

# TRACE & UNIFICATION GRAMMAR

Hans Ulrich Block and Stefanie Schachtl  
Siemens AG, Corporate Research, ZFE ST SN 74  
Otto Hahn-Ring 6  
D-8000 München 83  
Germany  
block@ztivax.zfe.siemens.de

## Abstract

This paper presents Trace & Unification Grammar (TUG), a declarative and reversible grammar formalism that brings together Unification Grammar (UG) and ideas of Government & Binding Theory (GB). The main part of the paper consists in a description of many free word order phenomena of German syntax. It will be shown that the use of traces allows for an easier and more elegant way of description than competing approaches like ID/LP-format rules as used e.g. in GPSG and HPSG. Grammars written in the TUG-formalism can be compiled to a very efficient parser. The occurrence of head movement, wh-movement and scrambling in one sentence does not lead to any decrease in parsing efficiency.

## 1 Introduction

The basic motivation in creating TUG formalism was to capture the empiric knowledge that represents the outcome of at least 15 years' linguistic discussion of German word order, while maintaining the efficiency that is required from a grammar formalism of today's standard. With reference to the basic work of [Len77] on marked and unmarked word order linguists in the generative tradition such as [Thi82], [dB84], [Fan87] and [Cze87] revealed an impressive list of descriptive phenomena that can be appropriately handled by the assumption of a configurational, i.e. VP containing description of German word order. Among these phenomena count asymmetries in the serialization behaviour of nominatively marked NPs in passive and the so called "psych" constructions<sup>1</sup>, the account for scrambling phenomena in Acl/ECM constructions<sup>2</sup> and the observations of [Cze87] wrt. to the voices of double accusative verbs in German. What all these works intend is to pronounce the danger a description of German is likely to run into, that "draws immediate conclusions as to the surface position of argument NPs on the basis of their surface cases" ([dB84]:59).

<sup>1</sup>[Len77] shows that verbs like *wundern*, *gefallen*, *gelingen*, etc. with non agentive subjects pattern with the passivized forms of "agentive" verbs in preferring the unmarked word order indirect object - subject. [dB84] extends this observation on copular construction with NP governing adjectives

<sup>2</sup>see also [dB84] for a short outline of this problematic issue

## 2 The TUG formalism

The basis of TUG is formed by a context free grammar that is augmented by PATR II-style feature equations. Besides this basis, the main features of TUG are feature typing, mixing of attribute-value-pair and (PROLOG-) term unification, flexible macros, unrestricted disjunction and special rule types for argument and head movement.

### 2.1 Basic rule types

As a very simple example we will look at the TUG version of the example grammar in [Shi84]<sup>3</sup>.

```
% type definition

s    => f.
np   => f(agr:agrmt).
vp   => f(agr:agrmt).
v    => f(agr:agrmt).

agrmt => f(number:number,person:person).

number => {singular,plural}.
person => {first,second,third}.

% rules

s ---> np, vp |
      np:agr = vp:agr.

vp ---> v, np |
      vp:agr = v:agr.
```

The two main differences to PATR II in the basic framework are that first, TUG is less flexible in that it has a "hard" contextfree backbone, whereas in PATR II categories of the context free part are placeholders for feature structures, their names being taken as the value of the *cat* feature in the structure. Second, TUG has a strict typing. For a feature path to be well defined, each of its attributes has to be declared in the type definition.

<sup>3</sup>For a slightly more detailed description of the basic features of TUG refer to ([Blo91]).

## 2.2 Movement rules

Further to these more standard UG-features, TUG provides special rule formats for the description of discontinuous dependencies, so called "movement rules". Two main types of movement are distinguished: argument movement and head movement. The format and processing of argument movement rules is greatly inspired by [CLW88] and [Che90], the processing of head movement is based on GPSG like slash features.

