9 Head-Driven Phrase Structure Grammar

Head-Driven Phrase Structure Grammar (HPSG) was developed by Carl Pollard and Ivan Sag in the mid-80's in Stanford and in the Hewlett Packard research laboratories in Palo Alto (Pollard & Sag 1987, 1994). Like LFG, HPSG is part of so-called West Coast linguistics. Another similarity to LFG is that HPSG aims to provide a theory of competence which is compatible with performance (Sag & Wasow (2011, 2015), also see Chapter 15).

The formal properties of the description language for HPSG grammars are well-understood and there are many systems for processing such grammars (Dörre & Seiffert 1991; Dörre & Dorna 1993; Popowich & Vogel 1991; Uszkoreit, Backofen, Busemann, Diagne, Hinkelman, Kasper, Kiefer, Krieger, Netter, Neumann, Oepen & Spackman 1994; Erbach 1995; Schütz 1996; Schmidt, Theofilidis, Rieder & Declerck 1996b; Schmidt, Rieder & Theofilidis 1996a; Uszkoreit, Backofen, Calder, Capstick, Dini, Dörre, Erbach, Estival, Manandhar, Mineur & Oepen 1996; Müller 1996c, 2004c; Carpenter & Penn 1996; Penn & Carpenter 1999; Götz, Meurers & Gerdemann 1997; Copestake 2002; Callmeier 2000; Dahllöf 2003; Meurers, Penn & Richter 2002; Penn 2004; Müller 2007a; Sato 2008; Kaufmann 2009).[1] Currently, the LKB system by Ann Copestake and the TRALE system, that was developed by Gerald Penn (Meurers, Penn & Richter 2002; Penn 2004), have the most users. The DELPH-IN consortium and various TRALE users have developed many small and some large fragments of various languages. The following is a list of implementations in different systems:

- Arabic (Haddar, Boukedi & Zalila 2010; Hahn 2011; Masum, Islam, Rahman & Ahmed 2012; Boukedi & Haddar 2014; Loukam, Balla & Laskri 2015; Arad Greshler, Herzig Sheinfux, Melnik & Wintner 2015),
- Bengali (Paul 2004; Islam, Hasan & Rahman 2012),
- Bulgarian (Simov, Osenova, Simov & Kouylekov 2004; Osenova 2010a,b, 2011),
- Cantonese (Fan, Song & Bond 2015),
- Danish (Ørsnes 1995, 2009b; Neville & Paggio 2004; Müller 2009c; Müller & Ørsnes 2011; Müller 2012a; Müller & Ørsnes 2015),

[1] Uszkoreit et al. (1996) and Bolc et al. (1996) compare systems that were available or were developed at the beginnings of the 1990s. Melnik (2007) compares LKB and TRALE. See also Müller (2015a: Sections 5.1).

- German (Kiss 1991; Netter 1993, 1996; Meurers 1994; Hinrichs et al. 1997; Kordoni 1999; Tseng 2000; Geißler & Kiss 1994; Keller 1994; Müller 1996c, 1999a; Müller & Kasper 2000; Crysmann 2003, 2005b,c; Müller 2007b; Kaufmann & Pfister 2007, 2008; Kaufmann 2009; Fokkens 2011),
- English (Copestake & Flickinger 2000; Flickinger, Copestake & Sag 2000; Flickinger 2000; Dahllöf 2002, 2003; De Kuthy & Meurers 2003a; Meurers, De Kuthy & Metcalf 2003; De Kuthy, Metcalf & Meurers 2004),
- Esperanto (Li 1996),
- French (Tseng 2003),
- Ga (Kropp Dakubu, Hellan & Beermann 2007; Hellan 2007),
- Sign Language (German, French, British, Greek) (Sáfár & Marshall 2002; Marshall & Sáfár 2004; Sáfár & Glauert 2010),
- Georgian (Abzianidze 2011),
- Greek (Kordoni & Neu 2005),
- Hausa (Crysmann 2005a, 2009, 2011, 2012),
- Hebrew (Melnik 2007; Haugereid, Melnik & Wintner 2013; Arad Greshler, Herzig Sheinfux, Melnik & Wintner 2015),
- Japanese (Siegel 2000; Siegel & Bender 2002; Bender & Siegel 2005),
- Yiddish (Müller & Ørsnes 2011),
- Korean (Kim & Yang 2003, 2004, 2006, 2009; Kim, Sells & Yang 2007; Song, Kim, Bond & Yang 2010; Kim, Yang, Song & Bond 2011),
- Maltese (Müller 2009b),
- Mandarin Chinese (Liu 1997; Ng 1997; Müller & Lipenkova 2009, 2013; Fan, Song & Bond 2015),
- Dutch (van Noord & Bouma 1994; Bouma, van Noord & Malouf 2001b; Fokkens 2011),
- Norwegian (Hellan & Haugereid 2003; Beermann & Hellan 2004; Hellan & Beermann 2006),
- Persian (Müller 2010b; Müller & Ghayoomi 2010),
- Polish (Przepiórkowski, Kupść, Marciniak & Mykowiecka 2002; Mykowiecka, Marciniak, Przepiórkowski & Kupść 2003),

- Portuguese (Branco & Costa 2008a,b; Costa & Branco 2010),
- Russian (Avgustinova & Zhang 2009),
- Sahaptin (Drellishak 2009),
- Spanish (Pineda & Meza 2005a,b; Bildhauer 2008; Marimon 2013),
- Turkish (Fokkens, Poulson & Bender 2009),
- Wambaya (Bender 2008a,c, 2010).

The first implemented HPSG grammar was a grammar of English developed in the Hewlett-Packard labs in Palo Alto (Flickinger, Pollard & Wasow 1985; Flickinger 1987). Grammars for German were developed in Heidelberg, Stuttgart and Saarbrücken in the LILOG project. Subsequently, grammars for German, English and Japanese were developed in Heidelberg, Saarbrücken and Stanford in the Verb*mobil* project. Verb*mobil* was the largest ever AI project in Germany. It was a machine translation project for spoken language in the domains of trip planning and appointment scheduling (Wahlster 2000).

Currently there are two larger groups that are working on the development of grammars: the DELPH-IN consortium (Deep Linguistic Processing with HPSG)[2] and the network CoGETI (Constraintbasierte Grammatik: Empirie, Theorie und Implementierung)[3]. Many of the grammar fragments that are listed above were developed by members of DELPH-IN and some were derived from the Grammar Matrix which was developed for the LKB to provide grammar writers with a typologically motivated initial grammar that corresponds to the properties of the language under development (Bender, Flickinger & Oepen 2002). The CoreGram project[4] is a similar project that is being run at the Freie Universität Berlin. It is developing grammars for German, Danish, Persian, Maltese, Mandarin Chinese, Spanish, French and Yiddish that share a common core. Constraints that hold for all languages are represented in one place and used by all grammars. Furthermore there are constraints that hold for certain language classes and again they are represented together and used by the respective grammars. So while the Grammar Matrix is used to derive grammars that individual grammar writers can use, adapt and modify to suit their needs, CoreGram really develops grammars for various languages that are used simultaneously and have to stay in sync. A description of the CoreGram can be found in Müller (2013a, 2015a).

There are systems that combine linguistically motivated analyses with statistics components (Brew 1995; Miyao et al. 2005; Miyao & Tsujii 2008) or learn grammars or lexica from corpora (Fouvry 2003; Cramer & Zhang 2009).

The following URLs point to pages on which grammars can be tested:

- http://www.delph-in.net/erg/
- http://hpsg.fu-berlin.de/Demos/

[2] http://www.delph-in.net/. 13.11.2015.
[3] http://wwwuser.gwdg.de/~cogeti/. 13.11.2015. Supported by the DFG under the grant number HO3279/3-1.
[4] http://hpsg.fu-berlin.de/Projects/CoreGram.html. 14.01.2016.

9.1 General remarks on the representational format

HPSG has the following characteristics: it is a lexicon-based theory, that is, the majority of linguistic constraints are situated in the descriptions of words or roots. HPSG is sign-based in the sense of Saussure (1916a): the form and meaning of linguistic signs are always represented together. Typed feature structures are used to model all relevant information. These structures can be described with feature descriptions such as in (1). Lexical entries, phrases and principles are always modeled and described with the same formal means. Generalizations about word classes or rule schemata are captured with inheritance hierarchies. Phonology, syntax and semantics are represented in a single structure. There are no separate levels of representation such as PF or LF in Government & Binding Theory. (1) shows an excerpt from the representation of a word such as *Grammatik* 'grammar'.

(1) Lexical item for the word *Grammatik* 'grammar':

$$\begin{bmatrix}
\text{PHONOLOGY} & \langle \textit{Grammatik} \rangle \\
\text{SYNTAX-SEMANTICS} \ldots & \begin{bmatrix}
\text{CATEGORY} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix}\text{CASE} & \boxed{1}\\ \textit{noun} & \end{bmatrix} \\ \text{SUBCAT} & \langle \text{DET}[\text{CASE } \boxed{1}] \rangle \\ \textit{category} & \end{bmatrix} \\
\text{CONTENT} \ldots & \begin{bmatrix}\text{INST} & X \\ \textit{grammatik} & \end{bmatrix} \\
\textit{local} &
\end{bmatrix} \\
\textit{word} &
\end{bmatrix}$$

One can see that this feature description contains information about the phonology, syntactic category and semantic content of the word *Grammatik*. To keep things simple, the value of PHONOLOGY (PHON) is mostly given as an orthographic representation. In fully fleshed-out theories, the PHON value is a complex structure that contains information about metrical grids and weak or strong accents. See Bird & Klein (1994), Orgun (1996), Höhle (1999), Walther (1999), Crysmann (2002: Chapter 6), and Bildhauer (2008) for phonology in the framework of HPSG. The details of the description in (1) will be explained in the following sections.

HPSG has adopted various insights from other theories and newer analyses have been influenced by developments in other theoretical frameworks. Functor-argument structures, the treatment of valence information and function composition have been adopted from Categorial Grammar. Function composition plays an important role in the analysis of verbal complexes in languages like German and Korean. The ID/LP format as well as the Slash mechanism for long-distance dependencies both come from GPSG. The analysis assumed here for verb position in German was developed in the framework of Government & Binding.

9.1.1 Representation of valence information

The phrase structure grammars discussed in Chapter 2 have the disadvantage that one requires a great number of different rules for the various valence types. (2) shows some examples of this kind of rules and the corresponding verbs.

(2) S → NP[*nom*], V — *X schläft* 'X is sleeping'
S → NP[*nom*], NP[*acc*], V — *X Y erwartet* 'X expects Y'
S → NP[*nom*], PP[*über*], V — *X über Y spricht* 'X talks about Y'
S → NP[*nom*], NP[*dat*], NP[*acc*], V — *X Y Z gibt* 'X gives Y to Z'
S → NP[*nom*], NP[*dat*], PP[*mit*], V — *X Y mit Z dient* 'X serves Y with Z'

In order for the grammar not to create any incorrect sentences, one has to ensure that verbs are only used with appropriate rules.

(3) a. * dass Peter das Buch schläft
that Peter the book sleeps

b. * dass Peter erwartet
that Peter expects

c. * dass Peter über den Mann erwartet
that Peter about the man expects

Therefore, verbs (and heads in general) have to be divided into valence classes. These valence classes have to then be assigned to grammatical rules. One must therefore further specify the rule for transitive verbs in (2) as follows:

(4) S → NP[*nom*], NP[*acc*], V[*nom_acc*]

Here, valence has been encoded twice. First, we have said something in the rules about what kind of elements can or must occur, and then we have stated in the lexicon which valence class the verb belongs to. In Section 5.5, it was pointed out that morphological processes need to refer to valence information. Hence, it is desirable to remove redundant valence information from grammatical rules. For this reason, HPSG – like Categorial Grammar – includes descriptions of the arguments of a head in the lexical entry of that head. There is a feature with a list-value, the SUBCAT feature, which contains descriptions of the objects that must combine with a head in order to yield a complete phrase. (5) gives some examples for the verbs in (2):

(5)

Verb	SUBCAT
schlafen 'to sleep'	⟨ NP[*nom*] ⟩
erwarten 'to expect'	⟨ NP[*nom*], NP[*acc*] ⟩
sprechen 'to speak'	⟨ NP[*nom*], PP[*über*] ⟩
geben 'to give'	⟨ NP[*nom*], NP[*dat*], NP[*acc*] ⟩
dienen 'to serve'	⟨ NP[*nom*], NP[*dat*], PP[*mit*] ⟩

SUBCAT is an abbreviation for subcategorization. It is often said that a head subcategorizes for certain arguments. See page 93 for more on the term *subcategorization*.

Figure 9.1 shows the analysis for (6a) and the analysis for (6b) is in Figure 9.2 on the facing page:

(6) a. [dass] Peter schläft
  that Peter sleeps

b. [dass] Peter Maria erwartet
  that Peter Maria expects
  'that Peter expects Maria'

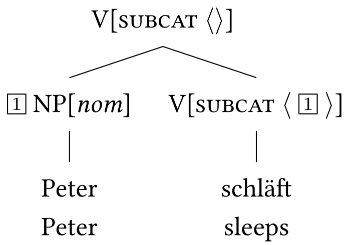

Figure 9.1: Analysis of *Peter schläft* 'Peter sleeps' in *dass Peter schläft* 'that Peter sleeps'

In Figures 9.1 and 9.2, one element of the SUBCAT list is combined with its head in each local tree. The elements that are combined with the selecting head are then no longer present in the SUBCAT list of the mother node. V[SUBCAT ⟨ ⟩] corresponds to a complete phrase (VP or S). The boxes with numbers show the structure sharing (see Section 6.4). Structure sharing is the most important means of expression in HPSG. It plays a central role for phenomena such as valence, agreement and long-distance dependencies. In the examples above, $\boxed{1}$ indicates that the description in the SUBCAT list is identical to another daughter in the tree. The descriptions contained in valence lists are usually partial descriptions, that is, not all properties of the argument are exhaustively described. Therefore, it is possible that a verb such as *schläft* 'sleeps' can be combined with various kinds of linguistic objects: the subject can be a pronoun, a proper name or a complex noun phrase, it only matters that the linguistic object in question has an empty SUBCAT list and bears the correct case.[5]

9.1.2 Representation of constituent structure

As already noted, feature descriptions in HPSG serve as the sole descriptive inventory of morphological rules, lexical entries and syntactic rules. The trees we have seen thus far are only visualizations of the part whole relations and do not have any theoretical status. There are also no rewrite rules in HPSG.[6] The job of phrase structure rules is handled

[5] Furthermore, it must agree with the verb. This is not shown here.

