5 Generalized Phrase Structure Grammar

Generalized Phrase Structure Grammar (GPSG) was developed as an answer to Transformational Grammar at the end of the 1970s. The book by Gazdar, Klein, Pullum & Sag (1985) is the main publication in this framework. Hans Uszkoreit has developed a largish GPSG fragment for German (1987). Analyses in GPSG were so precise that it was possible to use them as the basis for computational implementations. The following is a possibly incomplete list of languages with implemented GPSG fragments:

- German (Weisweber 1987; Weisweber & Preuss 1992; Naumann 1987, 1988; Volk 1988)
- English (Evans 1985; Phillips & Thompson 1985; Phillips 1992; Grover, Carroll & Briscoe 1993)
- French (Emirkanian, Da Sylva & Bouchard 1996)
- Persian (Bahrani, Sameti & Manshadi 2011)

As was discussed in Section 3.1.1, Chomsky (1957) argued that simple phrase structure grammars are not well-suited to describe relations between linguistic structures and claimed that one needs transformations to explain them. These assumptions remained unchallenged for two decades (with the exception of publications by Harman (1963) and Freidin (1975)) until alternative theories such as LFG and GPSG emerged, which addressed Chomsky's criticisms and developed non-transformational explanations of phenomena for which there were previously only transformational analyses or simply none at all. The analysis of local reordering of arguments, passives and long-distance dependencies are some of the most important phenomena that have been discussed in this framework. Following some introductory remarks on the representational format of GPSG in Section 5.1, I will present the GPSG analyses of these phenomena in some more detail.

5.1 General remarks on the representational format

This section has three parts. The general assumptions regarding features and the representation of complex categories is explained in Section 5.1.1, the assumptions regarding the linearization of daughters in a phrase structure rule is explained in Section 5.1.2. Section 5.1.3 introduces metarules, Section 5.1.4 deals with semantics, and Section 5.1.5 with adjuncts.

5.1.1 Complex categories, the Head Feature Convention, and $\overline{\mathrm{X}}$ rules

In Section 2.2, we augmented our phrase structure grammars with features. GPSG goes one step further and describes categories as sets of feature-value pairs. The category in (1a) can be represented as in (1b):

(1) a. NP(3,sg,nom)
b. { CAT n, BAR 2, PER 3, NUM sg, CASE nom }

It is clear that (1b) corresponds to (1a). (1a) differs from (1b) with regard to the fact that the information about part of speech and the $\overline{\mathrm{X}}$ level (in the symbol NP) are prominent, whereas in (1b) these are treated just like the information about case, number or person.

Lexical entries have a feature SUBCAT. The value is a number which says something about the kind of grammatical rules in which the word can be used. (2) shows examples for grammatical rules and lists some verbs which can occur in these rules.[1]

(2)

V2 → H[5]	(*kommen* 'come', *schlafen* 'sleep')
V2 → H[6], N2[CASE acc]	(*kennen* 'know', *suchen* 'search')
V2 → H[7], N2[CASE dat]	(*helfen* 'help', *vertrauen* 'trust')
V2 → H[8], N2[CASE dat], N2[CASE acc]	(*geben* 'give', *zeigen* 'show')
V2 → H[9], V3[+dass]	(*wissen* 'know', *glauben* 'believe')

These rules license VPs, that is, the combination of a verb with its complements, but not with its subject. The numbers following the category symbols (V or N) indicate the $\overline{\mathrm{X}}$ projection level. For Uszkoreit, the maximum number of projections of a verbal projection is three rather than two as is often assumed.

The H on the right side of the rule stands for *head*. The *Head Feature Convention* (HFC) ensures that certain features of the mother node are also present on the node marked with H (for details see Gazdar, Klein, Pullum & Sag 1985: Section 5.4 and Uszkoreit 1987: 67):

Principle 3 (Head Feature Convention)
The mother node and the head daughter must bear the same head features unless indicated otherwise.

In (2), examples for verbs which can be used in the rules are given in brackets. As with ordinary phrase structure grammars, one also requires corresponding lexical entries for verbs in GPSG. Two examples are provided in (3):

(3) V[5, VFORM *inf*] → einzuschlafen
V[6, VFORM *inf*] → aufzuessen

The first rule states that *einzuschlafen* 'to fall asleep' has a SUBCAT value of 5 and the second indicates that *aufzuessen* 'to finish eating' has a SUBCAT value of 6. It follows, then, that *einzuschlafen* can only be used in the first rule (2) and *aufzuessen* can only be

[1] The analyses discussed in the following are taken from Uszkoreit (1987).

used in the second. Furthermore, (3) contains information about the form of the verb (*inf* stands for infinitives with *zu* ‘to’).

If we analyze the sentence in (4) with the second rule in (2) and the second rule in (3), then we arrive at the structure in Figure 5.1.

(4) Karl hat versucht, [den Kuchen aufzuessen].
Karl has tried the cake to.eat.up
‘Karl tried to finish eating the cake.’

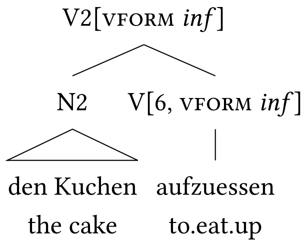

Figure 5.1: Projection of head features in GPSG

The rules in (2) say nothing about the order of the daughters which is why the verb (H[6]) can also be in final position. This aspect will be discussed in more detail in Section 5.1.2. With regard to the HFC, it is important to bear in mind that information about the infinitive verb form is also present on the mother node. Unlike simple phrase structure rules such as those discussed in Chapter 2, this follows automatically from the Head Feature Convention in GPSG. In (3), the value of vform is given and the HFC ensures that the corresponding information is represented on the mother node when the rules in (2) are applied. For the phrase in (4), we arrive at the category V2[vform *inf*] and this ensures that this phrase only occurs in the contexts it is supposed to:

(5) a. [Den Kuchen aufzuessen] hat er nicht gewagt.
the cake to.eat.up has he not dared
‘He did not dare to finish eating the cake.’

b. * [Den Kuchen aufzuessen] darf er nicht.
the cake to.eat.up be.allowed.to he not
Intended: ‘He is not allowed to finish eating the cake.’

c. * [Den Kuchen aufessen] hat er nicht gewagt.
the cake eat.up has he not dared
Intended: ‘He did not dare to finish eating the cake.’

d. [Den Kuchen aufessen] darf er nicht.
the cake eaten.up be.allowed.to he not
‘He is not allowed to finish eating the cake.’

gewagt 'dared' selects for a verb or verb phrase with an infinitive with *zu* 'to' but not a bare infinitive, while *darf* 'be allowed to' takes a bare infinitive.

