GERD LAU - HANS DIETER LUTZ

AUTOMATIC ANALYSIS OF THE GERMAN NOUN GROUP AND SOME PROBLEMS

0. Introduction.

Since 1971 the research group “Maschinelle Syntaxanalyse” (MasA) has been working as a part of the project “Linguistische Datenverarbeitung” (LDV) at the “Institut für deutsche Sprache”, supported by the “Bundesminister für Forschung und Technologie” of the Federal Republic of Germany.

The task of the MasA is to elaborate a system for analysing simple sentences of present-day-German. “Simple sentences”, in this case, means any type of sentence with only one finite verb in it.

1. The scope of the analysis.

The team works on the basis of a so-called verb grammar. In this type of grammar the verb represents the nucleus of a sentence, thus being opposed to the subject-predicate-grammar. Proposals for the elaboration of such a grammar have been presented by Helbig-Schenkel,¹ Heringer² and Engel.³

This type of grammar is based on a verb lexicon, which describes all the obligatory and facultative objects that depend on the verb, e.g.

\textit{treffen} with two obligatory objects in
\begin{quote}
Ich treffe ihn morgen
\end{quote}

¹ G. HELBIG, \textit{Der Begriff der Valenz als Mittel der strukturellen Sprachbeschreibung und des Fremdsprachenunterrichts}, in \textit{Deutsch als Fremdsprache}, I (1965), pp. 10-23.
² H. J. HERINGER, \textit{Theorie der deutschen Syntax}, München 1970.
³ U. ENGEL, \textit{Bemerkungen zur Dependenzgrammatik}, in \textit{Neue Grammatiktheorien und ihre Anwendung auf das heutige Deutsch. Sprache der Gegenwart}, XX (1971), pp. 111-155.
treffen with one obligatory object and one facultative object in
Er trifft (den Ball)

treffen auf with two obligatory objects in
Die Mannschaft traf auf einen schwachen Gegner.

After having detected the number of both the obligatory and facultative objects and specified these objects according to their syntactical functions, one can group the verbs into classes, e.g. the class of all the verbs with one obligatory object (this class contains for example the verb laufen), the class of all the verbs with two obligatory objects (containing for example treffen), the class of all the verbs with three obligatory objects (containing for example anbieten), the class of all the verbs with one obligatory and one facultative object (containing for example treffen), and so on. Each of these classes represents a so-called "Satzbauplan" (SBP). 4

The systematic levels of the grammar are
a) the word class-level (WK-level)
b) the group-level (noun group (NG), verbal group (VG))
c) the complex-level (noun complex (NK), verb complex (VX))
d) the object-level (nominative object (Oo), genitive object (Og), ...).

The following graph can be regarded as an illustration of these four systematic levels:

---

4 U. Engel, Die deutschen Satzbaupläne, in "Wirkendes Wort", XX (1970), pp. 361-392.

6 For the explanation of the word classes, especially DET2B, ADJ, NOM, DET4, and PRAEP see section 2; Oo stands for "nominative object", O1 for "accusative
The way of finding the word class of a given word form, i.e. the morpho-syntactical description of a word form, is delineated in the contribution of Werner Brecht.  

2. Description of NG-relevant word classes.

The following word classes are relevant for the constitution of a NG:

- a) preposition (postposition) \textit{PRAEP}, e.g. \textit{vor}
- b) determinantia \textit{DET}, e.g. \textit{allen}
- c) adverbs \textit{ADV}, e.g. \textit{sehr}
- d) non-inflected adjectives \textit{ADJU}, e.g. \textit{gut}
- e) inflected adjectives \textit{ADJ}, e.g. \textit{sichtbaren}
- f) nouns \textit{NOM}, e.g. \textit{Objekten}
- g) pronouns \textit{PRON}, e.g. \textit{uns}.

The word classes are defined by the following morphological and/or syntactical properties:

Prepositions are collected in a catalogue and specified by their position (in front of, behind, in front of or behind, in front of and behind the nucleus of a NG), by their government (accusative, genetive, dative) and by the sort of object that is introduced by the preposition (prepositional object, situative object, directional object).

