Issues in a German core grammar in HPSG

Frank Richter and Manfred Sailer
Seminar für Sprachwissenschaft
Universität Tübingen
SFB 340, B8

Tübingen, October 3rd, 1996

1 HPSG from a Logical Point of View

For specifying HPS grammars we use a logic of description based on SRL as defined in (King 1989) and (King 1994) and extend it by relations. Without going into technical details, the fragment grammar uses the following mathematical concepts:

1. A signature Σ which fixes the sort lattice, the attribute names, the relation names and the appropriateness function. Σ has to be defined explicitly by the linguist writing the grammar.

2. An interpretation I for the signature which provides a set of objects, assigns each object a maximally specific sort, interprets the relation names with respect to the arity function and interprets the attributes with respect to the appropriateness function. Thus the interpretation is tied to Σ.

3. The signature generates a set of descriptions $\mathfrak{D}$ which get a classical set theoretic interpretation.

4. A finite set of clauses C restricts the denotation of the relational symbols.

5. A grammar Γ is a triple consisting of a signature Σ, a theory θ, which is a set of descriptions, and a finite set of clauses C.

6. The linguistically intended models of Γ are called exhaustive models. An exhaustive model for Γ is such that every object in it satisfies every description of θ, and it contains every object that is permitted by θ. Losely speaking, we can say that exhaustive models are the biggest models permitted by θ which respect Σ and C.

We abbreviate descriptions as AVMs. We call the descriptions that belong to the theory θ of a grammar constraints.

We extend the AVM notation to the abbreviation of formal definitions of clauses. It should be noted, though, that while tags in traditional AVMs always abbreviate path equations, in clauses they sometimes may abbreviate variables. In the present context we can ignore this important formal distinction.

2 Major Principles for Signs

Among the constraints that are imposed on objects of sort *sign*, three restrictions are of special interest. It has to be said what the words in the grammar look like, how signs can be combined to form new phrases and how the word order in phrases is constrained:

- lexicon: words in the grammar
- ID principle: combining signs
- linearization module: word order in phrases

2.1 Words in the Fragment

In the logical framework outlined in section 1 the lexicon of a grammar is treated just like the other principles of grammar. This means that the lexicon is specified by means of a description in the theory. We call it the word principle.

(1) *The word principle*

$$\begin{bmatrix} \textit{word} \end{bmatrix} \quad \Longrightarrow \quad \mathrm{Lex}_1 \vee \mathrm{Lex}_2 \vee \ldots \vee \mathrm{Lex}_n \vee \mathrm{LR}$$

As an abbreviation for (1) we write $\delta_{lexicon}$. (1) says that every object of sort *word* must be in the denotation of (at least) one of the disjuncts on the righthand side of the implication. We call the descriptions Lex_1 to Lex_n, $n \in \mathbb{N}$, *lexical entries*. Alternatively, a word may be in the denotation of the output of a lexical rule. This is indicated by the last disjunct in the word principle, which is abbreviated as LR.[1]

(2) *Example for a lexical entry: the finite verb geht*

$$\boxed{1} \begin{bmatrix} \textit{word} & \\ \textsc{synsem loc cat} & \begin{bmatrix} \textit{cat} & \\ \textsc{head} & \begin{bmatrix} \textit{verb} & \\ \textsc{vform} & \textit{finite} \\ \textsc{mod} & \textit{minus} \end{bmatrix} \\ \textsc{val comps} & \langle \textsc{dp}[\textit{complete}] \rangle \end{bmatrix} \\ \texttt{construct-phon} & \left(\boxed{1}, \langle \textit{geht} \rangle\right) \end{bmatrix}$$

[1] For a formal treatment of lexical rules as descriptions see the DLR approach in (Calcagno and Meurers 1996).