This works in an analogous way for noun phrases: there are rules for nouns which do not take an argument as well as for nouns with certain arguments. Examples of rules for nouns which either require no argument or two PPs are given in (6) (Gazdar, Klein, Pullum & Sag 1985: 127):

(6) N1 $\rightarrow$ H[30] (*Haus* 'house', *Blume* 'flower')
N1 $\rightarrow$ H[31], PP[*mit*], PP[*über*] (*Gespräch* 'talk', *Streit* 'argument')

The rule for the combination of $\overline{\text{N}}$ and a determiner is as follows:

(7) N2 $\rightarrow$ Det, H1

N2 stands for NP, that is, for a projection of a noun phrase on bar level two, whereas H1 stands for a projection of the head daughter on the bar level one. The Head Feature Convention ensures that the head daughter is also a nominal projection, since all features on the head daughter apart from the $\overline{\text{X}}$ level are identified with those of the whole NP. When analyzing (8), the second rule in (6) licenses the $\overline{\text{N}}$ *Gesprächs mit Maria über Klaus*. The fact that *Gesprächs* 'conversation' is in the genitive is represented in the lexical item of *Gesprächs* and since *Gesprächs* is the head, it is also present at $\overline{\text{N}}$, following the Head Feature Convention.

(8) des Gespräch-s mit Maria über Klaus
the.GEN conversation-GEN with Maria about Klaus
'the conversation with Maria about Klaus'

For the combination of $\overline{\text{N}}$ with the determiner, we apply the rule in (7). The category of the head determines the word class of the element on the left-hand side of the rule, which is why the rule in (7) corresponds to the classical $\overline{\text{X}}$ rules that we encountered in (64c) on page 76. Since *Gesprächs mit Maria über Klaus* is the head daughter, the information about the genitive of $\overline{\text{N}}$ is also present at the NP node.

5.1.2 Local reordering

The first phenomenon to be discussed is local reordering of arguments. As was already discussed in Section 3.5, arguments in the middle field can occur in an almost arbitrary order. (9) gives some examples:

(9) a. [weil] der Mann der Frau das Buch gibt
because the.NOM man the.DAT woman the.ACC book gives
'because the man gives the book to the woman'

b. [weil] der Mann das Buch der Frau gibt
because the.NOM man the.ACC book the.DAT woman gives

c. [weil] das Buch der Mann der Frau gibt
because the.ACC book the.NOM man the.DAT woman gives

d. [weil] das Buch der Frau der Mann gibt
because the.ACC book the.DAT woman the.NOM man gives

e. [weil] der Frau der Mann das Buch gibt
because the.DAT woman the.NOM man the.ACC book gives

f. [weil] der Frau das Buch der Mann gibt
because the.DAT woman the.ACC book the.NOM man gives

In the phrase structure grammars in Chapter 2, we used features to ensure that verbs occur with the correct number of arguments. The following rule in (10) was used for the sentence in (9a):

(10) S → NP[nom] NP[dat] NP[acc] V_nom_dat_acc

If one wishes to analyze the other orders in (9), then one requires an additional five rules, that is, six in total:

(11) S → NP[nom] NP[dat] NP[acc] V_nom_dat_acc
S → NP[nom] NP[acc] NP[dat] V_nom_dat_acc
S → NP[acc] NP[nom] NP[dat] V_nom_dat_acc
S → NP[acc] NP[dat] NP[nom] V_nom_dat_acc
S → NP[dat] NP[nom] NP[acc] V_nom_dat_acc
S → NP[dat] NP[acc] NP[nom] V_nom_dat_acc

In addition, it is necessary to postulate another six rules for the orders with verb initial order:

(12) S → V_nom_dat_acc NP[nom] NP[dat] NP[acc]
S → V_nom_dat_acc NP[nom] NP[acc] NP[dat]
S → V_nom_dat_acc NP[acc] NP[nom] NP[dat]
S → V_nom_dat_acc NP[acc] NP[dat] NP[nom]
S → V_nom_dat_acc NP[dat] NP[nom] NP[acc]
S → V_nom_dat_acc NP[dat] NP[acc] NP[nom]

Furthermore, one would also need parallel rules for transitive and intransitive verbs with all possible valences. Obviously, the commonalities of these rules and the generalizations regarding them are not captured. The point is that we have the same number of arguments, they can be realized in any order and the verb can be placed in initial or final position. As linguists, we find it desirable to capture this property of the German language and represent it beyond phrase structure rules. In Transformational Grammar, the relationship between the orders is captured by means of movement: the Deep Structure corresponds to verb-final order with a certain order of arguments and the surface order is derived by means of move α. Since GPSG is a non-transformational theory, this kind of explanation is not possible. Instead, GPSG imposes restrictions on *immediate dominance* (ID), which differ from those which refer to *linear precedence* (LP): rules such as (13) are to be understood as dominance rules, which do not have anything to say about the order of the daughters (Pullum 1982).

(13) S → V, NP[nom], NP[acc], NP[dat]

The rule in (13) simply states that S dominates all other nodes. Due to the abandonment of ordering restrictions for the right-hand side of the rule, we only need one rule rather than twelve.

