



Deutsches
Forschungszentrum
für Künstliche
Intelligenz GmbH

Research
Report
RR-97-06

Scrambling in German – Extraction into the *Mittelfeld*

Stefan Müller

August, 1995

**Deutsches Forschungszentrum für Künstliche Intelligenz
GmbH**

Postfach 20 80
67608 Kaiserslautern, FRG
Tel.: + 49 (631) 205-3211
Fax: + 49 (631) 205-3210
E-Mail: info@dfki.uni-kl.de

Stuhlsatzenhausweg 3
66123 Saarbrücken, FRG
Tel.: + 49 (681) 302-5252
Fax: + 49 (681) 302-5341
E-Mail: info@dfki.de

WWW: <http://www.dfki.de>

Deutsches Forschungszentrum für Künstliche Intelligenz

DFKI GmbH

German Research Center for Artificial Intelligence

Founded in 1988, DFKI today is one of the largest nonprofit contract research institutes in the field of innovative software technology based on Artificial Intelligence (AI) methods. DFKI is focusing on the complete cycle of innovation — from world-class basic research and technology development through leading-edge demonstrators and prototypes to product functions and commercialization.

Based in Kaiserslautern and Saarbrücken, the German Research Center for Artificial Intelligence ranks among the important “Centers of Excellence” worldwide.

An important element of DFKI’s mission is to move innovations as quickly as possible from the lab into the marketplace. Only by maintaining research projects at the forefront of science can DFKI have the strength to meet its technology transfer goals.

DFKI has about 115 full-time employees, including 95 research scientists with advanced degrees. There are also around 120 part-time research assistants.

Revenues for DFKI were about 24 million DM in 1997, half from government contract work and half from commercial clients. The annual increase in contracts from commercial clients was greater than 37% during the last three years.

At DFKI, all work is organized in the form of clearly focused research or development projects with planned deliverables, various milestones, and a duration from several months up to three years.

DFKI benefits from interaction with the faculty of the Universities of Saarbrücken and Kaiserslautern and in turn provides opportunities for research and Ph.D. thesis supervision to students from these universities, which have an outstanding reputation in Computer Science.

The key directors of DFKI are Prof. Wolfgang Wahlster (CEO) and Dr. Walter Olthoff (CFO).

DFKI’s six research departments are directed by internationally recognized research scientists:

- ❑ Information Management and Document Analysis (Director: Prof. A. Dengel)
- ❑ Intelligent Visualization and Simulation Systems (Director: Prof. H. Hagen)
- ❑ Deduction and Multiagent Systems (Director: Prof. J. Siekmann)
- ❑ Programming Systems (Director: Prof. G. Smolka)
- ❑ Language Technology (Director: Prof. H. Uszkoreit)
- ❑ Intelligent User Interfaces (Director: Prof. W. Wahlster)

In this series, DFKI publishes research reports, technical memos, documents (eg. workshop proceedings), and final project reports. The aim is to make new results, ideas, and software available as quickly as possible.

Prof. Wolfgang Wahlster
Director

Scrambling in German – Extraction into the *Mittelfeld*

Stefan Müller

DFKI-RR-97-06

A short version of this paper appeared in Proceedings of the tenth Pacific Asia Conference on Language, Information and Computation. Hong Kong. 1995

© Deutsches Forschungszentrum für Künstliche Intelligenz 1995

This work may not be copied or reproduced in whole or part for any commercial purpose. Permission to copy in whole or part without payment of fee is granted for nonprofit educational and research purposes provided that all such whole or partial copies include the following: a notice that such copying is by permission of the Deutsche Forschungszentrum für Künstliche Intelligenz, Kaiserslautern, Federal Republic of Germany; an acknowledgement of the authors and individual contributors to the work; all applicable portions of this copyright notice. Copying, reproducing, or republishing for any other purpose shall require a licence with payment of fee to Deutsches Forschungszentrum für Künstliche Intelligenz.

ISSN 0946-008X

Scrambling in German – Extraction into the *Mittelfeld*

Stefan Müller*

Stefan.Mueller@dfki.de
<http://www.dfki.de/~stefan/>

August, 1995

Abstract

German is a language with a relatively free constituent order. During the last few years considerable efforts have been made in all syntactic frameworks to explain so-called scrambling phenomena.

The following paper deals with some tough cases of German constituent order which cannot be described by assuming flat sentence structures or word order domains. The phenomena discussed are stranded prepositions and PP complements of nouns and adjectives which can appear separated from their heads in the German *Mittelfeld*. The similarity to fronting of these elements is used to explain these phenomena by a generalized version of the head-filler schema used in the standard HPSG framework.

*Thanks to Frank Keller and Andreas Kathol for comments on earlier versions of this paper. Thanks to Uta Waller for proof-reading.

1 Introduction

There are two basic ideas how to describe scrambling in languages with relatively free constituent order in certain syntactic domains. Firstly, one can assume that a kind of movement takes place, i.e., there is a position in a string where something is missing (a trace) and there is a corresponding position at another location in the string where the missing constituent appears. The alternative is to allow constituents to appear in any order in some particular domain. This domain usually is the domain of the head of a phrase. In HPSG (Pollard and Sag, 1987; Pollard and Sag, 1994), order variation is commonly associated with ordering variations among sister constituents in a flat structure.

This concept was extended by Mike Reape (1994) to allow for complex domain formation operations which—in his approach—are driven by a feature called `UNIONED`. In the combination of signs, a functor can specify the `UNIONED` value of its arguments. The functor is either the head in a head-complement structure or the adjunct in a head-adjunct structure.

If one allows adjuncts to domain-union with their heads, the fact that adjuncts can appear at any position between complements in the *Mittelfeld* can be accounted for.

In the following I will give an account that employs both word order domains and the nonlocal-mechanism provided by HPSG. I will not use the `UNIONED` feature suggested by Reape since it can be shown that the clause union phenomena which Reape describes with domain-union can be accounted for with argument attraction along the lines of Hinrichs and Nakazawa (1989).

2 Phenomena

The German main sentence is partitioned into at least four topological fields: *Vorfeld*, *linke Satzklammer* (finite verb or complementizer), *Mittelfeld* and *rechte Satzklammer* or *Verbalkomplex* (verb cluster).

- (1) [Die Frau]_i hat _{-i} dem Mann das Buch gegeben.
 The woman has the man the book given
 ‘The woman has given the book to the man.’

Sentences with a verb in second position are assumed to be derived from sentences with verb-initial position by fronting of one constituent. It is possible to front almost all kinds of constituents. In (2) for instance, the indirect object is fronted.

- (2) [Dem Mann]_i hat die Frau _{-i} das Buch gegeben.
 The man has the woman the book given
 ‘The woman has given the book to the man.’

The fronted element can be extracted from an arbitrarily deep level.

- (3) a. [Von Maria] habe ich [ein Bild _{-i}] ins Photoalbum geklebt.
 of Maria have I a picture into the album sticked
 ‘I have sticked a picture of Maria into the album.’

- b. [Dagegen]_i hat Hans [ein Argument _{-i}] vorgebracht.
 against this has Hans an argument advanced
 ‘Hans has advanced an argument against this.’
- c. [Auf seinen Sohn]_i war Karl gestern sehr [_{-i} stolz].
 of his son was Karl yesterday very proud
 ‘Karl was very proud of his son yesterday.’

Whereas the fronting of one constituent is usually explained by HPSG’s NON-LOCAL mechanism, the free order of the complements of the verb in the *Mittelfeld* is standardly explained by allowing the complements of one head to appear in an arbitrary order. The restrictions on this order are factored out of the dominance rules. The restrictions for complements of verbs are very weak, so in (4) all permutations of the complements of *geben* are allowed.

- (4) a. Deshalb gab der Mann der Frau das Buch.
 Therefore gave the man the woman the book
 ‘Therefore the man gave the woman the book.’
- b. Deshalb gab der Mann das Buch der Frau.
- c. Deshalb gab das Buch der Mann der Frau.
- d. Deshalb gab das Buch der Frau der Mann.
- e. Deshalb gab der Frau der Mann das Buch.
- f. Deshalb gab der Frau das Buch der Mann.

(5b) shows that preposition stranding¹ is possible in certain cases.²

- (5) a. Hans hat ein Argument dagegen vorgebracht.
 Hans has an argument against this advanced
 ‘Hans has advanced an argument against this.’
- b. [Da]_i hat Hans [ein Argument [gegen _{-i}]] vorgebracht.
 this has Hans an argument against advanced
- c. Hans hat darin geschlafen.
 Hans has in there slept
 ‘Hans has slept in there.’
- d. [Da]_i hat Hans [drin _{-i}] geschlafen.
 There has Hans in slept

¹There is an ongoing argument about whether these prepositions are pre- or postpositions. As the phenomenon will be described lexically, this does not matter. From a historical point of view, the term postposition might be more appropriate (Oppenrieder, 1991, p. 170).

