cooperation in Science and Technology with Central and Eastern European Countries. Subsequent elaboration has been made possible by the Grant Agency of the Czech Republic, grant no. GAČR 405/93/2173, and by the Research Support Scheme of the Open Society Institute, project no. RSS 340/93. The work could not have been accomplished without this material support.

For intellectual support in many stimulating discussions, the author wishes to express gratitude to Tania Avgustinova, Eva Hajičová, Karel Oliva, Petr Sgall, Wojciech Skut, and Hans Uszkoreit. As the present report is so much based on the work of others, apologies are due to all whose ideas have been misrepresented or misinterpreted.

Last but not least, the author's thanks are also due to his family, for emotional support and a lot of patience.

grammar framework	**applicability of word order constraints**	**solution to crossing dependencies**	**interaction of constraints**
GPSG (Gazdar *et al.* 1985)	to sister pairs in local trees	slash mechanism	no
HPSG (Pollard & Sag 1987)	to sister pairs in local trees	slash mechanism, function *interleave-constituents*	no, but constraints may depend on the lexically specified 'obliqueness hierarchy'
GPSG with complex LP rules (Uszkoreit 1987b)	to sister pairs in local trees (= to head domains, due to flat structure)	slash mechanism (due to flat structure, German complex fronting impossible)	an LP rule may consist of an ordered hierarchy of disjunctive LP clauses
FUG with regular expressions (Kay 1985)	to two or more (non-adjacent) sisters in local trees (= head domains, due to flat structure)	slash mechanism	no
HPSG with sequence union (Reape 1989, 1990)	to any pair of terminal leaves in a recursive head domain	order of terminal leaves need not correspond to surface order	no
HPSG with LP constraints in lexical entries (Engelkamp *et al.* 1992a)	to any pair of terminal leaves in a recursive head domain	slash mechanism (due to flexible binary-branching structure solves the problem of Uszkoreit (1987b))	no (only a conjunction of LP constraints)
LP constraints in grammar rules (Uszkoreit *et al.* 1994)	to non-recursive head domain, bounded in finite clause	slash mechanism	yes; complex interactions of constraints applied through individual rules
LP constraints on lists (Oliva 1992a)	to non-recursive head domain, bounded in finite clause	?	possible: complex interaction of constraints applied to local trees through restrictions including conjunctions or disjunctions of sorts
Deep and surface order (FGD, Sgall *et al.* 1986)	to dependency slots	no crossing dependencies can occur in non-recursive head domains, otherwise movement rules	discourse-based deep order is transformed into surface order by movement rules

Figure 7: Comparison of different approaches.

Sources of examples

Br82 J. W. Bresnan, J. W., R. M. Kaplan, S. Peters and A. Zaenen (1982). Cross-serial dependencies in Dutch. *Linguistic Inquiry* 13, 613-635.

Caro *CAROLINA* (electronic weekly) 94/115.

HN89 E. Hinrichs and Ts. Nakazawa (1989). *Flipped out: AUX in German.* Chicago Linguistic Society 25.

HN93 E. Hinrichs and Ts. Nakazawa (1993). Linearizing AUXs in German verbal complexes. In *German Grammar in HPSG*, CSLI, Stanford.

HU86 H. Uszkoreit (1986). *Constraints on Order.* CSLI Report No. CSLI-86-46. Center for the Study of Language and Information, Stanford University, Stanford, USA.

HU87 H. Uszkoreit (1987). *Word Order and Constituent Structure in German.* CSLI Lecture Notes No. 8. CSLI, Stanford.

Ja87 P. Jacobson (1987). Phrase Structure, Grammatical Relations, and Discontinuous Constituents. In Huck & Ojeda (1987).

KJ87 A. S. Kroch & A. K. Joshi (1987). Analyzing Extraposition in Tree Adjoining Grammar. In Huck & Ojeda (1987).

KO91 K. Oliva (1991). *On Cases of "Fixed" Word Order in a "Free" Word Order Language.* CLAUS report No. 14, Computerlinguistik, Universität des Saarlandes, Saarbrücken.

Mc87 J. D. McCawley (1987). Some Additional Evidence for Discontinuity. In Huck & Ojeda (1987).

