Head-Driven Phrase Structure Grammar
An Introduction as Background for Grammar Implementation

Detmar Meurers
OSU, LING795K, Spring 2002

Why implement an HPSG theory?

Implementing grammars can be very valuable in terms of

a) providing feedback for a rigid and complete formalization of a linguistic theory.

b) stimulating system development to enhance the link between theory and implementation and to improve performance.

For this to work,

- all differences between a linguistic theory and its implementation need to be documented, and
- the system should support a clear, tractable, and formally meaningful way of implementing close to the linguistic theory.

  Recoding a linguistic theory in terms of some unrelated or lower-level computer language makes it difficult to obtain meaningful feedback from the implementation for linguistics.

The building blocks of HPSG grammars

1. From a linguistic perspective
2. From a formal perspective

HPSG grammars from a linguistic perspective

From a linguistic perspective, an HPSG grammar consists of

a) a lexicon
licensing basic words

b) lexical rules
licensing derived words

c) immediate dominance (ID) schemata
licensing constituent structure

d) linear precedence (LP) statements
constraining word order

e) a set of grammatical principles
expressing generalizations about linguistic objects

HPSG grammars from a formal perspective

From a formal perspective, an HPSG grammar consists of

- **the signature** as declaration of the domain, and
- **the theory** constraining the domain.

The signature

- defines the ontology ('declaration of what exists'):
  - which kind of objects are distinguished, and
  - which properties of which objects are modelled.
- consists of
  - the type (or sort) hierarchy and
  - the appropriateness conditions, defining which type has which appropriate attributes (or features) with which appropriate values.

Models of linguistic objects

- The objects are modelled by feature structures, which are depicted as directed graphs.
- Since these models represent objects in the world (and not knowledge about the world), they are total with respect to the ontology declared in the signature. Technically, one says that these feature structures are
  - *totally well-typed*: Every node has all the attributes appropriate for its type and each attributes has an appropriate value.
  - *sort-resolved*: Every node is of a maximally specific type.

Descriptions

A **description language** and its abbreviating **AVM notation** is used to talk about sets of objects. Descriptions consists of three building blocks:

- **Type** decriptions single out all objects of a particular type, e.g., *word*
- **Attribute-value pairs** describe objects that have a particular property. The attribute must be appropriate for the particular type of object, and the value can be any kind of description, e.g., $\left[\textsc{spouse}\ \left[\textsc{name}\ \textit{mary}\right]\right]$
- **Tags** (structure sharing) to specify **token identity**, e.g. $\boxed{1}$

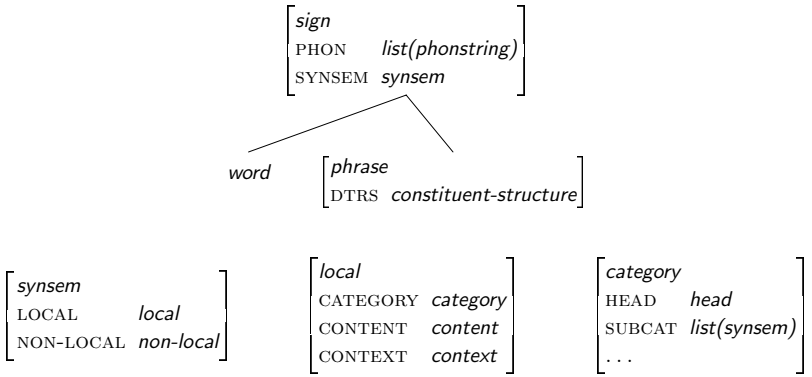
Descriptions (cont.)

Complex descriptions are obtained by combining descriptions with the help of conjunction ($\wedge$), disjunction ($\vee$) and negation ($\neg$). In the AVM notation, conjunction is implicit.

A **theory** (in the formal sense) is a set of description language statements, often referred to as the constraints.

- The theory singles out a subset of the objects declared in the signature, namely those which are grammatical.
- A linguistic object is admissible with respect to a theory iff it satisfies each of the descriptions in the theory and so does each of its substructures.

