

Structure Of Noun Phrases In German

W. P. Lehmann
University of Texas

Analyzing German materials with a view to mechanical translation in accordance with current linguistic practices, we find the noun phrase as one of the commonest constituents. The noun phrase is an immediate constituent of most subject-predicate sentences, and frequently an immediate constituent of a predicate as well, as may be illustrated with a fairly frequent type of sentence (1):¹

Jede sprachliche Veränderung und mithin auch die Entstehung jeder dialektischen Eigentümlichkeit hat ihre besondere Geschichte. 'Every linguistic change and consequently also the origin of every dialectal peculiarity has its special history.'

By our first analysis of this sentence we determine one immediate constituent, which consists of the material preceding *hat*, to be a compound noun phrase. (Since the position of the finite verb is rigidly circumscribed, analysis of most German clauses into two immediate constituents is a simple matter.) By our second analysis, we determine another noun phrase, the material after *hat*. While analysis and translation into English of this sentence should cause no difficulty for man or machine, those who have consulted materials written in technical German are aware that interpretation of noun phrases occasions their greatest difficulty for man; one might therefore expect that it would also occasion difficult problems in machine translation. Principles involved in its analysis will be discussed here as well as its structure. For the structure of noun phrases can fairly readily be summarized.

We may classify noun phrases in four groups: I. those consisting of nouns preceded by modifiers; II. those consisting of adjectives functioning as nouns, preceded by modifiers; III. those consisting of nouns or adjectives, that is, phrases of types I and II, followed

¹ Among materials analyzed for this sketch were sections from H. Paul's *Prinzipien der Sprachgeschichte* (5th ed. Halle, 1920), N. Trubetzkoy's *Grundzüge der Phonologie* (Prague, 1939) and Helmut Rehder's "Karl Jaspers" in *Philosophen des 20. Jahrhunderts* (Stuttgart, 1957). German examples cited may be located in these texts: (1) Paul, p. 42; (2) Rehder, p. 739; (3) Trubetzkoy, p. 15; (5) Paul, p. 8; (6) Paul, p. 14.

by modifiers; IV. those consisting of phrases of types I and II accompanied by positionals, usually prepositionals. The following is a brief summary of this classification, with examples.²

I. Nouns preceded by modifiers

A. ϕ	No			Bier	'beer'		
B. Aj	No	a. Li	No a. DE	No	das Bier	'the beer'	
				β . KE	No	sein Bier	'his beer'
				γ . Adj	No	Pauls Bier	'Paul's beer'
		b. De	No			dunkles Bier	'dark beer'
			(De De ...)	(Cnj)		De No kaltes dunkles Bier	'cold dark beer'
C. Li	Aj	No	[Amplification of I.B]				
	a. DE	Aj	No			das dunkle Bier	'the dark beer'
	b. KE	Aj	No			sein dunkles Bier	'his dark beer'
		c. ϕ				recht dunkles Bier	'quite dark beer'
			DE	Aj	No	das recht dunkle Bier	'the quite dark beer'
				KE		sein recht dunkles Bier	'his quite dark beer'
							(..) ((Adv) (Adv)De) ((Adv) (Adv)De) (Cnj) Adv De No cf. sentence (5) below

² The abbreviations used here have been chosen for mnemonic purposes. Abbreviations composed of two letters represent inflectional elements; of three, invariant elements; of four, immediate constituents of larger constructions. When all members of the abbreviation are capitalized, the abbreviation refers to a class which is named after one of its members. Statements on morphophonemic/morphographemic variation are not dealt with in this paper.

The simple sentence consists of Subj $\left\{ \begin{array}{l} \text{Pred} \\ \text{Verb Objt} \end{array} \right\}$

Abbreviations used are:

Adv	Adverbial	DE	<i>der</i> -Li	Objt	Object
Adj	Adjectival	KE	<i>kein</i> -Li	Ppn	Postpositional
Aj	Adjective	Li	Limiting Aj	Pred	Predicate
Cnj	Conjunctive	orl. B. a. γ	Adj	Prp	Prepositional
De	Descriptive Aj	No	Noun	Subj	Subject

Though I. A., II. A, etc. consist of but one word, they are included in the roster of phrases to provide a complete analysis of nouns which may occur in the subject position. Further, certain sub-classes of nouns are marked by the absence of *ein*, i.e. by an A rather than one form of B.a. constructions.