The _{-i} is used to indicate from which constituent the R-pronoun is extracted. In fact there is no trace, neither to the right nor to the left of a preposition.

²According to Duden (1984, p. 353) preposition stranding is common in northern variants of German.

However, as the examples in (6) show, preposition stranding in general is ungrammatical. The stranding is restricted to very few cases where so-called R-pronouns³ (*da* (*there*), *wo* (*where*) and *hier* (*here*)) are extracted from the preposition.

- (6) a. * [Diesen Vorschlag]_i hat Hans ein Argument [gegen _{-i}] vorgebracht.
 this proposal has Hans an argument against advanced
 ‘Hans has advanced an argument against this proposal.’
- b. * [Maria]_i redete Karl [mit _{-i}].
 Maria talked Karl with
 ‘Maria, Karl talked to.’

All these frontable elements can appear disconnected from their heads in non-canonical positions in the *Mittelfeld*.

- (7) a. , daß [von Michael Hanson]_i jetzt nur noch [wenige Bilder _{-i}]
 that of Michael Hanson now only few pictures
 veröffentlicht werden.
 published are
 ‘that now only a few pictures of Michael Hanson are published.’
- b. Deshalb hat Hans [dagegen]_i [ein Argument _{-i}] vorgebracht.
 Therefore has Hans against-this an argument advanced
 ‘Therefore Hans has advanced an argument against this.’
- c. Ich weiß, daß Karl [auf seinen Sohn]_i gestern sehr [_{-i} stolz]
 I know that Karl of his son yesterday very proud
 war.
 was
 ‘I know that Karl was very proud of his son yesterday.’
- d. Deshalb hat Karl [da]_i [ein Argument [gegen _{-i}]] vorgebracht.
 Therefore has Karl this an argument against advanced
 ‘Therefore Karl has advanced an argument against this.’

(7) shows that it is possible to scramble constituents that are not complements of the same head. However, this kind of scrambling is clause bound. Sentences like (8c,d) are ungrammatical.

- (8) a. , weil ich glaube, daß Maria den Mann liebt.
 because I believe that Maria the man loves
 ‘because I believe that Maria loves that man.’
- b. , weil ich, daß Maria den Mann liebt, glaube.
- c. * , weil ich [den Mann]_i glaube, [daß Maria _{-i} liebt].

³The term R-pronoun was first used by Riemsdijk (1978) for Dutch pronouns that contain a syntactically relevant /r/.

As Oppenrieder (1991) mentions, in earlier stages of German, the German R-pronouns contained an /r/ as well.

d. *, weil ich [den Mann]_i, [daß Maria _{-i} liebt,] glaube.

The sentences in (7) are not accounted for by the standard HPSG approach. The elements that are scrambled into the *Mittelfeld* are not complements of a verb but rather complements of nouns, prepositions or adjectives. In the following I will argue that all these sentences can be analyzed as instances of leftward movement by means of the NONLOC-mechanism provided by HPSG.

As was noted by Pollard, Kasper, and Levine (1994), the acceptability of extraction of NP complements depends on the predicate of which the NP is a complement. As (9–10) show, this acceptability contrast can be observed for extraction into the *Vorfeld* and scrambling into the *Mittelfeld*.

(9) a. , weil Peter einen Film über England gesehen/zerstört hat.
 because Peter a film about England seen/destroyed has
 ‘because Peter has seen/destroyed a film about England.’

b. , weil Peter über England einen Film gesehen hat.

c. # , weil Peter über England einen Film zerstört hat.

(10) a. Über England hat Peter einen Film gesehen.

b. # Über England hat Peter einen Film zerstört.

3 The Analysis

3.1 Schemata and Domain Formation

Instead of having 2 or 3 schemata for combining heads with their complements like Pollard and Sag (1994), I use only one very general head-complement schema. Schema 1 admits exactly one complement in the COMP-DTRS list, which leads to binary branching structures.

It is clear that it would not be of much use to be able to order the head daughter and the members of the COMP-DTRS list with respect to each other because this would not be sufficient to account for the scrambling of complements with respect to each other (see sentence (4)).

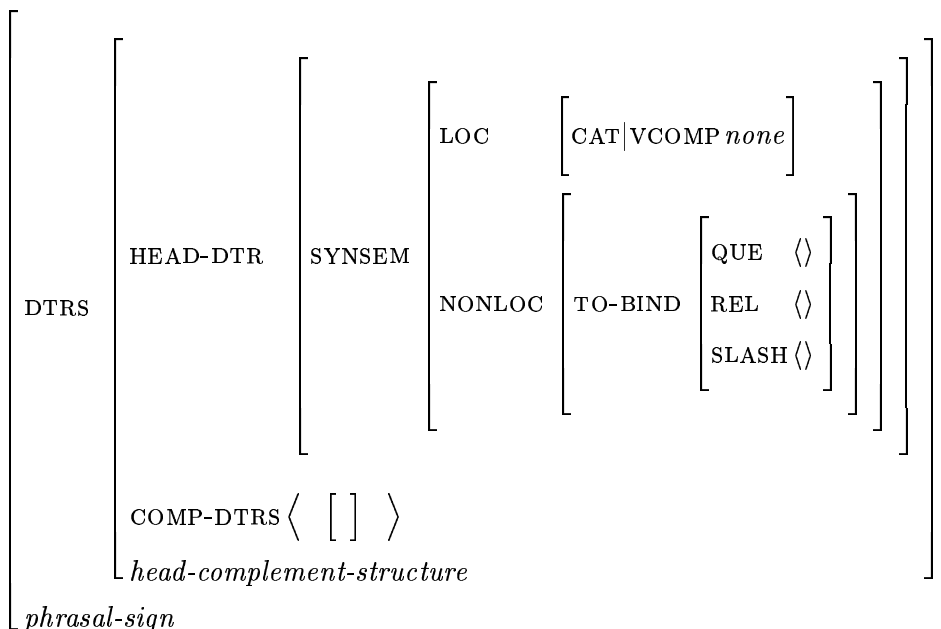
To allow for scrambling, complements are inserted into the domain of their heads by the following implication:

Domain Formation:

$$\left[\begin{array}{l} \text{DTRS} \\ \text{phrasal-sign} \end{array} \left[\text{head-complement-structure} \right] \right] \Rightarrow \left[\begin{array}{l} \text{DTRS} \\ \text{DOM} \end{array} \left[\begin{array}{l} \text{HEAD-DTR} | \text{DOM} \boxed{1} \\ \text{COMP-DTRS} \quad \boxed{2} \\ \boxed{1} \circ \boxed{2} \end{array} \right] \right] \quad (11)$$

A head contains an appropriate description of its syntactic and semantic properties already in its domain.

Schema 1 (Head-Complement Schema)



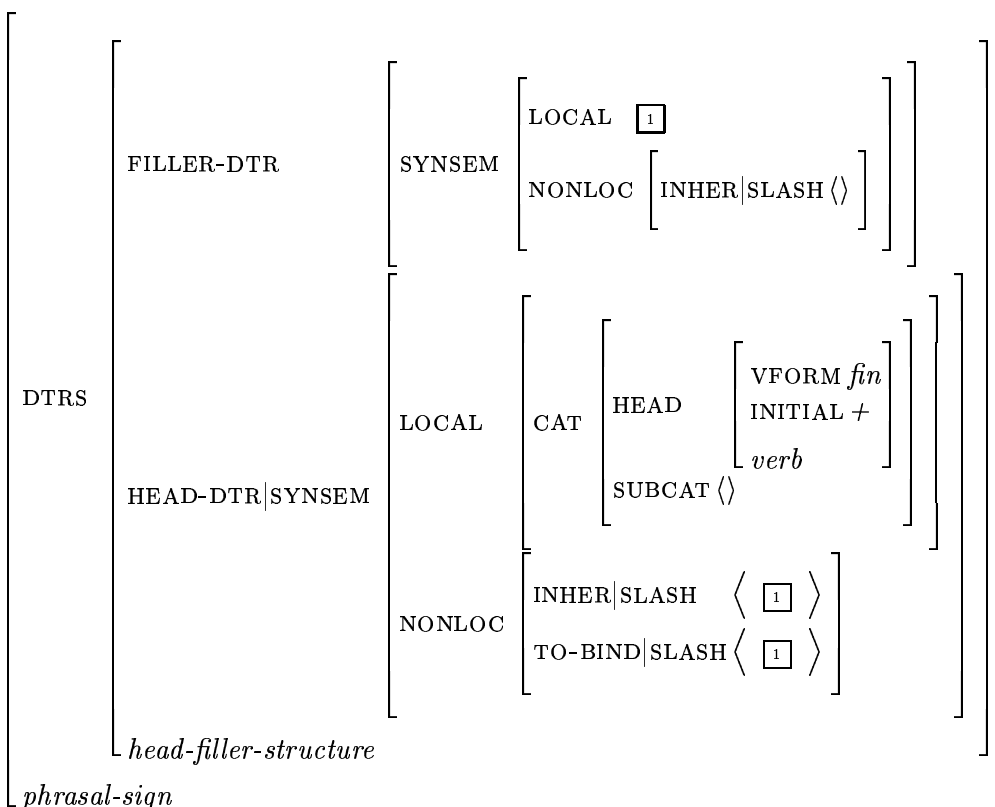
The \circlearrowleft is the *shuffle* relation as used by Reape (1994). The *shuffle* relation holds between three lists A, B, and C, iff C contains all elements of A and B and the order of the elements of A and the order of elements of B is preserved in C. So if a and b are elements of A and a precedes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