An ontology of linguistic objects



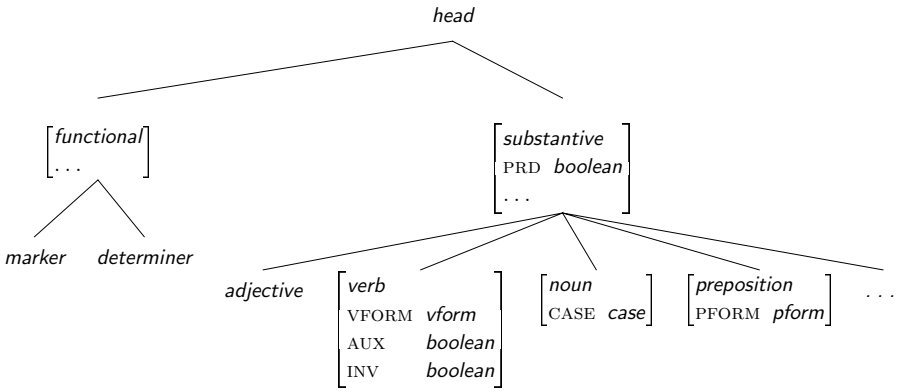
Motivating SUBCAT

(1) a. I *laugh*. (<NP>)
b. I *saw* him. (<NP NP>)
c. I *give* her the book. (<NP NP NP>)
d. I *said* that she left. (<NP S[that]>)

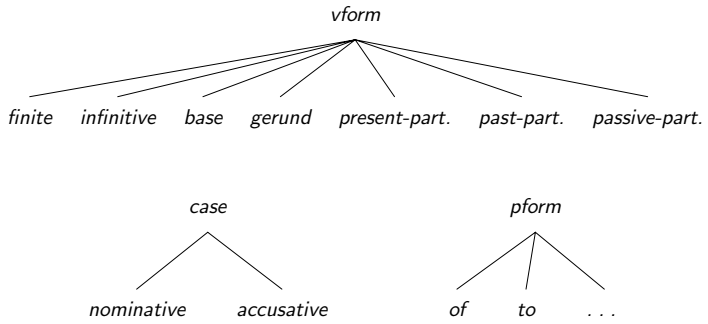
Cannot always be derived from semantics:

(2) a. Paul ate a steak. (<NP>)
b. Paul ate. (<NP NP>)
(3) a. Paul devoured a steak. (<NP>)
b. * Paul devoured (<NP NP>)

Part-of-speech



Properties of particular part-of-speech



Motivating VFORM

(4) a. Peter will *win* the race. (base form)
b. * Peter will *won* the race.
c. * Peter will *to win* the race.

(5) a. Peter has *won* the race. (past participle)
b. * Peter has *win* the race.
c. Peter has *to win* the race.
(→ different verb)

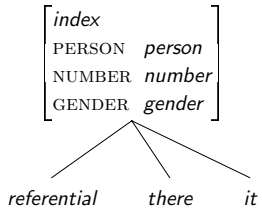
(6) a. Peter seems *to win* the race. (*to*-infinitive)
b. * Peter seems *win* the race.
c. * Peter seems *won* the race.

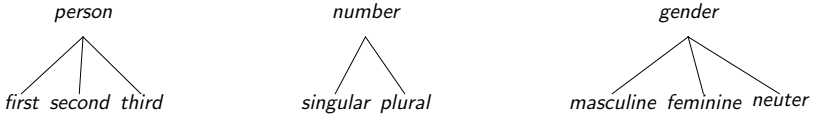
Motivating CASE

(7) a. *He* left. (nom)
b. * *Him* left.

