Research on Language and Computation **2:** 209–257, 2004.

Continuous or Discontinuous Constituents? A Comparison between Syntactic Analyses for Constituent Order and Their Processing Systems

STEFAN MÜLLER
DFKI GmbH, Stuhlsatzenhausweg 3, D-66123 Saarbrücken (E-mail: Stefan.Mueller@dfki.de)

Abstract. In this paper I discuss several possible analyses for constituent order in German. Approaches that assume continuous constituents are compared with an approach that assumes discontinuous constituents. I will show that certain proposals that have been made to analyze constituent order are either not adequate or cannot be implemented with currently available systems. For the proposals that can be implemented I will discuss the amount of work a parser has to do. I then compare two implementations of larger fragments of German: the Verb*mobil* grammar and the Babel grammar. It is shown that the amount of work to be done to parse the Verb*mobil* grammar is significantly higher than the work that has to be done parsing with the Babel grammar.

Key words: German, HPSG, implementation, linearization, parsing

1. Introduction

During the last years, several grammarians have argued for linguistic descriptions of language that use the concept of discontinuous constituents (Reape, 1991, 1992, 1994a; Pollard, Kasper and Levine, 1992, 1994; Kathol and Pollard, 1995; Kathol, 1995, 2000; Müller, 1995, 1997b 1999; Richter and Sailer, 2001a; Donohue and Sag, 1999; Penn, 1999; Campbell-Kibler, 2001). Usually constituent order freedom is taken as motivation for analyses with discontinuous constituents. As an example consider the two sentences in (1).

(1) a. weil der Mann das Buch der Frau gab.
because the man$_{nom}$ the book$_{acc}$ the woman$_{dat}$ gave
'because the man gave the woman the book.'

b. weil der Mann der Frau das Buch gab.
because the man$_{nom}$ the woman$_{dat}$ the book$_{acc}$ gave

In standard HPSG (Pollard and Sag, 1994) it is usually assumed that arguments are combined with their heads in the order dative, accusative, nominative. If one assumes binary branching structures, *der Frau gab* forms a constituent in (1a). In the next projection this constituent is combined with the accusative object and the resulting constituent is combined with the nominative argument. Now consider example b: Here, *der Frau* and *gab* are not adjacent. If one allows for discontinuous constituents *der Frau* and *gab* can form a discontinuous constituent. In the next projection step this constituent can be combined with *das Buch* and then with the nominative argument. Grammars that do not allow for discontinuous constituents have to analyze (1b) in a different way. How this can be done will be examined in section 2. Apart from such reorderings I will discuss accounts for the position of the finite verb, accounts for the analysis of the predicate complex, and proposals for extraposition.

I will show that grammars for German that assume continuous constituents are difficult to process since certain information that is needed to guide the parser (valence information, for instance) is missing during the analysis of phenomena that will be discussed throughout the paper. In section 3 it will be shown that such problems do not arise during the processing of grammars that allow for discontinuous constituents.

Section 2 consists of two parts: In the first part I will discuss the analyses that have been proposed in the literature[1] and in the second part I will deal with actual implementations. Most systems that are currently available for processing HPSG grammars use a phrase structure based backbone. I will show that the analyses that assume continuous constituents either require a huge quantity of phrase structure rules that is hardly maintainable for the grammar writer or they license an enormous amount of constituents which slow down certain parsers considerably.

In section 3 I will discuss an alternative approach that allows for discontinuous constituents. I will argue that if discontinuous constituents are allowed for describing language, one can write more compact grammars. I will discuss the advantages that a grammar for German has as far as parsing is concerned and I will show how the parser that I have implemented works.

Reape (1991) notes that it is possible to develop grammars based on the concept of discontinuous constituents which span every subset of an input string. The complexity of the parsing problem for such grammars is at least exponential in both time and space. As Reape (1991, p. 62) has argued for the processing of grammars with discontinuous constituents and as Carroll (1994) has demonstrated for different parsing strategies for grammars with continuous constituents, such theoretical values are not of much help when it comes to non-toy grammars that are supposed to be used in practical applications. I will therefore compare the parsing results of two actually implemented grammars for German: one that assumes continuous constituents (the Verb*mobil* grammar[2]) and one that allows for discontinuous constituents (the Babel grammar). Statistics that were computed from parses of 24,602 utterances taken from the Verb*mobil* corpus show that the average number of passive edges built up during a parse is considerably larger for the Verb*mobil* grammar than for the Babel grammar. I will also compare the runtimes for a toy-grammar for a PP attachment fragment and the runtimes for the Verb*mobil* corpus. The system that parses the grammar with continuous constituents is much faster in processing the toy grammar. The difference in parsing the

Verb*mobil* corpus with the complete grammars is significantly smaller so that it can be concluded that the extra computational resources that are needed for processing discontinuous constituents are justified.

2. Continuous Constituents

2.1. LINGUISTIC THEORY

In what follows I will discuss various constituent order phenomena: the position of the finite verb, the permutation of arguments of a single head, the formation of verbal complexes and permutation of arguments of the heads in the verbal complex, and extraposition.

In simple phrase structure based grammars all order variation had to be encoded in the grammar rules directly. HPSG, like GPSG, divides the grammar into Immediate Dominance and Linear Precedence rules (LP rules). So, in principle, linearization issues can be dealt with separately. In GPSG, LP statements play the role of constraining order in local trees. While in HPSG larger linearization domains are possible, as is discussed in section 3, most HPSG publications have implicitly adopted the GPSG conception of applying LP rules to local trees only.

2.1.1. *Position of the Finite Verb*

In this section I will discuss two approaches to the position of the finite verb. One assumes flat structures and therefore leads to a straightforward account for finite verb position and the other one uses binary branching structures, which makes it necessary to use a verb-movement approach. As will be shown, such a verb-movement approach is computationally expensive. The flat structure approach is straightforward, but it complicates other parts of the grammar as will be discussed in later sections.

Before I discuss the two alternative analyses, I want to show the three possible positions for the finite verb in German sentences: verb first position (2a), verb second position (2b), and verb last position (2c).

(2) a. Gab der Mann der Frau das Buch?
gave the man the woman the book
'Did the man give the woman the book?'

b. Der Mann gab der Frau das Buch.
'The man gave the woman the book.'

c. daß der Mann der Frau das Buch gab.
'that the man gave the woman the book.'

Verb second sentences are usually analyzed as derived from verb first sentences by the fronting of one constituent (Erdmann, 1886; Paul, 1919). Since fronting is usually treated in the same way in grammars with continuous and in grammars

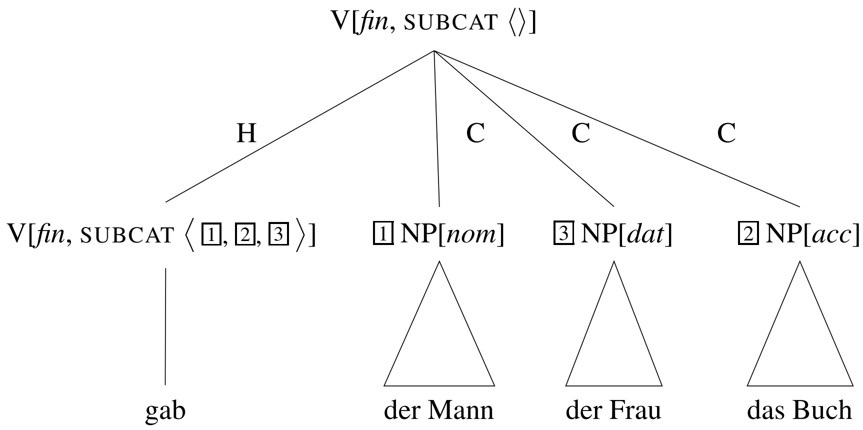

Figure 1. Flat structure: *der Mann der Frau das Buch gab* ('The man gave the woman the book').

with discontinuous constituents, I will ignore sentences like (2b) and focus on (2a) and (2c).

2.1.1.1 Flat structure

If a flat structure is assumed, the verb is in the same local tree as its adjuncts and arguments, and it can be serialized either as the leftmost or the rightmost daughter in the tree. In Figure 1, which shows an analysis for the sentence (2a), the verb *gab* ('gave') is serialized to the very left of its arguments, but a serialization of the verb at the right periphery is also admitted. Therefore examples like (2c) are accounted for.

2.1.1.2 Binary branching structures

If binary branching structures are assumed, the finite verb position is usually described by head movement analyses (Jacobs, 1986; Kiss and Wesche, 1991; Netter, 1992, 1998; Egli and von Heusinger, 1992; Frank, 1994; Kiss, 1995a; Müller and Kasper, 2000; Meurers, 2000).

As a motivation for such an analysis, Netter provides the structures in Figure 2 for the example in (3).

(3) Bringt Peter morgen die Ladung?
brings Peter tomorrow the load
'Is Peter bringing the load tomorrow?'

Without verb movement one would get two different structures for the sentence, the first structure being the mirror image of the second. The problem with these structures is that adverbs like *morgen* scope over different parts of the tree. Since

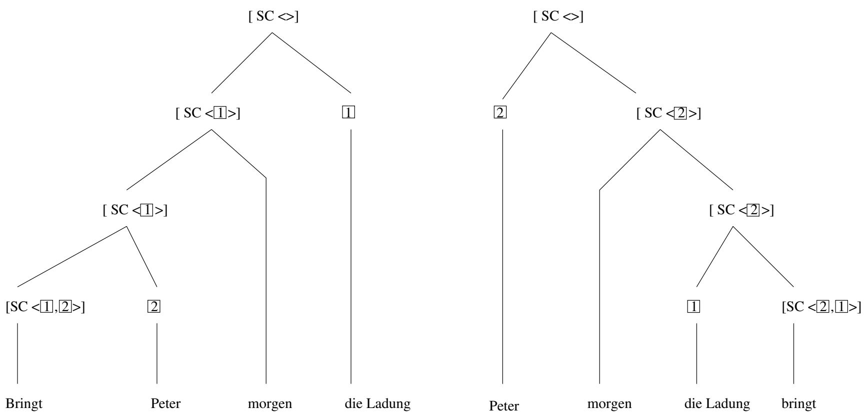

Figure 2. Unmotivated divergence of structures for verb-first and verb-last sentences.

scope in German is determined from left to right, the different structures in a verb first sentence would make wrong predictions.

(4) a. daß Peter oft wegen der Konferenz arbeitete.
that Peter often because.of the conference worked
'that Peter often worked because of the conference.'

b. daß Peter wegen der Konferenz oft arbeitete.

c. Arbeitete Peter oft wegen der Konferenz?

The two sentences in (4a) and (4b) have different readings. The first adjunct always scopes over the second. If (4c) had a structure as is shown in the first tree in Figure 2, the reading in (4b) is predicted, which is empirically wrong.[3] To cope with this problem, Netter suggests the analysis shown in Figure 3. A verbal trace (5) attracts all arguments of the verb ($\boxed{1}$).[4]

$$(5)\quad \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} \end{bmatrix} \\ \text{SUBCAT} & \left\langle \text{V[LEX+, SUBCAT } \boxed{1}\text{, CONT } \boxed{2}\text{]} \right\rangle \oplus \boxed{1} \end{bmatrix} \\ \text{CONT} & \boxed{2} \\ \textit{loc} & \end{bmatrix}$$

The trace functions as the head. It combines with the arguments of the verb ($\boxed{1}$ and $\boxed{2}$ in Figure 3), and after having done so with the verb itself.

The problem with such head-movement approaches is that the head-trace is dramatically underspecified. In particular, since the SUBCAT value is underspecified, any number and type of arguments can combine with the head-trace with the expectation that ultimately one can license this subcategorization list by finding the appropriate verbal filler for the head trace.[5] If the grammar contains other empty elements that can be combined with the verbal trace, a bottom-up parser will not terminate.

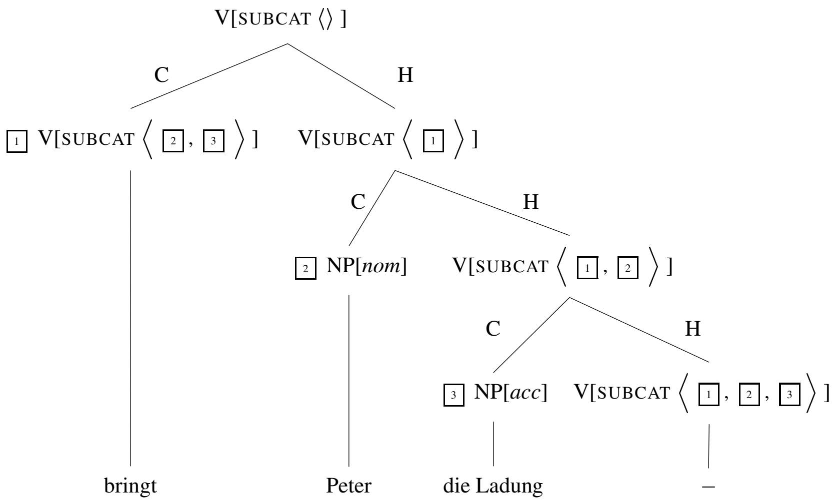

Figure 3. Analysis with verbal trace.

2.1.2. *Relatively Free Constituent Order in the* Mittelfeld

In German, arguments and adjuncts of a head can be ordered relatively freely. For instance, with a ditransitive verb like *geben* (*give*), all six permutations of the arguments are possible, provided appropriate context and intonation. This is exemplified by the verb last clauses in (6):

(6) a. weil der Mann der Frau das Buch gab.
because the man the woman the book gave
'because the man gave the woman the book.'
b. weil der Mann das Buch der Frau gab.
c. weil das Buch der Mann der Frau gab.
d. weil das Buch der Frau der Mann gab.
e. weil der Frau der Mann das Buch gab.
f. weil der Frau das Buch der Mann gab

In general, the possibility to permute constituents depends on a broad variety of interacting constraints such as animateness, weight, definiteness (see for instance Behaghel, 1930; Drach, 1937; Hoberg, 1981; Höhle, 1982; Uszkoreit, 1987).

Verbs like *kaufen* (*buy*) take four arguments (see Kunze, 1991), and as Wegener (1985) argued convincingly, some of the so-called "free datives" have to be analyzed as complements as well. Therefore *kaufen*, as is used in (7), has five arguments.

(7) Deshalb kauft Karl von Hans für fünf Mark seiner Frau ein Buch.
therefore buys Karl from Hans for five Marks his wife a book
'Therefore Karl buys a book for his wife from Hans for five Marks.'

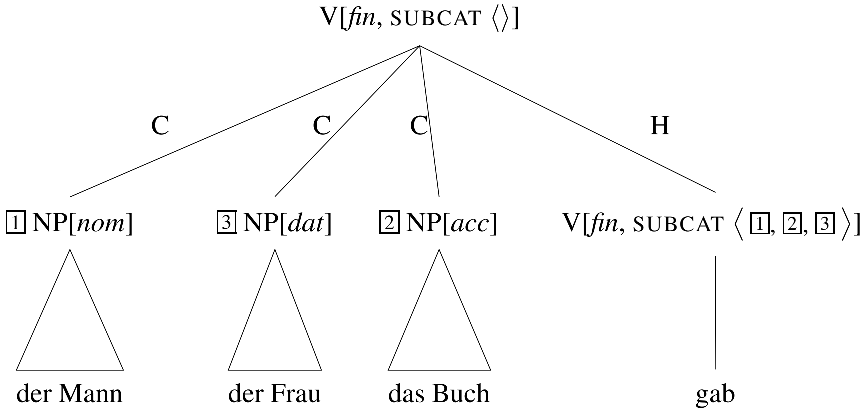

Figure 4. Flat structure: *der Mann der Frau das Buch gab* ('The man gave the woman the book').

In principle, all permutations of these five arguments are possible. For sentences with five arguments the number of possible permutations is $5! = 120$.

In the following I will discuss approaches that assume flat structures and approaches that assume binary branching structures.

2.1.2.1 Flat structures

To account for the constituent order freedom in the examples we saw in (6), Uszkoreit (1987) and Pollard (1996) suggested a flat structure. Uszkoreit and Pollard assume that linearization constraints must hold in a local tree and since all arguments in Figure 4 are in the same local tree, they can be permuted as long as no LP rule is violated.

