Ling 566
Oct 11, 2012

How the Grammar Works

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*

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

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

Lexical Entry for *cat*

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

- 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

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

Description of Word Structures for *cat*

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

cat

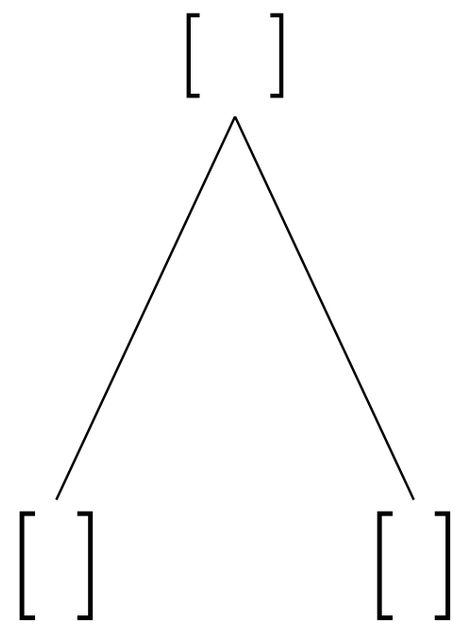
Description of Word Structures for *a*

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

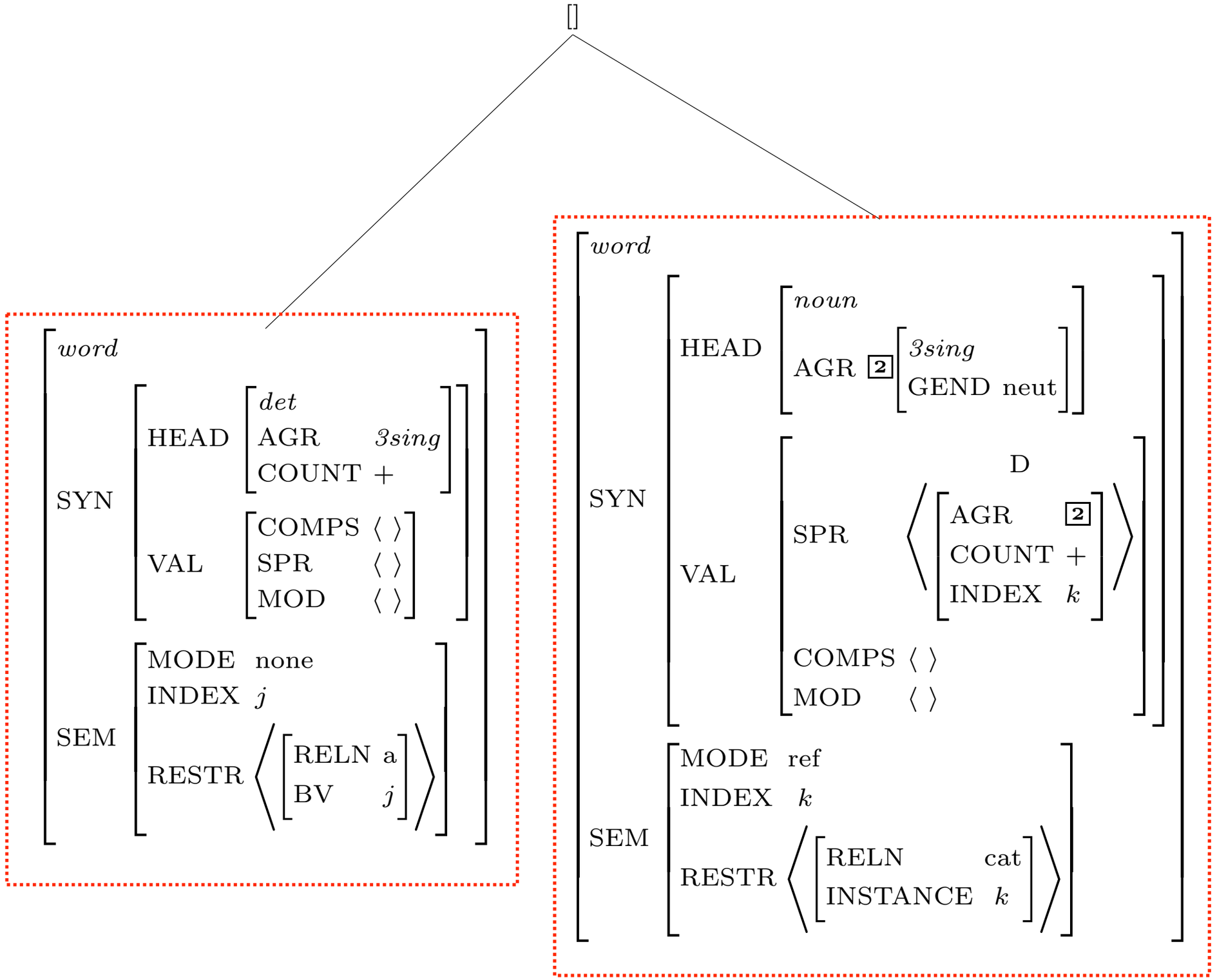
|

a

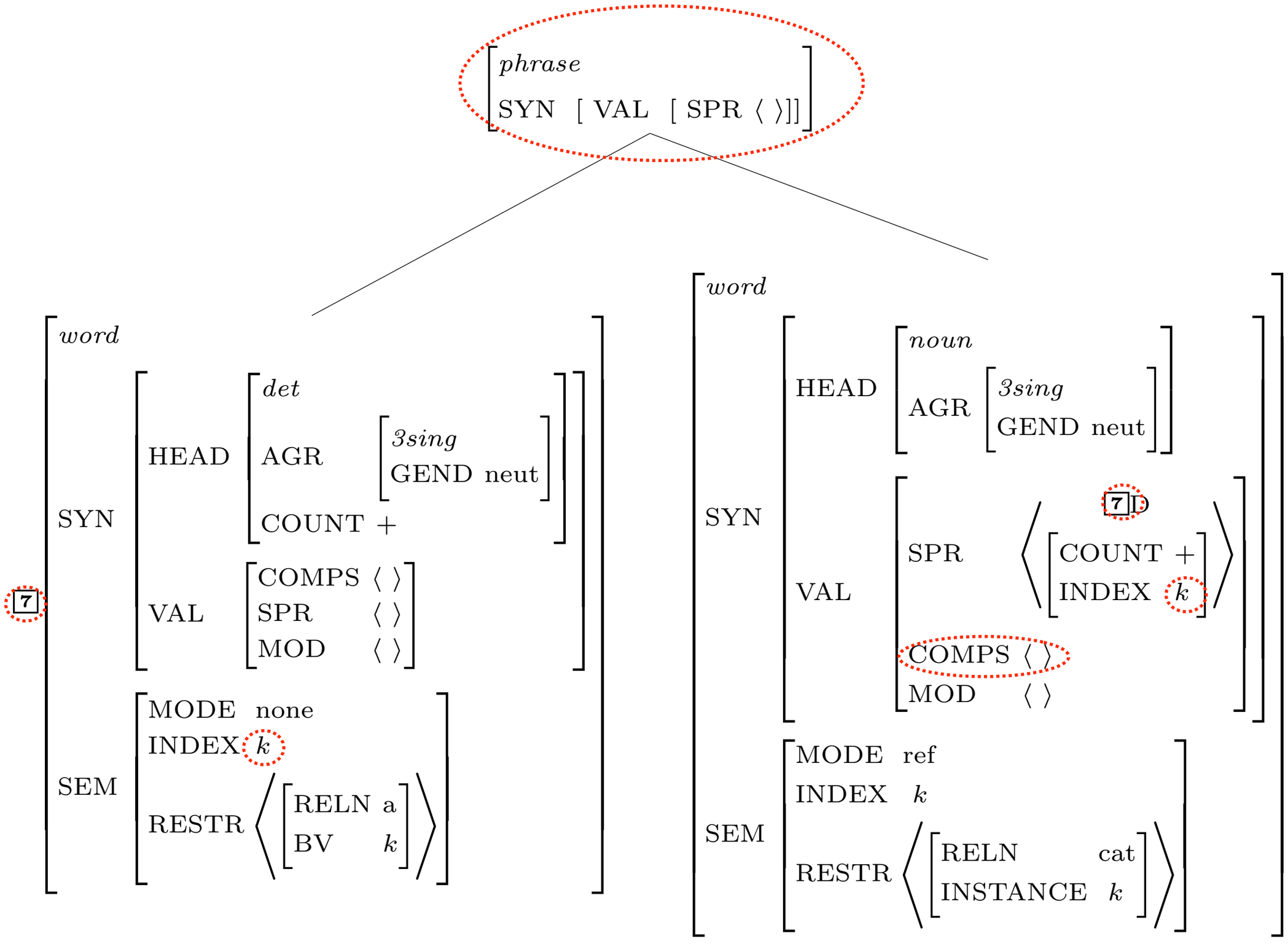
Building a Phrase



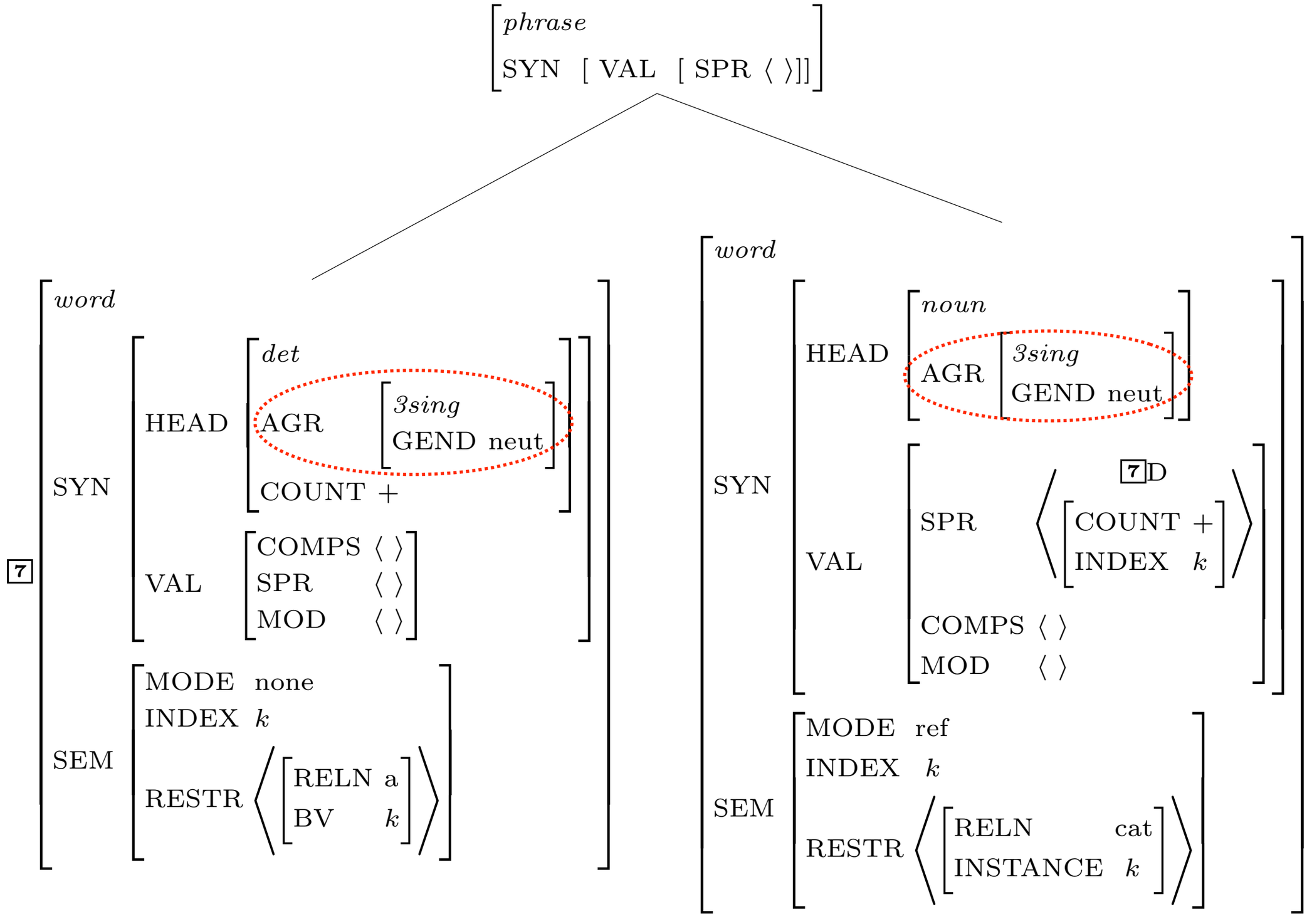
Constraints Contributed by Daughter Subtrees