[6] However, phrase structure rules are used in some computer implementations of HPSG in order to improve the efficiency of processing.

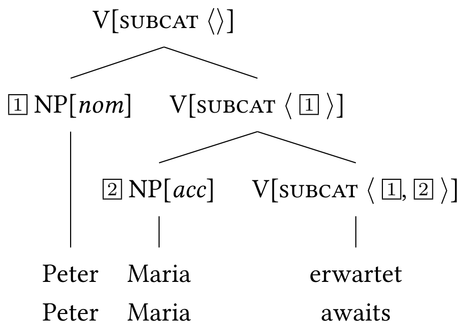

Figure 9.2: Analysis of *Peter Maria erwartet* 'Peter awaits Maria'

by feature descriptions. Information about dominance is represented using DTR features (head daughter and non-head daughter), information about precedence is implicitly contained in PHON. (7) shows the representation of PHON values in a feature description corresponding to the tree in Figure 9.3.

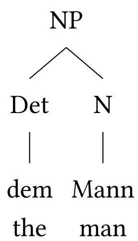

Figure 9.3: Analysis of *dem Mann* 'the man'

(7) $$\begin{bmatrix} \textsc{phon} & \langle\ \textit{dem Mann}\ \rangle \\ \textsc{head-dtr} & \begin{bmatrix} \textsc{phon}\ \langle\ \textit{Mann}\ \rangle \end{bmatrix} \\ \textsc{non-head-dtrs} & \left\langle \begin{bmatrix} \textsc{phon}\ \langle\ \textit{dem}\ \rangle \end{bmatrix} \right\rangle \end{bmatrix}$$

In (7), there is exactly one head daughter (HEAD-DTR). The head daughter is always the daughter containing the head. In a structure with the daughters *das* and *Bild von Maria*, the latter would be the head daughter. In principle, there can be multiple non-head daughters. If we were to assume a flat structure for a sentence with a ditransitive verb, as in Figure 2.1 on page 56, we would have three non-head daughters. It also makes sense to assume binary branching structures without heads (see Müller (2007b: Chapter 11) for an analysis of relative clauses). In such structures we would also have more than one non-head daughter, namely exactly two.

Before it is shown how it is ensured that only those head-argument structures are licensed in which the argument matches the requirements of the head, I will present the general structure of feature descriptions in HPSG. The structure presented at the start of this chapter is repeated in (8) with all the details relevant to the present discussion:

(8)
$$
\begin{bmatrix}
\text{PHON} & \langle \textit{Grammatik} \rangle \\
\text{SYNSEM} & \begin{bmatrix}
\text{LOC} & \begin{bmatrix}
\text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix}\text{CASE} & \boxed{1}\\ \textit{noun} & \end{bmatrix} \\ \text{SUBCAT} & \langle \text{DET}[\text{CASE}\ \boxed{1}]\rangle \\ \textit{category} & \end{bmatrix}\\
\text{CONT} & \begin{bmatrix} \text{IND} & \boxed{2} \begin{bmatrix}\text{PER} & \textit{third}\\ \text{NUM} & \textit{sg}\\ \text{GEN} & \textit{fem}\end{bmatrix}\\ \text{RELS} & \left\langle \begin{bmatrix}\text{INST} & \boxed{2}\\ \textit{grammatik} & \end{bmatrix}\right\rangle \\ \textit{mrs} & \end{bmatrix}\\
\textit{local} &
\end{bmatrix}\\
\text{NONLOC} & \begin{bmatrix}\text{INHER|SLASH} & \langle\rangle \\ \text{TO-BIND|SLASH} & \langle\rangle\end{bmatrix}
\end{bmatrix}\\
\textit{word} &
\end{bmatrix}
$$

In the outer layer, there are the features PHON and SYNSEM. As previously mentioned, PHON contains the phonological representation of a linguistic object. The value of SYNSEM is a feature structure which contains syntactic and semantic information that can be selected by other heads. The daughters of phrasal signs are represented outside of SYNSEM. This ensures that there is a certain degree of locality involved in selection: a head cannot access the internal structure of the elements which it selects (Pollard und Sag 1987: 143–145; 1994: 23). See also Sections 10.6.2.1 and 18.2 for a discussion of locality. Inside SYNSEM, there is information relevant in local contexts (LOCAL, abbreviated to LOC) as well as information important for long-distance dependencies (NONLOCAL or NONLOC for short). Locally relevant information includes syntactic (CATEGORY or CAT), and semantic (CONTENT or CONT) information. Syntactic information encompasses information that determines the central characteristics of a phrase, that is, the head information. This is represented under HEAD. Further details of this will be discussed in Section 9.1.4. Among other things, the part of speech of a linguistic object belongs to the head properties of a phrase. As well as HEAD, SUBCAT belongs to the information contained inside CAT. The semantic content of a sign is present under CONT. The type of the CONT value is *mrs*, which stands for *Minimal Recursion Semantics* (Copestake, Flickinger, Pollard &

Sag 2005). An MRS structure is comprised of an index and a list of relations which restrict this index. Of the NONLOCAL features, only SLASH is given here. There are further features for dealing with relative and interrogative clauses (Pollard & Sag 1994; Sag 1997; Ginzburg & Sag 2000; Holler 2005), which will not be discussed here.

As can be seen, the description of the word *Grammatik* 'grammar' becomes relatively complicated. In theory, it would be possible to list all properties of a given object directly in a single list of feature-value pairs. This would, however, have the disadvantage that the identity of groups of feature-value pairs could not be expressed as easily. Using the feature geometry in (8), one can express the fact that the CAT values of both conjuncts in symmetric coordinations such as those in (9) are identical.

(9) a. [der Mann] und [die Frau]
   the man and the woman

b. Er [kennt] und [liebt] diese Schallplatte.
   he knows and loves this record

c. Er ist [dumm] und [arrogant].
   he is dumb and arrogant

If valence and the part of speech information were not represented in one common substructure, we would have to state separately that utterances such as (9) require that both conjuncts have the same valence and part of speech.

After this general introduction of the feature geometry that is assumed here, we can now turn to the head-argument schema:

Schema 1 (Head-Argument Schema (binary branching, preliminary version))
head-argument-phrase ⇒

$$\begin{bmatrix} \text{SYNSEM}|\text{LOC}|\text{CAT}|\text{SUBCAT}\ \boxed{1} \\ \text{HEAD-DTR}|\text{SYNSEM}|\text{LOC}|\text{CAT}|\text{SUBCAT}\ \boxed{1} \oplus \langle\ \boxed{2}\ \rangle \\ \text{NON-HEAD-DTRS}\ \langle\ [\ \text{SYNSEM}\ \boxed{2}\ ]\ \rangle \end{bmatrix}$$

Schema 1 states the properties a linguistic object of the type *head-argument-phrase* must have. The arrow in Schema 1 stands for a logical implication and not for the arrow of rewrite rules as we know it from phrase structure grammars. '⊕' (*append*) is a relation which combines two lists. (10) shows possible splits of a list that contains two elements:

(10) $\langle x, y \rangle = \langle x \rangle \oplus \langle y \rangle$ or
$\langle\rangle \oplus \langle x, y \rangle$ or
$\langle x, y \rangle \oplus \langle\rangle$

The list $\langle x, y \rangle$ can be subdivided into two lists each containing one element, or alternatively into the empty list and $\langle x, y \rangle$.

Schema 1 can be read as follows: if an object is of the type *head-argument-phrase* then it must have the properties on the right-hand side of the implication. In concrete terms, this means that these objects always have a valence list which corresponds to $\boxed{1}$, that they have a head daughter with a valence list that can be divided into two sublists $\boxed{1}$ and $\langle \boxed{2} \rangle$ and also that they have a non-head daughter whose syntactic and semantic properties (SYNSEM value) are compatible with the last element of the SUBCAT list of the head daughter ($\boxed{2}$). (11) shows the corresponding feature description for the example in (6a).

(11)
$$\begin{bmatrix} \text{PHON} \ \langle \textit{Peter schläft} \rangle \\ \text{SYNSEM} \mid \text{LOC} \mid \text{CAT} \mid \text{SUBCAT} \ \langle\rangle \\ \text{HEAD-DTR} \ \begin{bmatrix} \text{PHON} \ \langle \textit{schläft} \rangle \\ \text{SYNSEM} \mid \text{LOC} \mid \text{CAT} \mid \text{SUBCAT} \ \langle \boxed{1}\ \text{NP}[\textit{nom}] \rangle \end{bmatrix} \\ \text{NON-HEAD-DTRS} \ \left\langle \begin{bmatrix} \text{PHON} \ \langle \textit{Peter} \rangle \\ \text{SYNSEM} \ \boxed{1} \end{bmatrix} \right\rangle \\ \textit{head-argument-phrase} \end{bmatrix}$$

NP[*nom*] is an abbreviation for a complex feature description. Schema 1 divides the SUBCAT list of the head daughter into a single-element list and what is left. Since *schläft* 'sleeps' only has one element in its SUBCAT list, what remains is the empty list. This remainder is also the SUBCAT value of the mother.

9.1.3 Linearization rules

Dominance schemata do not say anything about the order of the daughters. As in GPSG, linearization rules are specified separately. Linearization rules can make reference to the properties of daughters, their function in a schema (head, argument, adjunct, …) or both. If we assume a feature INITIAL for all heads, then heads which precede their arguments would have the INITIAL value '+' and heads following their arguments would have the value '−'. The linearization rules in (12) ensure that ungrammatical orders such as (13b,d) are ruled out.[7]

(12) a. Head[INITIAL +] < Argument
  b. Argument < Head[INITIAL−]

Prepositions have an INITIAL value '+' and therefore have to precede arguments. Verbs in final position bear the value '−' and have to follow their arguments.

(13) a. [in [den Schrank]]
  in the cupboard
  b. * [[den Schrank] in]
  the cupboard in

[7] Noun phrases pose a problem for (12): determiners have been treated as argument until now and were included in the SUBCAT list of the head noun. Determiners occur to the left of noun, whereas all other arguments of the noun are to the right. This problem can be solved either by refining linearization rules (Müller 1999a: 164–165) or by introducing a special valence feature for determiners (Pollard & Sag 1994: Section 9.4). For an approach using such a feature, see Section 9.6.1.

c. dass [er [ihn umfüllt]]
   that he it refills

d. * dass [er [umfüllt ihn]]
   that he refills him

9.1.4 Projection of head properties

As was explained in Section 1.5 certain properties of heads are important for the distribution of the whole phrase. For instance, the verb form belongs to the features that are important for the distribution of verbal projections. Certain verbs require a verbal argument with a particular form:

(14) a. [Dem Mann helfen] will er nicht.
   the man help wants he not
   'He doesn't want to help the man.'

b. [Dem Mann geholfen] hat er nicht.
   the man helped has he not
   'He hasn't helped the man.'

c. * [Dem Mann geholfen] will er nicht.
   the man helped wants he not

d. * [Dem Mann helfen] hat er nicht.
   the man help has he not

wollen 'to want' always requires an infinitive without *zu* 'to', while *haben* 'have' on the other hand requires a verb in participle form. *glauben* 'believe' can occur with a finite clause, but not with an infinitive without *zu*:

(15) a. Ich glaube, Peter kommt morgen.
   I believe Peter comes tomorrow
   'I think Peter is coming tomorrow.'

b. * Ich glaube, Peter morgen kommen.
   I believe Peter tomorrow come

c. * Ich glaube, morgen kommen.
   I believe tomorrow come

This shows that projections of verbs must not only contain information about the part of speech but also information about the verb form. Figure 9.4 on the next page shows this on the basis of the finite verb *gibt* 'gives'.

GPSG has the Head Feature Convention that ensures that head features on the mother node are identical to those on the node of the head daughter. In HPSG, there is a similar principle. Unlike GPSG, head features are explicitly contained as a group of features in the feature structures. They are listed under the path SYNSEM|LOC|CAT|HEAD. (16) shows the lexical item for *gibt* 'gives':

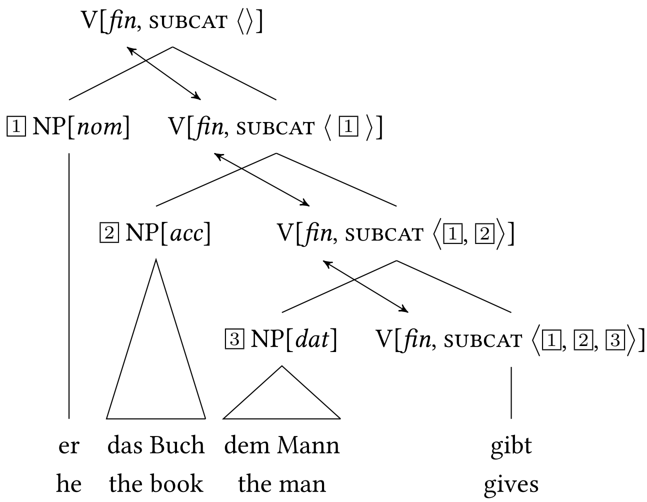

Figure 9.4: Projection of the head features of the verb

(16) *gibt* 'gives':

$$\begin{bmatrix} \textsc{phon} & \langle \textit{gibt} \rangle \\ \textsc{synsem}|\textsc{loc}|\textsc{cat} & \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textsc{vform} & \textit{fin} \\ \textit{verb} & \end{bmatrix} \\ \textsc{subcat} & \langle \text{NP}[\textit{nom}], \text{NP}[\textit{acc}], \text{NP}[\textit{dat}] \rangle \end{bmatrix} \\ \textit{word} & \end{bmatrix}$$

The *Head Feature Principle* takes the following form:

Principle 4 (*Head Feature Principle*)
The HEAD *value of any headed phrase is structure-shared with the* HEAD *value of the head daughter.*

Figure 9.5 on the facing page is a variant of Figure 9.4 with the structure sharing made explicit.