Adverbs are listed, and the list only includes words that meet the following conditions: non-inflexional; no predicative use; certain regularities of position.

Non-inflected adjectives are not listed; they are described by an algorithm, which is to find out the degree of comparison.

Inflected adjectives are not listed, either. They, too, are described...
by an algorithm with respect to case, number, gender, degree of comparison and kind of inflexion (strong, weak).

Nouns are not listed in the present state of our work. There exists an algorithm that provides the description of those words according to case, number, and gender.7

In general, the definition of the several word classes is the common one, as found in the traditional grammars. But there are two exceptions: the so-called determinantia and the pronouns.

The so-called determinantia are listed in a catalogue; they are grouped with respect to the possible arrangements of several determinantia within a NG, e.g.

\[ \text{vor allen diesen meinen schönen Büchern}. \]

The results of having found out certain regularities are both a number of distinct subclasses of determinantia and the formulation of context-sensitive rules for analysing arrangements of determinantia.

First, there are 10 subclasses, namely

- \( DET1A \) containing \( all, all-; \)
- \( DET1B \) containing \( manch, solch, welch; \)
- \( DET1C \) containing \( irgend, irgendwelch; \)
- \( DET2A \) containing \( dies-, jen-; \)
- \( DET2B \) containing \( d-, d-jenig; \)
- \( DET2C \) containing \( ein--; \)
- \( DET2D \) containing \( kein-, welch-, irgendein-, irgendwelch-, ebend-, ebend-selb-, ebendies-, d-selb; \)
- \( DET3 \) containing \( jed-, jeglich-, jedwed; \)
- \( DET4 \) containing \( mein-, dein-, sein-, uns(e)r-, eu(e)r-, ihr--; \)
- \( DET5 \) containing \( einig-, etlich-, manch-, mehrer-, ein paar. \)

Secondly, there are 5 rules for describing the regularities of determinantia arrangements within a NG having a noun as nucleus. For

7 In future it will be possible to incorporate a more effective algorithm which reverts to an almost complete lexicon of German nouns. This lexicon has been placed at our disposal by the research group for “Automatische Lemmatisierung” at the University of Saarbrücken, W-Germany.

8 The subclassification of the determinantia is a modification of the one given in ENGL, Regeln zur Wortstellung, in Forschungsberichte des Instituts für deutsche Sprache, Mannheim, V (1969), pp. 9-148, esp. 102ff.

9 The catalogue contains all word forms of the listed stems (indicated by “-”).
the demonstration of these rules we will use a simplified NG-structure with an abbreviated notation.

\[
\begin{align*}
R1: & \ \ (DET1A) \ (DET2A) \ (DET4) \ (ADJ) \ NOM \rightarrow NG \\
R2: & \ \ (DET1A) \ (DET2B) \ (ADJ) \ NOM \rightarrow NG \\
R3: & \ \ (DET1B) \ (DET2C) \ (ADJ) \ NOM \rightarrow NG \\
R4: & \ \ (DET1C)/(DET2D)/(DET5) \ (ADJ) \ NOM \rightarrow NG \\
R5: & \ \ (DET2C) \ (DET3) \ (ADJ) \ NOM \rightarrow NG.^{10}
\end{align*}
\]

It is important that there can be no iteration of a subclass of determinantia.

The rules mentioned above can be illustrated by the diagram on the next page.

The pronouns are considered as pro-forms of a NG with a nominal nucleus. A part of the pronouns can be grouped in several subclasses according to three significant features: possibility of using a preposition in front of the pronoun, possibility of putting determinantia in front of the pronoun, possibility of using a genitive or prepositional attribute in the post-nuclear field.

