

An Introduction to Grammar Engineering Using HPSG

Frederik Foyry

Computational Linguistics, Saarland University

foyyry@coli.uni-sb.de

Stephan Open

CSLI Stanford

oesli@csli.stanford.edu

Melanie Siegel

DFKI GmbH

siegel@dfki.de

Outline: What We Are About to Do (and Why)

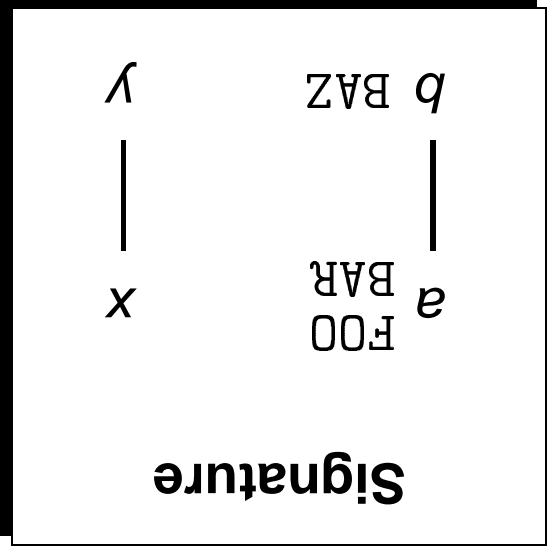
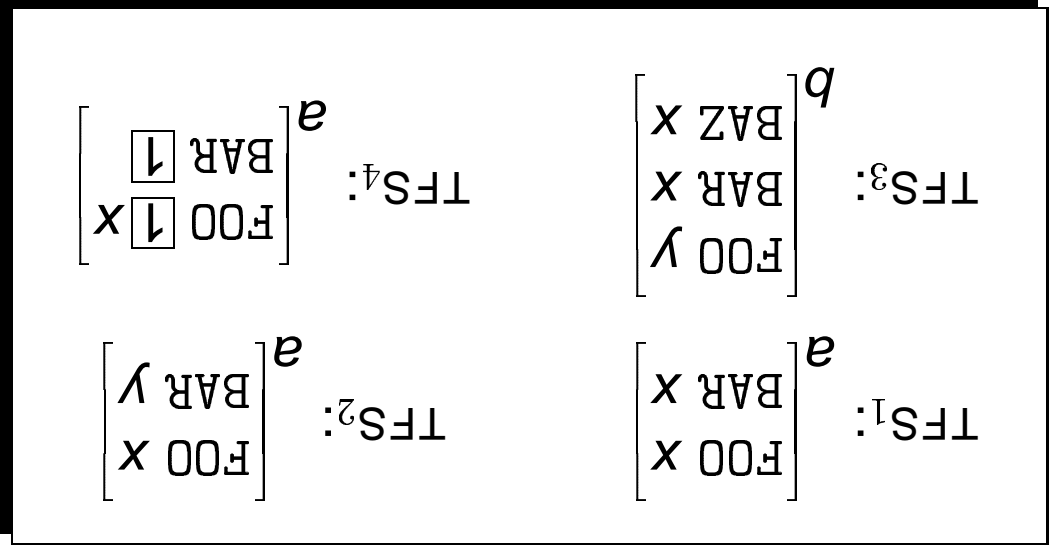
Course Outline

- familiarize with computational grammar development environment;
- learn how to formalize grammars in typed feature structures;
- adapt and develop sequence of trivial HPSG grammars in LKB;
- solve daily exercises: immediate gratification (risk of late hours).

Why Computational Grammars

- **research** formalize linguistic theories with complex interactions of language phenomena; identify cross-language generalizations;
- **education** teach frameworks or analyses in formal morphology, syntax, and semantics; support student experimentation;
- **applications** embed grammar-based natural language analysis in research prototypes and commercial applications.