The following section will deal with how this principle is formalized as well as how it can be integrated into the architecture of HPSG.

9.1.5 Inheritance hierarchies and generalizations

Up to now, we have seen one example of a dominance schema and more will follow in the coming sections, e. g. schemata for head-adjunct structures as well as for the binding off of long-distance dependencies. The Head Feature Principle is a general principle which must be met by all structures licensed by these schemata. As mentioned above, it must

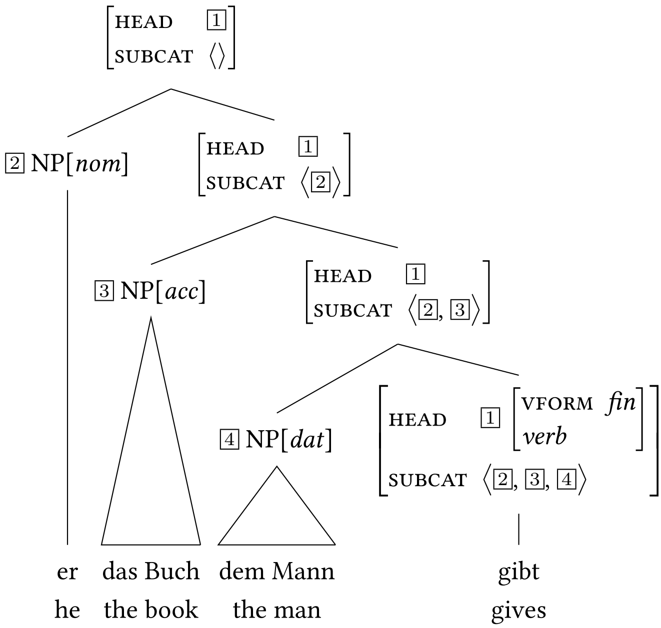

Figure 9.5: Projection of head features of a verb with structure sharing

be met by all structures with a head. Formally, this can be captured by categorizing syntactic structures into those with and those without heads and assigning the type *headed-phrase* to those with a head. The type *head-argument-phrase* – the type which the description in Schema 1 on page 269 has – is a subtype of *headed-phrase*. Objects of a certain type x always have all properties that objects have that are supertypes of x. Recall the example from Section 6.2: an object of the type *female person* has all the properties of the type *person*. Furthermore, objects of type *female person* have additional, more specific properties not shared by other subtypes of *person*.

If one formulates a restriction on a supertype, this automatically affects all of its subtypes. The Head Feature Principle hence can be formalized as follows:

(17) $$\textit{headed-phrase} \Rightarrow \begin{bmatrix} \text{SYNSEM}|\text{LOC}|\text{CAT}|\text{HEAD}\ \boxed{1} \\ \text{HEAD-DTR}|\text{SYNSEM}|\text{LOC}|\text{CAT}|\text{HEAD}\ \boxed{1} \end{bmatrix}$$

The arrow corresponds to the arrow that is used in implications in logic. (17) can be read as follows: if a structure is of type *headed-phrase*, then it must hold that the value of SYNSEM|LOC|CAT|HEAD is identical to the value of HEAD-DTR|SYNSEM|LOC|CAT|HEAD.

An extract from the type hierarchy under *sign* is given in Figure 9.6 on the next page. *word* and *phrase* are subclasses of linguistic signs. Phrases can be divided into phrases with heads (*headed-phrase*) and those without (*non-headed-phrase*). There are also subtypes for phrases of type *non-headed-phrase* and *headed-phrase*. We have already dis-

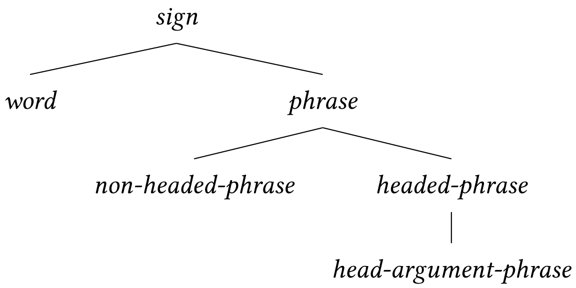

Figure 9.6: Type hierarchy for *sign*: all subtypes of *headed-phrase* inherit constraints

cussed *head-argument-phrase*, and other subtypes of *headed-phrase* will be discussed in the later sections. As well as *word* and *phrase*, there are the types *root* and *stem*, which play an important role for the structure of the lexicon and the morphological component. Due to space considerations, it is not possible to further discuss these types here, but see Chapter 22.

The description in (18) shows the Head-Argument Schema from page 269 together with the restrictions that the type *head-argument-phrase* inherits from *headed-phrase*.

(18) Head-Argument Schema + Head Feature Principle:

$$\begin{bmatrix} \text{SYNSEM}|\text{LOC}|\text{CAT} & \begin{bmatrix} \text{HEAD} & \boxed{1} \\ \text{SUBCAT} & \boxed{2} \end{bmatrix} \\ \text{HEAD-DTR}|\text{SYNSEM}|\text{LOC}|\text{CAT} & \begin{bmatrix} \text{HEAD} & \boxed{1} \\ \text{SUBCAT} & \boxed{2} \oplus \langle \boxed{3} \rangle \end{bmatrix} \\ \text{NON-HEAD-DTRS} & \langle \begin{bmatrix} \text{SYNSEM} & \boxed{3} \end{bmatrix} \rangle \\ \textit{head-argument-phrase} & \end{bmatrix}$$

(19) gives a description of a structure that is licensed by Schema 1. As well as valence information, the head information is specified in (19) and it is also apparent how the Head Feature Principle ensures the projection of features: the head value of the entire structure ($\boxed{1}$) corresponds to the head value of the verb *gibt* ‘gives’.

(19)
$$\begin{bmatrix}
\textsc{phon} \ \langle \textit{dem Mann gibt} \rangle \\
\textsc{synsem}|\textsc{loc}|\textsc{cat} \begin{bmatrix} \textsc{head} & \boxed{1} \\ \textsc{subcat} & \boxed{2} \ \langle \text{NP}[\textit{nom}], \text{NP}[\textit{acc}] \rangle \end{bmatrix} \\
\textsc{head-dtr} \begin{bmatrix} \textsc{phon} \ \langle \textit{gibt} \rangle \\ \textsc{synsem}|\textsc{loc}|\textsc{cat} \begin{bmatrix} \textsc{head} & \boxed{1} \begin{bmatrix} \textsc{vform} & \textit{fin} \\ \textit{verb} \end{bmatrix} \\ \textsc{subcat} & \boxed{2} \oplus \langle \boxed{3} \rangle \end{bmatrix} \\ \textit{word} \end{bmatrix} \\
\textsc{non-head-dtrs} \left\langle \begin{bmatrix} \textsc{phon} \ \langle \textit{dem Mann} \rangle \\ \textsc{synsem} \ \boxed{3} \begin{bmatrix} \textsc{loc}|\textsc{cat} \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textsc{cas} & \textit{dat} \\ \textit{noun} \end{bmatrix} \\ \textsc{subcat} & \langle\rangle \end{bmatrix} \end{bmatrix} \\ \textsc{head-dtr} \ \ldots \\ \textsc{non-head-dtrs} \ \ldots \end{bmatrix} \right\rangle \\
\textit{head-argument-phrase}
\end{bmatrix}$$

For the entire sentence *er das Buch dem Mann gibt* 'he the book to the man gives', we arrive at a structure (already shown in Figure 9.5) described by (20):

(20)
$$\begin{bmatrix}
\textsc{synsem}|\textsc{loc}|\textsc{cat} \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textsc{vform} & \textit{fin} \\ \textit{verb} \end{bmatrix} \\ \textsc{subcat} & \langle\rangle \end{bmatrix}
\end{bmatrix}$$

This description corresponds to the sentence symbol S in the phrase structure grammar on page 55, however (20) additionally contains information about the form of the verb.

Using dominance schemata as an example, we have shown how generalizations about linguistic objects can be captured, however, we also want to be able to capture generalizations in other areas of the theory: like Categorial Grammar, the HPSG lexicon contains a very large amount of information. Lexical entries (roots and words) can also be divided into classes, which can then be assigned types. In this way, it is possible to capture what all verbs, intransitive verbs and transitive verbs, have in common. See Figure 22.1 on page 663.

Now that some fundamental concepts of HPSG have been introduced, the following section will show how the semantic contribution of words is represented and how the meaning of a phrase can be determined compositionally.

9.1.6 Semantics

An important difference between theories such as GB, LFG and TAG, on the one hand, and HPSG and CxG on the other is that the semantic content of a linguistic object is modeled in a feature structure just like all its other properties. As previously mentioned, semantic information is found under the path SYNSEM|LOC|CONT. (21) gives an example of the CONT value for *Buch* 'book'. The representation is based on Minimal Recursion Semantics (MRS):[8]

(21)
$$\begin{bmatrix} \text{IND} & \boxed{1} \begin{bmatrix} \text{PER} & \mathit{3} \\ \text{NUM} & \mathit{sg} \\ \text{GEN} & \mathit{neu} \end{bmatrix} \\ \text{RELS} & \left\langle \begin{bmatrix} \text{INST} & \boxed{1} \\ \mathit{buch} & \end{bmatrix} \right\rangle \\ \mathit{mrs} & \end{bmatrix}$$

IND stands for index and RELS is a list of relations. Features such as person, number and gender are part of a nominal index. These are important in determining reference or coreference. For example, *sie* 'she' in (22) can refer to *Frau* 'woman' but not to *Buch* 'book'. On the other hand, *es* 'it' cannot refer to *Frau* 'woman'.

(22) Die Frau$_i$ kauft ein Buch$_j$. Sie$_i$ liest es$_j$.
the woman buys a book she reads it
'The woman buys a book. She reads it.'

In general, pronouns have to agree in person, number and gender with the element they refer to. Indices are then identified accordingly. In HPSG, this is done by means of structure sharing. It is also common to speak of *coindexation*. (23) provides some examples of coindexation of reflexive pronouns:

(23) a. Ich$_i$ sehe mich$_i$.
I see myself

b. Du$_i$ siehst dich$_i$.
you see yourself

c. Er$_i$ sieht sich$_i$.
he sees himself

d. Wir$_i$ sehen uns$_i$.
we see ourselves

e. Ihr$_i$ seht euch$_i$.
you see yourselves

f. Sie$_i$ sehen sich$_i$.
they see themselves

[8] Pollard & Sag (1994) and Ginzburg & Sag (2000) make use of Situation Semantics (Barwise & Perry 1983; Cooper, Mukai & Perry 1990; Devlin 1992). An alternative approach which has already been used in HPSG is Lexical Resource Semantics (Richter & Sailer 2004). For an early underspecification analysis in HPSG, see Nerbonne (1993).

The question of which instances of coindexation are possible and which are necessary is determined by Binding Theory. Pollard & Sag (1992, 1994) have shown that Binding Theory in HPSG does not have many of the problems that arise when implementing binding in GB with reference to tree configurations. There are, however, a number of open questions for Binding Theory in HPSG (Müller 1999a: Section 20.4).

(24) shows the CONT value for the verb *geben* 'give':

(24)
$$\begin{bmatrix} \text{IND} & \boxed{1}\ \textit{event} \\ \text{RELS} & \left\langle \begin{bmatrix} \text{EVENT} & \boxed{1} \\ \text{AGENT} & \textit{index} \\ \text{THEME} & \textit{index} \\ \text{GOAL} & \textit{index} \\ \textit{geben} & \end{bmatrix} \right\rangle \\ \textit{mrs} & \end{bmatrix}$$

It is assumed that verbs have an event variable of the type *event*, which is represented under IND just as with indices for nominal objects. Until now, we did not assign elements in the valence list to argument roles in the semantic representation. This connection is referred to as *linking*. (25) shows how linking works in HPSG. The referential indices of the argument noun phrases are structure shared with one of the semantic roles of the relation contributed by the head.

(25) *gibt* 'gives':
$$\begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{VFORM} & \textit{fin} \\ \textit{verb} & \end{bmatrix} \\ \text{SUBCAT} & \left\langle \text{NP}[\textit{nom}]_{\boxed{1}}, \text{NP}[\textit{acc}]_{\boxed{2}}, \text{NP}[\textit{dat}]_{\boxed{3}} \right\rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{IND} & \boxed{4}\ \textit{event} \\ \text{RELS} & \left\langle \begin{bmatrix} \text{EVENT} & \boxed{4} \\ \text{AGENT} & \boxed{1} \\ \text{THEME} & \boxed{2} \\ \text{GOAL} & \boxed{3} \\ \textit{geben} & \end{bmatrix} \right\rangle \\ \textit{mrs} & \end{bmatrix} \end{bmatrix}$$

Since we use general terms such as AGENT and PATIENT for argument roles, it is possible to state generalizations about valence classes and the realization of argument roles. For example, one can divide verbs into verbs taking an agent, verbs with an agent and theme, verbs with agent and patient etc. These various valence/linking patterns can be represented in type hierarchies and these classes can be assigned to the specific lexical entries, that is, one can have them inherit constraints from the respective types. A type constraint for verbs with agent, theme and goal takes the form of (26):

(26)
$$\left[\begin{array}{ll} \text{CAT}|\text{SUBCAT} & \left\langle []_{\boxed{1}}, []_{\boxed{2}}, []_{\boxed{3}} \right\rangle \\ \text{CONT} & \left[\begin{array}{ll} \text{IND} & \boxed{4}\ \textit{event} \\ \text{RELS} & \left\langle \left[\begin{array}{ll} \text{EVENT} & \boxed{4} \\ \text{AGENT} & \boxed{1} \\ \text{THEME} & \boxed{2} \\ \text{GOAL} & \boxed{3} \\ \multicolumn{2}{l}{\textit{agent-theme-goal-rel}} \end{array}\right] \right\rangle \\ \multicolumn{2}{l}{\textit{mrs}} \end{array}\right] \end{array}\right]$$

$[]_{\boxed{1}}$ stands for an object of unspecified syntactic category with the index $\boxed{1}$. The type for the relation *geben′* is a subtype of *agent-theme-goal-rel*. The lexical entry for the word *geben* 'give' or the root *geb-* has the linking pattern in (26).