MC94 Michael A. Covington (1994). *An Empirically Motivated Reinterpretation of Dependency Grammar.* Research Report. The University of Geargia, Athens, USA.

MEAN P. Sgall, E. Hajičová, and J. Panevová (1986). *The Meaning of the Sentence in its Semantic and Pragmatic Aspects.* Academia, Praha and Reidel, Dordrecht.

MT89 S. Makino & M. Tsutsui (1989). *A Dictionary of Basic Japanese Grammar.* The Japan Times, Tokyo.

Po91 I. Pollard and I. Sag (in press). *Information-Based Syntax and Semantics. Vol. 2: Agreement, Binding and Control.* CSLI Lecture Notes, Stanford.

RH90 M. Reape & M. Hepple (1990). Word Order and Constituency in West Continental Germanic. In Engdahl & Reape (1990).

Rp90 Reape, M. (1990). A theory of word order and discontinuous constituency in West Continental Germanic. In Engdahl & Reape (1990).

Rp91 Reape, M. (1991). Clause Union and Word Order Variation in Germanic. In M. Reape, *Word Order in Germanic and Parsing*, DYANA Deliverable R1.1.C, University of Edinburgh.

SS94 L. Smoljak & Z. Svěrák (1994). Záskok. In *Literární noviny* 11/1994.

St87 Stucky, S. U. (1987). Configurational Variation in English: A Study of Extraposition and Related Matters. In Huck and Ojeda (1987).

Šmil V. Šmilauer (1993). *Novočeská skladba.* Pedagogické nakladatelství, Praha.

WoSk Wojciech Skut, personal communication.

References

Avgustinova, T. and K. Oliva (1995). "On the Nature of the Wackernagel Position in Czech." to be published in *Proceedings of the First European Conference on Formal Description of Slavic Languages*, Universität Leipzig, 30 November – 2 December 1995

Covington, Michael A. (1990). Parsing discontinuous constituents in dependency grammar. *Computational Linguistics* 16:234–236.

Engdahl, E. and M. Reape, eds. (1990). *Parametric Variation in Germanic and Romance: Preliminary Investigations.* DYANA Deliverable R1.1.A, University of Edinburgh.

Engelkamp, J., G. Erbach and H. Uszkoreit (1992). Handling Linear Precedence Constraints by Unification. In *Proceedings of ACL '92.* Newark.

Gazdar, G. (1978). English as a context-free language. Unpublished manuscript.

Gazdar, G. and G.K. Pullum (1981). Subcategorization, Constituent Order and the Notion of "Head". In M. Moortgat, H.v.d. Hulst, and T. Hoekstra, eds. *The Scope of Lexical Rules*, Foris, Dordrecht.

Gazdar, G., E. Klein, G.K. Pullum, I. Sag (1985). *Generalized Phrase Structure Grammar.* Basil Blackwell, Oxford, UK.

Hajičová, E. (1991). A dependency-based parser for topic and focus. In M. Tomita, *Current Issues in Parsing Technology*, Kluwer, Dordrecht, 127–138.

Hajičová, E., P. Sgall & H. Skoumalová (1993). Indentifying Topic and Focus by an Automatic Procedure. In *Proceedings of EACL'93*, Utrecht University.

Harris, Z. S. (1956). String Analysis of Natural Language. Manuscript.

Hinrichs, E. & Ts. Nakazawa (1989a). Subcategorization and VP Structure in German. In E. Hinrichs & Ts. Nakazawa, *Aspects of German VP Structure – An HPSG Account.* SfS–Report–01-93, Seminar für Sprachwissenschaft, Computerlinguistik, Eberhard-Karls-Universität, Tübingen, 1993.

Hinrichs, E. & Ts. Nakazawa (1989b). Flipped out: AUX in German. Chicago Linguistic Society 25.

Huck, J.G. and A.E. Ojeda (eds.) (1987). *Syntax and Semantics. Vol.* 20; *Discontinuous Constinuency.* Academic Press, New York.

Hudson, Richard (1990). *English word grammar.* Basil Blackwell, Oxford.

Karlsson, F. (1990). Constraint Grammar as a framework for parsing running text. In *Proceedings of COLING–90*, Vol.3. Helsinki. 168–173.