### 2.2.1 Head Movement

A head movement rule defines a relation between two positions in a parse tree, one is the landing site, the other the trace position. Head movement is constrained by the condition that the trace is the head of a specified sister (the root node) of the landing site<sup>4</sup>. Trace and antecedent are identical with the exception that the landing site contains overt material, the trace doesn't.

To formulate head movement in TUG the following format is used. First, a head definition defines which category is the head of which other, e.g. for the V-projection line of the above grammar:

```
v is_head_of vp.  
vp is_head_of s.
```

Second, the landing site is defined by a rule like

```
s' ---> v+s | ...
```

where landing site and root node are linked by a +. To include recursive rules in the head path, heads are defined by the following head definitions. In a structure  $[M D_1 \dots D_n]$   $D_i$  is the head of  $M$  if either  $D_i$  is\_head\_of  $M$  is defined or  $D_i$  has the same category as  $M$  and either  $D_i$  is\_head\_of  $X$  or  $X$  is\_head\_of  $D_i$  is defined for any category  $X$ .

### 2.2.2 Argument Movement

Argument movement rules describe a relation between a landing site and a trace. The trace is always c-commanded by the landing site, its antecedent. Two different traces are distinguished, anaphoric traces and variable traces. Anaphoric traces must find their antecedent within the same bounding node, variable trace binding is constrained by subadjacency, e.a. the binding of the trace to its antecedent must not cross two bounding nodes. Anaphoric traces are found for example in English passive constructions  $[s [np \text{The book of this author}]_i \text{ was read } t_i]$  whereas variable traces are usually found in wh-constructions and topicalization. Similar to the proposal in [CLW88], argument movement is coded in TUG by a rule that describes the landing site, as for example in

<sup>4</sup>Here, "head of" is a transitive relation s.t. if  $x$  is head of  $y$  and  $y$  is head of  $z$  then  $x$  is head of  $z$ .

```
s2 ---> np:ante<trace(var,np:trace), s1 |  
      ante:fx = trace:fx,  
      ...
```

This rule states that  $np:ante^5$  is the antecedent of an  $np$ -trace that is dominated by  $s1$ .

The first argument in the trace-term indicates whether the landing site is for a variable ( $var$ ) or for an anaphoric ( $ana$ ) trace. Other than head movement, where trace and antecedent are by definition identical, the feature sharing of argument traces with their antecedents has to be defined in the grammar by feature equations ( $ante:fx = trace:fx, \dots$ ). Furthermore, it is not necessary that the antecedent and the trace have the same syntactic category.

The current version of the formalisms requires that the grammar contains a declaration on which categories are possible traces. In such a declaration it is possible to assign features to a trace, for example marking it as empty:

```
trace(np) | np:empty = yes.
```

Bounding nodes have to be declared as such in the grammar by statements of the form

```
bounding_node(np).  
bounding_node(s) | s:tense = yes.
```

As in the second case, bounding nodes may be defined in terms of category symbols and features.

The main difference of argument movement to other approaches for the description of discontinuities like extraposition grammars ([Per81]) is that argument movement is not restricted to nested rule application. This makes the approach especially attractive for a scrambling analysis of the relative free word order in the German *Mittelfeld* as explained in more detail below.

## 3 Some facts on German syntax

### 3.1 Basic assumptions

In the following we will sketch the basic structures of German syntax. According to the position of the finite verb, we distinguish sentences with the verb in the second (a), the first(b) and the last position (c).<sup>6</sup>

- (1) a. Karl fährt nach Hamburg.  
Karl goes to Hamburg
- b. Führt Karl nach Hamburg?  
Goes Karl to Hamburg
- c. ... daß Karl nach Hamburg fährt.  
... that Karl to Hamburg goes

<sup>5</sup>The notation  $Car:Index$  is used to distinguish two or more occurrences of the same category in the same rule in the equation part.  $ante$  and  $trace$  are arbitrary names used as index to refer to the two different  $aps$ .

<sup>6</sup>English literal translations are given in typewriter font.