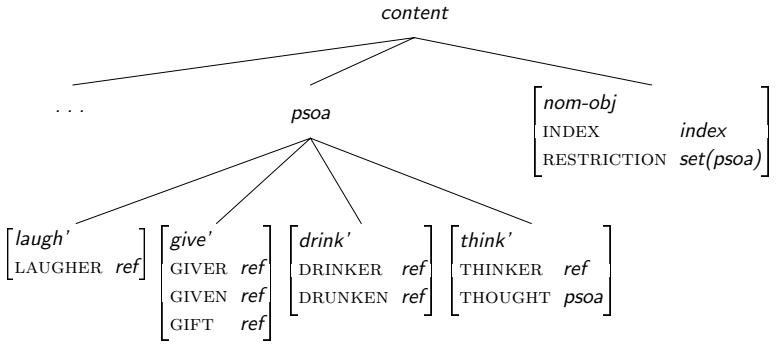
(8) a. She sees *him*. (acc)
b. * She sees *he*.

Indices

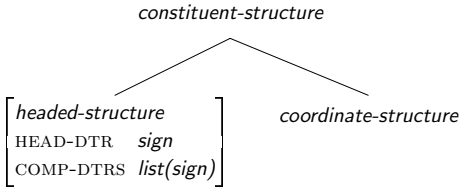




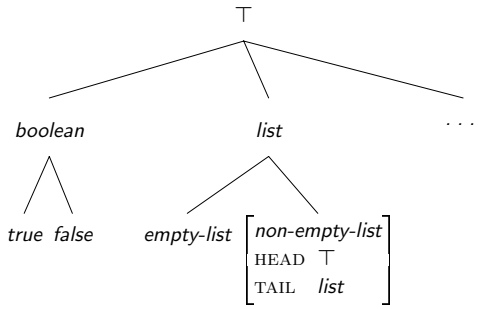
Semantic representations



Constituent structure



Auxiliary data structures



Abbreviations for describing lists

empty-list	is abbreviated as	*e-list*, <>
non-empty-list	is abbreviated as	*ne-list*
$\begin{bmatrix}\text{HEAD}\ \boxed{1}\\ \text{TAIL}\ \boxed{2}\end{bmatrix}$	is abbreviated as	$\langle \boxed{1} \mid \boxed{2} \rangle$
$\langle \ldots \boxed{1} \mid \langle\rangle \rangle$	is abbreviated as	$\langle \ldots \boxed{1} \rangle$
$\begin{bmatrix}\text{HEAD}\ \boxed{1}\\ \text{TAIL}\ \begin{bmatrix}\text{HEAD}\ \boxed{2}\\ \text{TAIL}\ \boxed{3}\end{bmatrix}\end{bmatrix}$	is abbreviated as	$\langle \boxed{1}, \boxed{2} \mid \boxed{3} \rangle$

Attention: $\langle \top \rangle$ and $\langle \boxed{1} \rangle$ describe all lists of length **one**!

Abbreviations of common AVMs

Pollard and Sag (1994) use some abbreviations to describe *synsem* objects:

Abbreviation	Abbreviated AVM
NP$_{\boxed{1}}$	$\begin{bmatrix} \textit{synsem} \\ \text{LOCAL} \begin{bmatrix} \text{CATEGORY} \begin{bmatrix} \text{HEAD} & \textit{noun} \\ \text{SUBCAT} & \langle\rangle \end{bmatrix} \\ \text{CONTENT}\vert\text{INDEX} \ \boxed{1} \end{bmatrix} \end{bmatrix}$
S:$\boxed{1}$	$\begin{bmatrix} \textit{synsem} \\ \text{LOCAL} \begin{bmatrix} \text{CATEGORY} \begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{SUBCAT} & \langle\rangle \end{bmatrix} \\ \text{CONTENT} \ \boxed{1} \end{bmatrix} \end{bmatrix}$
VP:$\boxed{1}$	$\begin{bmatrix} \textit{synsem} \\ \text{LOCAL} \begin{bmatrix} \text{CATEGORY} \begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{SUBCAT} & \left\langle \textit{synsem} \right\rangle \end{bmatrix} \\ \text{CONTENT} \ \boxed{1} \end{bmatrix} \end{bmatrix}$