Nevertheless, without any kind of restrictions on the right-hand side of the rule, there would be far too much freedom. For example, the following order would be permissible:

(14) * Der Frau der Mann gibt ein Buch.
the woman.DAT the.NOM man gives the.ACC book

Such orders are ruled out by so-called *Linear Precedence Rules* or LP-rules. LP-constraints are restrictions on local trees, that is, trees with a depth of one. It is, for example, possible to state something about the order of V, NP[nom], NP[acc] and NP[dat] in Figure 5.2 using linearization rules.

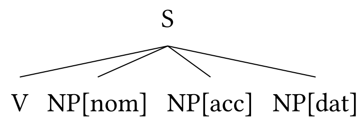

Figure 5.2: Example of a local tree

The following linearization rules serve to exclude orders such as those in (14):

(15) V[+MC] < X
X < V[−MC]

MC stands for *main clause*. The LP-rules ensure that in main clauses (+MC), the verb precedes all other constituents and follows them in subordinate clauses (−MC). There is a restriction that says that all verbs with the MC-value '+' also have to be (+FIN). This will rule out infinitive forms in initial position.

These LP rules do not permit orders with an occupied prefield or postfield in a local tree. This is intended. We will see how fronting can be accounted for in Section 5.4.

5.1.3 Metarules

We have previously encountered linearization rules for sentences with subjects, however our rules have the form in (16), that is, they do not include subjects:

(16) V2 → H[7], N2[CASE dat]
V2 → H[8], N2[CASE dat], N2[CASE acc]

These rules can be used to analyze the verb phrases *dem Mann das Buch zu geben* 'to give the man the book' and *das Buch dem Mann zu geben* 'to give the book to the man' as they appear in (17), but we cannot analyze sentences like (9), since the subject does not occur on the right-hand side of the rules in (16).

(17) a. Er verspricht, [dem Mann das Buch zu geben].
he promises the.DAT man the.ACC book to give
'He promises to give the man the book.'

b. Er verspricht, [das Buch dem Mann zu geben].
he promises the.ACC book the.DAT man to give
'He promises to give the book to the man.'

A rule with the format of (18) does not make much sense for a GPSG analysis of German since it cannot derive all the orders in (9) as the subject can occur between the elements of the VP as in (9c).

(18) S → N2 V2

With the rule in (18), it is possible to analyze (9a) as in Figure 5.3 and it would also be possible to analyze (9b) with a different ordering of the NPs inside the VP. The remaining examples in (9) cannot be captured by the rule in (18), however. This has to do with the

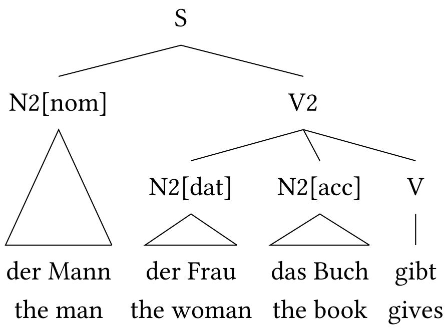

Figure 5.3: VP analysis for German (not appropriate in the GPSG framework)

fact that only elements in the same local tree, that is, elements which occur on the right-hand side of a rule, can be reordered. While we can reorder the parts of the VP and thereby derive (9b), it is not possible to place the subject at a lower position between the objects. Instead, a metarule can be used to analyze sentences where the subject occurs between other arguments of the verb. This rule relates phrase structure rules to other phrase structure rules. A metarule can be understood as a kind of instruction that creates another rule for each rule with a certain form and these newly created rules will in turn license local trees.

For the example at hand, we can formulate a metarule which says the following: if there is a rule with the form "V2 consists of something" in the grammar, then there also has to be another rule "V3 consists of whatever V2 consists + an NP in the nominative". In formal terms, this looks as follows:

(19) V2 → W ↦
V3 → W, N2[CASE nom]

W is a variable which stands for an arbitrary number of categories (W = *whatever*). The metarule creates the following rule in (20) from the rules in (16):

(20) V3 → H[7], N2[CASE dat], N2[CASE nom]
V3 → H[8], N2[CASE dat], N2[CASE acc], N2[CASE nom]

Now, the subject and other arguments both occur in the right-hand side of the rule and can therefore be freely ordered as long as no LP rules are violated.

5.1.4 Semantics

The semantics adopted by Gazdar, Klein, Pullum & Sag (1985: Chapter 9–10) goes back to Richard Montague (1974). Unlike a semantic theory which stipulates the combinatorial possibilities for each rule (see Section 2.3), GPSG uses more general rules. This is possible due to the fact that the expressions to be combined each have a semantic type. It is customary to distinguish between entities (*e*) and truth values (*t*). Entities refer to an object in the world (or in a possible world), whereas entire sentences are either true or false, that is, they have a truth value. It is possible to create more complex types from the types *e* and *t*. Generally, the following holds: if *a* and *b* are types, then $\langle\ a, b\ \rangle$ is also a type. Examples of complex types are $\langle\ e, t\ \rangle$ and $\langle\ e, \langle\ e, t\ \rangle\rangle$. We can define the following combinatorial rule for these kind of typed expressions:

(21) If α is of type $\langle\ b, a\ \rangle$ and β of type *b*, then $\alpha(\beta)$ is of type *a*.

This type of combination is also called *functional application.* With the rule in (21), it is possible that the type $\langle\ e, \langle\ e, t\ \rangle\rangle$ corresponds to an expression which still has to be combined with two expressions of type *e* in order to result in an expression of *t*. The first combination step with *e* will yield $\langle\ e, t\ \rangle$ and the second step of combination with a further *e* will give us *t*. This is similar to what we saw with λ-expressions on page 64: $\lambda y \lambda x$ *like*′(x, y) has to combine with a y and an x. The result in this example was *mögen*′(*max*′, *lotte*′), that is, an expression that is either true or false in the relevant world.