We call the sentence types V2-S (a), V1-S (b) and Ve-S (c) respectively. In V1-S and V2-S, compound verbforms build a so called *Satzklammer* between the finite and the nonfinite parts.

- (2) Karl ist mit dem Zug nach Hamburg  
Karl has with the train to Hamburg  
gefahren.  
gone  
"Karl has gone to Hamburg by train."

Traditionally, according to the position of the verbal elements, we call the position in front of the finite verb the *Vorfeld* and the positions within the *Satzklammer* the *Mittelfeld*.

In accordance with the mentioned configurational analysis of the german sentence, we suppose an unmarked "underlying" structure, that is similar to the order in the german subordinate clause. This structure contains four different positions for verb arguments, as exemplified by the following sentences.

- (3) a. [<sub>S2</sub> [<sub>S1</sub> daß [<sub>S</sub> [<sub>NP</sub> der Mann] [<sub>VP</sub> [<sub>NP</sub> der Frau] [<sub>V1</sub> [<sub>NP</sub> das Buch] [<sub>VK</sub> gegeben hat] ] ] ] ] ]  
that the man the woman the book given has  
"that the man has given the woman the book"  
b. [<sub>S2</sub> [<sub>S1</sub> daß [<sub>S</sub> [<sub>NP</sub> der Mann] [<sub>VP</sub> [<sub>V1</sub> [<sub>NP</sub> das Buch] [<sub>VK</sub> [PP in den Schrank] [<sub>VK</sub> gelegt hat] ] ] ] ] ] ]  
that the man the book into the bookshelf put has  
"that the man has put the book into the bookshelf"

A sentence always contains the v-projection line VK, V<sup>1</sup>, VP, S, even if the sentence contains less than three arguments.

- (4) a. [<sub>S2</sub> [<sub>S1</sub> daß [<sub>S</sub> [<sub>NP</sub> der Mann] [<sub>VP</sub> [<sub>V1</sub> [<sub>VK</sub> tanzt]]]]]  
that the man[nom] dances  
b. [<sub>S2</sub> [<sub>S1</sub> daß [<sub>S</sub> [<sub>VP</sub> [<sub>NP</sub> dem Mann] [<sub>V1</sub> [<sub>VK</sub> geholfen wird]]]]]  
that the man[dat] helped is  
"that the man is helped"  
c. [<sub>S2</sub> [<sub>S1</sub> daß [<sub>S</sub> [<sub>VP</sub> [<sub>V1</sub> [<sub>NP</sub> das Buch] [<sub>VK</sub> gelesen wird]]]]]  
that the book[nom] read is  
"that the book is read"  
d. [<sub>S2</sub> [<sub>S1</sub> daß [<sub>S</sub> [<sub>VP</sub> [<sub>V1</sub> [<sub>VK</sub> getanzt wird]]]]]  
that danced is  
"that there is dancing"

As is shown in (4a.) vs. (4c.) the nominative may be assigned to a VP-external or a VP-internal position. Adverbials are chomsky-adjoined to S, VP, V<sup>1</sup> and VK. An adjunction is only possible, if the right daughter is binarily branching.

- (5) a. [<sub>S2</sub> [<sub>S1</sub> daß [<sub>S</sub> heute [<sub>S</sub> [<sub>NP</sub> der Mann] [<sub>VP</sub> [<sub>V1</sub> [<sub>VK</sub> tanzt]]]]]]]  
b. [<sub>S2</sub> [<sub>S1</sub> daß [<sub>S</sub> [<sub>VP</sub> heute [<sub>VP</sub> [<sub>NP</sub> dem Mann] [<sub>V1</sub> [<sub>VK</sub> geholfen wird]]]]]]]  
c. [<sub>S2</sub> [<sub>S1</sub> daß [<sub>S</sub> [<sub>VP</sub> [<sub>V1</sub> heute [<sub>NP</sub> das Buch] [<sub>VK</sub> gelesen wird]]]]]]]  
d. [<sub>S2</sub> [<sub>S1</sub> daß [<sub>S</sub> [<sub>VP</sub> [<sub>V1</sub> [<sub>VK</sub> heute [<sub>VK</sub> getanzt wird]]]]]]]]]

For a word order that differs from the underlying structure the movement rules of TUG are used. V1-S are formed by head-movement of the finite verb to the position of the complementizer.