The Lexicon

The basic lexicon is defined by the *Word Principle* as part of the theory. It defines which of the ontologically possible words are grammatical:

$$\textit{word} \rightarrow \text{lexical-entry}_1 \vee \text{lexical-entry}_2 \vee \ldots$$

with each of the lexical entries being descriptions, such as e.g.:

$$\begin{bmatrix} \textit{word} \\ \text{PHON} & \textit{<laughs>} \\ \text{SYNSEM}\vert\text{LOC} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} \\ \text{VFORM} & \textit{fin} \end{bmatrix} \\ \text{SUBCAT} & \left\langle \text{NP}[\textit{nom}]_{\boxed{1}}[\textit{3rd,sing}] \right\rangle \end{bmatrix} \\ \text{CONTENT} \begin{bmatrix} \textit{laugh'} \\ \text{LAUGHER} \ \boxed{1} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

An example lexicon

$$\textit{word} \rightarrow \begin{bmatrix} \text{PHON} & \textit{<gives>} \\ \text{S}\vert\text{L} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} \\ \text{VFORM} & \textit{fin} \end{bmatrix} \\ \text{SUBCAT} & \left\langle \text{NP}[\textit{nom}]_{\boxed{1}}[\textit{sing}], \ \text{NP}[\textit{acc}]_{\boxed{2}}, \ \text{PP}[\textit{to}]_{\boxed{3}} \right\rangle \end{bmatrix} \\ \text{CONT} \begin{bmatrix} \textit{give'} \\ \text{GIVER} \ \boxed{1} \\ \text{GIFT} \ \boxed{2} \\ \text{GIVEN} \ \boxed{3} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

$$\bigvee \begin{bmatrix} \text{PHON} & \textit{<drinks>} \\ \text{S}\vert\text{L} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} \\ \text{VFORM} & \textit{fin} \end{bmatrix} \\ \text{SUBCAT} & \left\langle \text{NP}[\textit{nom}]_{\boxed{1}}[\textit{3rd,sing}], \ \text{NP}[\textit{acc}]_{\boxed{2}} \right\rangle \end{bmatrix} \\ \text{CONT} \begin{bmatrix} \textit{drink'} \\ \text{DRINKER} \ \boxed{1} \\ \text{DRUNKEN} \ \boxed{2} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

$$\bigvee \begin{bmatrix} \text{PHON} & \textit{<drink>} \\ \text{S}\vert\text{L} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} \\ \text{VFORM} & \textit{fin} \end{bmatrix} \\ \text{SUBCAT} & \left\langle \text{NP}[\textit{nom}]_{\boxed{1}}[\textit{plur}], \ \text{NP}[\textit{acc}]_{\boxed{2}} \right\rangle \end{bmatrix} \\ \text{CONT} \begin{bmatrix} \textit{drink'} \\ \text{DRINKER} \ \boxed{1} \\ \text{DRUNKEN} \ \boxed{2} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

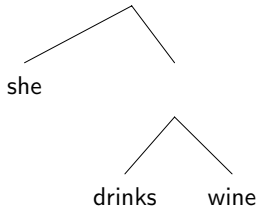
$$\bigvee \begin{bmatrix} \text{PHON} & \langle \textit{she} \rangle \\ \text{SYNSEM|LOC} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{noun} \\ \text{CASE} \ \textit{nom} \end{bmatrix} \\ \text{SUBCAT} & \langle\rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{INDEX} & \begin{bmatrix} \text{PER} & \textit{third} \\ \text{NUM} & \textit{sing} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

$$\bigvee \begin{bmatrix} \text{PHON} & \langle \textit{wine} \rangle \\ \text{SYNSEM|LOC} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \textit{noun} \\ \text{SUBCAT} & \langle\rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{INDEX} & \begin{bmatrix} \text{PER} & \textit{third} \\ \text{NUM} & \textit{sing} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

