(13) shows the rules for nominative noun phrases. We would need analogous rules for genitive, dative, and accusative. We would then require 48 symbols for determiners ($3 * 2 * 4 * 2$), 48 symbols for nouns and 48 rules rather than one.

2.2 Expanding PSG with features

Phrase structure grammars which only use atomic symbols are problematic as they cannot capture certain generalizations. We as linguists can recognize that NP_3_sg_nom stands for a noun phrase because it contains the letters NP. However, in formal terms this symbol is just like any other symbol in the grammar and we cannot capture the commonalities of all the symbols used for NPs. Furthermore, unstructured symbols do not capture the fact that the rules in (13) all have something in common. In formal terms, the only thing that the rules have in common is that there is one symbol on the left-hand side of the rule and two on the right.

We can solve this problem by introducing features which are assigned to category symbols and therefore allow for the values of such features to be included in our rules. For example, we can assume the features person, number and case for the category symbol NP. For determiners and nouns, we would adopt an additional feature for gender and one for inflectional class.

(14) NP(3,sg,nom) → D(fem,sg,nom,strong) N(fem,sg,nom,strong)
NP(3,sg,nom) → D(mas,sg,nom,strong) N(mas,sg,nom,strong)

If we were to use variables rather than the values in (14), we would get the following rules as in (15):

(15) NP(3,Num,Case) → D(Gen,Num,Case,Infl) N(Gen,Num,Case,Infl)

The values of the variables here are not important. What is important is that they match. The value of the person feature (the first position in the NP(3,Num,Case)) is fixed at '3' by the rule. These kind of restrictions on the values can, of course, be determined in the lexicon:

(16) NP(3,sg,nom) → es
D(mas,sg,nom,strong) → des

The rules in (10) can be collapsed into a single schema as in (17):

(17) S → NP(Per1,Num1,nom)
NP(Per2,Num2,dat)
NP(Per3,Num3,acc)
V(Per1,Num1,ditransitive)

The identification of Per1 and Num1 on the verb and on the subject ensures that there is subject-verb agreement. For the other NPs, the values of these features are irrelevant. The case of these NPs is explicitly determined.

2.3 Semantics

In the introductory chapter and the previous sections, we have been dealing with syntactic aspects of language and the focus will remain very much on syntax for the remainder of this book. It is, however, important to remember that we use language to communicate, that is, to transfer information about certain situations, topics or opinions. If we want to accurately explain our capacity for language, then we also have to explain the meanings that our utterances have. To this end, it is necessary to understand their syntactic structure, but this alone is not enough. Furthermore, theories of language acquisition that only concern themselves with the acquisition of syntactic constructions are also inadequate. The syntax-semantics interface is therefore important and every grammatical theory has to say something about how syntax and semantics interact. In the following, I will show how we can combine phrase structure rules with semantic information. To represent meanings, I will use first-order predicate logic and λ-calculus. Unfortunately, it is not possible to provide a detailed discussion of the basics of logic so that even readers without prior knowledge can follow all the details, but the simple examples discussed here should be enough to provide some initial insights into how syntax and semantics interact and furthermore, how we can develop a linguistic theory to account for this.

To show how the meaning of a sentence is derived from the meaning of its parts, we will consider (18a). We assign the meaning in (18b) to the sentence in (18a).

(18) a. Max schläft.
Max sleeps
'Max is sleeping.'

b. *schlafen′*(*max′*)

Here, we are assuming *schlafen′* to be the meaning of *schläft* 'sleeps'. We use prime symbols to indicate that we are dealing with word meanings and not actual words. At first glance, it may not seem that we have really gained anything by using *schlafen′* to represent the meaning of (18a), since it is just another form of the verb *schläft* 'sleeps'. It is, however, important to concentrate on a single verb form as inflection is irrelevant when it comes to meaning. We can see this by comparing the examples in (19a) and (19b):

(19) a. Jeder Junge schläft.
every boy sleeps
'Every boy sleeps.'

b. Alle Jungen schlafen.
all boys sleep
'All boys sleep.'

When looking at the meaning in (18b), we can consider which part of the meaning comes from each word. It seems relatively intuitive that *max′* comes from *Max*, but the trickier question is what exactly *schläft* 'sleeps' contributes in terms of meaning. If we think

about what characterizes a 'sleeping' event, we know that there is typically an individual who is sleeping. This information is part of the meaning of the verb *schlafen* 'to sleep'. The verb meaning does not contain information about the sleeping individual, however, as this verb can be used with various subjects:

(20) a. Paul schläft.
Paul sleeps
'Paul is sleeping.'

b. Mio schläft.
Mio sleeps
'Mio is sleeping.'

c. Xaver schläft.
Xaver sleeps
'Xaver is sleeping.'

We can therefore abstract away from any specific use of *schlafen′* and instead of, for example, *max′* in (18b), we use a variable (e. g. x). This x can then be replaced by *paul′*, *mio′* or *xaver′* in a given sentence. To allow us to access these variables in a given meaning, we can write them with a λ in front. Accordingly, *schläft* 'sleeps' will have the following meaning:

(21) λx *sleep′*(x)

The step from (18b) to (21) is referred to as *lambda abstraction*. The combination of the expression (21) with the meaning of its arguments happens in the following way: we remove the λ and the corresponding variable and then replace all instances of the variable with the meaning of the argument. If we combine (21) and *max′* as in (22), we arrive at the meaning in (18b).

(22) λx *sleep′*(x) *max′*

The process is called β-reduction or λ-conversion. To show this further, let us consider an example with a transitive verb. The sentence in (23a) has the meaning given in (23b):

(23) a. Max mag Lotte.
Max likes Lotte
'Max likes Lotte.'

b. *like′*(*max′*, *lotte′*)

The λ-abstraction of *mag* 'likes' is shown in (24):

(24) $\lambda y \lambda x$ *like′*(x, y)

Note that it is always the first λ that has to be used first. The variable y corresponds to the object of *mögen*. For languages like English it is assumed that the object forms a verb phrase (VP) together with the verb and this VP is combined with the subject. German

differs from English in allowing more freedom in constituent order. The problems that result for form meaning mappings are solved in different ways by different theories. The respective solutions will be addressed in the following chapters.

If we combine the representation in (24) with that of the object *Lotte*, we arrive at (25a), and following β-reduction, (25b):

(25) a. $\lambda y \lambda x$ *like′*(x, y) *lotte′*
b. λx *like′*(x, *lotte′*)

This meaning can in turn be combined with the subject and we then get (26a) and (26b) after β-reduction:

(26) a. λx *like′*(x, *lotte′*) *max′*
b. *like′*(*max′*, *lotte′*)

After introducing lambda calculus, integrating the composition of meaning into our phrase structure rules is simple. A rule for the combination of a verb with its subject has to be expanded to include positions for the semantic contribution of the verb, the semantic contribution of the subject and then the meaning of the combination of these two (the entire sentence). The complete meaning is the combination of the individual meanings in the correct order. We can therefore take the simple rule in (27a) and turn it into (27b):

(27) a. S → NP(nom) V
b. S(V′ NP′) → NP(nom, NP′) V(V′)

V′ stands for the meaning of V and NP′ for the meaning of the NP(nom). V′ NP′ stands for the combination of V′ and NP′. When analyzing (18a), the meaning of V′ is λx schlafen(x) and the meaning of NP′ is *max′*. The combination of V′ NP′ corresponds to (28a) or after β-reduction to (18b) – repeated here as (28b):

(28) a. λx *sleep′*(x) *max′*
b. *sleep′*(*max′*)

For the example with a transitive verb in (23a), the rule in (29) can be proposed:

(29) S(V′ NP2′ NP1′) → NP(nom, NP1′) V(V′) NP(acc, NP2′)

The meaning of the verb (V′) is first combined with the meaning of the object (NP2′) and then with the meaning of the subject (NP1′).

At this point, we can see that there are several distinct semantic rules for the phrase structure rules above. The hypothesis that we should analyze language in this way is called the *rule-by-rule hypothesis* (Bach 1976). A more general process for deriving the meaning of linguistic expression will be presented in Section 5.1.4.

2.4 Phrase structure rules for some aspects of German syntax

Whereas determining the direct constituents of a sentence is relative easy, since we can very much rely on the movement test due to the somewhat flexible order of constituents in German, it is more difficult to identify the parts of the noun phrase. This is the problem we will focus on in this section. To help motivate assumptions about $\overline{X}$ syntax to be discussed in Section 2.5, we will also discuss prepositional phrases.

2.4.1 Noun phrases

Up to now, we have assumed a relatively simple structure for noun phrases: our rules state that a noun phrase consists of a determiner and a noun. Noun phrases can have a distinctly more complex structure than (30a). This is shown by the following examples in (30):

(30) a. eine Frau
a woman

b. eine Frau, die wir kennen
a woman who we know

c. eine Frau aus Stuttgart
a woman from Stuttgart

d. eine kluge Frau
a smart woman

e. eine Frau aus Stuttgart, die wir kennen
a woman from Stuttgart who we know

f. eine kluge Frau aus Stuttgart
a smart woman from Stuttgart

g. eine kluge Frau, die wir kennen
a smart woman who we know

h. eine kluge Frau aus Stuttgart, die wir kennen
a smart woman from Stuttgart who we know

As well as determiners and nouns, noun phrases can also contain adjectives, prepositional phrases and relative clauses. The additional elements in (30) are adjuncts. They restrict the set of objects which the noun phrase refers to. Whereas (30a) refers to a being which has the property of being a woman, the referent of (30b) must also have the property of being known to us.

Our previous rules for noun phrases simply combined a noun and a determiner and can therefore only be used to analyze (30a). The question we are facing now is how we can modify this rule or which additional rules we would have to assume in order to

analyze the other noun phrases in (30). In addition to rule (31a), one could propose a rule such as the one in (31b).[5,6]

(31) a. NP → Det N
b. NP → Det A N

However, this rule would still not allow us to analyze noun phrases such as (32):

(32) alle weiteren schlagkräftigen Argumente
all further strong arguments
'all other strong arguments'

In order to be able to analyze (32), we require a rule such as (33):

(33) NP → Det A A N

It is always possible to increase the number of adjectives in a noun phrase and setting an upper limit for adjectives would be entirely arbitrary. Even if we opt for the following abbreviation, there are still problems:

(34) NP → Det A* N

The asterisk in (34) stands for any number of iterations. Therefore, (34) encompasses rules with no adjectives as well as those with one, two or more.

The problem is that according to the rule in (34) adjectives and nouns do not form a constituent and we can therefore not explain why coordination is still possible in (35):

(35) alle [[geschickten Kinder] und [klugen Frauen]]
all skillful children and smart women
'all the skillful children and smart women'

If we assume that coordination involves the combination of two or more word strings with the same syntactic properties, then we would have to assume that the adjective and noun form a unit.

The noun phrases with adjectives discussed thus far can be explained by the following rules:

(36) a. NP → Det $\overline{\text{N}}$
b. $\overline{\text{N}}$ → A $\overline{\text{N}}$
c. $\overline{\text{N}}$ → N

These rules state the following: a noun phrase consists of a determiner and a nominal element ($\overline{\text{N}}$). This nominal element can consist of an adjective and a nominal element (36b), or just a noun (36c). Since $\overline{\text{N}}$ is also on the right-hand side of the rule in (36b), we can apply this rule multiple times and therefore account for noun phrases with multiple adjectives such as (32). Figure 2.3 on the next page shows the structure of a noun phrase

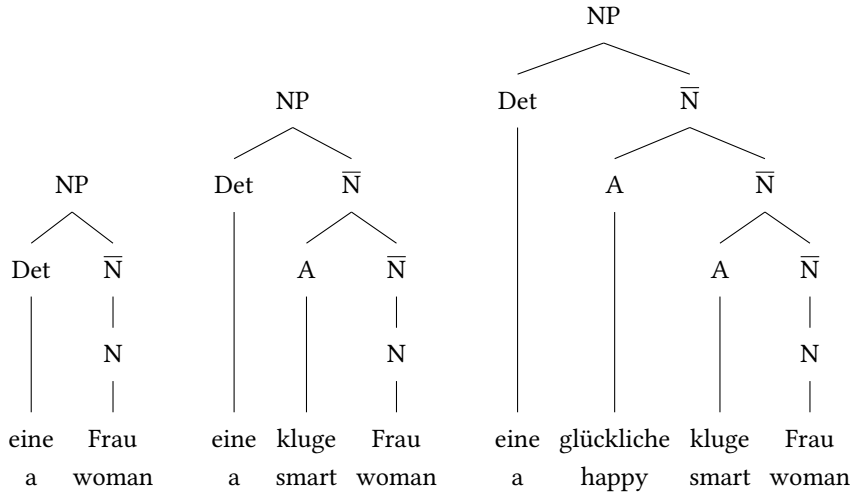


Figure 2.3: Noun phrases with differing numbers of adjectives

without an adjective and that of a noun phrase with one or two adjectives. The adjective *klug* 'smart' restricts the set of referents for the noun phrase. If we assume an additional adjective such as *glücklich* 'happy', then it only refers to those women who are happy as well as smart. These kinds of noun phrases can be used in contexts such as the following:

(37) A: Alle klugen Frauen sind unglücklich.
all smart women are unhappy
B: Nein, ich kenne eine glückliche kluge Frau.
no I know a happy smart woman

We observe that this discourse can be continued with *Aber alle glücklichen klugen Frauen sind schön* 'but all happy, smart women are beautiful' and a corresponding answer. The possibility to have even more adjectives in noun phrases such as *eine glückliche kluge Frau* 'a happy, smart woman' is accounted for in our rule system in (36). In the rule (36b), $\overline{\text{N}}$ occurs on the left as well as the right-hand side of the rule. This kind of rule is referred to as *recursive*.

We have now developed a nifty little grammar that can be used to analyze noun phrases containing adjectival modifiers. As a result, the combination of an adjective and noun is given constituent status. One may wonder at this point if it would not make sense to also assume that determiners and adjectives form a constituent, as we also have the following kind of noun phrases:

[5] See Eisenberg (2004: 238) for the assumption of flat structures in noun phrases.
[6] There are, of course, other features such as gender and number, which should be part of all the rules discussed in this section. I have omitted these in the following for ease of exposition.

(38) diese schlauen und diese neugierigen Frauen
these smart and these curious women

Here, we are dealing with a different structure, however. Two full NPs have been conjoined and part of the first conjunct has been deleted.

(39) diese schlauen ~~Frauen~~ und diese neugierigen Frauen
these smart women and these curious women

One can find similar phenomena at the sentence and even word level:

(40) a. dass Peter dem Mann das Buch ~~gibt~~ und Maria der Frau die Schallplatte
that Peter the man the book gives and Maria the woman the record
gibt
gives
'that Peters gives the book to the man and Maria the record to the woman'

b. be- und ent-laden
PRT and PRT-load
'load and unload'

Thus far, we have discussed how we can ideally integrate adjectives into our rules for the structure of noun phrases. Other adjuncts such as prepositional phrases or relative clauses can be combined with $\overline{\text{N}}$ in an analogous way to adjectives:

(41) a. $\overline{\text{N}} \rightarrow \overline{\text{N}}$ PP
b. $\overline{\text{N}} \rightarrow \overline{\text{N}}$ relative clause

With these rules and those in (36), it is possible – assuming the corresponding rules for PPs and relative clauses – to analyze all the examples in (30).

(36c) states that it is possible for $\overline{\text{N}}$ to consist of a single noun. A further important rule has not yet been discussed: we need another rule to combine nouns such as *Vater* 'father', *Sohn* 'son' or *Bild* 'picture', so-called *relational nouns*, with their arguments. Examples of these can be found in (42a–b). (42c) is an example of a nominalization of a verb with its argument:

(42) a. der Vater von Peter
the father of Peter
'Peter's father'

b. das Bild vom Gleimtunnel
the picture of.the Gleimtunnel
'the picture of the Gleimtunnel'

c. das Kommen des Installateurs
the coming of.the plumber
'the plumber's visit'

The rule that we need to analyze (42a,b) is given in (43):

(43) $\overline{\text{N}} \rightarrow \text{N PP}$

Figure 2.4 shows two structures with PP-arguments. The tree on the right also contains an additional PP-adjunct, which is licensed by the rule in (41a).