- (6) a. [<sub>S2</sub> [<sub>S1</sub> daß [<sub>S</sub> Peter [<sub>VP</sub> [<sub>V1</sub> die Bilder] [<sub>VK</sub> abmalt]]]]]  
that Peter the pictures copies  
"that Peter copies the pictures"  
b. [<sub>S2</sub> [<sub>S1</sub> malti [<sub>S</sub> Peter [<sub>VP</sub> [<sub>V1</sub> die Bilder] [<sub>VK</sub> ab t<sub>i</sub>]]]]]]]  
copies Peter the pictures  
"Does Peter copy the pictures?"

This can be formulated in a rather compact way in TUG by definition of the head relation and rules for the introduction of the landing site of the finite verb.

- (7) v is\_head\_of vk.  
vk is\_head\_of v1.  
v1 is\_head\_of vp.  
vp is\_head\_of s.  
(8) s1 ---> comp, s  
s1 ---> v+s

V2-S are formed by occupying the *Vorfeld*, i.e. the position immediately dominated by S<sup>2</sup> with either a verb argument by argument movement (a), by an adverbial (b) or by a *Vorfeld-es* (c).

- (9) a. [<sub>S2</sub> Peter<sub>i</sub> [<sub>S1</sub> malti [<sub>S</sub> t<sub>i</sub> [<sub>VP</sub> [<sub>V1</sub> die Bilder] [<sub>VK</sub> ab t<sub>i</sub>]]]]]]]  
Peter copies the pictures  
b. [<sub>S2</sub> Heute [<sub>S1</sub> malti [<sub>S</sub> Peter [<sub>VP</sub> [<sub>V1</sub> die Bilder] [<sub>VK</sub> ab t<sub>i</sub>]]]]]]]  
Today copies Peter the pictures  
c. [<sub>S2</sub> Es [<sub>S1</sub> malti [<sub>S</sub> Peter [<sub>VP</sub> [<sub>V1</sub> die Bilder] [<sub>VK</sub> ab t<sub>i</sub>]]]]]]]  
It copies Peter the pictures  
"Peter copies the pictures"

This facts can be described by the following rules:

- (10) s2 ---> es, s1  
s2 ---> pp, s1  
s2 ---> advp, s1  
s2 ---> np<trace(var,np), s1  
trace(var,np).

Free word order in the *Mittelfeld* is described by "moving" an argument to a chomsky-adjoined position on the V-projection. Here it obeys the same conditions as an adverbial and leaves a trace in the original argument position.

- (11) a. [<sub>S</sub><sup>2</sup> [<sub>S</sub><sup>1</sup> daß [<sub>S</sub> der Mann [<sub>VP</sub> [<sub>NP</sub> das Buch]<sub>i</sub> [<sub>VP</sub> der Frau [<sub>v</sub><sub>i</sub> [<sub>VK</sub> gegeben haß]]]]]]]]]  
 that the man the book the woman given has  
 b. [<sub>S</sub><sup>2</sup> [<sub>S</sub><sup>1</sup> daß [<sub>S</sub> [<sub>NP</sub> der Frau]<sub>j</sub> [<sub>S</sub> [<sub>NP</sub> das Buch]<sub>i</sub> [<sub>S</sub> der Mann [<sub>VP</sub> [<sub>v</sub><sub>i</sub> [<sub>VK</sub> gegeben haß]]]]]]]]]  
 that the woman[dat] the book[acc] the man[nom] given has  
 "that the man has given the woman the book"