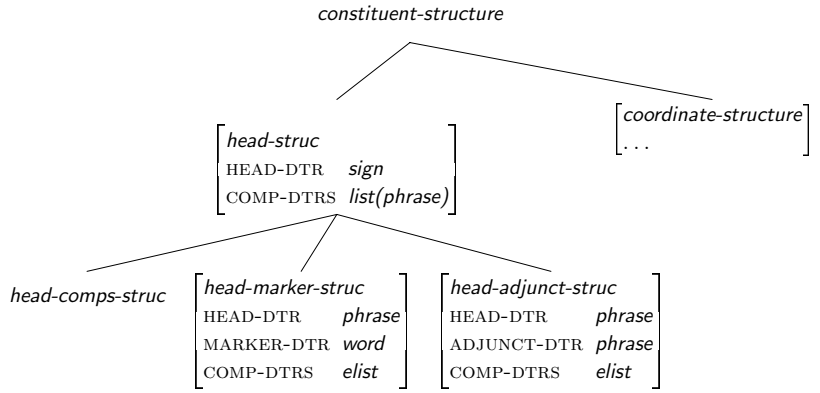
$$\bigvee \begin{bmatrix} \text{PHON} & \langle \textit{to} \rangle \\ \text{S|L} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{preposition} \\ \text{PFORM} \ \textit{to} \end{bmatrix} \\ \text{SUBCAT} & \left\langle \text{NP}[\textit{acc}]\boxed{1} \right\rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{INDEX} & \boxed{1} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

$$\bigvee \begin{bmatrix} \text{PHON} & \langle \textit{think} \rangle \\ \text{S|L} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} \\ \text{VFORM} \ \textit{fin} \end{bmatrix} \\ \text{SUBCAT} & \left\langle \text{NP}[\textit{nom}]_{\boxed{1}}[\textit{plur}], \ \text{S}[\textit{fin}]{:}\boxed{2} \right\rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \textit{think'} \\ \text{THINKER} \ \boxed{1} \\ \text{THOUGHT} \ \boxed{2} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

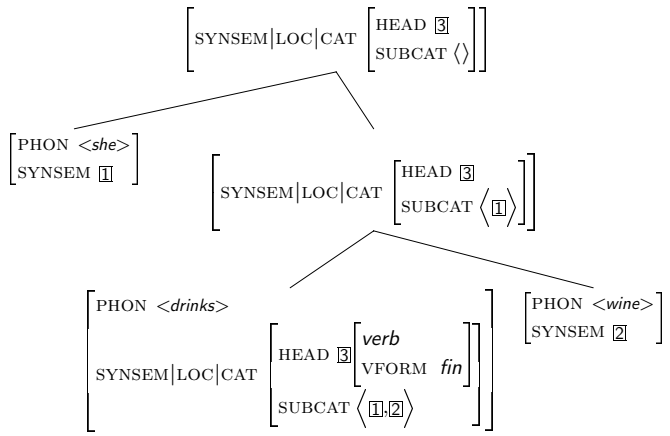
A very first sketch of an example



An ontology of phrases



Sketch of an example for head-complement structures



Head-Feature Principle:

$$\begin{bmatrix} \textit{phrase} \\ \text{DTRS}\ \ \textit{headed-structure} \end{bmatrix} \rightarrow \begin{bmatrix} \text{SYNSEM|LOC|CAT|HEAD} & \boxed{1} \\ \text{DTRS|HEAD-DTR|SYNSEM|LOC|CAT|HEAD} & \boxed{1} \end{bmatrix}$$

Subcat Principle:

$$\begin{bmatrix} \text{DTRS}\ \ \textit{headed-structure} \end{bmatrix} \rightarrow \begin{bmatrix} \text{SYNSEM|LOC|CAT|SUBCAT}\ \boxed{1} \\ \text{DTRS} \begin{bmatrix} \text{HEAD-DTR|SYNSEM|LOC|CAT|SUBCAT}\ \boxed{1} \oplus \boxed{2} \\ \text{COMP-DTRS}\ \ \textit{synsem2sign}\Big(\boxed{2}\Big) \end{bmatrix} \end{bmatrix}$$

with $\oplus$ standing for list concatenation, i.e., *append*, defined as follows

$$\begin{array}{lcl} \textit{e-list} \oplus \boxed{1} & := & \boxed{1}. \\ \begin{bmatrix} \text{FIRST}\ \boxed{1} \\ \text{REST}\ \ \boxed{2} \end{bmatrix} \oplus \boxed{3} & := & \begin{bmatrix} \text{FIRST}\ \boxed{1} \\ \text{REST}\ \ \boxed{2} \oplus \boxed{3} \end{bmatrix}. \end{array}$$