In German adjuncts may appear in any position between the arguments of the verbs:

(8) a. Gab der Mann der Frau das Buch gestern?
gave the man the woman the book yesterday
b. Gab der Mann der Frau gestern das Buch?
c. Gab der Mann gestern der Frau das Buch?
d. Gab gestern der Mann der Frau das Buch?

Usually it is assumed that the number of adjuncts per clause is not restricted.[8] In order to account for this the dominance schema has to license both the arguments of a verb and an infinite number of adjuncts. The meaning of the complete clause has to reflect the meaning of the adjuncts contained in it. Kasper (1994) suggested a head-argument-adjunct schema that comes together with relational constraints that compute the meaning of the mother. While such schemata are a valid linguistic description they do not lend themselves easily to implementations in phrase structure based systems. I will return to this in section 2.2.

2.1.2.2 Binary branching structures

The problems that one has with flat structures disappear if one uses binary branching structures. A head is combined with one argument in a head-complement structure. Lexical heads and head-complement structures can function as head in head-adjunct structures. (1) shows the possible positions for the adverb *gestern* ('yesterday') and the appropriate binary branching structures.

(9) weil [(gestern) [der Mann [(gestern) [das Buch [(gestern) [der
because yesterday the man yesterday the book yesterday the
Frau [(gestern) gab]]]]]]].
woman yesterday gave
'because the man gave the woman the book yesterday.'

Therefore it is trivial to account for the free appearance of adjuncts in the German *Mittelfeld*.

However, it is not trivial to account for the free ordering of arguments: In HPSG as developed by Pollard and Sag (1987, 1994) the elements in the SUBCAT list that represents the valence of a head are specified in an order that corresponds to the obliqueness hierarchy of Keenan and Comrie (1977) given in (10).

(10) SUBJECT => DIRECT OBJECT => INDIRECT OBJECT => OBLIQUES => GENITIVES => OBJECTS OF COMPARISON

A valence principle cancels off elements in a strict order beginning with the most oblique element. For our example with the ditransitive verb *geben* ('give') this means that the dative object is combined with the head first, then this projection is combined with the accusative, and finally the projection is combined with the subject. Figure 5 shows the analysis of the sentence (6b) – repeated here as (11).

(11) der Mann das Buch der Frau gab.
the man the book the woman gave
'The man gave the book to the woman.'

Since *der Mann*, and *das Buch*, and *der Frau* are not sisters in a local tree, they cannot be permuted freely. There are several solutions to this problem: Gunji (1986), Hinrichs and Nakazawa (1989b), Pollard (1996), Engelkamp, Erlbach and Uszkoreit (1992), and Kiss (2001) suggested using a set rather than a list to represent valence information. The valence principle is adapted appropriately and it is possible to combine two elements in head-complement structures if the argument is an arbitrary element of the SUBCAT set of the head.

The problem with such an approach is that one gets spurious ambiguities for constructions where the head is in the middle. An example of a case where spurious ambiguities arise is the conjunction in coordinated structures, if they are treated as the head of the construction, as suggested by Paritong (1992). The phrase *Karl and Mary* could have the two structures in (12):

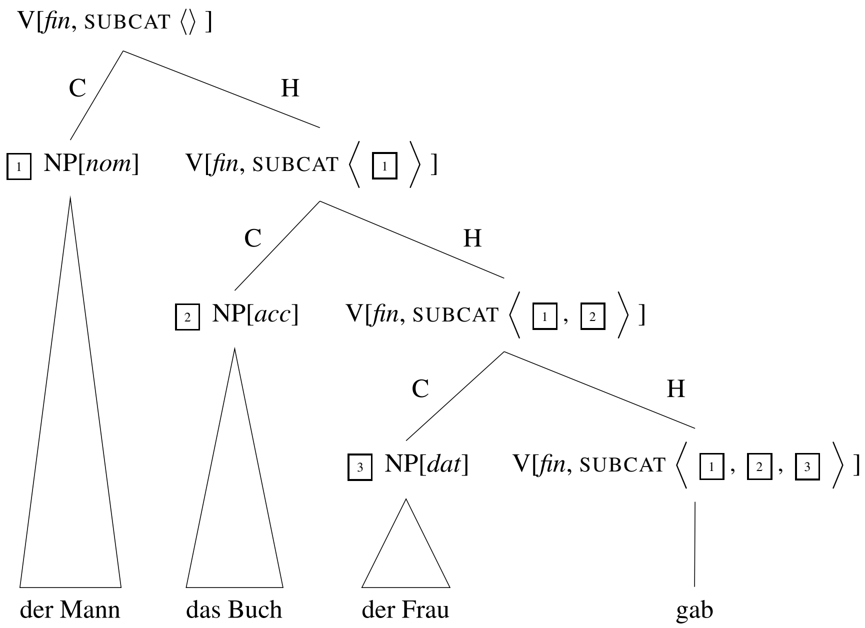

Figure 5. Binary branching structure: *der Mann das Buch der Frau gab*.

(12) a. [Karl [and Mary]]
b. [[Karl and] Mary]

In the first case the argument to the right of the head is combined with it first and then the one to the left, and in (12b) it is the other way round.

An alternative to the SUBCAT set approach is to assume that the valence representation is a list valued feature, but relax the requirement that elements have to be removed from this list in the order of their obliqueness. This approach has problems that are similar to those of the set based approach.

Another alternative was developed by Uszkoreit (1986). He suggested using a lexical rule that for each verb licenses lexical items with permuted elements in the SUBCAT list. In Uszkoreit's approach the order of elements in the SUBCAT list corresponds to the surface order of the elements.[7] This means that at least six lexical items are licensed for a ditransitive verb like *geben* ('give').[8] This considerably increases the lexical ambiguity. Furthermore, the approach has problems with spurious ambiguities that cannot be solved without stipulations. Consider the example in (13).

(13) Der Frau gab der Mann das Buch.
the woman$_{dat}$ gave the man$_{nom}$ the book$_{acc}$
'The man gave the woman the book.'

The fronting of constituents in German is usually analyzed in terms of a nonlocal dependency: A argument of the verb is removed from the SUBCAT list by the saturation by a trace, a unary projection, or by a lexical rule. The argument is introduced into a list (SLASH) and percolated up the tree and then realized to the left of the finite verb. The problem with (13) is that it can be analyzed with three lexical items that have the SUBCAT lists shown in (14), thus yielding spurious ambiguities.

(14) a. ⟨ NP[*nom*], NP[*dat*], NP[*acc*] ⟩
b. ⟨ NP[*dat*], NP[*nom*], NP[*acc*] ⟩
c. ⟨ NP[*dat*], NP[*acc*], NP[*nom*] ⟩

Only the *nom* and *acc* are realized to the right of the finite verb and *nom* and *acc* have the same order in all orderings in (14). With an analysis of extraction based on lexical rules, an order for rule application could be stipulated, i.e., the extraction lexical rule could be restricted to apply before the permutation lexical rule. An additional feature that blocks certain orders of rule applications is needed. With the other two approaches, i.e., with the trace or unary projection, one is forced to assume exception features that block the extraction of an element that was permuted by the lexical rule. A grammar that works without such technical features is, of course, preferred over a grammar that needs them.

Note that the problem would disappear if permuted elements were extraction islands. Then all reordered SUBCAT list elements could be marked as islands simultaneously blocking the extraction of the element in the SUBCAT list itself. I have shown in Müller (1999, p. 101) that permuted elements are not islands.

(15) a. [Zum Gartenvereinsvorsitzenden]$_i$ hätte wohl nur dieser Mann [das Talent _$_i$].
'Only this man had the talent to become the president of the garden society.'
b. [Zum Gartenvereinsvorsitzenden]$_i$ hätte [das Talent _$_i$] wohl nur dieser Mann.

In a lexical-rule based approach the two sentences in (15) are analyzed with a ⟨ NP[*nom*], NP[*acc*] ⟩ and a ⟨ NP[*acc*], NP[*nom*] ⟩ SUBCAT list respectively. In both cases the extraction out of the accusative NP is possible. Therefore the spurious ambiguities in analyses of (13) can only be avoided by ad hoc features.

Depending on the way nonlocal dependencies are introduced, this is a problem for the subcat-set approach and for the approach with the relaxed SUBCAT principle also. If the nonlocal dependency is introduced in syntax, i.e., by a trace or by a unary projection, it is not clear at which position in the tree the introduction has to happen.[9]

In this section I have shown that one needs either flat structures with an infinite number of daughters to account for the order of arguments and adjuncts, or one has to stipulate certain features to avoid spurious ambiguities. In section 3.1.1 I will show how a discontinuous grammar can analyze the data with binary branching structures without stipulating any additional features.

2.1.3. *The Predicate Complex*

In this section I want to discuss the formation of predicate complexes. In clauses with a predicate complex, dependents of different heads may be permuted in the *Mittelfeld* as if they were dependents of a single head. This is interesting since it means that extra machinery is needed for certain approaches to constituent order freedom in German. I will discuss argument attraction analyses that have been suggested by Hinrichs and Nakazawa (1989b) and I will show that the processing of finite clauses with the verb in initial position is problematic for most parsers regardless whether flat or binary branching structures are used.

Verbs that embed an infinitive without *zu* and verbs that select for participles form a complex with their verbal complement (Hinrichs and Nakazawa, 1989a). Furthermore, some of the verbs that select an infinitive with *zu* form a complex (Kiss, 1995). Hinrichs and Nakazawa (1989a) suggested analyzing these verbal complexes via argument attraction, essentially a lexical variant of a functional composition combining the two verbal functors (Johnson, 1986). For (16) this means that *zu lesen* and *versprochen* form a verbal complex.

(16) weil es ihm jemand zu lesen versprochen hat.
because it him somebody to read promised has
'since somebody promised him to read it.'

The complex inherits all arguments of the verbs that are involved in the complex formation, i.e., for the complex *zu lesen versprochen* we have *jemand*, *ihm*, and *es* as arguments. This verbal complex is in turn embedded under *hat*. *hat* inherits all arguments from the embedded verbal complex. Since all arguments are dependents of this verbal complex, it can be explained why these elements can be permuted in the same way as normal arguments of one single verb can be (see the discussion of (6)).[10] As an example for the formalization of argument attraction consider the lexical entry for *hat* in (17).

(17) $$\begin{bmatrix} \text{SUBCAT} & \boxed{1} \\ \text{VCOMP} & \langle\, \text{V}[\text{SUBCAT}\ \boxed{1}]\, \rangle \\ \textit{cat} & \end{bmatrix}$$

The feature VCOMP in (17) is a special valence feature that is used for the selection of elements that form a predicate complex with their head. For a motivation of such a special selectional feature see Chung (1993); Rentier (1994); Müller (1997b); Kathol (1998); Müller (2002).

The valence list of the auxiliary in (17) is totally underspecified. If the auxiliary embeds an intransitive verb, the subject of the intransitive verb becomes the subject of the verbal complex consisting of *hat* and the verb. If a transitive verb is embedded, both the subject and the object of the embedded verb become arguments of the verbal complex.

In contrast to the auxiliary, the verb *versprechen* has its own arguments and adds the arguments of the embedded verb or verbal complex:

(18)
$$\begin{bmatrix} \text{SUBCAT} & \langle \text{NP}[nom], \text{NP}[dat] \rangle \oplus \boxed{1} \\ \text{VCOMP} & \langle \text{V}[\text{SUBCAT } \boxed{1}] \rangle \\ cat & \end{bmatrix}$$

As will be discussed shortly, the fact that the valence of a verb that forms a verbal complex is not fully specified until it is combined with the embedded verb has consequences for approaches that assume continuous constituents.

Note that the number of arguments of the complex *kaufen lassen hat* in (19) is six.

(19) Hat er den Mann der Frau das Buch von Karl für fünf Mark kaufen
has he the man the woman the book from Karl for five Marks buy
lassen?
let
'Did he let the man buy the book for the woman from Karl for five marks?'

As illustrated by (20), this addition of arguments can be iterated.

(20) weil Hans Cecilia John das Nilpferd füttern helfen läßt.
because Hans Cecilia John the hippo feed help let
'because Hans lets Cecilia help John feed the hippo.'

Füttern is a transitive verb. *Helfen* takes a subject, a dative NP and a verbal complex. The complex *füttern helfen* has three arguments. *Lassen* takes a subject, an accusative object (which is identical to the subject of *füttern helfen*) and a verbal complex. The complex *füttern helfen läßt* has four arguments.

Restricting the number of arguments that a verbal complex may take is no less ad hoc than limiting the number of center self-embeddings of relative clauses. Restrictions on the number of arguments should not be part of a competence grammar. The consequence of this is that it is impossible to predict the number of arguments in a clause unless all verbs that take part in the formation of the predicate complex have been combined. In the discussion in the next two subsections I will show where the problems for continuous grammars are.

2.1.3.1 Flat structures

Hinrichs and Nakazawa (1994, p. 11) assume an immediate dominance schema for sentences with a predicate complex that is similar to the following rule.[11]

(21) $$\left[\text{SYNSEM|LOC|CAT} \begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{VAL|COMPS} & \langle\rangle \\ \text{NPCOMP} & + \end{bmatrix}\right] \rightarrow$$

H, C[SYNSEM|LOC|CAT|HEAD ¬ verb]$^+$, (C[SYNSEM|LOC|CAT|HEAD verb])

A simplified version of this rule that uses the notation adopted in the rest of the paper is shown in (22):

(22) H[SC <>] → H, C$^+$, (VC)

The C^+ stands for a set of at least one non-verbal argument. The VC stands for an optional verbal complex that is built by the verbal complex schema shown in (23).

(23) H → H, VC

A sentence with a verbal complex like (19) – repeated here as (24) – is analyzed in the following way: With the verbal complex ID rule (23), the verbal complex *kaufen lassen* is built.

(24) Hat er den Mann der Frau das Buch von Karl für fünf Mark kaufen
has he the man the woman the book from Karl for five Marks buy
lassen?
let
'Did he let the man buy the book for the woman from Karl for five marks?'

This verbal complex inherits all arguments from *kaufen* and *lassen*. In the lexical entry for *hat* it is specified that *hat* inherits all the arguments of the embedded verbal complex. The head argument schema (22) is used to combine *hat* with all its arguments and with the verbal complex *kaufen lassen*.

Now consider the problems that arise when one wants to process a rule like (22): The head can be an auxiliary like *hat*. In this case no information about the elements in the *Mittelfeld* is present, i.e., it is not known what C$^+$ stands for. This information is only present when we have parsed VC. One could argue that it would be possible to start parsing with the VC constituent, but while this would work for sentences with an auxiliary as a finite verb, it would fail for sentences like (25) where the finite verb in initial position is an argument composition verb that introduces its own arguments:

(25) Deshalb verspricht es ihm niemand zu lesen.
therefore promises it him nobody to read
'Therefore nobody promises him to read it.'

Since there should not be an upper limit for the number of elements in C$^+$, all parsers that do not process H and VC before processing C$^+$ will suffer from performance problems since the parsing process is not guided by the rule.

2.1.3.2 Binary branching structures

With binary branching structures no additional machinery is necessary to account for verbs in initial position in sentences with a verbal complex. A rule for the verbal complex like the one in (23) is sufficient. The remaining work is done by the verb movement analysis: The verb in first position takes a projection with an empty element that built the verbal complex:

(26) [Hat$_i$ [der Mann [die Frau [gesehen _$_i$]]]].

If an analysis without verb movement is assumed one has to assume that the verb in initial position combines with non-verbal arguments before it is combined with the embedded verbal complex:

(27) [[[Hat$_i$ der Mann] die Frau] gesehen].

With a lexical entry for the auxiliary *hat* like the one in (17) analyzing sentences in such a way is problematic since the valence requirements of the auxiliary are totally underspecified. If the grammar contains empty elements, a bottom-up parser will not terminate.

This means that in approaches with binary branching structures one either gets problems because of the underspecification of valence information in the verbal trace or because of the underspecification of valence information in head initial verbs that form a predicate complex.

2.1.4. *Extraposition*

In this section I will show that extraposition is a nonlocal phenomenon and that the number of extraposed elements should not be restricted in a competence grammar. The consequence of this is that bottom up parsers get problems, since the number of elements that are extraposed cannot be determined locally and therefore an infinite number of hypotheses about extraposed elements have to be considered by the parser.

