Grammar Formalisms
Head-Driven Phrase Structure Grammar (HPSG)

Laura Kallmeyer, Timm Lichte, Wolfgang Maier
Universität Tübingen

27.06.2007
Outline

- Feature Structure Basics
- The building blocks of HPSG
- Analyzing raising and control
- Analyzing long-distance dependencies
Subsumption and Unification

\[
\begin{array}{ll}
\text{ATTR}_1 & \text{val}_1 \\
\text{ATTR}_2 & \text{val}_2 \\
\ldots & \ldots \\
\text{ATTR}_n & \text{val}_n \\
\end{array}
\{<\text{ATTR}_1,\text{val}_1>, <\text{ATTR}_2,\text{val}_2>, \\
\ldots, <\text{ATTR}_n,\text{val}_n>\}
\]

**Subsumption** $\sqsubseteq$: $A \sqsubseteq B$, iff

if $t \in A$, then $t \in B$.

**Unification** $\sqcup$: $A \sqcup B = C$, iff

$C$ is the smallest feature structure such that $A \sqsubseteq C$ and $B \sqsubseteq C$. 
Feature structures as values:

- non-recursive:
  \[
  \begin{bmatrix}
  \text{INDEX} & \begin{bmatrix}
  \text{PER} & 1 \\
  \text{NUM} & \text{sg} \\
  \text{GEND} & \text{fem}
  \end{bmatrix}
  \end{bmatrix}
  \]

- recursive:
  \[
  \begin{bmatrix}
  \text{SUBCAT} & \left< \left[ \text{SUBCAT} \left[ \ldots \right] \right] \right>
  \end{bmatrix}
  \]

Other than FTAG, HPSG uses recursive feature structures!
Re-entrancies (or “structure sharing”):

- boxed numbers (1, 2, ...) indicate structure sharing
- within feature structures:

  $\begin{bmatrix}
  \text{ATTR}_1 & 1 \\
  \text{ATTR}_2 & 1 
  \end{bmatrix}
  \begin{bmatrix}
  \text{ATTR}_1 & \text{val}_1 \\
  \text{ATTR}_2 & 1 
  \end{bmatrix}
  \begin{bmatrix}
  \text{ATTR}_1 & 1 \\
  \text{ATTR}_2 & 1 
  \end{bmatrix}$

Re-entrancies produce cyclic structures!
Correspondence between feature structures and graphs:

\[
\begin{bmatrix}
\text{type1} \\
\text{ATTR}_1 \\
\text{type2} \\
\text{attr} \\
\end{bmatrix}
\begin{bmatrix}
\text{attr} \\
\text{attr} \\
\text{attr} \\
\text{attr} \\
\end{bmatrix}
\]
Types and type hierarchy

- **Types**: fixed bundles of attributes

  ![Types Diagram]

  - `index`
  - `synsem`
  - `PER 1`
  - `LOC 1`
  - `NUM 2`
  - `NONLOC 2`
  - `GEND 3`

  \( \rightsquigarrow \) specified in the **signature**

- **Type hierarchy**: order on types
  \( \approx \) attribute inheritance

  ![Type Hierarchy Diagram]

  - `word`
  - `phrase`

  \( \rightsquigarrow \) specified in the **signature**
The signature $\Sigma$

is a quadruple (in P&S, 94):

$$< \text{types} , \text{type hierarchy} , \text{attributes} , ((\text{types} \times \text{attributes}) \times \text{types}) >$$

The signature licenses possible feature structures.
(language independent)

A constraint $\tau \in \theta$

is based on $\Sigma$, but also contains variables, conjunction, disjunction, implication, negation, ...

Other name: Principle.

The constraints further restrict the set of well-formed feature structures.
(language independent and dependent)
A grammar $G$ is a signature $\Sigma$ and a set of constraints $\theta$.

Each $G$-licensed feature structure must be:

- **type-resolved**: each type is maximally specific.
- **totally well-typed**: exactly the legal attributes of a type must be present.
Phenomena, Model, Formal Theory

from Müller (2007)

Phenomenon: Ling. Objects

Model: Feature Structures

Formal Theory: Feature Descriptions

models

predicts

determines

licensed by the theory
The architecture of linguistic feature structures

- **PHON**: The surface of a sign.
- **SYNSEM**: The core of a sign. Contains syntactic and semantic information.
- **LOCAL**: Contains information local to the linguistic sign, such as agreement information
- **NONLOCAL**: Contains non-local references, e.g. for long-distance dependencies

