Ling 566
Oct 20, 2020

How the Grammar Works

Midterm feedback: Thank you!

- More examples, more polls, more info on connection to research & industry
- More cats!
- Answer keys
- Ability to ask post-lecture questions
- Examples like the homework problems
- Importance of study groups

Overview

- What we're trying to do
- The pieces of our grammar
- Two extended examples
- Reflection on what we've done, what we still have to do
- Reading questions

What We're Trying To Do

- Objectives
  - Develop a theory of knowledge of language
  - Represent linguistic information explicitly enough to distinguish well-formed from ill-formed expressions
  - Be parsimonious, capturing linguistically significant generalizations.
- Why Formalize?
  - To formulate testable predictions
  - To check for consistency
  - To make it possible to get a computer to do it for us

How We Construct Sentences

- The Components of Our Grammar
  - Grammar rules
  - Lexical entries
  - Principles
  - Type hierarchy (very preliminary, so far)
  - Initial symbol (S, for now)
- We combine constraints from these components.
  - Q: What says we have to combine them?

An Example

A cat slept.

- Can we build this with our tools?
- Given the constraints our grammar puts on well-formed sentences, is this one?

Lexical Entry for *a*

$$\left\langle \text{a} \;,\; \begin{bmatrix} \textit{word} \\ \text{SYN} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{det} & \\ \text{AGR} & \textit{3sing} \\ \text{COUNT} & + \end{bmatrix} \\ \text{VAL} & \begin{bmatrix} \text{COMPS} & \langle\,\rangle \\ \text{SPR} & \langle\,\rangle \\ \text{MOD} & \langle\,\rangle \end{bmatrix} \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \text{MODE} & \text{none} \\ \text{INDEX} & j \\ \text{RESTR} & \left\langle \begin{bmatrix} \text{RELN} & \text{a} \\ \text{BV} & j \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \right\rangle$$

- Is this a fully specified description?
- What features are unspecified?
- How many word structures can this entry license?

Lexical Entry for *cat*

$$\left\langle \text{cat} , \left[\begin{array}{l} \textit{word} \\ \text{SYN} \left[\begin{array}{ll} \text{HEAD} & \left[\begin{array}{ll} \textit{noun} & \\ \text{AGR} & \left[\begin{array}{ll} \textit{3sing} & \\ \text{GEND} & \text{neut} \end{array}\right] \end{array}\right] \\ \text{VAL} & \left[\begin{array}{ll} \text{SPR} & \left\langle \left[\begin{array}{ll} \text{COUNT} & + \\ \text{INDEX} & k \end{array}\right] \right\rangle \\ \text{COMPS} & \langle\,\rangle \\ \text{MOD} & \langle\,\rangle \end{array}\right] \end{array}\right] \\ \text{SEM} \left[\begin{array}{ll} \text{MODE} & \text{ref} \\ \text{INDEX} & k \\ \text{RESTR} & \left\langle \left[\begin{array}{ll} \text{RELN} & \text{cat} \\ \text{INSTANCE} & k \end{array}\right] \right\rangle \end{array}\right] \end{array}\right] \right\rangle$$

- Which feature paths are abbreviated?
- Is this a fully specified description?
- What features are unspecified?
- How many word structures can this entry license?

Effect of Principles: the SHAC

$$\left\langle \text{cat} \;,\; \left[\begin{array}{ll} \textit{word} & \\ \text{SYN} & \left[\begin{array}{ll} \text{HEAD} & \left[\begin{array}{ll} \textit{noun} & \\ \text{AGR} & \boxed{2}\left[\begin{array}{ll} \textit{3sing} & \\ \text{GEND} & \text{neut} \end{array}\right] \end{array}\right] \\ \text{VAL} & \left[\begin{array}{ll} \text{SPR} & \left\langle \left[\begin{array}{ll} \text{AGR} & \boxed{2} \\ \text{COUNT} & + \\ \text{INDEX} & k \end{array}\right] \right\rangle \\ \text{COMPS} & \langle\,\rangle \\ \text{MOD} & \langle\,\rangle \end{array}\right] \end{array}\right] \\ \text{SEM} & \left[\begin{array}{ll} \text{MODE} & \text{ref} \\ \text{INDEX} & k \\ \text{RESTR} & \left\langle \left[\begin{array}{ll} \text{RELN} & \text{cat} \\ \text{INSTANCE} & k \end{array}\right] \right\rangle \end{array}\right] \end{array}\right] \right\rangle$$

Description of Word Structures for *cat*

$$
\begin{bmatrix}
\textit{word} & & \\
\text{SYN} & \begin{bmatrix}
\text{HEAD} & \begin{bmatrix} \textit{noun} & \\ \text{AGR} \ \boxed{2} & \begin{bmatrix} \textit{3sing} \\ \text{GEND} \ \text{neut} \end{bmatrix} \end{bmatrix} \\
\text{VAL} & \begin{bmatrix}
\text{SPR} & \left\langle \begin{bmatrix} \text{AGR} & \boxed{2} \\ \text{COUNT} & + \\ \text{INDEX} & k \end{bmatrix} \right\rangle_{\text{D}} \\
\text{COMPS} & \langle\ \rangle \\
\text{MOD} & \langle\ \rangle
\end{bmatrix}
\end{bmatrix} \\
\text{SEM} & \begin{bmatrix}
\text{MODE} & \text{ref} \\
\text{INDEX} & k \\
\text{RESTR} & \left\langle \begin{bmatrix} \text{RELN} & \text{cat} \\ \text{INSTANCE} & k \end{bmatrix} \right\rangle
\end{bmatrix}
\end{bmatrix}
$$

cat

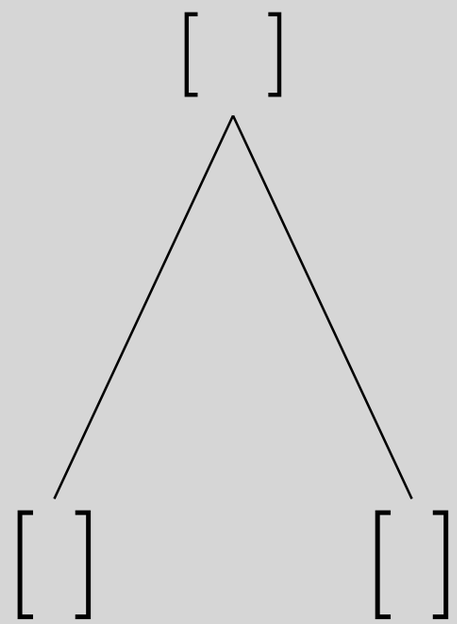
Description of Word Structures for *a*

$$\left[\begin{array}{ll} \textit{word} & \\ \text{SYN} & \left[\begin{array}{ll} \text{HEAD} & \left[\begin{array}{ll} \textit{det} & \\ \text{AGR} & \textit{3sing} \\ \text{COUNT} & + \end{array}\right] \\ \text{VAL} & \left[\begin{array}{ll} \text{COMPS} & \langle\ \rangle \\ \text{SPR} & \langle\ \rangle \\ \text{MOD} & \langle\ \rangle \end{array}\right] \end{array}\right] \\ \text{SEM} & \left[\begin{array}{ll} \text{MODE} & \text{none} \\ \text{INDEX} & j \\ \text{RESTR} & \left\langle \left[\begin{array}{ll} \text{RELN} & \text{a} \\ \text{BV} & j \end{array}\right] \right\rangle \end{array}\right] \end{array}\right]$$