- II. Adjectives preceded by modifiers
- | | |
|--------------|---|
| A. ϕ De | Dunkles 'dark (beer)' |
| B. Li De | sein Dunkles 'his dark (beer)' |
| C. Li Adv De | das recht Dunkle
'the quite dark (beer)' |
- III. I and II followed by modifiers
- | | |
|---------------------|---|
| A. I + (Li) (De) No | das dunkle Bier dieser
Brauerei 'the dark beer
of this brewery' |
| B. I + Adv | ein reichliches Material zum
Beleg 'copious material
as evidence' |
- IV. I and II with positionals; such phrases are Adv
- | | |
|------------------|--|
| A. Prp + I | mit einem solchen Bier
'with such a beer' |
| B. I + Ppn | diesem Gasthaus gegenüber
'opposite this inn' |
| C. Prp + I + Ppn | von dieser Grundlage aus
'from this foundation' |

Other classifications of noun phrases might be made. This one is based on morphological differentiation of the head of the phrases and immediate constituent analysis of sentences. Each of these four types of phrases is of course found with various complexities, as is indicated after I.B.b and I.C.c; but these do not present any essentially new features. Moreover, each component of II III and IV phrases must be analyzed in greater detail, as in the analyses after I labeled *a.a.*, *a. β .*, etc. For the sake of brevity such further analysis has been omitted here; but an example of such analysis is provided in the third last paragraph of this paper. Given a summary like that presented here, various procedures for mechanical translation could be devised, or various types of notation. Some of these will be touched on as further aspects of noun phrases are discussed.

At the present stage of machine translation, analysis of materials to be translated must be undertaken on the graphemic level, the morphological level, and the syntactic level, if we leave aside here lexical problems. Moreover, it is restricted to one style of German, since machine translation work will be directed at technical rather than at literary or colloquial German. While with Joos one may be dubious about the possibility of achieving machine translation from

the imperfect representation that is provided by the written language,³ our imperfect knowledge of the signals of spoken German and the stage of our translation equipment suggest that we must first attempt to achieve a suitable method of dealing with written texts, subsequently with spoken materials. If machine translation of written texts is impossible of achievement, other than in the dubious form of the so-called pragmatic approach, preliminary linguistic work will be useful for the machine translation of spoken texts and it will contribute to a description of German. For the treatments of the noun phrase in our extant grammars of German need some modification, based as they are on a traditional morphological approach. Useful descriptions of some types of phrases however have been published and others are utilized, though not systematically presented, in some pedagogical works on the German language; previous discussions of the machine translation of German also include valuable material on the noun phrase. These differ from the present study in dealing essentially with morphological data, while a syntactic analysis is presented here.

Although written German may be deficient in marking some of the signals of spoken German, it provides us with graphemic markers which may in part compensate for the absent phonological markers. The first of these is the sentence-final marker, normally the period, which is equivalent to phonemically significant final junctures. (Another, the interrogation point, also has a distinctive function, though different in extent from the German sentence-final junctures; and the exclamation point again has a significant function, though this may be redundant with morphological and syntactic markers of the imperative construction.) From these, with the help of morphological and syntactic markers, we can readily determine clauses, which as noted above can be broken into constituents. Two other

³ See his review of *Machine Translation of Languages* ed. by W. N. Locke and A. D. Booth (New York, 1955) in *Language* 32. 293-98 (1956).

⁴ See especially Victor A. Oswald Jr. and Stuart L. Fletcher Jr. "Proposals for the Mechanized Resolution of German Syntax Patterns", *Modern Language Forum* 36. 1-24 (1951) and Erwin Reifler, "The Mechanical Determination of Meaning" in *Machine Translation of Languages* 136-64. C. V. Pollard has utilized graphemic and syntactic markers in teaching technical German; see his *Key to German Translation* (Austin, 1954).

