Session 6:    SYNTAX


SYNTAX OF THE GERMAN NOUN PHRASE

Joseph R.  Applegate

Massachusetts Institute of Technology[1]


It is generally agreed that a successful mechanical translation routine must be based on an accurate grammatical description of both the source and target languages.   Furthermore,  the description should be presented in a form that can easily be adapted for computer programming.    In this paper,  an attempt will be made to present a brief syntactic description of the noun phrase in German.    Noun phrase here is defined as that sentence constituent which together with the verb phrase comprises the sentence.    Its case is always nominative.    In traditional terminology,   it is "the subject of the sentence".    Some of the descriptive statements presented in this paper will be applicable to other phrases containing nouns,   those that serve as objects of verbs for example,   but these or those noun phrases which occur after prepositions will not be considered here.

The model chosen for the description is that outlined by Chomsky in Syntactic Structures.  Such a grammar is constructed on the assumption that the sentences of a language can be described in terms of the ordered set of rules necessary for their production.   Such a grammar makes it possible to give a detailed description of the sentences of a language in rather precise terms.   A grammar of this type is especially useful for mechanical translation,  for it provides the basis not only for the preparation of a routine for the construction of sentences in the target language,   but also for the preparation of a routine for sentence recognition in the source language.    This point has already been discussed by Dr. Matthews in a preceding session.[2]

The data on which the description is based were collected in the following way.    An initial study was based on material   obtained from

---

[2] The Use of Grammars Within the Mechanical Translation Routine.

<u>A Grammar of the German Language</u> by George O.  Curme.    To test
the acceptability of sentences constructed according to rules obtained
from that grammar,   the sentences were submitted to native inform-
ants.    On the basis of their responses,   a new set of rules was formu-
lated.    Utterances were constructed according to this new set of
rules and again tested with native informants.    In addition,   text analy-
sis was used.    The texts included scientific articles and German news-
paper texts,  and they were searched to see whether patterns described
by the set of rules occurred as well as to determine whether any
patterns not described by rules also occurred.

     After an analysis of the data,   the following general statements
can be made about the noun-phrase structure.    There are four major
constituents in the noun phrase:    the head,   the determiner,   the
numeral and the modifier.    It has been suggested that the numeral and
the  prenominal   modifier,   if they occur,  may be added to the deter-
miner.    This question has not yet been resolved,  however.    Any of the
constituents may be omitted (or occur as  Ø).    The occurrence or non-
occurrence of a particular constituent as well as the form of a specific
constituent are affected by factors outside of the noun phrase as well
as by factors within the phrase itself.    An obvious example of this is
the fact that a noun phrase containing an inanimate noun does not occur
with certain verbs:  <u>sprechen</u>,   <u>sehen</u>,   <u>hören.</u>    A less obvious illustra-
tion of this is found in the distribution of <u>derjenige</u> which occurs with
<u>sein</u>  (or similar verbs).    The rules in our description should be for-
mulated in such a way that these relationships among the various
constituents of the noun phrase,   as well as their dependence on factors
either inside or outside of the noun phrase itself,   become clear.

     Each rule in the set consists of several subrules which represent
optional choices.    In some cases,   however,   the choices are restricted
by some choice made in a preceding rule.    The rules should therefore
be arranged so that rules that restrict a large number of subsequent
choices are placed at the beginning.    By applying the rules in the order
specified by the grammar,   it is possible to produce all of the noun
phrases in the language.    A complete grammar would,   of course,   in-
clude morphological rules,   and a dictionary from which the appropriate
lexical items could be chosen.

Session 6:   SYNTAX

Of the four constituents,   the head is the most important,  for it controls the selection of other constituents to a greater degree than any other.   The choice of head also determines the gender,   number, and person which are basic features of the noun phrase.   The head may be one of several types,   and it is these types that serve as the basis for the classification of noun phrases in the Appendix.   The first type of noun phrase has as its head a noun stem; the second,  a personal pronoun.   In the third type,  a clause serves as head.   The fourth type has an indefinite.   There are similarities among all four of these.   The fifth type is,  of course,   different,   and it may be necessary to consider it separately.

From the chart given in the Appendix,  we can easily see the effect of the choice of the noun-phrase head on the other constituents of the phrase.   If a head of Type I  is chosen,   a choice must be made as to whether the noun stem is to be considered a proper noun or a common noun.    This decision is not necessary if the head chosen is a personal pronoun,   a clause,   or the indefinite.    The rule immediately following the selection of a noun-phrase head would handle this choice.