Karttunen, L. and M. Kay (1985). Parsing in a free word order language. In Dowty, Karttunen, and Zwicky (eds.), *Parsing Natural Language*, Cambridge University Press, Cambridge, 279–306.

Kay, M. (1985). Parsing in Functional Unification Grammar. In D. Dowty *et al.* (eds.), *Natural Language Parsing.* Cambridge University Press, Cambridge, UK.

Melčuk, I. and N. V. Pertsov (1987). *Surface Syntax of English.* Benjamins, Amsterdam.

Netter, K. (1991). Clause Union Phenomena and Complex Predicates in German. In *Clause Structure and Word Union Variation in Germanic.* ESPRIT Basic Research Action 3175 DYANA, Deliverable R1.1.B., University of Edinburgh.

Oliva, K. (1991). *On Cases of "Fixed" Word Order in a "Free" Word Order Language.* CLAUS report No. 14, Computerlinguistik, Universität des Saarlandes, Saarbrücken.

Oliva, K. (1992a). *Word Order Constraints in Binary Branching Syntactic Structures.* CLAUS report No. 20, Computerlinguistik, Universität des Saarlandes, Saarbrücken.

Oliva, K. (1992b). The proper treatment of word order in HPSG. In *Proceedings of COLING 1992*, Nantes.

Pechmann, T. and H. Uszkoreit (1994).

Petkevič, V. (1993). Underlying Structures and Unification. *The Prague Bulletin of Mathematical Linguistics* 59–60.

Pfeiffer, O. E., M. Půček and P. Sgall (1994). Die Thema-Rhema Gliederung im Deutschen und ihre automatische Analyse. In Ursula Klenk (ed.) *Computation Linguae II* (ZDL-Beiheft 83), Franz Steiner Verlag, Stuttgart.

Pollard, I. & I. Sag (1987). *Information-Based Syntax and Semantics. Vol. 1: Fundamentals.* CSLI Lecture Notes No. 13. Stanford, California.

Pollard, I. & I. Sag (in press). *Information-Based Syntax and Semantics. Vol. 2: Agreement, Binding and Control.* CSLI Lecture Notes, Stanford, California, USA.

Reape, M. (1989). A Logical Treatment of Semi-Free Word Order and Bounded Discontinuous Constituency. In *Proceedings of the 4th Meeting of the European Chapter of the ACL.* Manchester.

Reape, M. (1990). A theory of word order and discontinuous constituency in West Continental Germanic. In Engdahl & Reape (1990).

Reape, M. (1991). Clause Union and Word Order Variation in Germanic. In M. Reape, *Word Order in Germanic and Parsing*, DYANA Deliverable R1.1.C, University of Edinburgh.

Sag, I.A. (1987). Grammatical Hierarchy and Linear Precedence. In Huck and Ojeda (1987).

Sgall, P., E. Hajičová, and J. Panevová (1986). *The Meaning of the Sentence in its Semantic and Pragmatic Aspects.* Academia, Praha and Reidel, Dordrecht.

Sleator, D. D. and D. Temperley (1992). Parsing English with a Link Grammar. Unpublished manuscript. School of Computer Science, Carnegie Mellon University, Pittsburgh, Pennsylvania.

Stucky, S. U. (1987). Configurational Variation in English: A Study of Extraposition and Related Matters. In Huck and Ojeda (1987).

Tesnière, Lucien (1959) *Éléments de syntaxe structurale.* Klincksieck, Paris.

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

Uszkoreit, H. (1986). *Constraints on Order.* CSLI Report No. CSLI-86-46. Center for the Study of Language and Information, Stanford University, Stanford.

Uszkoreit, H. (1987a). Linear Precedence in Discontinuous Constituents: Complex Fronting in German. In Huck and Ojeda, 1987.

Uszkoreit, H. (1987b). *Word Order and Constituent Structure in German.* CSLI Lecture Notes No. 8. CSLI, Stanford.

Uszkoreit, H., G. Erbach, and W. Skut (1994). *Linear Precedence Constraints in "Lean Formalisms", Part II.* Universität des Saarlandes, Computational Linguistics, and Deutsches Forschungszentrum fuer Künstliche Intelligenz, Saarbrücken.