The examples in (28) clearly show that extraposition is a dependency which is nonlocal in nature, although this is often denied (for instance by Jacobson, 1987, p. 62; Grewendorf, 1988, S. 281; Haider, 1996, p. 261; Rohrer 1996, p. 103):

(28) a. Karl hat mir [ein Bild [einer Frau _$_i$]] gegeben,
[die schon lange tot ist]$_i$.
'Karl gave me a picture of a woman who has been dead for a long time.'

b. Karl hat mir [eine Fälschung [des Bildes [einer Frau _$_i$]]] gegeben,
[die schon lange tot ist]$_i$.
'Karl gave me a forgery of the picture of a woman who has been dead for a long time.'

c. Karl hat mir [eine Kopie [einer Fälschung [des Bildes [einer Frau $\_{i}$]]]] gegeben, [die schon lange tot ist]$_i$.

'Karl gave me a copy of a forgery of the picture of a woman who has been dead for a long time.'

Relative clauses can be extraposed from an arbitrarily deeply embedded NP. The same holds for complement clauses, as I have shown in Müller (1999, p. 206):

(29) a. Ich habe [von [der Vermutung $\_{i}$]] gehört,
[daß es Zahlen gibt, die die folgenden Bedingungen erfüllen]$_i$.

'I have heard of the assumption that there are numbers for which the following conditions hold.'

b. Ich habe [von [einem Beweis [der Vermutung $\_{i}$]]] gehört,
[daß es Zahlen gibt, die die folgenden Bedingungen erfüllen]$_i$.

'I have heard of a proof of the assumption that there are numbers for which the following conditions hold.'

c. Ich habe [von [dem Versuch [eines Beweises [der Vermutung $\_{i}$]]]] gehört,
[daß es Zahlen gibt, die die folgenden Bedingungen erfüllen]$_i$.

'I have heard of the attempt to prove the assumption that there are numbers for which the following conditions hold.'

The example in (30) is a corpus example where a sentential complement is extraposed over two NP borders:

(30) Für das Volk der Deutschen Demokratischen Republik ist dabei [die einmütige Bekräftigung [der Auffassung $\_{i}$]] wichtig, [daß es die Interessen des Friedens und der Sicherheit erfordern, daß alle Staaten gleichberechtigte Beziehungen auf völkerrechtlicher Grundlage zur Deutschen Demokratischen Republik aufnehmen und die bestehenden europäischen Staatsgrenzen einschließlich der Oder-Neiße-Grenze als endgültig und unantastbar anerkennen.]$_i$ [12]

'The unanimous confirmation of the opinion that the interests of peace and security require that all countries establish relationships to the German Democratic Republic on the basis of international law and that all countries accept the existing state borders including the Oder Neiße border as final and inviolable is important for the people of the German Democratic Republic.'

There are several approaches to extraposition in HPSG that treat it as a nonlocal dependency using the same nonlocal mechanism that accounts for extraction to the left[13] (Keller, 1994, 1995; Bouma, 1996). Since extraposition behaves in some respect differently from extraction to the left, a different feature called EXTRA is used for these nonlocal dependencies. This makes it possible to specify constraints on both kinds of dislocation without any interferences.

To analyze a sentence like (31) an element is introduced into the EXTRA list of the verb *gearbeitet* that has the form of an adjunct. Depending on the actual analysis of extraposition that element in EXTRA is percolated to the phrase *Ich habe gearbeitet* and then bound off by its filler *an diesem Abend*.

(31) Ich habe gearbeitet [an diesem Abend].
I have worked at this evening
'I worked that evening.'

The introduction of elements into the EXTRA list can be done by a trace, by a lexical rule, by a unary grammar rule, or by some lexical specification as was suggested in Bouma, Malouf and Sag (2001). As the example (32) shows, the number of extraposed elements in German is not restricted to two as was claimed by Olsen (1982).

(32) Ich habe gearbeitet [an diesem Abend] [in der Kneipe] [als Kellnerin].[14]
I have worked at this evening in the pub as barmaid
'I worked as a barmaid in the pub that evening.'

The only limit on adding additional PPs seems to be a performance limit, which should not be modeled in a competence grammar. The PPs are adjuncts and are as such not subcategorized by the verb, i.e., they are not predictable from subcategorization information of lexical heads.[15]

When arguments are extraposed they correspond to an argument of a lexical head. Since the number of arguments of a head is finite, only a finite number of arguments of a head can be introduced into the EXTRA list. The situation is different for adjuncts: In principle an infinite number of adjuncts may be combined with a single head. Since we do not want to impose restrictions on the number of extraposed elements per clause, we cannot restrict the number of elements in the EXTRA list. Therefore a system that parses bottom-up and introduces elements into an EXTRA list as hypotheses that there will be extraposed adjuncts to the right would have to compute an infinite number of projections of a head. Van Noord and Bouma (1994) have shown how a lexical rule based approach can be combined with lazy evaluation techniques in a way that makes such grammars processable.

Kiss (2000) suggested an analysis for adjunct extraposition that is different from the analysis of nonlocal phenomena that is standardly assumed in HPSG: He assumes simple adjunction of extraposed adjuncts. In his paper he claims that the extraposition of arguments is not a genuine nonlocal phenomenon. This is empirically wrong as the examples in (29) show and since his account does not extend to the nonlocal cases of argument extraposition, it has to be rejected.

2.2. IMPLEMENTATIONS

Many implemented systems that can process HPSG-like grammars, like for instance, PAGE (Uszkoreit et al., 1994) LKB (Copestake, 1999), and PET (Callmeier, 2000) do not have a linearization component. Grammars that were developed in these systems use a context free phrase structure backbone. An example for such a grammar is the grammar developed in the Verb*mobil* project (Müller and Kasper, 2000).

There are other systems that allow for relational constraints like ALE (Carpenter and Penn, 1996; Penn and Carpenter, 1999), ALEP (Schütz, 1996), and ProFIT (Erbach, 1995, 1998). Examples for grammars in the latter systems are the grammar developed by Meurers (1994) and LS-Gram by Schmidt, Rieder and Theofilidis (1996). As discussed by Meurers (1994), relational constraints can be used to factor out specifications from the individual phrase structure rules and encode them as part of definite clauses attached to several rules. Apart from using this method to encode generalizations over several rules such as the universal principles of HPSG, it can also be used to express the LP constraints in the form of relational attachments to rather underspecified phrase structure rules. As Meurers points out, the serious disadvantage for processing with such a grammar is that the information in the relational attachments is not available to guide the parsing process. For efficiency reasons, proceeding in this way therefore does not seem to be an option for implementations of non-toy grammars.

Systems that do direct ID/LP parsing as suggested by Shieber (1984) are not available.

In the following I will examine the consequences that the proposals discussed in the last section have for implementations of non-toy grammars that use systems like the ones mentioned above for grammar development.

2.2.1. *Flat Structures*

The problem with flat structures for systems with a phrase structure based backbone is that the number of rules needed to license the wide variety of flat structures is quite big. There have to be rules for intransitive verbs, for transitive verbs, for ditransitive verbs, and for verbs with four arguments. If the verb appears in initial position, there may also be a verbal complex at the right periphery of the clause. In order to account for this, the number of rules has to be increased again.

As was noted in section 2.1.2.1, adverbs can be placed anywhere between the arguments. The number of adverbs is not restricted. If this has to be reflected in the grammar rules, the number of rules is infinite. Even if one restricts the number of adverbs in an ad hoc way, the set of rules will be huge.

2.2.1.1 Multiplying out the rules

If flat structures are used, there are two ways to account for the free constituent order in the *Mittelfeld*. Firstly, a SUBCAT list with fixed order can be used and all possible permutations are represented in the rules. This leads to a large number of phrase structure rules shown in (33).[16]

(33) a. H → H[SC <A>], A
b. H → A, H[SC <A>]
c. H → H[SC <A>], A, VC
d. H → H[SC <A, B>], A, B, VC
e. H → H[SC <A, B>], B, A, VC
f. H → A, B, H[SC <A, B>]
g. H → B, A, H[SC <A, B>]

(34) a. H → H[SC <A, B>], A, B
b. H → H[SC <A, B>], B, A
c. H → H[SC <A, B, C>], A, B, C, VC
d. H → H[SC <A, B, C>], A, C, B, VC
e. H → H[SC <A, B, C>], C, A, B, VC
...
f. H → A, B, C, H[SC <A, B, C>]
g. H → A, C, B, H[SC <A, B, C>]
h. H → C, A, B, H[SC <A, B, C>]
...

The first two rules in (33) are needed for intransitive verbs with the verb (H) in initial and in final position. The third rule (33c) stands for a sentence with an intransitive verb, a verbal complex and a finite verb in initial position. The sentences in (35) are examples.

(35) a. Hat er geschlafen?
has he slept
'Did he sleep?'
b. Hat er schlafen wollen?
has he sleep want
'Did he want to sleep?

The fourth and fifth rule (33d,e) stand for intransitive verbs in sentences where the verbs in the verbal complex add an argument.

(36) a. Hat das Kind ihn schlafen lassen?
'Did the child let him sleep?'
b. Hat ihn das Kind schlafen lassen?

The sixth and the seventh rule (33f,g) stand for sentences with intransitive verbs and a verbal complex in final position that adds one argument .

(37) a. daß das Kind ihn schlafen lassen hat.
'that the child let him sleep.'
b. daß ihn das Kind schlafen lassen hat.

In these rules the verbal complex functions as a normal head. The rule therefore can be used for transitive verbs in final position as well.

(38) a. daß keine Frau diesen Mann liebt.
'that no woman loves this man'
b. daß diesen Mann keine Frau liebt.

For transitive verbs one has to add rules for head initial position (34a–b), for a verbal complex that adds one argument with the head in initial position (34c–d), and rules for sentences with a verbal complex in final position (34f–h). The '...' stand for three more rules with appropriate permutations. So for verbs with two arguments we had to add $2! + (2 * 3!)$ rules. For a maximum of n arguments we get:

(39) $$2 + \sum_{i=1}^{n} i! + 2 * (i + 1)!$$

The 2 in (39) stands for the rules (33b–c), which are not part of the recursion. If one assumes a maximum of 5 elements in the SUBCAT list of lexical verbs (compare section 2.1.2), one gets 1901 rules. Such a huge quantity of rules is hardly maintainable by a grammar writer.

2.2.1.2 A Lexical Rule

If one uses lexical rules that license lexical items with all permutations of the elements in a SUBCAT list, like the one discussed in section 2.1.2.2, the following 17 phrase structure rules are sufficient if one artificially restricts the length of the subcategorization list to five elements:

(40)
a. H → H[SC <A>], A
b. H → A, H[SC <A>]
c. H → H[SC <A>], A, VC
d. H → H[SC <A, B>], A, B, VC
e. H → A, B, H[SC <A, B>]
f. H → H[SC <A, B>], A, B
g. H → H[SC <A, B, C>], A, B, C, VC
h. H → A, B, C, H[SC <A, B, C>]
i. H → H[SC <A, B, C>], A, B, C
j. H → H[SC <A, B, C, D>], A, B, C, D, VC
k. H → A, B, C, D, H[SC <A, B, C, D>]
l. H → H[SC <A, B, C, D>], A, B, C, D
m. H → H[SC <A, B, C, D, E>], A, B, C, D, E, VC
n. H → A, B, C, D, E, H[SC <A, B, C, D, E>]
o. H → H[SC <A, B, C, D, E>], A, B, C, D, E
p. H → H[SC <A, B, C, D, E, F>], A, B, C, D, E, F, VC
q. H → A, B, C, D, E, F, H[SC <A, B, C, D, E, F>]

The rules in (40) list the daughters in the same order as they appear in the SUBCAT lists of the respective heads. The permutations are accounted for by the lexical rule. So in contrast to (33) and (34) there is no permutation of SUBCAT elements in the grammar rules in (40).

For a maximum of *n* arguments we get.

(41) $2 + 3 * n$

If one assumes a maximum of 5 elements in the SUBCAT list of verbs, one gets 17 rules. See Meurers (1994) for such a proposal and for other rules that are necessary to analyze constructions that have not been discussed here. In Meurers' grammar the number of complements + subject is restricted to three, i.e., he has rules that are equivalent to (40a–i).

Since the number of adjuncts per head is not restricted, one would need an infinite number of rules or equivalent mechanisms that compute them on the fly to account for the adjuncts in the German *Mittelfeld*. To my knowledge, there is no HPSG grammar around that uses flat structures, that can be processed with a system without lazy evaluation techniques,[17] and that accounts for adjuncts in the *Mittelfeld*.

2.2.2. *Binary Branching Structures*

In grammars that use binary branching structures (LS-Gram, Verb*mobil*) the description of adjuncts is less problematic, since in such a setup adjuncts can be sisters of every binary tree or terminal node. In LS-Gram the scope facts were not accounted for: For verb first and verb last sentences, two different structures that branch in a different way are assumed. In section 2.1.1.2 I discussed the evidence that was provided by Netter against such a treatment. The Verb*mobil* grammar handles the positioning of the finite verb in a linguistically motivated way. This results in performance problems due to the enormous quantity of edges that a bottom-up parser has to compute because of the underspecification of valance features of the head-trace in the head-movement analysis (see section 2.1.1.2).[18]

In both LS-Gram and Verb*mobil* a lexical rule is used that permutes the elements of the SUBCAT list (see section 2.1.2.2). So for each valence feature two grammar rules are sufficient. In the case of the Verb*mobil* grammar we have the two binary branching head-complement rules in (1) and the two predicate complex rules in (2):

(42) a. H → X, H[SC <X>]
b. H → H[SC <X>], X

(43) a. H → VC, H[VCOMP <VC>]
b. H → H[VCOMP <VC>], VC

2.2.3. *Extraposition*

To my knowledge, there is no system/grammar that does not use lazy evaluation and handles extraposition in an adequate way. Despite of the known nonlocality of the phenomenon, extraposition is described by local means in both the LS-Gram and the Verb*mobil* grammar.

3. Discontinuous Constituents

In the next subsection, I want to explain how the mentioned phenomena can be analyzed with discontinuous constituents. Reape and Kathol developed grammar fragments for German that also use the concept of discontinuous constituents. Reape (1996, 1992, 1994) originally introduced constituent order domains into the HPSG framework to account for the permutations of arguments in sentences with verbal complexes like the one in (16). Reape's analysis of the verbal complex is problematic for a number of reasons that have been discussed in Kathol (1998) and Müller (1999) and that will not be repeated here.

Kathol (1995, 2000) and Kathol and Pollard (1995) developed analyses for the linearization of arguments, elements of the verbal complex, and extraposed constituents that are in many ways similar to analyses in the Babel grammar. I discussed some differences in Müller (1999, 2002).

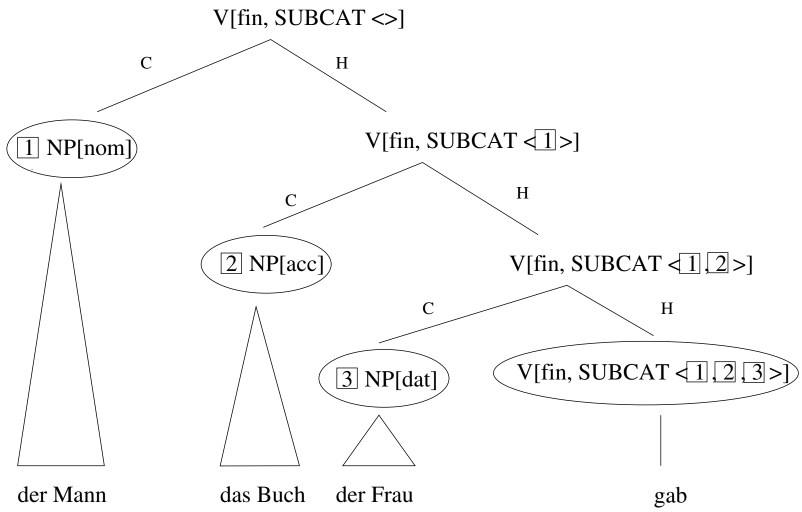

Figure 6. Dominance structure for: *der Mann das Buch der Frau gab*.

Since this paper focuses on contrasting a linearization approach with constructional approaches we will not discuss alternative analyses using discontinuous constituents here, but focus on one coherent proposal namely the one implemented in the Babel system. This also facilitates a comparison of the actual implementations.