Feature Structure Unification: The Logics



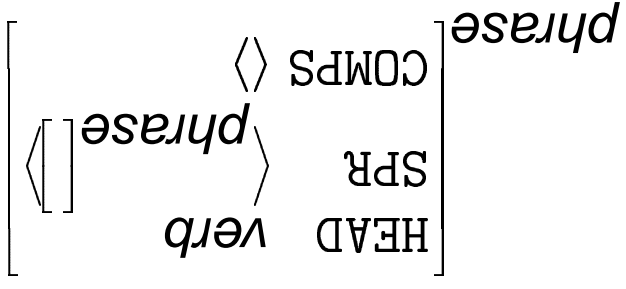
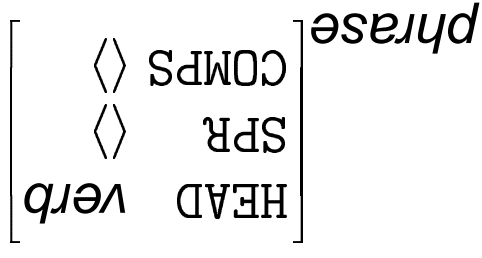
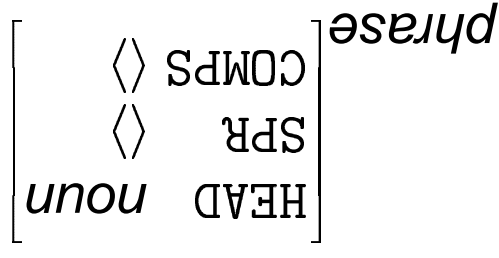
$$TFS_1 \sqcap TFS_2 \equiv TFS_2 \quad TFS_1 \sqcap TFS_3 \equiv TFS_3 \quad TFS_3 \sqcap TFS_4 \equiv TFS_4$$

$$\begin{bmatrix} \text{FOO } \downarrow y \\ \text{BAR } \downarrow \\ \text{BAZ } x \end{bmatrix}^b$$

UNIFICATION (' \sqcap ') ensures compatibility and combines all information

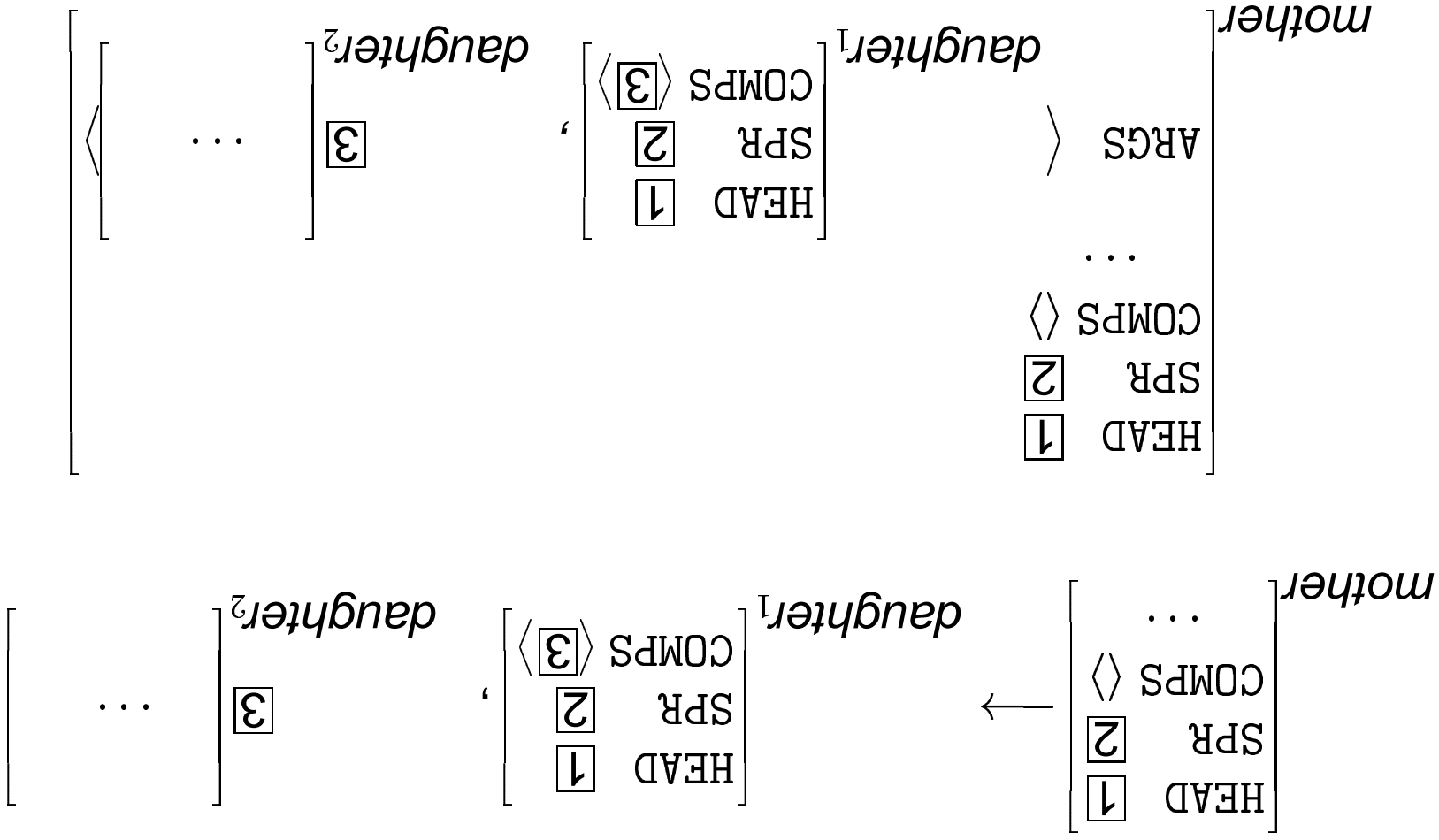
Structured Categories in a Unification Grammar