$$\left[\textit{phrasal-sign} \right] \Rightarrow \left[\begin{array}{l} \text{PHON}_{\boxed{1}} \oplus \dots \oplus \text{PHON}_{\boxed{n}} \\ \text{DOM } \left\langle \left[\begin{array}{l} \text{PHON}_{\boxed{1}} \\ \textit{sign} \end{array} \right], \dots, \left[\begin{array}{l} \text{PHON}_{\boxed{n}} \\ \textit{sign} \end{array} \right] \right\rangle \end{array} \right] \quad (12)$$

In (12), \oplus corresponds to the *append* relation.

The positioning of one constituent in the *Vorfeld* is accounted for by schema 2.

Schema 2 (Head-Filler Schema)



Like complements, fillers are inserted into the domain of their heads. It is thus possible to account for linearization phenomena in sentences with multiple extraposed constituents.⁴

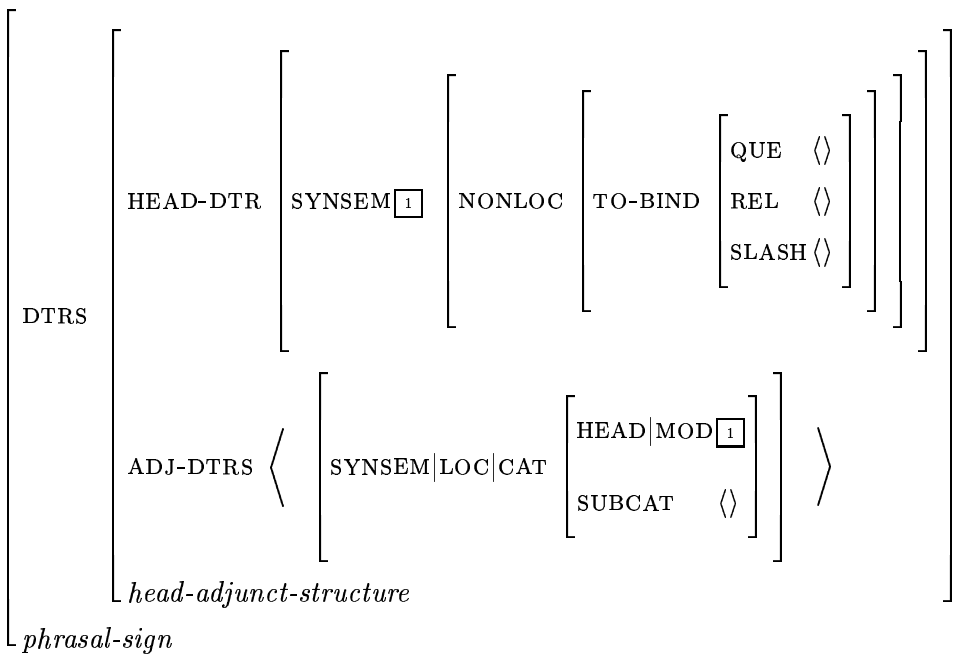
Schema 3 licenses head-adjunct structures. Again the adjunct daughter is inserted into the domain of the head. The free appearance of adverbs in the *Mittelfeld* can thus be explained.

Figure 1 shows the syntax tree for the sentence (13).

- (13) Gab der Frau der Mann gestern das Buch?
 Gave the woman the man yesterday the book
 'Did the man give the woman the book yesterday?'

⁴For details on extraposition and word order domains see (Müller, To appear) and (Kathol and Pollard, 1995).

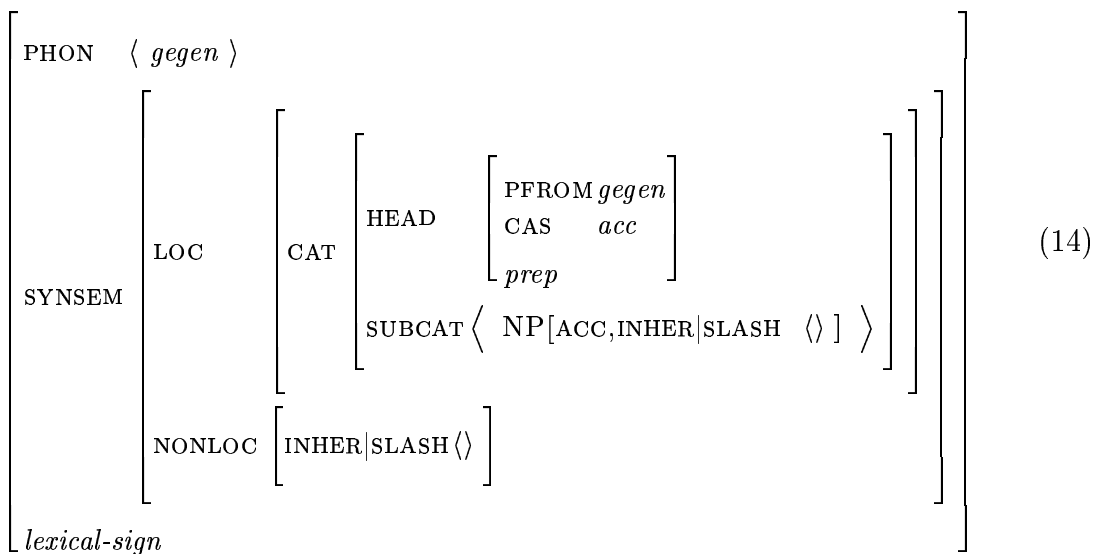
Schema 3 (Head-Adjunct Schema)



3.2 Stranded Prepositions

For the stranded preposition *gegen*, I assume the lexical entry in (15), which is generated by the lexical rule (17) from the lexical entry (14).

gegen:



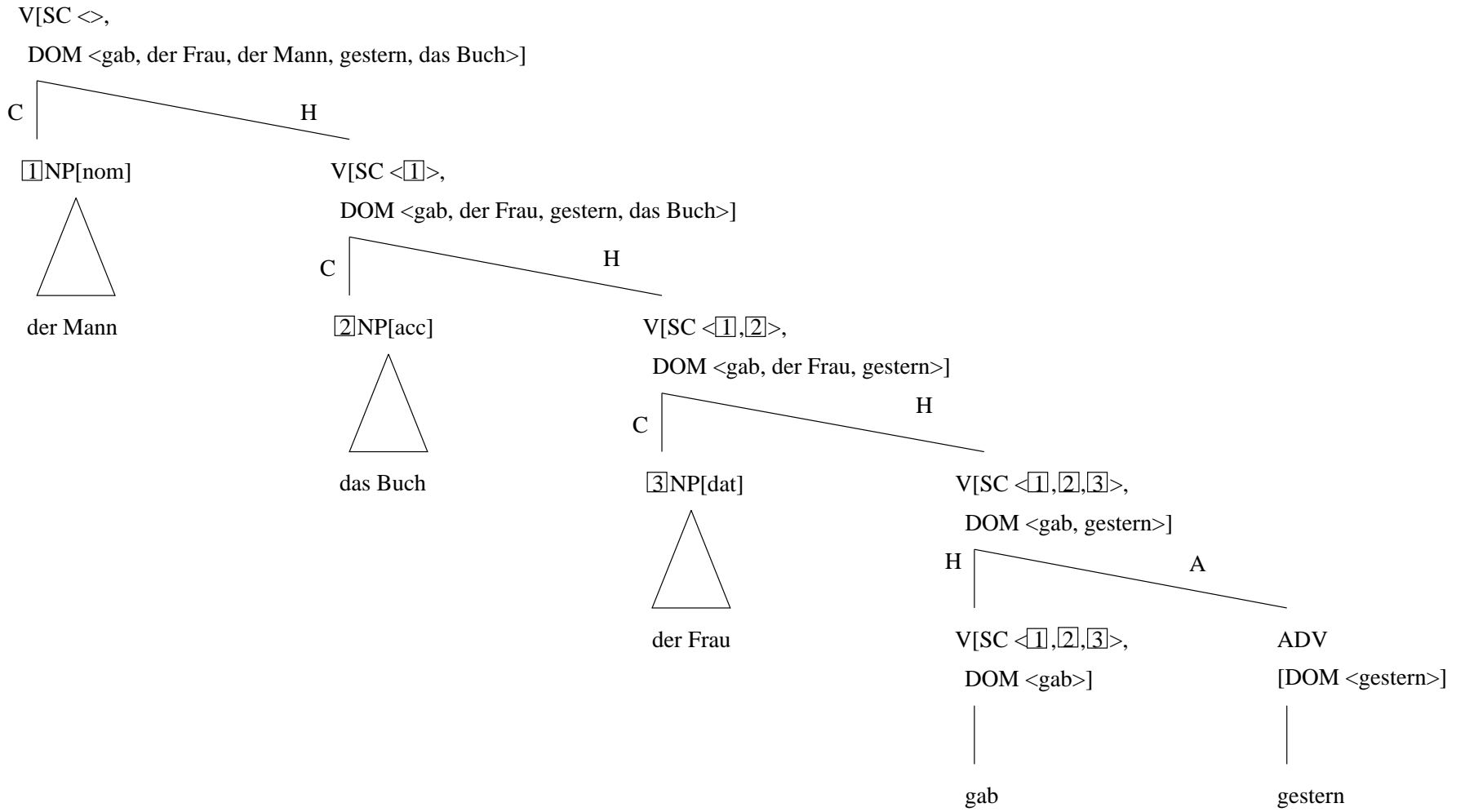


Figure 1: Analysis for *Gab der Frau der Mann gestern das Buch?*

da-gegen:

$$\left[\begin{array}{l} \text{PHON} \langle \textit{gegen} \rangle \\ \text{SYNSEM} \left[\begin{array}{l} \text{LOC} \left[\begin{array}{l} \text{CAT} \left[\begin{array}{l} \text{HEAD} \left[\begin{array}{l} \text{PFORM} \textit{gegen} \\ \textit{prep} \end{array} \right] \\ \text{SUBCAT} \langle \rangle \end{array} \right] \\ \text{NONLOC} \left[\text{INHER|SLASH} \langle \text{R-PRONOUN} \rangle \end{array} \right] \end{array} \right] \end{array} \right] \end{array} \right] \\ \textit{lexical-sign} \end{array} \right] \quad (15)$$

The LOCAL value of an R-pronoun is contained in the SLASH list⁵ and can percolate to the top of a finite verbal projection where it can be bound in filler position. Note that *da* is never a complement of *gegen*. It is either contained in the word (*dagegen*) or it is an element of the SLASH list. The regularities are captured by the lexical rules that produce the entries for *dagegen* (16) and *da-gegen* (17) from

⁵I assume the value of SLASH to be a list rather than a set for German, since there are no parasitic gaps in German. Sets as used in Pollard and Sag (1994) would only complicate the nonlocal mechanism in a grammar for German.

For instance an additional constraint would be needed to prevent the two accusative noun phrases in (i) from collapsing.

- (i) [Diesen Mann] hat Karl den Hund _{-i} beißen lassen.
 This man has Karl the dog bite let
 ‘Karl had the dog bite this man.’

If one assumes the set definition of Pollard and Moshier (1990), as Pollard and Sag (1994) do, constructions like (ii) would be possible.

- (ii) [Diesen Mann]_{i=j} hat Karl _{-i} _{-j} beißen lassen.

In (ii) two traces collapse and the sentence would have the semantics shown in (iii) instead of (iv).

- (iii) let(karl,bite(man,man))
 (iv) let(karl,bite(someone,man))

This would have to be avoided by an additional constraint on SLASH that parallels the Relative Uniqueness Principle of Pollard and Sag (1994, p. 212). However, if one assumes the SLASH value to be a list, such constraints are not necessary.

the input preposition *gegen*.

$$\left[\begin{array}{l} \text{PHON} \langle \boxed{1} \rangle \\ \text{SYNSEM} \left[\text{LOC|CAT|HEAD} \left[\textit{prep} \right] \right] \\ \textit{lexical-sign} \end{array} \right] \rightarrow \left[\begin{array}{l} \text{PHON} \ f(da(r), \boxed{1}) \\ \text{SYNSEM} \left[\text{LOC} \left[\begin{array}{l} \text{CAT} \left[\text{SUBCAT} \langle \rangle \right] \\ \text{CONT} \left[\textit{ppro} \right] \end{array} \right] \right] \\ \textit{lexical-sign} \end{array} \right] \quad (16)$$

$$\left[\begin{array}{l} \text{PHON} \langle \boxed{1} \rangle \\ \text{SYNSEM} \left[\text{LOC|CAT} \left[\begin{array}{l} \text{HEAD} \left[\textit{prep} \right] \\ \text{SUBCAT} \langle \left[\text{LOC} \boxed{2} \right] \rangle \right] \right] \\ \textit{lexical-sign} \end{array} \right] \rightarrow \left[\begin{array}{l} \text{PHON} \ f((dr), \boxed{1}) \\ \text{SYNSEM} \left[\text{LOC} \left[\begin{array}{l} \text{CAT} \left[\text{SUBCAT} \langle \rangle \right] \\ \text{NONLOC} \left[\text{INHER|SLASH} \langle \boxed{2} \text{ R-PRONOUN} \rangle \right] \end{array} \right] \right] \\ \textit{lexical-sign} \end{array} \right] \quad (17)$$

The lexical rules in (16) and (17) modify the PHON value of the input lexical entry. During the construction of the pro-PP form by (16), an *r* is inserted if the preposition starts with a vowel.

(18) a. Karl hat da*(r)über nachgedacht.

b. Karl hat davon geträumt.

(19) Da hat Karl *(dr)über nachgedacht.

With such a description of the phonological transformation, it is possible to explain the phenomenon in the lexicon. The syntactic analysis than can be done in a monotonous way. No transformations of PHON values is required. This differs from other accounts known from the GB framework.⁶

Note that the lexical rule (17) differs from the complement extraction lexical rule (CELR) one would need for a traceless analysis along the lines of Pollard and Sag (1994, Chapter 9). The CELR is shown in (20).⁷

$$\left[\begin{array}{l} \text{SYNSEM} \\ \text{lexical-sign} \end{array} \left[\begin{array}{l} \text{LOC} \left[\text{CAT|SUBCAT} \boxed{1} \oplus \langle \boxed{2} \rangle \oplus \boxed{3} \right] \\ \text{NONLOC} \left[\text{INHER|SLASH} \boxed{4} \right] \end{array} \right] \right] \rightarrow$$

(20)

$$\left[\begin{array}{l} \text{SYNSEM} \\ \text{lexical-sign} \end{array} \left[\begin{array}{l} \text{LOC} \left[\text{CAT|SUBCAT} \boxed{1} \oplus \boxed{3} \right] \\ \text{NONLOC} \left[\text{INHER|SLASH} \boxed{4} \oplus \langle \boxed{5} \rangle \right] \end{array} \right] \right]$$

Where $\boxed{2}$ corresponds to the structure in (21).

$$\left[\begin{array}{l} \text{LOC} \quad \boxed{5} \\ \text{NONLOC} \quad \text{INHER} \left[\begin{array}{l} \text{QUE} \langle \rangle \\ \text{REL} \langle \rangle \\ \text{SLASH} \langle \boxed{5} \rangle \end{array} \right] \end{array} \right] \right]$$

(21)

In (17), the SLASH value of the extracted complement is not specified. Therefore (17) can be applied to the lexical entry for *gegen* but the CELR cannot. Because preposition stranding in general is impossible in German the INHER|SLASH value in (14) is specified as $\langle \rangle$. With the instantiation of the complement in the input

⁶cf. (Oppenrieder, 1991)

⁷For details see (Müller, 1994).