There are 9 subclasses, namely

- **PRON1** containing man;
- **PRON2** containing du, er, es, ich, ihr, sie, wir;
- **PRON3** containing dein(er), dich, dir, einander, euch, euer, ihr(er), ihm, ihm(en), mein(er), mich, mir, sein(er), sich, sie, uns(er);
- **PRON4** containing all-, ebend-, ebend-selb, ebend-selb, einig-, ein paar, etlich-, etwas, irgendbein-, irgend-ebed, irgendwelch-, jeder-, jemand-, kein-, manch-, mehrer-, niemand-, nichts, w-, welch-;
- **PRON5A** containing d-, d-selb-, d-jenig-, dies-, jen-;
- **PRON5B** containing ein-;
- **PRON5C** containing jed-, jedwed-, jeglich-;
- **PRON5D** containing dein-, ev(e)r-, ihr-, mein-, sein-, uns(e)r-;
- **PRON6** containing selber, selbst.

---

\(^{10}\) The brackets stand for "facultative"; the slash stands for "exclusive or". A full description of the structure of a noun group with a nominal nucleus will be given in section 5.
| class | DET1A | DET1B | DET1C | DET2A | DET2B | DET2C | DET2D | DET3 | DET4 | DET5 | ADJ | NOM |
|-------|-------|-------|-------|-------|-------|-------|-------|------|------|------|-----|-----|
| R1:   | alle  |       | diese |       |       |       |       |      |      | treuen | Hund |     |
| R2:   | all   |       | die   |       |       |       |       | meine |      | treuen | Hund |     |
| R3:   | manch |       |       |       |       |       |       | treuer | Hund |       |     |     |
| R4:   | irgendwelch | |       |       |       |       |       | ebendiese | treuen | Hund |     |     |
| R5:   |       |       |       |       | ein   |       |       | einige | treue | Hund |     |     |
| rule  | DET1A | DET1B | DET1C | DET2A | DET2B | DET2C | DET2D | DET3 | DET4 | DET5 | ADJ | NOM |
But, there is another group of pronouns without the characteristics mentioned above, i.e. without pre-nuclear field, without a preposition in front of them, without post-nuclear field. These pronouns do not differ at the NG- or NK-level. But they differ at the O-level, when you try to detect if such a pronoun is a substitute for an object (depending on the verb) or a substitute for a so-called "Angabe" (which does not depend on a verb). Furthermore, pronouns that can be substitutes for an object, can be differentiated according to the kind of these objects.

So, there are six more subclasses, namely

- **PRON7A** containing *wobei, wohin, womit, wonach, woraus, worum, wovon*;
- **PRON7B** containing *wo, wohinter, woneben*;
- **PRON7C** containing *woher, wohinaus, wohinein, wohinüber, wohin*;
- **PRON7D** containing *woran, worauf, worin, wohinter, worunter*;
- **PRON7E** containing *wogegegen*;
- **PRON8** containing *inwiefern, inwieweit, wann, warum, weshalb, weshalb, wie, wieso, wodurch*.

Pronouns of **PRON7A** are substitutes for prepositional objects (O4); pronouns of **PRON7B** are substitutes for situative objects (O5); pronouns of **PRON7C** are substitutes for directional objects (O6); pronouns of **PRON7D** are substitutes for O4 or O5; pronouns of **PRON7E** are substitutes for so-called "Angaben".

The final division of pronouns in 15 subclasses makes it possible to set up rules for detecting and describing NGs with a pronominal nucleus. For the demonstration of these rules we will use a simplified and abbreviated notation.

\[
\begin{align*}
R6: & \text{ PRON1 } / \text{ PRON2 } \rightarrow \text{ NG}_0 \\
R7: & \text{ PRON6 } \rightarrow \text{ NG}_2 \rightarrow \text{ NG}_{\text{all}} \\
R8: & \text{ (PRAEP) } \{\text{PRON3 } / \text{ PRON4}\} \rightarrow \text{ NG}_2 \\
R9: & \text{ (PRAEP) } \{\{(DET1A) \text{ PRON5A}\} / \{(DET1B) \text{ PRON5B}\} / \{(DET1C) \text{ PRON5C}\} / \{(DET1A) \text{ DET2A}\} / \{(DET2B) \text{ PRON5D}\}\} \rightarrow \text{ NG}_2
\end{align*}
\]

The reason for these pronouns not being included in **PRON1** (which meets the same conditions at the first glance) is that **PRON1** covers always the position of a nominative object (O0), in contrast to the pronouns of **PRON7A** to **PRON8**.
This collection of rules shows the difference between PRO1 ... PRO6 and PRO7A ... PRO8. A comparison with the rules R1 to R5 illustrates that there are less possibilities to combine determinanlia in front of the nucleus of a pronominal NG. Other differences will be discussed in section 3.