**sign**

**PHON**

**phon**

**synsem**

**local**

**CATEGORY**  **category**

**CONTENT**  **content**

**CONTEXT**  **context**

**NONLOCAL**  **nonlocal**
An example 

```
word

PHON  ⟨she⟩

synsem

local

CATEGORY
HEAD
SUBCAT
⟨⟩

ppro

context

CONTENT
INDEX

CONTEXT

NONLOCAL  ...
CATEGORY: Contains *syntactic* information

\[
\begin{array}{c}
cat \\
\text{HEAD} \\
\text{SUBCAT}
\end{array}
\begin{array}{c}
\langle\rangle \\
\text{noun} \\
\text{CASE nom}
\end{array}
\]

- **HEAD**: Syntactic features (roughly part-of-speech)
- **SUBCAT**: Subcategorization information
CONTENT: Contains *semantic* information

\[
\begin{bmatrix}
 ppro \\
 INDEX & 1 \\
\end{bmatrix}
\begin{bmatrix}
 ref \\
 PER & 3rd \\
 NUM & sing \\
 GEND & fem \\
\end{bmatrix}
\]

CONTEXT: Contains *pragmatic* information

\[
\begin{bmatrix}
 context \\
 BACKGR \\
\end{bmatrix}
\begin{bmatrix}
 psoa \\
 RELN & female \\
 INST & 1 \\
\end{bmatrix}
\]
Abbreviations of feature structures (1)

\[
\begin{align*}
\text{synsem} & \quad \begin{pmatrix}
\text{CAT} & \text{HEAD} & \begin{pmatrix}
\text{noun} \\
\text{CASE} & \text{case}
\end{pmatrix} \\
\text{SUBCAT} & \langle \rangle \\
\text{CONT} & | \text{INDEX} & \underbrace{\text{i}}_\text{per}, \text{num}, \text{gend} \\
\text{GEND} & \text{per} \\
\text{PER} & \text{num}
\end{pmatrix} \\
\end{align*}
\]

\[\mapsto\]

\[
\text{NP}_i[\text{per, num, gend}][\text{case}]
\]

\[
\begin{align*}
\text{synsem} & \quad \begin{pmatrix}
\text{CAT} & \text{HEAD} & \begin{pmatrix}
\text{verb} \\
\text{vform} \quad \text{vform}
\end{pmatrix} \\
\text{SUBCAT} & \langle \text{synsem} \rangle \\
\text{CONT} & \underbrace{\text{i}}_\text{vform}
\end{pmatrix} \\
\end{align*}
\]

\[\mapsto\]

\[
\text{VP}[\text{vform}]:_i
\]
Abbreviations of feature structures (2)

\[
\begin{align*}
\text{HEAD} & \quad \left[ \begin{array}{c} \text{noun} \\ \text{CASE} \\ \text{nom} \end{array} \right] \quad \leadsto \quad \left[ \begin{array}{c} \text{HEAD} \\ \text{Noun} [\text{nom}] \end{array} \right] \\
\text{CONT} & \quad \left[ \begin{array}{c} \text{INDEX} \\ \text{REL} \\ \text{arg1} \end{array} \right] \quad \leadsto \quad \left[ \begin{array}{c} \text{INDEX} [\text{i}] \\ \text{CONTENT} \left[ \text{rel}(\text{1}) \right] \end{array} \right]
\end{align*}
\]

\[
\begin{align*}
\text{HEAD} & \quad \left[ \begin{array}{c} \text{verb} \\ \text{VFORM} \\ \text{fin} \end{array} \right] \quad \leadsto \quad \left[ \begin{array}{c} \text{HEAD} \\ \text{Verb} [\text{fin}] \end{array} \right] \\
\text{CONT} & \quad \left[ \begin{array}{c} \text{rel} \\ \text{ARG1} \end{array} \right] \quad \leadsto \quad \left[ \begin{array}{c} \text{CONTENT} \left[ \text{rel}(\text{1}, \ldots, \text{n}) \right] \end{array} \right]
\end{align*}
\]
Abbreviations in action

\[
\begin{align*}
\text{word} & \quad \text{PHON} \langle \textit{walks} \rangle \\
\text{HEAD} & \quad \text{verb} \\
\text{VFORM} & \quad \text{fin} \\
\text{SYNSEM} & \quad \text{LOC} \quad \text{CAT} \\
\text{SUBCAT} & \quad \text{LOC} \\
\text{CAT} & \quad \text{CONT} \\
\text{INDEX} & \quad \text{PER} \\
\text{NUM} & \quad \text{3} \\
\text{word} & \quad \text{PHON} \langle \textit{walks} \rangle \\
\text{SYNSEM} & \quad \text{LOC} \quad \text{CAT} \\
\text{SUBCAT} & \quad \langle \text{NP} \rangle \\
\end{align*}
\]
Structures of type phrase have the feature DTRS.

Tree notation:

```
C
NP[nom]

H
[HEAD Verb[fin]
  SUBCAT ⟨⟩]

[HEAD Verb[fin]
  SUBCAT ⟨NP[nom]⟩]
```

```
[HEAD Verb[fin]
  SUBCAT ⟨⟩]

Kim

walks
```
Principles

The Word Principle

\[ \text{word} \Rightarrow LE_1 \lor \ldots \lor LE_n \]

The Head Feature Principle

\[
\begin{align*}
\text{SYNSEM} &| \text{LOC} | \text{CAT} | \text{HEAD} & | 1 \\
\text{DTRS} &| \text{HEAD-DTR} | \text{SYNSEM} | \text{LOC} | \text{CAT} | \text{HEAD} & | 1
\end{align*}
\]

The Subcategorization Principle (binary version)

\[
\begin{align*}
\text{SYNSEM} &| \text{LOC} | \text{CAT} | \text{SUBCAT} & | 2 \\
\text{DTRS} & \left[ \text{HEAD-DTR} | \text{SYNSEM} | \text{LOC} | \text{CAT} | \text{SUBCAT} & | 2 | 1 \right] \\
\text{COMP-DTR} &| \text{SYNSEM} & | 1
\end{align*}
\]
The Immediate Dominance Schemata (1)

The ID Principle

Every headed phrase must satisfy exactly one of the ID schemata.

≈ phrase structure rules

For example (from P&S, 94):

- Head-Subject Schema
- Head-Complement Schema
- Head-Subject-Complement Schema
- Head-Marker Schema
- Head-Adjunct Schema
- Head-Filler Schema
The Immediate Dominance Schemata (1)

⇒ including Head Feature Principle and Subcategorization Principle!

**Head-Subject Schema:**

```
[ HEAD       1 ]
[ SUBCAT     ]
  \   /  \
 C /   \ H
  \   /  \
  [ phrase  2 ]
  [ HEAD     ]
  [ SUBCAT   ]
  [ 1 ]
```

**Head-Complement Schema:**

```
[ HEAD     1 ]
[ SUBCAT   ]
  \   /  \
 H /   \ C
  \   /  \
  \   /  \
  [ HEAD   ]
  [ SUBCAT ]
  [ 2 ]
  \   /  \
  [ C     ]
  \   /  \
  [ C     ]
  \   /  \
  [ C     ]
  \   /  \
  [ ...   ]
  \   /  \
  [ C     ]
  \   /  \
  [ HPSG  ]
```

HPSG 21
Raising and control (1) - A semantic distinction

(1) a. John seems to leave.
   b. John tries to leave.

"seems":

```
word
[HEAD Verb[fin]
  CAT [SUBCAT [1[3sg][nom], VP[inf, SUBCAT <1>]:3]]
  CONT seem(3)]
```

"tries":