For more on theories of linking in HPSG, see Davis (1996), Wechsler (1991) und Davis & Koenig (2000).

Up to now, we have only seen how the meaning of lexical entries can be represented. The Semantics Principle determines the computation of the semantic contribution of phrases: the index of the entire expression corresponds to the index of the head daughter, and the RELS value of the entire sign corresponds to the concatenation of the RELS values of the daughters plus any relations introduced by the dominance schema. The last point is important because the assumption that schemata can add something to meaning can capture the fact that there are some cases where the entire meaning of a phrase is more than simply the sum of its parts. Pertinent examples are often discussed as part of Construction Grammar. Semantic composition in HPSG is organized such that meaning components that are due to certain patterns can be integrated into the complete meaning of an utterance. For examples, see Section 21.10.

The connection between the semantic contribution of the verb and its arguments is established in the lexical entry. As such, we ensure that the argument roles of the verb are assigned to the correct argument in the sentence. This is, however, not the only thing that the semantics is responsible for. It has to be able to generate the various readings associated with quantifier scope ambiguities (see page 93) as well as deal with semantic embedding of predicates under other predicates. All these requirements are fulfilled by MRS. Due to space considerations, we cannot go into detail here. The reader is referred to the article by Copestake, Flickinger, Pollard & Sag (2005) and to Section 19.3 in the discussion chapter.

9.1.7 Adjuncts

Analogous to the selection of arguments by heads via SUBCAT, adjuncts can also select their heads using a feature (MODIFIED). Adjectives, prepositional phrases that modify nouns, and relative clauses select an almost complete nominal projection, that is, a noun that only still needs to be combined with a determiner to yield a complete NP. (27) shows part of the lexical entry for *interessantes* 'interesting':

(27) CAT value for *interessantes* 'interesting':

$$\begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{MOD} & \overline{\text{N}} \\ \textit{adj} & \end{bmatrix} \\ \text{SUBCAT} & \langle\rangle \end{bmatrix}$$

interessantes is an adjective that does not take any arguments and therefore has an empty SUBCAT list. Adjectives such as *treu* 'loyal' would have a dative NP in their SUBCAT list.

(28) ein dem König treues Mädchen
a the.DAT king loyal girl
'a girl loyal to the king'

The CAT value is given in (29):

(29) CAT value for *treues* 'loyal':

$$\begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{MOD} & \overline{\text{N}} \\ \textit{adj} & \end{bmatrix} \\ \text{SUBCAT} & \langle \text{NP}[\textit{dat}] \rangle \end{bmatrix}$$

dem König treues 'loyal to the king' forms an adjective phrase, which modifies *Mädchen*.

Unlike the selectional feature SUBCAT that belongs to the features under CAT, MOD is a head feature. The reason for this is that the feature that selects the modifying head has to be present on the maximal projection of the adjunct. The $\overline{\text{N}}$-modifying property of the adjective phrase *dem König treues* 'loyal to the king' has to be included in the representation of the entire AP just as it is present in the lexical entry for adjectives in (27) at the lexical level. The adjectival phrase *dem König treues* has the same syntactic properties as the basic adjective *interessantes* 'interesting':

(30) CAT value für *dem König treues*:

$$\begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{MOD} & \overline{\text{N}} \\ \textit{adj} & \end{bmatrix} \\ \text{SUBCAT} & \langle\rangle \end{bmatrix}$$

Since MOD is a head feature, the Head Feature Principle (see page 272) ensures that the MOD value of the entire projection is identical to the MOD value of the lexical entry for *treues* 'loyal'.

As an alternative to the selection of the head by the modifier, one could assume a description of all possible adjuncts on the head itself. This was suggested by Pollard & Sag (1987: 161). Pollard & Sag (1994: Section 1.9) revised the earlier analysis since the semantics of modification could not be captured.[9]

$\overline{\text{N}}$

AP[HEAD|MOD $\boxed{1}$] $\boxed{1}$ $\overline{\text{N}}$

interessantes Buch
interesting book

Figure 9.7: Head-adjunct structure (selection)

Figure 9.7 demonstrates selection in head-adjunct structures. Head-adjunct structures are licensed by the Schema 2.

Schema 2 (Head-Adjunct Schema (preliminary version))

head-adjunct-phrase ⇒

$$\begin{bmatrix} \text{HEAD-DTR}|\text{SYNSEM} & \boxed{1} \\ \text{NON-HEAD-DTRS} & \left\langle \begin{bmatrix} \text{SYNSEM}|\text{LOC}|\text{CAT} & \begin{bmatrix} \text{HEAD}|\text{MOD} & \boxed{1} \\ \text{SUBCAT} & \langle\rangle \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

The value of the selectional feature on the adjunct ($\boxed{1}$) is identified with the SYNSEM value of the head daughter, thereby ensuring that the head daughter has the properties specified by the adjunct. The SUBCAT value of the non-head daughter is the empty list, which is why only completely saturated adjuncts are allowed in head-adjunct structures. Phrases such as (31b) are therefore correctly ruled out:

(31) a. die Wurst in der Speisekammer
the sausage in the pantry

b. * die Wurst in
the sausage in

Example (31a) requires some further explanation. The preposition *in* (as used in (31a)) has the following CAT value:

(32) CAT value of *in*:

$$\begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{MOD} & \overline{\text{N}} \\ \textit{prep} & \end{bmatrix} \\ \text{SUBCAT} & \langle\ \text{NP}[\textit{dat}]\ \rangle \end{bmatrix}$$

[9] See Bouma, Malouf & Sag (2001a), however. Bouma, Malouf & Sag (2001a) pursue a hybrid analysis where there are adjuncts which select heads and also adjuncts that are selected by a head. *Minimal Recursion Semantics* (MRS) is the semantic theory underlying this analysis. Using this semantics, the problems arising for Pollard & Sag (1987) with regard to the semantics of modifiers are avoided.

After combining *in* with the nominal phrase *der Speisekammer* 'the pantry' one gets:

(33) CAT value for *in der Speisekammer* 'in the pantry':

$$\begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{MOD} & \overline{\text{N}} \\ \textit{prep} & \end{bmatrix} \\ \text{SUBCAT} & \langle\rangle \end{bmatrix}$$

This representation corresponds to that of the adjective *interessantes* 'interesting' and can – ignoring the position of the PP – also be used in the same way: the PP modifies a $\overline{\text{N}}$.

Heads that can only be used as arguments but do not modify anything have a MOD value of *none*. They can therefore not occur in the position of the non-head daughter in head-adjunct structures since the MOD value of the non-head daughter has to be compatible with the SYNSEM value of the head daughter.

9.2 Passive

HPSG follows Bresnan's argumentation (see Section 7.2) and takes care of the passive in the lexicon.[10] A lexical rule takes the verb stem as its input and licenses the participle form and the most prominent argument (the so-called designated argument) is suppressed.[11] Since grammatical functions are not part of theory in HPSG, we do not require any mapping principles that map objects to subjects. Nevertheless, one still has to explain the change of case under passivization. If one fully specifies the case of a particular argument in the lexical entries, one has to ensure that the accusative argument of a transitive verb is realized as nominative in the passive. (34) shows what the respective lexical rule would look like:

(34) Lexical rule for personal passives following Kiss (1992):

$$\begin{bmatrix} \text{PHON} & \boxed{1} \\ \text{SYNSEM}|\text{LOC}|\text{CAT} & \begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{SUBCAT} & \left\langle \text{NP}[\textit{nom}], \text{NP}[\textit{acc}]_{\boxed{2}} \right\rangle \oplus \boxed{3} \end{bmatrix} \\ \textit{stem} & \end{bmatrix} \mapsto$$

[10] Some exceptions to this are analyses influenced by Construction Grammar such as Tseng (2007) and Haugereid (2007). These approaches are problematic, however, as they cannot account for Bresnan's adjectival passives. For other problems with Haugereid's analysis, see Müller (2007c) and Section 21.3.6.

[11] For more on the designated argument, see Haider (1986a). HPSG analyses of the passive in German have been considerably influenced by Haider. Haider uses the designated argument to model the difference between so-called unaccusative and unergative verbs (Perlmutter 1978): unaccusative verbs differ from unergatives and transitives in that they do not have a designated argument. We cannot go into the literature on unaccusativity here. The reader is referred to the original works by Haider and the chapter on the passive in Müller (2007b).

$$\begin{bmatrix} \text{PHON}\ f(\boxed{1}) & \\ \text{SYNSEM|LOC|CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix}\text{VFORM} & \textit{passive-part}\end{bmatrix} \\ \text{SUBCAT} & \left\langle \text{NP}[\textit{nom}]_{\boxed{2}} \right\rangle \oplus \boxed{3} \end{bmatrix} \\ \textit{word} & \end{bmatrix}$$

This lexical rule takes a verb stem as its input[12], which requires a nominative argument, an accusative argument and possibly further arguments (if $\boxed{3}$ is not the empty list) and licenses a lexical entry that requires a nominative argument and possibly the arguments in $\boxed{3}$. The output of the lexical rule specifies the VFORM value of the output word. This is important as the auxiliary and the main verb must go together. For example, it is not possible to use the perfect participle instead of the passive participle since these differ in their valence in Kiss' approach:

(35) a. Der Mann hat den Weltmeister geschlagen.
the man has the world.champion beaten
'The man has beaten the world champion.'

b. * Der Mann wird den Weltmeister geschlagen.
the man is the world.champion beaten

c. Der Weltmeister wird geschlagen.
the world.champion is beaten
'The world champion is (being) beaten.'

There are a few conventions for the meaning of lexical rules: all information that is not mentioned in the output sign is taken over from the input sign. Thus, the meaning of the verb is not mentioned in the passive rule, which makes sense as the passive rule is a meaning preserving rule. The CONT values of the input and output are identical. It is important here that the linking information its retained. As an example consider the application of the rule to the verb stem *schlag-* 'beat':

(36) a. Input *schlag-* 'beat' :

$$\begin{bmatrix} \text{PHON}\ \langle\ \textit{schlag}\ \rangle & \\ \text{SYNSEM|LOC} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{SUBCAT} & \left\langle \text{NP}[\textit{nom}]_{\boxed{1}}, \text{NP}[\textit{acc}]_{\boxed{2}} \right\rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{IND} & \boxed{3}\ \textit{event} \\ \text{RELS} & \left\langle \begin{bmatrix} \text{EVENT} & \boxed{3} \\ \text{AGENT} & \boxed{1} \\ \text{PATIENT} & \boxed{2} \\ \textit{schlagen} & \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

[12] The term *stem* includes roots (*helf-* 'help-', products of derivation (*besing-* 'to extol') and compounds. The lexical rule can therefore also be applied to stems like *helf-* and derived forms such as *besing-*.

b. Output *geschlagen* 'beaten':

$$\begin{bmatrix} \text{PHON} & \langle \textit{geschlagen} \rangle \\ \text{SYNSEM}|\text{LOC} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{VFORM} & \textit{passive-part} \\ \textit{verb} & \end{bmatrix} \\ \text{SUBCAT} & \left\langle \text{NP}[\textit{nom}]_{\boxed{2}} \right\rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{IND} & \boxed{3}\ \textit{event} \\ \text{RELS} & \left\langle \begin{bmatrix} \text{EVENT} & \boxed{3} \\ \text{AGENT} & \boxed{1} \\ \text{PATIENT} & \boxed{2} \\ \textit{schlagen} & \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

The agent role is connected to the subject of *schlag-*. After passivization, the subject is suppressed and the argument that is connected to the patient role of *schlag-* becomes the subject of the participle. Argument linking is not affected by this and thus the nominative argument is correctly assigned to the patient role.

Lexical rules such as (34) can also be captured with feature descriptions (Meurers 2001):

(37)
$$\begin{bmatrix} \text{PHON} & f(\boxed{1}) \\ \text{SYNSEM}|\text{LOC}|\text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{VFORM} & \textit{passive-part} \end{bmatrix} \\ \text{SUBCAT} & \left\langle \text{NP}[\textit{nom}]_{\boxed{2}} \right\rangle \oplus \boxed{3} \end{bmatrix} \\ \text{LEX-DTR} & \begin{bmatrix} \text{PHON} & \boxed{1} \\ \text{SYNSEM}|\text{LOC}|\text{CAT} & \begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{SUBCAT} & \left\langle \text{NP}[\textit{nom}], \text{NP}[\textit{acc}]_{\boxed{2}} \right\rangle \oplus \boxed{3} \end{bmatrix} \\ \textit{stem} & \end{bmatrix} \\ \textit{acc-passive-lexical-rule} & \end{bmatrix}$$

What is on the left-hand side of the rule in (34), is contained in the value of LEX-DTR in (37). Since this kind of lexical rule is fully integrated into the formalism, feature structures corresponding to these lexical rules also have their own type. If the result of the application of a given rule is an inflected word, then the type of the lexical rule (*acc-passive-lexical-rule* in our example) is a subtype of *word*. Since lexical rules have a type, it is possible to state generalizations over lexical rules.

The lexical rules discussed thus far work well for the personal passive. For the impersonal passive, however, we would require a second lexical rule. Furthermore, we would have two different lexical items for the passive and the perfect, although the forms are always identical. In the following, I will discuss the basic assumptions that are needed for a theory of the passive that can sufficiently explain both personal and impersonal passives and thereby only require one lexical item for the participle form.

