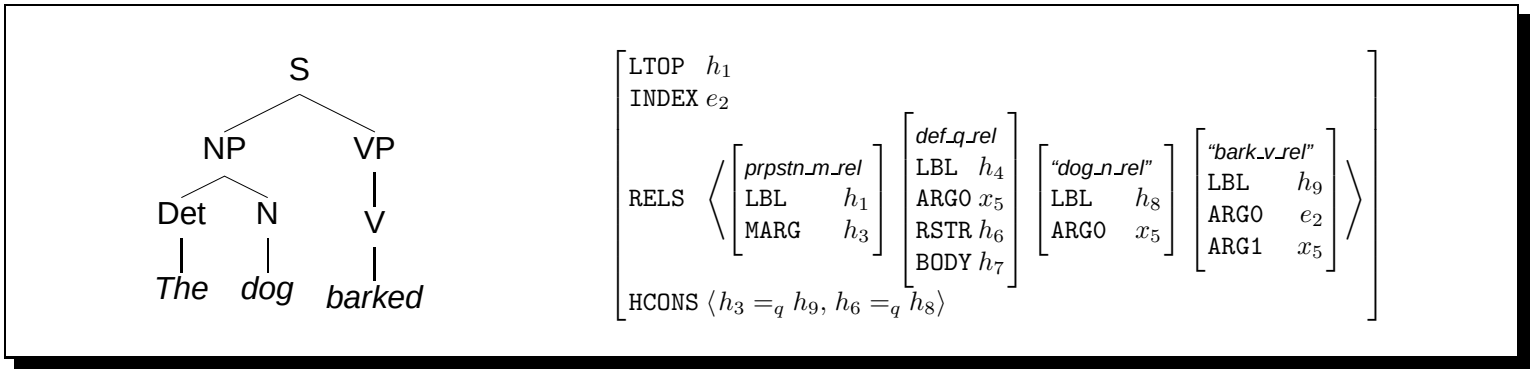


Computational Linguistics (INF2820 — Lexical Rules)

$$\textit{phrase}\begin{bmatrix}\texttt{HEAD} & \boxed{1} \\ \texttt{SPR} & \langle\rangle \\ \texttt{COMPS} & \boxed{3}\end{bmatrix} \longrightarrow \boxed{2}\textit{phrase}\begin{bmatrix}\texttt{SPR} & \langle\rangle \\ \texttt{COMPS} & \langle\rangle\end{bmatrix}, \textit{phrase}\begin{bmatrix}\texttt{HEAD} & \boxed{1} \\ \texttt{SPR} & \langle\boxed{2}\rangle \\ \texttt{COMPS} & \boxed{3}\end{bmatrix}$$