When we have fully specified the lexicon of our grammar, we have exactly one constraint in the theory:

$$\theta_{2.1} = \{\delta_{lexicon}\}$$

2.2 ID Schemata

In section 2.1 we have shown how we use a description to say which objects of sort *word* exist in our grammar. Suppose we have specified the lexicon. Then we have a finite or infinite number of *word* objects in our grammar, and all of them represent a well-formed word.

The next thing to do is to restrict the possible combinations of words and phrases when they form new phrases. For this purpose we have to define what the phrases which we get from combining signs have to look like.

To license objects of sort *phrase* (Pollard and Sag 1994) formulate so-called ID schemata. By convention ID schemata are disjoint disjuncts in the consequence of a description in the theory. Pollard and Sag (1994) define six ID schemata. Simplifying a bit, their ID principle is of the form

$$\begin{bmatrix} \textit{headed-phrase} \\ \end{bmatrix} \implies \textit{Desc}_{ID_1} \vee \textit{Desc}_{ID_2} \vee \textit{Desc}_{ID_3} \vee \textit{Desc}_{ID_4} \vee \textit{Desc}_{ID_5} \vee \textit{Desc}_{ID_6}$$

where *headed-phrase* is a subsort of *phrase*.[2]

In the fragment grammar we have three subsorts of the sort *phrase*[3], and we do not introduce a subsort *headed-phrase* for *phrase*. For each subsort of *phrase* there is exactly one ID schema. Thus our ID principle says that an object whose sort is one of the subsorts of *phrase* must be described by one of the following three ID schemata:

(3) *The ID schema for head-adjunct phrases*

$$\begin{bmatrix} \textit{ha-phrase} \\ \text{SYNSEM LOC CAT VAL SPR } \boxed{1} \end{bmatrix}$$

N:
$$\begin{bmatrix} \text{SYNSEM} & \begin{bmatrix} \text{STATUS} & \textit{complete} \\ \text{LOC CAT HEAD} & \begin{bmatrix} \textit{adverbial} \\ \text{MOD } \boxed{2} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

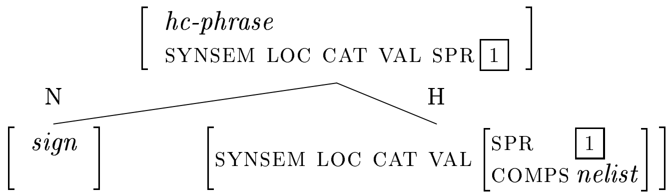
H:
$$\begin{bmatrix} \text{SYNSEM } \boxed{2} & \begin{bmatrix} \text{STATUS} & \textit{incomplete} \\ \text{LOC CAT VAL SPR } \boxed{1} \end{bmatrix} \end{bmatrix}$$

With (3) we state that in a *ha-phrase* the nonhead daughter has to be a projection of an adverbial, and the adjunct demands certain properties of the head daughter.

[2] Pollard and Sag (1994) don't use a sort *headed-phrase*. Instead they define an attribute DTRS on *phrase*, and use a sort hierarchy under *const-struc* (which is the value of the appropriateness function on $\langle$*phrase*, DTRS$\rangle$) to distinguish different kinds of phrases.

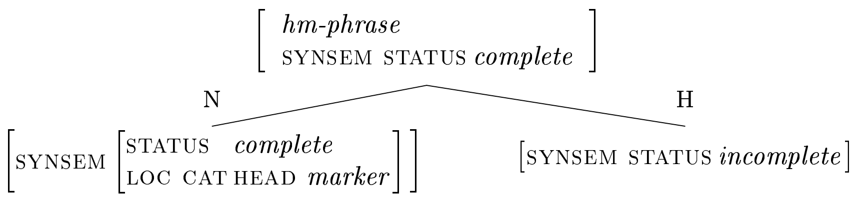
[3] Actually, we have four subsorts of *phrase*, but for simplicity we ignore *hspr-phrase* (head-specifier-phrase), which is only used in the analysis of DPs.