```
word
[HEAD Verb[fin]
  CAT [SUBCAT [NP[1[3sg][nom], VP[inf, SUBCAT <NP 1>]:3]]
  CONT try(1[ref], 3)]
```

⇒ no PRO and no ECM!
⇒ by and large: control and raising only differ semantically!
“to leave”:

\[
\begin{array}{c}
\text{CAT} \\
\text{CAT} \\
\text{CONT}
\end{array}
\begin{array}{c}
\text{word} \\
\text{HEAD} \\
\text{head}
\end{array}
\begin{array}{c}
\text{Verb[inf]} \\
\langle \text{NP} \rangle
\end{array}
\text{leave(1)}
\]

“John”:

\[
\begin{array}{c}
\text{CAT} \\
\text{CAT} \\
\text{CONT}
\end{array}
\begin{array}{c}
\text{word} \\
\text{HEAD} \\
\text{index}
\end{array}
\begin{array}{c}
\text{Noun[nom]} \\
\langle \rangle
\end{array}
\begin{array}{c}
\text{PER} \\
\text{NUM} \\
\text{GEND}
\end{array}
\begin{array}{c}
3 \\
\text{sg} \\
\text{masc}
\end{array}
\text{john(1)}
\]
(2) John seems to leave.

Raising and control (3) - The derivation
(3) Kim we know Sandy claims Dana hates __.

Trace (simplified version):

\[
\begin{bmatrix}
\text{PHON} & \langle \rangle \\
\text{SYNSEM} & \\
\text{LOCAL} & 1 \\
\text{NONLOCAL} & \\
\text{SLASH} & \{1\}
\end{bmatrix}
\]

- SLASH collects LOCAL-values of traces.
- LOCAL-values of traces and their fillers are structure shared.
Extraction - Principle and ID Schema

Nonlocal Feature Principle (for head-comp-struc)

$$\text{DTRS} \left\{ \begin{array}{c}
\text{SYNSEM} \mid \text{NONLOC} \mid \text{SLASH} \cup [2 \ldots 7] \\
\text{HEAD-DTR} \mid \text{SYNSEM} \mid \text{NONLOC} \mid \text{SLASH} \\
\text{COMP-DTR} \\
\ldots, \\
\text{SYNSEM} \mid \text{NONLOC} \mid \text{SLASH} \end{array} \right\}$$

Head-Filler Schema ($\approx$ P&S,94):

$$\begin{array}{c}
\text{F} \\
\text{SYNSEM} \mid \text{LOC} \{1\} \\
\text{H} \\
\text{LOC} \mid \text{CAT} \mid \text{HEAD} \\
\text{NONLOC} \mid \text{SLASH} \{2\{1,...\}\} \\
\text{VERB}[\text{fin}] \\
\text{SUBCAT} \langle \rangle \\
\end{array}$$
(4) Kim we know Sandy claims Dana hates __.
1. Finite sentences with complementizer (that, whether)

(5) *Who did the elephant whisper that the emu saw __; ?
   Who did the elephant say that the emu saw __; ?

- **Non-Bridge Verbs (whisper):** specify their sentential complement having an empty SLASH-set:

\[
\begin{array}{c}
\text{LOC | HEAD | SUBCAT} \\
\text{$\ldots, [\text{NONLOC | SLASH } \{\}, \ldots] $}
\end{array}
\]

- **Bridge Verbs (say):** don’t specify the sentential complement’s SLASH-set.
2.) Subjects from finite sentences with complementizer

(6) *Who; did Alice say that __; left.
    Who; did Alice say __; left.

Trace Principle:

SYNSEM values of traces must be a noninitial member of a (substantive head’s) SUBCAT list.

⇝ implementation within the Head-Subject Schema:

```
[ HEAD
  [ SUBCAT
    [ C
      [ NONLOC | SLASH
        { }
      ]
    ]
  ]

[ phrase
  [ HEAD
    [ SUBCAT
      [ H
        [ 1 ]
      ]
    ]
  ]
```
2.) **Subjects from finite sentences with complementizer**

(7) *Who; did Alice say that ___; left.  
    Who; did Alice say ___; left.

**Subject Extraction Lexical Rule (SELR):**

\[
\begin{align*}
\text{SYNSEM} & \mid \text{LOC} \mid \text{CAT} \mid \text{HEAD} \mid \text{SUBCAT} \arrow{\Downarrow} \\
\text{LOC} & \mid \text{CAT} \mid \text{HEAD} \mid \text{SUBCAT} \arrow{\Downarrow} \text{VP} \\
\text{NONLOC} & \mid \text{SLASH} \{ 1 \}
\end{align*}
\]
3.) Adjuncts

(8) *[Which movie]i; did Gorgette fall asleep after watching __i;.

⇝ Head-Adjunct Schema:
Lexical items are encoded as AVMs, such as the one we have already seen:

```
word
PHON ⟨she⟩
  synsem
    local
cat
  CATEGORY
  HEAD
  SUBCAT ⟨⟩
  ppro
  ref
  INDEX 1
  PER 3rd
  NUM sing
  GEND fem
  context
  BACKGR
  psoa
  RELN female
  INST 1
```

The lexicon
“Vertical” generalizations: Make use of underspecification possible through clever design of the type hierarchy
“Horizontal” generalizations: Lexical rules which “automatically” license lexical entries based on the present ones.

Example: Passivization (Kiss, 1992):

```
stem
CAT
  SUBCAT  [HEAD verb [NP[NOM], NP[ACC]_{1} \oplus 2]]

word
CAT
  SUBCAT  [HEAD [VFORM passiv-part]
             [NP[NOM]_{1} \oplus 2]]
```
For a linguistic theory following HPSG, we need at least the following:

- **Signature**
  - Type hierarchy
  - Feature geometry

- **Principles**,
  - general *principles*, restrictions on linguistic objects
  - *ID schemata*, licensing the combination of lexical items

- **Lexicon**:
  - *Lexical entries* (AVMs)
  - *Lexical rules* (AVMs), licensing more lexical items
Please see the reference list on the course web page at
http://www.sfb441.uni-tuebingen.de/emmy/gf/references.pdf