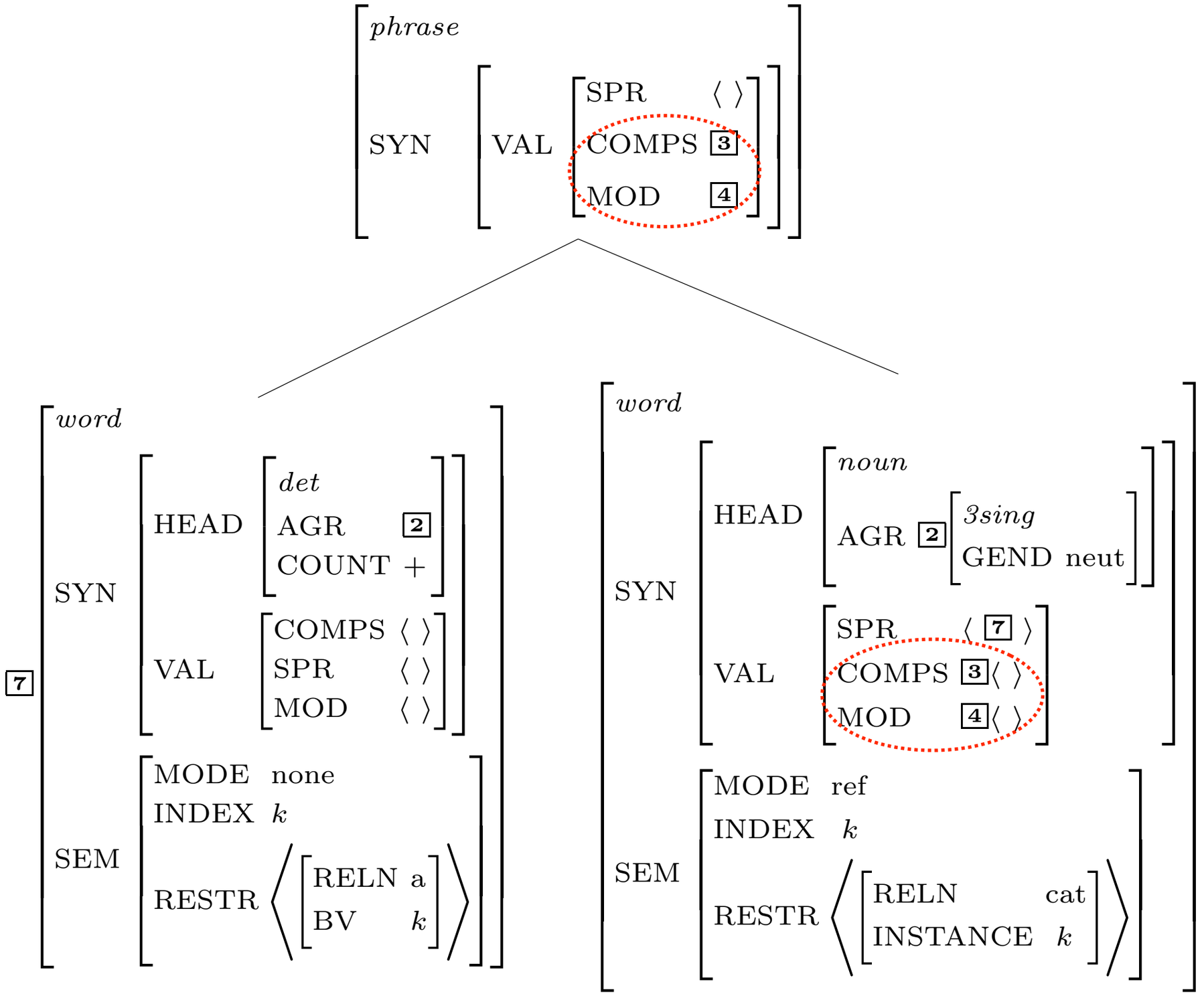
Constraints Contributed by the Grammar Rule