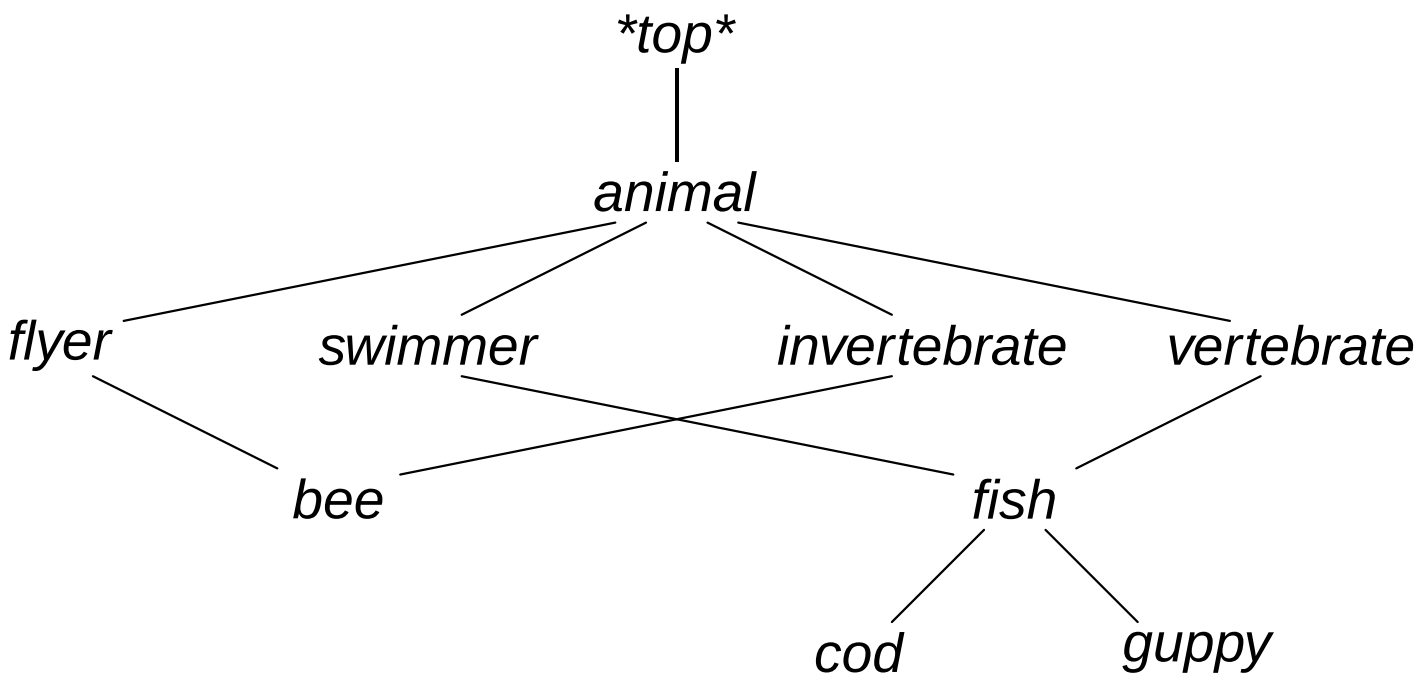
Stephan Oepen

Universitetet i Oslo

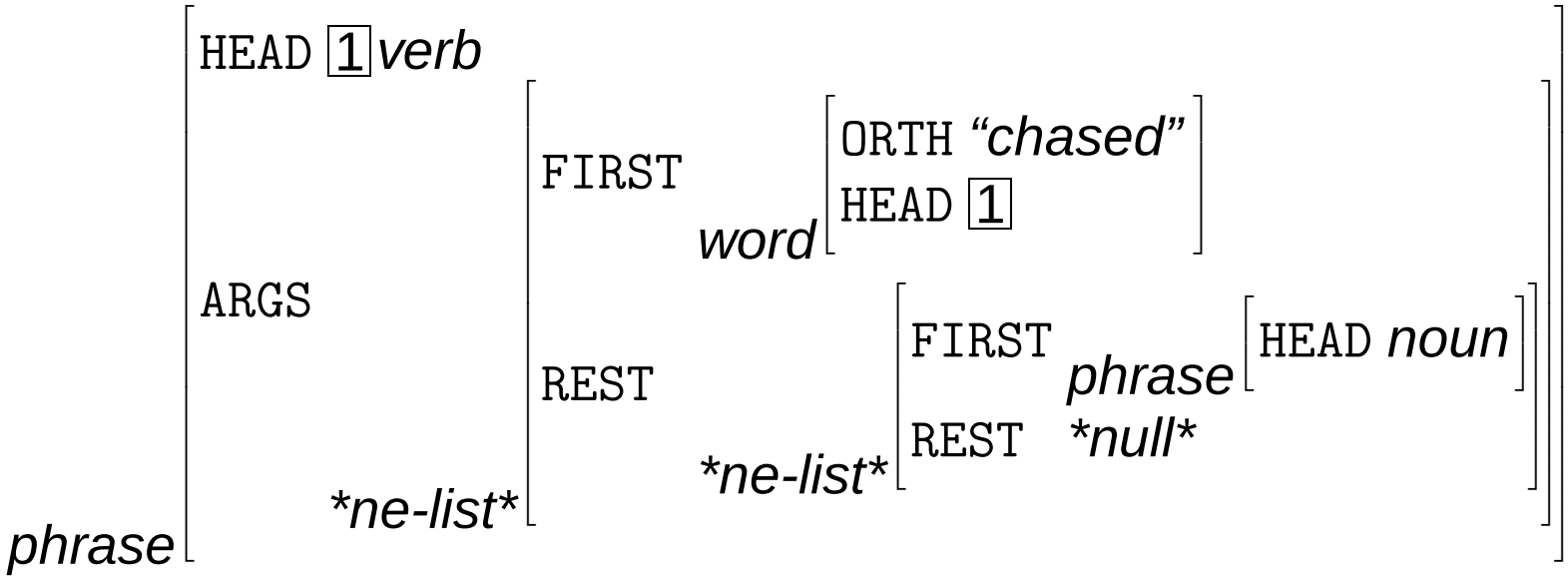
oe@ifi.uio.no

Multiple Inheritance

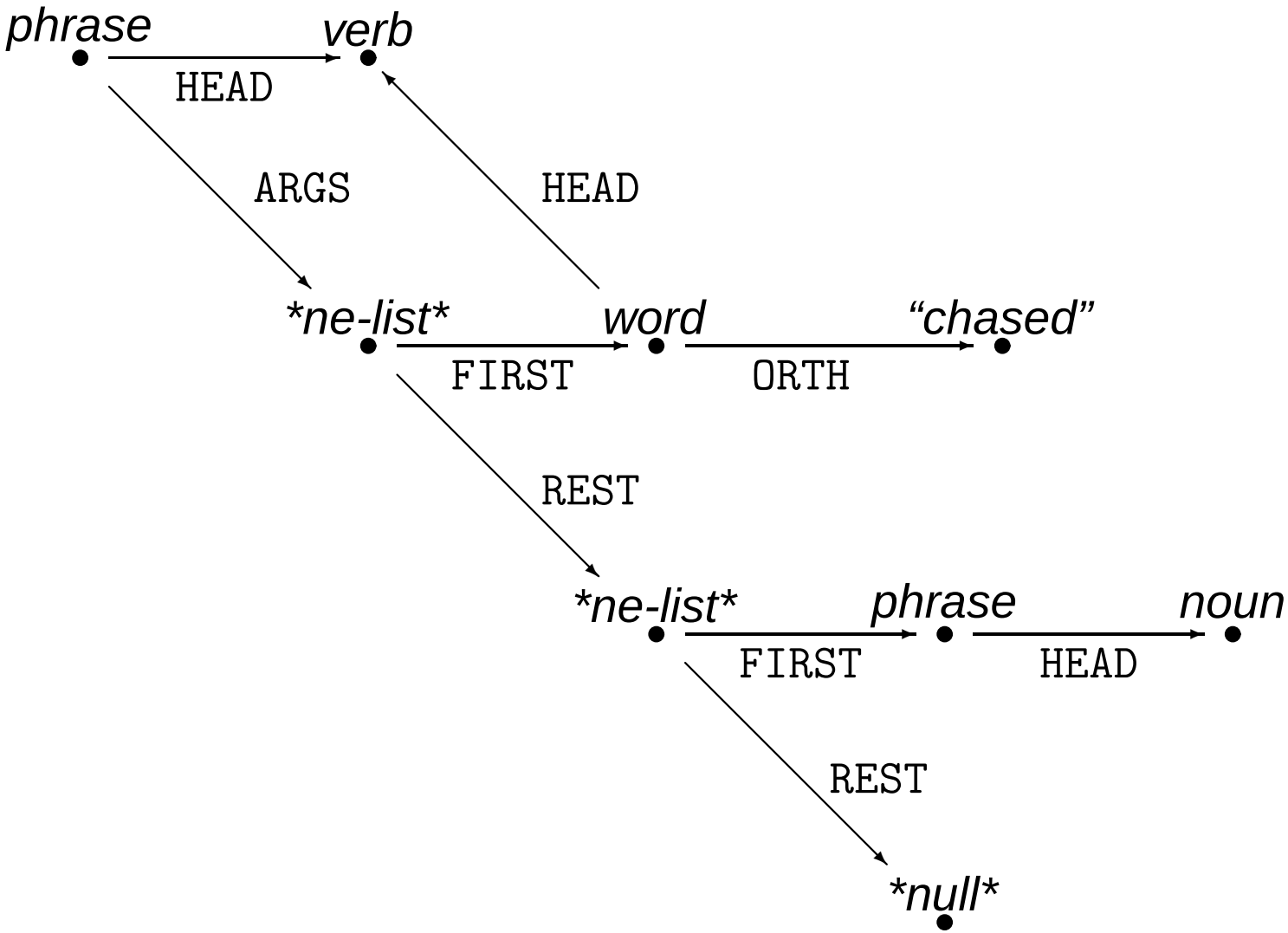
- *flyer* and *swimmer* no common descendants: they are incompatible;
- *flyer* and *bee* stand in hierarchical relationship: they unify to subtype;
- *flyer* and *invertebrate* have a unique greatest common descendant.