In Gazdar et al. (1985), an additional type is assumed for worlds in which an expression is true or false. For reasons of simplicity, I will omit this here. The types that we need for sentences, NPs and N′s, determiners and VPs are given in (22):

(22) a. TYP(S) = *t*
b. TYP(NP) = $\langle\ \langle\ e, t\ \rangle, t\ \rangle$
c. TYP(N′) = $\langle\ e, t\ \rangle$
d. TYP(Det) = $\langle$ TYP(N′), TYP(NP) $\rangle$
e. TYP(VP) = $\langle\ e, t\ \rangle$

A sentence is of type *t* since it is either true or false. A VP needs an expression of type *e* to yield a sentence of type *t*. The type of the NP may seem strange at first glance, however, it is possible to understand it if one considers the meaning of NPs with quantifiers. For sentences such as (23a), a representation such as (23b) is normally assumed:

(23) a. All children laugh.
b. $\forall x$ *child′*(x) $\rightarrow$ *laugh′*(x)

The symbol $\forall$ stands for the universal quantifier. The formula can be read as follows. For every object, for which it is the case that it has the property of being a child, it is also the case that it is laughing. If we consider the contribution made by the NP, then we see that the universal quantifier, the restriction to children and the logical implication come from the NP:

(24) $\forall x$ *child′*(x) $\rightarrow$ P(x)

This means that an NP is something that must be combined with an expression which has exactly one open slot corresponding to the x in (24). This is formulated in (22b): an NP corresponds to a semantic expression which needs something of type $\langle\ e, t\ \rangle$ to form an expression which is either true or false (that is, of type *t*).

An N′ stands for a nominal expression for the kind λx child(x). This means if there is a specific individual which one can insert in place of the x, then we arrive at an expression that is either true or false. For a given situation, it is the case that either John has the property of being a child or he does not. An N′ has the same type as a VP.

TYP(N′) and TYP(NP) in (22d) stand for the types given in (22c) and (22b), that is, a determiner is semantically something which has to be combined with the meaning of N′ to give the meaning of an NP.

Gazdar, Klein, Pullum & Sag (1985: 209) point out a redundancy in the semantic specification of grammars which follow the rule-by-rule hypothesis (see Section 2.3) since, instead of giving rule by rule instructions with regard to combinations, it suffices in many cases simply to say that the functor is applied to the argument. If we use types such as those in (22), it is also clear which constituent is the functor and which is the argument. In this way, a noun cannot be applied to a determiner, but rather only the reverse is possible. The combination in (25a) yields a well-formed result, whereas (25b) is ruled out.

(25) a. Det′(N′)
b. N′(Det′)

The general combinatorial principle is then as follows:

(26) Use functional application for the combination of the semantic contribution of the daughters to yield a well-formed expression corresponding to the type of the mother node.

The authors of the GPSG book assume that this principle can be applied to the vast majority of GPSG rules so that only a few special cases have to be dealt with by explicit rules.

5.1.5 Adjuncts

For nominal structures in English, Gazdar et al. (1985: 126) assume the $\overline{\mathrm{X}}$ analysis and, as we have seen in Section 2.4.1, this analysis is applicable to nominal structures in German. Nevertheless, there is a problem regarding the treatment of adjuncts in the verbal domain if one assumes flat branching structures, since adjuncts can freely occur between arguments:

(27) a. weil der Mann der Frau das Buch *gestern* gab
because the man the woman the book yesterday gave
'because the man gave the book to the woman yesterday'

b. weil der Mann der Frau *gestern* das Buch gab
because the man the woman yesterday the book gave

c. weil der Mann *gestern* der Frau das Buch gab
because the man yesterday the woman the book gave

d. weil *gestern* der Mann der Frau das Buch gab
because yesterday the man the woman the book gave

For (27), one requires the following rule:

(28) V3 → H[8], N2[CASE dat], N2[CASE acc], N2[CASE nom], AdvP

Of course, adjuncts can also occur between the arguments of verbs from other valence classes:

(29) weil (oft) die Frau (oft) dem Mann (oft) hilft
because often the woman often the man often helps
'because the woman often helps the man'

Furthermore, adjuncts can occur between the arguments of a VP:

(30) Der Mann hat versucht, der Frau heimlich das Buch zu geben.
the man has tried the woman secretly the book to give
'The man tried to secretly give the book to the woman.'

In order to analyze these sentences, we can use a metarule which adds an adjunct to the right-hand side of a V2 (Uszkoreit 1987: 146).

(31) V2 → W ↦
V2 → W, AdvP

By means of the subject introducing metarule in (19), the V3-rule in (28) is derived from a V2-rule. Since there can be several adjuncts in one sentence, a metarule such as (31) must be allowed to apply multiple times. The recursive application of metarules is often ruled out in the literature due to reasons of generative capacity (see Chapter 17) (Thompson 1982; Uszkoreit 1987: 146). If one uses the Kleene star, then it is possible to formulate the adjunct metarule in such as way that it does not have to apply recursively (Uszkoreit 1987: 146):

(32) $V2 \rightarrow W \mapsto$
$V2 \rightarrow W, AdvP^*$

If one adopts the rule in (32), then it is not immediately clear how the semantic contribution of the adjuncts can be determined.[2] For the rule in (31), one can combine the semantic contribution of the AdvP with the semantic contribution of the V2 in the input rule. This is of course also possible if the metarule is applied multiple times. If this metarule is applied to (33a), for example, the V2-node in (33a) contains the semantic contribution of the first adverb.

(33) a. V2 → V, NP, AdvP
b. V2 → V, NP, AdvP, AdvP

The V2-node in (33b) receives the semantic representation of the adverb applied to the V2-node in (33a).

Weisweber & Preuss (1992) have shown that it is possible to use metarules such as (31) if one does not use metarules to compute a set of phrase structure rules, but rather directly applies the metarules during the analysis of a sentence. Since sentences are always of finite length and the metarule introduces an additional AdvP to the right-hand side of the newly licensed rule, the metarule can only be applied a finite number of times.