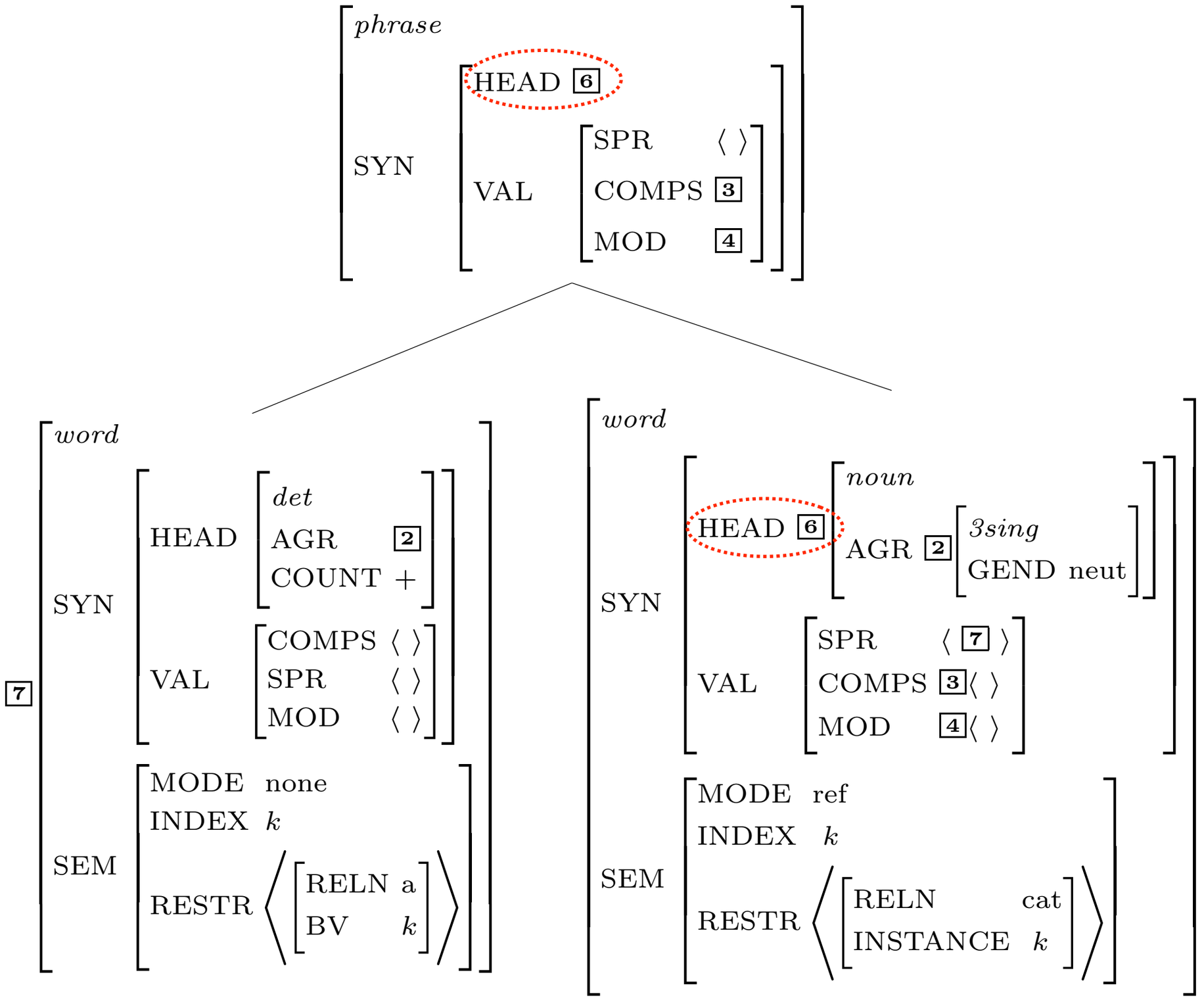
A Constraint Involving the SHAC



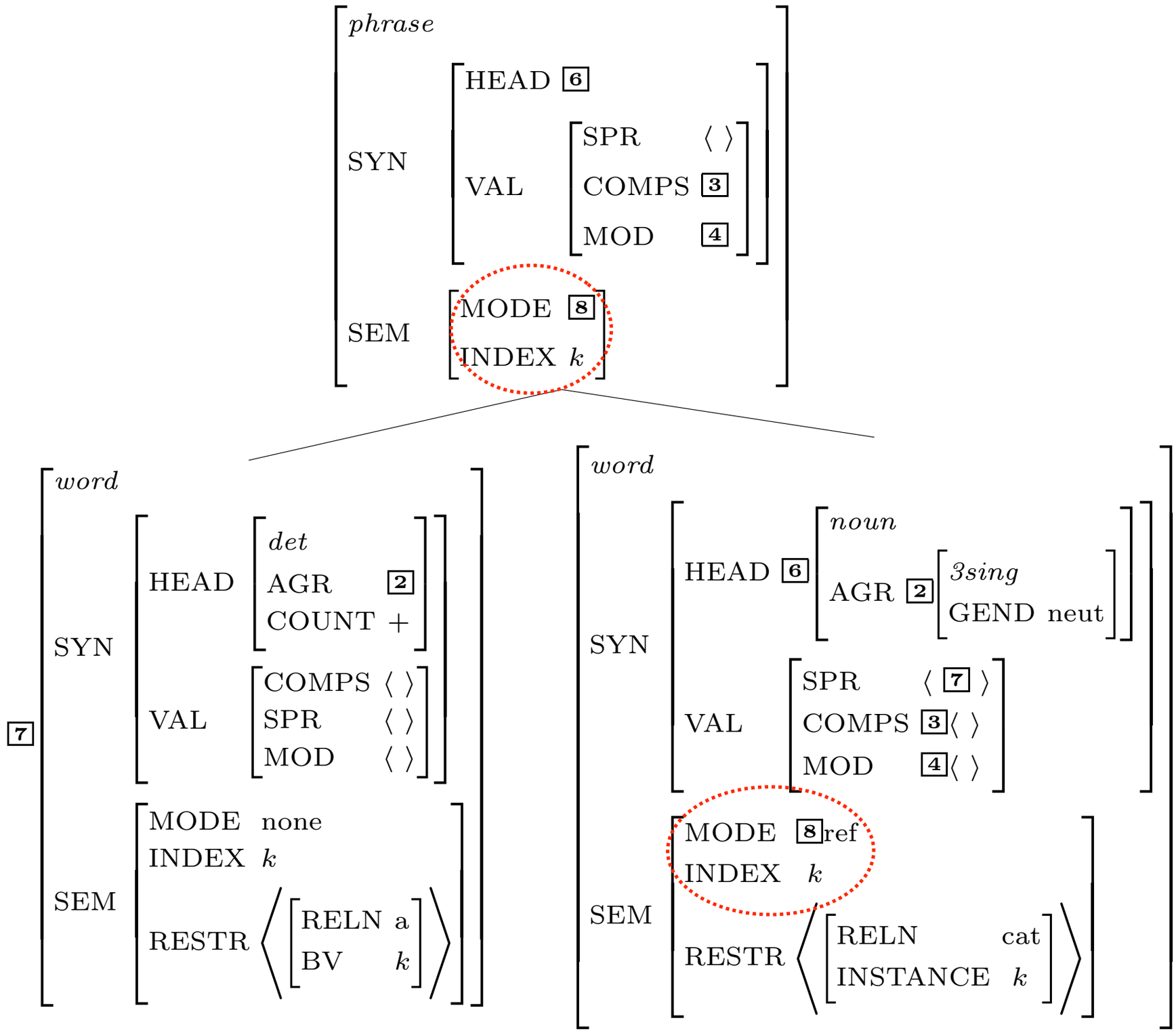
Effects of the Valence Principle



Effects of the Head Feature Principle



Effects of the Semantic Inheritance Principle



Effects of the Semantic Compositionality Principle

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

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

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

Is the Mother Node Now Completely Specified?