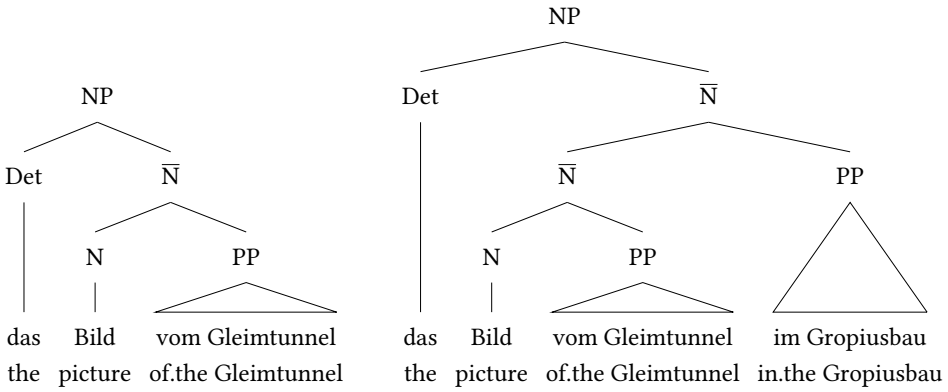


Figure 2.4: Combination of a noun with PP complement *vom Gleimtunnel* to the right with an adjunct PP

In addition to the previously discussed NP structures, there are other structures where the determiner or noun is missing. Nouns can be omitted via ellipsis. (44) gives an example of noun phrases, where a noun that does not require a complement has been omitted. The examples in (45) show NPs in which only one determiner and complement of the noun has been realized, but not the noun itself:

(44) a. eine kluge _
a smart
'a smart one'

b. eine kluge große _
a smart tall
'a smart tall one'

c. eine kluge _ aus Hamburg
a smart from Hamburg
'a smart one from Hamburg'

d. eine kluge _, die alle kennen
a smart who everyone knows
'a smart one who everyone knows'

(45) a. (Nein, nicht der Vater von Klaus), der _ von Peter war gemeint.
no not the father of Klaus the of Peter was meant
'No, it wasn't the father of Klaus, but rather the one of Peter that was meant.'

b. (Nein, nicht das Bild von der Stadtautobahn), das _ vom Gleimtunnel
no not the picture of the motorway the of.the Gleimtunnel
war beeindruckend.
was impressive
'No, it wasn't the picture of the motorway, but rather the one of the Gleimtunnel that was impressive.'

c. (Nein, nicht das Kommen des Tischlers), das _ des Installateurs ist
no not the coming of.the carpenter the of.the plumber is
wichtig.
important
'No, it isn't the visit of the carpenter, but rather the visit of the plumber that is important.'

The underscore marks the position where the noun would normally occur. In English, the pronoun *one* must often be used in the corresponding position, but in German the noun is simply omitted. (See Fillmore, Lee-Goldmann & Rhomieux (2012: Section 4.12) for English examples without the pronoun *one*.) In phrase structure grammars, this can be described by a so-called *epsilon production*. These rules replace a symbol with nothing (46a). The rule in (46b) is an equivalent variant which is responsible for the term *epsilon production*:

(46) a. $N \rightarrow$
b. $N \rightarrow \epsilon$

The corresponding trees are shown in Figure 2.5. Going back to boxes, the rules in (46)

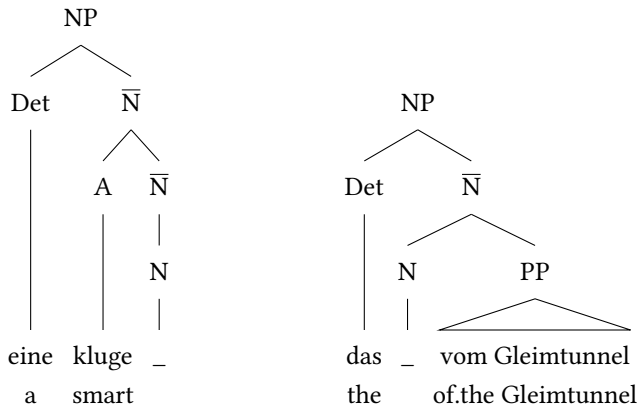


Figure 2.5: Noun phrases without an overt head

correspond to empty boxes with the same labels as the boxes of ordinary nouns. As we

have considered previously, the actual content of the boxes is unimportant when considering the question of where we can incorporate them. In this way, the noun phrases in (30) can occur in the same sentences. The empty noun box also behaves like one with a genuine noun. If we do not open the empty box, we will not be able to ascertain the difference to a filled box.

It is not only possible to omit the noun from noun phrases, but the determiner can also remain unrealized in certain contexts. (47) shows noun phrases in plural:

(47) a. Frauen
women

b. Frauen, die wir kennen
women who we know

c. kluge Frauen
smart women

d. kluge Frauen, die wir kennen
smart women who we know

The determiner can also be omitted in singular if the noun denotes a mass noun:

(48) a. Getreide
grain

b. Getreide, das gerade gemahlen wurde
grain that just ground was
'grain that has just been ground'

c. frisches Getreide
fresh grain

d. frisches Getreide, das gerade gemahlen wurde
fresh grain that just ground was
'fresh grain that has just been ground'

Finally, both the determiner and the noun can be omitted:

(49) a. Ich helfe klugen.
I help smart
'I help smart ones.'

b. Dort drüben steht frisches, das gerade gemahlen wurde.
there over stands fresh that just ground was
'Over there is some fresh (grain) that has just been ground.'

Figure 2.6 on the next page shows the corresponding trees.

It is necessary to add two further comments to the rules we have developed up to this point: up to now, I have always spoken of adjectives. However, it is possible to have very complex adjective phrases in pre-nominal position. These can be adjectives with complements (50a,b) or adjectival participles (50c,d):

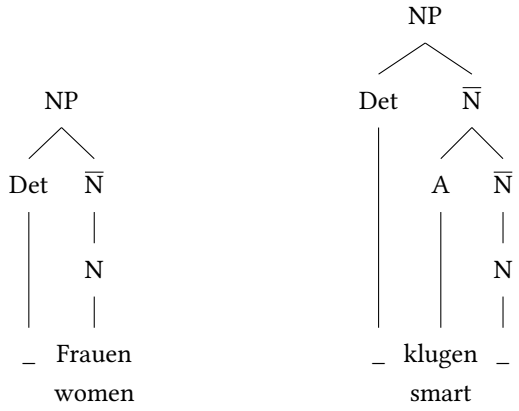


Figure 2.6: Noun phrases without overt determiner

(50) a. der seiner Frau treue Mann
the his.DAT wife faithful man
'the man faithful to his wife'

b. der auf seinen Sohn stolze Mann
the on his.ACC son proud man
'the man proud of his son'

c. der seine Frau liebende Mann
the his.ACC woman loving man
'the man who loves his wife'

d. der von seiner Frau geliebte Mann
the by his.DAT wife loved man
'the man loved by his wife'

Taking this into account, the rule (36b) has to be modified in the following way:

(51) $\overline{\text{N}} \rightarrow \text{AP}\ \overline{\text{N}}$

An adjective phrase (AP) can consist of an NP and an adjective, a PP and an adjective or just an adjective:

(52) a. $\text{AP} \rightarrow \text{NP A}$
b. $\text{AP} \rightarrow \text{PP A}$
c. $\text{AP} \rightarrow \text{A}$

There are two imperfections resulting from the rules we have developed thus far. These are the rules for adjectives or nouns without complements in (52c) as well as (36c) – repeated here as (53):

(53) $\overline{\text{N}} \rightarrow \text{N}$

If we apply these rules, then we will generate unary branching subtrees, that is trees with a mother that only has one daughter. See Figure 2.6 for an example of this. If we maintain the parallel to the boxes, this would mean that there is a box which contains another box which is the one with the relevant content.

In principle, nothing stops us from placing this information directly into the larger box. Instead of the rules in (54), we will simply use the rules in (55):

(54) a. $\text{A} \rightarrow$ kluge
b. $\text{N} \rightarrow$ Mann

(55) a. $\text{AP} \rightarrow$ kluge
b. $\overline{\text{N}} \rightarrow$ Mann

(55a) states that *kluge* 'smart' has the same properties as a full adjective phrase, in particular that it cannot be combined with a complement. This is parallel to the categorization of the pronoun *er* 'he' as an NP in the grammars (2) and (6).

Assigning $\overline{\text{N}}$ to nouns which do not require a complement has the advantage that we do not have to explain why the analysis in (56b) is possible as well as (56a) despite there not being any difference in meaning.

(56) a. $[_{\text{NP}}$ einige $[_{\overline{\text{N}}}$ kluge $[_{\overline{\text{N}}}$ $[_{\overline{\text{N}}}$ $[_{\text{N}}$ Frauen] und $[_{\overline{\text{N}}}$ $[_{\text{N}}$ Männer]]]]]]
some smart women and men

b. $[_{\text{NP}}$ einige $[_{\overline{\text{N}}}$ kluge $[_{\overline{\text{N}}}$ $[_{\text{N}}$ $[_{\text{N}}$ Frauen] und $[_{\text{N}}$ Männer]]]]]
some smart women and men

In (56a), two nouns have projected to $\overline{\text{N}}$ and have then been joined by coordination. The result of coordination of two constituents of the same category is always a new constituent with that category. In the case of (56a), this is also $\overline{\text{N}}$. This constituent is then combined with the adjective and the determiner. In (56b), the nouns themselves have been coordinated. The result of this is always another constituent which has the same category as its parts. In this case, this would be N. This N becomes $\overline{\text{N}}$ and is then combined with the adjective. If nouns which do not require complements were categorized as $\overline{\text{N}}$ rather than N, we would not have the problem of spurious ambiguities. The structure in (57) shows the only possible analysis.

(57) $[_{\text{NP}}$ einige $[_{\overline{\text{N}}}$ kluge $[_{\overline{\text{N}}}$ $[_{\overline{\text{N}}}$ Frauen] und $[_{\overline{\text{N}}}$ Männer]]]]
some smart women and men

2.4.2 Prepositional phrases

Compared to the syntax of noun phrases, the syntax of prepositional phrases (PPs) is relatively straightforward. PPs normally consist of a preposition and a noun phrase whose case is determined by that preposition. We can capture this with the following rule:

(58) PP → P NP

This rule must, of course, also contain information about the case of the NP. I have omitted this for ease of exposition as I did with the NP-rules and AP-rules above.

The Duden grammar (Eisenberg et al. 2005: §1300) offers examples such as those in (59), which show that certain prepositional phrases serve to further define the semantic contribution of the preposition by indicating some measurement, for example:

(59) a. [[Einen Schritt] vor dem Abgrund] blieb er stehen.
one step before the abyss remained he stand
'He stopped one step in front of the abyss.'

b. [[Kurz] nach dem Start] fiel die Klimaanlage aus.
shortly after the take.off fell the air.conditioning out
'Shortly after take off, the air conditioning stopped working.'

c. [[Schräg] hinter der Scheune] ist ein Weiher.
diagonally behind the barn is a pond
'There is a pond diagonally across from the barn.'

d. [[Mitten] im Urwald] stießen die Forscher auf einen alten Tempel.
middle in.the jungle stumbled the researchers on an old temple
'In the middle of the jungle, the researches came across an old temple.'

To analyze the sentences in (59a,b), one could propose the following rules in (60):

(60) a. PP → NP PP
b. PP → AP PP

These rules combine a PP with an indication of measurement. The resulting constituent is another PP. It is possible to use these rules to analyze prepositional phrases in (59a,b), but it unfortunately also allows us to analyze those in (61):

(61) a. * [$_{PP}$ einen Schritt [$_{PP}$ kurz [$_{PP}$ vor dem Abgrund]]]
one step shortly before the abyss

b. * [$_{PP}$ kurz [$_{PP}$ einen Schritt [$_{PP}$ vor dem Abgrund]]]
shortly one step before the abyss

Both rules in (60) were used to analyze the examples in (61). Since the symbol PP occurs on both the left and right-hand side of the rules, we can apply the rules in any order and as many times as we like.

We can avoid this undesired side-effect by reformulating the previously assumed rules:

(62) a. PP → NP $\overline{\text{P}}$
b. PP → AP $\overline{\text{P}}$
c. PP → $\overline{\text{P}}$

d. $\overline{\text{P}} \rightarrow \text{P NP}$

Rule (58) becomes (62d). The rule in (62c) states that a PP can consist of $\overline{\text{P}}$. Figure 2.7 shows the analysis of (63) using (62c) and (62d) as well as the analysis of an example with an adjective in the first position following the rules in (62b) and (62d):

(63) vor dem Abgrund
before the abyss
'in front of the abyss'

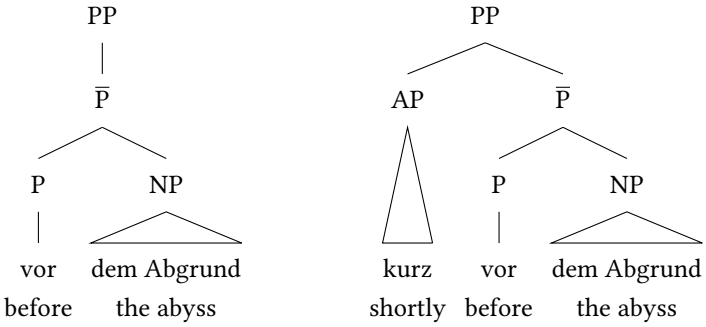


Figure 2.7: Prepositional phrases with and without measurement

At this point, the attentive reader is probably wondering why there is no empty measurement phrase in the left figure of Figure 2.7, which one might expect in analogy to the empty determiner in Figure 2.6. The reason for the empty determiner in Figure 2.6 is that the entire noun phrase without the determiner has a meaning similar to those with a determiner. The meaning normally contributed by the visible determiner has to somehow be incorporated in the structure of the noun phrase. If we did not place this meaning in the empty determiner, this would lead to more complicated assumptions about semantic combination: we only really require the mechanisms presented in Section 2.3 and these are very general in nature. The meaning is contributed by the words themselves and not by any rules. If we were to assume a unary branching rule such as that in the left tree in Figure 2.7 instead of the empty determiner, then this unary branching rule would have to provide the semantics of the determiner. This kind of analysis has also been proposed by some researchers. See Chapter 19 for more on empty elements.

Unlike determiner-less NPs, prepositional phrases without an indication of degree or measurement do not lack any meaning component for composition. It is therefore not necessary to assume an empty indication of measurement, which somehow contributes to the meaning of the entire PP. Hence, the rule in (62c) states that a prepositional phrase consists of $\overline{\text{P}}$, that is, a combination of P and NP.

2.5 $\overline{\text{X}}$ theory

If we look again at the rules that we have formulated in the previous section, we see that heads are always combined with their complements to form a new constituent (64a,b), which can then be combined with further constituents (64c,d):

(64) a. $\overline{\text{N}} \rightarrow \text{N PP}$
b. $\overline{\text{P}} \rightarrow \text{P NP}$
c. $\text{NP} \rightarrow \text{Det } \overline{\text{N}}$
d. $\text{PP} \rightarrow \text{NP } \overline{\text{P}}$

Grammarians working on English noticed that parallel structures can be used for phrases which have adjectives or verbs as their head. I discuss adjective phrases at this point and postpone the discussion of verb phrases to Chapter 3. As in German, certain adjectives in English can take complements with the important restriction that adjective phrases with complements cannot realize these pre-nominally in English. (65) gives some examples of adjective phrases:

(65) a. He is proud.
b. He is very proud.
c. He is proud of his son.
d. He is very proud of his son.

Unlike prepositional phrases, complements of adjectives are normally optional. *proud* can be used with or without a PP. The degree expression *very* is also optional.

The rules which we need for this analysis are given in (66), with the corresponding structures in Figure 2.8 on the next page.

(66) a. $\text{AP} \rightarrow \overline{\text{A}}$
b. $\text{AP} \rightarrow \text{AdvP } \overline{\text{A}}$
c. $\overline{\text{A}} \rightarrow \text{A PP}$
d. $\overline{\text{A}} \rightarrow \text{A}$

As was shown in Section 2.2, it is possible to generalize over very specific phrase structure rules and thereby arrive at more general rules. In this way, properties such as person, number and gender are no longer encoded in the category symbols, but rather only simple symbols such as NP, Det and N are used. It is only necessary to specify something about the values of a feature if it is relevant in the context of a given rule. We can take this abstraction a step further: instead of using explicit category symbols such as N, V, P and A for lexical categories and NP, VP, PP and AP for phrasal categories, one can simply use a variable for the word class in question and speak of X and XP.

This form of abstraction can be found in so-called $\overline{\text{X}}$ theory (or X-bar theory, the term *bar* refers to the line above the symbol.), which was developed by Chomsky (1970) and refined by Jackendoff (1977). This form of abstract rules plays an important role in many

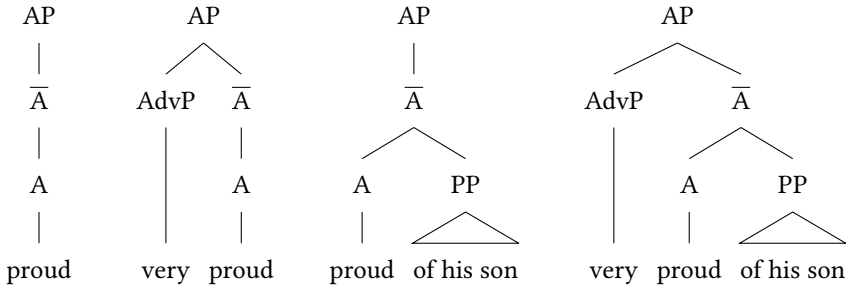


Figure 2.8: English adjective phrases

different theories. For example: Government & Binding (Chapter 3), Generalized Phrase Structure Grammar (Chapter 5) and Lexical Functional Grammar (Chapter 7). In HPSG (Chapter 9), $\overline{X}$ theory also plays a role, but not all restrictions of the $\overline{X}$ schema have been adopted.