graphemic markers are of great significance for the analysis of the noun phrase: the comma and the capital letter. Since commas stand before any new clause in German, they assist greatly in determining the constructions which must be subjected to immediate constituent analysis, in distinguishing homonymic pronouns from adjectives, and so on. We may note at once, however, that graphemic markers will not of themselves permit analysis of all clauses. A simple clause, determined by a comma and a period, like (2): *welche mögliche Vollkommenheit simulieren*. 'which simulate possible perfection' can only be analyzed on the basis of morphological data outside the noun phrase, i.e. by the morphological datum that a noun phrase consisting of *welche* followed by the *e* of *mögliche* before *Vollkommenheit* would require final *t* on *simulier-* if we were dealing with a unit noun phrase *welche mögliche Vollkommenheit*; the syntactic datum that *simulieren* stands at the end of the clause requires us to make our immediate constituent analysis after *welche*. Moreover, the graphemically significant capital letters in German provide information unavailable in the spoken language, for they mark nouns, as may be illustrated by the following sentences: *Er geht zu messen*. 'He is going for the purpose of measuring.', and *Er geht zu Messen*. 'He goes to fairs.'⁵ It might be assumed then that the final extent of noun phrases is marked essentially by capital letters, and that the chief machine translation problem in the determination of noun phrases is one of order, not of extent. For noun phrases of type I this assumption holds. Graphemic devices, capital letters, are instrumental in determining the final limit of immediate constituents in successive cuts, marking nouns; syntactic and morphological sub-classes mark their prior limit. Accordingly in describing in greater detail noun phrases of this type, we must determine the classes and sub-classes of their components.

Homonymity of phrases like *der Wagen* 'the car' and *der Wagen* 'of the scales' suggests that we cannot make adequate analyses from the sub-class of the noun modifiers alone, but that we also must classify nouns into various sub-classes—indicating the sub-class by the form of the noun we choose to list. Through comparison of minimally contrasting clauses like: *Er hat den Wagen gekauft*. 'He

⁵ Note however the qualification discussed by Reifler, *Machine Translation of Languages*, 150-51; since capital letters are required at the beginnings of sentences, nouns in this position are not distinctively marked.

bought the car.’ and *Er hat die Wagen gekauft*. ‘He bought the cars.’, where the contrast between the noun modifiers alone marks the difference between singular and plural, it is clear that distinguishing data may further be presented in the sub-classes of modifiers. Distinctions marked in this way for noun phrases of type I.B. (except for a. γ) number seven for nouns of masculine gender, six for nouns of neuter gender, five for nouns of feminine gender.

Although we make provision for these five to seven forms, we will still need to rely on syntactic, or semantic, markers to distinguish between homonymic nominal phrases like *den alten Wagen* ‘the old car (acc.)’ and *den alten Wagen* ‘(in) the old cars’ or *die neue Bank, -e* ‘the new bench’ and *die neue Bank, -en* ‘the new bank.’ Homonyms like *die Wagen* ‘cars (nom.)’ and ‘cars (acc.)’ will be distinguished on the basis of our initial cut of the sentence into two immediate constituents; if the homonymous nominative and accusative plural falls in the constituent accompanying the verb, it usually stands in the accusative. This procedure will also distinguish the homonymous nominative and accusative singular of feminine nouns. Though listings of nouns and their modifiers in their five to seven possible forms will then account for all of the information provided by the morphological markers of phrases of type I.B, additional distinctions are indicated by syntactic markers. For type I.A too, we have to rely heavily on syntactic markers. Type I.C requires on the other hand even further syntactic analysis, for in this sub-type lie many of the, difficulties met in technical German.

Simple examples of type I.C are contained in sentence (1): *jede sprachliche Veränderung* and *jeder dialektischen Eigentümlichkeit*. *jed-* belongs to a syntactic sub-class, the DE-words, which marks the prior limit of a noun phrase. Mutually exclusive with the definite article, it contains information which defines the function of a following noun. But while noun phrases of type I.C are usually bounded by adjectives, generally limiting, and nouns, the elements contained within these bounds must be analyzed for their form; for the endings of the enclosed elements disclose the proper analysis of noun phrases longer in extent than two words. If in accordance with certain rules, the enclosed elements end in *-e/-en*, the phrase can be equated to corresponding English noun phrases and managed

as simply as are the noun phrases of our first sentence.⁶ If, however, the enclosed elements end in different graphemes or if they belong to sub-classes incapable of forming type I.C noun phrases, we are generally confronted with a I.C.c construction. The following sentence (3) illustrates an impossible grapheme.