- All (constituent) categories in the grammar are typed feature structures;
- specific TFS configurations may correspond to 'traditional' categories;
- labels like 'S' or 'NP' are mere abbreviations, not elements of the theory.



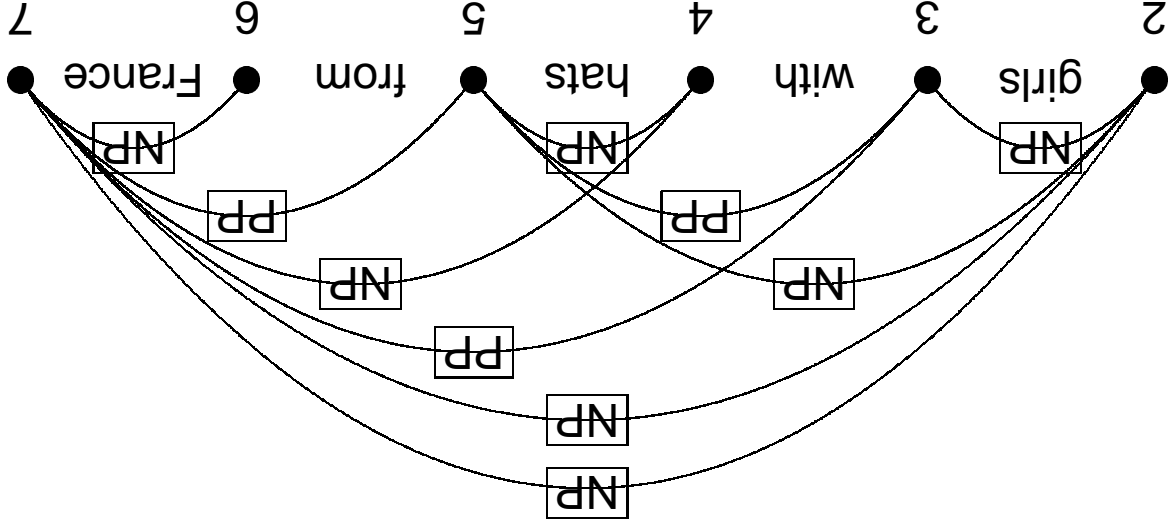
'NP', 'S', 'VP'

The Format of Grammar Rules in the LKB



Bottom-Up Chart Parsing in the LKB

- Initialize chart: retrieve all lexical entries for all words in the input string;
- Parsing: apply all rules to all adjacent tuples of edges in the chart;
- Add new edge for each successful instantiation of a grammar rule.



Literature and Further Pointers

On-line User Manual for the LKB

- at COLI, use the command 'lkbdoc' from the xterm command line;
- anywhere, go to `http://www-csl1.stanford.edu/~aac/lkb.html`.

Course Materials: Exercises and Solutions

`http://www.coli.uni-sb.de/~fourvy/teaching/`

Background Reading

- Sag, Ivan A. and Thomas Wasow (1999). *Syntactic Theory. A Formal Introduction*. Stanford, CA: CSLI Publications.
- Ann A. Copestake (2002). *Implementing Typed Feature Structure Grammars*. Stanford, CA: CSLI Publications.
- LINGO project at CSLI Stanford: `http://lingo.stanford.edu/`.