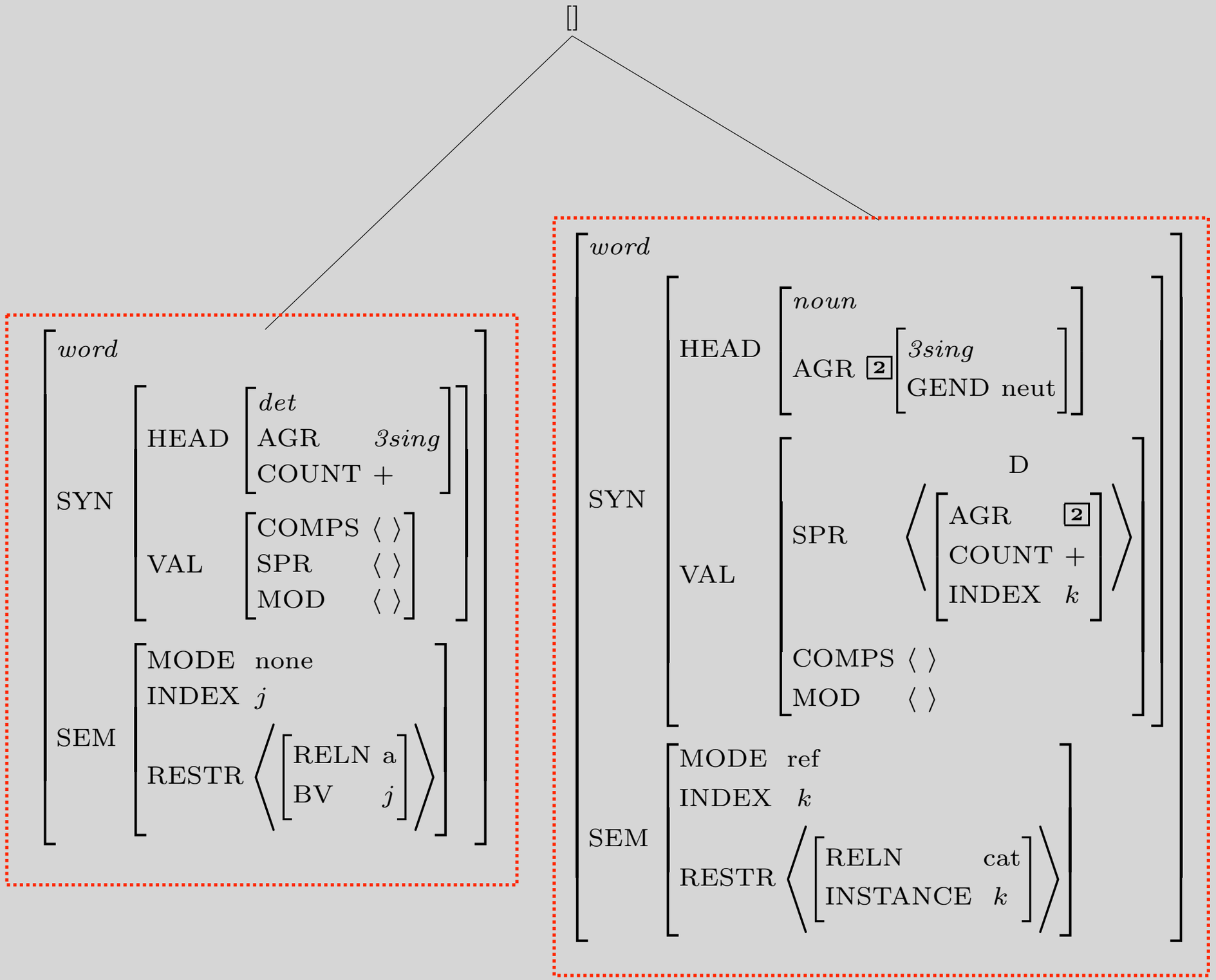
$$|$$

$$a$$

Building a Phrase



Constraints Contributed by Daughter Subtrees



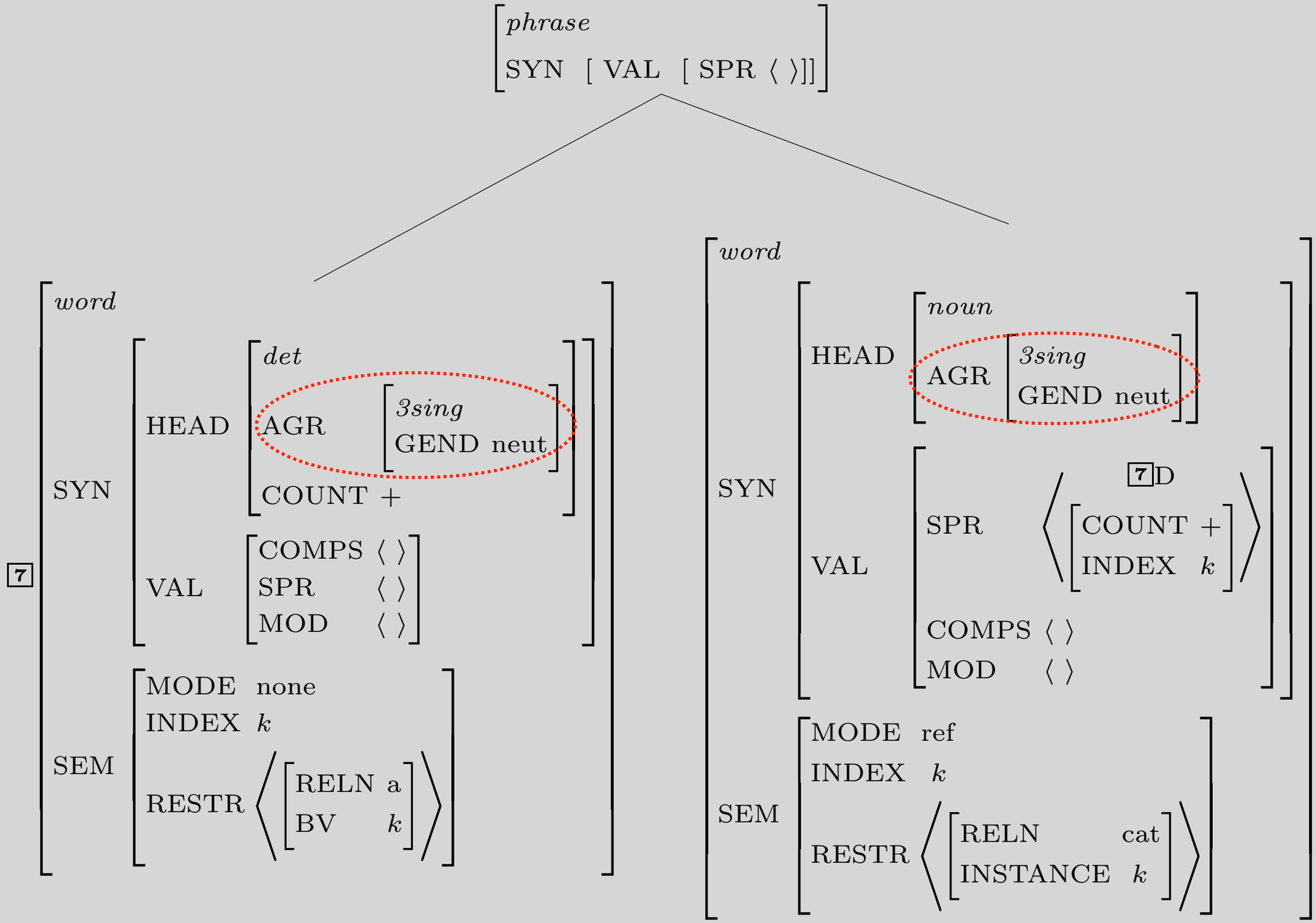
Constraints Contributed by the Grammar Rule

$$\begin{bmatrix} phrase \\ \text{SYN} \;\; [\,\text{VAL} \;\; [\,\text{SPR} \;\langle\,\rangle]] \end{bmatrix}$$

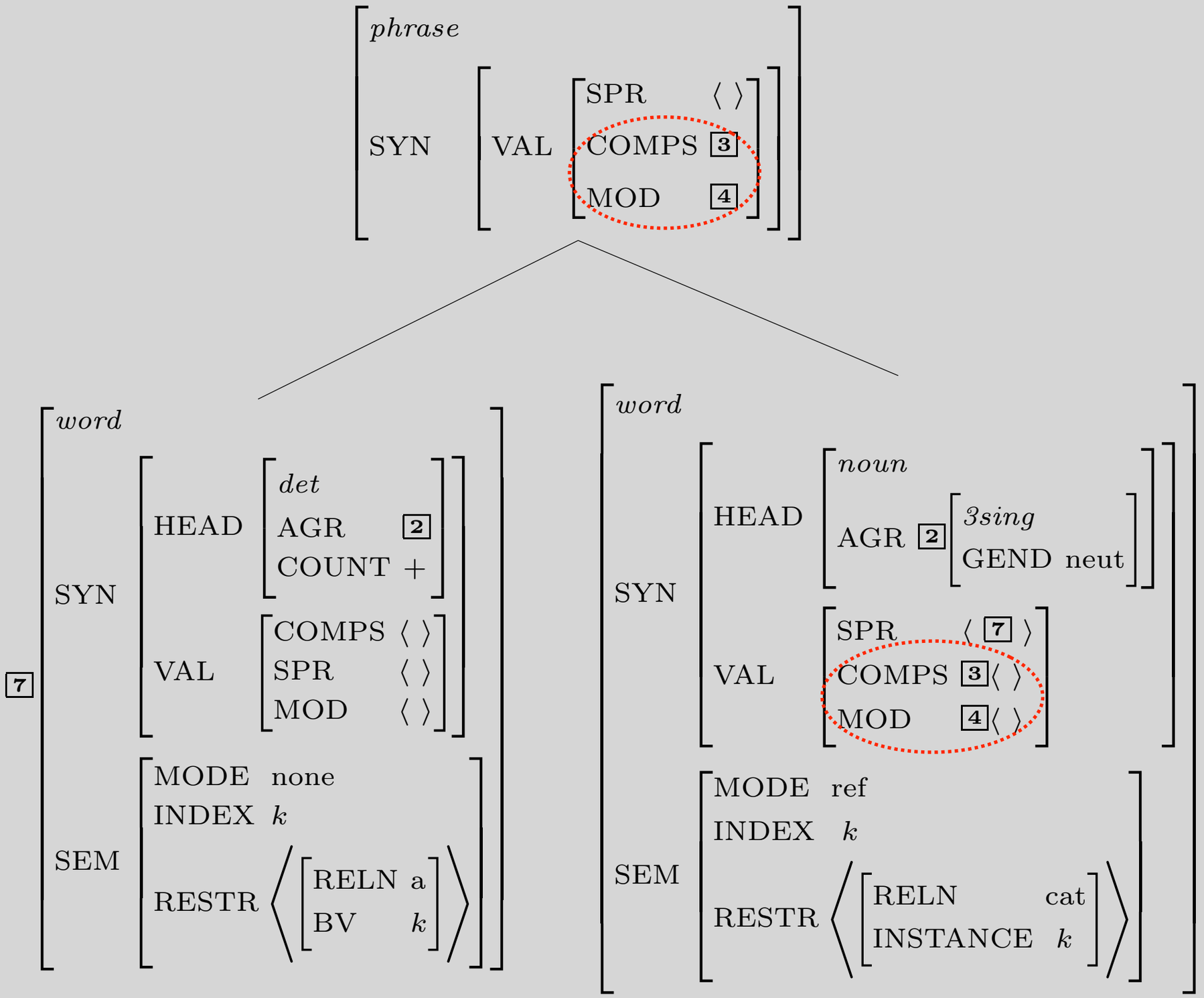
$$\begin{bmatrix} word \\ \text{SYN} \;\; \begin{bmatrix} \text{HEAD} \;\; \begin{bmatrix} det \\ \text{AGR} \;\; \begin{bmatrix} 3sing \\ \text{GEND} \;\; neut \end{bmatrix} \\ \text{COUNT} \;\; + \end{bmatrix} \\ \boxed{7} \\ \text{VAL} \;\; \begin{bmatrix} \text{COMPS} \;\langle\,\rangle \\ \text{SPR} \;\langle\,\rangle \\ \text{MOD} \;\langle\,\rangle \end{bmatrix} \end{bmatrix} \\ \text{SEM} \;\; \begin{bmatrix} \text{MODE} \;\; none \\ \text{INDEX} \;\; k \\ \text{RESTR} \;\; \left\langle \begin{bmatrix} \text{RELN} \;\; a \\ \text{BV} \;\; k \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

$$\begin{bmatrix} word \\ \text{SYN} \;\; \begin{bmatrix} \text{HEAD} \;\; \begin{bmatrix} noun \\ \text{AGR} \;\; \begin{bmatrix} 3sing \\ \text{GEND} \;\; neut \end{bmatrix} \end{bmatrix} \\ \text{VAL} \;\; \begin{bmatrix} \text{SPR} \;\; \left\langle \boxed{7} \begin{bmatrix} \text{COUNT} \;\; + \\ \text{INDEX} \;\; k \end{bmatrix} \right\rangle \\ \text{COMPS} \;\langle\,\rangle \\ \text{MOD} \;\langle\,\rangle \end{bmatrix} \end{bmatrix} \\ \text{SEM} \;\; \begin{bmatrix} \text{MODE} \;\; ref \\ \text{INDEX} \;\; k \\ \text{RESTR} \;\; \left\langle \begin{bmatrix} \text{RELN} \;\; cat \\ \text{INSTANCE} \;\; k \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

A Constraint Involving the SHAC



Effects of the Valence Principle



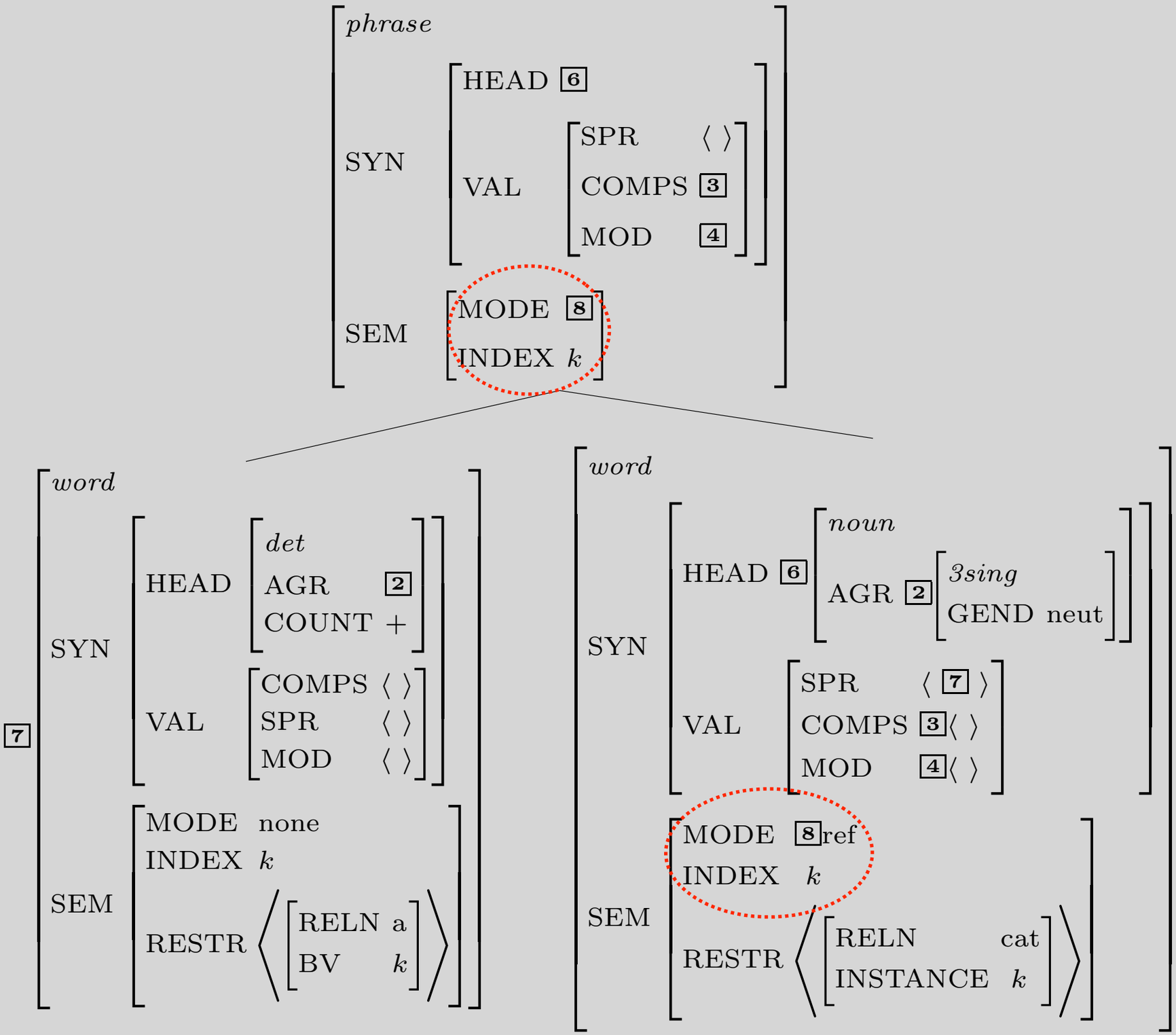
Effects of the Head Feature Principle

$$\begin{bmatrix} \textit{phrase} \\ \text{SYN} \begin{bmatrix} \text{HEAD } \boxed{6} \\ \text{VAL} \begin{bmatrix} \text{SPR} & \langle\,\rangle \\ \text{COMPS} & \boxed{3} \\ \text{MOD} & \boxed{4} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

$$\boxed{7}\begin{bmatrix} \textit{word} \\ \text{SYN} \begin{bmatrix} \text{HEAD} \begin{bmatrix} \textit{det} \\ \text{AGR} & \boxed{2} \\ \text{COUNT} & + \end{bmatrix} \\ \text{VAL} \begin{bmatrix} \text{COMPS} & \langle\,\rangle \\ \text{SPR} & \langle\,\rangle \\ \text{MOD} & \langle\,\rangle \end{bmatrix} \end{bmatrix} \\ \text{SEM} \begin{bmatrix} \text{MODE} & \text{none} \\ \text{INDEX} & k \\ \text{RESTR} & \left\langle \begin{bmatrix} \text{RELN} & \text{a} \\ \text{BV} & k \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

$$\begin{bmatrix} \textit{word} \\ \text{SYN} \begin{bmatrix} \text{HEAD } \boxed{6} \begin{bmatrix} \textit{noun} \\ \text{AGR } \boxed{2} \begin{bmatrix} \textit{3sing} \\ \text{GEND} & \text{neut} \end{bmatrix} \end{bmatrix} \\ \text{VAL} \begin{bmatrix} \text{SPR} & \langle\,\boxed{7}\,\rangle \\ \text{COMPS} & \boxed{3}\langle\,\rangle \\ \text{MOD} & \boxed{4}\langle\,\rangle \end{bmatrix} \end{bmatrix} \\ \text{SEM} \begin{bmatrix} \text{MODE} & \text{ref} \\ \text{INDEX} & k \\ \text{RESTR} & \left\langle \begin{bmatrix} \text{RELN} & \text{cat} \\ \text{INSTANCE} & k \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