vielmehr kommt es auf die mit ihrer Hilfe zu erlangenden
allgemeinen Begriffe an. 'rather, it depends on the general con-
cepts which may be obtained with its aid.'

mit signals as impossible analysis of the segment from *die* to *Be-
griffe* as a simple noun phrase; the *-en* ending on *allgemeinen*, how-
ever, suggests assumption of the intervening material before
erlangenden as adverbial. Just as a descriptive adjective ending
in *-e* or *-en* without an immediately preceding limiting adjective
generally marks this construction, so do two limiting adjectives—
unless they indicate that one of the apparent limiting adjectives is
a pronoun, as in the following sentence (4):

Ihre Betonung stimmte mit der Genetive der konsonantischen
Stämme überein. 'Their accentuation coincided with that of the
genitives of the consonant stems.'

While constructions of type I.C.c by their variety and complexity
often provide man with the sole diversion he derives from reading
technical German, their basis structure is that of the examples after
I.C.c on the chart, as we may illustrate from the following moderately
complex example (5):

[Es werden] mit Hilfe des menschlichen Leibes bearbeitete
oder auch nur von dem Orte ihrer Entstehung zu irgend einem
Dienste verrückte Naturgegenstände ...[übertragen] 'natural objects
handled with the aid of the human body or merely moved from the
place of their origin for some use...'

The material of this example which is not enclosed in brackets
consists of the construction given after I.C.c, with two adjectives

⁶ On the other hand, an *-e* or *-en* ending after Li may not indicate De, as
in the following examples: Paul 17 ...von den oben Seite acht angegebenen
vier Kategorien... 'Of the four categories given above on page eight';
Paul 18 ...die aus ihnen bestehenden Sätze... 'the sentences consisting
of them...'

modifying *Naturgegenstände*, *bearbeitete* and *verrückte*. Each adjective is modified by an adverbial. The adverbial preceding *bearbeitete* is a type IV noun phrase, *mit Hilfe*, accompanied by a III, (*Hilfe*) *des menschlichen Leibes*; the material preceding *verrückte* is a type IV accompanied by a type III, followed by a type IV, *von dem Orte ihrer Entstehung zu irgend einem Dienste*. Each of these phrases in turn incorporates or is a type described after I: (*mit Hilfe*) a type I.A; *des menschlichen Leibes* a type I.C.a; *dem Orte* a type I.B. a.a; *ihrer Entstehung* and *irgend einem Dienste* a type I.B. a.β. If one wished to proceed in reverse, to apply successive rules for the expansion of a given symbol (as has been suggested by some workers in machine translation) this could readily be done. One can handle such complex phrases in a variety of ways, by treating them as transformations or by using whatever other analytical device seems most economical.

The other noun phrases may similarly be analyzed with the procedures outlined above. Simple examples of II (which conforms to I, with the difference that the nominal retains the inflections of the descriptive adjective) were presented on the chart. A variant is found in the following sentence (6) ending in *entsprungenen*:

Um die einer in ihr selbst entsprungenen entsprechende Vorstellungsverbindung in einer anderen Seele hervorzurufen,...
 'In order to produce in another mind the association of ideas which corresponds to one which has arisen in oneself,...

Examples of III and IV were given in the chart, and further illustrated in sentence (5) above.

A complete description of the noun phrase in German would be much more explicit with regard to sub-types of these four types, to variations in forms, and to frequency of the various types. It would also deal with problems such as the order of adjectivals, especially that of descriptive adjectives. Their order presents little difficulty when one moves from German to English, as is evident from the following extensive noun phrase of the type beloved by linguists:

All the nice little new two-story white brick houses...
 All die netten kleinen neuen weißen zwei-stöckigen
 Backsteinhäuser..

The order of the modifiers is virtually identical in the two languages; the propensity of German to make compound nouns as illustrated in this phrase would be a problem outside the scope of analysis of the noun phrase, one in the sphere of morphological analysis. Further problems such as the non-correspondance of limiting adjectives in some German and English noun phrases, methods of distinguishing between adjectives like *der*, *jeder*, *seiner* and pronouns homonymous with them would also require full description in a complete analysis of the noun phrase in German.