9.2.1 Valence information and the Case Principle

In Section 3.4.1, the difference between structural and lexical case was motivated. In the HPSG literature, it is assumed following Haider (1986a) that the dative is a lexical case. Lexical cases are specified in the descriptions of the selected arguments. Arguments with structural case are also specified as taking structural cases in the lexicon, but the actual case value is not provided. In order for the grammar not to make any false predictions, it has to be ensured that the structural cases receive a unique value dependent on their environment. This is handled by the Case Principle:[13]

Principle 5 (Case Principle)

- *In a list containing the subject as well as complements of a verbal head, the least oblique element with structural case receives nominative.*
- *All other elements in the list with structural case receive accusative.*
- *In nominal environments, elements with structural case are assigned genitive.*

(38) shows prototypical valence lists for finite verbs:

(38) a. *schläft* 'sleeps': SUBCAT $\langle$ NP[*str*]$_j$ $\rangle$
b. *unterstützt* 'supports': SUBCAT $\langle$ NP[*str*]$_j$, NP[*str*]$_k$ $\rangle$
c. *hilft* 'helps': SUBCAT $\langle$ NP[*str*]$_j$, NP[*ldat*]$_k$ $\rangle$
d. *schenkt* 'gives': SUBCAT $\langle$ NP[*str*]$_j$, NP[*str*]$_k$, NP[*ldat*]$_l$ $\rangle$

str stands for *structural* and *ldat* for lexical dative. It is commonly assumed in HPSG that elements in the valence list are ordered corresponding to the Obliqueness Hierarchy in Keenan & Comrie (1977) and Pullum (1977):

SUBJECT => DIRECT OBJECT => INDIRECT OBJECT => OBLIQUES => GENITIVES => OBJECTS OF COMPARISON

This hierarchy corresponds to the different syntactic activeness of grammatical functions. Elements that occur further left tend to to occur in specific syntactic constructions more often. Examples for syntactic constructions where obliqueness plays a role are the following:

- Ellipsis (Klein 1985)
- Topic Drop (Fries 1988)
- Free relatives (Bausewein 1990; Pittner 1995; Müller 1999b)

[13] The Case Principle has been simplified here. Cases of so-called 'raising' require special treatment. For more details, see Meurers (1999c), Przepiórkowski (1999a) and Müller (2007b: Chapter 14, Chapter 17). The Case Principle given in these publications is very similar to the one proposed by Yip, Maling & Jackendoff (1987) and can therefore also explain the case systems of the languages discussed by Yip, Maling and Jackendoff, notably the complicated case system of Icelandic.

- Passive (Keenan & Comrie 1977)
- Depictive predicates (Müller 2004d, 2002a, 2008a)
- Binding Theory (Grewendorf 1985; Pollard und Sag: 1992; 1994: Chapter 6)

The Case Principle ensures that the subjects of the verbs listed above have to be realized in the nominative and also that objects with structural case are assigned accusative.

With the difference between structural and lexical case, it is possible to formulate a passive-lexical rule that can account for both the personal and the impersonal passive:

(39) Lexical rule for personal and impersonal passive (simplified):

$$\begin{bmatrix} \textsc{phon}\ \boxed{1} \\ \textsc{synsem}|\textsc{loc}|\textsc{cat} \begin{bmatrix} \textsc{head} & \textit{verb} \\ \textsc{subcat} & \langle \text{NP}[\textit{str}] \rangle \oplus \boxed{2} \end{bmatrix} \\ \textit{stem} \end{bmatrix} \mapsto$$

$$\begin{bmatrix} \textsc{phon}\ f(\boxed{1}) \\ \textsc{synsem}|\textsc{loc}|\textsc{cat} \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textsc{vform} & \textit{ppp} \end{bmatrix} \\ \textsc{subcat} & \boxed{2} \end{bmatrix} \\ \textit{word} \end{bmatrix}$$

This lexical rule does exactly what we expect it to do from a pretheoretical perspective on the passive: it suppresses the most prominent argument with structural case, that is, the argument that corresponds to the subject in the active clause. The standard analysis of verb auxiliary constructions assumes that the main verb and the auxiliary forms a verbal complex (Hinrichs & Nakazawa 1994; Pollard 1994; Müller 1999a, 2002a; Meurers 2000; Kathol 2000). The arguments of the embedded verb are taken over by the auxiliary. After combining the participle with the passive auxiliary, we arrive at the following SUBCAT lists:

(40) a. *geschlafen wird* 'slept is': SUBCAT $\langle\ \rangle$
b. *unterstützt wird* 'supported is': SUBCAT $\langle$ NP[*str*]$_k$ $\rangle$
c. *geholfen wird*: 'helped is' SUBCAT $\langle$ NP[*ldat*]$_k$ $\rangle$
d. *geschenkt wird*: 'given is' SUBCAT $\langle$ NP[*str*]$_k$, NP[*ldat*]$_l$ $\rangle$

(40) differs from (38) in that a different NP is in first position. If this NP has structural case, it will receive nominative case. If there is no NP with structural case, as in (40c), the case remains as it was, that is, lexically specified.

We cannot go into the analysis of the perfect here. It should be noted, however, that the same lexical item for the participle is used for (41).

(41) a. Er hat den Weltmeister geschlagen.
he has the world.champion beaten
'He has beaten the world champion.'

b. Der Weltmeister wurde geschlagen.
the world.champion was beaten
'The world champion was beaten.'

It is the auxiliary that determines which arguments are realized (Haider 1986a; Müller 2007b: Chapter 17). The lexical rule in (39) licenses a form that can be used both in passive and perfect. Therefore, the VFORM value is of *ppp*, which stands for *participle*.

One should note that this analysis of the passive works without movement of constituents. The problems with the GB analysis do not arise here. Reordering of arguments (see Section 9.4) is independent of passivization. The accusative object is not mentioned at all unlike in GPSG, Categorial Grammar or Bresnan's LFG analysis from before the introduction of Lexical Mapping Theory. The passive can be analyzed directly as the suppression of the subject. Everything else follows from interaction with other principles of grammar.

9.3 Verb position

The analysis of verb position that I will present here is based on the GB-analysis. In HPSG, there are a number of different approaches to describing verb position, however in my opinion, the HPSG variant of the GB analysis is the only adequate one (Müller 2005b,c, 2015b). The analysis of (42) can be summarized as follows: in the verb-initial clauses, there is a trace in verb-final position. There is a special form of the verb in initial position that selects a projection of the verb trace. This special lexical item is licensed by a lexical rule. The connection between the verb and the trace is treated like long-distance dependencies in GPSG via identification of information in the tree or feature structure (structure sharing).

(42) Kennt$_k$ jeder diesen Mann $\__k$?
knows everyone this man
'Does everyone know this man?'

Figure 9.8 on the next page gives an overview of this. The verb trace in final position behaves just like the verb both syntactically and semantically. The information about the missing word is represented as the value of the feature DOUBLE SLASH (abbreviated: DSL). This is a head feature and is therefore passed up to the maximal projection (VP). The verb in initial position has a VP in its SUBCAT list which is missing a verb (VP//V). This is the same verb that was the input for the lexical rule and that would normally occur in final position. In Figure 9.8, there are two maximal verb projections: *jeder diesen Mann* $\__k$ with the trace as the head and *kennt jeder diesen Mann* $\__k$ with *kennt* as the head.

This analysis will be explained in more detail in what follows. For the trace in Figure 9.8, one could assume the following lexical entry:

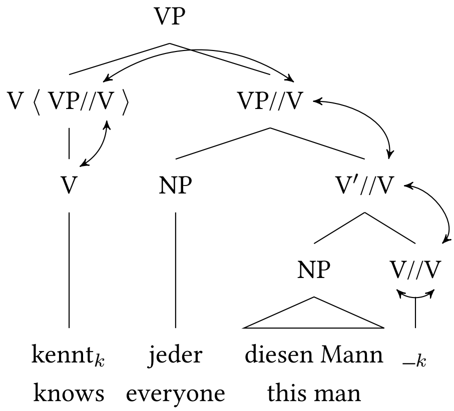

Figure 9.8: Analysis of verb position in HPSG

(43) Verb trace for *kennt* 'knows':

$$\left[\begin{array}{ll}\textsc{phon} & \langle\rangle \\ \textsc{synsem}|\textsc{loc} & \left[\begin{array}{ll}\textsc{cat} & \left[\begin{array}{ll}\textsc{head} & \left[\begin{array}{ll}\textsc{vform} & \textit{fin}\\ \textit{verb} & \end{array}\right] \\ \textsc{subcat} & \left\langle \text{NP}[\textit{nom}]_{\boxed{1}}, \text{NP}[\textit{acc}]_{\boxed{2}} \right\rangle \end{array}\right] \\ \textsc{cont} & \left[\begin{array}{ll}\textsc{ind} & \boxed{3} \\ \textsc{rels} & \left\langle \left[\begin{array}{ll}\textsc{event} & \boxed{3}\\ \textsc{experiencer} & \boxed{1}\\ \textsc{theme} & \boxed{2}\\ \textit{kennen} & \end{array}\right] \right\rangle \end{array}\right] \end{array}\right] \end{array}\right]$$

This lexical entry differs from the normal verb *kennt* only in its PHON value. The syntactic aspects of an analysis with this trace are represented in Figure 9.9 on the following page.

The combination of the trace with *diesen Mann* 'this man' and *jeder* èverbody' follows the rules and principles that we have encountered thus far. This begs the immediate question as to what licenses the verb *kennt* in Figure 9.9 and what status it has.

If we want to capture the fact that the finite verb in initial position behaves like a complementizer (Höhle 1997), then it makes sense to give head status to *kennt* in Figure 9.9 and have *kennt* select a saturated, verb-final verbal projection. Finite verbs in initial position differ from complementizers in that they require a projection of a verb trace, whereas complementizers need projections of overt verbs:

(44) a. dass [jeder diesen Mann kennt]
that everybody this man knows
'that everybody knows this man'

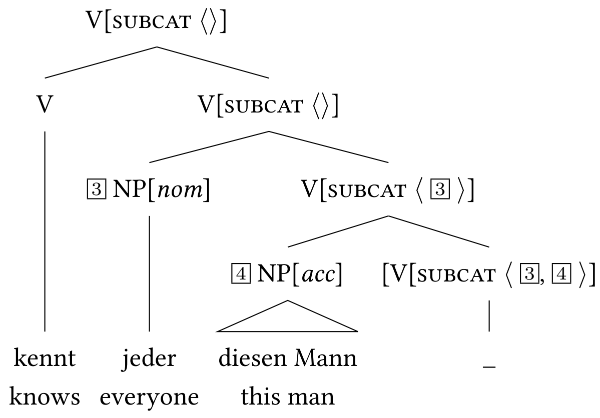

Figure 9.9: Analysis of *Kennt jeder diesen Mann?* ‘Does everyone know this man?’

b. Kennt [jeder diesen Mann _]
knows everybody this man
‘Does everybody know this man?’

It is normally not the case that *kennen* ‘know’ selects a complete sentence and nothing else as would be necessary for the analysis of *kennt* as the head in (44b). Furthermore, we must ensure that the verbal projection with which *kennt* is combined contains the verb trace belonging to *kennt*. If it could contain a trace belonging to *gibt* ‘gives’, for example, we would be able to analyze sentences such as (45b):

(45) a. Gibt [der Mann der Frau das Buch $\_{gibt}$]?
gives the man the woman the book
‘Does the man give the woman the book?’

b. * Kennt [der Mann der Frau das Buch $\_{gibt}$]?
knows the man the woman the book

In the preceding discussion, the dependency between the fronted verb and the verb trace was expressed by coindexation. In HPSG, identity is always enforced by structure sharing. The verb in initial position must therefore require that the trace has exactly those properties of the verb that the verb would have had, were it in final position. The information that must be shared is therefore all locally relevant syntactic and semantic information, that is, all information under LOCAL. Since PHON is not part of the LOCAL features, it is not shared and this is why the PHON values of the trace and verb can differ. Up to now, one crucial detail has been missing in the analysis: the LOCAL value of the trace cannot be directly structure shared with a requirement of the initial verb since the verb *kennt* can only select the properties of the projection of the trace and the SUBCAT list of the selected projection is the empty list. This leads us to the problem that was pointed out in the discussion of (45b). It must therefore be ensured that all information about the verb trace is available on the highest node of its projection. This can be achieved by introducing a head feature whose value is identical to the LOCAL value of the trace. This feature is referred to as DSL. As was already mentioned above, DSL stands for *double slash*. It is called so because it has a similar function to the SLASH feature, which we will encounter in the following section.[14] (46) shows the modified entry for the verb trace:

(46) Verb trace of *kennt* (preliminary version):

$$\begin{bmatrix} \textsc{phon} & \langle\rangle \\ \textsc{synsem}|\textsc{loc}\ \boxed{1} & \begin{bmatrix} \textsc{cat} & \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textsc{vform} & \textit{fin} \\ \textsc{dsl} & \boxed{1} \\ \multicolumn{2}{l}{\textit{verb}} \end{bmatrix} \\ \textsc{subcat} & \left\langle \text{NP}[\textit{nom}]_{\boxed{2}}, \text{NP}[\textit{acc}]_{\boxed{3}} \right\rangle \end{bmatrix} \\ \textsc{cont} & \begin{bmatrix} \textsc{ind} & \boxed{4} \\ \textsc{rels} & \left\langle \begin{bmatrix} \textsc{event} & \boxed{4} \\ \textsc{experiencer} & \boxed{2} \\ \textsc{theme} & \boxed{3} \\ \multicolumn{2}{l}{\textit{kennen}} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Through sharing of the LOCAL value and the DSL value in (46), the syntactic and semantic information of the verb trace is present at its maximal projection, and the verb in initial position can check whether the projection of the trace is compatible.

The special lexical item for verb-initial position is licensed by the following lexical rule:[15]

[14] The feature DSL was proposed by Jacobson (1987a) in the framework of Categorial Grammar to describe head movement in English inversions. Borsley (1989) adopted this idea and translated it into HPSG terms, thereby showing how head movement in a HPSG variant of the CP/IP system can be modeled using DSL. The introduction of the DSL feature to describe head movement processes in HPSG is motivated by the fact that, unlike long-distance dependencies as will be discussed in Section 9.5, this kind of movement is local.