5.2 Passive as a metarule

The German passive can be described in an entirely theory-neutral way as follows:[3]

- The subject is suppressed.
- If there is an accusative object, this becomes the subject.

This is true for all verb classes which can form the passive. It does not make a difference whether the verbs takes one, two or three arguments:

(34) a. weil er noch gearbeitet hat
because he.NOM still worked has
'because he has still worked'

[2] In LFG, an adjunct is entered into a set in the functional structure (see Section 7.1.6). This also works with the use of the Kleene Star notation. From the f-structure, it is possible to compute the semantic denotation with corresponding scope by making reference to the c-structure. In HPSG, Kasper (1994) has made a proposal which corresponds to the GPSG proposal with regard to flat branching structures and an arbitrary number of adjuncts. In HPSG, however, one can make use of so-called relational constraints. These are similar to small programs which can create relations between values inside complex structures. Using such relational constraints, it is then possible to compute the meaning of an unrestricted number of adjuncts in a flat branching structure.

[3] This characterization does not hold for other languages. For instance, Icelandic allows for dative subjects. See Zaenen, Maling & Thráinsson (1985).

b. weil noch gearbeitet wurde
because still worked was
'because there was still working there'

(35) a. weil er an Maria gedacht hat
because he.NOM on Maria thought has
'because he thought of Maria'

b. weil an Maria gedacht wurde
because on Maria thought was
'because Maria was thought of'

(36) a. weil sie ihn geschlagen hat
because she.NOM him.ACC beaten has
'because she has beaten him'

b. weil er geschlagen wurde
because he.NOM beaten was
'because he was beaten'

(37) a. weil er ihm den Aufsatz gegeben hat
because he.NOM him.DAT the.ACC essay given has
'because he has given him the essay'

b. weil ihm der Aufsatz gegeben wurde
because him.DAT the.NOM essay given was
'because he was given the essay'

In a simple phrase structure grammar, we would have to list two separate rules for each pair of sentences making reference to the valence class of the verb in question. The characteristics of the passive discussed above would therefore not be explicitly stated in the set of rules. In GPSG, it is possible to explain the relation between active and passive rules using a metarule: for each active rule, a corresponding passive rule with suppressed subject is licensed. The link between active and passive clauses can therefore be captured in this way.

An important difference to Transformational Grammar/GB is that we are not creating a relation between two trees, but rather between active and passive rules. The two rules license two unrelated structures, that is, the structure of (38b) is not derived from the structure of (38a).

(38) a. weil sie ihn geschlagen hat
because she.NOM him.ACC beaten has
'because she has beaten him'

b. weil er geschlagen wurde
because he.NOM beaten was
'because he was beaten'

The generalization with regard to active/passive is captured nevertheless.

In what follows, I will discuss the analysis of the passive given in Gazdar, Klein, Pullum & Sag (1985) in some more detail: the authors suggest the following metarule for English (p. 59):[4]

(39) VP → W, NP ↦
VP[pas] → W, (PP[*by*])

This rule states that verbs which take an object can occur in a passive VP without this object. Furthermore, a *by*-PP can be added. If we apply this metarule to the rules in (40), then this will yield the rules listed in (41):

(40) VP → H[2], NP
VP → H[3], NP, PP[*to*]

(41) VP[pas] → H[2], (PP[*by*])
VP[pas] → H[3], PP[*to*], (PP[*by*])

It is possible to use the rules in (40) to analyze verb phrases in active sentences:

(42) a. [$_{S}$ The man [$_{VP}$ devoured the carcass]].
b. [$_{S}$ The man [$_{VP}$ handed the sword to Tracy]].

The combination of a VP with the subject is licensed by an additional rule (S → NP, VP).

With the rules in (41), one can analyze the VPs in the corresponding passive sentences in (43):

(43) a. [$_{S}$ The carcass was [$_{VP[pas]}$ devoured (by the man)]].
b. [$_{S}$ The sword was [$_{VP[pas]}$ handed to Tracy (by the man)]].

At first glance, this analysis may seem odd as an object is replaced inside the VP by a PP which would be the subject in an active clause. Although this analysis makes correct predictions with regard to the syntactic well-formedness of structures, it seems unclear how one can account for the semantic relations. It is possible, however, to use a lexical rule that licenses the passive participle and manipulates the semantics of the output lexical item in such a way that the *by*-PP is correctly integrated semantically (Gazdar et al. 1985: 219).

We arrive at a problem, however, if we try to apply this analysis to German since the impersonal passive cannot be derived by simply suppressing an object. The V2-rules for verbs such as *arbeiten* 'work' and *denken* 'think' as used for the analysis of (34a) and (35a) have the following form:

(44) V2 → H[5]
V2 → H[13], PP[*an*]

[4] See Weisweber & Preuss (1992: 1114) for a parallel rule for German which refers to accusative case on the left-hand side of the metarule.

There is no NP on the right-hand side of these rules which could be turned into a *von*-PP. If the passive is to be analyzed as suppressing an NP argument in a rule, then it should follow from the existence of the impersonal passive that the passive metarule has to be applied to rules which license finite clauses, since information about whether there is a subject or not is only present in rules for finite clauses.[5] In this kind of system, the rules for finite sentences (V3) are the basic rules and the rules for V2 would be derived from these.

It would only make sense to have a metarule which applies to V3 for German since English does not have V3 rules which contain both the subject and its object on the right-hand side of the rule.[6] For English, it is assumed that a sentence consists of a subject and a VP (see Gazdar et al. 1985: 139). This means that we arrive at two very different analyses for the passive in English and German, which do not capture the descriptive insight that the passive is the suppression of the subject and the subsequent promotion of the object in the same way. The central difference between German and English seems to be that English obligatorily requires a subject,[7] which is why English does not have an impersonal passive. This is a property independent of passives, which affects the possibility of having a passive structure, however.