So, for scrambling, we basically need the following rules:

- s ---> np<trace(ana,np), s  
 vp ---> np<trace(ana,np), vp

### 3.2 Alternative approaches

Whereas most concurrent theories adopt the view that an argument phrase in the *Vorfeld* is linked to the argument position by a trace - be it by movement or by the slash-feature - the relative free word order in the *Mittelfeld* is often accounted for by the distinction of phrase structure rules into immediate dominance (ID) rules and linear precedence (LP) rules. ID rules define the hierarchical structure of constituents, LP rules the linear ordering of daughters constituents. In this paradigm the German *Mittelfeld* including the finite verb typically is supposed to form a flat structure, generated by an ID rule like s ---> np[nom], np[akk], np[dat], v[fin], vk[infin]<sup>7</sup>. The elements on the right hand side can then be (partially) ordered by LP statements of the form v[fin] < np, np < vk (a finite verb precedes an NP, a VK follows an NP). As no LP statement is made for the NPs, the rules generate all possible permutations of NPs.

- (13) [s hat der Mann der Frau das Buch gegeben]  
 [s hat der Frau der Mann das Buch gegeben]  
 [s hat das Buch der Mann der Frau gegeben]  
 [s hat der Frau das Buch der Mann gegeben]  
 [s hat der Mann das Buch der Frau gegeben]

So, where TUG supposes a fixed unmarked word order, from which marked orders are derived by movement rules (scrambling), GPSG and HPSG suppose unordered ID rules and express constraints on order explicitly by LP statements. The same holds for the position of the finite verb in the different German sentence types. As for movement to the *Vorfeld* the GPSG/HPSG approach using the slash feature and the TUG approach are rather similar, as in fact movement is implemented in TUG by structure sharing.

### 3.3 Some more facts on German syntax

In the following paragraph we will outline a few descriptive phenomena where we think the use of traces

<sup>7</sup>But see [Rea89] for an alternative approach using LP-statements that does not have to assume a flat structure.

as in TUG allows for more elegant formulations of the facts.

### 3.3.1 Preposition stranding

In many German dialects, prepositional proforms like *damit* or *dagegen* and question forms like *womit* or *wogegen* can be used discontinuously. Both the slash analysis and the movement to the *Vorfeld* are able to describe appropriately sentences like (14b.), where the pronoun part of a prepositional proform is detached from its preposition.

- (14) a. Damit kann er diese Theorie widerlegen.  
 Therewith can he this theory defeat  
 b. Da kann er diese Theorie mit widerlegen.  
 There can he this theory with defeat  
 "With this he can defeat the theory"

Whereas the preceding data give strong evidence for movement into the *Vorfeld*, preposition stranding in German is not restricted to that position. The *da* of a discontinuous proform can also occur in the *Mittelfeld*.

- (15) a. Er kann diese Theorie damit widerlegen.  
 He can this theory therewith defeat.  
 b. Er kann da diese Theorie mit widerlegen.  
 He can there this theory with defeat.

Furthermore, it kann be combined in one sentence with another discontinuity, e. g. discontinuous *was für*.

- (16) a. Was kann er da für eine Theorie mit widerlegen.  
 What can he there for a theory with defeat  
 "What a theory can he defeat with this"  
 b. Was<sub>i</sub> kann er da<sub>j</sub> t<sub>i</sub> für eine Theorie t<sub>j</sub> mit widerlegen.

Evidently, a formalism whose slash feature allows only for one discontinuous constituent has to describe *da ... mit* by LP rules. Therefore, *da* and *mit* would have to be of the same constituent as *diese Theorie*. But in 14 (a) *damit* obviously forms one constituent. It is unclear how in a flat structure the obligatoriness of *da* can be expressed, an effect that naturally falls out of the trace approach.