3. Two types of NGs.

It is no new observation that there are at least two different types of NGs in German.

The first type covers all NGs having a noun as nucleus of the NG, e.g. die Mitbestimmung,

the second type covers all NGs with a pronoun as nucleus of the NG, e.g. sie, die unsere, woran.

The structure of the NG depends on what kind of nucleus appears.

In a NG with a pronominal nucleus, there can be no adverbs, no non-inflected adjectives, and no adjective – in opposition to a NG with a noun as a nucleus.

In a pronominal NG there are less possibilities of arranging determinantia than in a nominal NG (see rules R1-R15). The reason is quite evident: by suppressing the noun (nucleus) in a nominal NG, the last determinans turns into a pronoun, i.e. it represents the nucleus in a pronominal NG. Compare the following two examples:

\[
\begin{array}{c}
\text{alle diese Bücher} \\
\text{DET1A DET2A NOM} \\
\uparrow \\
\text{nucleus}
\end{array}
\quad vs. \quad 
\begin{array}{c}
\text{alle diese} \\
\text{DET1A PRON5A} \\
\uparrow \\
\text{nucleus}
\end{array}
\]

12 The "?" indicates that the case of the NG is not predictable in general. "A" is an abbreviation of "Angabe". The braces are to denote larger units belonging to each other. For the explanation of other signs see footnote 10.
There is another difference between a nominal NG and a pronominal one: in nominal NGs one can find other NGs embedded:

*die Ansprachen haltenden Parlamentarier*, or (rather complex)
*die nach Marx alle menschlichen Handlungen wesentlich und zu aller Zeit zum Nachteil der Mehrheit der Bevölkerung bestimmenden Kapitalgesetze*.

In both cases the first *die* belongs to the last word, to *Parlamentarier* and *Kapitalgesetze*; *die Parlamentarier* and *die Kapitalgesetze* form the hierarchically highest NG. If you transform the determinants of this highest NG into a pronoun such as *sie*, there is no possibility of embedding other NGs in this pronominal NG. This explains the difference between a nominal NG and a pronominal one, too.

4. *Criteria for the analysis.*

The different elements of a German NG underlie different kinds of syntagmatic relations. The number of these relations varies corresponding to the two types of NGs.

Four kinds of relations are relevant for the nominal NG:

- congruity between adjectives and nouns (concerning case, number, and gender)
- between determinants and nouns (concerning case, number, and gender)
- between prepositions and nouns (concerning the case of a noun and the possible government of a preposition)
- between determinants and adjectives (concerning case, number, and gender);
- inflectional relation between determinants and adjectives (concerning the inflexion of a determinant and the inflexion of a following adjective);
- degree of comparison of the adjective of the nominal NG (providing the connection with further parts of the syntactical analysis);
- compatibility of subclasses of determinants according to the rules R1 to R5; the rule of compatibility of nominal NG has the form:
Only two kinds of relations are relevant for the pronominal NG: congruity between determinantia and pronouns (concerning case, number, and gender) between prepositions and pronouns (concerning the case of a pronoun and the possible government of a preposition); compatibility of subclasses of determinantia according to the rules R6 to R9; the rule of compatibility of a pronominal NG has the form:

\[ ((det = 1) \Rightarrow ((M(WK, I, n-i) = DET1A \land (M(WK, I, n-i + 1) = DET2A \lor DET2B)) \lor \]
\[ ((M(WK, I, n-i) = DET1B \land (M(WK, I, n-i + 1) = DET2C)) \lor \]
\[ ((det = 2) \Rightarrow (M(WK, I, n-i) = DET1A \land (M(WK, I, n-i + 1) = DET2A \lor DET2B))).^{13} \]

Summing up, there are only two kinds of relations of the pronominal NG, but four kinds of relations of the nominal NG; the set of the various congruity relations noticeable at the nominal NG differs from the one noticeable at the pronominal NG; and the compatibility

---

13 This notation is a conglomerate of matrix-notation, logical signs, and of programmer's conventions. It can be read: "If the value of the counter "det" equals 1, then the value of the place "WK" of the matrix "M" at the level "I" at the position "n-i" must have been "DET1A" and the value of the place "WK" of the matrix "M" at the level "I" at the position "n-i+1" must have been "DET2A or DET2B or DET4, or the value of..." and so on.