The problem with the GPSG analysis is the fact that valence is encoded in phrase structure rules and that subjects are not present in the rules for verb phrases. In the following chapters, we will encounter approaches from LFG, Categorial Grammar, HPSG, Construction Grammar, and Dependency Grammar which encode valence separately from phrase structure rules and therefore do not have a principled problem with impersonal passive.

See Jacobson (1987b: 394–396) for more problematic aspects of the passive analysis in GPSG and for the insight that a lexical representation of valence – as assumed in Categorial Grammar, GB, LFG and HPSG – allows for a lexical analysis of the phenomenon, which is however unformulable in GPSG for principled reasons having to do with the fundamental assumptions regarding valence representations.

5.3 Verb position

Uszkoreit (1987) analyzed verb-initial and verb-final order as linearization variants of a flat tree. The details of this analysis have already been discussed in Section 5.1.2.

An alternative suggestion in a version of GPSG comes from (Jacobs 1986: 110): Jacobs' analysis is a rendering of the verb movement analysis in GB. He assumes that there is an

[5] GPSG differs from GB in that infinitive verbal projections do not contain nodes for empty subjects. This is also true for all other theories discussed in this book with the exception of Tree-Adjoining Grammar.

[6] Gazdar et al. (1985: 62) suggest a metarule similar to our subject introduction metarule on page 187. The rule that is licensed by their metarule is used to analyze the position of auxiliaries in English and only licenses sequences of the form AUX NP VP. In such structures, subjects and objects are not in the same local tree either.

[7] Under certain conditions, the subject can also be omitted in English. For more on imperatives and other subject-less examples, see page 524.

empty verb in final position and links this to the verb in initial position using technical means which we will see in more detail in the following section.

5.4 Long-distance dependencies as the result of local dependencies

One of the main innovations of GPSG is its treatment of long-distance dependencies as a sequence of local dependencies (Gazdar 1981b). This approach will be explained taking constituent fronting to the prefield in German as an example. Until now, we have only seen the GPSG analysis for verb-initial and verb-final position: the sequences in (45) are simply linearization variants.

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

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

What we want is to derive the verb-second order in the examples in (46) from V1 order in (45b).

(46) a. Der Mann gibt der Frau das Buch.
  the man gives the woman the book
  'The man gives the woman the book.'

b. Der Frau gibt der Mann das Buch.
  the woman gives the man the book
  'The man gives the woman the book.'

For this, the metarule in (47) has to be used. This metarule removes an arbitrary category X from the set of categories on the right-hand side of the rule and represents it on the left-hand side with a slash ('/'):[8]

(47) V3 → W, X ↦
  V3/X → W

This rule creates the rules in (49) from (48):

(48) V3 → H[8], N2[CASE dat], N2[CASE acc], N2[CASE nom]

(49) V3/N2[CASE nom] → H[8], N2[CASE dat], N2[CASE acc]
  V3/N2[CASE dat] → H[8], N2[CASE acc], N2[CASE nom]
  V3/N2[CASE acc] → H[8], N2[CASE dat], N2[CASE nom]

[8] An alternative to Uszkoreit's trace-less analysis (1987: 77), which is explained here, consists of using a trace for the extracted element as in GB.

The rule in (50) connects a sentence with verb-initial order with a constituent which is missing in the sentence:

(50) V3[+FIN] → X[+TOP], V3[+MC]/X

In (50), X stands for an arbitrary category which is marked as missing in V3 by the '/'. X is referred to as a *filler*.

The interesting cases of values for X with regard to our examples are given in (51):

(51) V3[+FIN] → N2[+TOP, CASE nom], V3[+MC]/N2[CASE nom]
V3[+FIN] → N2[+TOP, CASE dat], V3[+MC]/N2[CASE dat]
V3[+FIN] → N2[+TOP, CASE acc], V3[+MC]/N2[CASE acc]

(51) does not show actual rules. Instead, (51) shows examples for insertions of specific categories into the X-position, that is, different instantiations of the rule.

The following linearization rule ensures that a constituent marked by [+TOP] in (50) precedes the rest of the sentence:

(52) [+TOP] < X

TOP stands for *topicalized*. As was mentioned on page 109, the prefield is not restricted to topics. Focused elements and expletives can also occur in the prefield, which is why the feature name is not ideal. However, it is possible to replace it with something else, for instance *prefield*. This would not affect the analysis. X in (52) stands for an arbitrary category. This is a new X and it is independent from the one in (50).

Figure 5.4 on the facing page shows the interaction of the rules for the analysis of (53).[9]

(53) Dem Mann gibt er das Buch.
the.DAT man gives he,NOM the.ACC book
'He gives the man the book.'

The metarule in (47) licenses a rule which adds a dative object into slash. This rule now licenses the subtree for *gibt er das Buch* 'gives he the book'. The linearization rule V[+MC] < X orders the verb to the very left inside of the local tree for V3. In the next step, the constituent following the slash is bound off. Following the LP-rule [+TOP] < X, the bound constituent must be ordered to the left of the V3 node.

The analysis given in Figure 5.4 may seem too complex since the noun phrases in (53) all depend on the same verb. It is possible to invent a system of linearization rules which would allow one to analyze (53) with an entirely flat structure. One would nevertheless still need an analysis for sentences such as those in (37) on page 110 – repeated here as (54):

[9] The FIN feature has been omitted one some of the nodes since it is redundant: +MC-verbs always require the FIN value '+'.

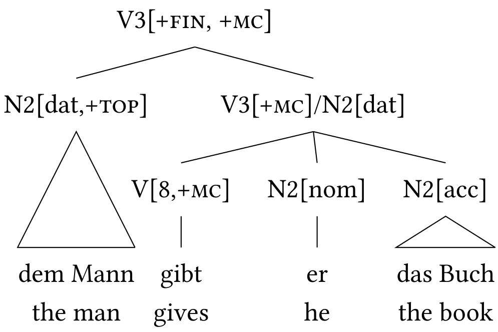

Figure 5.4: Analysis of fronting in GPSG

(54) a. [Um zwei Millionen Mark]$_i$ soll er versucht haben, [eine Versicherung $\_{i}$ zu betrügen].[10]
around two million Deutsche.Marks should he tried have an 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.[11]
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.[12]
who believe you that I seen have
'Who do you think I saw?'

d. [Gegen ihn]$_i$ falle es den Republikanern hingegen schwerer, [[Angriffe $\_{i}$] zu lancieren].[13]
against him fall it the Republicans however more.difficult attacks to launch
'It is, however, more difficult for the Republicans to launch attacks against him.'