Effects of the Semantic Inheritance Principle



Effects of the Semantic Compositionality Principle

$$\left[\begin{array}{ll} \textit{phrase} & \\ \text{SYN} & \left[\begin{array}{ll} \text{HEAD} & \boxed{6} \\ \text{VAL} & \left[\begin{array}{ll} \text{SPR} & \langle\ \rangle \\ \text{COMPS} & \boxed{3} \\ \text{MOD} & \boxed{4} \end{array}\right] \end{array}\right] \\ \text{SEM} & \left[\begin{array}{ll} \text{MODE} & \boxed{8} \\ \text{INDEX} & k \\ \text{RESTR} & \boxed{A} \oplus \boxed{B} \end{array}\right] \end{array}\right]$$

$$\boxed{7}\left[\begin{array}{ll} \textit{word} & \\ \text{SYN} & \left[\begin{array}{ll} \text{HEAD} & \left[\begin{array}{ll} \textit{det} & \\ \text{AGR} & \boxed{2} \\ \text{COUNT} & + \end{array}\right] \\ \text{VAL} & \left[\begin{array}{ll} \text{COMPS} & \langle\ \rangle \\ \text{SPR} & \langle\ \rangle \\ \text{MOD} & \langle\ \rangle \end{array}\right] \end{array}\right] \\ \text{SEM} & \left[\begin{array}{ll} \text{MODE} & \text{none} \\ \text{INDEX} & k \\ \text{RESTR} & \boxed{A}\left\langle\left[\begin{array}{ll} \text{RELN} & \text{a} \\ \text{BV} & k \end{array}\right]\right\rangle \end{array}\right] \end{array}\right]$$

$$\left[\begin{array}{ll} \textit{word} & \\ \text{SYN} & \left[\begin{array}{ll} \text{HEAD} & \boxed{6}\left[\begin{array}{ll} \textit{noun} & \\ \text{AGR} & \boxed{2}\left[\begin{array}{ll} \textit{3sing} & \\ \text{GEND} & \text{neut} \end{array}\right] \end{array}\right] \\ \text{VAL} & \left[\begin{array}{ll} \text{SPR} & \langle\ \boxed{7}\ \rangle \\ \text{COMPS} & \boxed{3}\langle\ \rangle \\ \text{MOD} & \boxed{4}\langle\ \rangle \end{array}\right] \end{array}\right] \\ \text{SEM} & \left[\begin{array}{ll} \text{MODE} & \boxed{8}\text{ref} \\ \text{INDEX} & k \\ \text{RESTR} & \boxed{B}\left\langle\left[\begin{array}{ll} \text{RELN} & \text{cat} \\ \text{INSTANCE} & k \end{array}\right]\right\rangle \end{array}\right] \end{array}\right]$$

Is the Mother Node Now Completely Specified?

$$\left[\begin{array}{ll} \textit{phrase} & \\ \text{SYN} & \left[\begin{array}{ll} \text{HEAD} & \boxed{6} \\ \text{VAL} & \left[\begin{array}{ll} \text{SPR} & \langle\,\rangle \\ \text{COMPS} & \boxed{3} \\ \text{MOD} & \boxed{4} \end{array}\right] \end{array}\right] \\ \text{SEM} & \left[\begin{array}{ll} \text{MODE} & \boxed{8} \\ \text{INDEX} & k \\ \text{RESTR} & \boxed{\text{A}} \oplus \boxed{\text{B}} \end{array}\right] \end{array}\right]$$

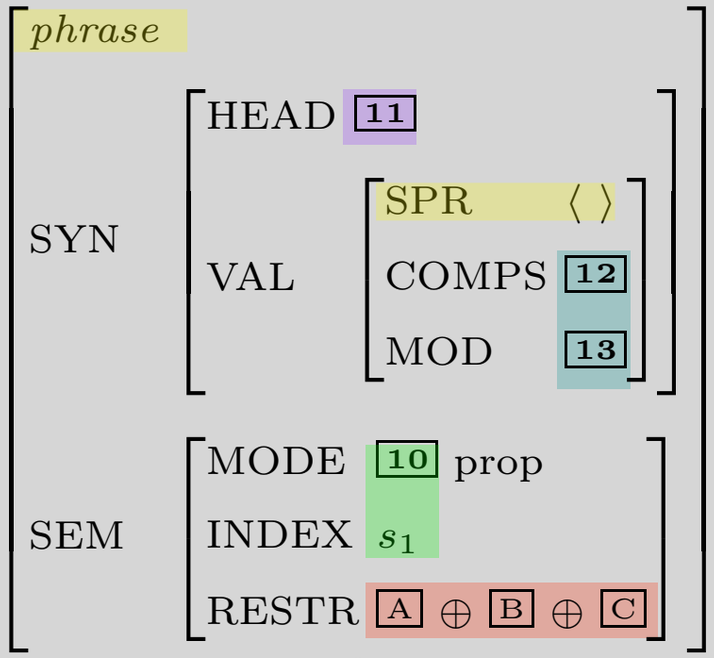
$$\boxed{7}\left[\begin{array}{ll} \textit{word} & \\ \text{SYN} & \left[\begin{array}{ll} \text{HEAD} & \left[\begin{array}{ll} \textit{det} & \\ \text{AGR} & \boxed{2} \\ \text{COUNT} & + \end{array}\right] \\ \text{VAL} & \left[\begin{array}{ll} \text{COMPS} & \langle\,\rangle \\ \text{SPR} & \langle\,\rangle \\ \text{MOD} & \langle\,\rangle \end{array}\right] \end{array}\right] \\ \text{SEM} & \left[\begin{array}{ll} \text{MODE} & \text{none} \\ \text{INDEX} & k \\ \text{RESTR} & \boxed{\text{A}}\left\langle\left[\begin{array}{ll} \text{RELN} & \text{a} \\ \text{BV} & k \end{array}\right]\right\rangle \end{array}\right] \end{array}\right]$$

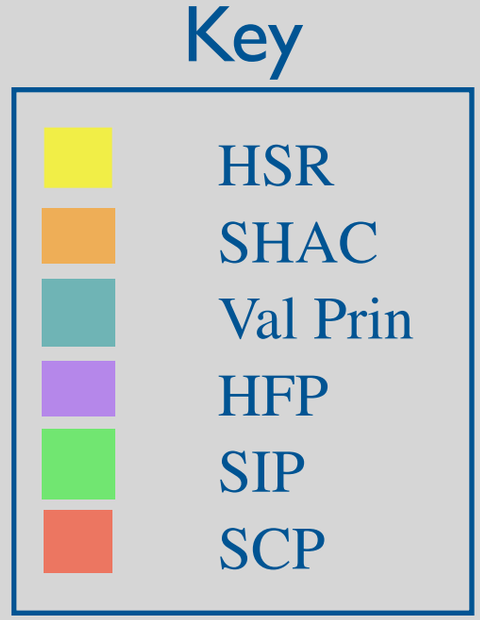
$$\left[\begin{array}{ll} \textit{word} & \\ \text{SYN} & \left[\begin{array}{ll} \text{HEAD} & \boxed{6}\left[\begin{array}{ll} \textit{noun} & \\ \text{AGR} & \boxed{2}\left[\begin{array}{ll} \textit{3sing} & \\ \text{GEND} & \text{neut} \end{array}\right] \end{array}\right] \\ \text{VAL} & \left[\begin{array}{ll} \text{SPR} & \langle\,\boxed{7}\,\rangle \\ \text{COMPS} & \boxed{3}\langle\,\rangle \\ \text{MOD} & \boxed{4}\langle\,\rangle \end{array}\right] \end{array}\right] \\ \text{SEM} & \left[\begin{array}{ll} \text{MODE} & \boxed{8}\text{ref} \\ \text{INDEX} & k \\ \text{RESTR} & \boxed{\text{B}}\left\langle\left[\begin{array}{ll} \text{RELN} & \text{cat} \\ \text{INSTANCE} & k \end{array}\right]\right\rangle \end{array}\right] \end{array}\right]$$