The Babel system (Müller, 1996) can process a grammar like the one described in the following directly. The details of the Babel parser will be discussed in section 3.2.

3.1. THE ANALYSES

3.1.1. *Relatively Free Constituent Order in the* Mittelfeld

To parse the sentences we saw in the examples (6), a dominance structure is built that is shown in Figure 6 on the next page. The elements that are circled are inserted into a list which is called constituent order domain (DOM). In this list the elements can appear in any order provided that no LP-constraint is violated.

This is formalized in the following way: The lexical representation of heads contain an element in their domain list that has identical phonology and syntactic and semantic information.[19]

(44)
$$\begin{bmatrix} \text{PHON} & \boxed{1} \\ \text{SYNSEM} & \boxed{2} \\ \text{DOM} & \left\langle \begin{bmatrix} \text{PHON} & \boxed{1} \\ \text{SYNSEM} & \boxed{2} \\ \text{DOM} & \langle\rangle \\ \multicolumn{2}{l}{\textit{lexical-sign}} \end{bmatrix} \right\rangle \\ \multicolumn{2}{l}{\textit{lexical-sign}} \end{bmatrix}$$

If a head is combined with arguments or adjuncts or other dependent elements, these – the non-head daughters – are inserted as ordered unbreakable units into the domain of the head. This is formalized in (45).

(45) Construction of Domains:

$$\left[\textit{headed-structure}\right] \Rightarrow \begin{bmatrix} \text{HEAD-DTR|DOM} & \boxed{1} \\ \text{NON-HEAD-DTRS} & \boxed{2} \\ \text{DOM} & \boxed{1} \bigcirc \boxed{2} \end{bmatrix}$$

The $\bigcirc$ is the *shuffle* relation as used by Reape (1994). The *shuffle* relation holds between three lists A, B, and C, iff C contains all elements of A and B and the order of the elements of A and the order of elements of B is preserved in C. So if a and b are elements of A and a precedes b in A, it has to precede b in C too.

The PHON value of a phrasal sign is the concatenation of the PHON values of its domain elements:

(46)
$$\left[\textit{phrasal-sign}\right] \Rightarrow \begin{bmatrix} \text{PHON} & \boxed{1} \oplus \ldots \oplus \boxed{n} \\ \text{DOM} & \left\langle \begin{bmatrix} \text{PHON} & \boxed{1} \\ \textit{sign} & \end{bmatrix}, \ldots, \begin{bmatrix} \text{PHON} & \boxed{n} \\ \textit{sign} & \end{bmatrix} \right\rangle \end{bmatrix}$$

In (46), $\oplus$ corresponds to the *append* relation.

If a sentence like (6b) is analyzed, it gets exactly the same dominance structure as (6a). The examples are repeated here as (47b) and (47a), respectively.

(47) a. weil der Mann der Frau das Buch gab.
because the man the woman the book gave
'because the man gave the woman the book.'

b. weil der Mann das Buch der Frau gab.

The dominance structures are displayed in Figure 7 on the following page and in Figure 8 on the next page, respectively.

The only difference between the two analyses is that the constituents in (47b) are continuous, whereas for (47a) we get the discontinuous constituent *der Frau gab*.

Like arguments, adverbs are inserted into the domain of their head. Their free appearance in the *Mittelfeld* is therefore explained.

3.1.2. *Position of the Finite Verb*

In a domain based approach the verb is in the same linearization domain as its adjuncts and arguments. It can be serialized either to their left in verb initial position or to their right in verb final position. The dominance structure is identical in both cases. An example that involves a discontinuous projection is shown in Figure 9. Figure 10 shows the analysis in a tree structure where the leaves correspond to the surface order of the constituents. The scope facts are explained in a

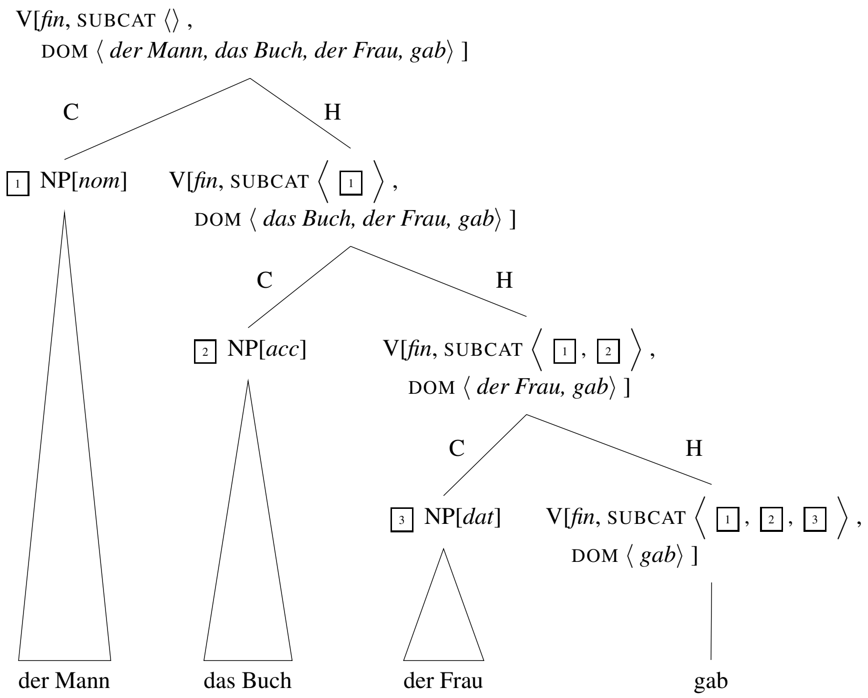

Figure 7. Dominance structure of *weil der Mann das Buch der Frau gab*.

similar way: Only the order of the combinations of adjuncts and heads is important. The order is the same regardless whether the verb is in final or in initial position. Linearization rules ensure that an adjunct in the *Mittelfeld* always has scope over all other adjuncts in the *Mittelfeld* to its right, i.e., the last adjunct in a linearization domain is combined with the verb it modifies first.

3.1.3. *The Predicate Complex*

The predicate complex is licensed by a single binary branching ID schema that combines a head with the element that is selected via VCOMP (see (23)). In the case of sentences with a predicate complex forming finite verb in initial position, the finite verb and the verbal complex at the right periphery form a discontinuous constituent. In an analysis of (25) – repeated here as (48) – *verspricht* and *zu lesen* form a discontinuous verbal complex.

(48) Deshalb verspricht es ihm niemand zu lesen.
therefore promises it him nobody to read
'Therefore nobody promises him to read it.'

After the formation of the verbal complex it is clear what the arguments of this complex head are. This is an advantage compared to systems that process grammars for continuous constituents from left to right or from right to left, since as was

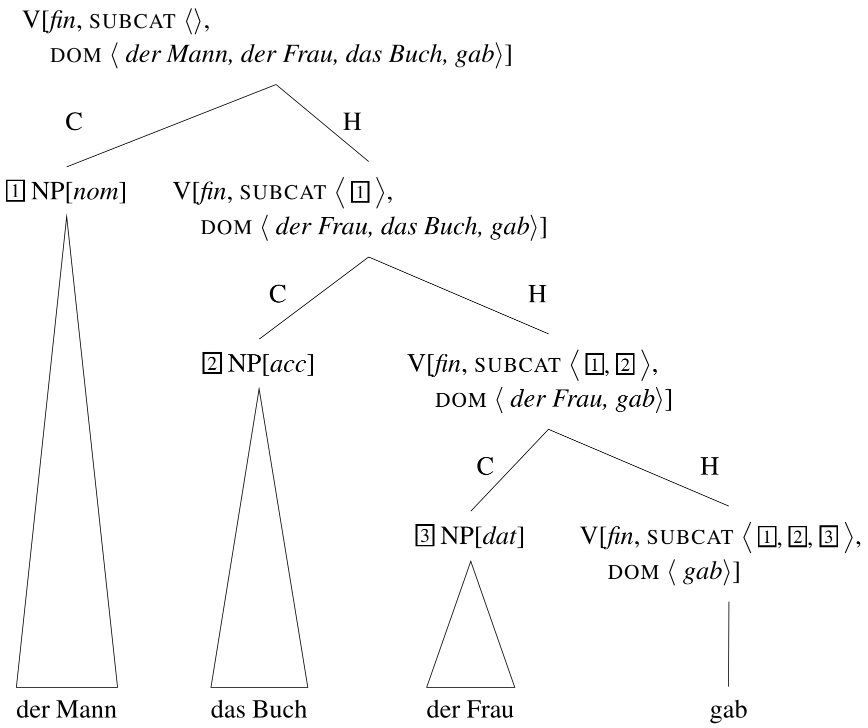

Figure 8. Dominance structure of *weil der Mann der Frau das Buch gab*.

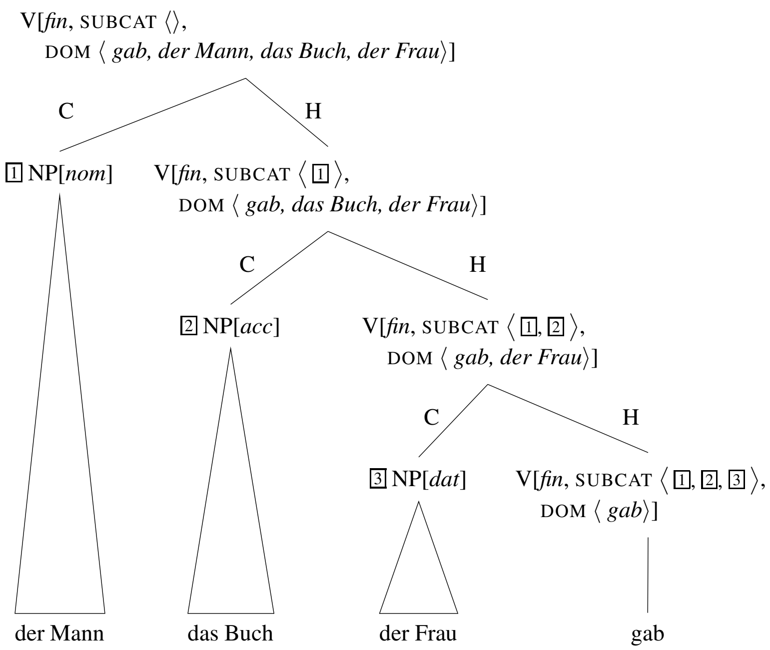

Figure 9. Dominance structure of a verb first example with discontinuous constituents: *Gab der Mann das Buch der Frau?*

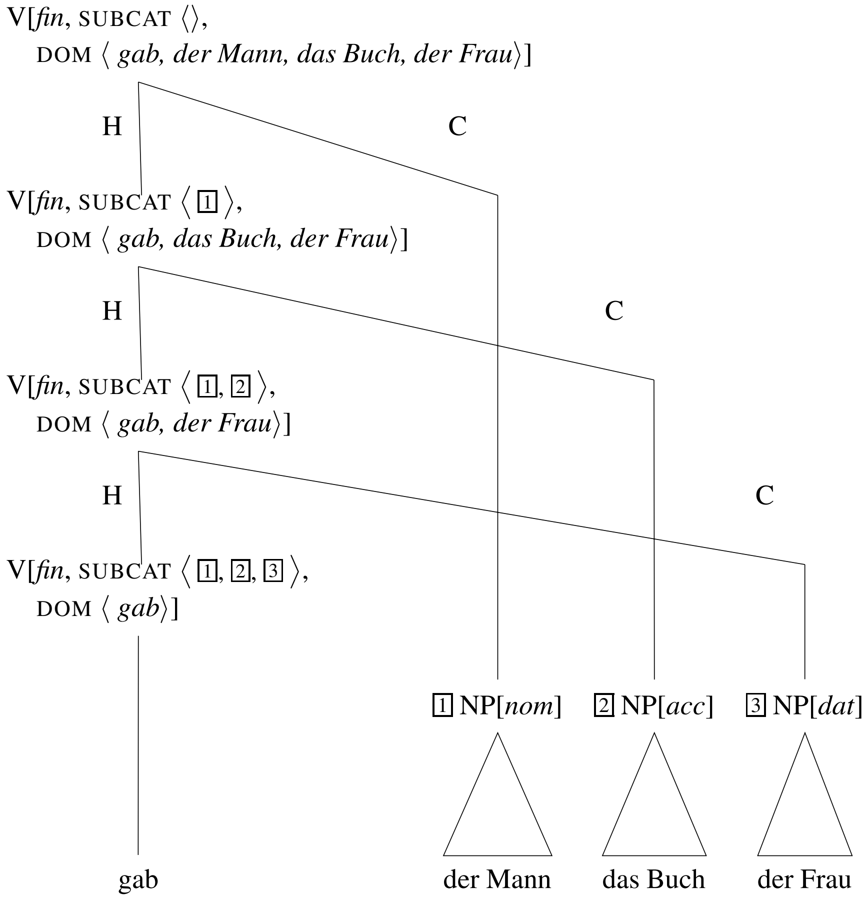

Figure 10. Verb first with discontinuous constituents: *Gab der Mann das Buch der Frau?*

shown in section 2.1.3, in such systems the number and the properties of the elements between the finite verb in initial position and the predicate complex are unknown until all parts of the predicate complex have been processed.

3.1.4. *Extraposition*

Kathol and Pollard (1995) suggested analyzing an extraposed element and its head as a discontinuous constituent. In the process of constituent order domain formation, the extraposed constituent is inserted in the higher domain separately. In grammars that do not allow for discontinuous constituents extraposition is handled by percolation of elements in a special list (EXTRA). These elements encode hypotheses that there will be extraposed material to the right. As was discussed in section 2.1.4 a competence grammar should not restrict the number of extraposed elements. Since an infinite number of adjuncts may modify a head and since certain types of adjuncts may be extraposed, an infinite number of elements have to be introduced into the EXTRA list. With an approach that allows for discontinuous constituents, such an explicit encoding of hypotheses about extraposed material is not necessary. A parser that works bottom-up combines just the material that is present in the string.

3.2. PROCESSING

In the Babel system, I use a bottom up chart parser that is generalized for grammars that employ the concept of discontinuous constituents. The parser is designed to process grammars with binary branching rules but can be easily generalized. The parser takes a specification like (49) and produces two rules from this that are shown in (50) and (51).

(49)
$$\begin{bmatrix} \text{SYNSEM|LOC|CAT|SUBCAT } \boxed{1} \\ \text{DTRS} \begin{bmatrix} \text{HEAD-DTR} & \begin{bmatrix} \text{SYNSEM|LOC|CAT|SUBCAT } \boxed{1} \oplus \langle \boxed{2} \rangle \\ \text{DOM } \boxed{3} \end{bmatrix} \\ \text{NON-HEAD-DTRS} & \left\langle \boxed{4} \begin{bmatrix} \text{SYNSEM } \boxed{2} \end{bmatrix} \right\rangle \end{bmatrix} \\ \text{DOM } \textit{lp-insert}(\boxed{4}, \boxed{3}) \end{bmatrix}$$

(50)
$$\begin{bmatrix} \text{SYNSEM|LOC|CAT|SUBCAT } \boxed{1} \\ \text{DOM } \textit{lp-insert}(\boxed{4}, \boxed{3}) \\ \text{CODE } \textit{or}(\boxed{5}, \boxed{6}) \end{bmatrix} \rightarrow \begin{bmatrix} \text{SYNSEM|LOC|CAT|SUBCAT } \boxed{1} \oplus \langle \boxed{2} \rangle \\ \text{DOM } \boxed{3} \\ \text{CODE } \boxed{5} \end{bmatrix}, \; \boxed{4} \begin{bmatrix} \text{SYNSEM } \boxed{2} \\ \text{CODE } \boxed{6} \end{bmatrix} \; \wedge \textit{and}(\boxed{5}, \boxed{6}) = 0$$

(51)
$$\begin{bmatrix} \text{SYNSEM|LOC|CAT|SUBCAT } \boxed{1} \\ \text{DOM } \textit{lp-insert}(\boxed{4}, \boxed{3}) \\ \text{CODE } \textit{or}(\boxed{5}, \boxed{6}) \end{bmatrix} \rightarrow \boxed{4} \begin{bmatrix} \text{SYNSEM } \boxed{2} \\ \text{CODE } \boxed{5} \end{bmatrix}, \; \begin{bmatrix} \text{SYNSEM|LOC|CAT|SUBCAT } \boxed{1} \oplus \langle \boxed{2} \rangle \\ \text{DOM } \boxed{3} \\ \text{CODE } \boxed{6} \end{bmatrix} \; \wedge \textit{and}(\boxed{5}, \boxed{6}) = 0$$

lp-insert is a relational constraint that takes an element and a list as arguments. The constraint either puts the element as a whole into the list or the element can be split into more objects that are then in turn inserted into the list. The splitting of elements is needed for the implementation of extraposition that was suggested by Kathol and Pollard (1995) (see section 3.1.4). LP-constraints are tested only for the elements that are newly inserted.