(67) shows a possible instantiation of $\overline{X}$ rules, where the category X has been used in place of N, as well as examples of word strings which can be derived by these rules:

(67)	$\overline{X}$ rule	with specific categories	example strings
	$\overline{\overline{X}} \rightarrow \overline{\overline{\text{specifier}}}\ \overline{X}$	$\overline{\overline{N}} \rightarrow \overline{\overline{\text{DET}}}\ \overline{N}$	the [picture of Paris]
	$\overline{X} \rightarrow \overline{X}\ \overline{\overline{\text{adjunct}}}$	$\overline{N} \rightarrow \overline{N}\ \overline{\overline{\text{REL\_CLAUSE}}}$	[picture of Paris] [that everybody knows]
	$\overline{X} \rightarrow \overline{\overline{\text{adjunct}}}\ \overline{X}$	$\overline{N} \rightarrow \overline{\overline{A}}\ \overline{N}$	beautiful [picture of Paris]
	$\overline{X} \rightarrow X\ \overline{\overline{\text{complement}}}*$	$\overline{N} \rightarrow N\ \overline{\overline{P}}$	picture [of Paris]

Any word class can replace X (e. g. V, A or P). The X without the bar stands for a lexical item in the above rules. If one wants to make the bar level explicit, then it is possible to write X^0. Just as with the rule in (15), where we did not specify the case value of the determiner or the noun but rather simply required that the values on the right-hand side of the rule match, the rules in (67) require that the word class of an element on the right-hand side of the rule (X or $\overline{X}$) matches that of the element on the left-hand side of the rule ($\overline{X}$ or $\overline{\overline{X}}$).

A lexical element can be combined with all its complements. The '*' in the last rule stands for an unlimited amount of repetitions of the symbol it follows. A special case is zerofold occurrence of complements. There is no PP complement of *Bild* 'picture' present in *das Bild* 'the picture' and thus N becomes $\overline{N}$. The result of the combination of a lexical element with its complements is a new projection level of X: the projection level 1, which is marked by a bar. $\overline{X}$ can then be combined with adjuncts. These can occur to the left or right of $\overline{X}$. The result of this combination is still $\overline{X}$, that is the projection level is not changed by combining it with an adjunct. Maximal projections are marked by two bars. One can also write XP for a projection of X with two bars. An XP consists of a specifier

and $\overline{X}$. Depending on one's theoretical assumptions, subjects of sentences (Haider 1995, 1997a; Berman 2003a: Section 3.2.2) and determiners in NPs (Chomsky 1970: 210) are specifiers. Furthermore, degree modifiers (Chomsky 1970: 210) in adjective phrases and measurement indicators in prepositional phrases are also counted as specifiers.

Non-head positions can only host maximal projections and therefore complements, adjuncts and specifiers always have two bars. Figure 2.9 gives an overview of the minimal and maximal structure of phrases.

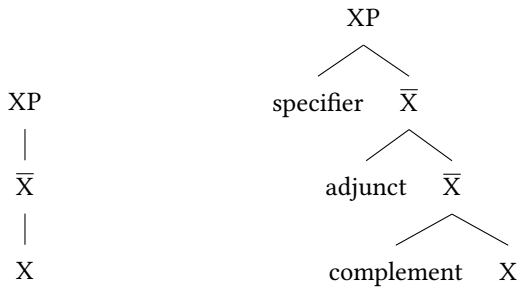


Figure 2.9: Minimal and maximal structure of phrases

Some categories do not have a specifier or have the option of having one. Adjuncts are optional and therefore not all structures have to contain an $\overline{X}$ with an adjunct daughter. In addition to the branching shown in the right-hand figure, adjuncts to XP and head-adjuncts are sometimes possible. There is only a single rule in (67) for cases in which a head precedes the complements, however an order in which the complement precedes the head is of course also possible. This is shown in Figure 2.9.

Figure 2.10 on the next page shows the analysis of the NP structures *das Bild* 'the picture' and *das schöne Bild von Paris* 'the beautiful picture of Paris'. The NP structures in Figure 2.10 and the tree for *proud* in Figure 2.8 show examples of minimally populated structures. The left tree in Figure 2.10 is also an example of a structure without an adjunct. The right-hand structure in Figure 2.10 is an example for the maximally populated structure: specifier, adjunct, and complement are present.

The analysis given in Figure 2.10 assumes that all non-heads in a rule are phrases. One therefore has to assume that there is a determiner phrase even if the determiner is not combined with other elements. The unary branching of determiners is not elegant but it is consistent.[7] The unary branchings for the NP *Paris* in Figure 2.10 may also seem somewhat odd, but they actually become more plausible when one considers more complex noun phrases:

[7] For an alternative version of $\overline{X}$ theory which does not assume elaborate structure for determiners see Muysken (1982a).

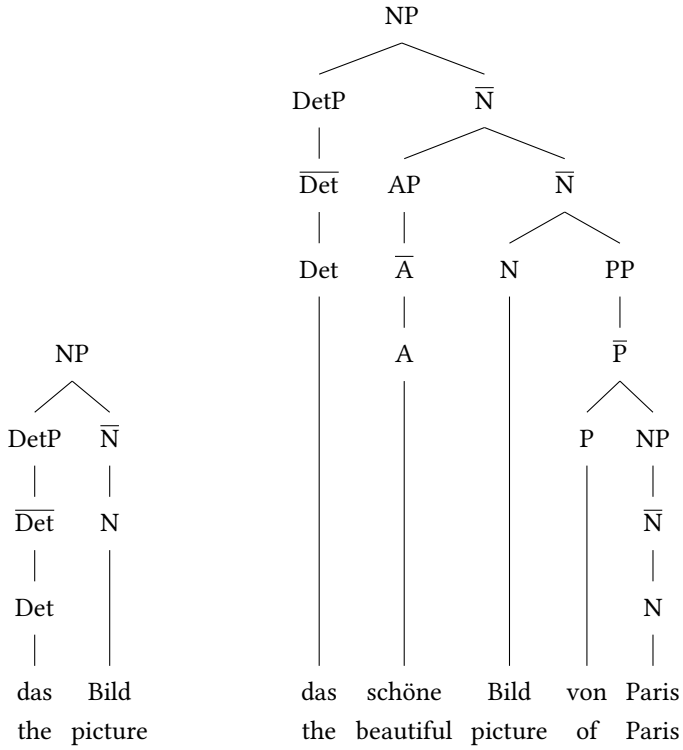


Figure 2.10: $\overline{\text{X}}$ analysis of *das Bild* 'the picture' and *das schöne Bild von Paris* 'the beautiful picture of Paris'

(68) a. das Paris der dreißiger Jahre
the Paris of.the thirty years
'30's Paris'

b. die Maria aus Hamburg
the Maria from Hamburg
'Maria from Hamburg'

Unary projections are somewhat inelegant but this should not concern us too much here, as we have already seen in the discussion of the lexical entries in (55) that unary branching nodes can be avoided for the most part and that it is indeed desirable to avoid such structures. Otherwise, one gets spurious ambiguities. In the following chapters, we will discuss approaches such as Categorial Grammar and HPSG, which do not assume unary rules for determiners, adjectives and nouns.

Furthermore, other $\overline{\text{X}}$ theoretical assumptions will not be shared by several theories discussed in this book. In particular, the assumption that non-heads always have to

be maximal projections will be disregarded. Pullum (1985) and Kornai & Pullum (1990) have shown that the respective theories are not necessarily less restrictive than theories which adopt a strict version of the $\overline{X}$ theory. See also the discussion in Section 13.1.2.

Comprehension questions

1. Why are phrase structure grammars that use only atomic categories inadequate for the description of natural languages?
2. Assuming the grammar in (6), state which steps (replacing symbols) one has to take to get to the symbol V in the sentence (69).

   (69) er das Buch dem Mann gibt
   he the book the man gives
   'He gives the book to the man.'

   Your answer should resemble the analysis in (3).
3. Give a representation of the meaning of (70) using predicate logic:

   (70) a. Ulrike kennt Hans.
   Ulrike knows Hans
   b. Joshi freut sich.
   Joshi is.happy REFL
   'Joshi is happy.'

Exercises

1. On page 57, I claimed that there is an infinite number of grammars we could use to analyze (1). Why is this claim correct?
2. Try to come up with some ways in which we can tell which of these possible grammars is or are the best?
3. A fragment for noun phrase syntax was presented in Section 2.4.1. Why is the interaction of the rules in (71) problematic?

   (71) a. $NP \rightarrow Det\ \overline{N}$
   b. $\overline{N} \rightarrow N$
   c. $Det \rightarrow \epsilon$
   d. $N \rightarrow \epsilon$
4. Why is it not a good idea to mark *books* as NP in the lexicon?

5. Can you think of some reasons why it is not desirable to assume the following rule for nouns such as *books*:

   (72) NP → Modifier* books Modifier*

   The rule in (72) combines an unlimited number of modifiers with the noun *books* followed by an unlimited number of modifiers. We can use this rule to derive phrases such as those in (73):

   (73) a. books
   b. interesting books
   c. interesting books from Stuttgart

   Make reference to coordination data in your answer. Assume that symmetric coordination requires that both coordinated phrases or words have the same syntactic category.

6. Fillmore et al. (2012) suggested treating nounless structures like those in (74) as involving a phrasal construction that combines the determiner *the* with an adjective.

   (74) a. Examine the plight of the very poor.
   b. Their outfits range from the flamboyant to the functional.
   c. The unimaginable happened.

   (75) shows a phrase structure rule that corresponds to their construction:

   (75) NP → the Adj

   Adj stands for something that can be a single word like *poor* or complex like *very poor*.

   Revisit the German data in (44) and (45) and explain why such an analysis and even a more general one as in (76) would not extend to German.

   (76) NP → Det Adj

7. Why can $\overline{X}$ theory not account for German adjective phrases without additional assumptions? (This task is for (native) speakers of German only.)

8. Come up with a phrase structure grammar that can be used to analyze the sentence in (77), but also rules out the sentences in (78).

(77) a. Der Mann hilft der Frau.
the.NOM man helps the.DAT woman
'The man helps the woman.'

b. Er gibt ihr das Buch.
he.NOM gives her.DAT the book
'He gives her the book.'

c. Er wartet auf ein Wunder.
he.NOM waits on a miracle
'He is waiting for a miracle.'

(78) a. * Der Mann hilft er.
the.NOM man helps he.NOM

b. * Er gibt ihr den Buch.
he.NOM gives her.DAT the.M book.N

9. Consider which additional rules would have to be added to the grammar you developed in the previous exercise in order to be able to analyze the sentences in (79):

(79) a. Der Mann hilft der Frau jetzt.
the.NOM man helps the.DAT woman now
'The man helps the woman now.'

b. Der Mann hilft der Frau neben dem Bushäuschen.
the.NOM man helps the.DAT woman next to.the bus.shelter
'The man helps the woman next to the bus shelter.'

c. Er gibt ihr das Buch jetzt.
he.NOM gives her.DAT the.ACC book now
'He gives her the book now.'

d. Er gibt ihr das Buch neben dem Bushäuschen.
he.NOM gives her.DAT the.ACC book next to.the bus.shelter
'He gives her the book next to the bus shelter.'

e. Er wartet jetzt auf ein Wunder.
he.NOM waits now on a miracle
'He is waiting for a miracle now.'

f. Er wartet neben dem Bushäuschen auf ein Wunder.
he.NOM waits next to.the.DAT bus.shelter on a miracle
'He is waiting for a miracle next to the bus shelter.'

10. Install a Prolog system (e. g. SWI-Prolog[8]) and try out your grammar. Details for the notation can be found in the corresponding handbook under the key word Definite Clause Grammar (DCG).

Further reading

The expansion of phrase structure grammars to include features was proposed as early as 1963 by Harman (1963).

The phrase structure grammar for noun phrases discussed in this chapter covers a large part of the syntax of noun phrases but cannot explain certain NP structures. Furthermore, it has the problem, which exercise 3 is designed to show. A discussion of these phenomena and a solution in the framework of HPSG can be found in Netter (1998) and Kiss (2005).

The discussion of the integration of semantic information into phrase structure grammars was very short. A detailed discussion of predicate logic and its integration into phrase structure grammars – as well as a discussion of quantifier scope – can be found in Blackburn & Bos (2005).

[8] http://www.swi-prolog.org/

3 Transformational Grammar – Government & Binding

Transformational Grammar and its subsequent incarnations (such as Government and Binding Theory and Minimalism) were developed by Noam Chomsky at MIT in Boston (Chomsky 1957, 1965, 1975, 1981a, 1986a, 1995b). Manfred Bierwisch (1963) was the first to implement Chomsky's ideas for German. In the 60s, the decisive impulse came from the *Arbeitsstelle Strukturelle Grammatik* 'Workgroup for Structural Grammar', which was part of the Academy of Science of the GDR. (See Bierwisch (1992) and Vater (2010) for an historic overview.) As well as Bierwisch's work, the following other works focussing on German, which have been written as part of this research program should also be mentioned: Fanselow (1987), Fanselow & Felix (1987), von Stechow & Sternefeld (1988), Grewendorf (1988), Haider (1993), Sternefeld (2006).

The variants of Chomskyan theories are often grouped under the heading *Generative Grammar*. This term comes from the fact that phrase structure grammars, coupled with the additional assumptions of Chomsky, can generate sets of well-formed expressions (see p. 56). It is such a set of sentences that constitutes a language (in the formal sense) and one can test if a sentence forms part of a language by checking if a particular sentence is in the set of sentences generated by a given grammar or not. In this sense, simple phrase structure grammars and, with corresponding formal assumptions, GPSG, LFG, HPSG and Construction Grammar (CxG) are generative theories. In recent years, a different view of the formal basis of theories such as LFG, HPSG and CxG has emerged such that the aforementioned theories are now *model theoretic* theories rather than generative-enumerative ones[1] (See Chapter 14 for discussion). In 1957, Chomsky defined the term *Generative Grammar* in the following way (also see Chomsky 1995b: 162):

> A grammar of a language purports to be a description of the ideal speaker-hearer's intrinsic competence. If the grammar is, furthermore, perfectly explicit – in other words, if it does not rely on the intelligence of the understanding reader but rather provides an explicit analysis of his contribution – we may call it (somewhat redundantly) a *generative grammar*. (Chomsky 1965: 4)

In this sense, all grammatical theories discussed in this book would be viewed as generative grammars. To differentiate further, sometimes the term *Mainstream Generative Grammar* (MGG) is used (Culicover & Jackendoff 2005: 3) for Chomskyan models. In this

[1] Model theoretic approaches are always constraint-based and the terms *model theoretic* and *constraint-based* are sometimes used synonymously.

chapter, I will discuss a well-developed and very influential version of Chomskyan grammar, GB theory. More recent developments following Chomsky's Minimalist Program are dealt with in Chapter 4.

3.1 General remarks on the representational format

This section provides an overview of general assumptions. I introduce the concept of transformations in Section 3.1.1. Section 3.1.2 provides background information about assumptions regarding language acquisition, which shaped the theory considerably, Section 3.1.3 introduces the so-called T model, the basic architecture of GB theory. Section 3.1.4 introduces the $\overline{\text{X}}$ theory in the specific form that is used in GB and Section 3.1.5 shows how this version of the $\overline{\text{X}}$ theory can be applied to English. The discussion of the analysis of English sentences is an important prerequisite for the understanding of the analysis of German, since many analyses in the GB framework are modeled in parallel to the analyses of English. Section 3.1.6 introduces the analysis of German clauses in a parallel way to what has been done for English in Section 3.1.5.

3.1.1 Transformations

In the previous chapter, I introduced simple phrase structure grammars. Chomsky (1957: Chapter 5) criticized these kind of rewrite grammars since – in his opinion – it is not clear how one can capture the relationship between active and passive sentences or the various ordering possibilities of constituents in a sentence. While it is of course possible to formulate different rules for active and passive sentences in a phrase structure grammar, it would not adequately capture the fact that the same phenomenon occurs in the example pairs in (1)–(3):

(1) a. weil dort noch jemand arbeitet
because there still somebody works
'because somebody is still working there'

b. weil dort noch gearbeitet wurde
because there still worked was
'because work was still being done there'

(2) a. weil er den Weltmeister schlägt
because he the world.champion beats
'because he beats the world champion'

b. weil der Weltmeister geschlagen wurde
because the world.champion beaten was
'because the world champion was beaten'

(3) a. weil der Mann der Frau den Schlüssel stiehlt
because the man the woman the key steals
'because the man is stealing the key from the woman'

b. weil der Frau der Schlüssel gestohlen wurde
because the woman the key stolen was
'because the key was stolen from the woman'

Chomsky (1957: 43) suggests a transformation that creates a connection between active and passive sentences. The passive transformation for English that he suggested has the form in (4):

(4) NP V NP → 3 [$_{AUX}$ be] 2en [$_{PP}$ [$_{P}$ by] 1]
1 2 3

This transformational rule maps a tree with the symbols on the left-hand side of the rule onto a tree with the symbols on the right-hand side of the rule. Accordingly, 1, 2 and 3 on the right of the rule correspond to symbols, which are under the numbers on the left-hand side. *en* stands for the morpheme which forms the participle (*seen*, *been*, ..., but also *loved*). Both trees for (5a,b) are shown in Figure 3.1.

(5) a. John loves Mary.
b. Mary is loved by John.