(4) *The ID schema for head-complement phrases*



A phrase can only be a head-complement phrase, if the head daughter has a nonempty COMPS-list. In other words, the head of the construction is still not saturated. In interaction with the Subcategorization Principle of the fragment, (4) ensures that the nonhead daughter in head-complement phrases is in fact a complement of the head of the construction.

(5) *The ID schema for head-marker phrases*



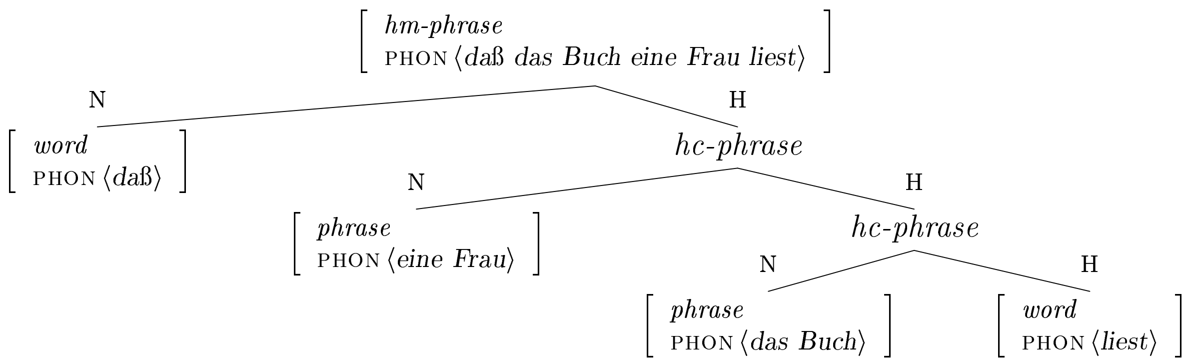
In head-marker phrases, the nonhead daughter must be a marker.
With the word principle and the ID principle we have now two constraints in the theory of the fragment grammar:

$$\theta_{2.2} = \{\delta_{lexicon}, \delta_{ID-principle}\}$$

2.3 Linearization in Permutation Domains

By virtue the two principles of section 2.1 and 2.2 we have a lexicon of words, and we know how to combine signs. We have not said anything yet about the word order in the phrases we get when we combine signs.

(6) *A verbfinal sentence*



Linearization concerns all objects of sort *phrase*:

(7) *The linearization principle*

$$\begin{bmatrix} phrase \end{bmatrix} \Longrightarrow \boxed{2}\begin{bmatrix} phrase \\ \text{PHON } \boxed{1} \\ \texttt{build-phon}(\boxed{2},\boxed{1}) \\ \texttt{lp}(\boxed{2},\boxed{1}) \end{bmatrix}$$

The relation `build-phon` basically says that the phonology of a phrase is an arbitrary permutation of the phonologies of its daughters. The word order statements that are called with the relation `lp` restrict the actual word order. As a consequence of this architecture, the entire word order module of the fragment grammar is formalized in clauses rather than in the theory. The only principle in the theory of grammar that concerns word order is the one given in (7).

Independent of the distinction between verbfinal, verbfirst and verbsecond sentences, we obtain a uniform syntactic sentence structure. For adverbials we use an underspecified semantics.

Including the linearization principle, there are now three constraints in the theory of the fragment:

$$\theta_{2.3} = \{\delta_{lexicon}, \delta_{ID-principle}, \delta_{linearization}\}$$

As to the word order in sentences, the linearization module of the fragment closely follows the traditional theory of topological fields in German as presented in (Höhle 1986, p.330)[4]:

(8) *Topological fields for verbfinal sentences*

C	X	VK	Y
ob	wohl jemand den Hund	füttern würde	morgen abend?
daß	du mir ja nicht den Hund	fütterst!	
in dessen Hütte	eine Katze	gefunden wurde	
wer		kommt	

(9) *Topological fields for verbfirst sentences*

FINIT	X	VK	Y
würde	denn jemand den Hund	füttern	morgen abend
kommen	die Leute denn auch,		die du eingeladen hast
hättest	du mir doch	zugehört	damals
gib	mir mal einer den Hammer!		