M is a three-dimensional matrix which gets built up during the morphological analysis of an input sentence, and which contains all computed information of all the words of this sentence. "I" stands for "interpretation" and is a variable, running from 1 to 50. At the moment, "WK" stands for "word class" with 40 different specifications. "n" indicates the position of the noun in a given sentence, and "i" is a position counter; so, if n = 3 and i = 1, then n - i = 2, i.e. the position in question is the next to the noun on the left side.
rule for the nominal NG is much more complex than that for the pronominal NG.

5. An algorithm as representation of a hyperrule.

The preceding outlines displayed the complexity of contextual restrictions concerning German noun groups. It sounds extremely clumsy to talk about these restrictions in a natural language. We therefore developed a formal notation for the communication between linguists and programmers of our team. It consists of a matrix-notation, sentential connectives, functions and flow chart conventions. The compatibility rules were a small cut-out. For more details see our forthcoming publication.14

In the following, we will discuss the nominal NG only, because its structure and its difficulties are more interesting than the structure of the pronominal NG.

At the present state of our algorithm the following constructions have been excluded: coordination of adjectives, fusion of prepositions and certain determinants (i.e. am being an + dem), postponed prepositions, prepositions like um ... willen, and embedded participle constructions.

The reason for excluding embedded participle constructions is that, in our opinion, complex constructions like these can be analyzed after the successful interpretation of simple sentences only. The other restrictions mentioned above are not so important, i.e. the solution of these problems are, more or less, not so difficult, and, in fact, one of our colleagues has made a proposal for the treatment of all kinds of prepositions.

At the moment, we will consider the following hyperrule and we will interpret it as a statement about the word order only:

(1) \[(PRAEP) (DET) (DET) (DET) (((ADV*) ADJU*) ADJ)* NOM \rightarrow NG.\]15

14 ARBETSGRUPPE MASA, Zur maschinellen Syntaxanalyse I. Morphosyntaktische Voraussetzungen für eine maschinelle Sprachanalyse des Deutschen, in Forschungsberichte des Instituts für deutsche Sprache, Mannheim, XVIII, 1, XVIII, 2, 1974.

15 The asterisk indicates the possibility of iterating. For the explanation of the other signs see footnote 10.
The program that is intended to analyze nominal NGs first looks for a NOM, proceeding from the left to the right. After having found a NOM, the whole algorithm will be called up in order to search an ADJ to the left of the NOM. ADV* and ADJU* can appear only in case an ADJ has been found. Otherwise the algorithm turns to DET, which tests the word order and compatibility of the DETs if there exists a DET. There are no more than three DETs admitted. PRAEP is the word class which stands at the extreme left in German NGs (according to our restrictions); it is the last one to be looked for.

In many cases the output of this analysis depends on homography. An example may prove this as well as the fact that there remain lots of ambiguities at this level of analysis:

\begin{itemize}
  \item [description 1] nahe so sehr gründlich vergifteten Wäldern
  \item [description 2] \textit{PRAEP ADV ADV ADJU ADJ NOM}
  \item [description 3] \textit{VERB ADV ADV ADJU ADJ NOM}
  \item [description 4] \textit{VERB ADV ADV ADJU ADJ NOM}
  \item [description 5] \textit{VERB ADV ADV ADJU ADJ NOM}
  \item [description 6] \textit{VERB ADV ADV ADJU VERB NOM}
\end{itemize}

\begin{itemize}
  \item [description 1] vor so sehr gründlich vergifteten Wäldern
  \item [description 2] \textit{PRAEP ADV ADV ADJU ADJ NOM}
\end{itemize}

If PRAEP is a homograph, ADV* and ADJU* may or may not be constituents of the nominal NG (see (2)). If PRAEP is no homograph and if there is no DET, the whole string must be a NG (see (3)). As homographs may exist, the algorithm is constructed in the following way: in a first step the longest nominal NG possible is built up, and in a second step this nominal NG is minimized in all possible interpretations (in the case of (2) there are five interpretations which cover shorter nominal NGs than the first one).
6. Problems concerning the identification of a "nominal complex" (consisting of a NG or of a number of NGs).