The sentences in (54) cannot be explained by local reordering as the elements in the prefield are not dependent on the highest verb, but instead originate in the lower clause. Since only elements from the same local tree can be reordered, the sentences in (54) cannot be analyzed without postulating some kind of additional mechanism for long-distance dependencies.[14]

[10] taz, 04.05.2001, p. 20.
[11] Spiegel, 8/1999, p. 18.
[12] Scherpenisse (1986: 84).
[13] taz, 08.02.2008, p. 9.
[14] One could imagine analyses that assume the special mechanism for nonlocal dependencies only for sen-

Before I conclude this chapter, I will discuss yet another example of fronting, namely one of the more complex examples in (54). The analysis of (54c) consists of several steps: the introduction, percolation and finally binding off of information about the long-distance dependency. This is shown in Figure 5.5. Simplifying somewhat, I assume that

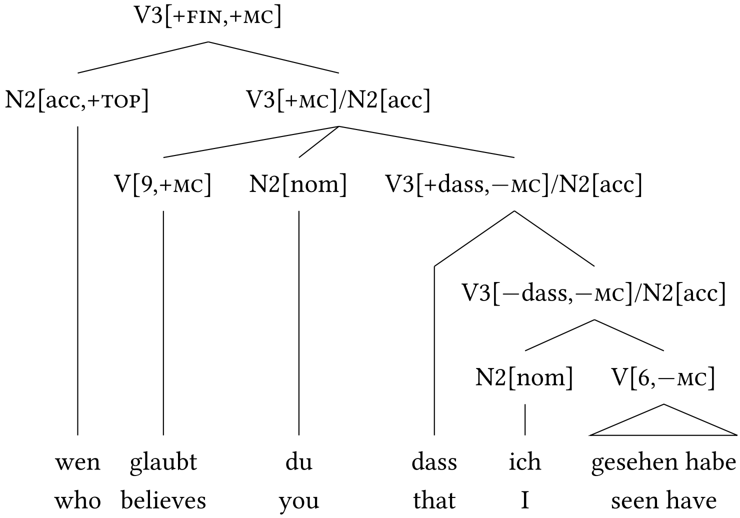

Figure 5.5: Analysis of long-distance dependencies in GPSG

gesehen habe 'have seen' behaves like a normal transitive verb.[15] A phrase structure rule licensed by the metarule in (47) licenses the combination of *ich* 'I' and *gesehen habe* 'has seen' and represents the missing accusative object on the V3 node. The complementizer *dass* 'that' is combined with *ich gesehen habe* 'I have seen' and the information about the fact that an accusative NP is missing is percolated up the tree. This percolation is controlled by the so-called *Foot Feature Principle*, which states that all foot features of all the daughters are also present on the mother node. Since the slash feature is a foot feature, the categories following the '/' percolate up the tree if they are not bound off in the local tree. In the final step, the V3/N2[acc] is combined with the missing N2[acc]. The result is a complete finite declarative clause of the highest projection level.

tences that really involve dependencies that are nonlocal. This was done in HPSG by Kathol (1995) and Wetta (2011) and by Groß & Osborne (2009) in Dependency Grammar. I discuss the Dependency Grammar analyses in detail in Section 11.7.1 and show that analyses that treat simple V2 sentences as ordering variants of non-V2 sentences have problems with the scope of fronted adjuncts, with coordination of simple sentences and sentences with nonlocal dependencies and with so-called multiple frontings.

[15] See Nerbonne (1986a) and Johnson (1986), for analyses of verbal complexes in GPSG.

5.5 Summary and classification

Some twenty years after Chomsky's criticism of phrase structure grammars, the first large grammar fragment in the GPSG framework appeared and offered analyses of phenomena which could not be described by simple phrase structure rules. Although works in GPSG essentially build on Harman's 1963 idea of a transformation-less grammar, they also go far beyond this. A special achievement of GPSG is, in particular, the treatment of long-distance dependencies as worked out by Gazdar (1981b). By using the SLASH-mechanism, it was possible to explain the simultaneous extraction of elements from conjuncts (Across the Board Extraction, Ross 1967). The following examples from Gazdar (1981b: 173) show that gaps in conjuncts must be identical, that is, a filler of a certain category must correspond to a gap in every conjunct:

(55) a. The kennel which Mary made and Fido sleeps in has been stolen. (= S/NP & S/NP)

b. The kennel in which Mary keeps drugs and Fido sleeps has been stolen. (= S/PP & S/PP)

c. * The kennel (in) which Mary made and Fido sleeps has been stolen. (= S/NP & S/PP)

GPSG can plausibly handle this with mechanisms for the transmission of information about gaps. In symmetric coordination, the SLASH elements in each conjunct have to be identical. On the one hand, a transformational approach is not straightforwardly possible since one normally assumes in such analyses that there is a tree and something is moved to another position in the tree thereby leaving a trace. However, in coordinate structures, the filler would correspond to two or more traces and it cannot be explained how the filler could originate in more than one place.

While the analysis of Across the Board extraction is a true highlight of GPSG, there are some problematic aspects that I want to address in the following: the interaction between valence and morphology, the representation of valence and partial verb phrase fronting, and the expressive power of the GPSG formalism.