[4]For the present purpose we can ignore the topological fields KOORD and K_L

(10) *Topological fields for verbsecond sentences*

K	FINIT	X	VK	Y
die Frau	hat	das Buch	gelesen	gestern abend
es	hört	schon	auf	
wer	kommt?			
dann	hör	doch	zu!	

In our adaptation of the theory of topological fields for HPSG we introduce the notion 'opener'. Technically, an opener is a word whose head bears the attribute FRONTED. To see what this means consider the following table:

(11) *Topological fields and linearization*

Opener	Middlefield	Verbal Complex	Extraposed Material
C	X	VK	Y
FINIT	X	VK	Y
K FINIT	X	VK	Y

(11) implies the linearization statement which covers all three sentence types of German:

(12) *Word order of the opener in sentences*

In a sentence, the opener's phonology precedes the phonology of all other elements.

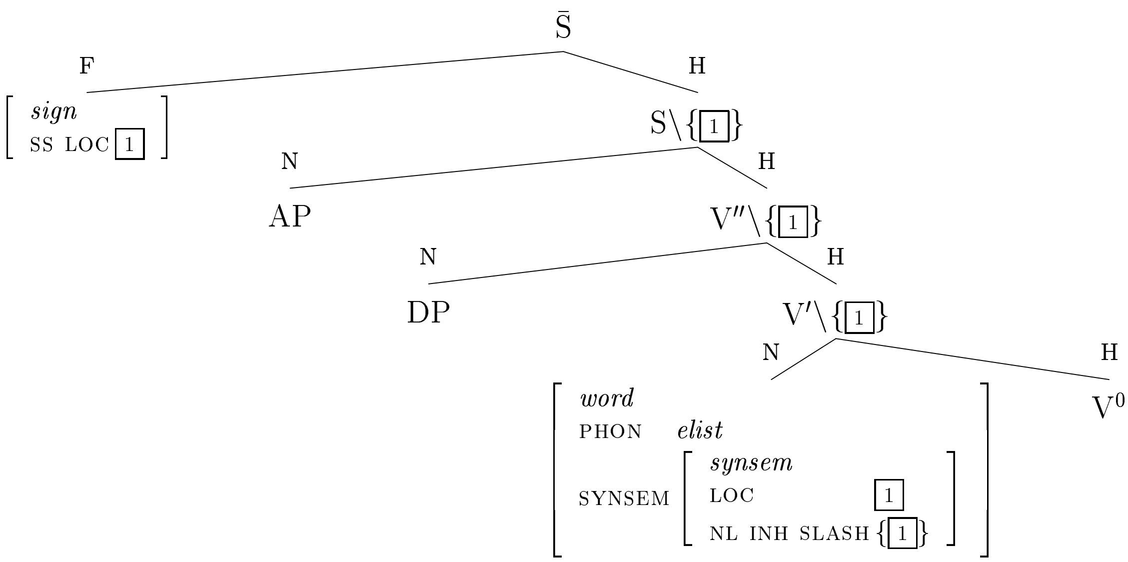
(12) comprises everything that has to be said about word order at the left periphery of German sentences.

3 Movement as Relation

Properties of long movement as in (Pollard and Sag 1994):

1. A signature that comprises attributes and sorts that are needed for the movement principles.
2. A lexical entry for traces that shares LOCAL and NONLOCAL values.
3. A percolation mechanism that passes on the slash values in the desired way.
4. A specific ID schema for fillers.
5. A principle that binds off slashes at head filler phrases.

(13) *The traditional analysis of movement in HPSG*



We modify the lexical entry of traces according to a simplified ontology that does not comprise slashes, and we dismiss the rest of the traditional theory of movement in HPSG entirely. In effect, an essentially local treatment of movement is replaced by a global treatment. A relation `move-`α is defined which establishes the connection between fillers and gaps. As the name of the relation indicates, we are aiming at a theory of movement in HPSG that exploits some of the concepts that have been developed in GB theory.
To capture long movement, one more principle is introduced into the theory of the fragment:

(14) *The frontedbesetzung principle* $\delta_{frontedbesetzung}$

A finite sentence respects frontedbesetzung.