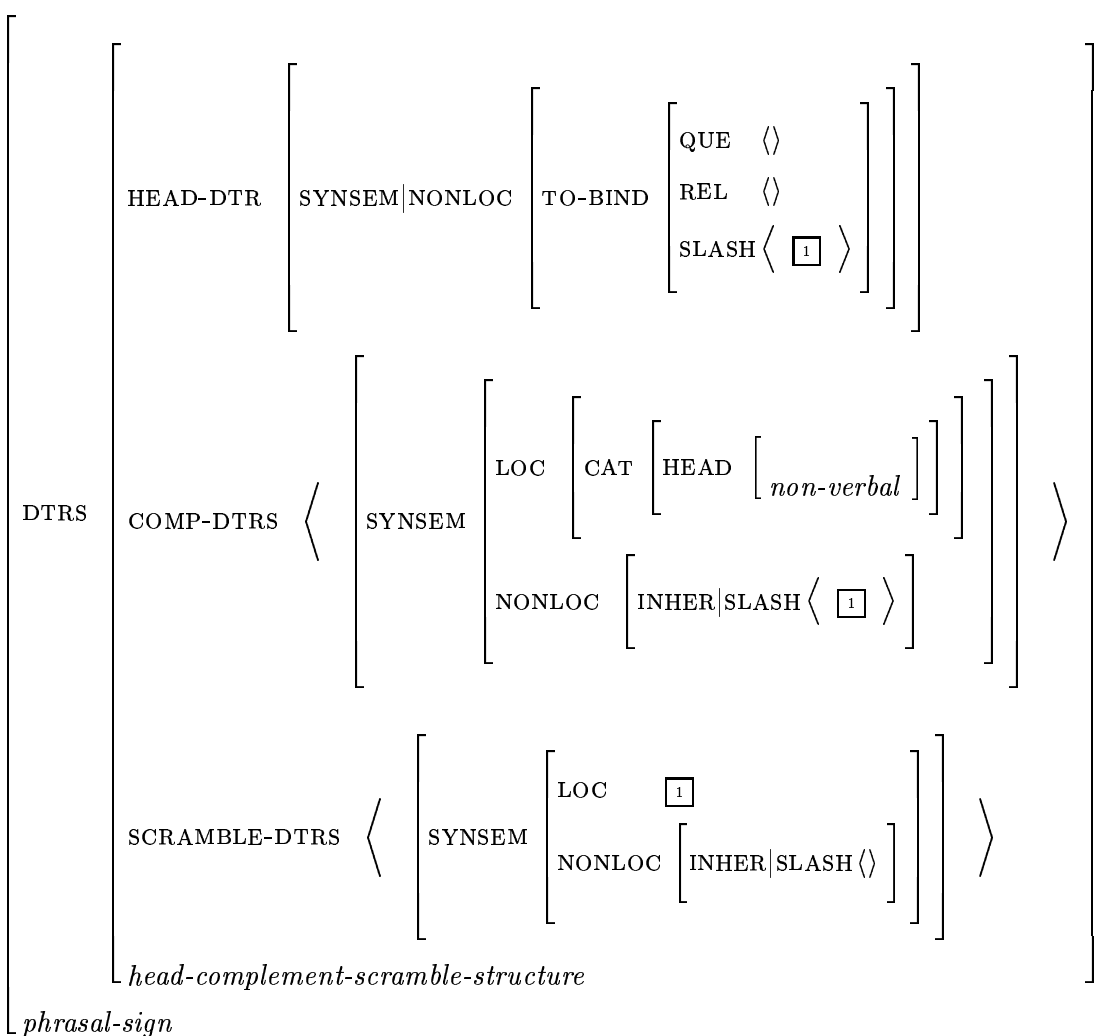
description, the CELR cannot apply and is in this respect equivalent to the trace analysis assumed in (Pollard and Sag, 1994, Chapter 1–8).

As the output description of the preposition stranding lexical rule further instantiates the extracted element and therefore only allows for the extraction of R-pronouns, the difference between (5b) and (6a) is explained.

3.3 The SLASH-Embedding Schema

The sentences in (7) have in common that a member of a nonverbal complement's SLASH list appears together with this complement in the domain of the verb. This can be accounted for by the following schema.

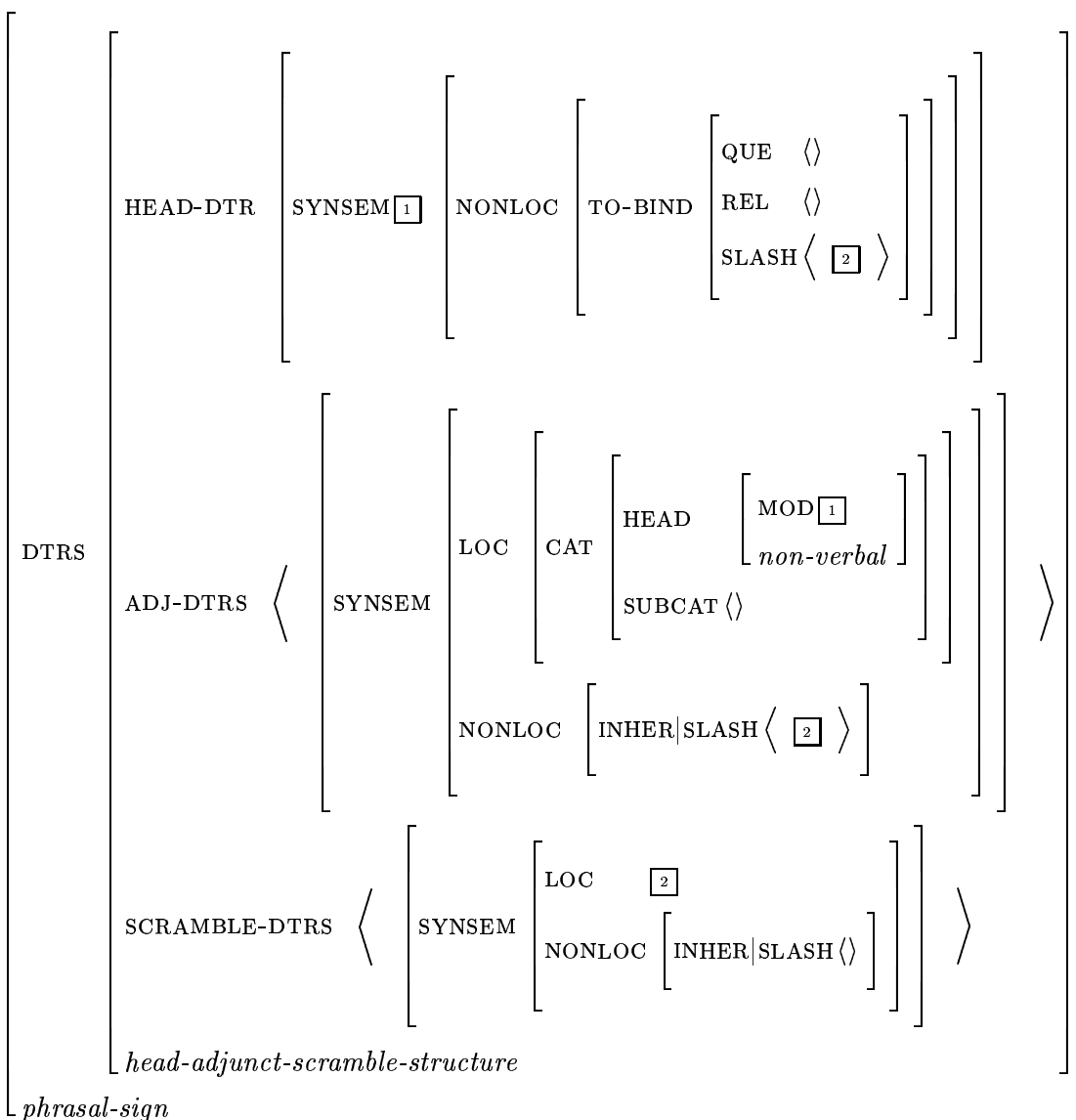
Schema 4 (SLASH-Embedding Schema (Complements))



In this schema, the valence of the head daughter is reduced by one element, specifically the element given as the value of the COMP-DTRS list. This complement has a non-empty SLASH list. One element of the SLASH list is identical to the local value of the SCRAMBLE-DTR ($\boxed{1}$). This element of the SLASH list of the complement daughter is bound by the NONLOCAL feature principle because the TO-BIND|SLASH value of the head daughter is $\boxed{1}$. The specification of the complement daughter as nonverbal reflects the fact that scrambling is clause bound.

An analogous schema is needed for head-adjunct structures to analyze sentences like (5d).