The suggestion to percolate information about the verb trace as part of the head information comes from Oliva (1992).

[15] The lexical rule analysis cannot explain sentences such as (i):

(i) Karl kennt und liebt diese Schallplatte.
Karl knows and loves this record

This has to do with the fact that the lexical rule cannot be applied to the result of coordination, which constitutes a complex syntactic object. If we apply the lexical rule individually to each verb, then we arrive at variants of the verbs which would each select verb traces for *kennen* 'to know' and *lieben* 'to love'. Since the CAT values of the conjuncts are identified with each other in coordinations, coordinations involving the V1 variants of *kennt* and *liebt* would be ruled out since the DSL values of the selected VPs contain the meaning of the respective verbs and are hence not compatible (Müller 2005b: 13). Instead of a lexical rule, one must assume a unary syntactic rule that applies to the phrase *kennt und liebt* 'knows and loves'. As we have seen, lexical rules in the HPSG formalization assumed here correspond to unary rules such that the difference between (47) and a corresponding syntactic rule is mostly a difference in representation.

(47) Lexical rule for verbs in initial position:

$$\begin{bmatrix} \textsc{synsem}|\textsc{loc}\ \boxed{1} & \begin{bmatrix} \textsc{cat}|\textsc{head} & \begin{bmatrix} \textsc{vform} & \textit{fin} \\ \textsc{initial} & - \\ \textit{verb} & \end{bmatrix} \end{bmatrix} \end{bmatrix} \mapsto$$

$$\begin{bmatrix} \textsc{synsem}|\textsc{loc}|\textsc{cat} & \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textsc{vform} & \textit{fin} \\ \textsc{initial} & + \\ \textsc{dsl} & \textit{none} \\ \textit{verb} & \end{bmatrix} \\ \textsc{subcat} & \left\langle \begin{bmatrix} \textsc{loc}|\textsc{cat} & \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textsc{dsl} & \boxed{1} \\ \textit{verb} & \end{bmatrix} \\ \textsc{subcat} & \langle\rangle \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

The verb licensed by this lexical rule selects a maximal projection of the verb trace which has the same local properties as the input verb. This is achieved by the coindexation of the LOCAL values of the input verb and the DSL values of the selected verb projection. Only finite verbs in final position (INITIAL−) can be the input for this rule. The output is a verb in initial-position (INITIAL+).

The corresponding extended analysis is given in Figure 9.10. V1-LR stands for the verb-initial lexical rule.

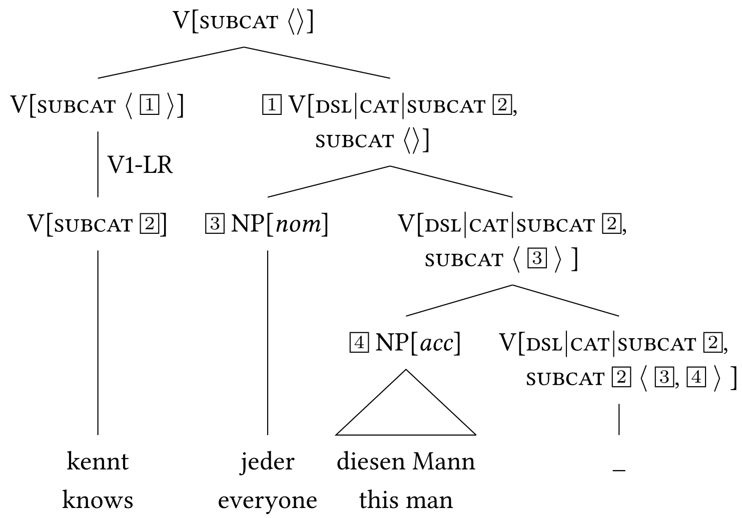

Figure 9.10: Visualization of the analysis of *Kennt jeder diesen Mann?* 'Does everyone know this man?'

The lexical rule in (47) licenses a verb that selects a VP (1 in Figure 9.10). The DSL value of this VP corresponds to the LOCAL value of the verb that is the input of the lexical rule. Part of the DSL value is also the valence information represented in Figure 9.10 ($\boxed{2}$). Since DSL is a head feature, the DSL value of the VP is identical to that of the verb trace and since the LOCAL value of the verb trace is identified with the DSL value, the SUBCAT information of the verb *kennen* is also available at the trace. The combination of the trace with its arguments proceeds exactly as with an ordinary verb.

It would be unsatisfactory if we had to assume a special trace for every verb. Fortunately, this is not necessary as a general trace as in (48) will suffice for the analysis of sentences with verb movement.

(48) General verb trace following Meurers (2000: 206–208):

$$\begin{bmatrix} \text{PHON} \ \langle\rangle \\ \text{SYNSEM}|\text{LOC} \ \boxed{1} \begin{bmatrix} \text{CAT}|\text{HEAD}|\text{DSL} \ \boxed{1} \end{bmatrix} \end{bmatrix}$$

This may seem surprising at first glance, but if we look closer at the interaction of the lexical rule (47) and the percolation of the DSL feature in the tree, then it becomes clear that the DSL value of the verb projection and therefore the LOCAL value of the verb trace is determined by the LOCAL value of the input verb. In Figure 9.10, *kennt* is the input for the verb movement lexical rule. The relevant structure sharing ensures that, in the analysis of (42), the LOCAL value of the verb trace corresponds exactly to what is given in (46).

The most important points of the analysis of verb position are summarized below:

- A lexical rule licenses a special lexical item for each finite verb.
- This lexical item occupies the initial position and requires as its argument a complete projection of a verb trace.
- The projection of the verb trace must have a DSL value corresponding to the LOCAL value of the input verb of the lexical rule.
- Since DSL is a head feature, the selected DSL value is also present on the trace.
- As the DSL value of the trace is identical to its LOCAL value, the LOCAL value of the trace is identical to the LOCAL value of the input verb in the lexical rule.

After discussing the analysis of verb-first sentences, we will now turn to local reordering.

9.4 Local reordering

There are several possibilities for the analysis of constituent order in the middle field: one can assume completely flat structures as in GPSG (Kasper 1994), or instead assume binary branching structures and allow for arguments to be saturated in any order. A compromise was proposed by Kathol (2001) and Müller (1999a, 2002a, 2004c): binary branching structures with a special list that contains the arguments and adjuncts belonging to one head. The arguments and adjuncts are allowed to be freely ordered inside such lists. See Reape (1994) and Section 11.7.2.2 of this book for the formal details of these approaches. Both the completely flat analysis and the compromise have proved to be on the wrong track (see Müller 2005b, 2014c and Müller 2007b: Section 9.5.1) and therefore, I will only discuss the analysis with binary branching structures.

Figure 9.11 shows the analysis of (49a).

(49) a. [weil] jeder diesen Mann kennt
because everyone this man knows

b. [weil] diesen Mann jeder kennt
because this man everyone knows

'because everyone knows this man'

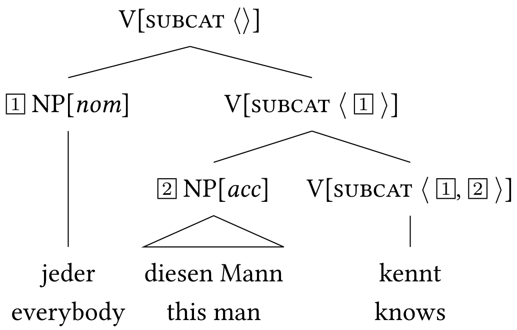

Figure 9.11: Analysis of constituent order in HPSG: normal order

The arguments of the verb are combined with the verb starting with the last element of the SUBCAT list, as explained in Section 9.1.2. The analysis of the marked order is shown in Figure 9.12 on the next page. Both trees differ only in the order in which the elements are taken off from the SUBCAT list: in Figure 9.11, the last element of the SUBCAT list is discharged first and in Figure 9.12 the first one is.

The following schema is a revised version of the Head-Argument Schema:

Schema 3 (Head-Argument Schema (binary branching))
head-argument-phrase ⇒

$$\begin{bmatrix} \text{SYNSEM}|\text{LOC}|\text{CAT}|\text{SUBCAT} \; \boxed{1} \oplus \boxed{3} \\ \text{HEAD-DTR}|\text{SYNSEM}|\text{LOC}|\text{CAT}|\text{SUBCAT} \; \boxed{1} \oplus \langle \boxed{2} \rangle \oplus \boxed{3} \\ \text{NON-HEAD-DTRS} \; \langle \; [\; \text{SYNSEM} \; \boxed{2} \;] \; \rangle \end{bmatrix}$$

Whereas in the first version of the Head-Argument Schema it was always the last element from the SUBCAT list that was combined with the head, the SUBCAT list is divided

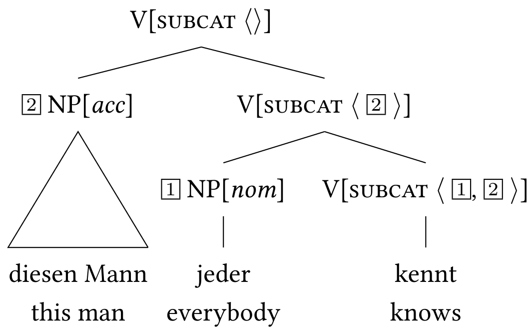

Figure 9.12: Analysis of constituent order in HPSG: marked order

into three parts using *append*: a list of arbitrary length ($\boxed{1}$), a list consisting of exactly one element (⟨ $\boxed{2}$ ⟩) and a further list of arbitrary length ($\boxed{3}$). The lists $\boxed{1}$ and $\boxed{3}$ are combined and the result is the SUBCAT value of the mother node.

Languages with fixed constituent order (such as English) differ from languages such as German in that they discharge the arguments starting from one side (for more on the subject in English, see Section 9.6.1), whereas languages with free constituent order can combine arguments with the verb in any order. In languages with fixed constituent order, either $\boxed{1}$ or $\boxed{3}$ is always the empty list. Since German structures are not restricted with regard to $\boxed{1}$ or $\boxed{3}$, that is $\boxed{1}$ and $\boxed{3}$ can either be the empty list or contain elements, the intuition is captured that there are less restrictions in languages with free constituent order than in languages with fixed order. We can compare this to the Kayneian analysis from Section 4.6.1, where it was assumed that all languages are derived from the base order [specifier [head complement]] (see Figure 4.20 on page 150 for Laenzlinger's analysis of German as an SVO-language (Laenzlinger 2004)). In these kinds of analyses, languages such as English constitute the most basic case and languages with free ordering require some considerable theoretical effort to get the order right. In comparison to that, the analysis proposed here requires more theoretical restrictions if the language has more restrictions on permutations of its constituents. The complexity of the licensed structures does not differ considerably from language to language under an HPSG approach. Languages differ only in the type of branching they have.[16,17]

The analysis presented here utilizing combination of arguments in any order is similar to that of Fanselow (2001) in the framework of GB/MP as well as the Categorial Grammar

[16] This does not exclude that the structures in question have different properties as far as their processability by humans is concerned. See Gibson (1998); Hawkins (1999) and Chapter 15.

[17] Haider (1997b: 18) has pointed out that the branching type of VX languages differs from those of XV languages in analyses of the kind that is proposed here. This affects the c-command relations and therefore has implications for Binding Theory in GB/MP. However, the direction of branching is irrelevant for HPSG analyses as Binding Principles are defined using o-command (Pollard & Sag 1994: Chapter 6) and o-command makes reference to the Obliqueness Hierarchy, that is, the order of elements in the SUBCAT list rather than the order in which these elements are combined with the head.

analyses of Hoffman (1995: Section 3.1) and Steedman & Baldridge (2006). Gunji proposed similar HPSG analyses for Japanese as early as 1986.

9.5 Long-distance dependencies

The analysis of long-distance dependencies utilizes techniques that were originally developed in GPSG: information about missing constituents is passed up the tree (or feature structure).[18] There is a trace at the position where the fronted element would normally occur. Figure 9.13 shows the analysis of (50).

(50) [Diesen Mann]$_j$ kennt$_i$ $\_{j}$ jeder $\_{i}$.
this man knows everyone
'Everyone knows this man.'

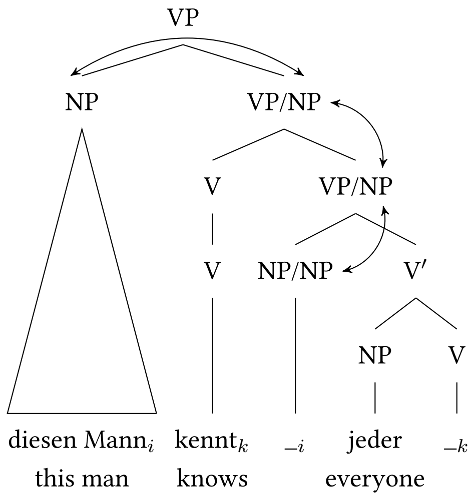

Figure 9.13: Analysis of long-distance dependencies in HPSG

Unlike verb movement, which was discussed in Section 9.3, constituent movement is nonlocal, which is why the two movement types are modeled with different features (SLASH vs. DSL). DSL is a head feature and, like all other head features, projects to the highest node of a projection (for more on the Head Feature Principle, see page 272).

[18] Nothing is actually 'passed up' in a literal sense in feature structures or trees. For expository purposes, it makes sense to explain the analysis as if the structure were built bottom-up, but linguistic knowledge is independent of the direction of processing. In recent computer implementations, structure building is mostly carried out bottom-up but there were other systems which worked top-down. The only thing that is important in the analysis of nonlocal dependencies is that the information about the missing element on all intermediate nodes is identical to the information in the filler and the gap.

SLASH, on the other hand, is a feature that belongs to the NONLOC features represented under SYNSEM|NONLOC. The value of the NONLOC feature is a structure with the features INHERITED (or INHER for short) and TO-BIND. The value of INHER is a structure containing information about elements involved in a long-distance dependency. (51) gives the structure assumed by Pollard & Sag (1994: 163):[19]

(51) $$\begin{bmatrix} \text{QUE} & \textit{list of npros} \\ \text{REL} & \textit{list of indices} \\ \text{SLASH} & \textit{list of local structures} \\ \textit{nonloc} & \end{bmatrix}$$

QUE is important for the analysis of interrogative clauses as is REL for the analysis of relative clauses. Since these will not feature in this book, they will be omitted in what follows. The value of SLASH is a list of *local* objects.