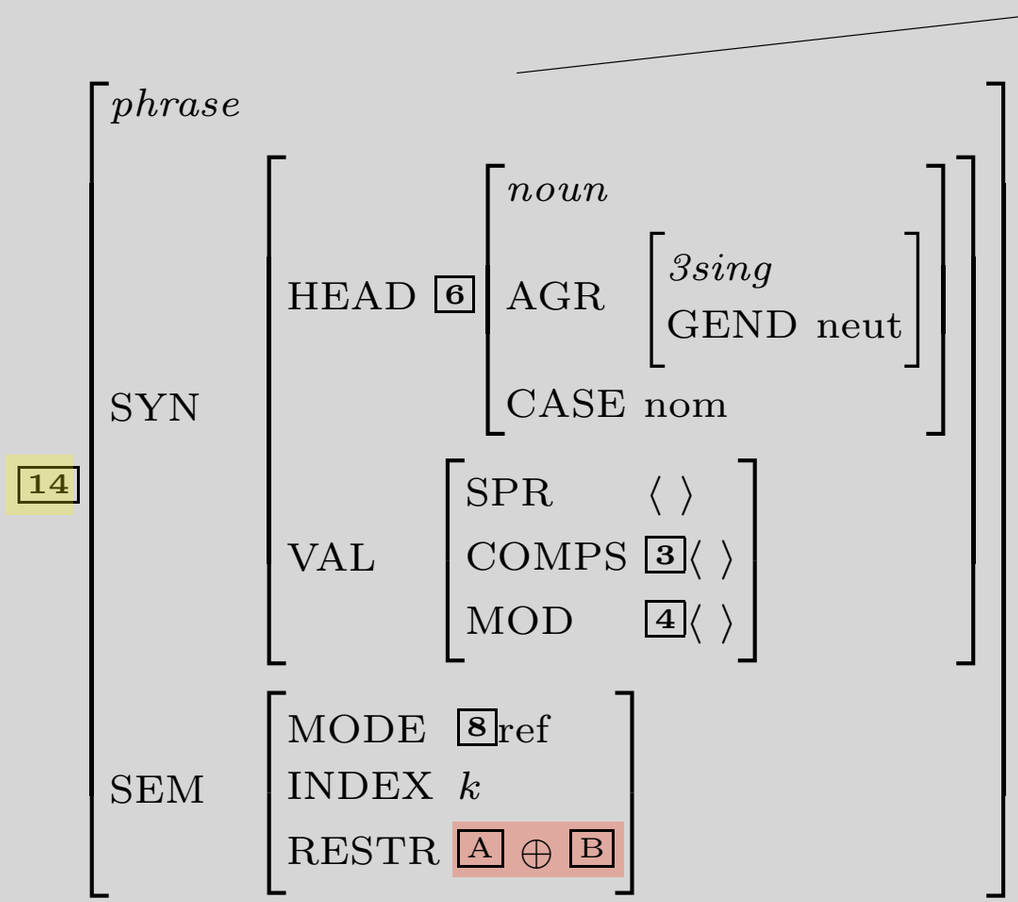
Lexical Entry for *slept*

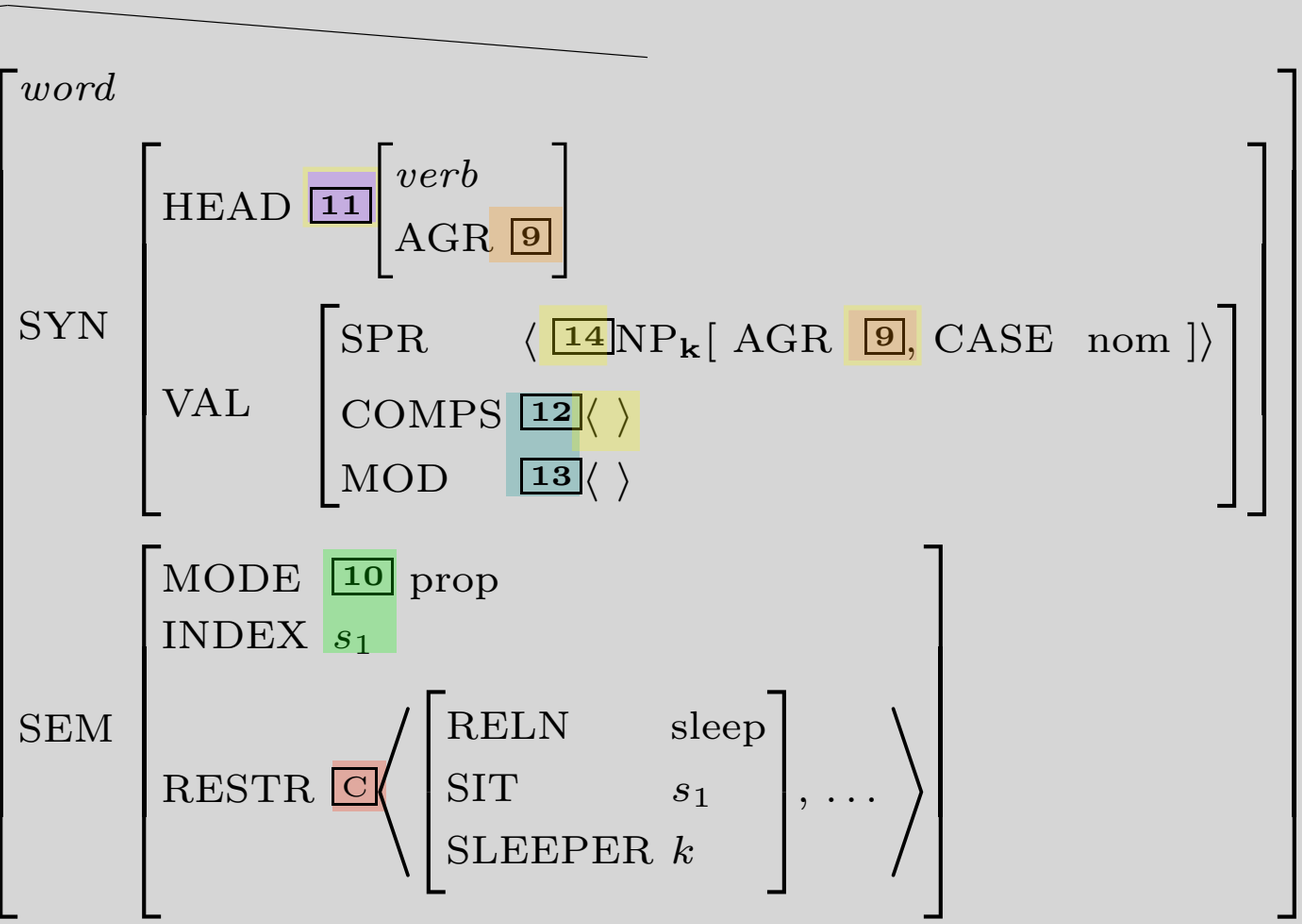
$$\left\langle \text{slept}, \left[\begin{array}{ll} \textit{word} & \\ \text{SYN} & \left[\begin{array}{ll} \text{HEAD} & \textit{verb} \\ \text{VAL} & \left[\begin{array}{ll} \text{SPR} & \left\langle \left[\begin{array}{ll} \text{AGR} & \boxed{9} \\ \text{CASE} & \text{nom} \end{array}\right] \right\rangle_{\text{NP}_m} \\ \text{COMPS} & \langle\ \rangle \\ \text{MOD} & \langle\ \rangle \end{array}\right] \end{array}\right] \\ \text{SEM} & \left[\begin{array}{ll} \text{INDEX} & s_1 \\ \text{MODE} & \text{prop} \\ \text{RESTR} & \left\langle \left[\begin{array}{ll} \text{RELN} & \text{sleep} \\ \text{SIT} & s_1 \\ \text{SLEEPER} & m \end{array}\right], \ldots \right\rangle \end{array}\right] \end{array}\right] \right\rangle$$

Another Head-Specifier Phrase









Is this description fully specified?

$$\begin{bmatrix} \textit{phrase} & \\ \text{SYN} & \begin{bmatrix} \text{HEAD } \boxed{11} & \\ \text{VAL} & \begin{bmatrix} \text{SPR} & \langle\,\rangle \\ \text{COMPS} & \boxed{12} \\ \text{MOD} & \boxed{13} \end{bmatrix} \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \text{MODE} & \boxed{10}\ \text{prop} \\ \text{INDEX} & s_1 \\ \text{RESTR} & \boxed{\text{A}} \oplus \boxed{\text{B}} \oplus \boxed{\text{C}} \end{bmatrix} \end{bmatrix}$$

$$\boxed{14}\begin{bmatrix} \textit{phrase} & \\ \text{SYN} & \begin{bmatrix} \text{HEAD } \boxed{6} & \begin{bmatrix} \textit{noun} & \\ \text{AGR} & \begin{bmatrix} \textit{3sing} & \\ \text{GEND} & \text{neut} \end{bmatrix} \\ \text{CASE} & \text{nom} \end{bmatrix} \\ \text{VAL} & \begin{bmatrix} \text{SPR} & \langle\,\rangle \\ \text{COMPS} & \boxed{3}\langle\,\rangle \\ \text{MOD} & \boxed{4}\langle\,\rangle \end{bmatrix} \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \text{MODE} & \boxed{8}\text{ref} \\ \text{INDEX} & k \\ \text{RESTR} & \boxed{\text{A}} \oplus \boxed{\text{B}} \end{bmatrix} \end{bmatrix}$$

$$\begin{bmatrix} \textit{word} & \\ \text{SYN} & \begin{bmatrix} \text{HEAD } \boxed{11} & \begin{bmatrix} \textit{verb} & \\ \text{AGR} & \boxed{9} \end{bmatrix} \\ \text{VAL} & \begin{bmatrix} \text{SPR} & \langle\, \boxed{14}\text{NP}_{\mathbf{k}}[\ \text{AGR}\ \boxed{9},\ \text{CASE}\ \ \text{nom}\ ]\rangle \\ \text{COMPS} & \boxed{12}\langle\,\rangle \\ \text{MOD} & \boxed{13}\langle\,\rangle \end{bmatrix} \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \text{MODE} & \boxed{10}\ \text{prop} \\ \text{INDEX} & s_1 \\ \text{RESTR} & \boxed{\text{C}}\left\langle \begin{bmatrix} \text{RELN} & \text{sleep} \\ \text{SIT} & s_1 \\ \text{SLEEPER} & k \end{bmatrix}, \ldots \right\rangle \end{bmatrix} \end{bmatrix}$$

Does the top node satisfy the initial symbol?

$$\left[\begin{array}{ll} \textit{phrase} & \\ \text{SYN} & \left[\begin{array}{ll} \text{HEAD} & \boxed{11} \\ \text{VAL} & \left[\begin{array}{ll} \text{SPR} & \langle\,\rangle \\ \text{COMPS} & \boxed{12} \\ \text{MOD} & \boxed{13} \end{array}\right] \end{array}\right] \\ \text{SEM} & \left[\begin{array}{ll} \text{MODE} & \boxed{10}\ \text{prop} \\ \text{INDEX} & s_1 \\ \text{RESTR} & \boxed{\text{A}} \oplus \boxed{\text{B}} \oplus \boxed{\text{C}} \end{array}\right] \end{array}\right]$$

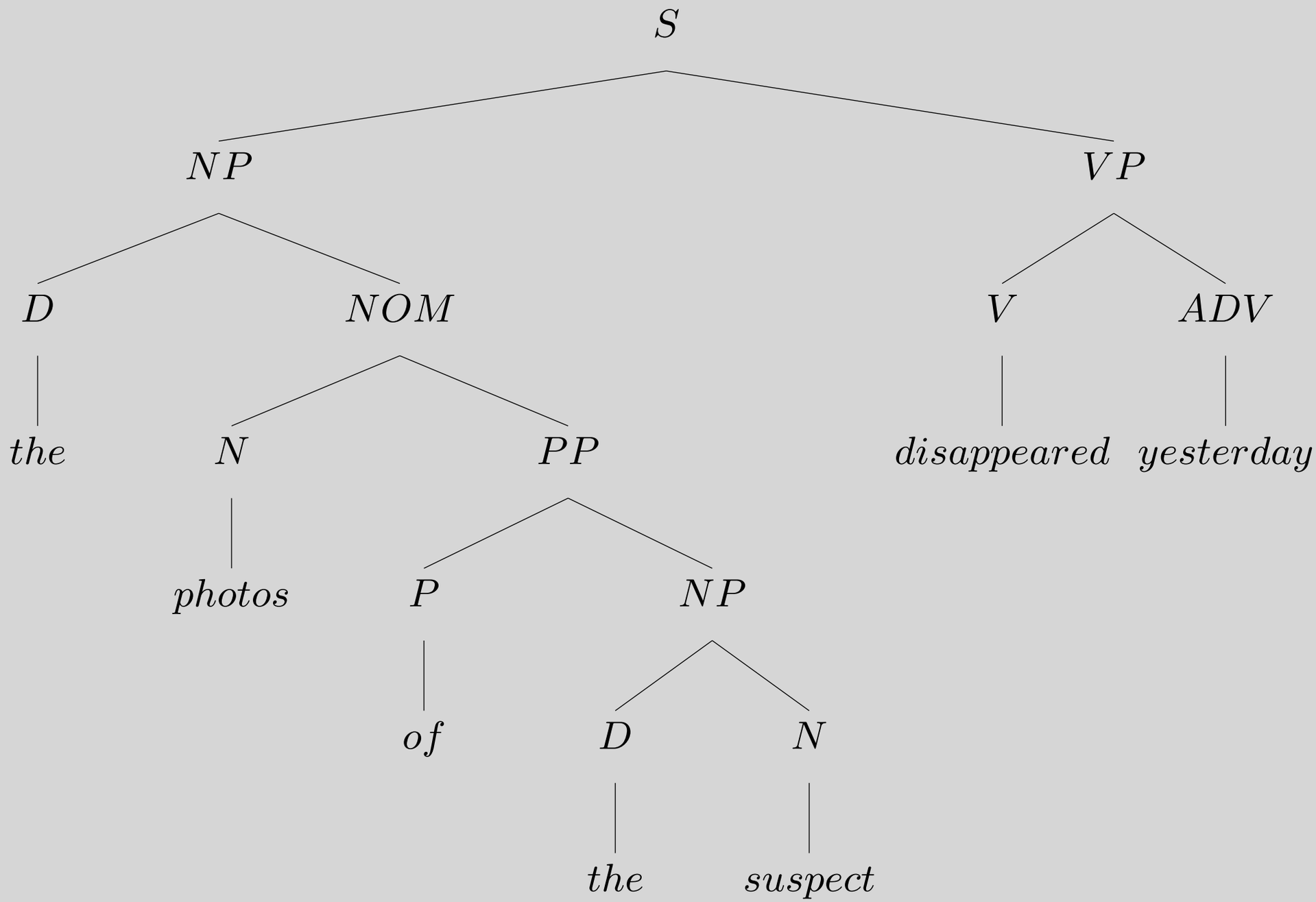
$$\boxed{14}\left[\begin{array}{ll} \textit{phrase} & \\ \text{SYN} & \left[\begin{array}{ll} \text{HEAD} & \boxed{6}\left[\begin{array}{ll} \textit{noun} & \\ \text{AGR} & \left[\begin{array}{ll} \textit{3sing} & \\ \text{GEND} & \text{neut} \end{array}\right] \\ \text{CASE} & \text{nom} \end{array}\right] \\ \text{VAL} & \left[\begin{array}{ll} \text{SPR} & \langle\,\rangle \\ \text{COMPS} & \boxed{3}\langle\,\rangle \\ \text{MOD} & \boxed{4}\langle\,\rangle \end{array}\right] \end{array}\right] \\ \text{SEM} & \left[\begin{array}{ll} \text{MODE} & \boxed{8}\text{ref} \\ \text{INDEX} & k \\ \text{RESTR} & \boxed{\text{A}} \oplus \boxed{\text{B}} \end{array}\right] \end{array}\right]$$

$$\left[\begin{array}{ll} \textit{word} & \\ \text{SYN} & \left[\begin{array}{ll} \text{HEAD} & \boxed{11}\left[\begin{array}{ll} \textit{verb} & \\ \text{AGR} & \boxed{9} \end{array}\right] \\ \text{VAL} & \left[\begin{array}{ll} \text{SPR} & \langle\ \boxed{14}\text{NP}_{\mathbf{k}}[\ \text{AGR}\ \boxed{9},\ \text{CASE}\ \ \text{nom}\ ]\rangle \\ \text{COMPS} & \boxed{12}\langle\,\rangle \\ \text{MOD} & \boxed{13}\langle\,\rangle \end{array}\right] \end{array}\right] \\ \text{SEM} & \left[\begin{array}{ll} \text{MODE} & \boxed{10}\ \text{prop} \\ \text{INDEX} & s_1 \\ \text{RESTR} & \boxed{\text{C}}\left\langle\left[\begin{array}{ll} \text{RELN} & \text{sleep} \\ \text{SIT} & s_1 \\ \text{SLEEPER} & k \end{array}\right], \ldots\right\rangle \end{array}\right] \end{array}\right]$$