Schema 5 (SLASH-Embedding Schema (Adjuncts))



The scramble daughter is inserted by one of the following constraints for structures of type *head-complement-scramble-structure* and *head-adjunct-scramble-structure*.

$$\left[\begin{array}{l} \text{DTRS} \left[\text{head-complement-scramble-structure} \right] \\ \text{phrasal-sign} \end{array} \right] \Rightarrow \left[\begin{array}{l} \text{DTRS} \left[\begin{array}{l} \text{HEAD-DTR|DOM} \boxed{1} \\ \text{COMP-DTRS} \boxed{2} \\ \text{SCRAMBLE-DTRS} \boxed{3} \end{array} \right] \\ \text{DOM} \boxed{1} \circ \boxed{2} \circ \boxed{3} \end{array} \right] \quad (22)$$

$$\left[\begin{array}{l} \text{DTRS} \\ \text{phrasal-sign} \end{array} \left[\text{head-adjunct-scramble-structure} \right] \right] \Rightarrow \left[\begin{array}{l} \text{DTRS} \\ \text{DOM} \end{array} \left[\begin{array}{l} \text{HEAD-DTR|DOM} \boxed{1} \\ \text{ADJ-DTRS} \quad \boxed{2} \\ \text{SCRAMBLE-DTRS} \boxed{3} \\ \boxed{1} \circ \boxed{2} \circ \boxed{3} \end{array} \right] \right] \quad (23)$$

Scramble daughters can occur in the *Mittelfeld* only. For them, the same LP constraints hold as for complement daughters. In addition, the scramble daughter has to appear to the left of the constituent from which it is extracted.

If one assumes that a feature structure of type *headed-structure* has features for HEAD-DTR, FILLER-DTRS, ADJ-DTRS, COMP-DTRS and SCRAMBLE-DTRS one could generalize the domain formation in the following way.⁸

$$\left[\begin{array}{l} \text{DTRS} \\ \text{phrasal-sign} \end{array} \left[\text{headed-structure} \right] \right] \Rightarrow \left[\begin{array}{l} \text{DTRS} \\ \text{DOM} \end{array} \left[\begin{array}{l} \text{HEAD-DTR|DOM} \boxed{1} \\ \text{FILLER-DTRS} \quad \boxed{2} \\ \text{ADJ-DTRS} \quad \boxed{3} \\ \text{COMP-DTRS} \quad \boxed{4} \\ \text{SCRAMBLE-DTRS} \boxed{5} \\ \boxed{1} \circ \boxed{2} \circ \boxed{3} \circ \boxed{4} \circ \boxed{5} \end{array} \right] \right] \quad (24)$$

The value of the ...-DTRS is a list with at most one element. In a head complement structure for instance the value of ADJ-DTRS is the empty list.

This implication states that all daughters of a head appear in the domain of their head. The relative order of the daughters in respect to each other is constrained by LP rules. The same holds for the order of the daughters in respect to the head of the phrase.

3.4 Binary Branching Structures

It is conceptionally not very nice to have non-binary branching structures in a grammar where all other rules are binary branching. It is therefore reasonable to look for ways to describe the phenomena discussed above with binary branching structures.

For an account with binary branching structures three assumptions are necessary:

1. The SLASH list can contain more than one element.
2. The SLASH list of a mother sign is built up from the elements of the daughter sign in a particular order.
3. At most one element can be scrambled into the *Mittelfeld*.

⁸The formation of the verbal cluster is an exception. So the implicative constraint (24) has to be formulated for a subtype of *headed-structure*.

The first assumption seems to be supported by the following sentences by Fanselow (1987).

- (25) a. [Einladen]_i weiß ich nicht [_{w-s} [wen]_j [_s ich _{-j} [_{-i} könnte]]].
 invite know I not who I could
 ‘I don’t know who I could invite.’
- b. [Radios]_i weiß ich nicht [_{w-s} [wer]_j [_s _{-j} _{-i} repariert]].
 radios know I not who repairs
 ‘I don’t know who repairs radios.’

These sentences might not be the best stylistically, but they are grammatical nevertheless. If one assumes that *w*-sentences are analyzed similarly to relative clauses, then the phrase *ich könnte* contains two elements in its SLASH list: the *wh*-word *wen* and *einladen*.

The following sentence seems to be an example which contradicts assumption three.

- (26) Er hat mit Maria über dieses Thema ein ausführliches Gespräch
 He had with Maria about this topic a detailed talk
 geführt.
 carried on
 ‘He had a detailed talk with Maria on this topic.’

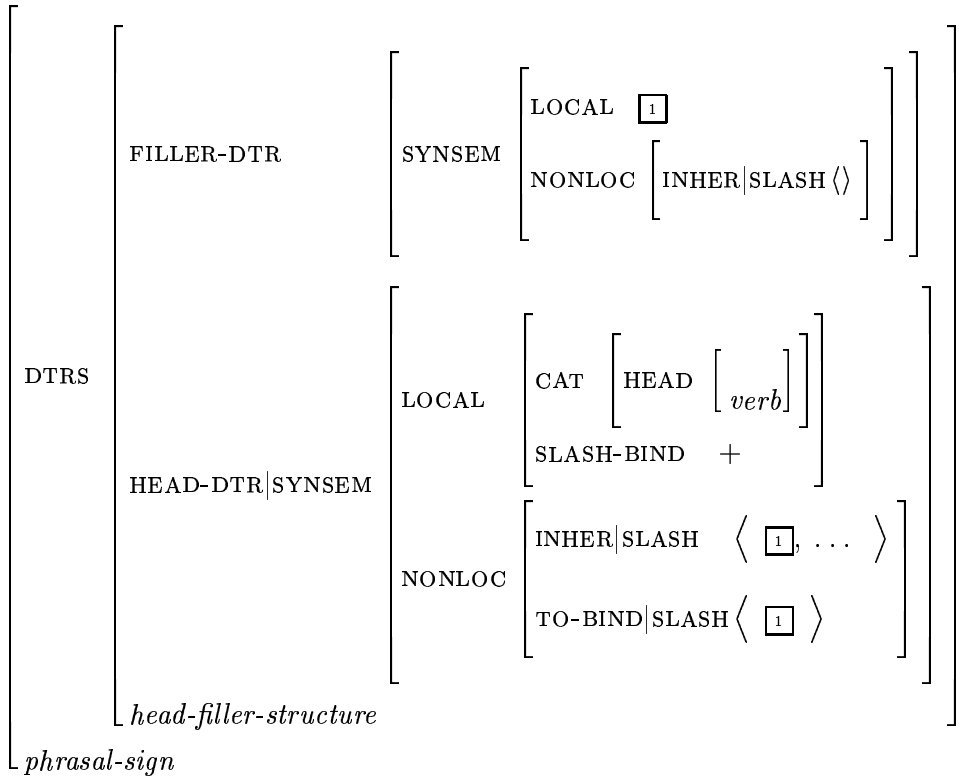
However, this sentence is not an instance of insertion of slash elements into the *Mittelfeld*, but rather a support verb construction (Krenn and Erbach, 1994). The arguments of *Gespräch* are raised to *geführt*. As arguments of the verb, the PPs can appear in the order they do appear in (26).

Schema 6 is a generalization of the Head-Filler Schema. It fulfills the function of the schemata 2, 4 and 5. It admits both verb final and verb initial phrases as head daughters. The head daughter might be not saturated.

The only restriction is that the value of SLASH-BIND has to be +. SLASH-BIND is a feature that becomes instantiated by a relational constraint that implements both the restrictions for movement into the *Vorfeld* and the restrictions for movement into the *Mittelfeld*.

SLASH-BIND has a value of type *slash-bind*. This type is partitioned into + and -. + has the two subtypes *vorfeld* (*vf*) and *mittelfeld* (*mf*). The assignment of

Schema 6 (Generalized Head-Filler Schema)



the value of SLASH-BIND is controlled by the following principle.⁹

Principle 1 *The value of SLASH-BIND is*

- + *if the sign is saturated and verb initial and there is an element in the SLASH list of the nonverbal non-head-daughter.*
- vorfeld* *if the non-head-daughter in a headed structure is verbal, the head is verb initial and saturated and the SLASH list of the resulting sign is not empty;*
if the resulting sign is saturated and verb initial and contains an element in the SLASH list that does not originate from the non-head-daughter
- mittelfeld* *if in a headed structure the resulting sign is not saturated and the nonverbal non-head-daughter contains an element in SLASH*
if the resulting sign is verb final and the nonverbal non-head-daughter contains an element in SLASH
- *otherwise.*

These constraints correspond to the constraints expressed in the preliminary version of the head-filler schema and the SLASH-insertion schema. The ordering of

⁹This principle is inspired by Keller (1995). He uses a similar principle to define landing sites for extraposed constituents.

bound fillers is constrained by the following three LP statements:

$$\begin{aligned}
 & FILLER < V[SLASH-BIND \textit{vf}] \\
 & V[SLASH-BIND \textit{mf}, INI+] < FILLER \\
 & FILLER < HEAD[INI-]
 \end{aligned}
 \tag{27}$$

Figure 2 shows how this works for the sentence (28b).