The rules in (50) and (51) are almost equivalent to rules that a grammar writer had to write instead of (49) in a system that uses a phrase structure backbone. The only difference is that the elements on the right-hand side may be not adjacent.

The rules in (50) and (51) are used for two cases that happen during a parse. The first case is that a constituent is entered into the chart that is a head. Then the chart is checked for an edge that matches the complement description. The other case is that a complement is entered into the chart. Then the parser searches for a head that is looking for an appropriate complement. Parsing is a simple recursive process: A word is taken from the string and the appropriate lexical description is entered into the chart. The first element of a rule is unified with the entered word and if this succeeds, the chart is searched for the second element of the rule. The result of the combination is a new description of a linguistic object which is entered into the chart. This is done for all grammar rules. The results of combinations are entered into the chart in the same way as lexical items are entered. This recursive process terminates after the last word of the input string is consumed.

This works in much the same way as the chart parsers for continuous constituents that are known so far. The only difference is that the chart is not an $n * n$ table with start and end points of vertices but a list. Two constituents that are combined may not border each other. The only thing that can be said about their position in the string without looking at constraints in the grammar is that they are not allowed to overlap, i.e., material in the string may not be used twice. To enforce this constraint, boolean vectors are used that were first suggested by Johnson (1985) and that were applied to the problem of parsing HPSG grammars that allow for discontinuous constituents by Reape (1991). A position in the string is represented by a bit. To parse utterances of the length n, one uses bit vectors of length n: If the position in a certain string is occupied by a word, the bit is set to 1, otherwise it is 0. To give an example, consider the sentence (47a) – repeated here as (52a).

(52) a. weil der Mann der Frau das Buch gab.
because the man the woman the book gave
'because the man gave the woman the book.'

b. der Frau gab

For *weil* we have the bit 0, for *der* the bit 1, for *Mann* the bit 2, and so on. The representation for the complete utterance in (52a) is shown in (53a).

(53) a. 1 1 1 1 1 1 1 1

b. 0 0 0 1 1 0 0 1

As was discussed in section 3.1 we need *der Frau gab* as a constituent in the analysis of (52a) (see Figure 8). The bit vector representation of this discontinuous phrase is shown in (53b).

If two passive edges in the chart are to be combined, the logical *and* of their bit vectors has to be 0 since edges are not allowed to contain the same material.

For instance, *der Mann der Frau* ('the man of the woman') can be analyzed as a complex NP with a possessive genitive. The bit vector representation of this NP as part of the utterance (52a) is (54).

(54) 0 1 1 1 1 0 0 0

The logical *and* of the bit vector for *der Frau das Buch gab* which is shown in (55a) and the bit vector for *der Mann der Frau* which was given in (54) is (55b).

(55) a. 0 0 0 1 1 1 1 1
b. 0 0 0 1 1 0 0 0

In (56b), the bits for all words that are used in both strings are 1. Since (55b) contains two bits that are not 0, the value of the logical *and* is not 0 and therefore the combination of *der Mann der Frau* and *der Frau das Buch gab* is not computed by the parser.

If the combination of two edges succeeds, the resulting edge has as its code the logical *or* of the code of the edges on the right hand side of the rule. As an example, consider the combination of *der Frau gab* which has the bit vector in (53b) with *das Buch* which has the representation in (56):

(56) 0 0 0 0 0 1 1 0

The bit vector for *der Frau das Buch gab* contains a 1 at every place where we have a 1 in one of the two constituents that are combined. The result of the logical *or* was already given as (55a).

The checking of the logical *and* and the computation of the new bit vector for the result of a rule application is integrated into the rules in (50) and (51).

A parser for grammars that parse continuous constituents can use the string positions of chart items to search for elements that are adjacent to an element that is to be combined with other chart items. This is not possible with parsers that allow for discontinuous constituents. However, Reape suggested an implementation of the bit vectors as Prolog terms and appropriately computed Prolog terms can be used for indexing: Suppose we are looking for elements that can be combined with *der Frau gab*. This means that we search for an element that has a zero at the places where we have a one in (56b) and that has a zero or a one at the places where we have a zero in (56b). The representation in (57) matches exactly those elements that are candidates for a combination with the constituent *der Frau gab*.

(57) _ _ _ 0 0 _ _ 0

Each '_' is compatible with any value. With this representation of the bit vectors, the non-overlapping test can be performed using simple Prolog unification. This test is applied before any feature structures are unified.

The Babel parser does not need lazy evaluation techniques like the parser described by Kasper, Calcagno and Davis (1998) and therefore the system runs on all Prolog versions that have the functionality described in Clocksin and Mellish (1984).

The parser does not use any elaborated method to reduce the search space. One method that could be applied in the Babel parser is head corner parsing (possibly together with selective memoization (van Noord, 1997). Van Noord (1993, 1997) provides algorithms that can be used for parsing grammars that allow for discontinuous constituents. Since the focus of my research is linguistic theory and its formalization, the parsing algorithm used in the Babel system is only a minimal algorithm necessary to parse linearization grammars. While this means that the parse times of Babel are not optimal, it also means that the number of edges that is constructed by the Babel parser is comparable to the number of edges that is constructed by systems like LKB, PAGE, and PET, which also work bottom-up without any special strategies to restrict the number of passive edges that are entered into the chart.

Oepen and Carroll (2000) describe techniques for local ambiguity packing that dramatically reduce the number of chart items. Oepen and Caroll do not store edges in the chart if they are subsumed by edges that have already been entered into the chart. Such techniques can also be applied to the parser described here. As Caroll and Oepen mentioned, it is important not to duplicate the parse tree construction in feature structures, but leave it to the parser. Systems like PAGE, LKB, and PET cut out the DAUGHTERS features before they store newly constructed elements in the chart. The daughters features can also be cut out in linearization grammars, but the domain elements are important for enforcing linear precedence constraints. For better illustration of this point let us assume a linearization rule that says that PPs have to follow NPs.

(58) weil über diesen Witz keiner lacht.
because about this joke nobody laughs
'because nobody laughs about this joke.'

The first element that is constructed is *über diesen Witz lacht*. When this phrase is combined with *keiner*, the linearization component has to be able to check whether the NP is to the left or to the right of the PP, i.e., it has to have access to the domain elements of the head. Access to domain elements of domain elements is not required though. Therefore it is not necessary to reconstruct the complete recursive structure of the domain elements in order to check LP constraints.

The number of passive edges created by the PET system and by the Babel system that will be compared in section 4.2 was obtained without local ambiguity packing.

4. Comparison

In the following, the Verb*mobil* grammar (Müller and Kasper, 2000) will be compared with a grammar that uses discontinuous constituents: the grammar of the

Babel system (Müller, 1996), which was developed earlier and which was extended to cover the Verb*mobil* data. Both grammars provide an analysis for approximately 80 % of the grammatical input in the Verb*mobil* domain (appointment scheduling and trip planning). The remaining 20 % are cases of ellipsis (4.7 %), asymmetric coordinations and gapping (1.4 %). 13.9 % of the utterances that could not be parsed consists of a set of minor phenomena that are not treated in the grammars, each below 1 %. One example of the latter are examples with a missing determiner as shown in (59).

(59) a. Achte geht sehr schlecht.
eighth goes very bad
'The eighth (of a certain month) doesn't suit me well at all.'

b. Das Hotel Cristal Hannover ist vielleicht ein bisschen teuer, hat
the hotel Cristal Hanover is perhaps a bit expensive has
dafür aber Sauna, Solarium, Bar und Bistro.
therefore but sauna solarium bar and bistro
'The hotel Cristal Hanover might be a bit expensive, but it does have a sauna, a solarium, a bar, and a bistro, after all.'

Sentences like (59) could easily be covered if determiners were made optional for the numeral in (59a) and the nouns in (59b), but since (59a) seems to be a marked domain specific utterance that should not be admitted in a German grammar and since allowing the omitance of the determiner for nouns like *Sauna* would admit ungrammatical sentence like (69), I have chosen not to do so.

(60) * Ich gehe in Sauna.
I go into sauna
Intended: 'I go into the sauna.'

For a list and some discussion of the utterances not covered by the Verb*mobil* grammar see Müller and Kasper (2000).

Both grammars are competence grammars of German that are supposed to reject ungrammatical input. The robustness that is needed for the processing of spoken input was taken care of in separate modules of the processing system of Verb*mobil*. When an utterance was not parseable, all maximal projections were passed to a component that did robust semantic processing (Pinkal, Rupp and Worm, 2000). Single words that did not belong to any maximal projection were also passed to this component. Of course the same strategy can be used for the grammar with discontinuous constituents.

4.1. PROPERTIES OF THE COMPARED GRAMMARS

The grammar for the system that can only process continuous constituents is more restricted. From a linguistic point of view this is not satisfying, but from a practical point of view these restrictions turned out not to be very limiting in the Verb*mobil* domain. I expect this to be different for other domains though. For instance, in newspaper text a broad variety of grammatical constructions can be found that was less important for covering the Verb*mobil* data. In the following subsections I will discuss some restrictions that have been imposed on the Verb*mobil* grammar. Since the Verb*mobil* grammar was used in a system that had strict time limits, several restrictions were necessary. Therefore certain costly phenomena were not covered (see Müller and Kasper (2000) for some discussion). In the Verb*mobil* project the PAGE system was used for processing the German, an English and a Japanese grammar. It was a project requirement that all three grammars could be processed by the same system. It was therefore impossible to use a system that could process grammars that allow for discontinuous constituents. After the end of Verb*mobil* I adapted the grammar for LKB (LKB has type inference and no disjunction) and the grammar can now be processed with the PET system, which is a reimplementation of LKB in C. PET needs less memory and is more efficient than the LISP systems. It was therefore possible to increase the coverage of the grammar.

The purpose of the Babel system is grammar verification only. Therefore no compromises had to be made in the first place.

In what follows I will provide a list of differences between both grammars. The list of differences is not complete. Its purpose is to show that the Babel system provides analyses of phenomena which are very costly in general and which are not covered by the Verb*mobil* grammar. This is relevant when it comes to the comparison of the results in section 4.2.

4.1.1. *Position of the Finite Verbs and Number of Arguments*

Both grammars use only binary and unary branching rules. The grammar with continuous constituents accounts for the finite verb position by a head movement analysis that combines Netter's analysis (discussed in section 2.1.1.2) and a proposal made by Kiss (1995). Unary projections which are a partial evaluation of the verbal trace with respect to the rules in the grammar are used instead of an empty element since such rules can better be controlled by the parser. A lexical rule licenses a special lexical item for verbs that occur sentence initially. Such verbs are subcategorized for one complement, namely the verb final projection that corresponds to Netter's projection of a verb final trace (see Kiss, 1995, p. 72). The length of the SUBCAT list of the verbal projection that is introduced by the unary rule is limited to five elements and the syntactic properties of these elements are specified in a way that excludes all combinations of the verbal projection with unsaturated arguments and with elements that can never be arguments of verbs (for instance determiners). The property of being a possible argument of a verb is encoded in the type hierarchy. This is an additional effort that is necessary for grammars with verb movement analyses and is not motivated linguistically.

Note that despite the limitation of possible arguments to five for lexical verbs, there is no problem with recursion over verbal complexes that add arguments (see section 2.1.3). During the analysis of (61), the verbal complex *füttern helfen* is built first.

(61) Läßt Hans Cecilia John das Nilpferd füttern helfen?
lets Hans Cecilia John the hippo feed help
'Did Hans let Cecilia help John to feed the hippo?'

It has three arguments. It is projected by a unary verb movement rule that introduces the option of an additional argument that might be added by the verb in initial position.

4.1.2. *Extraposition*

As discussed in section 2.2.3, existing grammars with continuous constituents usually only treat local extraposition. This alone is empirically wrong. Furthermore, they use rules that are triggered by material in the right periphery of the main clause.

(62) a. Ich schlage vor, daß wir uns am Montag treffen.
I suggest PART that we us on Monday meet.
'I suggest we meet on Monday.'

b. Sie hatten doch vorgeschlagen, sich am Montag zu treffen.
you had but suggested self on Monday to meet
'But you suggested we should meet on Monday.'

In (62) the right periphery is marked by a particle or by a non-finite verb. As (63) shows, the right periphery may be unmarked.

(63) Der Mann gibt der Frau das Buch, die er liebt.
the man gives the woman the book who he loves
'The man gives the book to the woman he loves.'

Due to performance problems, such cases are not handled by the Verb*mobil* grammar.

4.1.3. *Extraposition with Antecedent*

The grammar with continuous constituents does not handle sentences like those in (64), in which a constituent has been extraposed leaving a correlate expression behind in the *Mittelfeld*.

(64) a. *Das* wäre dann zum Beispiel möglich, daß wir um achtzehn Uhr
this would then for instance possible that we at eighteen clock
vierzig abfliegen.
forty fly
'So it would be possible to depart at twenty to seven.'

b. *Das* würde mir auch passen, am dreiundzwanzigsten und
this would me also suit at.the twentythird and
vierundzwanzigsten April den Bericht abzufassen.
twentyfourth April the report to.write
'Writing the report on April the 23rd and 24th would suit me too.'

c. Denn wir sind ja *da*zu angehalten, möglichst wenig Spesen
for we are yes there.to encouraged possibly few expenses
auszugeben.
to.spent
'For we are encouraged to have as few expenses as possible.'

d. Oder sollten wir es vielleicht *so* machen, daß wir uns gleich
or should we it possibly so make that we us immediately
mittwochs abends in Hannover treffen?
on.Wednesday at.evening in Hanover meet
'Or would it be better to meet each other in Hanover on Wednesday evening?'

Pollard and Sag (1994, pp. 149–150) suggested handling similar constructions in English via a lexical rule that introduces an expletive *it* and a sentential complement for each verb that selects a sentential subject. As illustrated by (65) below, there are good reasons for assuming that the construction is more general, at least in German.

(65) a. Der Mann hat *es* gehaßt, daß er nie seinen Zug gekriegt hat.
the man has it hated that he never his train got has
'The man hated always missing his train.'

b. Der Mann hat *es* gehaßt, immer zu spät zu kommen.
the man has it hated always too late to come
'The man hated always arriving late.'

c. *Es* hat den Mann geärgert, daß er nie seinen Zug gekriegt hat.
it has the man annoyed that he never his train got has
'The man was annoyed that he never caught his train.'

d. *Es* hat den Mann geärgert, immer zu spät zu kommen.
it has the man annoyed always too late to come
'The man was annoyed that he always arrived late.'

e. Der Mann hat sich *da*rüber geärgert, daß er nie seinen Zug kriegt.
the man has self there.about annoyed that he never his train gets
'The man was annoyed that he never caught his train.'

f. Der Mann hat sich *da*rüber geärgert, immer zu spät zu kommen.
the man has self there.about annoyed always too late to come
'The man was annoyed that he always arrived late.'

Antecedent *es*/*das* and *da(r)*+Preposition can function as arguments where the head selects for subjects, objects, or prepositional complements with an appropriate meaning. They can appear together with subordinated clauses and with *zu* infinitives. If one follows the lexical rule based approach, this means that for a large class of verbs there have to be new lexical items. Note that the antecedent elements can be permuted as other subjects or complements in the *Mittelfeld*. So if lexical rules for permutation are used one gets an enormous quantity of additional lexical items.

The alternative to a lexical rule based approach would be to assume that the clauses are dependents of the antecedent element. But with this assumption one is forced to implement a proper treatment of extraposition as a nonlocal dependency.

In the case of pronominal adverbs as in (64c) it can be shown that the clause has the *da* as its antecedent. If this is to be captured adequately a proper treatment of extraposition is also necessary.