An auxiliary relation: synsem2sign/2

The call to *synsem2sign/2* is needed to relate the *synsem* objects on the SUBCAT to *sign* objects containing those *synsem* objects as value of their SYNSEM attribute. It is defined as follows:

$$\begin{array}{lcl} \textit{synsem2sign}\Big(\textit{e-list}\Big) & := & \textit{e-list}. \\ \textit{synsem2sign}\left(\begin{bmatrix} \text{FIRST}\ \boxed{1} \\ \text{REST}\ \ \boxed{2} \end{bmatrix}\right) & := & \begin{bmatrix} \text{FIRST}\ \begin{bmatrix} \text{SYNSEM}\ \boxed{1} \end{bmatrix} \\ \text{REST}\ \ \textit{synsem2sign}\Big(\boxed{2}\Big) \end{bmatrix}. \end{array}$$

Immediate Dominance Principle (for English):

$$\begin{bmatrix} \textit{phrase} \\ \text{DTRS}\ \ \textit{headed-struc} \end{bmatrix} \rightarrow \begin{bmatrix} \text{SYNSEM|LOC|CAT} & \begin{bmatrix} \text{HEAD} & \left(\begin{bmatrix} \textit{verb} \\ \text{INV}\ - \end{bmatrix} \vee \neg\ \textit{verb} \right) \\ \text{SUBCAT}\ \langle\rangle \end{bmatrix} \\ \text{DTRS} & \begin{bmatrix} \textit{head-comp-struc} \\ \text{HEAD-DTR}\ \ \textit{phrase} \\ \text{COMP-DTRS}\ \langle \textit{sign} \rangle \end{bmatrix} \end{bmatrix}$$ (Head-Subject)

$$\vee \begin{bmatrix} \text{SYNSEM|LOC|CAT} & \begin{bmatrix} \text{HEAD} & \left(\begin{bmatrix} \textit{verb} \\ \text{INV}\ - \end{bmatrix} \vee \neg\ \textit{verb} \right) \\ \text{SUBCAT}\ \langle \textit{synsem} \rangle \end{bmatrix} \\ \text{DTRS} & \begin{bmatrix} \textit{head-comps-struc} \\ \text{HEAD-DTR}\ \ \textit{word} \end{bmatrix} \end{bmatrix}$$ (Head-Complement)

$$\vee \begin{bmatrix} \text{SYNSEM|LOC|CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} \\ \text{INV}\ + \end{bmatrix} \\ \text{SUBCAT}\ \langle\rangle \end{bmatrix} \\ \text{DTRS} & \begin{bmatrix} \textit{head-comps-struc} \\ \text{HEAD-DTR}\ \ \textit{word} \end{bmatrix} \end{bmatrix}$$ (Head-Subject-Complement)

$\vee$ . . . continued on next page

Immediate Dominance Principle (for English):

$$\begin{bmatrix} \textit{phrase} \\ \text{DTRS}\ \ \textit{headed-struc} \end{bmatrix} \rightarrow \begin{array}{l} \vdots \\ \vdots \\ \vdots \end{array}$$

$$\vee \begin{bmatrix} \text{DTRS} & \begin{bmatrix} \textit{head-marker-struc} \\ \text{MARKER-DTR|SYNSEM|LOC|CAT|HEAD}\ \ \textit{marker} \end{bmatrix} \end{bmatrix}$$ (Head-Marker)

$$\vee \begin{bmatrix} \text{DTRS} & \begin{bmatrix} \textit{head-adjunct-struc} \\ \text{ADJ-DTR|SYNSEM|LOC|CAT|HEAD|MOD}\ \boxed{1} \\ \text{HEAD-DTR|SYNSEM}\ \ \boxed{1} \end{bmatrix} \end{bmatrix}$$ (Head-Adjunct)