(28) a. Da hat Karl nicht mit gerechnet.

b. Hat Karl da nicht mit gerechnet?

In (28a) and (28b) *hat*, *nicht*, and *gerechnet* are combined. Then this phrase is saturated by *mit*. The slash-bind value of *hat nicht mit gerechnet* is *mittelfeld*. The SLASH-element (*da*) can be bound of in the next derivation step. It will then be inserted into the *Mittelfeld* (sentence (28b)). Otherwise, it percolates up the tree and gets bound of in the *Vorfeld* position (sentence (28a)).

The SLASH list works like a stack. The SLASH list of a complement daughter is appended to the SLASH list of its head daughter by the nonlocal feature principle. Therefore, for instance, a PP that is extracted from an NP becomes the first element of the slash list when the NP functions as a complement. The PP then can be bound of by schema 6.

3.5 Linearization

As was shown in section 2, extracted elements have to appear to the left of the element they are extracted from.

- (29) * Karl glaubt, daß [ein Bild _{-i}] Max [von Maria]_i malt.
 Karl believes that a picture Max of Maria paints
 ‘Karl believes that Max paints a picture of Maria.’

The following LP statement rules out the above sentence.

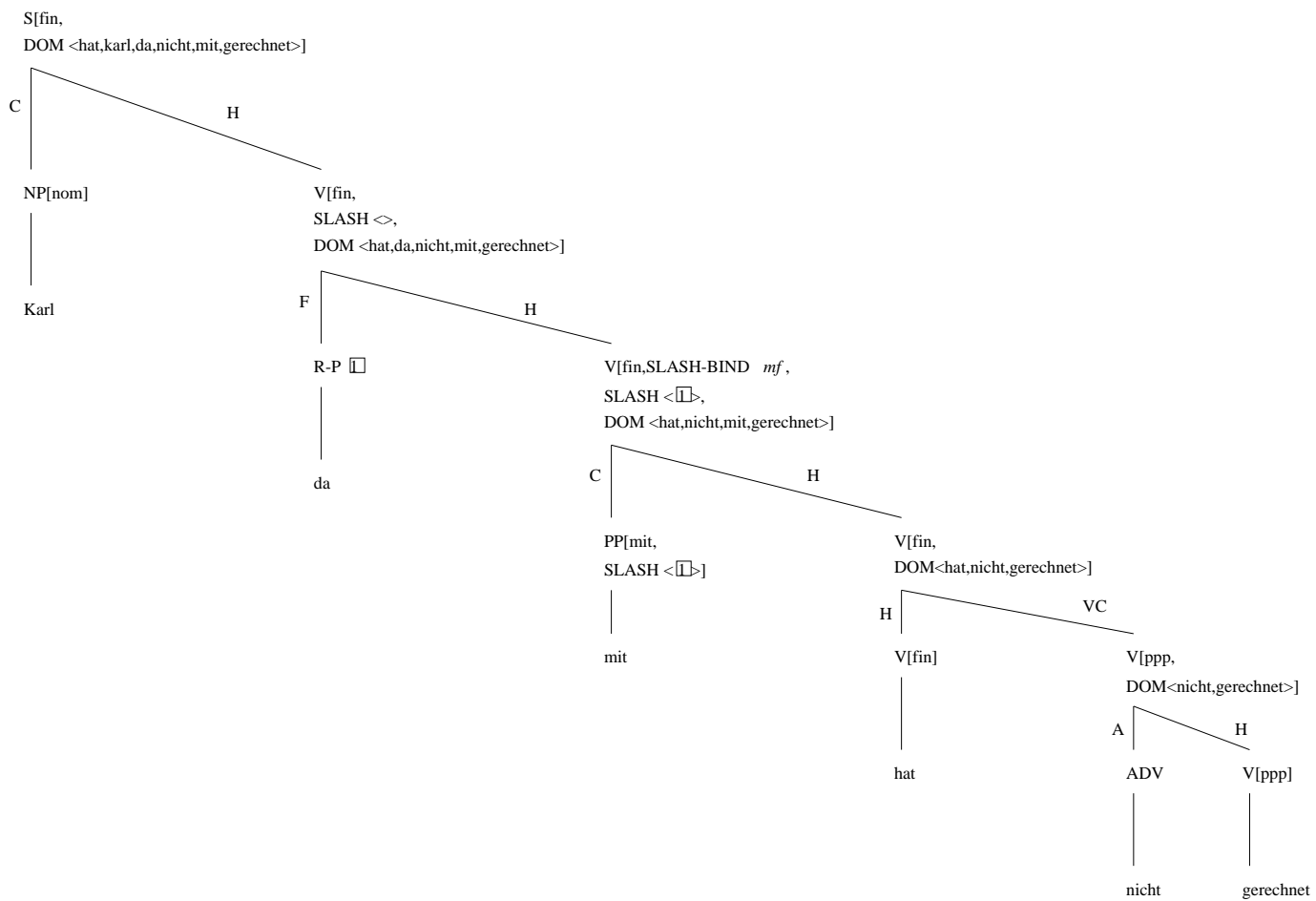
$$\left[\text{SYNSEM|LOC } \boxed{1} \right] < \left[\text{SYNSEM|NONLOC|INHER|SLASH } \langle \boxed{1} \rangle \right]
 \tag{30}$$

As Trissler (1993) notes, the generalization about the data in (31) seems to be that stranded prepositions cannot be extraposed over a finite verb.

- (31) a. Da hab ich schon gehört von.
 there have I already heard about
 ‘I have heard about this already.’

b. Wo hat er nichts zu tun mit?
 where has he nothing to do with
 ‘What has he got nothing to do with?’

- c. *, weil ich da gestern schon gehört habe von.
 because I there yesterday already heard have about

Figure 2: Analysis of *Hat Karl da nicht mit gerechnet?*

- d. * Wo will er nichts zu tun haben mit?
Where wants he nothing to do have with
- e. *, weil ich da gestern schon hörte von.
because I there yesterday already heard about

If this generalization holds, one can stipulate an additional LP constraint that rules out sentences like (31c–e).

$$\text{PP}[\text{SLASH } \langle \quad \rangle] < \text{V}[\text{LEX+}, \text{fn}, \text{INI-}] \quad (32)$$

4 Problems

It is unclear how sentences like (33) could be accounted for.

- (33) Karl ist sich gestern seiner Sache wie immer sehr sicher gewesen.
Karl has himself yesterday his ground like always very sure been
'As always Karl was very sure of himself yesterday.'

In (33) adverbs are scrambled between complements (*sich* and *seiner Sache*) of the adjective *sicher*.¹⁰

One possible explanation for sentences like this would be that the copula attracts the complements of its arguments. *Sich* and *seiner Sache* would then be complements of *gewesen* and the appearance of adjuncts between these NPs would be normal.

5 Alternatives

5.1 SALSH-to-Argument-Attraction

Pollard, Kasper, and Levine (1994) draw the conclusion that scrambling is a lexical phenomenon from the following sentences.¹¹

- (34) a. , weil Peter einen Film über England gesehen/zerstört hat.
because Peter a film about England seen/destroyed has
'because Peter has seen/destroyed a film about England.'

¹⁰Note that *sich* in (33) is a reflexive NP.

- (i) * Karl war ihm seiner Sache sicher.
Karl was him his ground sure

Reflexives in constructions that are lexical reflexive can never be fronted in German.

- (ii) Karl erholt sich gut.
Karl recreates self well.
'Karl recreates well.'

- (iii) * Sich erholt Karl gut.

¹¹The # in (35b) indicates that the sentence is not grammatical with the meaning where the *über*-PP is the topic of the film.

b. , weil Peter über England einen Film gesehen hat.

c. # , weil Peter über England einen Film zerstört hat.

They assume that special lexical entries for verbs that allow for scrambling exist. They have the following form:

$$\left[\begin{array}{l} \text{LOCAL|CAT|SUBCAT} \langle \dots, \boxed{1} \text{ PP, NP[SLASH } \{ \boxed{1} \}] \rangle \\ \text{NONLOC|TO-BIND|SLASH } \{ \boxed{1} \} \end{array} \right] \quad (35)$$

The mechanism is similar to the one proposed by Hinrichs and Nakazawa (1989). The SLASH-element of a complement is inserted into the subcat list of the governing verb. There would be such an entry for *sehen* but none for *zerstören*.