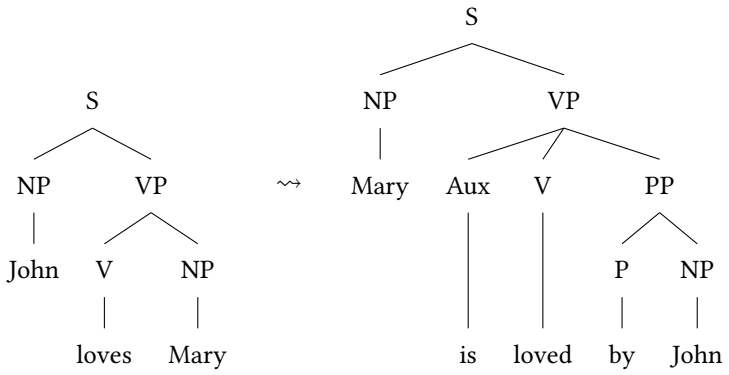


Figure 3.1: Application of passive transformation

The symbols on the left of transformational rules do not necessarily have to be in a local tree, that is, they can be daughters of different mothers as in Figure 3.1.

Rewrite grammars were divided into four complexity classes based on the properties they have. The simplest grammars are assigned to the class 3, whereas the most complex are of type 0. The so-called context-free grammars we have dealt with thus far are of

type 2. Transformational grammars which allow symbols to be replaced by arbitrary other symbols are of type 0 (Peters & Ritchie 1973). Research on the complexity of natural languages shows that the highest complexity level (type 0) is too complex for natural language. It follows from this – assuming that one wants to have a restrictive formal apparatus for the description of grammatical knowledge (Chomsky 1965: 62) – that the form and potential power of transformations has to be restricted.[2] Another criticism of early versions of transformational grammar was that, due to a lack of restrictions, the way in which transformations interact was not clear. Furthermore, there were problems associated with transformations which delete material (see Klenk 2003: Section 3.1.4). For this reason, new theoretical approaches such as Government & Binding (Chomsky 1981a) were developed. In this model, the form that grammatical rules can take is restricted (see Section 3.1.4). Elements moved by transformations are still represented in their original position, which makes them recoverable at the original position and hence the necessary information is available for semantic interpretation. There are also more general principles, which serve to restrict transformations.

After some initial remarks on the model assumed for language acquisition in GB theory, we will take a closer look at phrase structure rules, transformations and constraints.

3.1.2 The hypothesis regarding language acquisition: Principles & Parameters

Chomsky (1965: Section I.8) assumes that linguistic knowledge must be innate since the language system is, in his opinion, so complex that it would be impossible to learn a language from the given input using more general cognitive principles alone (also, see Section 13.8). If it is not possible to learn language solely through interaction with our environment, then at least part of our language ability must be innate. The question of exactly what is innate and if humans actually have an innate capacity for language remains controversial and the various positions on the question have changed over the course of the last decades. Some notable works on this topic are Pinker (1994), Tomasello (1995), Wunderlich (2004), Hauser, Chomsky & Fitch (2002) and Chomsky (2007). For more on this discussion, see Chapter 13.

Chomsky (1981a) also assumes that there are general, innate principles which linguistic structure cannot violate. These principles are parameterized, that is, there are options. Parameter settings can differ between languages. An example for a parameterized principle is show in (6):

(6) Principle: A head occurs before or after its complement(s) depending on the value of the parameter POSITION.

The Principles and Parameters model assumes that a significant part of language acquisition consists of extracting enough information from the linguistic input in order to be able to set parameters. Chomsky (2000: 8) compares the setting of parameters to flipping

[2] For more on the power of formal languages, see Chapter 17.

a switch. For a detailed discussion of the various assumptions about language acquisition in the P&P-model, see Chapter 21.6. Speakers of English have to learn that heads occur before their complements in their language, whereas a speaker of Japanese has to learn that heads follow their complements. (7) gives the respective examples:

(7) a. be showing pictures of himself
b. zibun -no syasin-o mise-te iru
REFL from picture showing be

As one can see, the Japanese verb, noun and prepositional phrases are a mirror image of the corresponding phrases in English. (8) provides a summary and shows the parametric value for the position parameter:

(8)	Language	Observation	Parameter: head initial
	English	Heads occur before complements	+
	Japanese	Heads occur after complements	−

Investigating languages based on their differences with regard to certain assumed parameters has proven to be a very fruitful line of research in the last few decades and has resulted in an abundance of comparative cross-linguistic studies.

After these introductory comments on language acquisition, the following sections will discuss the basic assumptions of GB theory.

3.1.3 The T model

Chomsky criticized simple PSGs for not being able to adequately capture certain correlations. An example of this is the relationship between active and passive sentences. In phrase structure grammars, one would have to formulate active and passive rules for intransitive, transitive and ditransitive verbs. The fact that the passive can otherwise be consistently described as the suppression of the most prominent argument is not captured by phrase structure rules. Chomsky therefore assumes that there is an underlying structure, the so-called *Deep Structure*, and that other structures are derived from this. The general architecture of the so-called T model is discussed in the following subsections.

3.1.3.1 D-structure and S-structure

During the derivation of new structures, parts of the Deep Structure can be deleted or moved. In this way, one can explain the relationship between active and passive sentences. As the result of this kind of manipulation of structures, also called transformations, one derives a new structure, the *Surface Structure*, from the original Deep Structure. Since the Surface Structure does not actually mirror the actual use of words in a sentence in some versions of the theory, the term *S-structure* is sometimes used instead as to avoid misunderstandings.

(9) *Surface Structure* = S-structure
Deep Structure = D-structure

Figure 3.2 gives an overview of the GB architecture: phrase structure rules and the lexicon license the D-structure from which the S-structure is derived by means of transformations.

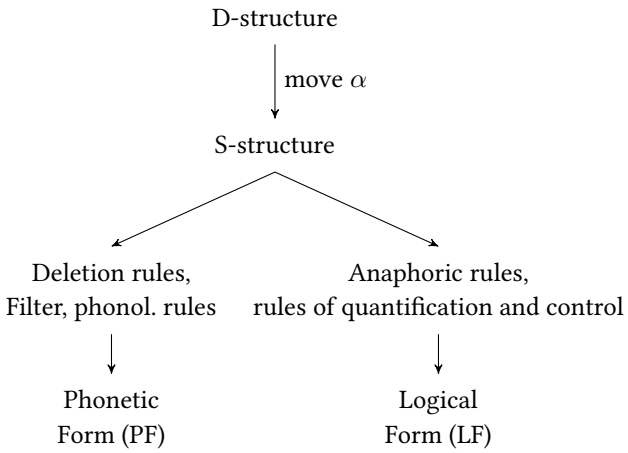


Figure 3.2: The T model

S-structure feeds into Phonetic Form (PF) and Logical Form (LF). The model is referred to as the *T-model* (or Y-model) because D-structure, S-structure, PF and LF form an upside-down T (or Y). We will look at each of these individual components in more detail.

Using phrase structure rules, one can describe the relationships between individual elements (for instance words and phrases, sometimes also parts of words). The format for these rules is $\overline{\text{X}}$ syntax (see Section 2.5). The lexicon, together with the structure licensed by $\overline{\text{X}}$ syntax, forms the basis for D-structure. D-structure is then a syntactic representation of the selectional grid (= valence classes) of individual word forms in the lexicon.

The lexicon contains a lexical entry for every word which comprises information about morphophonological structure, syntactic features and selectional properties. This will be explained in more detail in Section 3.1.3.4. Depending on one's exact theoretical assumptions, morphology is viewed as part of the lexicon. Inflectional morphology is, however, mostly consigned to the realm of syntax. The lexicon is an interface for semantic interpretation of individual word forms.

The surface position in which constituents are realized is not necessarily the position they have in D-structure. For example, there are the following ordering variants for a sentence with a ditransitive verb in (10):

(10) a. [dass] der Mann der Frau das Buch gibt
that the.NOM man the.DAT woman the.ACC book gives
'that the man gives the woman the book'

b. Gibt der Mann der Frau das Buch?
gives the.NOM man the.DAT woman the.ACC book
'Does the man give the woman the book?'

c. Der Mann gibt der Frau das Buch.
the.NOM man gives the.DAT woman the.ACC book
'The man gives the woman the book.'

The following transformational rules for the movements above are assumed: (10b) is derived from (10a) by fronting the verb, and (10c) is derived from (10b) by fronting the nominative noun phrase. In GB theory, there is only one very general transformation: Move-α = "Move anything anywhere!". The nature of what exactly can be moved where and for which reason is determined by principles. Examples of such principles are the Theta Criterion and the Case Filter, which will be dealt with below.

The relations between a predicate and its arguments that are determined by the lexical entries have to be accessible for semantic interpretation at all representational levels. For this reason, the base position of a moved element is marked with a trace. This means for instance that the position in which the fronted *gibt* 'gives' originated is indicated in (11b). The respective marking is referred to as a *trace* or a *gap*. Such empty elements may be frightening when one encounters them first, but I already motivated the assumption of empty elements in nominal structures in Section 2.4.1 (page 70).

(11) a. [dass] der Mann der Frau das Buch gibt
that the man the woman the book gives
'that the man gives the woman the book'

b. Gibt$_i$ der Mann der Frau das Buch $\_i$?
gives the man the woman the book
'Does the man give the woman the book?'

c. [Der Mann]$_j$ gibt$_i$ $\_j$ der Frau das Buch $\_i$.
the man gives the woman the book
'The man gives the woman the book.'

(11c) is derived from (11a) by means of two movements, which is why there are two traces in (11c). The traces are marked with indices so it is possible to distinguish the moved constituents. The corresponding indices are then present on the moved constituents. Sometimes, *e* (for *empty*) or *t* (for *trace*) is used to represent traces.

The S-structure derived from the D-structure is a surface-like structure but should not be equated with the structure of actual utterances.

3.1.3.2 Phonetic Form

Phonological operations are represented at the level of Phonetic Form (PF). PF is responsible for creating the form which is actually pronounced. For example, so-called *wanna*-contraction takes place at PF (Chomsky 1981a: 20–21).

(12) a. The students want to visit Paris.
b. The students wannna visit Paris.

The contraction in (12) is licensed by the optional rule in (13):

(13) want + to → wanna

3.1.3.3 Logical Form

Logical Form is the syntactic level which mediates between S-structure and the semantic interpretation of a sentence. Some of the phenomena which are dealt with by LF are anaphoric reference of pronouns, quantification and control.

Syntactic factors play a role in resolving anaphoric dependencies. An important component of GB theory is Binding Theory, which seeks to explain what a pronoun can or must refer to and when a reflexive pronoun can or must be used. (14) gives some examples of both personal and reflexive pronouns:

(14) a. Peter kauft einen Tisch. Er gefällt ihm.
Peter buys a table.M he likes him
'Peter is buying a table. He likes it/him.'
b. Peter kauft eine Tasche. Er gefällt ihm.
Peter buys a bag.F he likes him
'Peter is buying a bag. He likes it/him.'
c. Peter kauft eine Tasche. Er gefällt sich.
Peter buys a bag.F he likes himself
'Peter is buying a bag. He likes himself.'

In the first example, *er* 'he' can refer to either Peter, the table or something/someone else that was previously mentioned in the context. *ihm* 'him' can refer to Peter or someone in the context. Reference to the table is restricted by world knowledge. In the second example, *er* 'he' cannot refer to *Tasche* 'bag' since *Tasche* is feminine and *er* is masculine. *er* 'he' can refer to Peter only if *ihm* 'him' does not refer to Peter. *ihm* would otherwise have to refer to a person in the wider context. This is different in (14c). In (14c), *er* 'he' and *sich* 'himself' must refer to the same object. This is due to the fact that the reference of reflexives such as *sich* is restricted to a particular local domain. Binding Theory attempts to capture these restrictions.

LF is also important for quantifier scope. Sentences such as (15a) have two readings. These are given in (15b) and (15c).

(15) a. Every man loves a woman.
b. $\forall x \exists y(man(x) \rightarrow (woman(y) \wedge love(x,y)))$
c. $\exists y \forall x(man(x) \rightarrow (woman(y) \wedge love(x,y)))$

The symbol $\forall$ stands for a *universal quantifier* and $\exists$ stands for an *existential quantifier*. The first formula corresponds to the reading that for every man, there is a woman who he loves and in fact, these can be different women. Under the second reading, there is exactly one woman such that all men love her. The question of when such an ambiguity arises and which reading is possible when depends on the syntactic properties of the given utterance. LF is the level which is important for the meaning of determiners such as *a* and *every*.

Control Theory is also specified with reference to LF. Control Theory deals with the question of how the semantic role of the infinitive subject in sentences such as (16) is filled.

(16) a. Der Professor schlägt dem Studenten vor, die Klausur noch mal zu
the professor suggests the student PRT the test once again to
schreiben.
write
'The professor advises the student to take the test again.'
b. Der Professor schlägt dem Studenten vor, die Klausur nicht zu bewerten.
the professor suggests the student PRT the test not to grade
'The professor suggests to the student not to grade the test.'
c. Der Professor schlägt dem Studenten vor, gemeinsam ins Kino zu gehen.
the professor suggests the student PRT together into cinema to go
'The professor suggests to the student to go to the cinema together.'

3.1.3.4 The lexicon

The meaning of words tells us that they have to be combined with certain roles like "acting person" or "affected thing" when creating more complex phrases. For example, the fact that the verb *know* needs two arguments belongs to its semantic contribution. The semantic representation of the contribution of the verb *know* in (17a) is given in (17b):

(17) a. Maria knows the man.
b. *know'*(x,y)

Dividing heads into valence classes is also referred to as *subcategorization*: *know* is subcategorized for a subject and an object. This term comes from the fact that a head is already categorized with regard to its part of speech (verb, noun, adjective, …) and then further subclasses (e. g. intransitive or transitive verb) are formed with regard to valence information. Sometimes the phrase *X subcategorizes for Y* is used, which means *X selects*

Y. know is referred to as the predicate since *know′* is the logical predicate. The subject and object are the arguments of the predicate. There are several terms used to describe the total selectional requirements such as *argument structure*, *valence frames*, *subcategorization frame*, *thematic grid* and *theta grid* or θ-grid.

Adjuncts modify semantic predicates and when the semantic aspect is emphasized they are also called *modifiers*. Adjuncts are not present in the argument structure of predicates.

Following GB assumptions, arguments occur in specific positions in the clause – in so-called argument positions (e. g. the sister of an X^0 element, see Section 2.5). The Theta Criterion states that elements in argument positions have to be assigned a semantic role – a so-called theta role – and each role can only be assigned once (Chomsky 1981a: 36):

Principle 1 (Theta Criterion)

- *Each theta role is assigned to exactly one argument position.*
- *Every phrase in an argument position receives exactly one theta role.*

The arguments of a head are ordered, that is, one can differentiate between higher- and lower-ranked arguments. The highest-ranked argument of verbs and adjectives has a special status. Since GB assumes that it is often (and always in some languages) realized in a position outside of the verb or adjective phrase, it is often referred to as the *external argument*. The remaining arguments occur in positions inside of the verb or adjective phrase. These kind of arguments are dubbed *internal arguments* or *complements*. For simple sentences, this often means that the subject is the external argument.

When discussing types of arguments, one can identify three classes of theta roles:

- Class 1: agent (acting individual), the cause of an action or feeling (stimulus), holder of a certain property
- Class 2: experiencer (perceiving individual), the person profiting from something (beneficiary) (or the opposite: the person affected by some kind of damage), possessor (owner or soon-to-be owner of something, or the opposite: someone who has lost or is lacking something)
- Class 3: patient (affected person or thing), theme

If a verb has several theta roles of this kind to assign, Class 1 normally has the highest rank, whereas Class 3 has the lowest. Unfortunately, the assignment of semantic roles to actual arguments of verbs has received a rather inconsistent treatment in the literature. This problem has been discussed by Dowty (1991), who suggests using proto-roles. An argument is assigned the proto-agent role if it has sufficiently many of the properties that were identified by Dowty as prototypical properties of agents (e. g. animacy, volitionality).

The mental lexicon contains *lexical entries* with the specific properties of syntactic words needed to use that word grammatically. Some of these properties are the following:

- form
- meaning (semantics)
- grammatical features:
  syntactic word class + morphosyntactic features
- theta grid

(18) shows an example of a lexical entry:

(18)

form	*hilft* 'helps'
semantics	*helfen′*
grammatical features	verb, 3rd person singular indicative present active
theta grid	
theta roles	<u>agent</u> beneficiary
grammatical particularities	dative

Assigning semantic roles to specific syntactic requirements (beneficiary = dative) is also called *linking*.

Arguments are ordered according to their ranking: the highest argument is furthest left. In the case of *helfen*, the highest argument is the external argument, which is why the agent is underlined. With so-called unaccusative verbs,[3] the highest argument is not treated as the external argument. It would therefore not be underlined in the corresponding lexical entry.

3.1.4 $\overline{\text{X}}$ theory

In GB, it is assumed that all syntactic structures licensed by the core grammar[4] correspond to the $\overline{\text{X}}$ schema (see Section 2.5).[5] In the following sections, I will comment on the syntactic categories assumed and the basic assumptions with regard to the interpretation of grammatical rules.

[3] See Perlmutter (1978) for a discussion of unaccusative verbs. The term *ergative verb* is also common, albeit a misnomer. See Burzio (1981, 1986) for the earliest work on unaccusatives in the Chomskyan Framework and Grewendorf (1989) for German. Also, see Pullum (1988) on the usage of these terms and for a historical evaluation.

[4] Chomsky (1981a: 7–8) distinguishes between a regular area of language which is determined by a grammar which can be acquired using genetically determined language-specific knowledge and a periphery, to which irregular parts of language such as idioms (e. g. *to pull the wool over sb.'s eyes*) belong. See Section 16.3.