Reentrancy in a Typed Feature Structure (AVM)



Reentrancy in a Typed Feature Structure (Graph)



An Ambiguous Example

Kim *shoveled* *snow* *on* *lifts.*

A Highly Ambiguous Example

The manager placed his bid on my desk.

Our Grammars: Table of Contents

Type Description Language (TDL)

- `types.tdl` type definitions: hierarchy of grammatical knowledge;
- `lexicon.tdl` instances of (lexical) types plus orthography;
- `rules.tdl` instances of construction types; used by the parser;
- `lrules.tdl` lexical rules, applied before non-lexical rules;
- `irules.tdl` lexical rules that require orthographemic variation;
- `roots.tdl` grammar start symbol(s): 'selection' of final results.

Auxiliary Files (Grammar Configuration for LKB)

- `labels.tdl` TFS templates abbreviating node labels in trees;
- `globals.lsp`, `user-fns.lsp` parameters and interface functions;
- `mrsglobals.lsp` MRS parameters (path to semantics et al.)

The Format of Grammar Rules in the LKB

$$\textit{phrase}\begin{bmatrix}\texttt{HEAD} & \boxed{1} \\ \texttt{SPR} & \langle\rangle \\ \texttt{COMPS} & \boxed{3}\end{bmatrix} \longrightarrow \boxed{2}\textit{phrase}\begin{bmatrix}\texttt{SPR} & \langle\rangle \\ \texttt{COMPS} & \langle\rangle\end{bmatrix}, \textit{phrase}\begin{bmatrix}\texttt{HEAD} & \boxed{1} \\ \texttt{SPR} & \langle\boxed{2}\rangle \\ \texttt{COMPS} & \boxed{3}\end{bmatrix}$$

$$\textit{phrase}\begin{bmatrix}\texttt{HEAD} & \boxed{1} \\ \texttt{SPR} & \langle\rangle \\ \texttt{COMPS} & \boxed{3} \\ \texttt{ARGS} & \left\langle \boxed{2}\textit{phrase}\begin{bmatrix}\texttt{SPR} & \langle\rangle \\ \texttt{COMPS} & \langle\rangle\end{bmatrix}, \textit{phrase}\begin{bmatrix}\texttt{HEAD} & \boxed{1} \\ \texttt{SPR} & \langle\boxed{2}\rangle \\ \texttt{COMPS} & \boxed{3}\end{bmatrix}\right\rangle\end{bmatrix}$$