In the Babel grammar, I assume that *es daß er nie seinen Zug gekriegt hat* forms a discontinuous constituent, the *es* being the head of the construction (see Müller, 1999). The *es* and the clause or an infinitive are combined directly forming a discontinuous constituent. No introduction of hypotheses about extraposed constituents is necessary. If the sentence does not contain a *daß*-clause, nothing happens.

4.1.4. *Optional Coherence*

In section 2.1.3, I discussed the sentence (16), repeated here as (66a).

(66) a. weil es ihm jemand zu lesen versprochen hat.
since it him somebody to read promised has
'since somebody promised him to read it.'

b. weil ihm jemand versprochen hat, es zu lesen.
since him somebody promised has it to read
'since somebody promised him to read it.'

In (66a), the verbs *zu lesen*, *versprochen*, and *hat* form a verbal complex, while in (66b) they do not. The construction in (66a) is called a coherent construction and the one in (66b) an incoherent one. *Versprechen* is a verb that allows for both coherent and incoherent constructions. Verbs that govern an infinitive without *zu* obligatory construct coherently (see Bech (1955) for terminology and the general observations and Kiss (1995) for an HPSG analysis).

As I argued in Müller (1999), the optional coherence should not be treated as a subcase of coherence. In the Babel grammar an analysis for both constructions is implemented, i.e., both sentences in (66) can be analyzed. In the grammar with continuous constituents there is only the lexical entry for the incoherent construction. The other lexical entry is not specified in the grammar because of performance considerations. Therefore (66a) cannot be analyzed.

4.1.5. *Preposition Stranding*

In the Verb*mobil* grammar only a subset of the instances of preposition stranding is handled.

(67) a. Da müßten wir wohl noch einen Termin für festmachen.[20]
there must we well yet an appointment for solid.make
'We still have to schedule an appointment for this.'

b. Ah ja, da hatte ich schon von gehört.[21]
ah yes there had I already of heard
'Ah yes, I had already heard about that.'

The cases in (67) where the pronoun *da* is placed in the sentence initial position are accounted for by a nonlocal dependency.

(68) a. Ich hoffe, Sie haben da auch Lust zu.[22]
I hope you have there also lust to
'I hope you also feel like doing this.'

b. Also, da muß ich erst mal gucken, ob ich da überhaupt zuständig für bin.[23]
well there must I first look whether I there at.all responsible for am
'Well, first I have to check whether I am responsible for this.'

Examples like (68) where the pronoun is inserted into the *Mittelfeld* are not analyzed. In the Babel grammar I use a special grammar rule to account for sentences like (68) (Müller, 1997a, 1999). The analysis requires two elements to be admitted in SLASH. In the Verb*mobil* grammar only one element is allowed in SLASH. This reduces the search space for the Verb*mobil* grammar.

4.1.6. *Modal Infinitives*

Modal infinitives with *sein* are not covered in the grammar with continuous constituents.

(69) a. Deswegen wäre vielleicht das zweite von Ihnen genannte vorzuziehen.[24]
therefore would possibly the second by you called prefer
'Therefore the second one you mentioned might be preferable.'

b. Das ist leider nicht zu machen.[25]
this is unfortunately not to make
'Unfortunately this cannot be done.'

The auxiliary/copula *sein* is a highly ambiguous word and each new entry for the copula increases parse times and the search space for all sentences that contain this word. The Babel grammar contains such an additional auxiliary.

4.1.7. *Depictives*

The Babel grammar can handle depictive secondary predicates.

(70) Ich schreibe den Termin unbesehen ein.
I write the date unseen in
'I'll enter the date unchecked.'

Depictives can refer to subjects, direct objects, and indirect objects (Müller, 2001, 2002). The Babel grammar uses a lexical rule that licenses four additional lexical items for each predicative adjective. This, of course, yields a considerable enlargement of the search space of every sentence that contains a predicative adjective, whether it is used as depictive predicate or not.

4.2. RESULTS

After having shown that the Verb*mobil* grammar has less coverage than the Babel grammar, especially for expensive phenomena, I will now show that the work that has to be done by the parser in Babel is less than what is needed for parsing with the Verb*mobil* grammar.

Both systems parsed the utterances of the CDs 1, 3, 4, 14, 15, 20, and 32 of the Verb*mobil* project. These CDs contain 24.602 utterances. Due to space limitations for the parsing process, a limit for the number of passive edges was set (100.000 for the PET system and 9.000 for Babel). Both parsers fail on the same number of sentences because of this restriction. Both grammars in average license six readings per utterance. The average number of readings is not completely equal since the edge limit effects the average number of readings. A highly ambiguous sentence maybe parseable in one system, but nonparseable in the other one.

4.2.1. *Search Space*

Figure 11 shows the number of passive edges that are created during a full parse relative to the utterance length.[26,27] The figure shows that the number of passive edges grows more quickly in a system with continuous constituents. The figure only displays the results for sentences up to 21 words. The curve continues to grow in a similar fashion for longer sentences, but it is not smooth. This is due to the fact that the corpus consists of spoken language and longer utterances are not very frequent. The average utterance length is 7 words. Figure 12 shows the number of sentences with 17 to 24 words.

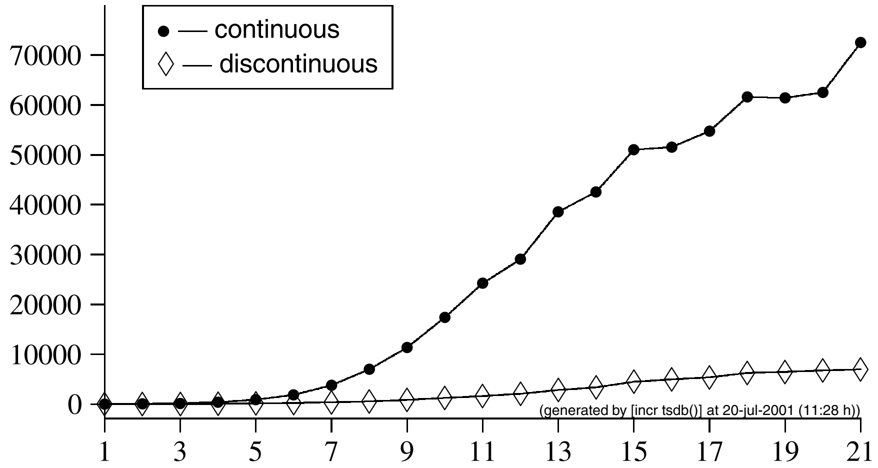

Figure 11. Passive edges generated by the grammars during a full parse.

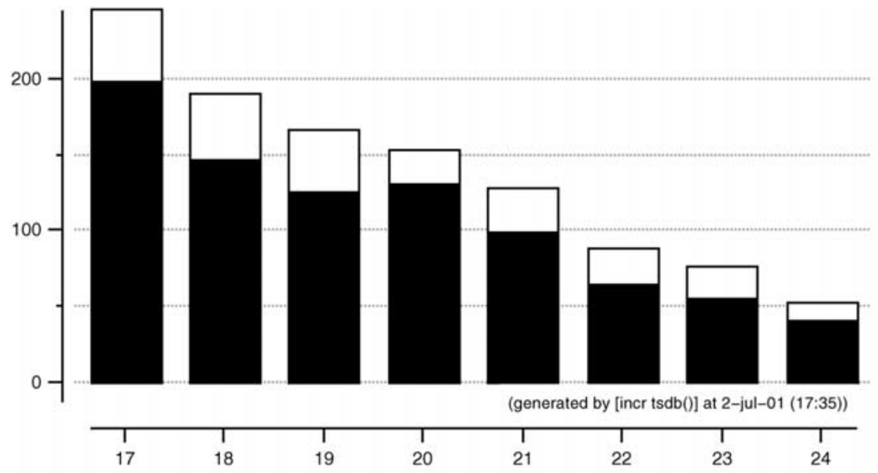

Figure 12. Number of sentences of a certain length: black = grammatical, white = ungrammatical.

4.2.2. *Parse Time*

The Figures 13 and 14 show the parse times for sentences that were not affected by edge limits for Babel and PET, respectively.

They show that PET is faster by a factor of approximately 13 for longer sentences. However, since the two systems are implemented in different programming languages the results are not really comparable. A reviewer suggested the implementation of two toy grammars that license NPs and PPs. The parse times of these grammars can be compared and the differences show differences in the implementations of the two parsers. Figures 15 and 16 show the parse times for two toy-PP-grammars.

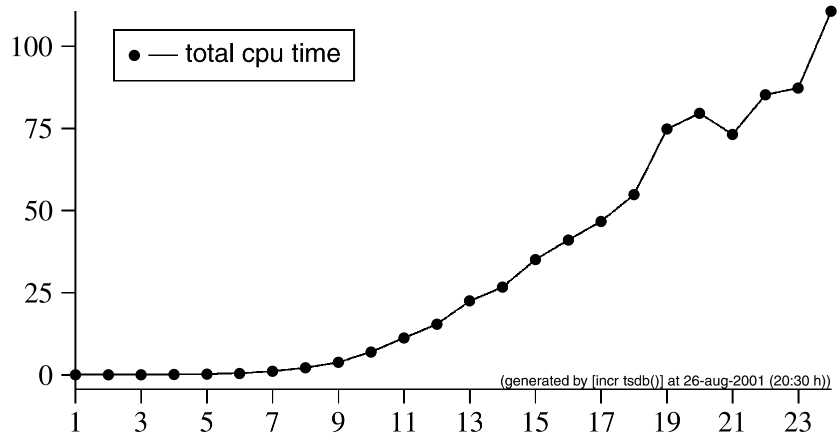

Figure 13. Parse time in seconds for Babel (Verb*mobil* data).

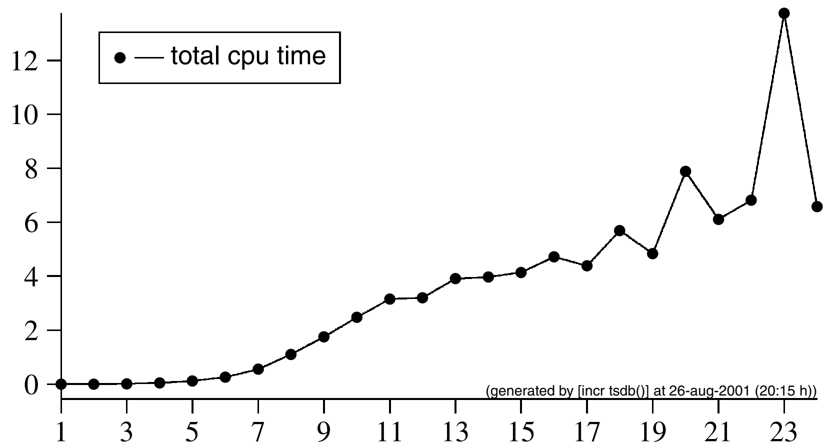

Figure 14. Parse time in seconds for PET without packing (Verb*mobil* data).

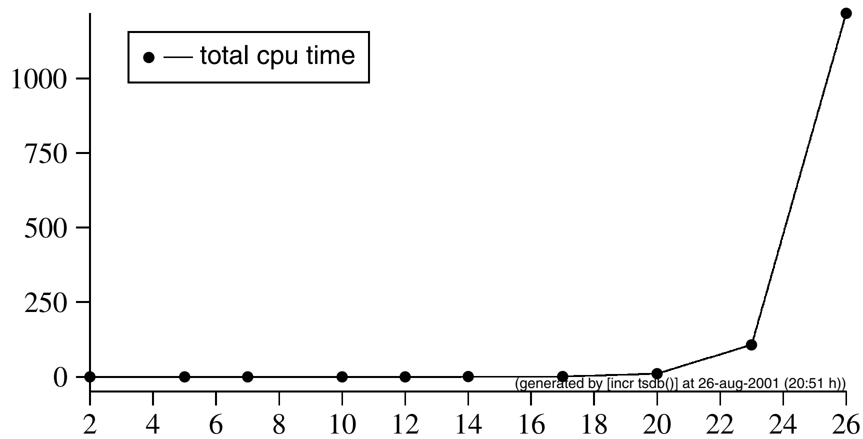

Figure 15. Parse time in seconds for Babel (PP toy-grammar)

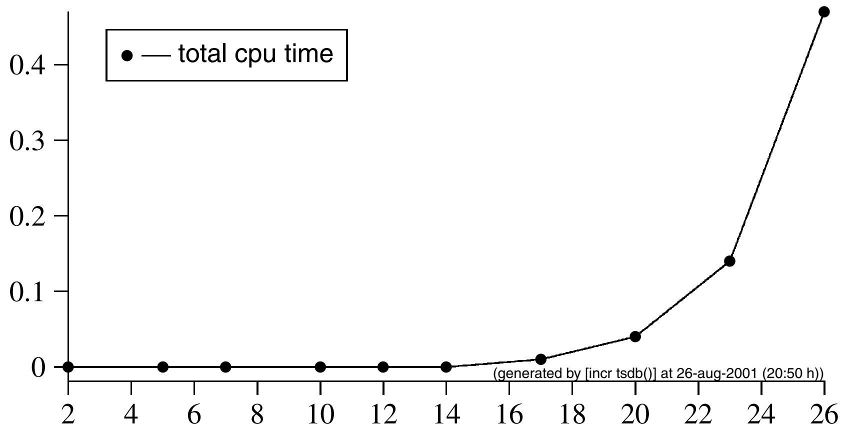

Figure 16. Parse time in seconds for PET without packing (PP toy-grammar)

For sentences with 9 prepositions (sentence length 26), PET is 2591 times faster than Babel.

As Figures 13 and 14 show, the difference between Babel and PET is much smaller (around 13) when the full grammars are processed. Therefore it seems promising to generalize systems like PET so that grammars that allow for discontinuous constituents to be processed.

5. Summary and Outlook

It has been shown that the domain-based grammar is better suited to parse sentences with a sentence length of up to 21 words. The search space for such sentences is considerably smaller. The Verb*mobil* corpus contains sentences with a length of up to 60 words, but the number of sentences with more than 21 words is not significant (2.54% of 24,602 utterances). It remains to be seen how different grammars behave when it comes to sentences longer than the ones one finds in spoken language.

An interesting direction for further research is to implement a grammar that can be processed with one system in two different modes: with and without discontinuous constituents. Since the part of the Verb*mobil* grammar that accounts for constituent order does not interfere with other parts of the grammar it can be separated easily and the other linearization module can be plugged in. A system that can process both kinds of grammars is currently under development (Fouvry and Meurers, 2000; Daniels and Meurers, 2002). With such a system it will become possible to compare runtimes, which is not possible now since there are too many varying parameters such as system specific memory requirements.

Acknowledgements

Thanks to Uli Callmeier, Dan Flickinger, Anette Frank, Bob Kasper, Tibor Kiss, Kai Lebeth, Detmar Meurers, Gerald Penn, Adrian Simpson, and anonymous reviewers for discussion and/or comments on earlier versions of this paper. I gave talks about constituent order and parsing at the Workshop on Efficient Processing with High-Level Grammar Formalisms, Schloß Dagstuhl, 1999, and the ESSLLI 2000 workshop on Linguistic Theory and Grammar Implementation in Birmingham, and at the Institut für Linguistik/Computerlinguistik in Potsdam.

I thank Potsdam for the invitation and the audiences of all three talks for discussion.

The research carried out for this paper was in part supported by a research grant from the German Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (BMBF) to the Project Verbmobil, Part 2.4, Grant Number 01 IV 701 V0) and in part by a research grant to the DFKI project WHITEBOARD (“Multilevel Annotation for Dynamic Free Text Processing”), FKZ 01 IW 002.

The paper was completed at the Institut für Germanistische Sprachwissenschaft of the Friedrich Schiller University Jena.

Notes

[1] Due to space limitations I cannot provide an introduction to HPSG here. I therefore have to presuppose some familiarity with HPSG in general and with the literature on HPSG for German in particular. For HPSG in general see (Pollard and Sag, 1987, 1994). Issues concerning German syntax in the HPSG framework are discussed in Nerbonne, Netter and Pollard (1994) and Meurers and Kiss (2001) and in some of the literature cited elsewhere in this article.

[2] On Verb*mobil* see Wahlster (2000).