[5] Chomsky (1970: 210) allows for grammatical rules that deviate from the $\overline{\text{X}}$ schema. It is, however, common practice to assume that languages exclusively use $\overline{\text{X}}$ structures.

3.1.4.1 Syntactic categories

The categories which can be used for the variable X in the $\overline{X}$ schema are divided into lexical and functional categories. This correlates roughly with the difference between open and closed word classes. The following are lexical categories:

- V = verb
- N = noun
- A = adjective
- P = preposition/postposition
- Adv = adverb

Lexical categories can be represented using binary features and a cross-classification:[6]

Table 3.1: Representation of four lexical categories using two binary features

	−V	+V
−N	P = [−N, −V]	V = [−N, +V]
+N	N = [+N, −V]	A = [+N, +V]

Adverbs are viewed as intransitive prepositions and are therefore captured by the decomposition in the table above.

Using this cross-classification, it is possible to formulate generalizations. One can, for example, simply refer to adjectives and verbs: all lexical categories which are [+V] are either adjectives or verbs. Furthermore, one can say of [+N] categories (nouns and adjectives) that they can bear case.

Apart from this some authors have tried to associate the head position with the feature values in Table 3.1 (see e. g. Grewendorf 1988: 52; Haftka 1996: 124; G. Müller 2011a: 238). With prepositions and nouns, the head precedes the complement in German:

(19) a. *für* Marie
for Marie

b. *Bild* von Maria
picture of Maria

With adjectives and verbs, the head is final:

[6] See Chomsky (1970: 199) for a cross-classification of N, A and V, and Jackendoff (1977: Section 3.2) for a cross-classification of P, albeit with differing feature assignment.

(20) a. dem König *treu*
the king loyal
'Loyal to the king'
b. der [dem Kind *helfende*] Mann
the the child helping man
'the man helping the child'
c. dem Mann *helfen*
the man help
'help the man'

This data seems to suggest that the head is final with [+V] categories and initial with [−V] categories. Unfortunately, this generalization runs into the problem that there are also postpositions in German. These are, like prepositions, not verbal, but do occur after the NP they require:

(21) a. des Geldes *wegen*
the money because
'because of the money'
b. die Nacht *über*
the night during
'during the night'

Therefore, one must either invent a new category, or abandon the attempt to use binary category features to describe ordering restrictions. If one were to place postpositions in a new category, it would be necessary to assume another binary feature.[7] Since this feature can have either a negative or a positive value, one would then have four additional categories. There are then eight possible feature combinations, some of which would not correspond to any plausible category.

For functional categories, GB does not propose a cross-classification. Usually, the following categories are assumed:

[7] Martin Haspelmath has pointed out that one could assume a rule that moves a post-head argument into a pre-head position (see Riemsdijk 1978: 89 for the discussion of a transformational solution). This would be parallel to the realization of prepositional arguments of adjectives in German:

(i) a. auf seinen Sohn stolz
on his son proud
'proud of his son'
b. stolz auf seinen Sohn
proud of his son

But note that the situation is different with postpositions here, while all adjectives that take prepositional objects allow for both orders, this is not the case for prepositions. Most prepositions do not allow their object to occur before them. It is an idiosyncratic feature of some postpositions that they want to have their argument to the left.

C Complementizer (subordinating conjunctions such as *dass* 'that')
I Finiteness (as well as Tense and Mood);
also Infl in earlier work (inflection),
T in more recent work (Tense)
D Determiner (article, demonstrative)

3.1.4.2 Assumptions and rules

In GB, it is assumed that all rules must follow the $\overline{\text{X}}$ format discussed in Section 2.5. In other theories, rules which correspond to the $\overline{\text{X}}$ format are used along other rules which do not. If the strict version of $\overline{\text{X}}$ theory is assumed, this comes with the assumption of *endocentricity*: every phrase has a head and every head is part of a phrase (put more technically: every head projects to a phrase).

Furthermore, as with phrase structure grammars, it is assumed that the branches of tree structures cannot cross (*Non-Tangling Condition*). This assumption is made by the majority of theories discussed in this book. There are, however, some variants of TAG, HPSG, Construction Grammar, and Dependency Grammar which allow crossing branches and therefore discontinuous constituents (Becker, Joshi & Rambow 1991; Reape 1994; Bergen & Chang 2005;Heringer 1996: 261; Eroms 2000: Section 9.6.2).

In $\overline{\text{X}}$ theory, one normally assumes that there are at most two projection levels (X' and X''). However, there are some versions of Mainstream Generative Grammar and other theories which allow three or more levels (Jackendoff 1977; Uszkoreit 1987). In this chapter, I follow the standard assumption that there are two levels, that is, phrases have at least three projection levels:

- X^0 = head
- X' = intermediate projection (X bar)
- XP = highest projection (= X'' = $\overline{\overline{\text{X}}}$), also called *maximal projection*

3.1.5 CP and IP in English

Most work in Mainstream Generative Grammar is heavily influenced by previous publications dealing with English. If one wants to understand GB analyses of German and other languages, it is important to first understand the analyses of English and, for this reason, this will be the focus of this section. The CP/IP system is also assumed in LFG grammars of English and thus the following section also provides a foundation for understanding some of the fundamentals of LFG presented in Chapter 7.

In earlier work, the rules in (22a) and (22b) were proposed for English sentences (Chomsky 1981a: 19).

(22) a. $\text{S} \rightarrow \text{NP VP}$

b. S → NP Infl VP

Infl stands for *Inflection* as inflectional affixes are inserted at this position in the structure. The symbol AUX was also used instead of Infl in earlier work, since auxiliary verbs are treated in the same way as inflectional affixes. Figure 3.3 shows a sample analysis of a sentence with an auxiliary, which uses the rule in (22b).

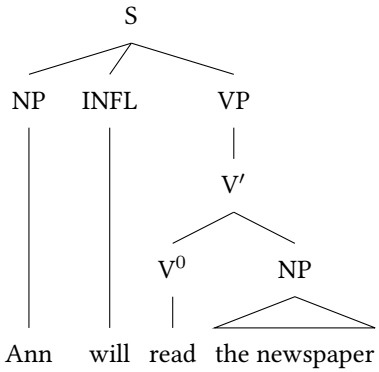


Figure 3.3: English sentence with an auxiliary verb following Chomsky (1981a: 19)

Together with its complements, the verb forms a structural unit: the VP. The constituent status of the VP is supported by several constituent tests and further differences between subjects and objects regarding their positional restrictions.

The rules in (22) do not follow the $\overline{\text{X}}$ template since there is no symbol on the right-hand side of the rule with the same category as one on the left-hand side, that is, there is no head. In order to integrate rules like (22) into the general theory, Chomsky (1986a: 3) developed a rule system with two layers above the verb phrase (VP), namely the CP/IP system. CP stands for *Complementizer Phrase*. The head of a CP can be a complementizer. Before we look at CPs in more detail, I will discuss an example of an IP in this new system. Figure 3.4 on the following page shows an IP with an auxiliary in the I^0 position. As we can see, this corresponds to the structure of the $\overline{\text{X}}$ template: I^0 is a head, which takes the VP as its complements and thereby forms I′. The subject is the specifier of the IP.

The sentences in (23) are analyzed as complementizer phrases (CPs), the complementizer is the head:

(23) a. that Ann will read the newspaper
b. that Ann reads the newspaper

In sentences such as (23), the CPs do not have a specifier. Figure 3.5 on the next page shows the analysis of (23a).

Yes/no-questions in English such as those in (24) are formed by moving the auxiliary verb in front of the subject.

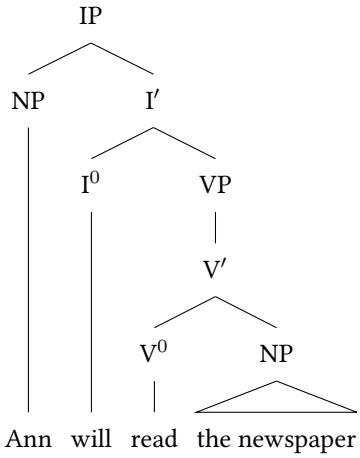


Figure 3.4: English sentence with auxiliary verb in the CP/IP system

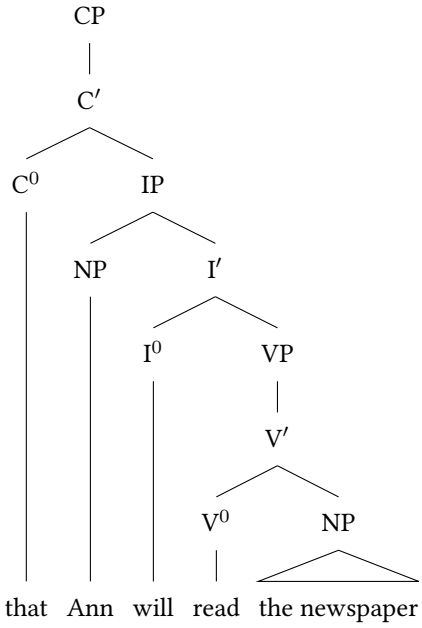


Figure 3.5: English complementizer phrase

(24) Will Ann read the newspaper?

Let us assume that the structure of questions corresponds to the structure of sentences with complementizers. This means that questions are also CPs. Unlike the sentences in (23), however, there is no subordinating conjunction. In the D-structure of questions, the C^0 position is empty and the auxiliary verb is moved to this position. Figure 3.6 shows an analysis of (24).

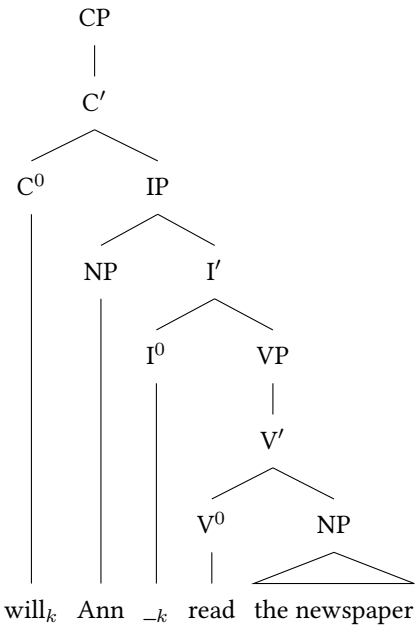


Figure 3.6: English polar question

The original position of the auxiliary is marked by the trace $\__k$, which is coindexed with the moved auxiliary.

wh-questions are formed by the additional movement of a constituent in front of the auxiliary. Figure 3.7 on the next page shows the analysis of (25):

(25) What will Ann read?

As before, the movement of the object of *read* is indicated by a trace. This is important when constructing the meaning of the sentence. The verb assigns some semantic role to the element in its object position. Therefore, one has to be able to "reconstruct" the fact that *what* actually originates in this position. This is ensured by coindexation of the trace with *what*.

Until now, I have not yet discussed sentences without auxiliaries such as (23b). In order to analyze these kinds of sentences, one has to assume that the inflectional affix is

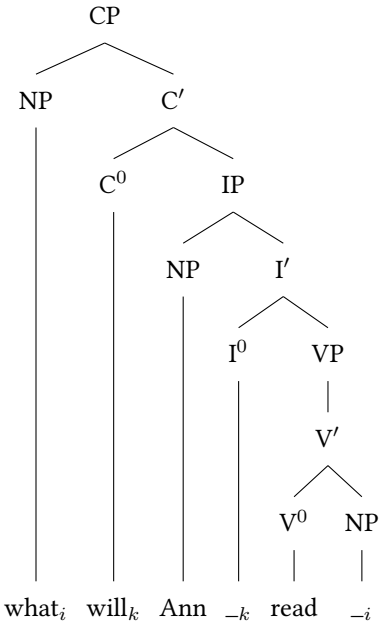


Figure 3.7: English *wh*-question

present in the I^0 position. An example analysis is given in Figure 3.8 on the facing page. Since the inflectional affix precedes the verb, some kind of movement operation still needs to take place. For theory-internal reasons, one does not wish to assume movement operations to positions lower in the tree, hence the verb has to move to the affix and not the other way around.

Following this excursus on the analysis of English sentences, we can now turn to German.

3.1.6 The structure of the German clause

The CP/IP model has been adopted by many scholars for the analysis of German.[8] The categories C, I and V, together with their specifier positions, can be linked to the topological fields as shown in Figure 3.9 on the next page.

Note that SpecCP and SpecIP are not category symbols. They do not occur in grammars with rewrite rules. Instead, they simply describe positions in the tree.

As shown in Figure 3.9, it is assumed that the highest argument of the verb (the subject

[8] For GB analyses without IP, see Bayer & Kornfilt (1989), Höhle (1991a: 157), Haider (1993, 1997a) and Sternefeld (2006: Section IV.3). Haider assumes that the function of I is integrated into the verb. In LFG, an IP is assumed for English (Bresnan 2001: Section 6.2; Dalrymple 2001: Section 3.2.1, but not for German (Berman 2003a: Section 3.2.3.2). In HPSG, no IP is assumed.

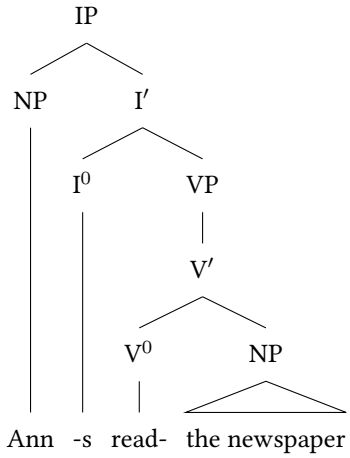
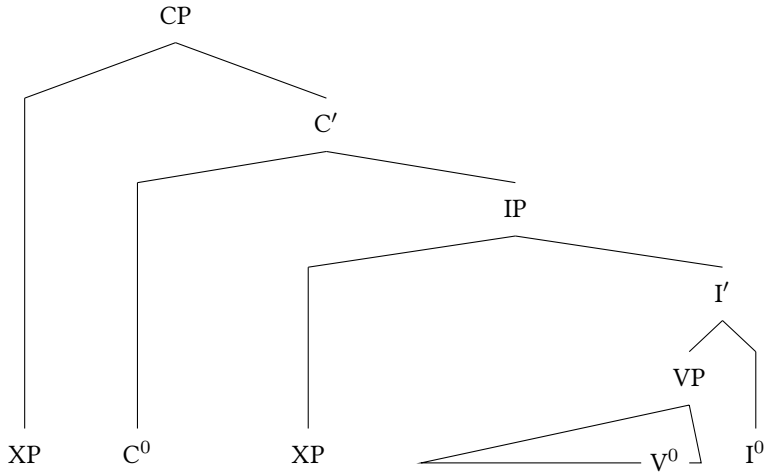


Figure 3.8: English sentence without auxiliary



SpecCP prefield	C^0 left SB	IP (without I^0, V^0) middlefield		V^0, I^0 right SB
		SpecIP subject position	phrases inside the VP	

Figure 3.9: CP, IP and VP and the topological model of German

in simple sentences) has a special status. It is taken for granted that the subject always occurs outside of the VP, which is why it is referred to as the external argument. The VP itself does not have a specifier. In more recent work, however, the subject is generated in the specifier of the VP (Fukui & Speas 1986; Koopman & Sportiche 1991). In some languages, it is assumed that it moves to a position outside of the VP. In other languages such as German, this is the case at least under certain conditions (e. g. definiteness, see Diesing 1992). I am presenting the classical GB analysis here, where the subject is outside the VP. All arguments other than the subject are complements of the V, that are realized within the VP, that is, they are internal arguments. If the verb requires just one complement, then this is the sister of the head V^0 and the daughter of V' according to the $\overline{X}$ schema. The accusative object is the prototypical complement.

Following the $\overline{X}$ template, adjuncts branch off above the complements of V'. The analysis of a VP with an adjunct is shown in Figure 3.10.

(26) weil der Mann morgen den Jungen trifft
because the man tomorrow the boy meets
'because the man is meeting the boy tomorrow'

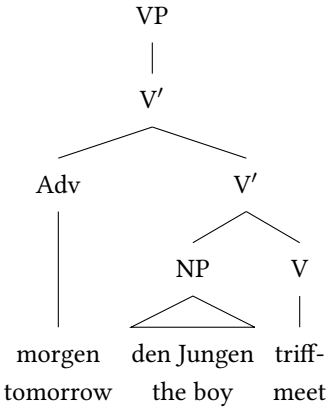


Figure 3.10: Analysis of adjuncts in GB theory

3.2 Verb position

In German, the position of the heads of VP and IP (V^0 and I^0) are to the right of their complements and V^0 and I^0 form part of the right sentence bracket. The subject and all other constituents (complements and adjuncts) all occur to the left of V^0 and I^0 and form the middle field. It is assumed that German – at least in terms of D-structure – is an SOV language (= a language with the base order Subject–Object–Verb). The analysis of German as an SOV language is almost as old as Transformational Grammar itself. It

was originally proposed by Bierwisch (1963: 34).[9] Unlike German, Germanic languages like Danish, English and Romance languages like French are SVO languages, whereas Welsh and Arabic are VSO languages. Around 40 % of all languages belong to the SOV languages, around 35 % are SVO (Dryer 2013c).