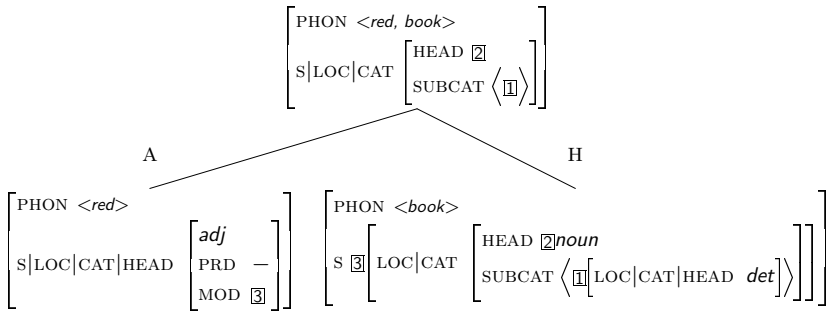
Lexical entry of an attributive adjective

$$\begin{bmatrix} \textit{word} \\ \text{PHON}\ \ \textit{<red>} \\ \text{SYNSEM} \begin{bmatrix} \textit{synsem} \\ \text{LOCAL} \begin{bmatrix} \textit{local} \\ \text{CATEGORY} \begin{bmatrix} \textit{cat} \\ \text{HEAD} \begin{bmatrix} \textit{adj} \\ \text{PRD}\ - \\ \text{MOD} \begin{bmatrix} \textit{synsem} \\ \text{LOCAL} \begin{bmatrix} \textit{local} \\ \text{CATEGORY} \begin{bmatrix} \textit{cat} \\ \text{HEAD}\ \ \textit{noun} \\ \text{SUBCAT} \left\langle \begin{bmatrix} \textit{synsem} \\ \text{LOCAL} \begin{bmatrix} \textit{local} \\ \text{CATEGORY} \begin{bmatrix} \textit{cat} \\ \text{HEAD}\ \ \textit{det} \end{bmatrix} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix} \\ \text{CONTENT} \begin{bmatrix} \textit{nom-obj} \\ \text{INDEX}\ \boxed{1}\textit{index} \\ \text{RESTR}\ \boxed{2}\textit{list} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{CONTENT} \begin{bmatrix} \textit{nom-obj} \\ \text{INDEX}\ \boxed{1} \\ \text{RESTR} \left\langle \begin{bmatrix} \textit{red-rel} \\ \text{ARG1}\ \boxed{1} \end{bmatrix} \mid \boxed{2} \right\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

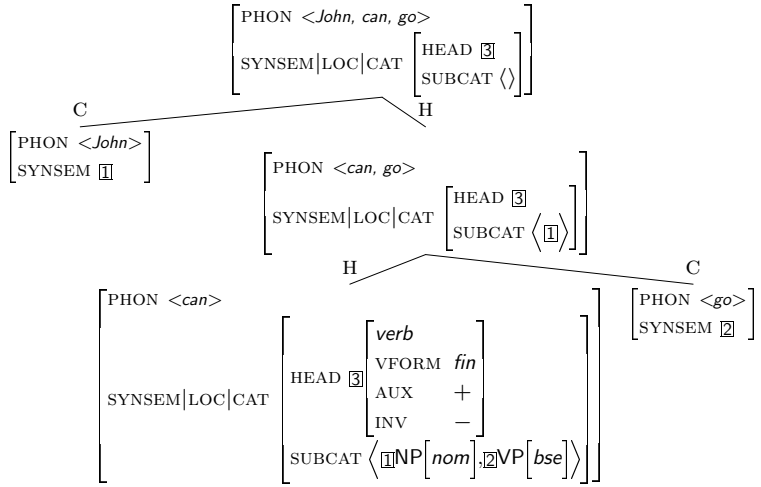
Lexical entry of an attributive adjective
Version without redundant specifications

$$\begin{bmatrix} \textit{word} \\ \text{PHON}\ \ \textit{<red>} \\ \text{SYNSEM|LOC} \begin{bmatrix} \text{CAT|HEAD} & \begin{bmatrix} \textit{adj} \\ \text{PRD}\ - \\ \text{MOD|LOC} \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD}\ \ \textit{noun} \\ \text{SUBCAT} \left\langle \begin{bmatrix} \text{LOC|CAT|HEAD}\ \ \textit{det} \end{bmatrix} \right\rangle \end{bmatrix} \\ \text{CONTENT} & \begin{bmatrix} \text{INDEX}\ \boxed{1}\textit{index} \\ \text{RESTR}\ \boxed{2}\textit{list} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{INDEX}\ \boxed{1} \\ \text{RESTR} \left\langle \begin{bmatrix} \textit{red-rel} \\ \text{ARG1}\ \boxed{1} \end{bmatrix} \mid \boxed{2} \right\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