### 3.3.2 The position of pronouns

The word order of personal pronouns in the German sentence is rather restricted as exemplified by the following sentences.

- (17) a. daß der Chef ihn ihr vorstellt  
 that the boss him[acc] her[dat]  
 presents  
 "that the boss presents her to him"  
 b. daß der Chef ihr den neuen Mitarbeiter vorstellt  
 that the boss her[dat] the new colleague presents  
 "that the boss presents the new colleague to her"  
 c. \*daß der Chef den neuen Mitarbeiter ihr vorstellt  
 \*that the boss the new colleague[acc] her[dat] presents  
 d. \*daß der Chef ihr ihn vorstellt  
 \*that the boss her[dat] him[acc] presents  
 e. daß ihn ihr der Chef vorstellt  
 that him[acc] her[dat] the boss presents  
 f. \*?daß ihn der Chef ihr vorstellt  
 \*that him[acc] the boss her[dat] presents  
 g. \*daß ihn ihr er vorstellt  
 \*that him[acc] her[dat] he[nom] presents

Even though it might be possible to describe these restrictions by a set of LP statements, our impression is that this kind of analysis obscures the rather simple pronoun word order. Furthermore, we cannot see how LP rules could allow for (a) and (e) while excluding (f)<sup>8</sup>. *Ihn* may precede the nominativ NP, but only if there is no dative pronoun following the latter.<sup>9</sup>

In our analysis, personal pronouns have a fixed position in the sentence either between S<sup>1</sup> and S or between S and VP.

- (18) a. [<sub>S<sup>2</sup></sub> [<sub>S<sup>1</sup></sub> daß es; ihr; [<sub>S</sub> der Mann [<sub>VP</sub> t; [<sub>V<sup>1</sup></sub> t; gib<sup>d</sup>]]]]]]  
 b. [<sub>S<sup>2</sup></sub> [<sub>S<sup>1</sup></sub> daß [<sub>S</sub> der Mann es; ihr; [<sub>VP</sub> t; [<sub>V<sup>1</sup></sub> t; gib<sup>d</sup>]]]]]]

### 3.4 ACI-constructions

In ACI-constructions however a personal pronoun may very well follow a non-nominativ NP. Compare (17c.) and (19).

<sup>8</sup> (f) seems to be somehow acceptable in some dialects but completely ungrammatical in others.

<sup>9</sup> As far as we understand it, also a solution by sequence union [Rea89] could not account for these facts.

- (19) a. Gestern hat Karl den Jungen ihr helfen lassen.

Yesterday has Karl the boy her[dat] help let  
 "Yesterday Karl has made the boy help her"  
 b. Gestern hat Karl den Jungen [<sub>VP</sub> ihr helfen] lassen.

To save an LP analysis as indicated above we would have to say that *ihr helfen* is part of another constituent as *den Jungen*, therefore the LP statements do not hold between *den Jungen* and *ihr*. The structure of the sentence (19a) might then be sketched as in (19b).

But now consider the following sentence:

- (20) Gestern hat sich Karl eine Lösung einfallen lassen.  
 Yesterday has Karl himself a solution come.to.mind let

This sentence, although it represents a regular ACI-construction and consequently must have an embedded VP constituent, shows the same serialization wrt. the pronoun as (17), where all NPs belonged to the same constituent. ID and LP rules therefore lead to a contradiction in handling (19) and (20)<sup>10</sup>. A TUG description on the other hand can make use of a scrambling analysis in (20). The pronoun *sich* leaves a trace in the pronoun position of the embedded construction and appears in the appropriate pronoun position of the matrix constituent, thereby maintaining the relevant serialization conditions in both constituents. And a further restriction may be implemented when configurational relations are available: the scrambling of a dative pronoun out of an ACI-construction as shown in (20) is sensitive as to whether the accusative of the ACI-verb replaces an external or internal nominative of the embedded verb. Finite *einfallen* marks its nominative internally, finite *helfen* marks it externally. Only internal nominative-accusatives allow being scrambled over, cf. (20) with (21).