As with the analysis of verb movement, it is assumed that there is a trace in the position where the accusative object would normally occur and that this trace shares the properties of that object. The verb can therefore satisfy its valence requirements locally. Information about whether there has been combination with a trace and not with a genuine argument is represented inside the complex sign and passed upward in the tree. The long-distance dependency can then be resolved by an element in the prefield higher in the tree.

Long-distance dependencies are introduced by the trace, which has a feature corresponding to the LOCAL value of the required argument in its SLASH list. (52) shows the description of the trace as is required for the analysis of (50):

(52) Trace of the accusative object of *kennen* (preliminary):

$$\begin{bmatrix} \text{PHON} & \langle\rangle \\ \text{SYNSEM} & \begin{bmatrix} \text{LOC} & \boxed{1} \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{CAS} & \textit{acc} \\ \textit{noun} & \end{bmatrix} \\ \text{SUBCAT} & \langle\rangle \end{bmatrix} \end{bmatrix} \\ \text{NONLOC} & \begin{bmatrix} \text{INHER}|\text{SLASH} & \langle \boxed{1} \rangle \\ \text{TO-BIND}|\text{SLASH} & \langle\rangle \end{bmatrix} \end{bmatrix} \\ \textit{word} & \end{bmatrix}$$

Since traces do not have internal structure (no daughters), they are of type *word*. The trace has the same properties as the accusative object. The fact that the accusative object is not present at the position occupied by the trace is represented by the value of SLASH.

The following principle is responsible for ensuring that NONLOC information is passed up the tree.

[19] Pollard & Sag assume that the values of QUE, REL, and SLASH are sets rather than lists. The math behind sets is rather complicated, which is why I assume lists here.

Principle 6 (Nonlocal Feature Principle)
In a headed phrase, for each nonlocal feature, the INHERITED *value of the mother is a list that is the concatenation of the* INHERITED *values of the daughters minus the elements in the* TO-BIND *list of the head daughter.*

The Head-Filler Schema licenses the highest node in Figure 9.14 on the facing page:

Schema 4 (Head-Filler Schema)

head-filler-phrase ⇒

$$\begin{bmatrix} \text{HEAD-DTR}|\text{SYNSEM} & \begin{bmatrix} \text{LOC}|\text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{VFORM} & \textit{fin} \\ \text{INITIAL} & + \\ \textit{verb} & \end{bmatrix} \\ \text{SUBCAT} & \langle\rangle \end{bmatrix} \\ \text{NONLOC} & \begin{bmatrix} \text{INHER}|\text{SLASH} & \langle \boxed{1} \rangle \\ \text{TO-BIND}|\text{SLASH} & \langle \boxed{1} \rangle \end{bmatrix} \end{bmatrix} \\ \text{NON-HEAD-DTRS} & \left\langle \begin{bmatrix} \text{SYNSEM} & \begin{bmatrix} \text{LOC} & \boxed{1} \\ \text{NONLOC}|\text{INHER}|\text{SLASH} & \langle\rangle \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

This schema combines a finite, verb-initial clause (INITIAL+) that has an element in SLASH with a non-head daughter whose LOCAL value is identical to the SLASH element. In this structure, no arguments are saturated. Nothing can be extracted from the filler daughter itself, which is ensured by the specification of the SLASH value of the non-head daughter. Figure 9.14 on the next page shows a more detailed variant of the analysis of fronting to the prefield.

The verb movement trace for *kennt* 'knows' is combined with a nominative NP and an extraction trace. The extraction trace stands for the accusative object in the above example. The accusative object is described in the SUBCAT list of the verb ($\boxed{4}$). Following the mechanism for verb movement, the valence information that was originally contained in the entry for *kennt* ($\langle \boxed{3}, \boxed{4} \rangle$) is present on the verb trace. The combination of the projection of the verb trace with the extraction trace works in exactly the same way as for non-fronted arguments. The SLASH value of the extraction trace is passed up the tree and bound off by the Head-Filler Schema.

(52) provides the lexical entry for a trace that can function as the accusative object of *kennen* 'to know'. As with the analysis of verb movement, it is not necessary to have numerous extraction traces with differing properties listed in the lexicon. A more general entry such as the one in (53) will suffice:

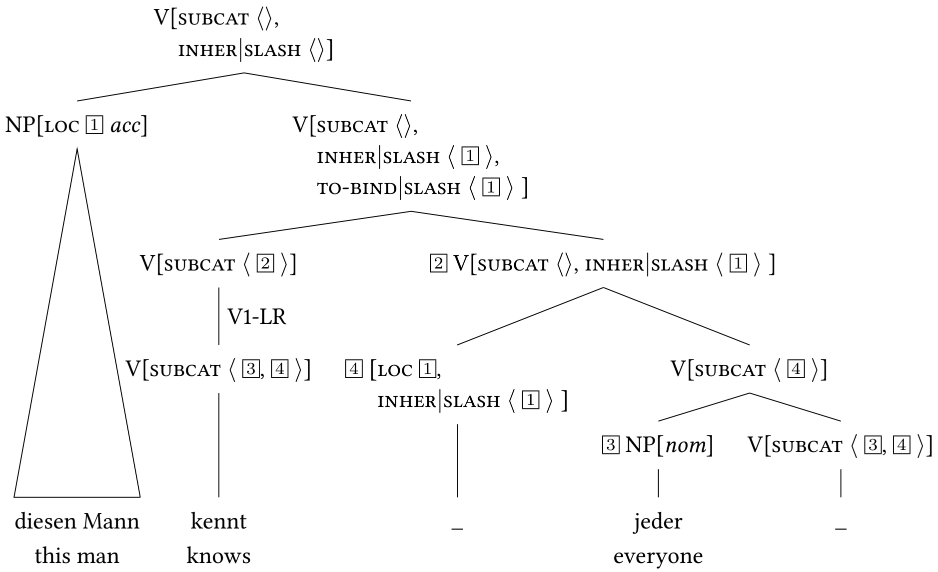

Figure 9.14: Analysis of *Diesen Mann kennt jeder.* 'Everyone knows this man.' combined with the verb movement analysis for verb-initial order

(53) Extraction trace:

$$\left[\begin{array}{ll} \text{PHON} & \langle\rangle \\ \text{SYNSEM} & \left[\begin{array}{ll} \text{LOC} & \boxed{1} \\ \text{NONLOC} & \left[\begin{array}{ll} \text{INHER|SLASH} & \langle \boxed{1} \rangle \\ \text{TO-BIND|SLASH} & \langle\rangle \end{array}\right] \end{array}\right] \\ \textit{word} & \end{array}\right]$$

This has to do with the fact that the head can satisfactorily determine the LOCAL properties of its arguments and therefore also the local properties of the traces that it combines with. The identification of the object in the SUBCAT list of the head with the SYNSEM value of the trace coupled with the identification of the information in SLASH with information about the fronted element serves to ensure that the only elements that can be realized in the prefield are those that fit the description in the SUBCAT list of the head. The same holds for fronted adjuncts: since the LOCAL value of the constituent in the prefield is identified with the LOCAL value of the trace via the SLASH feature, there is then sufficient information available about the properties of the trace.

The central points of the preceding analysis can be summarized as follows: information about the local properties of a trace is contained in the trace itself and then present on all nodes dominating it until one reaches the filler. This analysis can offer an explanation for so-called extraction path marking languages where certain elements show inflection depending on whether they are combined with a constituent out of which something has been extracted in a long-distance dependency. Bouma, Malouf & Sag (2001a) cite Irish, Chamorro, Palauan, Icelandic, Kikuyu, Ewe, Thompson Salish, Moore, French, Spanish, and Yiddish as examples of such languages and provide corresponding references. Since information is passed on step-by-step in HPSG analyses, all nodes intervening in a long-distance dependency can access the elements in that dependency.

9.6 New developments and theoretical variants

This section discusses refinements of the representation of valence information in Subsection 9.6.1 and briefly mentions an important variant of HPSG, namely Linearization-based HPSG in Subsection 9.6.2.

9.6.1 Specifier, complements and argument structure

In this chapter, SUBCAT was assumed as the only valence feature. This corresponds to the state of theory in Pollard & Sag (1994: Chapter 1–8). It has turned out to be desirable to assume at least one additional valence feature and a corresponding schema for the combination of constituents. This additional feature is called SPECIFIER (SPR) and is used in grammars of English (Pollard & Sag 1994: Chapter 9) and German (Müller 2007b: Section 9.3) for the combination of a determiner with a noun. It is assumed that the noun selects its determiner. For the noun *Zerstörung* 'destruction', we have the following CAT value:

(54)
$$\begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{INITIAL} & + \\ \textit{noun} & \end{bmatrix} \\ \text{SPR} & \langle \text{Det} \rangle \\ \text{SUBCAT} & \langle \text{NP[GEN]}, \text{PP[}\textit{durch}\text{]} \rangle \end{bmatrix}$$

Schema 5 can be used just like the Head-Argument Schema for the combination of noun and determiner.

Schema 5 (Specifier-Head Schema)
head-specifier-phrase ⇒
$$\begin{bmatrix} \text{SYNSEM|LOC|CAT|SPR } \boxed{1} & \\ \text{HEAD-DTR|SYNSEM|LOC|CAT} & \begin{bmatrix} \text{SPR} & \boxed{1} \oplus \langle \boxed{2} \rangle \\ \text{SUBCAT} & \langle\rangle \end{bmatrix} \\ \text{NON-HEAD-DTRS } \langle \text{[SYNSEM } \boxed{2}\text{]} \rangle & \end{bmatrix}$$

The analysis of the NP in (55) with the Specifier Schema is shown in Figure 9.15 on the facing page.

(55) die Zerstörung der Stadt durch die Soldaten
the destruction of.the city by the soldiers

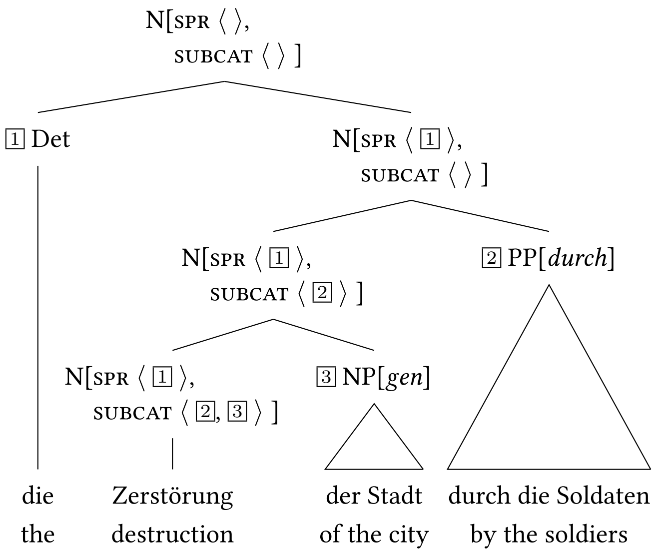

Figure 9.15: NP analysis with valence features SPR

Following the linearization rules discussed in Section 9.1.3, it is ensured that the noun occurs before the complements as the INITIAL value of the noun is '+'. The LP-rule in (56) leads to the determiner being ordered to the left of the noun.

(56) specifier < head

In grammars of English, the SPR feature is also used for the selection of the subject of verbs (Sag, Wasow & Bender 2003: Section 4.3). In a sentence such as (57), the verb is first combined with all its complements (the elements in the SUBCAT or COMPS in newer works) and is then combined with the subject in a second step by applying Schema 5.

(57) Max likes ice cream.

As we have seen in Section 9.4, it makes sense to represent subjects and arguments in the same valence list for the analysis of finite sentences. In this way, the fact can be captured that the order in which a verb is combined with its arguments is not fixed. While the different orders could also be captured by assuming that the subject is selected via SPR, the fact that scrambling is a phenomenon that affects all arguments in the same way would not be covered in a SPR-based analysis. Furthermore, the extraction out of subjects is impossible in languages like English, but it is possible in German (for references and attested examples see p. 538). This difference can be captured by assuming that subjects are selected via SPR in English and that extraction out of elements in the SPR list is prohibited. Since subjects in German are represented on the COMPS list, the fact that they pattern with the objects in terms of possible extractions is captured.

A further expansion from Pollard & Sag (1994: Chapter 9) is the introduction of an additional list that is called ARG-ST in newer works. ARG-ST stands for Argument Structure. The ARG-ST list corresponds to what we encountered as SUBCAT list in this chapter. It contains the arguments of a head in an order corresponding to the Obliqueness Hierarchy. The elements of the list are linked to argument roles in the semantic content of the head (see Section 9.1.6). Binding Theory operates on the ARG-ST list. This level of representation is probably the same for most languages: in every language there are semantic predicates and semantic arguments. Most languages make use of syntactic categories that play a role in selection, so there is both syntactic and semantic selection.[20] Languages differ with regard to how these arguments are realized. In English, the first element in the valence list is mapped to the SPR list and the remaining arguments to the SUBCAT (or COMPS list in more recent work). In German, the SPR list of verbs remains empty. (58) shows some relevant examples for German and English.