Frontedbesetzung turns out to be a relation which itself calls another relation, `move-`α.

(15) *The relevant clause of the definition of* `frontedbesetzung`

$$\texttt{frontedbesetzung}\left(\boxed{1}\begin{bmatrix} \textit{phrase} \end{bmatrix}\right) :=$$

$$\texttt{contains-fr-word}\left(\boxed{1}\begin{bmatrix} \textit{phrase} \end{bmatrix}, \begin{bmatrix} \text{SS LOC CAT HD FR CONST} \begin{bmatrix} \textit{sign} \\ \text{SS LOC}\ \boxed{2}\begin{bmatrix} \textit{loc} \\ \text{CAT HD MOD}\ \textit{minus} \end{bmatrix} \end{bmatrix} \end{bmatrix}\right)$$

$$\wedge\ \texttt{move-}\alpha\left(\boxed{1}\begin{bmatrix} \textit{phrase} \end{bmatrix}, \boxed{2}\begin{bmatrix} \textit{loc} \end{bmatrix}, \begin{bmatrix} \top \end{bmatrix}\right)$$

(16) *The relation* `move-`α

$$\texttt{move-}\alpha\left(\boxed{1}\begin{bmatrix} \textit{phrase} \end{bmatrix},\ \boxed{2}\begin{bmatrix} \textit{loc} \end{bmatrix},\ \boxed{3}\begin{bmatrix} \textit{word} & \\ \text{PHON} & \textit{elist} \\ \text{SYNSEM LOC} & \boxed{2} \end{bmatrix}\right) := \texttt{movedom}(\boxed{1}, \boxed{3})$$

`move-`α uses `movedom` to establish the link between the antecedent and its trace.

(17) *The clauses of the relation* `movedom`

$$\texttt{movedom}\left(\begin{bmatrix} \textit{phrase} \\ \text{HEAD-DTR}\ \boxed{1} \end{bmatrix},\ \boxed{2}\right) := \texttt{movedom}(\boxed{1}, \boxed{2})$$

$$\texttt{movedom}\left(\begin{bmatrix} \textit{phrase} \\ \text{NONHEAD-DTR}\ \boxed{1} \end{bmatrix},\ \boxed{2}\right) := \texttt{movedom}(\boxed{1}, \boxed{2})$$

$$\texttt{movedom}\left(\begin{bmatrix} \textit{phrase} \\ \text{NONHEAD-DTR}\ \boxed{1} \end{bmatrix},\ \boxed{1}\right) := \top$$

With (14) we have discussed four constraints in the theory of the fragment:

$$\theta_3 = \{\delta_{lexicon}, \delta_{ID-principle}, \delta_{linearization}, \delta_{frontedbesetzung}\}$$

References

Calcagno, M. and Meurers, D.: 1996, A Semantics for Lexical Rules in HPSG, Handout presented at the Lexical Rule Workshop at the Universität Tübingen, July 7th.

Höhle, T. N.: 1986, Der Begriff "Mittelfeld". Anmerkungen über die Theorie der topologischen Felder, *in* A. Schöne (ed.), *Kontroversen alte und neue. Akten des 7. Internationalen Germanistenkongresses Göttingen 1985*, Vol. 3, pp. 329–340.

King, P. J.: 1989, *A logical Formalism for Head-Driven Phrase Structure Grammar*, PhD thesis, University of Manchester.

King, P. J.: 1994, An expanded logical formalism for Head-Driven Phrase Structure Grammar, *Arbeitspapiere des SFB 340 59*, Universität Tübingen.

Pollard, C. and Sag, I.: 1994, *Head-Driven Phrase Structure Grammar*, University of Chicago Press.