The assumption of verb-final order as the base order is motivated by the following observations:[10]

1. Verb particles form a close unit with the verb.

   (27) a. weil er morgen an-fängt
   because he tomorrow PRT-starts
   'because he is starting tomorrow'

   b. Er fängt morgen an.
   he starts tomorrow PRT
   'He is starting tomorrow.'

   This unit can only be seen in verb-final structures, which speaks for the fact that this structure reflects the base order.

   Verbs which are derived from a noun by back-formation (e. g. *uraufführen* 'to perform something for the first time', can often not be divided into their component parts and V2 clauses are therefore ruled out (This was first mentioned by Höhle (1991b) in unpublished work. The first published source is Haider (1993: 62)):

   (28) a. weil sie das Stück heute ur-auf-führen
   because they the play today PREF-PART-lead
   'because they are performing the play for the first time today'

   b. * Sie ur-aufführen heute das Stück.
   they PREF-PART-lead today the play

   c. * Sie führen heute das Stück ur-auf.
   they lead today the play PREF-PART

   The examples show that there is only one possible position for the verb. This order is the one that is assumed to be the base order.

2. Verbs in non-finite clauses and in finite subordinate clauses with a conjunction are always in final position (I am ignoring the possibility of extraposing constituents):

[9] Bierwisch attributes the assumption of an underlying verb-final order to Fourquet (1957). A German translation of the French manuscript cited by Bierwisch can be found in Fourquet (1970: 117–135). For other proposals, see Bach (1962), Reis (1974), Koster (1975) and Thiersch (1978: Chapter 1). Analyses which assume that German has an underlying SOV pattern were also suggested in GPSG (Jacobs 1986: 110), LFG (Berman 1996: Section 2.1.4) and HPSG (Kiss & Wesche 1991; Oliva 1992; Netter 1992; Kiss 1993; Frank 1994; Kiss 1995; Feldhaus 1997, Meurers 2000; Müller 2005b, 2015b).

[10] For points 1 and 2, see Bierwisch (1963: 34–36). For point 4 see Netter (1992: Section 2.3).

(29) a. Der Clown versucht, Kurt-Martin die Ware zu geben.
the clown tries Kurt-Martin the goods to give
'The clown is trying to give Kurt-Martin the goods.'

b. dass der Clown Kurt-Martin die Ware gibt
that the clown Kurt-Martin the goods gives
'that the clown gives Kurt-Martin the goods'

3. If one compares the position of the verb in German to Danish (Danish is an SVO language like English), then one can clearly see that the verbs in German form a cluster at the end of the sentence, whereas they occur before any objects in Danish (Ørsnes 2009a: 146):

(30) a. dass er ihn gesehen$_3$ haben$_2$ muss$_1$
that he him seen have must

b. at han må$_1$ have$_2$ set$_3$ ham
that he must have seen him
'that he must have seen him'

4. The scope relations of the adverbs in (31) depend on their order: the left-most adverb has scope over the two following elements.[11] This was explained by assuming the following structure:

[11] At this point, it should be mentioned that there seem to be exceptions from the rule that modifiers to the left take scope over those to their right. Kasper (1994: 47) discusses examples such as (i), which go back to Bartsch & Vennemann (1972: 137).

(i) a. Peter liest gut wegen der Nachhilfestunden.
Peter reads well because.of the tutoring

b. Peter liest wegen der Nachhilfestunden gut.
Peter reads because.of the tutoring well
'Peter can read well thanks to the tutoring.'

As Koster (1975: Section 6) and Reis (1980: 67) have shown, these are not particularly convincing counterexamples as the right sentence bracket is not filled in these examples and therefore the examples are not necessarily instances of normal reordering inside of the middle field, but could instead involve extraposition of the PP. As noted by Koster and Reis, these examples become ungrammatical if one fills the right bracket and does not extrapose the causal adjunct:

(ii) a. * Hans hat gut wegen der Nachhilfestunden gelesen.
Hans has well because.of the tutoring read

b. Hans hat gut gelesen wegen der Nachhilfestunden.
Hans has well read because.of the tutoring
'Hans has been reading well because of the tutoring.'

However, the following example from Crysmann (2004: 383) shows that, even with the right bracket occupied, one can still have an order where an adjunct to the right has scope over one to the left:

(iii) Da muß es schon erhebliche Probleme mit der Ausrüstung gegeben haben, da wegen
there must it already serious problems with the equipment given have since because.of

(31) a. weil er [absichtlich [nicht lacht]]
because he intentionally not laughs
'because he is intentionally not laughing'

b. weil er [nicht [absichtlich lacht]]
because he not intentionally laughs
'because he is not laughing intentionally'

It is interesting to note that scope relations are not affected by verb position. If one assumes that sentences with verb-second order have the underlying structure in (31), then this fact requires no further explanation. (32) shows the derived S-structure for (31):

(32) a. Er lacht$_i$ [absichtlich [nicht _$_i$]].
he laughs intentionally not
'He is intentionally not laughing.'

b. Er lacht$_i$ [nicht [absichtlich _$_i$]].
he laughs not intentionally
'He is not laughing intentionally.'

After motivating and briefly sketching the analysis of verb-final order, I will now look at the CP/IP analysis of German in more detail. C^0 corresponds to the left sentence bracket and can be filled in two different ways: in subordinate clauses introduced by a conjunction, the subordinating conjunction (the complementizer) occupies C^0 as in English. The verb remains in the right sentence bracket, as illustrated by (33).

(33) dass jeder diesen Mann kennt
that everybody this man knows
'that everybody knows this man'

Figure 3.11 on the next page gives an analysis of (33). In verb-first and verb-second clauses, the finite verb is moved to C^0 via the I^0 position: $V^0 \rightarrow I^0 \rightarrow C^0$. Figure 3.12 on page 109 shows the analysis of (34):

(34) Kennt jeder diesen Mann?
knows everybody this man
'Does everybody know this man?'

---

schlechten Wetters ein Reinhold Messmer niemals aufgäbe.
bad weather a Reinhold Messmer never would.give.up
'There really must have been some serious problems with the equipment because someone like Reinhold Messmer would never give up just because of some bad weather.'

Nevertheless, this does not change anything regarding the fact that the corresponding cases in (31) and (32) have the same meaning regardless of the position of the verb. The general means of semantic composition may well have to be implemented in the same way as in Crysmann's analysis.

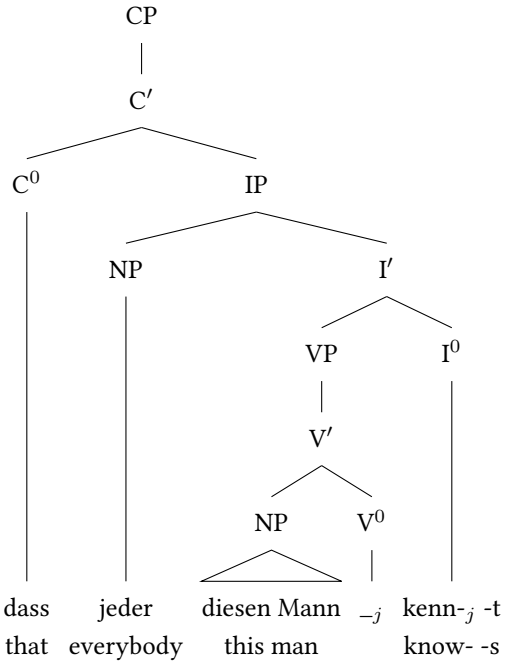


Figure 3.11: Sentence with a complementizer in C^0

The C^0 position is empty in the D-structure of (34). Since it is not occupied by a complementizer, the verb can move there.

3.3 Long-distance dependencies

The SpecCP position corresponds to the prefield and can be filled by any XP in declarative clauses in German. In this way, one can derive the sentences in (36) from (35) by moving a constituent in front of the verb:

(35) Gibt der Mann dem Kind jetzt den Mantel?
gives the.NOM man the.DAT child now the.ACC coat
'Is the man going to give the child the coat now?'

(36) a. Der Mann gibt dem Kind jetzt den Mantel.
the.NOM man gives the.DAT child now the.ACC coat
'The man is giving the child the coat now.'

b. Dem Kind gibt der Mann jetzt den Mantel.
the.DAT child gives the.NOM man now the.ACC coat

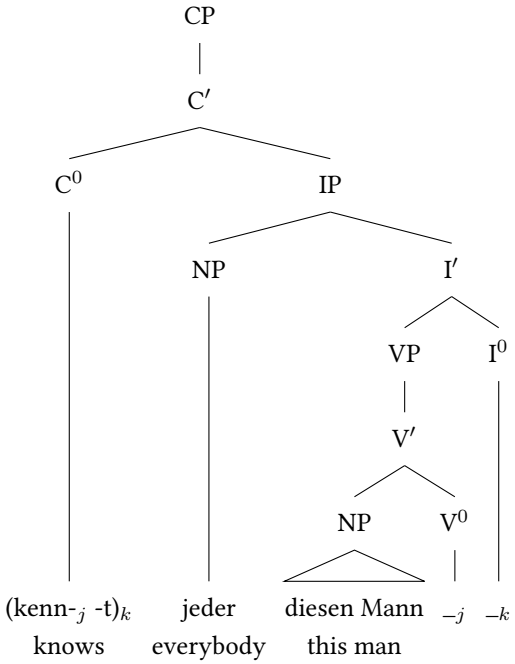


Figure 3.12: Verb position in GB

c. Den Mantel gibt der Mann dem Kind jetzt.
the.ACC coat gives the.NOM man the.DAT child now

d. Jetzt gibt der Mann dem Kind den Mantel.
now gives the.NOM man the.DAT child the.ACC coat

Since any constituent can be placed in front of the finite verb, German is treated typologically as one of the verb-second languages (V2). Thus, it is a verb-second language with SOV base order. English, on the other hand, is an SVO language without the V2 property, whereas Danish is a V2 language with SVO as its base order (see Ørsnes (2009a) for Danish).

Figure 3.13 on the next page shows the structure derived from Figure 3.12.
The crucial factor for deciding which phrase to move is the *information structure* of the sentence, that is, material connected to previously mentioned or otherwise-known information is placed further left (preferably in the prefield) and new information tends to occur to the right. Fronting to the prefield in declarative clauses is often referred to as *topicalization*. But this is rather a misnomer, since the focus (informally: the constituent being asked for) can also occur in the prefield. Furthermore, expletive pronouns can occur there and these are non-referential and as such cannot be linked to preceding or known information, hence expletives can never be topics.

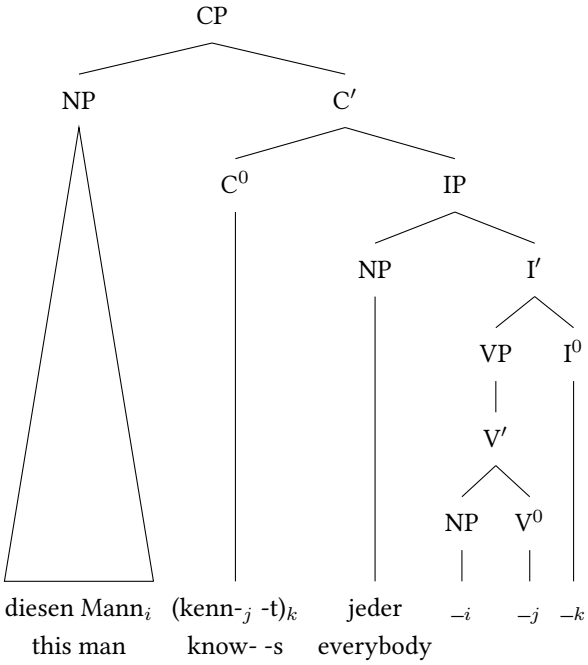


Figure 3.13: Fronting in GB theory

Transformation-based analyses also work for so-called *long-distance dependencies*, that is, dependencies over several phrase boundaries:

(37) a. [Um zwei Millionen Mark]$_i$ soll er versucht haben, [eine
around two million Deutsche.Marks should he tried have an
Versicherung $\__i$ zu betrügen].[12]
insurance.company to deceive
'He apparently tried to cheat an insurance company out of two million Deutsche Marks.'

b. „Wer$_i$, glaubt er, daß er $\__i$ ist?" erregte sich ein Politiker vom Nil.[13]
who believes he that he is retort REFL a politician from.the Nile
' "Who does he think he is?", a politician from the Nile exclaimed.'

c. Wen$_i$ glaubst du, daß ich $\__i$ gesehen habe.[14]
who believe you that I seen have
'Who do you think I saw?'

[12] taz, 04.05.2001, p. 20.
[13] Spiegel, 8/1999, p. 18.
[14] Scherpenisse (1986: 84).

d. [Gegen ihn]$_i$ falle es den Republikanern hingegen schwerer, [[Angriffe
against him fall it the Republicans however more.difficult attacks
_$_i$] zu lancieren].[15]
to launch
'It is, however, more difficult for the Republicans to launch attacks against him.'

The elements in the prefield in the examples in (37) all originate from more deeply embedded phrases. In GB, it is assumed that long-distance dependencies across sentence boundaries are derived in steps (Grewendorf 1988: 75–79), that is, in the analysis of (37c), the interrogative pronoun is moved to the specifier position of the *dass*-clause and is moved from there to the specifier of the matrix clause. The reason for this is that there are certain restrictions on movement which must be checked locally.

3.4 Passive

Before I turn to the analysis of the passive in Section 3.4.2, the first subsection will elaborate on the differences between structural and lexical case.

3.4.1 Structural and lexical case

The case of many case-marked arguments is dependent on the syntactic environment in which the head of the argument is realized. These arguments are referred to as arguments with *structural case*. Case-marked arguments, which do not bear structural case, are said to have *lexical case*.[16]

The following are examples of structural case:[17]

(38) a. Der Installateur kommt.
the.NOM plumber comes
'The plumber is coming.'

b. Der Mann lässt den Installateur kommen.
the man lets the.ACC plumber come
'The man is getting the plumber to come.'

[15] taz, 08.02.2008, p. 9.

[16] Furthermore, there is a so-called *agreeing case* (see page 44) and *semantic case*. Agreeing case is found in predicatives. This case also changes depending on the structure involved, but the change is due to the antecedent element changing its case. Semantic case depends on the function of certain phrases (e. g. temporal accusative adverbials). Furthermore, as with lexical case of objects, semantic case does not change depending on the syntactic environment. For the analysis of the passive, which will be discussed in this section, only structural and lexical case will be relevant.

[17] Compare Heinz & Matiasek (1994: 200).
(38b) is a so-called AcI construction. AcI stands for *Accusativus cum infinitivo*, which means "accusative with infinitive". The logical subject of the embedded verb (*kommen* 'to come' in this case) becomes the accusative object of the matrix verb *lassen* 'to let'. Examples for AcI-verbs are perception verbs such as *hören* 'to hear' and *sehen* 'to see' as well as *lassen* 'to let'.

c. das Kommen des Installateurs
the coming of.the plumber
'the plumber's visit'

In the first example, the subject is in the nominative case, whereas *Installateur* 'plumber' is in accusative in the second example and even in the genitive in the third following nominalization. The accusative case of objects is normally structural case. This case becomes nominative under passivization:

(39) a. Karl schlägt den Weltmeister.
Karl beats the.ACC world.champion
'Karl beats the world champion.'

b. Der Weltmeister wird geschlagen.
the.NOM world.champion is beaten
'The world champion is being beaten.'

Unlike the accusative, the genitive governed by a verb is a lexical case. The case of a genitive object does not change when the verb is passivized.

(40) a. Wir gedenken der Opfer.
we remember the.GEN victims

b. Der Opfer wird gedacht.
the.GEN victims are remembered
'The victims are being remembered.'

(40b) is an example of the so-called *impersonal passive*. Unlike example (39b), where the accusative object became the subject, there is no subject in (40b). See Section 1.7.1.

Similarly, there is no change in case with dative objects:

(41) a. Der Mann hat ihm geholfen.
the man has him.DAT helped
'The man has helped him.'

b. Ihm wird geholfen.
him.DAT is helped
'He is being helped.'

It still remains controversial as to whether some or all of the datives in verbal environments should be treated as instances of structural case. For reasons of space, I will not recount this discussion but instead refer the interested reader to Chapter 14 of Müller (2007b). In what follows, I assume – like Haider (1986a: 20) – that the dative is in fact a lexical case.

3.4.2 Case assignment and the Case Filter

In GB, it is assumed that the subject receives case from (finite) I and that the case of the remaining arguments comes from V (Chomsky 1981a: 50; Haider 1984: 26; Fanselow & Felix 1987: 71–73).

Principle 2 (Case Principle)

- *V assigns objective case (accusative) to its complement if it bears structural case.*
- *When finite, INFL assigns case to the subject.*

The Case Filter rules out structures where case has not been assigned to an NP.

Figure 3.14 on the following page shows the Case Principle in action with the example in (42a).[18]

(42) a. [dass] der Mann der Frau den Jungen zeigt
that the man the.DAT woman the.ACC boy shows
'that the man shows the boy to the woman'

b. [dass] der Junge der Frau gezeigt wird
that the boy.NOM the.DAT woman shown is
'that the boy is shown to the woman'

The passive morphology blocks the subject. The object that would get accusative in the active receives a semantic role in passives, but it does not get case. Therefore, it has to move to a position where case can be assigned to it (Chomsky 1981a: 124). Figure 3.15 on the next page shows how this works for example (42b).