Dative Shift: A Productive Process

$\{hand_1, give_1, send_1, ...\}$

$$\begin{bmatrix} \texttt{HEAD} & verb \\ \texttt{SPR} & \langle \cdots \rangle \\ \texttt{COMPS} & \left\langle \begin{bmatrix} \texttt{HEAD} & noun \\ \texttt{SPR} & \langle\rangle \\ \texttt{COMPS} & \langle\rangle \end{bmatrix}, \begin{bmatrix} phrase \\ \texttt{HEAD} & noun \\ \texttt{SPR} & \langle\rangle \\ \texttt{COMPS} & \langle\rangle \end{bmatrix} \right\rangle \end{bmatrix}$$

$\{hand_2, give_2, send_2, ...\}$

$$\begin{bmatrix} \texttt{HEAD} & verb \\ \texttt{SPR} & \langle \cdots \rangle \\ \texttt{COMPS} & \left\langle \begin{bmatrix} \texttt{HEAD} & noun \\ \texttt{SPR} & \langle\rangle \\ \texttt{COMPS} & \langle\rangle \end{bmatrix}, \begin{bmatrix} phrase \\ \texttt{HEAD} & prep\begin{bmatrix}\texttt{PFORM}\ to\end{bmatrix} \\ \texttt{SPR} & \langle\rangle \\ \texttt{COMPS} & \langle\rangle \end{bmatrix} \right\rangle \end{bmatrix}$$

Lexical Variation: Lexical Rules

- Dative shift, passivization, et al. are systematic processes in the lexicon;
- use of *monotonic* inheritance is insufficient to relate $give_1$ and $give_2$;
- *lexical rules* are unary grammar rules that operate 'within the lexicon';
- take as input a lexical sign (*expression*) and output a derived lexical sign.

Rough Approximation of Passive Lexical Rule

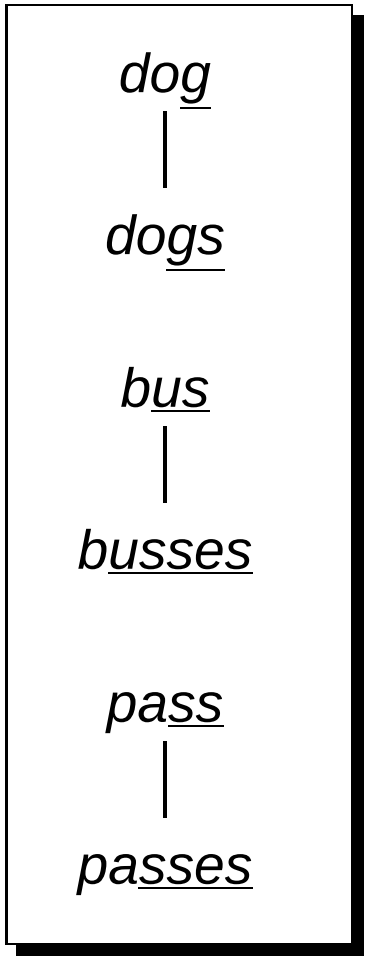
$$\left[\begin{array}{ll} \texttt{HEAD} & \boxed{1} \\ \texttt{SPR} & \langle \boxed{2} \rangle \\ \texttt{COMPS} & \boxed{3} \\ \texttt{ARGS} & \left\langle \left[\begin{array}{ll} \texttt{HEAD} & \boxed{1} \\ \texttt{SPR} & \left\langle \textit{phrase}\left[\texttt{HEAD}\ \textit{noun}\right] \right\rangle \\ \texttt{COMPS} & \left[\begin{array}{ll} \texttt{FIRST} & \boxed{2}\ \textit{phrase}\left[\texttt{HEAD}\ \textit{noun}\right] \\ \texttt{REST} & \boxed{3} \end{array}\right] \end{array}\right] \right\rangle \end{array}\right]$$