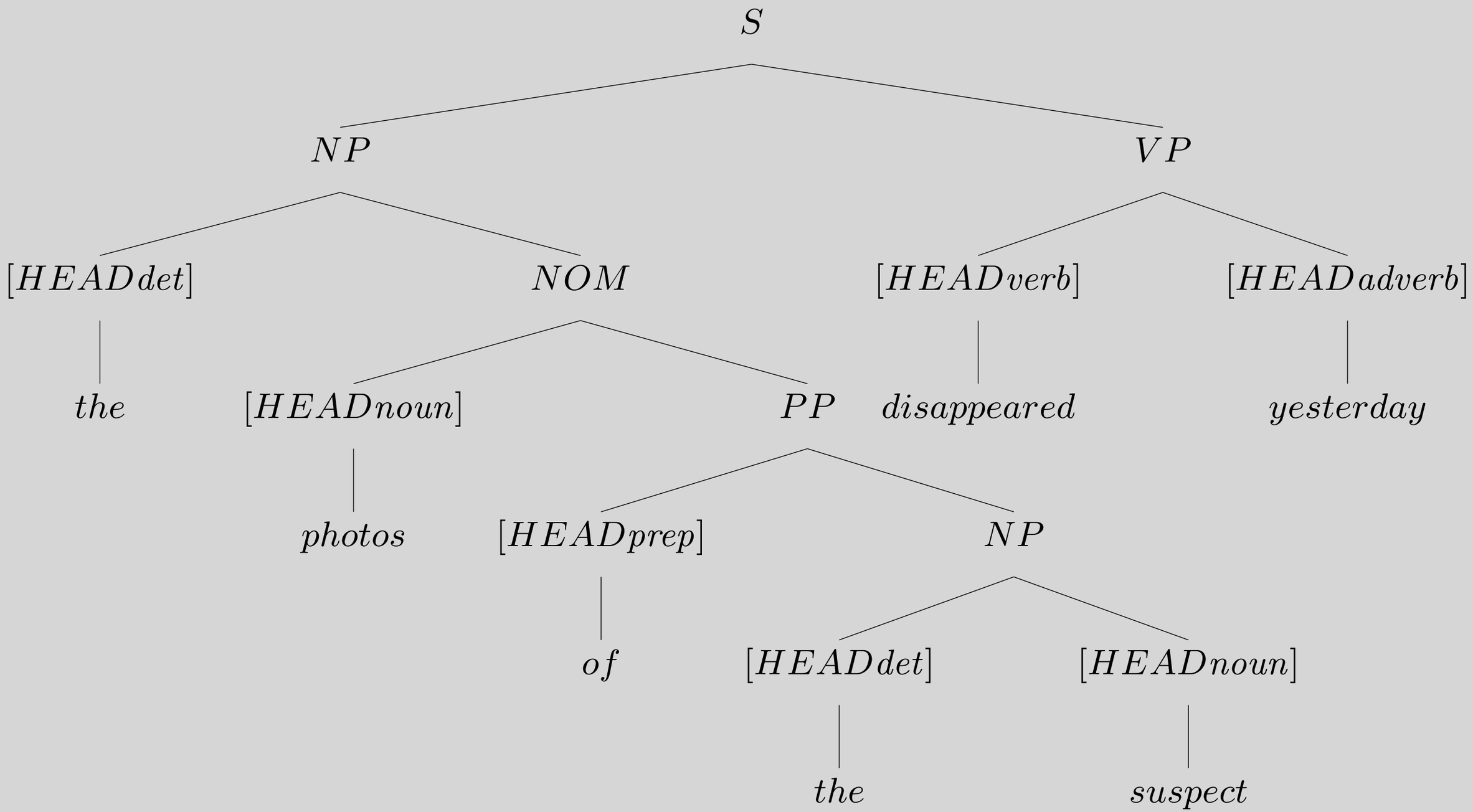
RESTR of the S node

$$\left\langle \begin{bmatrix} \text{RELN} & \text{a} \\ \text{BV} & k \end{bmatrix}, \begin{bmatrix} \text{RELN} & \text{cat} \\ \text{INST} & k \end{bmatrix}, \begin{bmatrix} \text{RELN} & \text{sleep} \\ \text{SIT} & s_1 \\ \text{SLEEPER} & k \end{bmatrix}, \ldots \right\rangle$$