This movement-based analysis works well for English since the underlying object always has to be moved:

(43) a. The mother gave [the girl] [a cookie].
b. [The girl] was given [a cookie] (by the mother).
c. * It was given [the girl] [a cookie].

(43c) shows that filling the subject position with an expletive is not possible, so the object really has to move. However, Lenerz (1977: Section 4.4.3) showed that such a movement is not obligatory in German:

(44) a. weil das Mädchen dem Jungen den Ball schenkte
because the.NOM girl the.DAT boy the.ACC ball gave
'because the girl gave the ball to the boy'

[18] The figure does not correspond to $\overline{X}$ theory in its classic form, since *der Frau* 'the woman' is a complement which is combined with V′. In classical $\overline{X}$ theory, all complements have to be combined with V^0. Furthermore, in the following figures the verb has been left in V^0 for reasons of clarity. In order to create a well-formed S-structure, the verb would have to move to its affix in I^0.

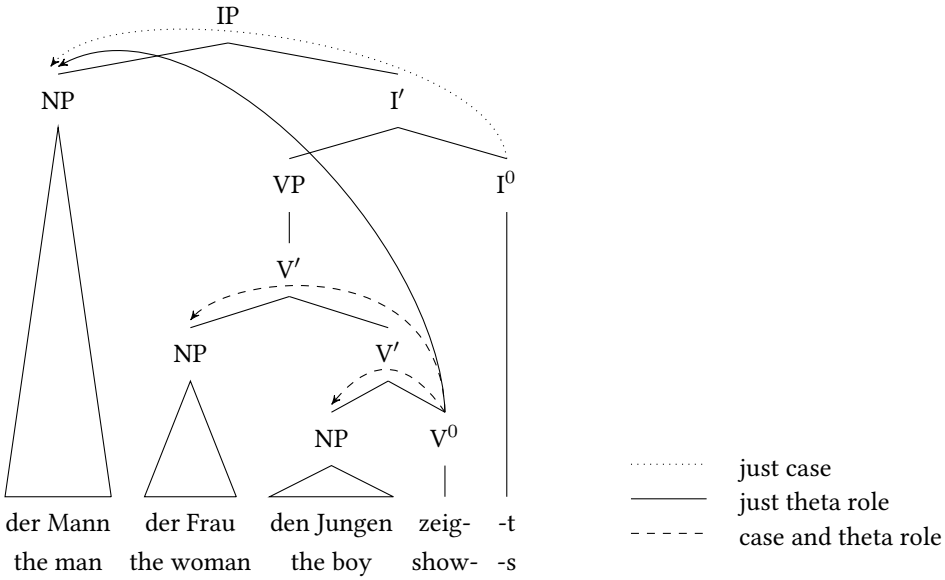


Figure 3.14: Case and theta-role assignment in active clauses

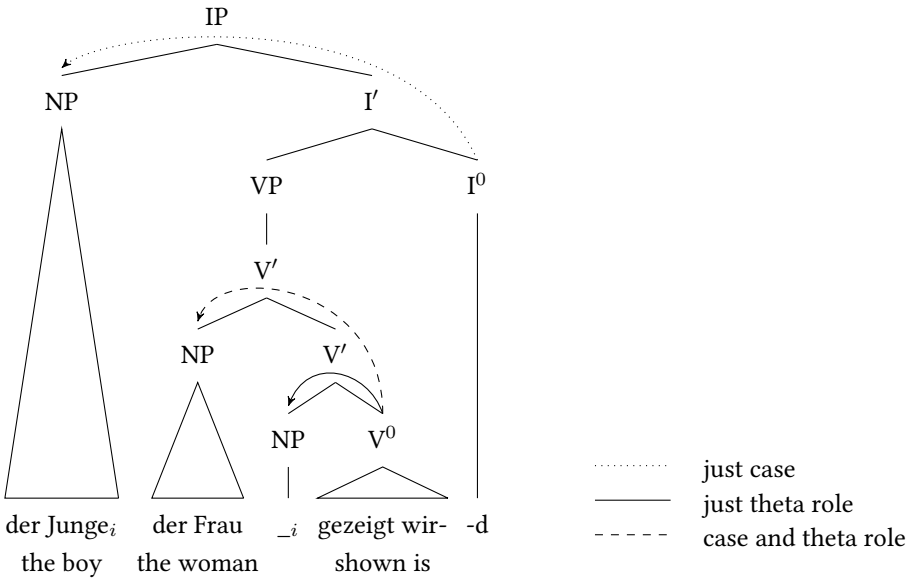


Figure 3.15: Case and theta-role assignment in passive clauses

b. weil dem Jungen der Ball geschenkt wurde
because the.DAT boy the.NOM ball given was
'because the ball was given to the boy'

c. weil der Ball dem Jungen geschenkt wurde
because the.NOM ball the.DAT boy given was

In comparison to (44c), (44b) is the unmarked order. *der Ball* 'the ball' in (44b) occurs in the same position as *den Ball* in (44a), that is, no movement is necessary. Only the case differs. (44c) is, however, somewhat marked in comparison to (44b). The analysis which has been proposed for cases such as (44b) involves abstract movement: the elements stay in their positions, but are connected to the subject position and receive their case information from there. Grewendorf (1993: 1311) assumes that there is an empty expletive pronoun in the subject position of sentences such as (44b) as well as in the subject position of sentences with an impersonal passive such as (45):[19]

(45) weil heute nicht gearbeitet wird
because today not worked is
'because there will be no work done today'

A silent expletive pronoun is something that one cannot see or hear and that does not carry any meaning. For discussion of these kind of empty elements, see Section 13.1.3 and Chapter 19.

In the following chapters, I describe alternative treatments of the passive that do without mechanisms such as empty elements that are connected to argument positions and seek to describe the passive in a more general, cross-linguistically consistent manner as the suppression of the most prominent argument.

A further question which needs to be answered is why the accusative object does not receive case from the verb. This is captured by a constraint, which goes back to Burzio (1986: 178–185) and is therefore referred to as *Burzio's Generalization*.[20]

[19] See Koster (1986: 11–12) for a parallel analysis for Dutch as well as Lohnstein (2014) for a movement-based account of the passive that also involves an empty expletive for the analysis of the impersonal passive.

[20] Burzio's original formulation was equivalent to the following: a verb assigns accusative, if and only if it assigns a semantic role to its subject.

This claim is problematic from both sides. In (i), the verb does not assign a semantic role to the subject, however there is nevertheless accusative case:

(i) Mich friert.
me.ACC freezes
'I am freezing.'

One therefore has to differentiate between structural and lexical accusative and modify Burzio's Generalization accordingly. The existence of verbs like *begegnen* 'to bump into' is problematic for the other side of the implication. *begegnen* has a subject but still does not assign accusative but rather dative:

(ii) Peter begegnete einem Mann.
Peter met a.DAT man
'Peter met a man.'

(46) Burzio's Generalization (modified):
If V does not have an external argument, then it does not assign (structural) accusative case.

Koster (1986: 12) has pointed out that the passive in English cannot be derived by Case Theory since if one allowed empty expletive subjects for English as well as German and Dutch, then it would be possible to have analyses such as the following in (47) where np is an empty expletive:

(47) np was read the book.

Koster rather assumes that subjects in English are either bound by other elements (that is, non-expletive) or lexically filled, that is, filled by visible material. Therefore, the structure in (47) would be ruled out and it would be ensured that *the book* would have to be placed in front of the finite verb so that the subject position is filled.

3.5 Local reordering

Arguments in the middle field can, in principle, occur in an almost arbitrary order. (48) exemplifies this:

(48) a. [weil] der Mann der Frau das Buch gibt
because the man the woman the book gives
'because the man gives the book to the woman'
b. [weil] der Mann das Buch der Frau gibt
because the man the book the woman gives
c. [weil] das Buch der Mann der Frau gibt
because the book the man the woman gives
d. [weil] das Buch der Frau der Mann gibt
because the book the woman the man gives
e. [weil] der Frau der Mann das Buch gibt
because the woman the man the book gives
f. [weil] der Frau das Buch der Mann gibt
because the woman the book the man gives

Burzio (1986: 185) assumes that one-place intransitive verbs have the potential to assign accusative. He supports this claim by pointing out the existence of the resultative constructions, in which additional accusatives can be realized. (iii) is an example:

(iii) He talked my head off.

However, there are also verbs such as *verschwinden* 'to disappear' which never assign accusative, not even in such constructions.

See Haider (1999) and Webelhuth (1995: 89) as well as the references cited there for further problems with Burzio's Generalization.

In (48b–f), the constituents receive different stress and the number of contexts in which each sentence can be uttered is more restricted than (48a) (Höhle 1982). The order in (48a) is therefore referred to as the *neutral order* or *unmarked order*.

Two proposals have been made for analyzing these orders: the first suggestion assumes that the five orderings in (48b–f) are derived from a single underlying order by means of move α (Frey 1993). As an example, the analysis of (48c) is given in Figure 3.16. The accusative object *das Buch* 'the book' is moved to the left and adjoined to the top-

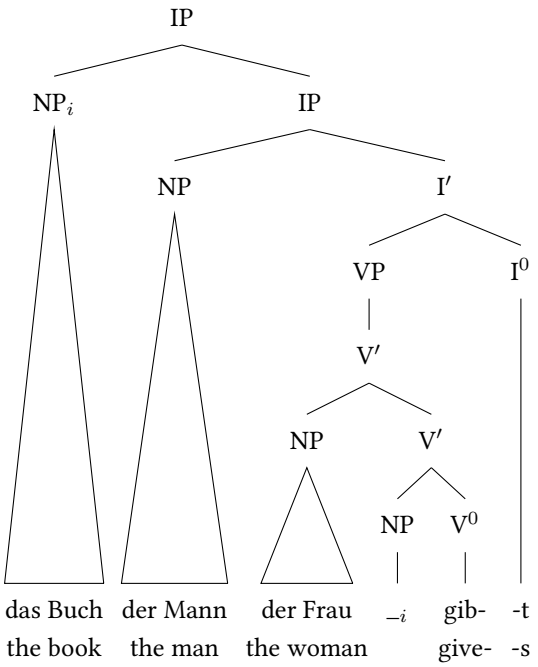


Figure 3.16: Analysis of local reordering as adjunction to IP

most IP.

An argument that has often been used to support this analysis is the fact that scope ambiguities exist in sentences with reorderings which are not present in sentences in the base order. The explanation of such ambiguities comes from the assumption that the scope of quantifiers can be derived from their position in the surface structure as well as their position in the deep structure. If the position in both the surface and deep structure are the same, that is when there has not been any movement, then there is only one reading possible. If movement has taken place, however, then there are two possible readings (Frey 1993):

(49) a. Es ist nicht der Fall, daß er mindestens einem Verleger fast jedes Gedicht
it is not the case that he at.least one publisher almost every poem

anbot.
offered

'It is not the case that he offered at least one publisher almost every poem.'

b. Es ist nicht der Fall, daß er fast jedes Gedicht$_i$ mindestens einem
it is not the case that he almost every poem at.least one
Verleger $\__i$ anbot.
publisher offered

'It is not the case that he offered almost every poem to at least one publisher.'

It turns out that approaches assuming traces run into problems as they predict certain readings for sentences with multiple traces which do not exist (see Kiss 2001: 146 and Fanselow 2001: Section 2.6). For instance in an example such as (50), it should be possible to interpret *mindestens einem Verleger* 'at least one publisher' at the position of $\__i$, which would lead to a reading where *fast jedes Gedicht* 'almost every poem' has scope over *mindestens einem Verleger* 'at least one publisher'.

(50) Ich glaube, dass mindestens einem Verleger$_i$ fast jedes Gedicht$_j$ nur dieser
I believe that at.least one publisher almost every poem only this
Dichter $\__i$ $\__j$ angeboten hat.
poet offered has

'I think that only this poet offered almost every poem to at least one publisher.'

This reading does not exist, however.

Sauerland & Elbourne (2002: 308) discuss analogous examples from Japanese, which they credit to Kazuko Yatsushiro. They develop an analysis where the first step is to move the accusative object in front of the subject. Then, the dative object is placed in front of that and then in a third movement, the accusative is then moved once more. The last movement can take place to construct either the S-structure[21] or as a movement to construct the phonological form. In the latter case, this movement will not have any semantic effects. While this analysis can predict the correct available readings, it does require a number of additional movement operations with intermediate steps.

The alternative to a movement analysis is so-called *base generation*: the starting structure generated by phrase structure rules is referred to as the *base*. One variant of base generation assumes that the verb is combined with one argument at the time and that θ-roles are assigned in parallel to head-argument combinations. The order in which arguments are combined with the verb is not specified, which means that all of the orders in (48) can be generated directly without any transformations (compare this to the grammar in (6) on page 57). This kind of analysis has been proposed for GB by Fanselow (2001).[22] For the discussion of different approaches to describing constituent position, see Fanselow (1993).

[21] The authors are working in the Minimalist framework. This means there is no longer S-structure strictly speaking. I have simply translated the analysis into the terms used here.

[22] The base generation analysis is the natural analysis in the HPSG framework. It has already been developed by Gunji in 1986 for Japanese and will be discussed in more detail in Section 9.4. Sauerland & Elbourne (2002: 313–314) claim that they show that syntax has to be derivational, that is, a sequence of syntactic

3.6 Summary and classification

Works in GB and some contributions to the Minimalist Program (see Chapter 4) have led to a number of new discoveries in both language-specific and cross-linguistic research. In the following, I will focus on some aspects of German syntax.

The analysis of verb movement developed in Transformational Grammar by Bierwisch (1963: 34), Reis (1974), Koster (1975), Thiersch (1978: Chapter 1) and den Besten (1983) has become the standard analysis in almost all grammar models (possibly with the exception of Construction Grammar and Dependency Grammar).

The work by Lenerz on constituent order (1977) has influenced analyses in other frameworks (the linearization rules in GPSG and HPSG go back to Lenerz' descriptions). Haider's work on constituent order, case and passive (1984; 1985b; 1985a; 1986a; 1990b; 1993) has had a significant influence on LFG and HPSG analyses of German.

The entire configurationality discussion, that is, whether it is better to assume that the subject of finite verbs in German is inside or outside the VP, was important (for instance Haider 1982; Grewendorf 1983; Kratzer 1984; Webelhuth 1985; Sternefeld 1985b; Scherpenisse 1986; Fanselow 1987; Grewendorf 1988; Dürscheid 1989; Webelhuth 1990; Oppenrieder 1991; Wilder 1991; Haider 1993; Grewendorf 1993; Frey 1993; Lenerz 1994; Meinunger 2000) and German unaccusative verbs received their first detailed discussion in GB circles (Grewendorf 1989; Fanselow 1992a). The works by Fanselow and Frey on constituent order, in particular with regard to information structure, have advanced German syntax quite considerably (Fanselow 1988, 1990, 1993, 2000a, 2001, 2003b,c, 2004a; Frey 2000, 2001, 2004a, 2005). Infinitive constructions, complex predicates and partial fronting have also received detailed and successful treatments in the GB/MP frameworks (Bierwisch 1963; Evers 1975; Haider 1982, 1986b, 1990a, 1991, 1993; Grewendorf 1983, 1987, 1988; den Besten 1985; Sternefeld 1985b; Fanselow 1987, 2002; von Stechow & Sternefeld 1988; Bayer & Kornfilt 1989, G. Müller 1996a, 1998; Vogel & Steinbach 1998). In the area of secondary predication, the work by Susanne Winkler (1997) is particularly noteworthy.

This list of works from subdisciplines of grammar is somewhat arbitrary (it corresponds more or less to my own research interests) and is very much focused on German. There are, of course, a wealth of other articles on other languages and phenomena, which should be recognized without having to be individually listed here.

In the remainder of this section, I will critically discuss two points: the model of language acquisition of the Principles & Parameters model and the degree of formalization inside Chomskyan linguistics (in particular the last few decades and the consequences this has). Some of these points will be mentioned again in Part II.

3.6.1 Explaining language acquisition

One of the aims of Chomskyan research on grammar is to explain language acquisition. In GB, one assumed a very simple set of rules, which was the same for all languages

trees has to be derived. I am of the opinion that this cannot generally be shown to be the case. There is, for example, an analysis by Kiss (2001) which shows that scope phenomena can be explained well by constraint-based approaches.

($\overline{X}$ theory) as well as general principles that hold for all languages, but which could be parameterized for individual languages or language classes. It was assumed that a parameter was relevant for multiple phenomena. The Principles & Parameters model was particularly fruitful and led to a number of interesting studies in which commonalities and differences between languages were uncovered. From the point of view of language acquisition, the idea of a parameter which is set according to the input has often been cricitized as it cannot be reconciled with observable facts: after setting a parameter, a learner would have to immediately have mastered certain aspects of that language. Chomsky (1986b: 146) uses the metaphor of switches which can be flipped one way or the other. As it is assumed that various areas of grammar are affected by parameters, setting one parameter should have a significant effect on the rest of the grammar of a given learner. However, the linguistic behaviour of children does not change in an abrupt fashion as would be expected (Bloom 1993: 731; Haider 1993: 6; Abney 1996: 3; Ackerman & Webelhuth 1998: Section 9.1; Tomasello 2000, 2003; Newmeyer 2005). Furthermore, it has not been possible to prove that there is a correlation between a certain parameter and various grammatical phenomena. For more on this, see Chapter 16.