- (21) \*Gestern hat ihr Karl den Jungen helfen lassen.  
 \*Yesterday has her[dat] Karl the boy help let

#### 3.4.1 Scrambling and Thematic Structure

Another aspect of the configurational differentiation between external and internal arguments can be made use of in analysing the thematic structure of a sentence. So e.g. the thematic differentiation between wide and narrow scope of a verbal argument depends on its appearing in marked or unmarked position. Compare the readings of (22) and (23) vs. (26), where boldface marks the posodically prominent syllable:

<sup>10</sup> A description like the one in [Rea89] would presumably have to state that verbs like *helfen* are combined in ACI-constructions by concatenation, verbs like *einfallen* by sequence union.

- (22) Ich glaube daß der Kollege dem Vorstand widersprochen hat.  
**I think that the colleague has the board contradicted**  
 "I think that the colleague has contradicted to the board."
- (23) Ich glaube daß dem Vorstand der Kollege widersprochen hat.  
**I think that the board the colleague contradicted has**  
 "I think that the colleague has contradicted to the board."

(22) is a possible answer to the question *what happened on the meeting yesterday* whereas (23) is only allowed in specific contexts such as *Wer hat dem Vorstand widersprochen?* *Widersprechen* like *helfen* marks its nominative externally, therefore the dative NP follows the nominative NP in the unmarked word order as shown in (22). If scrambling applies as in (23) this correlates with a change in the thematic and prosodic structure. In the TUG framework this would be achieved in combining the scrambling mechanism with a feature structure that indicates the desired thematic interpretation of the sentence:

- (24) `vp ---> np<trace(ana,np:trace),`  
`vp:h`  
`h:scope=narrow.`
- (25) `vp ---> np, v1`  
`np:scope=vp:scope.`

A straightforward implementation of this observation in the ID/LP format however would come to a halt in the case of (26):

- (26) Ich glaube daß dem Vorstand die Lösung eingefallen ist.  
**I think that the board the solution come.to.mind has**  
 "I think that the solution has come to mind to the board."

Although the nominative follows the dative in (26) the sentence has wide scope interpretation and unmarked prosodic structure<sup>11</sup>. (26) again is a likely answer to the general question *Was ist gestern passiert?* This clearly contradicts (22)-(23) if only the surface case marking of the arguments can be referred to in the generalization. On the other hand this result is exactly what would be expected by a TUG analysis: Since *entfallen* marks its nominative internally, no scrambling is involved in (26) vs. (23).

<sup>11</sup> John Pheby in [HFM81] postulates the distinction between marked and unmarked prosodic structure in German. [vSU86] combine this with a configurational syntax. See also [Uhm91] for a reformulation of the relevant observations in the framework of [Pie80].

## 4 Parsing with TUG

TUG can be processed by a parser and a generator. Before parsing and generation, the grammar is compiled to a more efficient form. The first compilation step that is common to generation and parsing transforms the attribute-value-pair structure to (PROLOG) term structure. This transformation makes use of the type definitions. For parsing, TUG is processed by a Tomita parser [Tom86]. For usage in that parser the result of the transformation to PROLOG further undergoes several transformations (expansion of head movement rules, transformation of argument movement rules, elimination of empty productions, conversion to LR(K) format and computation of LR tables). This compilation leads to a considerable increase in processing speed that makes it possible to use TUG for the syntactic description in real-time systems. Especially the separate compilation of head movement and argument movement leads to run time grammars that do not show the usual decrease in efficiency due to empty productions (traces). In fact, a compiled TUG does not contain empty productions any longer. Parsing time for simple sentences of about 10 words using a grammar of German with rather broad coverage is between 1 and 2 secs. on the average on a SUN SPARC 1 workstation running Quintus Prolog, even if the sentence contains verb fronting, argument movement to the *Vorfeld* and scrambling in the *Mittelfeld*.