(58) a.
$$\begin{bmatrix} \text{PHON} & \langle \textit{schlag} \rangle \\ \text{SYNSEM}|\text{LOC} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{SPR} & \langle\rangle \\ \text{SUBCAT} & \boxed{1} \\ \text{ARG-ST} & \boxed{1} \left\langle \text{NP}[\textit{str}]_{\boxed{2}}, \text{NP}[\textit{str}]_{\boxed{3}} \right\rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{IND} & \boxed{4}\ \textit{event} \\ \text{RELS} & \left\langle \begin{bmatrix} \text{EVENT} & \boxed{4} \\ \text{AGENT} & \boxed{2} \\ \text{PATIENT} & \boxed{3} \\ \textit{schlagen} & \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

b.
$$\begin{bmatrix} \text{PHON} & \langle \textit{beat} \rangle \\ \text{SYNSEM}|\text{LOC} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{SPR} & \langle \boxed{1} \rangle \\ \text{SUBCAT} & \boxed{2} \\ \text{ARG-ST} & \left\langle \boxed{1}\, \text{NP}[\textit{str}]_{\boxed{3}} \right\rangle \oplus \boxed{2} \left\langle \text{NP}[\textit{str}]_{\boxed{4}} \right\rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{IND} & \boxed{5}\ \textit{event} \\ \text{RELS} & \left\langle \begin{bmatrix} \text{EVENT} & \boxed{5} \\ \text{AGENT} & \boxed{3} \\ \text{PATIENT} & \boxed{4} \\ \textit{beat} & \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

One can view the ARG-ST list as the equivalent to Deep Structure in GB theory: semantic

[20] Koenig & Michelson (2012) argue for an analysis of Oneida (a Northern Iroquoian language) that does not include a representation of syntactic valence. If this analysis is correct, syntactic argument structure would not be universal, but would be characteristic for a large number of languages.

roles are assigned with reference to this list. The difference is that there is no ordered tree that undergoes transformations. The question of whether all languages can be derived from either VO or OV order therefore becomes irrelevant.

9.6.2 Linearization-based HPSG

The schemata that were presented in this chapter combine adjacent constituents. The assumption of adjacency can be dropped and discontinuous constituents maybe permitted. Variants of HPSG that allow for discontinuous constituents are usually referred to as *Linearization-based HPSG*. The first formalization was developed by Mike Reape (1991, 1992, 1994). Proponents of linearization approaches are for instance Kathol (1995, 2000); Donohue & Sag (1999); Richter & Sailer (1999b); Crysmann (2008); Beavers & Sag (2004); Sato (2006); Wetta (2011). I also suggested linearization-based analyses (Müller 1999a, 2002a) and implemented a large-scale grammar fragment based on Reape's ideas (Müller 1996c). Linearization-based approaches to the German sentence structure are similar to the GPSG approach in that it is assumed that verb and arguments and adjuncts are members of the same linearization domain and hence may be realized in any order. For instance, the verb may precede arguments and adjuncts or follow them. Hence, no empty element for the verb in final position is necessary. While this allows for grammars without empty elements for the analysis of the verb position, it is unclear how examples with apparent multiple frontings can be accounted for, while such data can be captured directly in the proposal suggested in this chapter. The whole issue is discussed in more detail in Müller (2015b). I will not explain Reape's formalization here, but defer its discussion until Section 11.7.2.2, where the discontinuous, non-projective structures of some Dependency Grammars are compared to linearization-based HPSG approaches. Apparent multiple frontings and the problems they pose for simple linearization-based approaches are discussed in Section 11.7.1.

9.7 Summary and classification

In HPSG, feature descriptions are used to model all properties of linguistic objects: roots, words, lexical rules and dominance schemata are all described using the same formal tools. Unlike GPSG and LFG, there are no separate phrase structure rules. Thus, although HPSG stands for Head-Driven Phrase Structure Grammar, it is not a phrase structure grammar. In HPSG implementations, a phrase structure backbone is often used to increase the efficiency of processing. However, this is not part of the theory and not linguistically necessary.

HPSG differs from Categorial Grammar in that it assumes considerably more features and also in that the way in which features are grouped plays an important role for the theory.

Long-distance dependencies are not analyzed using function composition as in Categorial Grammar, but instead as in GPSG by appealing to the percolation of information in the tree. In this way, it is possible to analyze pied-piping constructions such as those discussed in Section 8.6 with just one lexical item per relative pronoun, whose relevant local properties are identical to those of the demonstrative pronoun. The relative clause in (59) would be analyzed as a finite clause from which a PP has been extracted:

(59) der Mann, [$_{\text{RS}}$ [$_{\text{PP}}$ an den] [$_{\text{S/PP}}$ wir gedacht haben]]
the man on who we thought have
'the man we thought of'

For relative clauses, it is required that the first daughter contains a relative pronoun. This can, as shown in the English examples on page 256, be in fact very deeply embedded. Information about the fact that *an den* 'of whom' contains a relative pronoun is provided in the lexical entry for the relative pronoun *den* by specifying the value of NONLOC|INHER|REL. The Nonlocal Feature Principle passes this information on upwards so that the information about the relative pronoun is contained in the representation of the phrase *an den*. This information is bound off when the relative clause is put together (Pollard & Sag 1994: Chapter 5; Sag 1997). It is possible to use the same lexical entry for *den* in the analyses of both (59) and (60) as – unlike in Categorial Grammar – the relative pronoun does not have to know anything about the contexts in which it can be used.

(60) der Mann, [$_{\text{RS}}$ [$_{\text{NP}}$ den] [$_{\text{S/NP}}$ wir kennen]]
the man that we know
'the man that we know'

Any theory that wants to maintain the analysis sketched here will have to have some mechanism to make information available about the relative pronoun in a complex phrase If we have such a mechanism in our theory – as is the case in LFG and HPSG – then we can also use it for the analysis of long-distance dependencies. Theories such as LFG and HPSG are therefore more parsimonious with their descriptive tools than other theories when it comes to the analysis of relative phrases.

In the first decade of HPSG history (Pollard & Sag 1987, 1994; Nerbonne, Netter & Pollard 1994), despite the differences already mentioned here, HPSG was still very similar to Categorial Grammar in that it was a strongly lexicalized theory. The syntactic make-up and semantic content of a phrase was determined by the head (hence the term *head-driven*). In cases where head-driven analyses were not straight-forwardly possible, because no head could be identified in the phrase in question, then it was commonplace to assume empty heads. An example of this is the analysis of relative clauses in Pollard & Sag (1994: Chapter 5). Since an empty head can be assigned any syntactic valence and an arbitrary semantics (for discussion of this point, see Chapter 19) , one has not really explained anything as one needs very good reasons for assuming an empty head, for example that this empty position can be realized in other contexts. This is, however, not the case for empty heads that are only proposed in order to save theoretical assumptions. Therefore, Sag (1997) developed an analysis of relative clauses without any empty elements. As in the analyses sketched for (59) and (60), the relative phrases are combined directly with the partial clause in order to form the relative clause. For the various observable types of relative clauses in English, Sag proposes different dominance rules. His analysis constitutes a departure from strong lexicalism: in Pollard & Sag (1994), there are six dominance schemata, whereas there are 23 in Ginzburg & Sag (2000).

The tendency to a differentiation of phrasal schemata can also be observed in the proceedings of recent conferences. The proposals range from the elimination of empty elements to radically phrasal analyses (Haugereid 2007, 2009).[21]

Even if this tendency towards phrasal analyses may result in some problematic analyses, it is indeed the case that there are areas of grammar where phrasal analyses are required (see Section 21.10). For HPSG, this means that it is no longer entirely head-driven and is therefore neither Head-Driven nor Phrase Structure Grammar.

HPSG makes use of typed feature descriptions to describe linguistic objects. Generalizations can be expressed by means of hierarchies with multiple inheritance. Inheritance also plays an important role in Construction Grammar. In theories such as GPSG, Categorial Grammar and TAG, it does not form part of theoretical explanations. In implementations, macros (abbreviations) are often used for co-occurring feature-value pairs (Dalrymple, Kaplan & King 2004). Depending on the architecture assumed, such macros are not suitable for the description of phrases since, in theories such as GPSG and LFG, phrase structure rules are represented differently from other feature-value pairs (however, see Asudeh, Dalrymple & Toivonen (2008, 2013) for macros and inheritance used for c-structure annotations). Furthermore, there are further differences between types and macros, which are of a more formal nature: in a typed system, it is possible under certain conditions to infer the type of a particular structure from the presence of certain features and of certain values. With macros, this is not the case as they are only abbreviations. The consequences for linguistic analyses made by this differences are, however, minimal.

HPSG differs from GB theory and later variants in that it does not assume transformations. In the 80s, representational variants of GB were proposed, that is, it was assumed that there was no D-structure from which an S-structure is created by simultaneous marking of the original position of moved elements. Instead, one assumed the S-structure with traces straight away and the assumption that there were further movements in the mapping of S-structure to Logical Form was also abandoned (Koster 1978; Haider 1993: Section 1.4; Frey 1993: 14). This view corresponds to the view in HPSG and many of the analyses in one framework can be translated into the other.

In GB theory, the terms subject and object do not play a direct role: one can use these terms descriptively, but subjects and objects are not marked by features or similar devices. Nevertheless it is possible to make the distinction since subjects and objects are usually realized in different positions in the trees (the subject in specifier position of IP and the object as the complement of the verb). In HPSG, subject and object are also not primitives of the theory. Since valence lists (or ARG-ST lists) are ordered, however, this means that it is possible to associate the ARG-ST elements to grammatical functions: if there is a subject, this occurs in the first position of the valence list and objects follow.[22]

[21] For discussion, see Müller (2007c) and Section 21.3.6.

[22] When forming complex predicates, an object can occur in first position. See Müller (2002a: 157) for the long passive with verbs such as *erlauben* 'allow'. In general, the following holds: the subject is the first

For the analysis of (61b) in a transformation-based grammar, the aim is to connect the base order in (61a) and the derived order in (61b). Once one has recreated the base order, then it is clear what is the subject and what is the object. Therefore, transformations applied to the base structure in (61a) have to be reversed.

(61) a. [weil] jeder diesen Mann kennt
because everyone this man knows
'because everyone knows this man'

b. [weil] diesen Mann jeder kennt
because this man everyone knows

In HPSG and also in other transformation-less models, the aim is to assign arguments in the order in (61b) to descriptions in the valence list. The valence list (or ARG-ST in newer approaches) corresponds in a sense to Deep Structure in GB. The difference is that the head itself is not included in the argument structure, whereas this is the case with D-structure.

Bender (2008c) has shown how one can analyze phenomena from non-configurational languages such as Wambaya by referring to the argument structure of a head. In Wambaya, words that would normally be counted as constituents in English or German can occur discontinuously, that is an adjective that semantically belongs to a noun phrase and shares the same case, number and gender values with other parts of the noun phrase can occur in a position in the sentence that is not adjacent to the remaining noun phrase. Nordlinger (1998) has analyzed the relevant data in LFG. In her analysis, the various parts of the constituent refer to the f-structure of the sentence and thus indirectly ensure that all parts of the noun phrase have the same case. Bender adopts a variant of HPSG where valence information is not removed from the valence list after an argument has been combined with its head, but rather this information remains in the valence list and is passed up towards the maximal projection of the head (Meurers 1999c; Przepiórkowski 1999b; Müller 2007b: Section 17.4). Similar proposals were made in GB by Higginbotham (1985: 560) and Winkler (1997). By projecting the complete valence information, it remains accessible in the entire sentence and discontinuous constituents can refer to it (e. g. via MOD) and the respective constraints can be formulated.[23] In this analysis, the argument structure in HPSG corresponds to f-structure in LFG. The extended head domains of LFG, where multiple heads can share the same f-structure, can also be modeled in HPSG. To this end, one can utilize function composition as it was presented in the chapter on Categorial Grammar (see Chapter 8.5.2). The exact way in which this is translated into HPSG cannot be explained here due to space restrictions. The reader is referred to the original works by Hinrichs & Nakazawa (1994) and the explanation in Müller (2007b: Chapter 15).

argument with structural case.

[23] See also Müller (2008a) for an analysis of depictive predicates in German and English that makes reference to the list of realized or unrealized arguments of a head, respectively. This analysis is also explained in Section 18.2.

Valence information plays an important role in HPSG. The lexical item of a verb in principle predetermines the set of structures in which the item can occur. Using lexical rules, it is possible to relate one lexical item to other lexical items. These can be used in other sets of structures. So one can see the functionality of lexical rules in establishing a relation between sets of possible structures. Lexical rules correspond to transformations in Transformational Grammar. This point is discussed in more detail in Section 19.5. The effect of lexical rules can also be achieved with empty elements. This will also be the matter of discussion in Section 19.5.

In GPSG, metarules were used to license rules that created additional valence patterns for lexical heads. In principle, metarules could also be applied to rules without a lexical head. This is explicitly ruled out by Flickinger (1983) and Gazdar et al. (1985: 59) using a special constraint. Flickinger, Pollard & Wasow (1985: 265) pointed out that this kind of constraint is unnecessary if one uses lexical rules rather than metarules since the former can only be applied to lexical heads.

For a comparison of HPSG and Stabler's Minimalist Grammars, see Section 4.6.4.

Comprehension questions

1. What status do syntactic trees have in HPSG?

2. How does case assignment take place in the analysis of example (62)?

   (62) Dem Mann wurde ein Buch geschenkt.
   the.DAT man was a.NOM book given
   'The man was given a book.'

3. What is *linking* and how is it accounted for in HPSG?

Exercises

1. Give a feature description for (63) ignoring *dass*.

   (63) [dass] Max lacht
   that Max laughs

2. The analysis of the combination of a noun with a modifying adjective in Section 9.1.7 was just a sketch of an analysis. It is, for example, not explained how one can ensure that the adjective and noun agree in case. Consider how it would be possible to expand such an analysis so that the adjective-noun combination in (64a) can be analyzed, but not the one in (64b):

   (64) a. eines interessanten Mannes
   an.GEN interesting.GEN man.GEN

   b. * eines interessanter Mannes
   an.GEN interesting.NOM man.GEN

Further reading

Here, the presentation of the individual parts of the theory was – as with other theories – kept relatively short. For a more comprehensive introduction to HPSG, including motivation of the feature geometry, see Müller (2007b). In particular, the analysis of the passive was sketched in brief here. The entire story including the analysis of unaccusative verbs, adjectival participles, modal infinitives as well as diverse passive variants and the long passive can be found in Müller (2002a: Chapter 3) and Müller (2007b: Chapter 17).

Overviews of HPSG can be found in Levine & Meurers (2006), Przepiórkowski & Kupść (2006), Bildhauer (2014) and Müller (2015c).