[3] This argument is only valid if scope is determined with respect to the tree structure. In HPSG scope could also be determined with reference to the serialization of the involved elements. In section 2.1.3.2 about the predicate complex, I will show that one gets problems similar to those discussed in this section if one does not assume verb movement.

[4] I adapted Netter’s trace so that the order of elements on the SUBCAT list corresponds to the order that is assumed by Pollard and Sag (1994) and throughout this paper.

[5] Note, that the problem discussed below is not a general problem of argument attraction approaches. The problem arises only when underspecified valance lists are instantiated by saturation of arguments. If the element from which arguments are attracted is combined with its head before any other argument is combined with this head, all valance lists are instantiated.

[6] See Steinitz (1969) for a different position. She admits at most one (possibly complex) adverbial of a certain type in a clause. Steinitz allows several adverbials of the same kind to form a constituent. So for instance in (i) the PPs *in der Schönhauser Allee* and *unter der Laterne* form a constituent specifying the location of the event expressed by the verb (pp. 126–131).

(i) Das Auto hielt in der Schönhauser Allee unter einer Laterne.
the car stopped in the Schönhauser Allee under a streetlamp

Note however that it is possible to have adverbials of the same type in a sentence that refer to different verbs:

(ii) weil er gestern morgen sterben wollte.
because he yesterday tomorrow die wanted
‘because he wanted yesterday to die tomorrow.’

If we do not restrict the number of verbs per verbal complex (see the discussion of (20) in section 2.1.3), it follows the the number of adverbials per sentence is not restricted either.

[7] Jacobs (1986, p. 120) seems to have a similar analysis in mind (see also Egli and Egli-Gerer, 1991, p. 107).

[8] Uszkoreit actually assumes more than six lexical entries for *geben* since his lexical rule instantiates all features that are relevant for linearization. All these different permutations and instantiations give rise to 18 different lexical items for a ditransitive verb.

[9] This is basically equivalent to the problem of linearizing traces in a system that uses a phrase structure based backbone to process a grammar with flat structures.

[10] The possibility to scramble the arguments of the embedded verb with the arguments of the matrix verb has been denied for various classes of embedded verbs. However, that basically all permutations are possible becomes clear if one takes the data that was discussed by Bech (1955, p. 136), Bierwisch (1963, p. 25), Jacobs (1991, p. 20), and Haider (1991, p. 5). The permutation of elements with the same morphological case is restricted by performance factors in both simplex and complex clauses. See Kuno (1980, p. 175) for Japanese and Müller (1999, pp. 172–173) for German. A detailed discussion of predicate complex constructions and more references can be found in Müller (2002).

[11] In Hinrichs and Nakazawa (1998) they assume the rules in (i).

(i) a. H[SC <>] → C*, H_{word}
 b. H → H[VFORM ¬ fin], VC

[12] Neues Deutschland, 06.12.1969, p. 1.

[13] Extraction to the left is usually called topicalization, but for German the term fronting is more appropriate since expletives and fixed parts of idioms can be fronted and they are not topics. On the distributional properties of idioms see Müller (2002).

[14] Spiegel, 23/1997, p. 122.

[15] Note that some recent versions of HPSG treat adjuncts as dependents. Adjuncts are introduced into the SUBCAT list of their head by a lexical rule (van Noord and Bouma, 1994) or a relational constraint (Bouma, Malouf and Sag, 2001). With such an approach the problem of non-predictable extraposed elements is shifted to another place and lazy evaluation techniques are needed to process the grammar. Lazy evaluation means that the execution in one branch of the program is delayed until enough information is available. Not all systems that are used for processing HPSG-like grammars have built-in machinery that allows for lazy evaluation.

[16] A reviewer pointed out to me that ALE provides the possibility to represent the rules in (i) by a single rule of the form in (ii).

(i) a. H → H[SC <>]
 b. H → A, H[SC <A>]
 c. H → A, B, H[SC <A,B>]
 d. H → A, B, C, H[SC <A,B,C>]
 e. ...

(ii) H → X, H[SC X]

At the moment such a schematic rule is applied in parsing, the number of daughters has to be known though. This essentially makes it necessary that the head is found first to determine the length of the SUBCAT list and thereby the number of daughters involved. Therefore in ALE rules like (33a,c,d) cannot be abbreviated in this way. It is also not possible to collapse rules like (33e) and (33g) since here the arguments are explicitly referred to in order to permute them.

Note that in systems which do not provide for relational constraints the number of daughters has to be fixed anyway, since information has to be percolated from the daughters to the mother (for instance nonlocal features and semantic indices).

[17] See also note 15.

[18] Geißler (1994) suggests a special lexical rule that licenses two output elements: a lexical item for the initial verb and an instantiated verbal trace for head-movement. Johnson and Kay (1992) suggest lexical entries that consist of two elements: a phonologically filled element and a corresponding empty element. These approaches are problematic for sentences with high lexical ambiguity. Suppose we have four lexical entries for a verb like *wissen* ('know') (intransitive, transitive, with clausal complement, with interrogative clause as complement). In an account that treats permutation in the *Mittelfeld* via lexical rules, we have to license lexical items with the appropriate permutations of valence lists. For all these lexical items we have to license empty elements for verb movement. In the case of *wissen* we get seven empty elements. For longer sentences this will result in a combinatorial explosion. Note also that there may be more than one finite verb in a sentence. If we have several verbs with identical valence requirements they license several identical traces. So we need additional machinery that ensures that the trace is used with the head that licensed it. Furthermore, verbs in first/third person plural have a form that is identical to their infinitive form. Therefore when we parse an example like (i) we get empty elements for *will* ('want'), *wissen* ('know'), and *kommt* ('come').

(i) Ich will wissen, wer kommt.
I want know who comes
'I want to know who is coming.'

Whether the approaches suggested by Kay, Johnson, and Geißler lead to a performance improvement is therefore an open issue. It depends on the lexical ambiguity and on other parts of the grammar. The empirical tests that are needed to determine the usefulness for the Verb*mobil* grammar could not be made since neither LKB nor PET allows for phonologically empty elements. In these systems grammars with empty elements have to be rewritten in a way where the empty elements are replaced by unary rules. In such systems, the equivalent to Geißler's, Kay's, and Johnson's dynamically licensed empty elements is a dynamically licensed grammar rule, i.e., a grammar rule that can only be used if certain lexical entries are present in a string. No such thing exists.

[19] Note that the head does not contain itself in its domain list. A representation as in (i) would yield unintuitive representations since these domains can be understood as lists of daughters of a head and this would mean that a head has itself as a daughter.

(i) $\boxed{1}\begin{bmatrix}\text{DOM} \ \left\langle \boxed{1} \right\rangle \\ \textit{lexical-sign}\end{bmatrix}$

[20] Verb*mobil* Corpus, CD 14.
[21] Verb*mobil* Corpus, CD 04.
[22] Verb*mobil* Corpus, CD 20.
[23] Verb*mobil* Corpus, CD 04.
[24] Verb*mobil* Corpus, CD 20.
[25] Verb*mobil* Corpus, CD 01.
[26] Note that the curve for the Verb*mobil* grammar differs dramatically from what was given in Müller (2000). The reason is that I switched to another system for processing continuous grammars. As was mentioned at the beginning of this section, I now use the PET system which is a reimplementation of LKB in C. PET needs less memory and is much faster. This gave me the opportunity to extend the grammar in an adequate way: I now treat adjunct extraction as a non-local dependency. This is a very costly analysis and therefore (computational) linguists working on implementations usually stipulate special structures for fronted adjuncts (Kolb and Thiersch, 1991; van Noord, Bouma, Koeling and Nederhof, 1999). Note that adjunct extraction is not just needed for sentences with an adjunct in the *Vorfeld*, but also for interrogative and relative clauses that have an adjunct as the interrogative or relative phrase, respectively. Apart from changing the adjunct analysis, I introduced an analysis for matching free relatives, and I completed the analysis of complex fronting. All these phenomena were not implemented in the grammar reported about in Müller (2000). In order to parse a reasonable quantity of the test utterances, the edge limit has been raised to 100.000.
[27] The figure was generated by TSDB++ (Oepen and Flickinger, 1998).

References

Abney, S.P., Hinrichs E.W. (eds.) (1995) *Proceedings of the Seventh Conference of the European Chapter of the Association for Computational Linguistics*. Dublin.

Bech G. (1955) *Studien über das deutsche Verbum infinitum*. Historisk-filologiske Meddekekser udgivet af Det Kongelige Danske Videnskabernes Selskab. Bind 35, no. 2, 1955; Bind 36, no. 6, 1957, Kopenhagen, reprinted 1983, Tübingen: Max Niemeyer Verlag.

Behaghel O. (1930) Von deutscher Wortstellung. *Zeitschrift für Deutschkunde*, pp. 81–89.

Bierwisch M. (1963) *Grammatik des deutschen Verbs*, studia grammatica II. Berlin: Akademie Verlag.

Bouma, G. (1996) Extraposition as a Nonlocal Dependency. In Kruijff G-J., Morrill G., Oehrle D. (eds.), *Proceedings of Formal Grammar 96*, pp. 1–14, Prag. http://www.let.rug.nl/˜gosse/papers/. 07.24.97.

Bouma G., Malouf R. and Sag I.A. (2001) Satisfying Constraints on Extraction and Adjunction. *Natural Language and Linguistic Theory*, 1/19, pp. 1–65, ftp://csli-ftp.stanford.edu/linguistics/sag/bms-nllt.ps. 09.18.2002.

Bunt H., van Horck A (eds.) (1996) *Discontinuous Constituency*, Natural Language Processing, No. 6. Mouton de Gruyter, Berlin/New York.

Callmeier U. (2000) PET – A Platform for Experimentation with Efficient HPSG Processing Techniques. *Journal of Natural Language Engineering* 1/6, pp. 99–108 (Special Issue on Efficient Processing with HPSG: Methods, Systems, Evaluation).

Campbell-Kibler K. (2001) Bech's Problem, Again: The Dutch Word *er*. In *HPSG-2001, Abstracts*, NTNU, Trondheim, Department of Linguistics, Trondheim.

Carpenter B., Penn G. (1996) Efficient Parsing of Compiled Typed Attribute Value Logic Grammars. In Bunt H., Tomita M. (eds.), *Recent Advances in Parsing Technology*, Text, Speech and Language Technology, No. 1, Kluwer Academic Publishers. Dordrecht/Boston/London.

Carroll J. (1994) Relating Complexity to Practical Performance in Parsing with Wide-coverage Unification Grammars. In Pustejovsky J. (ed.), *Proceedings of the Thirty-Second Annual Meeting of the ACL*, Association for Computational Linguistics, Morristown, NJ, pp. 287–294. e-Print-Archive: http://xxx.lanl.gov/abs/cmp-lg/9405033. 04.07.99.

Chung, C. (1993) Korean Auxiliary Verb Constructions Without VP Nodes. In Kuno S., Lee I-H., Whitman J., Maling J., Kang Y-S., Kim Y.J. (eds.), *Proceedings of the 1993 Workshop on Korean Linguistics*, Harvard Studies in Korean Linguistics, No. 5, Harvard University, Department of Linguistics, Cambridge, Massachusetts, pp. 274–286.

Clocksin W.F., Mellish C.S. (1984) *Programming in Prolog*. Springer Verlag, Berlin, Heidelberg.

Copestake A. (1999) The (New) LKB System. ftp://www-csli.stanford.edu/˜aac/newdoc.pdf. 06.24.99.

Daniels M., Meurers W.D. (2002) Improving the Efficiency of Parsing with Discontinuous Constituents. In Wintner S. (ed.), *Proceedings of NLULP'02: The 7th International Workshop on Natural Language Understanding and Logic Programming*, Datalogiske Skrifter, No. 92, Roskilde Universitetscenter, Copenhagen, pp. 49–68.

Donohue C., Sag I.A. (1999) Domains in Warlpiri. In *Sixth International Conference on HPSG – Abstracts. 04–06 August*, Edinburgh, pp. 101–106. http://www-csli.stanford.edu/˜sag/papers/warlpiri.ps. 08.18.2002.

Drach E. (1937) *Grundgedanken der deutschen Satzlehre*. Wissenschaftliche Buchgesellschaft, Darmstadt.

Egli U., Egli-Gerber R. (1991) Sprachsysteme – logische und historische Grundlagen der erweiterten Phrasenstrukturgrammatik. Arbeitspapier 28, Universität Konstanz, Fachgruppe Sprachwissenschaft, http://ling.uni-konstanz.de/pages/publ/abstracts/ap028.html. 02.02.2002.

Egli U., von Heusinger K. (1992) Das PATR-II-Format in Anwendung auf das Deutsche. Arbeitspapier 42, Universität Konstanz, Fachgruppe Sprachwissenschaft.

Engelkamp J., Erbach G., Uszkoreit H. (1992) Handling Linear Precedence Constraints by Unification. In Thomson H.S. (ed.), *30th Annual Meeting of the Association for Computational*

Linguistics. Proceedings of the Conference, Association for Computational Linguistics, erschien auch als CLAUS-Re-port, Nr. 19, Universität des Saarlandes, Newark, Delaware, pp. 201–208.

Erbach G. (1995) ProFIT: Prolog with Features, Inheritance and Templates. In Abney and Hinrichs (eds.).

Erbach G. (1998) *Bottom-Up Earley Deduction for Preference-Driven Natural Language Processing*. Saarbrücken Dissertations in Computational Linguistics and Language Technology, No. 4. Deutsches Forschungszentrum für Künstliche Intelligenz und Universität des Saarlandes, Saarbrücken.

Erdmann O. (1886) *Grundzüge der deutschen Syntax nach ihrer geschichtlichen Entwicklung*, Vol. 1. Verlag der J. G. Cotta'schen Buchhandlung, Stuttgart. Reprint Georg Olms Verlag, Hildesheim, 1985.

Fouvry F, Meurers W.D. (2000). Towards a Platform for Linearization Grammars. In Hinrichs et al. (eds.), pp. 153–168.

Frank A. (1994) Verb Second by Lexical Rule or by Underspecification. Arbeitspapiere des SFB 340 No. 43, IBM Deutschland GmbH, Heidelberg, ftp://ftp.ims.uni-stuttgart.de/pub/papers/anette/v2-usp.ps.gz. 08.20.2002.

Geißler S. (1994) *Lexikalische Regeln in der IBM-Basisgrammatik*. Verbmobil Report 20, IBM Informationssysteme GmbH, Institut für Logik und Linguistik, Heidelberg.

Görz G. (ed.) (1992) *Konvens 92.1. "Konferenz Verarbeitung natürlicher Sprache"*. Nürnberg 7–9, Oktober 1992. Informatik aktuell, Berlin, Heidelberg, New York, Springer-Verlag.

Grewendorf G. (1988) *Aspekte der deutschen Syntax. Eine Rektions-Bindungs-Analyse*. Studien zur deutschen Grammatik, No. 33, Gunter Narr Verlag, Tübingen.

Gunji T. (1986) Subcategorization and Word Order. In Poser W.J. (ed.), *Papers from the Second International Workshop on Japanese Syntax*, CSLI Publications, Stanford, pp. 1–21.

Haider H. (1991) *Fakultativ kohärente Infinitivkonstruktionen im Deutschen*. Arbeitspapiere des SFB 340 No. 17, IBM Deutschland GmbH, Heidelberg.

Haider H. (1996) Downright Down to the Right. In *On Extraction and Extraposition in German*, Linguistik Aktuell/Linguistics Today, No. 11, Benjamins, Amsterdam, pp. 245–271.

Hinrichs E.W., Kathol A., Nakazawa T. (eds.) (1998) *Complex Predicates in Nonderivational Syntax*, Vol. 30 of *Syntax and Semantics*, Academic Press, San Diego.

Hinrichs E.W., Meurers, W.D., Wintner S. (eds.) (2000) *Proceedings of the ESSLLI-2000 Workshop on Linguistic Theory and Grammar Implementation*, Birmingham, UK, August 14–18.

Hinrichs E.W., Nakazawa T. (1989a) Flipped Out: AUX in German. In *Aspects of German VP Structure*, SfS-Report-01-93, Eberhard-Karls Universität, Tübingen.

Hinrichs E.W., Nakazawa T. (1989b) Subcategorization and VP Structure in German. In *Aspects of German VP Structure*, SfS-Report-01-93, Eberhard-Karls Universität, Tübingen.