The Principles and Parameters model nevertheless remains interesting for cross-linguistic research. Every theory has to explain why the verb precedes its objects in English and follows them in Japanese. One can name this difference a parameter and then classify languages accordingly, but whether this is actually relevant for language acquisition is being increasingly called in question.

3.6.2 Formalization

In his 1963 work on Transformational Grammar, Bierwisch writes the following:[23]

> It is very possible that the rules that we formulated generate sentences which are outside of the set of grammatical sentences in an unpredictable way, that is, they violate grammaticality due to properties that we did not deliberately exclude in our examination. This is meant by the statement that a grammar is a hypothesis about the structure of a language. A systematic check of the implications of a grammar that is appropriate for natural languages is surely a task that cannot be done by hand any more. This task could be solved by implementing the grammar as a calculating task on a computer so that it becomes possible to verify to which degree the result deviates from the language to be described. (Bierwisch 1963: 163)

[23] Es ist also sehr wohl möglich, daß mit den formulierten Regeln Sätze erzeugt werden können, die auch in einer nicht vorausgesehenen Weise aus der Menge der grammatisch richtigen Sätze herausfallen, die also durch Eigenschaften gegen die Grammatikalität verstoßen, die wir nicht wissentlich aus der Untersuchung ausgeschlossen haben. Das ist der Sinn der Feststellung, daß eine Grammatik eine Hypothese über die Struktur einer Sprache ist. Eine systematische Überprüfung der Implikationen einer für natürliche Sprachen angemessenen Grammatik ist sicherlich eine mit Hand nicht mehr zu bewältigende Aufgabe. Sie könnte vorgenommen werden, indem die Grammatik als Rechenprogramm in einem Elektronenrechner realisiert wird, so daß überprüft werden kann, in welchem Maße das Resultat von der zu beschreibenden Sprache abweicht.

Bierwisch's claim is even more true in light of the empirical progress made in the last decades. For example, Ross (1967) identified restrictions for movement and long-distance dependencies and Perlmutter (1978) discovered unaccusative verbs in the 70s. For German, see Grewendorf (1989) and Fanselow (1992a). Aside from analyses of these phenomena, restrictions on possible constituent positions have been developed (Lenerz 1977), as well as analyses of case assignment (Yip, Maling & Jackendoff 1987; Meurers 1999c; Przepiórkowski 1999b) and theories of verbal complexes and the fronting of parts of phrases (Evers 1975; Grewendorf 1988; Hinrichs & Nakazawa 1994; Kiss 1995; G. Müller 1998; Meurers 1999b; Müller 1999a, 2002a; De Kuthy 2002). All these phenomena interact!

Consider another quote:

> A goal of earlier linguistic work, and one that is still a central goal of the linguistic work that goes on in computational linguistics, is to develop grammars that assign a reasonable syntactic structure to every sentence of English, or as nearly every sentence as possible. This is not a goal that is currently much in fashion in theoretical linguistics. Especially in Government-Binding theory (GB), the development of large fragments has long since been abandoned in favor of the pursuit of deep principles of grammar. The scope of the problem of identifying the correct parse cannot be appreciated by examining behavior on small fragments, however deeply analyzed. Large fragments are not just small fragments several times over – there is a qualitative change when one begins studying large fragments. As the range of constructions that the grammar accommodates increases, the number of undesired parses for sentences increases dramatically. (Abney 1996: 20)

So, as Bierwisch and Abney point out, developing a sound theory of large fragment of a human language is a really demanding task. But what we aim for as theoretical linguists is much more: the aim is to formulate restrictions which ideally hold for all languages or at least for certain language classes. It follows from this, that one has to have an overview of the interaction of various phenomena in not just one but several languages. This task is so complex that individual researchers cannot manage it. This is the point at which computer implementations become helpful as they immediately flag inconsistencies in a theory. After removing these inconsistencies, computer implementations can be used to systematically analyze test data or corpora and thereby check the empirical adequacy of the theory (Müller, 1999a: Chapter 22; 2015a; 2014d; Oepen & Flickinger 1998; Bender 2008b, see Section 1.2).

More than 50 years after the first important published work by Chomsky, it is apparent that there has not been one large-scale implemented grammatical fragment on the basis of Transformational Grammar analyses. Chomsky has certainly contributed to the formalization of linguistics and developed important formal foundations which are still relevant in the theory of formal languages in computer science and in theoretical computational linguistics (Chomsky 1959). However, in 1981, he had already turned his back on rigid formalization:

> I think that we are, in fact, beginning to approach a grasp of certain basic principles of grammar at what may be the appropriate level of abstraction. At the same

> time, it is necessary to investigate them and determine their empirical adequacy by developing quite specific mechanisms. We should, then, try to distinguish as clearly as we can between discussion that bears on leading ideas and discussion that bears on the choice of specific realizations of them. (Chomsky 1981a: 2–3)

This is made explicit in a letter to *Natural Language and Linguistic Theory*:

> Even in mathematics, the concept of formalization in our sense was not developed until a century ago, when it became important for advancing research and understanding. I know of no reason to suppose that linguistics is so much more advanced than 19th century mathematics or contemporary molecular biology that pursuit of Pullum's injunction would be helpful, but if that can be shown, fine. For the present, there is lively interchange and exciting progress without any sign, to my knowledge, of problems related to the level of formality of ongoing work. (Chomsky 1990: 146)

This departure from rigid formalization has led to there being a large number of publications inside Mainstream Generative Grammar with sometimes incompatible assumptions to the point where it is no longer clear how one can combine the insights of the various publications. An example of this is the fact that the central notion of government has several different definitions (see Aoun & Sportiche (1983) for an overview[24]).

This situation has been cricitized repeatedly since the 80s and sometimes very harshly by proponents of GPSG (Gazdar, Klein, Pullum & Sag 1985: 6; Pullum 1985, 1989a; Pullum 1991: 48; Kornai & Pullum 1990).

The lack of precision and working out of the details[25] and the frequent modification of basic assumptions[26] has led to insights gained by Mainstream Generative Grammar rarely being translated into computer implementations. There are some implementations that are based on Transformational Grammar/GB/MP models or borrow ideas from Mainstream Generative Grammar (Petrick 1965; Zwicky, Friedman, Hall & Walker 1965; Kay 1967; Friedman 1969; Friedman, Bredt, Doran, Pollack & Martner 1971; Morin 1973; Marcus 1980; Abney & Cole 1986; Kuhns 1986; Correa 1987; Stabler 1987, 1992, 2001; Kolb & Thiersch 1991; Fong 1991; Crocker & Lewin 1992; Lohnstein 1993; Fordham & Crocker 1994; Nordgård 1994; Veenstra 1998; Fong & Ginsburg 2012),[27] but these implementations often do not use transformations or differ greatly from the theoretical assumptions of the publications. For example, Marcus (1980: 102–104) and Stabler (1987: 5) use special purpose rules for auxiliary movement.[28] These rules reverse the order of *John* and *has*

[24] A further definition can be found in Aoun & Lightfoot (1984). This is, however, equivalent to an earlier version as shown by Postal & Pullum (1986: 104–106).

[25] See e. g. Kuhns (1986: 550), Crocker & Lewin (1992: 508), Kolb & Thiersch (1991: 262), Kolb (1997: 3) and Freidin (1997: 580), Veenstra (1998: 25, 47), Lappin et al. (2000a: 888) and Stabler (2010: 397, 399, 400) for the latter.

[26] See e. g. Kolb (1997: 4), Fanselow (2009) and the quote from Stabler on page 178.

[27] See Fordham & Crocker (1994) for a combination of a GB approach with statistical methods.

[28] Nozohoor-Farshi (1986, 1987) has shown that Marcus' parser can only parse context-free languages. Since natural languages are of a greater complexity (see Chapter 17) and grammars of corresponding complexity are allowed by current versions of Transformational Grammar, Marcus' parser can be neither an adequate

for the analysis of sentences such as (51a) so that we get the order in (51b), which is then parsed with the rules for non-inverted structures.

(51) a. Has John scheduled the meeting for Wednesday?
b. John has scheduled the meeting for Wednesday?

These rules for auxiliary movement are very specific and explicitly reference the category of the auxiliary. This does not correspond to the analyses proposed in GB in any way. As we have seen in Section 3.1.5, there are no special transformational rules for auxiliary movement. Auxiliary movement is carried out by the more general transformation move α and the associated restrictive principles. It is not unproblematic that the explicit formulation of the rule refers to the category *auxiliary* as is clear when one views Stabler's GB-inspired phrase structure grammar:

(52) a. s → switch(aux_verb,np), vp.
b. s([First|L0],L,X0,X) :- aux_verb(First),
np(L0,L1,X0,X1),
vp([First|L1],L,X1,X).

The rule in (52a) is translated into the Prolog predicate in (52b). The expression [First|L0] after the s corresponds to the string, which is to be processed. The '|'-operator divides the list into a beginning and a rest. *First* is the first word to be processed and L0 contains all other words. In the analysis of (51a), First is *has* and L0 is *John scheduled the meeting for Wednesday*. In the Prolog clause, it is then checked whether First is an auxiliary (aux_verb(First)) and if this is the case, then it will be tried to prove that the list L0 begins with a noun phrase. Since *John* is an NP, this is successful. L1 is the sublist of L0 which remains after the analysis of L0, that is *scheduled the meeting for Wednesday*. This list is then combined with the auxiliary (First) and now it will be checked whether the resulting list *has scheduled the meeting for Wednesday* begins with a VP. This is the case and the remaining list L is empty. As a result, the sentence has been successfully processed.

The problem with this analysis is that exactly one word is checked in the lexicon. Sentences such as (53) can not be analyzed:[29]

(53) Could or should we pool our capital with that of other co-ops to address the needs of a regional "neighborhood"?[30]

In this kind of sentence, two modal verbs have been coordinated. They then form an X^0 and – following GB analyses – can be moved together. If one wanted to treat these cases as Stabler does for the simplest case, then we would need to divide the list of words to be processed into two unlimited sub-lists and check whether the first list contains an auxiliary or several coordinated auxiliaries. We would require a recursive predicate

implementation of the Chomskyan theory in question nor a piece of software for analyzing natural language in general.

[29] For a discussion that shows that the coordination of lexical elements has to be an option in linguistic theories, see Abeillé (2006).

[30] http://www.cooperativegrocer.coop/articles/index.php?id=595. 28.03.2010.

aux_verbs which somehow checks whether the sequence *could or should* is a well-formed sequence of auxiliaries. This should not be done by a special predicate but rather by syntactic rules responsible for the coordination of auxiliaries. The alternative to a rule such as (52a) would be the one in (54), which is the one that is used in theories like GPSG (Gazdar et al. 1985: 62), LFG (Falk 1984: 491), some HPSG analyses (Ginzburg & Sag 2000: 36), and Construction Grammar (Fillmore 1999):

(54) s → v(aux+), np, vp.

This rule would have no problems with coordination data like (53) as coordination of multiple auxiliaries would produce an object with the category v(aux+) (for more on coordination see Section 21.6.2). If inversion makes it necessary to stipulate a special rule like (52a), then it is not clear why one could not simply use the transformation-less rule in (54).

In the MITRE system (Zwicky et al. 1965), there was a special grammar for the surface structure, from which the deep structure was derived via reverse application of transformations, that is, instead of using one grammar to create deep structures which are then transformed into other structures, one required two grammars. The deep structures that were determined by the parser were used as input to a transformational component since this was the only way to ensure that the surface structures can actually be derived from the base structure (Kay 2011: 10).

There are other implementations discussed in this chapter that differ from transformation-based analyses. For example, Kolb & Thiersch (1991: 265, Section 4) arrive at the conclusion that a declarative, constraint-based approach to GB is more appropriate than a derivational one. Johnson (1989) suggests a *Parsing as Deduction* approach which reformulates sub-theories of GB ($\overline{X}$ theory, Theta-Theory, Case Theory, …) as logical expressions.[31] These can be used independently of each other in a logical proof. In Johnson's analysis, GB theory is understood as a constraint-based system. More general restrictions are extracted from the restrictions on S- and D-structure which can then be used directly for parsing. This means that transformations are not directly carried out by the parser. As noted by Johnson, the language fragment he models is very small. It contains no description of *wh*-movement, for example (p. 114).

Probably the most detailed implementation of the GB/Barriers tradition is Stabler's Prolog implementation (1992). Stabler's achievement is certainly impressive, but his book confirms what has been claimed thus far: Stabler has to simply stipulate many things which are not explicitly mentioned in *Barriers* (e. g. using feature-value pairs when formalizing $\overline{X}$ theory, which was borrowed from GPSG) and some assumptions cannot be properly formalized and are simply ignored (see Briscoe (1997) for details).

GB analyses which fulfill certain requirements can be reformulated so that they no longer make use of transformations. These transformation-less approaches are also called *representational*, whereas the transformation-based approaches are referred to as *derivational*. For representational analyses, there are only surface structures augmented by

[31] See Crocker & Lewin (1992: 511) and Fordham & Crocker (1994: 38) for another constraint-based Parsing-as-Deduction approach.

traces but none of these structures is connected to an underlying structure by means of transformations (see Koster 1978; 1987: 235; Kolb & Thiersch 1991; Haider 1993: Section 1.4; Frey 1993: 14; Lohnstein 1993: 87–88, 177–178; Fordham & Crocker 1994: 38; Veenstra 1998: 58, for example). These analyses can be implemented in the same way as corresponding HPSG analyses (see Chapter 9) as computer-processable fragments and this has in fact been carried out for example for the analysis of verb position in German,[32] but such implemented analyses differ from GB analyses with regard to their basic architecture and in small, but important details such as how one deals with the interaction of long-distance dependencies and coordination (Gazdar 1981b). For a critical discussion and classification of movement analyses in Transformational Grammar, see Borsley (2012).

Following this somewhat critical overview, I want to add a comment in order not to be misunderstood: I do demand that all linguistic work shall be completely formalized. There is simply not space for this in a, say, thirty page essay. Furthermore, I do not believe that all linguists should carry out formal work and implement their analyses as computational models. However, there has to be *somebody* who works out the formal details and these basic theoretical assumptions should be accepted and adopted for a sufficient amount of time by the research community in question.

Comprehension questions

1. Give some examples of functional and lexical categories.
2. How can one represent lexical categories with binary features and what advantages does this have?

Exercises

1. Draw syntactic trees for the following examples:

   (55) a. dass die Frau den Mann liebt
   that the.NOM woman the.ACC man loves
   'that the woman loves the man'

   b. dass der Mann geliebt wird
   that the.NOM man loved is
   'that the man is loved'

[32] This shows that ten Hacken's contrasting of HPSG with GB and LFG (ten Hacken 2007: Section 4.3) and the classification of these frameworks as belonging to different research paradigms is completely mistaken. In his classification, ten Hacken refers mainly to the model-theoretic approach that HPSG assumes. However, LFG also has a model-theoretic formalization (Kaplan 1995). Furthermore, there is also a model-theoretic variant of GB (Rogers 1998). For further discussion, see Chapter 14.

c. Der Mann wird geliebt.
the.NOM man is loved
'The man is loved.'

d. dass der Mann der Frau hilft
that the.NOM man the.DAT woman helps
'that the man helps the woman'

e. Der Mann hilft der Frau.
the man.NOM helps the.DAT woman
'The man is helping the woman.'

For the passive sentences, use the analysis where the subject noun phrase is moved from the object position, that is, the analysis without an empty expletive as the subject.

Further reading

For Sections 3.1–3.5, I used material from Peter Gallmann from 2003. This has been modified, however, at various points. I am solely responsible for any mistakes or inadequacies. For current materials by Peter Gallmann, see http://www.syntax-theorie.de.

In the book *Syntaktische Analyseperspektiven*, Lohnstein (2014) presents a variant of GB which more or less corresponds to what is discussed in this chapter (CP/IP, movement-based analysis of the passive). The chapters in said book have been written by proponents of various theories and all analyze the same newspaper article. This book is extremely interesting for all those who wish to compare the various theories out there.

Haegeman (1990) is a comprehensive introduction to GB. Those who do read German may consider the textbooks by Fanselow & Felix (1987), von Stechow & Sternefeld (1988) and Grewendorf (1988) since they are also addressing the phenomena that are covered in this book.

In many of his publications, Chomsky discusses alternative, transformation-less approaches as "notational variants". This is not appropriate, as analyses without transformations can make different predictions to transformation-based approaches (e. g. w. r. t. coordination and extraction. See Section 5.5 for a discussion of GPSG in this respect). In Gazdar (1981a), one can find a comparison of GB and GPSG as well as a discussion of the classification of GPSG as a notational variant of Transformational Grammar with contributions from Noam Chomsky, Gerald Gazdar and Henry Thompson.

Borsley (1999) and Kim & Sells (2008) have parallel textbooks for GB and HPSG in English. For the comparison of Transformational Grammar and LFG, see Bresnan & Kaplan (1982). Kuhn (2007) offers a comparison of modern deriviational analyses with constraint-based LFG and HPSG approaches. Borsley (2012) contrasts analyses of long-distance dependencies in HPSG with movement-based analyses as in GB/Minimalism. Borsley discusses four types of data which are problematic for movement-based approaches: extraction without fillers, extraction with multiple gaps (see also the discussion of (57)

on p. 173 and of (55) on p. 199 of this book), extractions where fillers and gaps do not match and extraction without gaps.