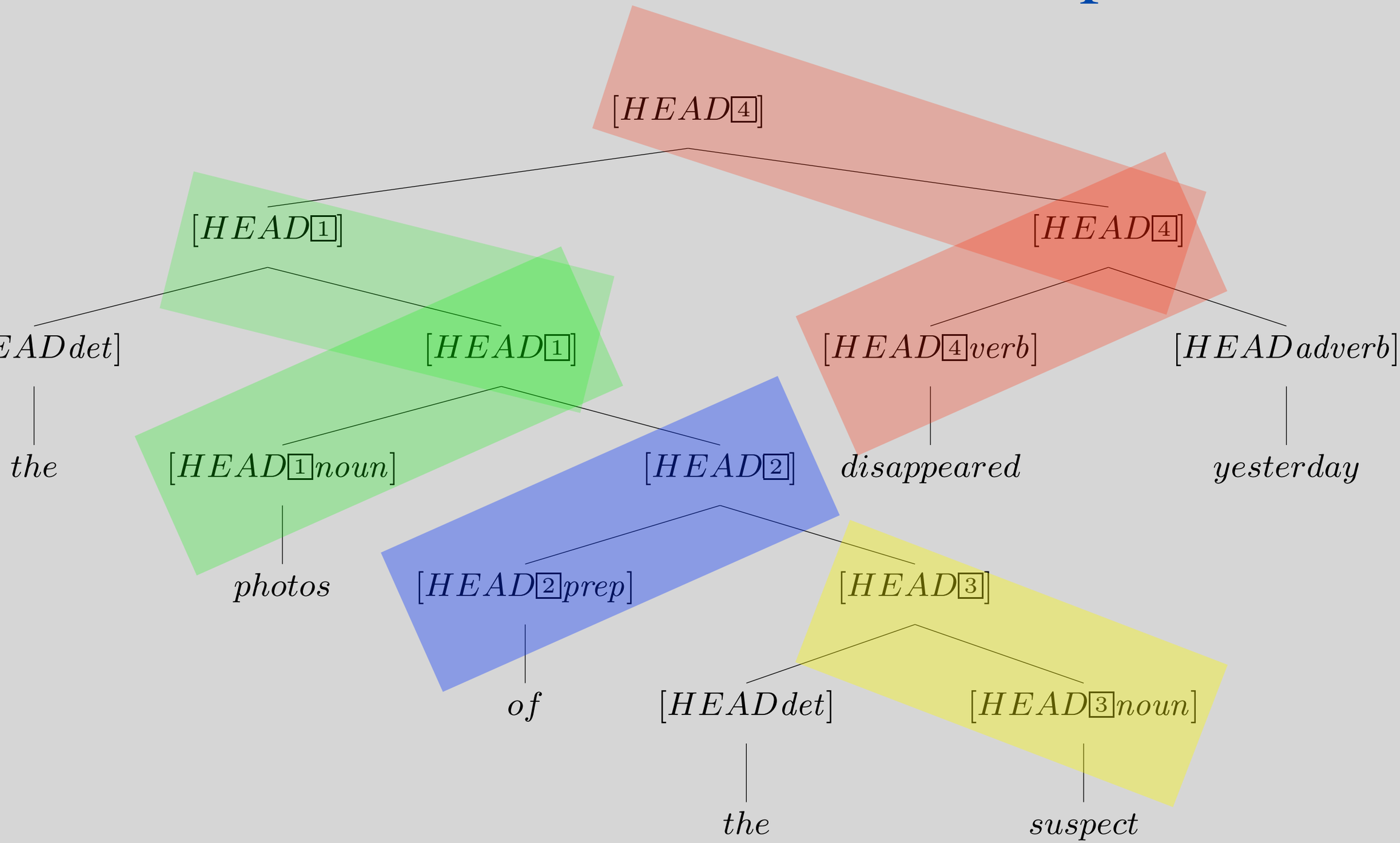
Another Example



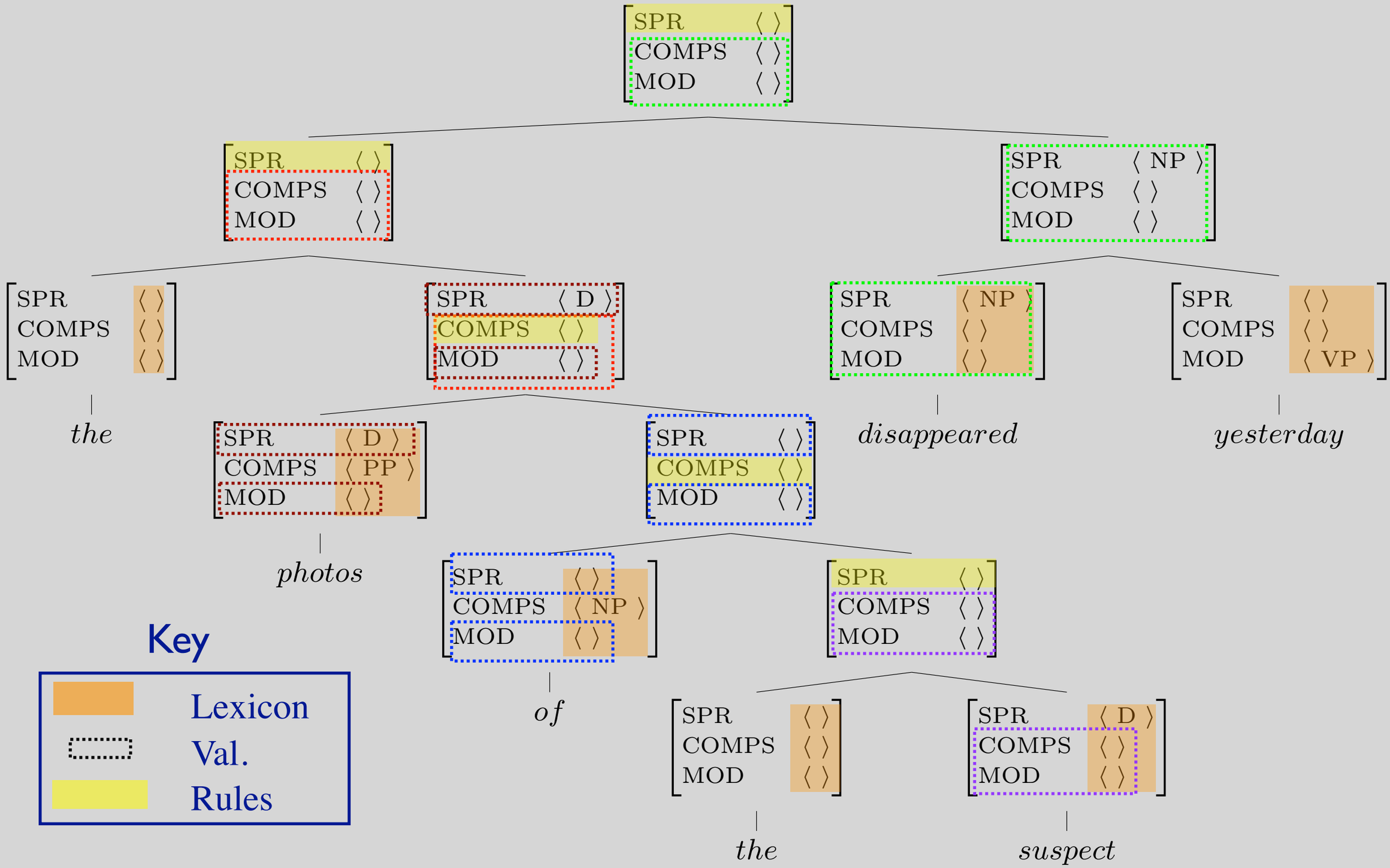
Head Features from Lexical Entries



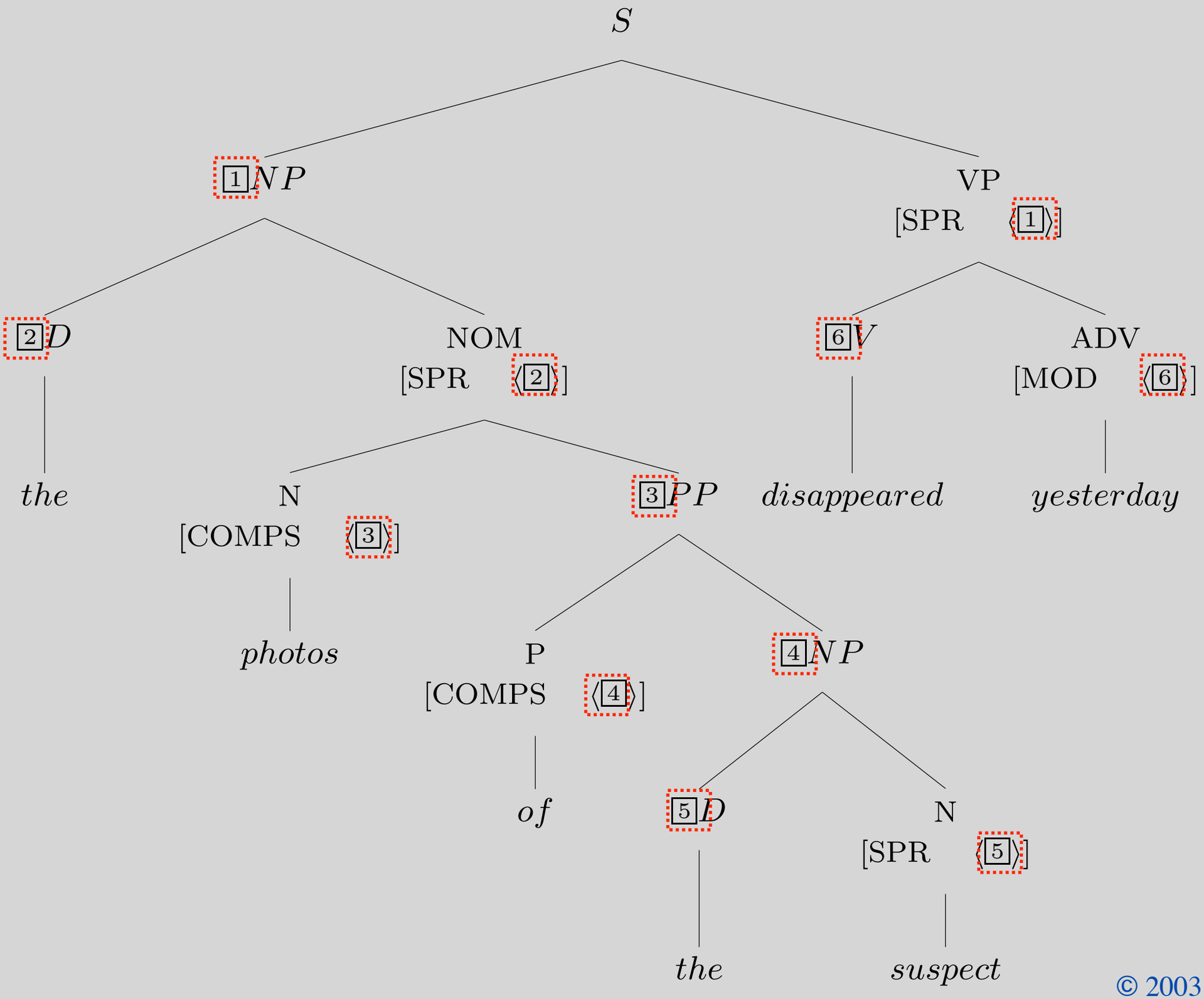
Head Features from Lexical Entries, plus HFP



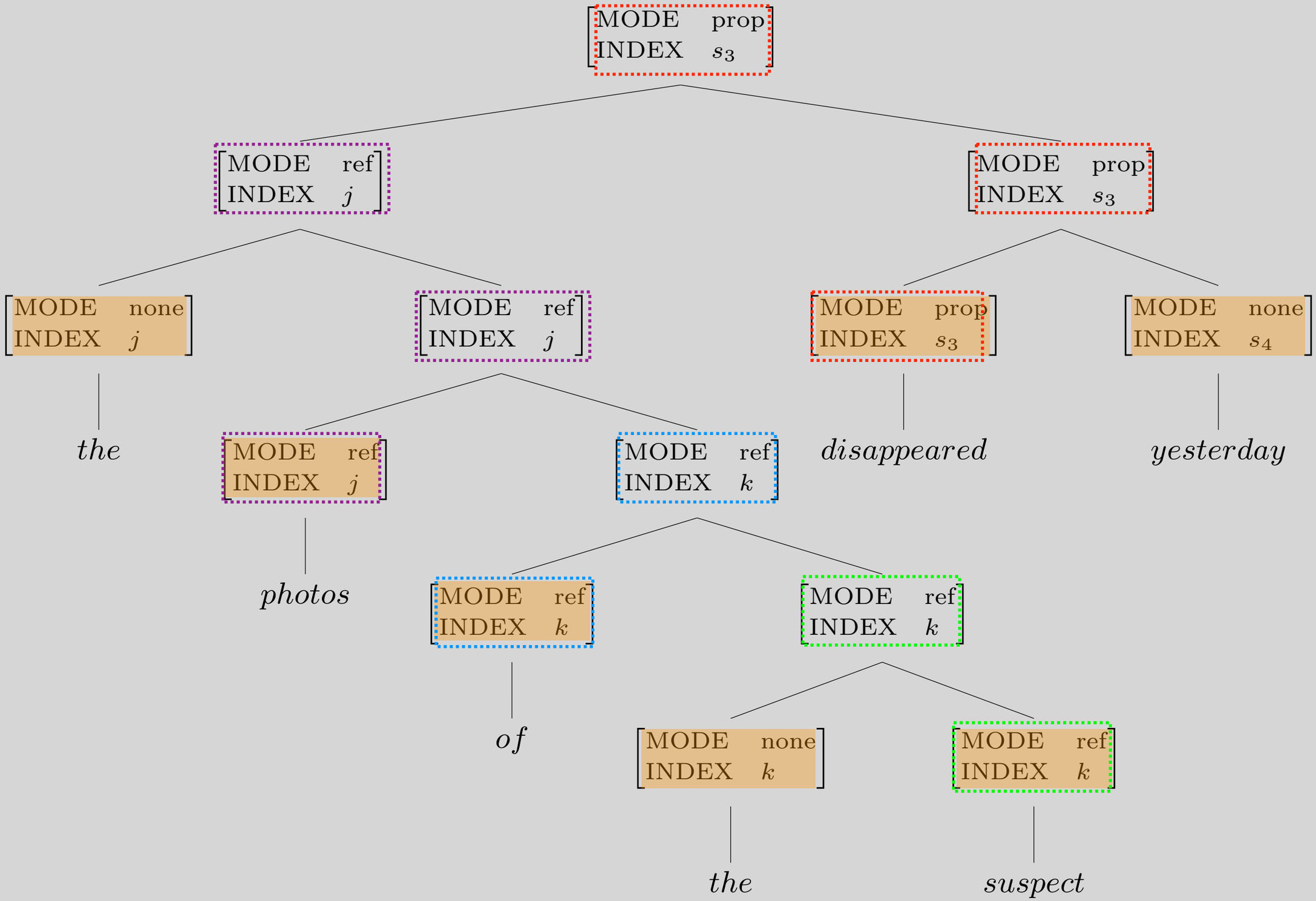
Valence Features: Lexicon, Rules, and the Valence Principle