Hinrichs E.W., Nakazawa T. (1994) Partial-VP and Split-NP Topicalization in German – An HPSG Analysis. In Hinrichs E.W., Meurers W.D., Nakazawa T. (eds.), *Partial-VP and Split-NP Topicalization in German – An HPSG Analysis and its Implementation*, Arbeitspapiere des SFB 340, No. No. 58, Eberhard-Karls Universität, Tübingen.

Hinrichs E.W., Nakazawa T. (1998) Third Construction and VP Extraposition in German: An HPSG Analysis. In Hinrichs et al. (eds.), pp. 115–157.

Hoberg U. (1981) *Die Wortstellung in der geschriebenen deutschen Gegenwartssprache*. Heutiges Deutsch. Linguistische Grundlagen. Forschungen des Instituts für deutsche Sprache, No. 10, Max Huber Verlag, München.

Höhle T.N. (1982) Explikation für "normale Betonung" und "normale Wortstellung". In Abraham W. (ed.), *Satzglieder im Deutschen – Vorschläge zur syntaktischen, semantischen und pragmatischen Fundierung*, Studien zur deutschen Grammatik, No. 15, Gunter Narr Verlag, Tübingen, pp. 75–153.

Jacobs J. (1986) The Syntax of Focus and Adverbials in German. In Abraham W., de Meij S. (eds.), *Topic, Focus, and Configurationality. Papers from the 6th Groningen Grammar Talks,* *Groningen, 1984*, Linguistik Aktuell/Linguistics Today, No. 4, John Benjamins Publishing Company, Amsterdam/Philadelphia, pp. 103–127.

Jacobs J. (1991) Bewegung als Valenztransfer. SFB 282: Theorie des Lexikons 1, Heinrich Heine Uni/BUGH, Düsseldorf/Wuppertal.

Jacobson P. (1987) Phrase Structure, Grammatical Relations, and Discontinuous Constituents. In Huck G.J., Ojeda A.E. (eds.), *Discontinuous Constituency*, Vol. 20 of *Syntax and Semantics*, Academic Press, New York, pp. 27–69.

Johnson M. (1985) Parsing with Discontinuos Constituents. In Mann W.C. (ed.), *Proceedings of the Twenty-Third Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, Chicago, IL, pp. 127–1321.

Johnson M. (1986) A GPSG Account of VP Structure in German. *Linguistics*, 24/5, pp. 871–882.

Johnson M., Kay M. (1992) *Parsing and Empty Nodes*. Arbeitspapiere des SFB 340 No. 25, IBM Deutschland GmbH, Heidelberg.

Kasper R.T. (1994) Adjuncts in the Mittelfeld. In Nerbonne et al. (eds.), pp, 39–70.

Kasper R.T., Calcagno M., Davis P.C. (1998) Know When to Hold 'Em: Shuffling Deterministically in a Parser for Nonconcatenative Grammars. In *Proceedings of the COLING-ACL 98*, Montreal, pp. 663–669.

Kathol A. (1995) *Linearization-Based German Syntax*. PhD thesis, Ohio State University.

Kathol A. (1998) Constituency and Linearization of Verbal Complexes. In Hinrichs et al. (eds.), pp., 221–270.

Kathol A. (2000) *Linear Syntax*. Oxford University Press, New York/Oxford.

Kathol A., Pollard C.J. (1995) Extraposition via Complex Domain Formation. In *Proceedings of the Thrirty-Third Annual Meeting of the ACL*, Association for Computational Linguistics, Boston, http://linguistics.berkeley.edu/˜kathol/Papers/ACL95.ps.gz. 06.29.99.

Kay M. (1980) Algorithm Schemata and Data Structures in Syntactic Processing. Technical Report CSL-80-12, Xerox Corporation, Palo Alto Research Center.

Keenan E.L., Comrie, B. (1977) Noun Phrase Accessibility and Universal Grammar. *Linguistic Inquiry*, 8, pp. 63–99.

Keller F. (1994) Extraposition in HPSG. Verbmobil-Report 30, IBM Deutschland Informationssysteme GmbH, Institut für Logik und Linguistik, Heidelberg ftp://ftp.ims.uni-stuttgart.de/pub/papers/keller/verbmobil.ps.gz. 04.15.97.

Keller F. (1995) Towards an Account of Extraposition in HPSG. In Abney and Hinrichs (eds.). ftp://ftp.ims.uni-stuttgart.de/pub/papers/keller/eacl95.ps.gz. 04.15.97.

Kiss T. (1995) *Infinite Komplementation. Neue Studien zum deutschen Verbum infinitum*. Linguistische Arbeiten, No. 333. Max Niemeyer Verlag, Tübingen.

Kiss T. (2000) Semantic Constraints on Relative Clause Extraposition, Bochum, ms.

Kiss T. (2001) Configurational and Relational Scope Determination in German. In Meurers and Kiss (eds.), pp. 131–166.

Kiss T., Wesche B. (1991) Verb Order and Head Movement. In Herzog O., Rollinger C-R. (eds.), *Text Understanding in LILOG*, Lecture Notes in Artificial Intelligence, No. 546, Springer-Verlag, Berlin Heidelberg New York, pp. 216–242.

Kolb H-P., Thiersch C.L. (1991) Levels and Empty Categories in a Principles and Parameters Based Approach to Parsing. In Haider H., Netter K. (eds.), *Representation and Derivation in the Theory of Grammar*, Studies in Natural Language and Linguistic Theory, No. 22, Kluwer Academic Publishers, Dordrecht/Boston/London.

Kuno S. (1980) A Further Note on Tonoike's Intra-Subjectivization Hypothesis. In Otsu Y., Farmer A. (eds.), *Theoretical Issues in Japanese Linguistics*, Vol. 2 of *MIT Working Papers in Linguistics*, MIT, Cambridge, pp. 171–184.

Kunze J. (1991) *Kasusrelationen und semantische Emphase*, studia grammatica XXXII. Akademie Verlag, Berlin.

Meurers W.D. (1994) On Implementing an HPSG Theory. In Hinrichs E.W., Meurers W.D., Nakazawa T. (eds.), *Partial-VP and Split-NP Topicalization in German – An HPSG Analysis and Its Implementation*, Arbeitspapiere des SFB 340, No. 58, Eberhard-Karls Universität, Tübingen. http://ling.osu.edu/~dm/on-implementing.html. 06.12.96.

Meurers W.D. (2000) Lexical Generalizations in the Syntax of German Non-Finite Constructions. Arbeitspapiere des SFB 340 No. 145, Eberhard-Karls Universität, Tübingen, http://www.ling.ohio-state.edu/~dm/papers/diss.html. 08.19.2002.

Meurers W.D., Kiss T. (eds.) (2001) *Constraint-Based Approaches to Germanic Syntax*. Studies in Constraint-Based Lexicalism, No. 7, CSLI Publications, Stanford. http://cslipublications.stanford.edu/site/1575863049.html. 08.18.2002.

Müller S. (1995) Scrambling in German – Extraction into the *Mittelfeld*. In T'sou B.K., Bong Yeung Lai T. (eds.), *Proceedings of the Tenth Pacific Asia Conference on Language, Information and Computation*, City University of Hong Kong, pp. 79–83. http://www.dfki.de/~stefan/Pub/scrambling.html. 09.09.2002.

Müller S. (1996) The Babel-System – An HPSG Prolog Implementation. In *Proceedings of the Fourth International Conference on the Practical Application of Prolog*, London, pp. 263–277. http://www.dfki.de/~stefan/Pub/babel.html. 09.09.2002.

Müller S. (1997a). Scrambling in German – Extraction into the *Mittelfeld*. Research Report RR-97-06, Deutsches Forschungszentrum für Künstliche Intelligenz, Saarbrücken, a shorter version appeared in *Proceedings of the Tenth Pacific Asia Conference on Language, Information and Computation*. City University of Hong Kong, 1995. http://www.dfki.de/~stefan/Pub/scrambling.html. 09.09.2002.

Müller S. (1997b) Yet another Paper about Partial Verb Phrase Fronting in German. Research Report RR-97-07, Deutsches Forschungszentrum für Künstliche Intelligenz, Saarbrücken, a shorter version appeared in *Proceedings of COLING 96*, pp. 800–805. http://www.dfki.de/~stefan/Pub/pvp.html. 09.09.2002.

Müller S. (1999) *Deutsche Syntax deklarativ. Head-Driven Phrase Structure Grammar für das Deutsche*. Linguistische Arbeiten, No. 394. Max Niemeyer Verlag, Tübingen. http://www.dfki.de/~stefan/Pub/hpsg.html. 09.09.2002.

Müller S. (2000) Continuous or Discontinuous Constituents? In Hinrichs et al. (eds.), pp. 133–152, http://www.dfki.de/~stefan/Pub/discont.html. 09.09.2002.

Müller S. (2001) An HPSG Analysis of German Depictive Secondary Predicates. In *Proceedings of Formal Grammar 2001/MOL 7*, Electronic Notes in Theoretical Computer Science, No. 53. Helsinki, http://www.dfki.de/ stefan/Pub/depiktive.html. 09.09.2002.

Müller S. (2002) *Complex Predicates: Verbal Complexes, Resultative Constructions, and Particle Verbs in German*. Studies in Constraint-Based Lexicalism, No. 13. CSLI Publications, Stanford. http://www.dfki.de/~stefan/Pub/complex.html. 09.09.2002.

Müller S., Kasper W. (2000) HPSG Analysis of German. In Wahlster (ed.), pp. 238–253.

Nerbonne J., Netter K., Pollard, C.J. (eds.) (1994) *German in Head-Driven Phrase Structure Grammar*. CSLI Lecture Notes, No. 46. CSLI Publications, Stanford.

Netter K. (1992) On Non-Head Non-Movement. An HPSG Treatment of Finite Verb Position in German. In Görz (ed.), pp. 218–227.

Netter K. (1998) *Functional Categories in an HPSG for German*. Saarbrücken Dissertations in Computational Linguistics and Language Technology, No. 3. Deutsches Forschungszentrum für Künstliche Intelligenz und Universität des Saarlandes, Saarbrücken.

Oepen S., Carroll J.A. (2000) Ambiguity Packing in Constraint-Based Parsing – Practical Results. In *Proceedings of the 1st Conference of the North American Chapter of the Association for Computational Linguistics (NAACL'00)*, Seattle, WA, pp. 162–169.

Oepen S., Flickinger D.P. (1998) Towards Systematic Grammar Profiling. Test Suite Technology Ten Years After. *Journal of Computer Speech and Language*, 12/4, pp. 411–436 (Special Issue on Evaluation).

Olsen S. (1982) On the Syntactic Description of German: Topological Fields vs. X′-Theory. In Welte W. (ed.), *Sprachtheorie und angewandte Linguistik. Festschrift für Alfred Wollmann*, Tübinger Beiträge zur Linguistik, No. 195, Gunter Narr Verlag, Tübingen, pp. 29–45.

Paritong M. (1992) Constituent Coordination in HPSG. In Görz (ed.), pp. 228–237.

Paul H. (1919) *Deutsche Grammatik. Teil IV: Syntax*, Vol. 3. Max Niemeyer Verlag, Halle an der Saale. 2nd unchanged edition (1968) Max Niemeyer Verlag, Tübingen.

Penn G. (1999) Linearization and *WH*-Extraction: Evidence from Serbo-Croatian. In Borsley R.D., Przepiórkowski A. (eds.), *Slavic in Head-Driven Phrase Structure Grammar*, Studies in Constraint-Based Lexicalism, No. 2, CSLI Publications, Stanford, pp. 149–182.

Penn G., Carpenter B. (1999) ALE for Speech: a Translation Prototype. In *Proceedings of the 6th Conference on Speech Communication and Technology (EUROSPEECH)*, Budapest, Hungary.

Pinkal M., Rupp C.J. and Worm, K.L. (2000) Robust Semantic Processing of Spoken Language. In Wahlster (ed.), pp. 322–336.

Pollard C.J. (1996) On Head Non-Movement. In Bunt and on Horck (eds.), pp. 279–305. Published version of a Ms. dated January 1990.

Pollard C.J., Kasper R.T., Levine R.D. (1992) *Linearization Grammar*. Research Proposal to the National Science Foundation.

Pollard C.J., Kasper R.T., Levine R.D. (1994) *Studies in Constituent Ordering: Toward a Theory of Linearization in Head-Driven Phrase Structure Grammar*, research Proposal to the National Science Foundation. ftp://csli-ftp.stanford.edu/linguistics/sag/linearization-prop.ps.gz. 08.20.2002.

Pollard C.J., Sag I.A. (1987) *Information-Based Syntax and Semantics*. CSLI Lecture Notes, No. 13. CSLI Publications, Stanford.

Pollard C.J., Sag I.A. (1994) *Head-Driven Phrase Structure Grammar*. Studies in Contemporary Linguistics, Chicago. University of Chicago Press, London.

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

Reape M. (1992) *A Formal Theory of Word Order: A Case Study in West Germanic*. PhD thesis, University of Edinburgh.

Reape M. (1994) Domain Union and Word Order Variation in German. In Nerbonne et al. (eds.), pp. 151–198.

Reape M. (1996) Getting Things in Order. In Bunt and van Horck (eds.), pp. 209–253. Published version of a Ms. dated January 1990.

Rentier G. (1994) Dutch Cross Serial Dependencies in HPSG. In *Proceedings of COLING 94*, Kyoto, Japan, pp. 818–822. http://xxx.lanl.gov/abs/cmp-lg/9410016. 08.18.2002.

Richter F., Sailer M. (2001) On the Left Periphery of German. In Meurers and Kiss (eds.).

Rohrer C. (1996) Fakultativ kohärente Infinitkonstruktionen im Deutschen und deren Behandlung in der Lexikalisch Funktionalen Grammatik. In Harras G., Bierwisch M. (eds.), *Wenn die Semantik arbeitet. Klaus Baumgärtner zum 65. Geburtstag*, Max Niemeyer Verlag, Tübingen, pp. 89–108.

Schmidt P., Rieder S., Theofilidis A. (1996) Final Documentation of the German LS-GRAM Lingware. Deliverable DC-WP6e (German), IAI, Saarbrücken.

Schütz, J. (1996) The ALEP Formalism in a Nutshell. http://www.iai.uni-sb.de/alep/alep-docs.html. 06.09.99.

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

Staff, COLING (ed.) (1994) *Proceedings of COLING 94*. ACL – Association for Computational Linguistics, Kyoto, Japan.

Steinitz R. (1969) *Adverbial-Syntax*, studia grammatica X. Akademie Verlag, unter Mitarbeit von Ewald Lang, Berlin.

Uszkoreit H. (1986) *Linear Precedence in Discontinuous Constituents: Complex Fronting in German*. Report No. CSLI-86-47, Center for the Study of Language and Information, Stanford.

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

Uszkoreit H., Backofen R., Busemann S., Diagne A.K., Hinkelman E.A., Kasper W., Kiefer B., Krieger H-U., Netter K., Neumann G., Oepen S., Spackman, S.P. (1994) DISCO – An HPSG-Based NLP System and its Application for Appointment Scheduling. In Staff (ed.), pp. 436–440.

van Noord G. (1993) *Reversibility in Natural Language Processing*. PhD thesis, Rijksuniversiteit te Utrecht.

van Noord G. (1997) An Efficient Implementation of the Head-Corner Parser. *Computational Linguistics*, 23/3.

van Noord G., Bouma G. (1994) The Scope of Adjuncts and the Processing of Lexical Rules. In Staff (ed.), pp. 250–256. http://grid.let.rug.nl/˜vannoord/papers/coling94.ps.gz. 08.18.2002.

van Noord G., Bouma G., Koeling R., Nederhof M-J. (1999) Robust Grammatical Analysis for Spoken Dialogue Systems. *Natural Language Engineering*, 1/5, pp. 45–93. e-Print-Archive: http://xxx.lanl.gov/abs/cmp-lg/9906026. 07.02.99.

Wahlster W. (ed.) (2000) *Verbmobil: Foundations of Speech-to-Speech Translation*. Artificial Intelligence. Springer-Verlag, New York/Heidelberg, Berlin.

Wegener H. (1985) *Der Dativ im heutigen Deutsch*. Studien zur deutschen Grammatik, No. 28. Gunter Narr Verlag, Tübingen.