Currently I am uncertain as to what kind of constraint rules out sentences like (34c), but it is certainly not the impossibility of scrambling PP-complements of arguments of *zerstören*. As (36) shows, such a scrambling is possible. It seems to me that there are semantic constraints that allow (36) but rule out (34c). In general, it is easier to extract *von*-PPs than other PPs.

- (36) Maria ist sehr traurig, weil Peter [von ihrer Puppe]_i gestern
 Maria is very sad because Peter of her doll yesterday
 auch noch [den Arm _{-i}] zerstört hat.
 also yet the arm destroyed had
 ‘Maria is very sad because yesterday Peter destroyed her doll’s arm
 as well.’

The same contrast as in (34) where phrases are inserted into the *Mittelfeld* can be found for fronting.

- (37) a. Über England hat Peter einen Film gesehen.
 b. # Über England hat Peter einen Film zerstört.

The fronted PP in (37b) specifies the location of the destruction not the topic of the film. This is not explained by an additional entry for *gesehen* that can be used in constructions like (34b).

5.2 Domain Union

Pollard, Kasper, and Levine (1994) suggest to describing the word order in (38) by means of domain union.

- (38) , weil da die Leute Geld für gesammelt haben.
 because there the people money for collected have
 ‘because the people have collected money for that.’

They assume an optional domain union of the verb with its prepositional complement. Together with an LP rule that states that lexical domain elements have to follow phrasal ones in a domain with a verbal head in final position it is possible both to account for sentences like (38) and to rule out (39).

- (39) * weil mit keiner [der Ermordung der Geiseln] gerechnet
 because with nobody the killing of-the hostages counted
 hatte.
 had
 ‘because nobody expected the killing of the hostages.’

In (38) *da-für* is inserted into the domain of *gesammelt haben*. Other members of this domain may appear between *da* and *für*. (39) is ruled out by the LP constraint which is violated because *mit* as a lexical element appears to the left of a phrasal element (*der Ermordung der Geiseln*).

However, it is unclear what rules out sentences like (40).

- (40) * weil der Ermordung der Geiseln keiner mit gerechnet hat.
 because the killing of-the hostages nobody with counted has

The biggest problem with this account is that it is not possible to analyze sentences like (7d). In (7d) the *da* appears to the left of the entire NP. This means that the NP would have to be domain-unioned into the domain of the verb (*hat vorgebracht*). This would allow scrambling of adjuncts between determiners and nominal heads which is totally ungrammatical.

- (41) * Deshalb hat da Karl ein wieder Argument gegen vorgebracht.
 Therefore has there Karl an again argument against advanced
 ‘Therefore Karl has again advanced an argument against this.’

An additional argument against the domain union analysis is relative clause constructions. In dialects of German that accept *wo* (where) as a relative word, sentences like (42) are possible.

- (42) unnützes Zeug, [wo]_i sie immer mal wieder [Geld [für _{-i}]] ausgibt,
 useless stuff where she ever once again money for spends
 ‘useless stuff that she spends money on time and time again

It is reasonable to describe German relative clause constructions analogously to their English counterpart (Pollard and Sag, 1994, Chapter 5). A relative phrase is extracted from a finite sentence with the finite verb in final position (Müller, To appear). Sentences like (42) are explainable without any additional assumptions with the extraction analysis given in the previous sections.

6 Conclusion

A unified account for scrambling in German has been presented. The proposed scrambling schema is superior to both slash-to-argument attraction and domain union analyses.

The schema is part of an implemented fragment of German¹² (Müller, To appear). The fragment covers several types of word order phenomena, for instance partial verb phrase fronting, auxiliary flip, scrambling in coherent constructions, and free

¹²http://www.dfki.de/~stefan/Babel/e_babel.html

appearance of adjuncts in the *Mittelfeld*. The schema fits nicely into the grammar and gives the right results even in more complicated cases like (43), which is an example of partial verb phrase fronting and preposition stranding.

- (43) Gerechnet hatte Karl da nicht mit.
 counted had Karl there not with
 ‘Karl had not expected that to happen.’

References

- Duden. 1984. *Grammatik der deutschen Gegenwartssprache*, volume 4. Mannheim, Wien, Zürich: Dudenverlag.
- Fanselow, Gisbert. 1987. *Konfiguralität*. Studien zur deutschen Grammatik, number 29. Tübingen: Gunter Narr Verlag.
- Hinrichs, Erhard W. and Tsuneko Nakazawa. 1989. Subcategorization and VP structure in German. In *Aspects of German VP Structure*, SfS-Report-01-93. Eberhard-Karls-Universität Tübingen.
- Kathol, Andreas and Carl J. Pollard. 1995. Extraposition via complex domain formation. In *Proceedings of the Thirty-Third Annual Meeting of the ACL*, Boston. Association for Computational Linguistics. <http://linguistics.berkeley.edu/~kathol/Papers/ACL95.ps.gz>. 09.21.97.
- Keller, Frank. 1995. Towards an account of extraposition in HPSG. In *Proceedings of the Seventh Conference of the European Chapter of the Association for Computational Linguistics*, Dublin. <ftp://ftp.ims.uni-stuttgart.de/pub/papers/keller/eacl95.ps.gz>. 04.15.97.
- Krenn, Brigitte and Gregor Erbach. 1994. Idioms and support verb constructions. In John Nerbonne, Klaus Netter, and Carl J. Pollard, editors, *German in Head-Driven Phrase Structure Grammar*, CSLI Lecture Notes, number 46. Center for the Study of Language and Information, Stanford, chapter 11, pages 365–396. <http://coli.uni-sb.de/claus/claus28.html>. 09.11.97.
- Müller, Stefan. 1994. Problems with complement extraction lexical rules. http://www.dfki.de/~stefan/Pub/e_celr.html. 06.16.1998. ms. Humboldt University Berlin.
- Müller, Stefan. To appear. *Deutsche Syntax deklarativ. Head-Driven Phrase Structure Grammar für das Deutsche*. Linguistische Arbeiten. Tübingen: Max Niemeyer Verlag. http://www.dfki.de/~stefan/Pub/e_hpsg.html. 06.16.1998.
- Oppenrieder, Wilhelm. 1991. Preposition Stranding im Deutschen? – Da will ich nichts von hören! In Gisbert Fanselow and Sascha W. Felix, editors, *Strukturen und Merkmale syntaktischer Kategorien*. Gunter Narr Verlag, Tübingen, pages 159–173.

Pollard, Carl J., Robert T. Kasper, and Robert D. Levine. 1994. Studies in constituent ordering: Toward a theory of linearization in Head-Driven Phrase Structure Grammar. Research Proposal to the National Science Foundation.

Pollard, Carl J. and Drew Moshier. 1990. Unifying partial descriptions of sets. In P. Hanson, editor, *Information, Language and Cognition, Vancouver Studies in Cognitive Science I*. University of British Columbia Press, Vancouver, pages 285–322.

Pollard, Carl J. and Ivan A. Sag. 1987. *Information-Based Syntax and Semantics Volume 1 Fundamentals*. CSLI Lecture Notes, number 13. Stanford: Center for the Study of Language and Information.

Pollard, Carl J. and Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Studies in Contemporary Linguistics. Chicago, London: University of Chicago Press.

Reape, Mike. 1994. Domain union and word order variation in German. In John Nerbonne, Klaus Netter, and Carl J. Pollard, editors, *German in Head-Driven Phrase Structure Grammar*, CSLI Lecture Notes, number 46. Center for the Study of Language and Information, Stanford, chapter 5, pages 151–198.

Rentier, Gerrit. 1994. A lexicalist approach to Dutch preposition stranding. In Harald Trost, editor, *Tagungsband Konvens 94, Verarbeitung natürlicher Sprache*, pages 280–289.

Riemsdijk, Henk van. 1978. *A Case Study in Syntactic Markedness. The Binding Nature of Prepositional Phrases*. Lisse: The Peter de Ridder Press.

Trissler, Susanne. 1993. P-stranding. In Franz-Josef d’Avis, Sigrid Beck, Uli Lutz, Jürgen Pafel, and Susanne Trissler, editors, *Extraktion im Deutschen I*, Arbeitspapiere des SFB 340, number 34. Universität Tübingen.

**Scrambling in German –
Extraction into the *Mittelfeld***

Stefan Müller

RR-97-06
Research Report