A nominal complex NK is a noun group or a string of noun groups which are part of a sentence. Yet, this gives no information about the role the NK plays in the sentence: a NK can be either an object which depends on the finite verb, or an "Angabe". NKs in the nominative case can serve as objects only, i.e. as subjects of sentences. The following schedule displays that any NK with an O-value between 1 and 6 can be either an object or an "Angabe":

| O | Object | "Angabe" |
|---|--------|----------|
| 1 | Er wiegt zwei Kilo | Er läuft zwei Meter |
| 2 | Er gedenkt des Morgens | Er läuft des Morgens |
| 3 | Er gibt mir Wein | Er steigt mir auf den Fuß |
|   |                    | Er trägt mir den Koffer |
| 4 | Er rechnet mit dem Vortrag des Bruders | Er verdient mit dem Vortrag des Bruders |
| 5 | Er wohnt im ersten Stock | Er schläft im ersten Stock |
| 6 | Er geht ins Kino hinein | Er schläft in den Tag hinein |

Thus we cannot derive any knowledge about a NK's function from its grammatical description only. But semantical criteria can provide some solutions at this level of analysis:

(4) Er rechnet mit dem Vortrag seines Bruders.

(4) contains the verb rechnen mit + 1 with the "Satzbauplan" SBP Ø4. But (4) contains also the verb rechnen with SBP Ø. As the semantic features of Vortrag tell us that Vortrag is no means of calculating, the SBP Ø will be excluded as an incorrect derivation.
(5) Er rechnet mit der Rechenmaschine.

(5) contains the same verbs as (4). As Rechenmaschine is a means of calculating, the NK is also an "Angabe". Therefore both SBP $\emptyset$ and SBP $\emptyset 4$ are possible analyses.

But we do not yet follow semantics. We try to analyze with syntactic means as long as possible.

If we combine two NGs in order to build a NK, the right NG serves as an attribute:

\[ \text{der Vertrag seines Bruders} \]
\[ \text{NG} \quad \text{NG} \]
\[ \text{NK} \]

An attribute is a NG with one of the following O-values:

- $O = 2$ .... genitive phrase: *der Vertrag seines Bruders*
- $O = 4$ .... prepositional phrase: *der Mann mit dem Kropf*
- $O = 5$ .... situative phrase: *der Baum im Garten*
- $O = 6$ .... directional phrase: *die Fahrt in die Stadt*

The rules RR1 and RR2 define an algorithm which is to build NKs of NG-combinations. Further on, RR3 and RR5 change a NK into an object, and RR4 an NK into an "Angabe".

- $L_i$ ... leftmost position of a NG, or a NK
- $R_i$ ... rightmost position of a NG, or a NK
- $O_i$ ... value of the "object class" of a NG, or NK

**RR1:** \[ \text{NG}_i (L_i, R_i, O_i) = \text{NK}_j (L_j, R_j, O_j) \]

**RR2:** \[ \text{NG}_i (L_i, R_i, O_i) + \text{NK}_j (L_j, R_j, O_j) = \text{NK}_x (L_x, R_x, O_x) \]

on condition that 1) $R_i + 1 = L_j$;
2) $O_j$ = either 2 or 4 or 5 or 6;
3) NG does not contain a preposition behind the noun.

**RR3:** \[ \text{NK}_j (L_j, R_j, O_j) = O_j (L_j, R_j, O_j) \]

(i.e. a derivation of an object.)
RR4: \( NK_j (L_j, R_j, O_j) = A_j (L_j, R_j, O_j) \) (i.e. a derivation of an "Angabe".)