## 5 Conclusion

We have presented Trace & Unification Grammar, a grammar formalism that tries to bridge the gap between UG and GB theory with the aim of adopting many of the linguistics descriptions of German found in the linguistics literature. Besides German, the presented grammar formalism has also been used successfully to describe a smaller subset of Chinese. We have compared TUG descriptions of some phenomena in German syntax to approaches that do not make use of movement rules but use ID/LP rules instead and shown that in all these cases TUG provides a simple and elegant description whereas the ID/LP approach in most cases even fails to describe the data at all. Furthermore we have briefly mentioned the way TUG can be compiled to an efficient parser. Empirical tests have shown that using movement rules and traces does not lead to a considerable decrease in parsing speed.

## Acknowledgements

We would like to thank Ping Peng, Manfred Gehrke, Rudi Hunze and Ludwig Schmid for many discussions on the TUG-formalism. The work described here is partly sponsored by the Federal Ministry of Research and Technology in the ASL project under no. 011V102AO.

## References

- [Blo91] Hans Ulrich Block. Compiling Trace & Unification Grammar for Parsing and Generation. In *Proceedings of the Reversible Grammar Workshop, ACL*, Berkeley, 1991.
- [Che90] H.-H. Chen. A logic-based government-binding parser for mandarin chinese. In *13th International Conference on Computational Linguistics (COLING-90)*, pages 1-6, 1990.
- [CLW88] H.-H. Chen, I.-P. Lin, and C.-P. Wu. A new design of prolog-based bottom-up parsing system with government-binding theory. In *12th International Conference on Computational Linguistics (COLING-88)*, pages 112-116, 1988.
- [Cze87] Hartmut Czepluch. Lexikalische Argumentstruktur und syntaktische Projektionen: zur Beschreibung grammatischer Relationen. *ZS*, 6(1):3-36, 1987.
- [dB84] Hans den Besten. The ergative hypothesis and free word order in Dutch and German. In Jindrich Toman, editor, *Studies in German Grammar*, pages 23-65. 1984.
- [Fan87] Gisbert Fanselow. *Konfiguralität*. Narr, Tübingen, 1987.
- [HFM81] Karl E. Heidolph, Walter Flämig, and Wolfgang Motsch. *Grundzüge einer deutschen Grammatik*. Akademie Verlag, Berlin, 1981.
- [Len77] Jürgen Lenerz. *Zur Abfolge nominaler Satzglieder im Deutschen*. Narr, Tübingen, 1977.
- [Per81] Fernando Pereira. Extraposition grammar. *Computational Linguistics*, 7:243-256, 1981.
- [Pie80] Janet Pierrehumbert. *The Phonology and Phonetics of English Intonation*. PhD thesis, M.I.T., 1980.
- [Rea89] Mike Reape. A logical treatment of semi-free word order and bounded discontinuous constituency. In *Proceedings of the 4th Conference of the European Chapter of the ACL (EACL-89)*, pages 103-110, 1989.
- [Shi84] Stuart M. Shieber. The Design of a Computer Language for Linguistic Information. In *10th International Conference on Computational Linguistics (COLING-84)*, pages 362-366, 1984.
- [Thi82] Craig Thiersch. A note on scrambling and the existence of vp. *Wiener Linguistische Gazette*, (27-28):83-95, 1982.
- [Tom86] M. Tomita. *Efficient Parsing for Natural Language: A fast Algorithm for Practical Systems*. Kluwer Academic Publishers, Boston, 1986.
- [Uhm91] Susaune Uhlmann. *Fokusphonologie*. Niemeyer, Tübingen, 1991.
- [vSU86] Armin von Stechow and Susaune Uhlmann. Some remarks on focus projection. In Werner Abraham and Sjaak de Meij, editors, *Topic, Focus, and Configurationality*, pages 295-320. John Benjamins, 1986.