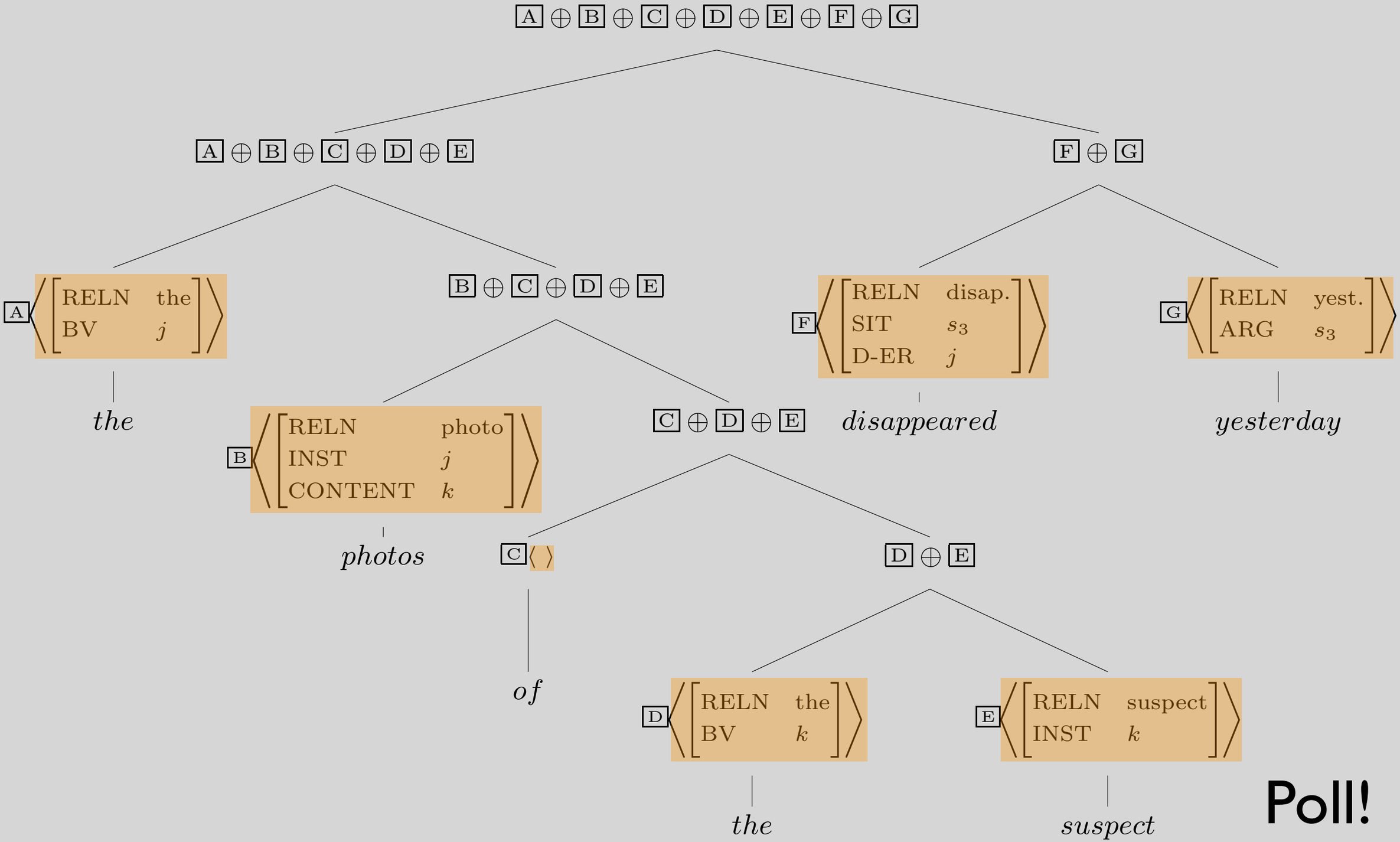
Required Identities: Grammar Rules



Two Semantic Features: the Lexicon & SIP

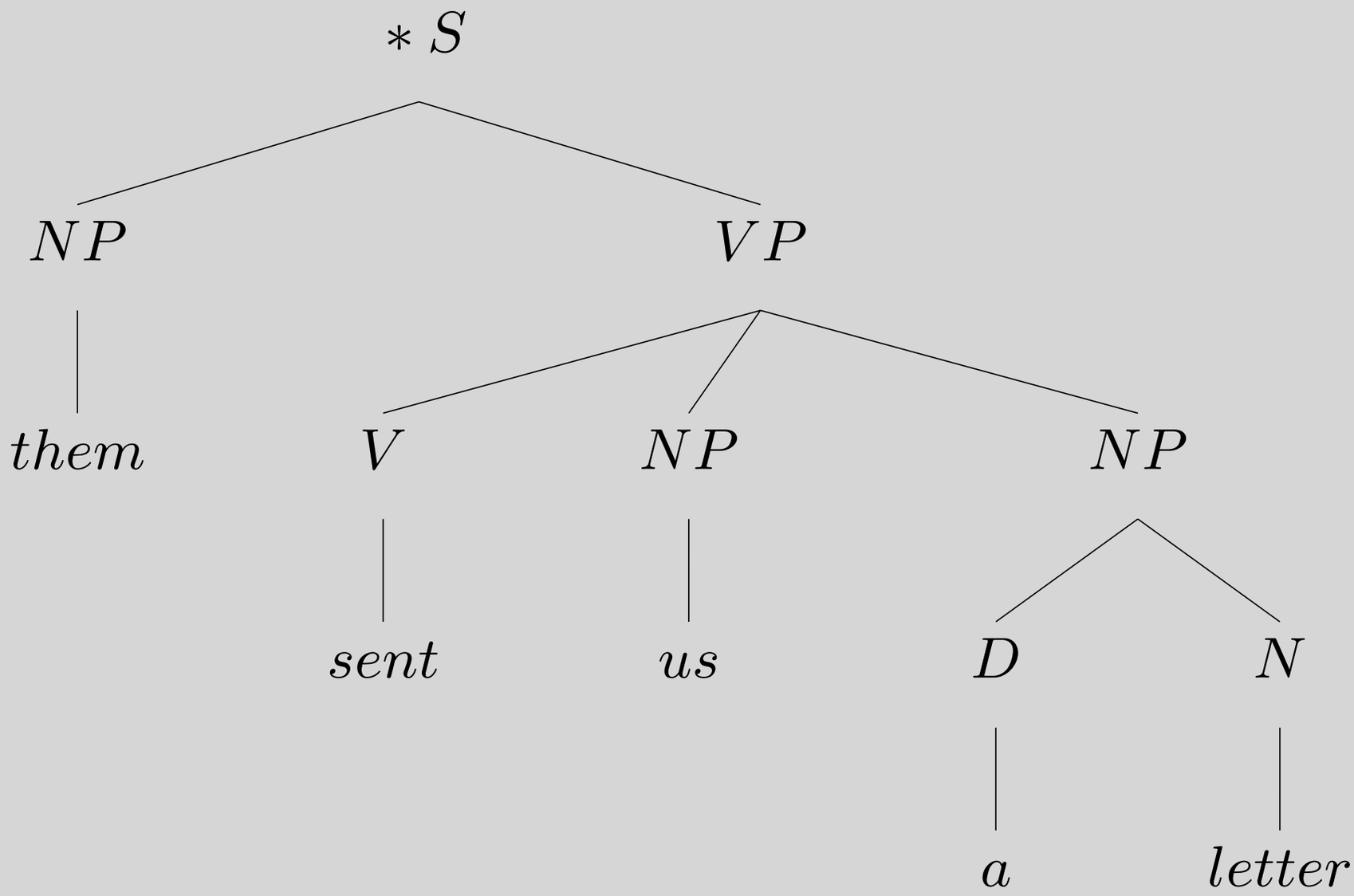


RESTR Values and the SCP



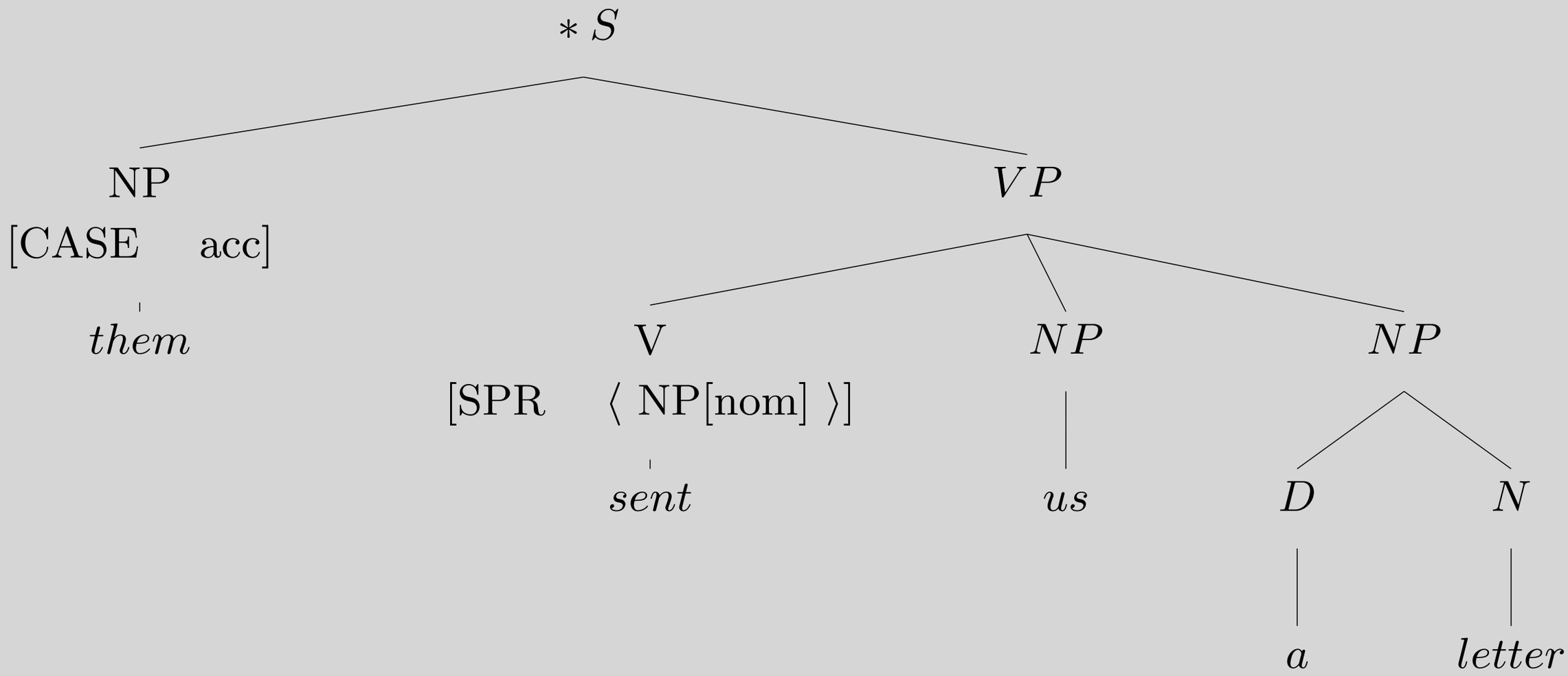
Poll!

An Ungrammatical Example



What's wrong with this sentence?

An Ungrammatical Example

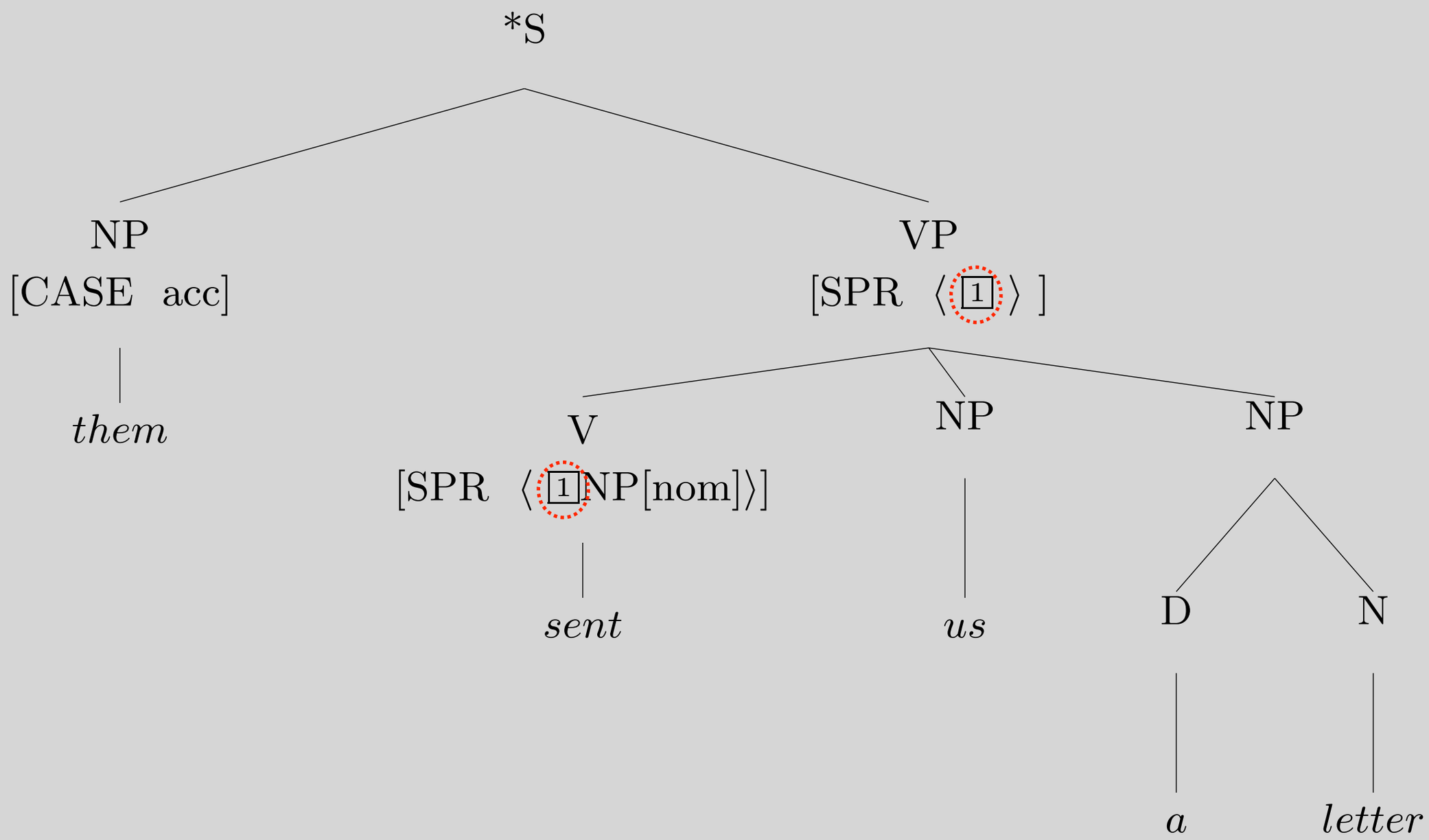


What's wrong with this sentence?