If the head chosen is of Type II,  a personal pronoun,  a decision must be made as to the person of the phrase.    The phrase may be either first,   second,  or third person.   If one of the other heads is chosen,  however,  there is no alternative; the phrase must be third person.

After the choice of person,  the selection of the gender and num-ber of the phrase must be determined.   Gender and number are com-bined in a four-part system,   so that plural number becomes the fourth gender.   For noun stems and third-person pronouns,  any of the four genders may be chosen.   For first- and second-person pronouns,  the choice is restricted to one of three (masculine,  feminine,   or plural), and for clauses the choice is restricted to two.   If the clause is per-sonal,  it must be masculine; if it is impersonal,  it must be neuter.

Following the selection of gender,  the determination must be made of whether the noun phrase is to be modified or not.   The general form of the modifier is considered to be the relative clause. The relative clause may be abbreviated to give the postnominal adjectives,   prepositional phrases,   or appositives if the head of the noun phrase is a noun stem or a personal pronoun.   If the head is a

282

personal clause,  there can be no modifier.    If,   on the other hand,
the head is an impersonal clause,   the modifier may be a relative
clause of a specific form (introduced by <u>was</u>) but it may not be
abbreviated.    For the indefinite head,   there can be no modifier.    A
further set of rules makes it possible to change the relative clause
used as the modifier of a common noun to a prenominal modifier.
This set of rules cannot be applied unless the head is a common noun,
however.

Finally the choice of determiner is made.    This choice is made
partly on the basis of the previous choices.    The determiners can be
divided into several classes,  and the choice of a specific class de-
pends on whether the noun is singular or plural,   modified or un-
modified,   used with the verb sein or an active verb.    Other factors
affecting the choice of determiners are the type of noun used as head,
and whether the noun is to be deleted or not; that is,   certain deter-
miners may occur as pronouns in specific sentences while others may
not.    If the head of the phrase is not a common noun,   no determiner
can occur.

By examining the chart,   we can see that there are similarities
among types IB,  II,  III,   and IV.    These similarities are most marked
in the selection of modifiers and determiners.    It is possible that
these similarities of structure have semantic significance.    The
similarities may be recognized as similarities in the set of rules
chosen for the construction of these phrases.    A recognition routine
based on the analysis of sentences in terms of the set of rules needed
to construct them would,   of course,   indicate these structural similari-
ties quite clearly.    It is at this point that additional analysis is
necessary.    In order to obtain a complete specifier for a given sen-
tence,   it will be necessary to indicate not only the particular set of
rules used in the generation of a sentence,   but also a statement of the
factors that determined the choice of that particular set of rules.    It
should be possible to find similar semantic factors operating whenever
there is close similarity in the structure of the sentences produced.

Noun-Phrase Types

I.
    A.   <u>Der Mann</u> geht nach Hause.

    B.   <u>Herr Schmidt</u> geht nach Hause.

II.       <u>Er</u>  geht nach Hause.

III.     <u>Dass die Neutronenintensität auf das   6 1/2-fache des Normal-
wertes anstieg</u>,    ergab   sich hierbei.
 <u>Was er macht,</u>    ist  mir   unbekannt.
 <u>Wer das getan hat,</u>   ist mir unbekannt.

IV.     <u>Man</u> muss hier noch erwähnen,   dass die englische Besetzung
der Kanalzone fünf Wochen vorher beendet wurde.

V.      <u>Es</u>  wird  getanzt.

|  | TYPE I A | TYPE I B | TYPE II 1st | TYPE II 2nd | TYPE II 3rd | III | III | IV | V |
|---|---|---|---|---|---|---|---|---|---|
| Person | 3rd | 3rd | 1st | 2nd | 3rd | 3rd | 3rd | 3rd | 3rd |
| Gender-Number | 1,2,3,4 | 1,2,3,4 | 1,2,4 | 1,2,4 | 1,2,3,4 | 1 | 3 | 3,1 | - |
| Modifier |  |  |  |  |  |  |  |  |  |
| a. Clause | + | + | + | + | + | - | + | + | - |
| b. Abbreviated Clause | + | + | + | + | + | - | - | - | - |
| c. Prenominal | + | - | - | - | - | - | - | - | - |
| Determiner | + | - | - | - | - | - | - | - | - |

A = common noun.
B = proper noun.
+ = optionally present.
- = does not occur.

285