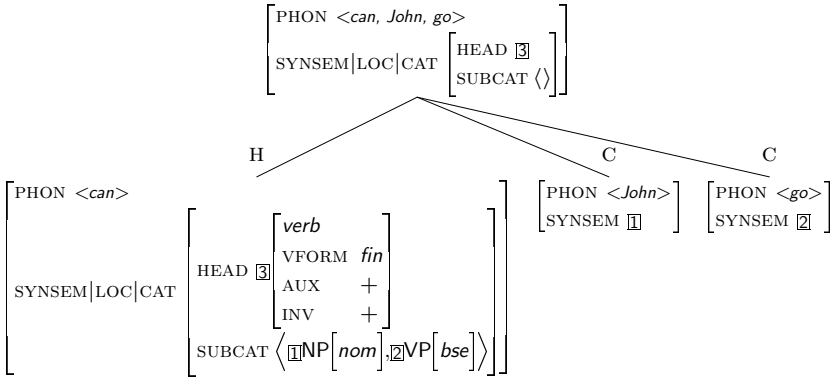
Sketch of an example for a head-adjunct structure



Sketch of an example with an auxiliary



Sketch of an example with an inverted auxiliary



SPEC Principle:

$$\left[\begin{array}{l}\textit{phrase}\\ \text{DTRS}\ \left[\left(\text{MARKER-DTR} \vee \text{COMP-DTRS}|\text{FIRST}\right)|\text{SYNSEM}|\text{LOC}|\text{CAT}|\text{HEAD}\ \textit{functional}\right]\end{array}\right]$$

$$\rightarrow \left[\begin{array}{l}\text{DTRS}\ \left[\left(\text{MARKER-DTR} \vee \text{COMP-DTRS}|\text{FIRST}\right)|\text{SYNSEM}|\text{LOC}|\text{CAT}|\text{HEAD}|\text{SPEC}\ \boxed{1}\right]\\ \text{HEAD-DTR}|\text{SYNSEM}\ \boxed{1}\end{array}\right]$$

Marking Principle:

$$\begin{bmatrix} \textit{phrase} \\ \text{DTRS}\ \ \textit{headed-structure} \end{bmatrix} \rightarrow \begin{bmatrix} \text{SYNSEM|LOC|CAT|MARKING}\ \boxed{1} \\ \text{DTRS}\ \begin{bmatrix} \textit{head-mark-struc} \\ \text{MARKER-DTR|SYNSEM|LOC|CAT|MARKING}\ \boxed{1} \end{bmatrix} \end{bmatrix} \vee \begin{bmatrix} \text{SYNSEM|LOC|CAT|MARKING}\ \boxed{1} \\ \text{DTRS}\ \begin{bmatrix} \neg\textit{head-mark-struc} \\ \text{HEAD-DTR|SYNSEM|LOC|CAT|MARKING}\ \boxed{1} \end{bmatrix} \end{bmatrix}$$

Lexical entry of the marker *that*

$$\begin{bmatrix} \text{PHON} & \textit{<that>} \\ \text{SYNSEM|LOC|CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{mark} \\ \text{SPEC} & \begin{bmatrix} \text{LOC|CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} \\ \text{VFORM}\ \ \textit{fin} \vee \textit{bse} \end{bmatrix} \\ \text{MARKING} & \textit{unmarked} \\ \text{SUBCAT} & \langle\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{SUBCAT} & \langle\rangle \\ \text{MARKING} & \textit{that} \end{bmatrix} \end{bmatrix}$$

Sketch of an example for a head-marker structure