Orthographemic Variation: Inflectional Rules

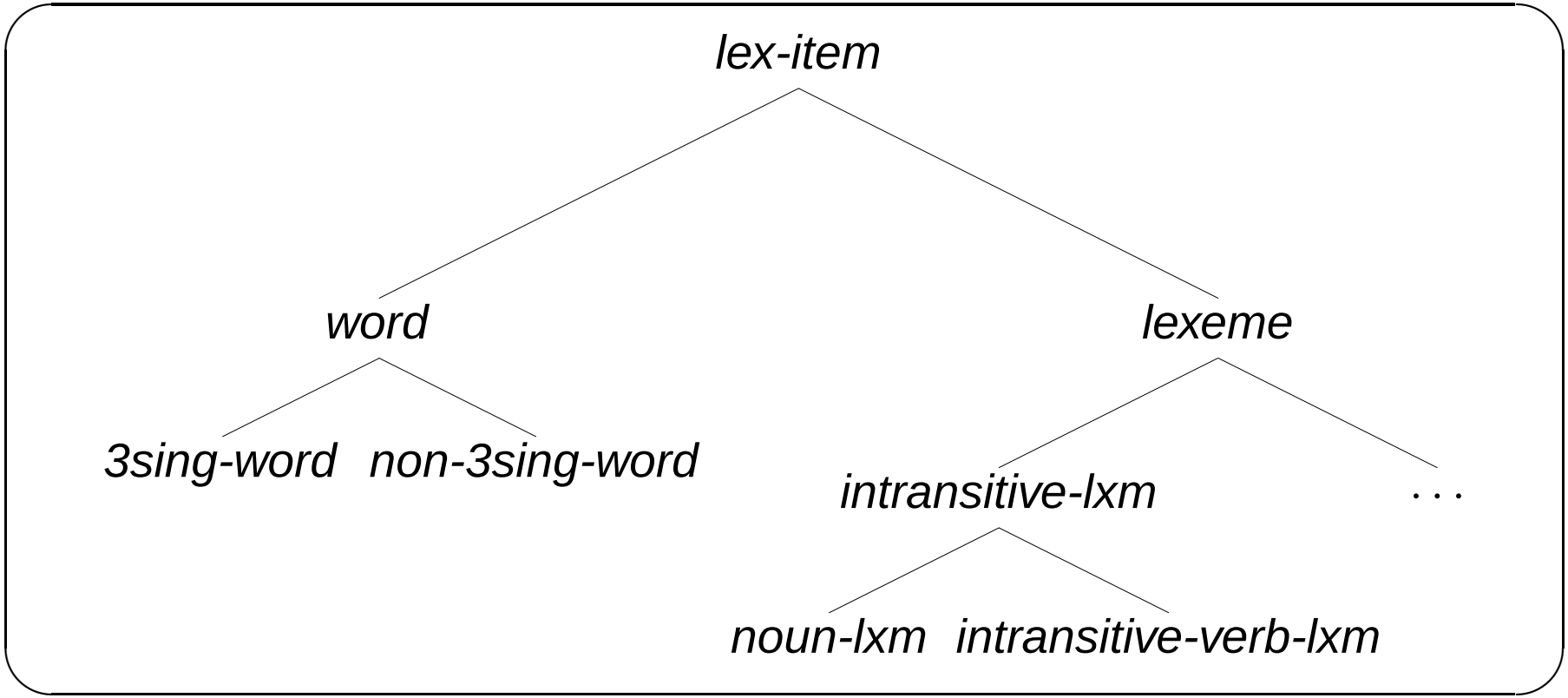
```
%(letter-set (!s abcdefghijklmnopqrtuvwxyz))

noun-non-3sing_irule :=
%suffix (!s !ss) (!ss !ssses) (ss sses)
non-3sing-word &
[ HEAD [ AGR non-3sing ],
  ARGS < noun-lxm > ].

noun-3sing_irule :=
3sing-word &
[ ORTH #1,
  ARGS < noun-lxm & [ ORTH #1 ] > ].
```



The Lexeme vs. Word Distinction



- Lexical entries are *uninflected*; cannot enter syntax by themselves;
- inflectional rules 'make' *word* from *lexeme*, possibly with 'null' suffix.