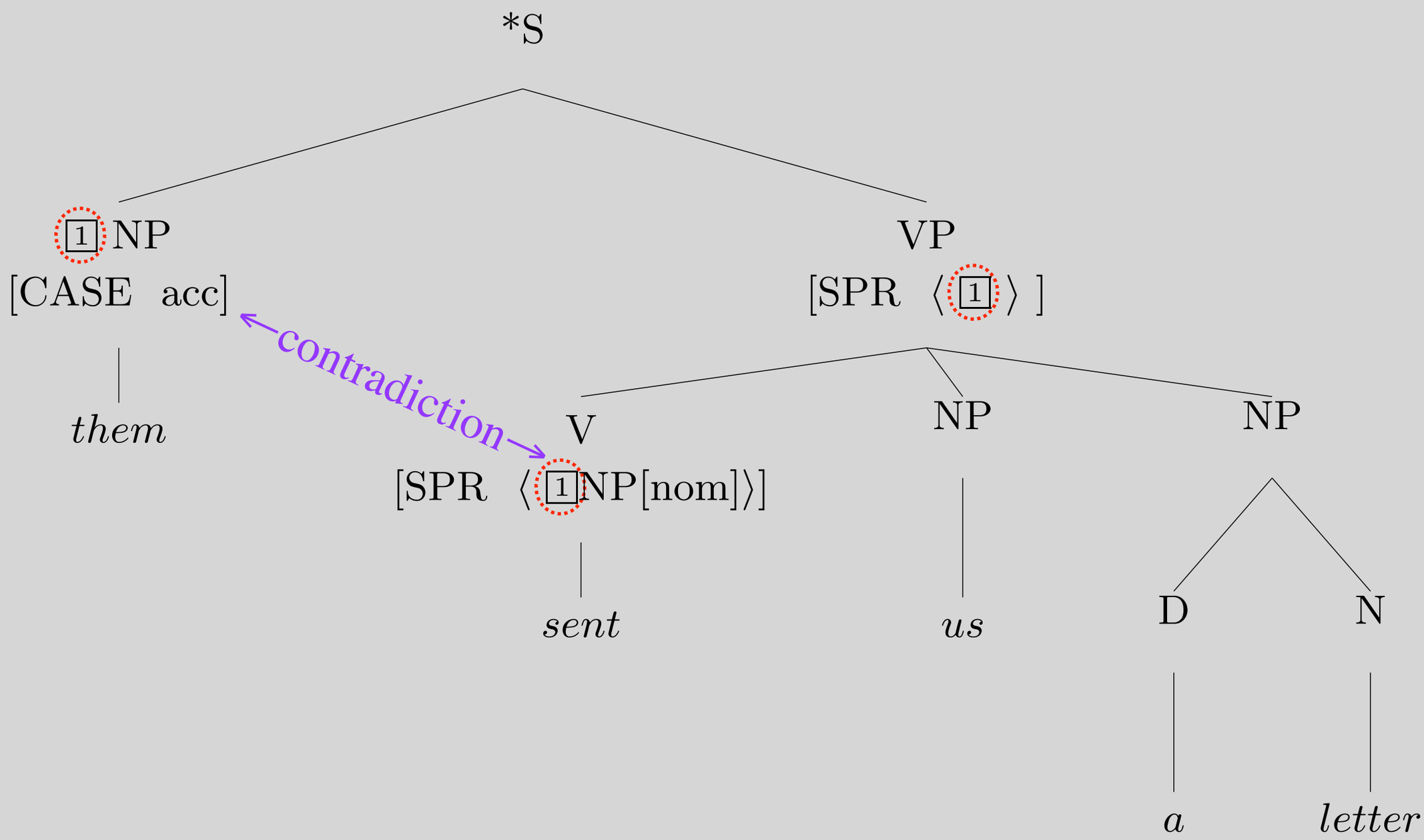
So what?

An Ungrammatical Example

The Valence Principle



An Ungrammatical Example



Reading Questions

- One major question I have is, oddly enough, where do we go from here? More specifically, what are the high-level cases where this grammar fails to account for English grammar, and where do future grammar frameworks fit in? Does the grammar developed thus far encompass any natural language?

Poll!

Exercise in Critical Thinking

- Our grammar has come a long way since Ch 2, as we've added ways of representing different kinds of information:
  - generalizations across categories
  - semantics
  - particular linguistic phenomena: valence, agreement, modification
- What else might we add? What facts about language are as yet unrepresented in our model?

Overview

- What we're trying to do
- The pieces of our grammar
- Two extended examples
- Reflection on what we've done, what we still have to do
- Reading questions

Reading Questions

- On the distinction between lexical entries and word structures. Do we consider lexical entries to be the most "general" (underspecified) version of the word structures that they license? If so, when do we draw the line for underspecification?

Reading Questions

- When is it better to have multiple entries into a RESTR list for a given word, rather than just a bigger predication for that word? For example, "us" has 3, but could maybe be created with fewer.

- So, in looking at the lexical semantic description of "us" (group+member+speaker), I'm again left feeling like the RESTR features are a bit arbitrary and I'm wondering - are linguists the ones who go out and define these, or can they be extracted more algorithmically? If not the latter, then how can this grammar learn from modern evolutions of language?

- In (11), the index 1 identifying the speaker of 'us' is never used to reference anything beyond the internal structure of the RESTR of 'us' when part of 'They sent us a letter', as it just identifies the speaker as a part of the group 'us'. Is 1 specified because there are scenarios in which another element in a sentence requires a reference to the speaker? For example, in the sentence 'I bought dinner for us', would the SEM of 'I' involve the index of the speaker of 'us'?

Reading Questions

- For (8), is it necessary to keep the ADDRESSEE role when 'letter' doesn’t happen to take a PP complement and there is no constituent to take on the role? In this case, would there be issues if we didn’t include that feature in the letter predication?

- In (8), the addressee (indexed m) was never specified. This was justified because we never said that SEM had to be fully realized. I understand how this is useful for passive structures but intuitively it feels like this wouldn't work every time. Are there any examples of where not needing SEM to be fully realized causes issues?

Reading Questions

- In the RESTR list for letter, it makes sense to me that a letter would necessarily have an addressee, so I get why we might want to note that even if it's unspecified in the sentence. Could we theoretically do something similar with other "mandatory" aspects of a letter that may not be made explicit in the sentence, like what it's about ("CONTENT", maybe)? I guess I'm curious to hear more discussion of the formation of RESTR lists and determining which semantic roles are salient.

Reading Questions

- It makes sense to me that in a sentence where a letter's addressee is unspecified to keep the addressee feature, but how would we semantically represent phrases like "open letter"?

- For (32), the only difference between these two predications is the label for NAMED. In this specific case of Lee, while j stands for the SENDEE argument of the verb, m stands for the ADDRESSEE role with respect to the noun letters. What is the convention here for j and m? How do we generate such

We [send [two letters to Lee]]
We [send [two letters] [to Lee]]

Reading Questions

- BV so far has been used with lexical entry for *a*. Is there other example of using BV? When do we know that we need it?

Reading Questions

- The tree to 'to Lee' in (28) shows that the semantics for the preposition 'to' is an empty list with an index to its complement, and so the semantics for Lee get pulled up to the mother node, and the entire phrase 'to Lee' just denotes 'Lee'.

- The motivation for giving 'to' no semantics is that prepositions in English function more like vestigial case markers than like 'real' words, but I would argue that both case markers and prepositions add semantics to phrases and sentences. For instance, the sentences "We send two letters from/about/for Lee" all have different interpretations, relying entirely on the preposition. The chapter handles the subtle difference in meaning of "send two letters to Lee" by changing the lexical entry for Lee when adding it to the tree, but I wonder if we could have added something to the semantics of 'to' that would allow it to optionally denote just 'Lee', but also optionally denote 'Lee-as-a-recipient'.

- This is a little emblematic of my skepticism that our semantics is really explaining anything when we (as English speakers who know the intended sentence meanings) are the ones setting the variables so that they interact in the correct way. Shouldn't there be a way to decide based on the syntax whether we index Lee as j or m, to derive the correct interpretation?

Reading Questions

- On page 174, because the word sent is in past tense, the lexical entry has an underspecified AGR value. I was wondering if our course will discuss tense more and if we will learn rules about them in our grammar? I was thinking about this during the last homework as well because two of the lexical entries we had to do (grew and seemed) were also in past tense, and I wondered if we needed to treat them differently in the lexical entries than their present tense forms.

Reading Questions

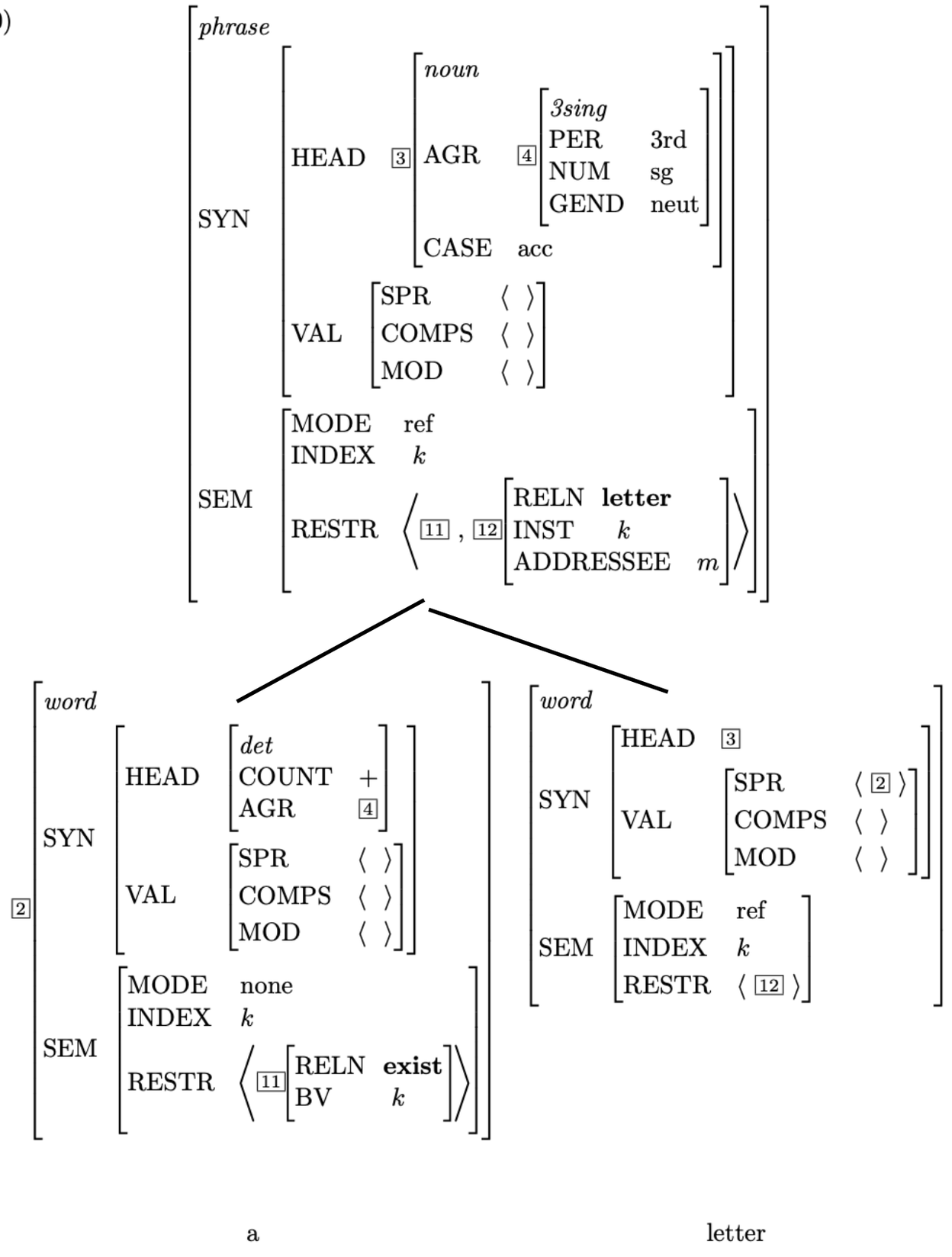
- I think I am still a bit confused by the direct sum notation on the RESTR values. Can we give a few more examples of when TO and when NOT TO use them?

Reading Questions

- Up until now, we have been identifying the HEAD value of the mother with that of the head daughter. (10) does the exact opposite. Similarly, in (10), the RESTR is a sum of all the daughters' RESTR values, yet the RESTR value of the head daughter is identified with that of the mother. Is this a consequence of how our trees aren't inherently bottom-up/ top-down?

- To me this is problematic because it seems to imply that features like RESTR and HEAD don't have a specific point of origin. That is, as long as the values match between mother and daughter where needed, it doesn't matter where the values are expressed. Up until now, I felt that these values came from the lexicon and thus intuitively felt that they must be specified as close to the word level as possible.

(10)



Reading Questions

- In (10), it mentions that following HSR created a link in a chain of identities, that letter identifies the INDEX of the element on its SPR list with its own INDEX and INST values. Then combined with node a the INDEX of a is identified. Does the order of the identification matter?

- In page 168, "The job of determining well-formedness can be distributed among the various pieces of our grammatical system because the licensing mechanism requires simultaneous satisfaction of all of the relevant constraints" makes me wonder if a tree is built in a multithreading fashion? If so, will there be any order among the four rules (e.g. HSR, HCR, HMR, and coordination rule)?

Reading Questions

- What's the difference between grammar rules and principles? It seems they are the same in the analysis of verifying a well-formed tree.

Reading Questions

- How can this theory of syntax has been/can be applied in computational settings. More specifically, comparing this to syntactic formalisms such CFG and semantic formalisms such as WordNet, this theory that are we covering seems to have lexical entries with a lot more information than those formalisms.

- In practice, what are some of the advantages of this theory over more established formalisms?

- With more information (ie features), I would not be surprised if it can parse sentences more precisely (ie not overgeneralize) based on lexical entries. However, another important aspect might be recall - how well can unseen sentences be parsed?

- Finally, the number of features seems to mean that annotation might be relatively more difficult than something like the Penn Tree Bank. How is this challenge handled?

Reading Questions

- How does the Semantic Compositionality Principle work in context heavy language like Japanese in which some of the lexical entries with the necessary RESTR values are not present in the sentence? Does the top node contain the sum of its daughters as well as its imaginary daughters?