5.5.1 Valence and morphology

The encoding of valence in GPSG is problematic for several reasons. For example, morphological processes take into account the valence properties of words. Adjectival derivation with the suffix *-bar* '-able' is only productive with transitive verbs, that is, with verbs with an accusative object which can undergo passivization:

(56) a. lös-bar (nominative, accusative)
solv-able

b. vergleich-bar (nominative, accusative, PP[mit])
compar-able

c. * schlaf-bar (nominative)
sleep-able

d. * helf-bar (nominative, dative)
help-able

A rule for derivations with *-bar-* '-able' must therefore make reference to valence information. This is not possible in GPSG grammars since every lexical entry is only assigned a number which says something about the rules in which this entry can be used. For *-bar*-derivations, one would have to list in the derivational rule all the numbers which correspond to rules with accusative objects, which of course does not adequately describe the phenomenon. Furthermore, the valence of the resulting adjective also depends on the valence of the verb. For example, a verb such as *vergleichen* 'compare' requires a *mit* (with)-PP and *vergleichbar* 'comparable' does too (Riehemann: 1993: 7, 54; 1998: 68). In the following chapters, we will encounter models which assume that lexical entries contain information as to whether a verb selects for an accusative object or not. In such models, morphological rules which need to access the valence properties of linguistic objects can be adequately formulated.

The issue of interaction of valence and derivational morphology will be taken up in Section 21.2.2 again, where approaches in LFG and Construction Grammar are discussed that share assumptions about the encoding of valence with GPSG.

5.5.2 Valence and partial verb phrase fronting

Nerbonne (1986a) and Johnson (1986) investigate fronting of partial VPs in the GPSG framework. (57) gives some examples: in (57a) the bare verb is fronted and its arguments are realized in the middle field, in (57b) one of the objects is fronted together with the verb and in (57c) both objects are fronted with the verb.

(57) a. Erzählen wird er seiner Tochter ein Märchen können.
tell will he his daughter a fairy.tale can

b. Ein Märchen erzählen wird er seiner Tochter können.
a fairy.tale tell will he his daughter can

c. Seiner Tochter ein Märchen erzählen wird er können.
his daughter a fairy.tale tell will he can

'He will be able to tell his daughter a fairy tale.'

The problem with sentences such as those in (57) is that the valence requirements of the verb *erzählen* 'to tell' are realized in various positions in the sentence. For fronted constituents, one requires a rule which allows a ditransitive to be realized without its arguments or with one or two objects. Furthermore, it has to be ensured that the arguments that are missing in the prefield are realized in the remainder of the clause. It is not legitimate to omit obligatory arguments or realize arguments with other properties like a different case, as the examples in (58) show:

(58) a. Verschlungen hat er es nicht.
devoured has he.NOM it.ACC not

'He did not devour it.'

b. * Verschlungen hat er nicht.
devoured has he.NOM not

c. * Verschlungen hat er ihm nicht.
devoured has he.NOM him.DAT not

The obvious generalization is that the fronted and unfronted arguments must add up to the total set belonging to the verb. This is scarcely possible with the rule-based valence representation in GPSG. In theories such as Categorial Grammar (see Chapter 8), it is possible to formulate elegant analyses of (58) (Geach 1970). Nerbonne and Johnson both suggest analyses for sentences such as (58) which ultimately amount to changing the representation of valence information in the direction of Categorial Grammar.

Before I turn to the expressive power of the GPSG formalism, I want to note that the problems that we discussed in the previous subsections are both related to the representation of valence in GPSG. We already run into valence-related problems when discussing the passive in Section 5.2: since subjects and objects are introduced in phrase structure rules and since there are some languages in which subject and object are not in the same local tree, there seems to be no way to describe the passive as the suppression of the subject in GPSG.

5.5.3 Generative capacity

In GPSG, the system of linearization, dominance and metarules is normally restricted by conditions we will not discuss here in such a way that one could create a phrase structure grammar of the kind we saw in Chapter 2 from the specification of a GPSG grammar. Such grammars are also called context-free grammars. In the mid-80s, it was shown that context-free grammars are not able to describe natural language in general, that is it could be shown that there are languages that need more powerful grammar formalisms than context-free grammars (Shieber 1985; Culy 1985; see Pullum (1986) for a historical overview). The so-called *generative capacity* of grammar formalisms is discussed in Chapter 17.

Following the emergence of constraint-based models such as HPSG (see Chapter 9) and unification-based variants of Categorial Grammar (see Chapter 8 and Uszkoreit 1986a), most authors previously working in GPSG turned to other frameworks. The GPSG analysis of long-distance dependencies and the distinction between immediate dominance and linear precedence are still used in HPSG and variants of Construction Grammar to this day. See also Section 12.2 for a Tree Adjoining Grammar variant that separates dominance from precedence.

Comprehension questions

1. What does it mean for a grammar to be in an ID/LP format?

2. How are linear variants of constituents in the middle field handled by GPSG?

3. Think of some phenomena which have been described by transformations and consider how GPSG has analyzed these data using other means.

Exercises

1. Write a small GPSG grammar which can analyze the following sentences:

(59) a. [dass] der Mann ihn liest
that the.NOM man him.ACC reads
'that the man reads it'

b. [dass] ihn der Mann liest
that him.ACC the.NOM man reads
'that the man reads it'

c. Der Mann liest ihn.
the.NOM man reads him.ACC
'The man reads it.'

Include all arguments in a single rule without using the metarule for introducing subjects.

Further reading

The main publication in GPSG is Gazdar, Klein, Pullum & Sag (1985). This book has been critically discussed by Jacobson (1987b). Some problematic analyses are contrasted with alternatives from Categorial Grammar and reference is made to the heavily Categorial Grammar influenced work of Pollard (1984), which counts as one of the predecessors of HPSG. Some of Jacobson's suggestions can be found in later works in HPSG.

Grammars of German can be found in Uszkoreit (1987) and Busemann (1992). Gazdar (1981b) developed an analysis of long-distance dependencies, which is still used today in theories such as HPSG.

A history of the genesis of GPSG can be found in Pullum (1989b).