RR5: \( NK_j (L_j, R_j, \emptyset) = O_j (L_j, R_j, \emptyset) \) (i.e. the derivation of the subject.)

Consider an example of a possible application of RR1-RR5 in order to combine three NGs and thus make a single NK:

\[
\begin{array}{c}
die \text{ Garage} \\
| \text{ vor dem Haus} \\
\text{NG (3, 5, 5)} \quad \text{NG (1, 2, \emptyset)} \\
\text{NG (6, 7, 2)} \\
\text{NK (3, 7, 2)} \\
\text{NK (1, 7, \emptyset)} \\
\end{array}
\]

\[
\begin{array}{c}
des \text{ Freundes} \\
\end{array}
\]

\text{Exceptional case I: Prepositions in front of, and behind, a NG or NK.}

\( (6) \quad \text{um des Freundes willen} \)

\( (7) \quad \text{um des Freundes in Paris willen} \)

Evidently (7) is no string of two NGs. Therefore, RR1 and RR2 are not applicable to this type of NK, and we must construct a special algorithm for the analysis of it.

\text{Exceptional case II: Preposition behind a NG or a NK.}

In German there are some prepositions which can appear in front of, or behind, a NG, for instance nach, wegen, ...

\text{Er zittert wegen des Kampfes oder Er zittert des Kampfes wegen.}

When we consider (8), (9) and (10) we see that a postposition in a NG at the same time is a postposition of a NK.
We need a special algorithm for the analysis of NKs of the type (8)-(10). It must provide, for instance, the following analyses of (11):

(11) *Er entsinnt sich des Beifalls wegen des Gesanges.*

| Analysis a) | NG(4,5,2) | NG(6,8,4) |
| Analysis b) | NG(4,6,4) | NG(7,8,2) |

It is possible to unite both NGs of a) into one NK by application of RR1 and RR2. But in b) this is impossible; condition 3. of RR2 is not met. Therefore both NGs in b) must be changed into separate NKs by RR1.

Another quite complicated example:

(12) *Er entledigt sich des Freundes des Boxens wegen des Messers*

| Analysis a) | NG(4,5,2) | NG(6,7,2) | NG(8,10,4) |
| Analysis b) | NG(4,8,2) | NK(9,10,2) |

RR1-RR5 and the special algorithms for exceptional case I and II are the theoretical basis to find all possible combinations of NGs and to derive objects and “Angaben” from NKs.

Next we check the syntax for useful hints so as to achieve a correct analysis of NKs of a given sentence, and hints so as to decide, which NKs are objects and which are “Angaben.”
In (13) and (14) we find NGs of the same structure. But the correct analysis of (13) differs completely from that of (14).

(13) weil der Mann mit der Verletzung der Elle im linken Arm schläft

NK (2, 11, Ø)

O (2, 11, Ø)

(14) weil der Mann mit der Frau des Morgens im alten Cafe tanzt

NK (2, 3, Ø) NK (4, 6, 4) NK (7, 8, 2) NK (9, 11, 5)

O (2, 3, Ø) A (4, 6, 4) A (7, 8, 2) A (9, 11, 5)

As long as we do not find any hints, the rules must be applied in a systematic order, so as to have all possible analyses as a result. A semantic component then can choose the subset of reasonable analyses out of the whole set. Thus the analysis of (13) would have to be an analysis of (14), and vice versa.

It would need too much space to print the set of theoretically possible analyses of (15) here. The reader easily sees, in how many different ways RR1-RR5 can be applied. The whole set of possible analyses is greater than 100! We just show a single analysis. We note the rule being applied in the leftmost column. Now we try to use the "Satzbauplan" as a means of improving the exhaustive analysis-procedure. In (15) wartete is the single finite verb. It has the SBP 0 4 auf + 1. That means, it takes both an object in the nominative case and a prepositional object. Having found a certain SBP, our algorithm runs as follows:

1. Look for a NG ( , , Ø) and make it an object. (According to RR1, RR5.)
2. Look for a NG ( , , 4) which contains the preposition auf, and make it an object. (According to RR1, RR3.)
3. Apply RR1, RR2 and RR4 to the NGs which are not objects.
4. Try to analyze one of the NGs which are not objects as an attribute of an object. If this is possible, restart at 3.; if not, stop!
| (15) | weil der Mann | mit der Frau | des Morgens | in der Küche | auf den Bediensteten | wartete |
|------|--------------|-------------|-------------|--------------|---------------------|---------|
|      | $NG_3(2,3,0)$ | $NG_4(4,6,4)$ | $NG_4(7,8,2)$ | $NG_4(9,11,5)$ | $NG_4(12,14,4)$ | $NG_4(12,14,6)$ |
| RR1($NG_3$) | $NK_3(2,3,0)$ |             |             |              |                     |         |
| RR1($NG_4$) |             | $NK_4(4,6,4)$ |             |              |                     |         |
| RR1($NG_3$) |             |             | $NK_3(7,8,2)$ |              |                     |         |
| RR1($NG_4$) |             |             |             | $NK_4(9,11,5)$ |                     |         |
| RR1($NG_4$) |             |             |             |              | $NK_4(12,13,4)$ |         |
| RR1($NG_4$) |             |             |             |              |                     |         |
| RR5($NK_3$) | $O_3(2,3,0)$ |             |             |              |                     |         |
| RR4($NK_3$) |             | $A(4,6,4)$ |             |              |                     |         |
| RR4($NK_3$) |             |             | $A(7,8,2)$ |              |                     |         |
| RR4($NK_3$) |             |             |             | $A(9,11,5)$ |                     |         |
| RR3($NK_3$) |             |             |             |              | $O(12,14,4)$ |         |
(15) weil der Mann
mit der Frau des Morgens in der Küche auf den Bediensteten wartete

O(2,3,0) A(4,6,4) A(7,8,2) A(9,11,5) O(12,14,4)

mit dem Kropf

O(2,6,0) A(7,8,2) A(9,11,5) O(12,14,4)

mit dem Hut meines Vaters

O(2,8,0) A(9,11,5) O(12,14,4)

meines Bruders in der Schweiz

O(2,11,6) O(12,14,4)

mit dem Besitzer des Hauses

O(2,3,0) A(4,8,4) A(9,11,5) O(12,14,4)

O(2,11,6) O(12,14,4)

O(2,3,0) A(4,11,4) O(12,14,4)

O(2,3,0) A(4,6,4) A(7,11,2) O(12,14,4)

O(2,6,0) A(7,11,2) O(12,14,4)

(The last two analyses are not fine semantically, but maybe somebody else can find examples, for which these analyses are correct.)
In the case of (15), this algorithm produces no more than 8 analyses. That means that the unshortened result of more than 100 analyses can be reduced to 8 by using an improved version of analysis. This shows that the SBP gives us the syntactical information which is necessary to make the correct analysis.

On the opposite page we show that all 8 analyses are possible with the SBP $\emptyset 4 \text{ auf } +1$ and the structure of NGs as in (15).

Sometimes one can find more than one finite verb in a sentence:

\begin{equation}
\text{Er liebe bestimmte nahe Wälder}
\end{equation}

\begin{tabular}{c|c|c|c}
verb & 1st position & 2nd position & end \\
\hline
liebe & - & + & - \\
bestimmte & - & + & - \\
nahe & - & + & - \\
\end{tabular}

In these cases one has to find out which of the verbs is the actual finite verb. In German the finite verb can appear at the first position in the sentence, at the second position, and at the end. It is easy to fill up the following matrix according to the verbs in (16):

Neither verb$_1$, verb$_2$ or verb$_3$ is standing at the first position or at the end of the sentence. From that we can conclude that each of them must be at the second position of the sentence. We still know another criterion to find the right verb: The string in front of it, at position 1, must be an NK. In (16) only verb$_1$ meets this condition. Therefore liebe is the finite verb.

Only if one knows a lot of syntactical heuristics like these, one is able to decide at which level of sentence-analysis semantics should be employed. We think there will be need for an interaction between a syntactical and semantical component.