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

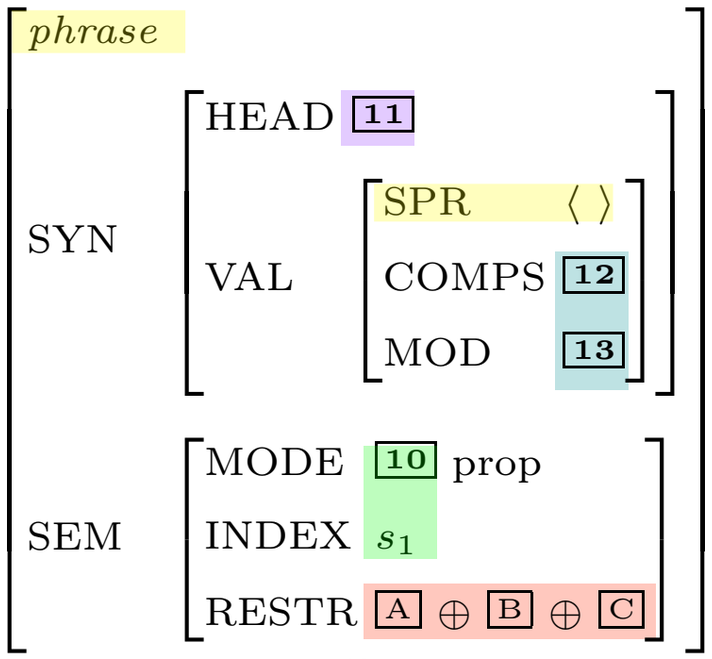
$$\boxed{7}\begin{bmatrix} \textit{word} & & \\ \text{SYN} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{det} & \\ \text{AGR} & \boxed{2} \\ \text{COUNT} & + \end{bmatrix} \\ \text{VAL} & \begin{bmatrix} \text{COMPS} & \langle\ \rangle \\ \text{SPR} & \langle\ \rangle \\ \text{MOD} & \langle\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

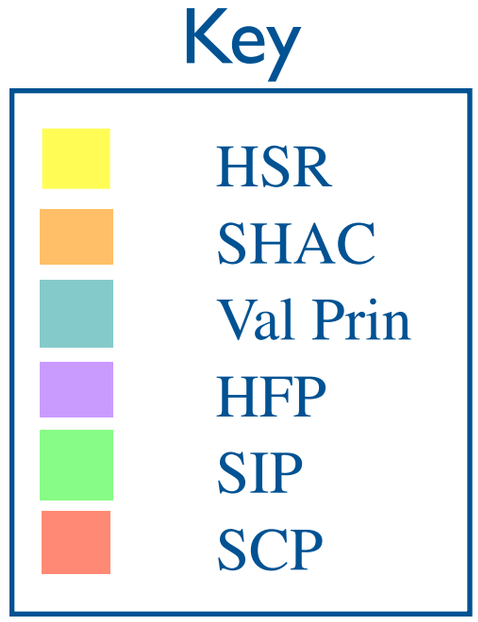
$$\begin{bmatrix} \textit{word} & & \\ \text{SYN} & \begin{bmatrix} \text{HEAD} & \boxed{6}\begin{bmatrix} \textit{noun} & \\ \text{AGR} & \boxed{2}\begin{bmatrix} \textit{3sing} & \\ \text{GEND} & \text{neut} \end{bmatrix} \end{bmatrix} \\ \text{VAL} & \begin{bmatrix} \text{SPR} & \langle\ \boxed{7}\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

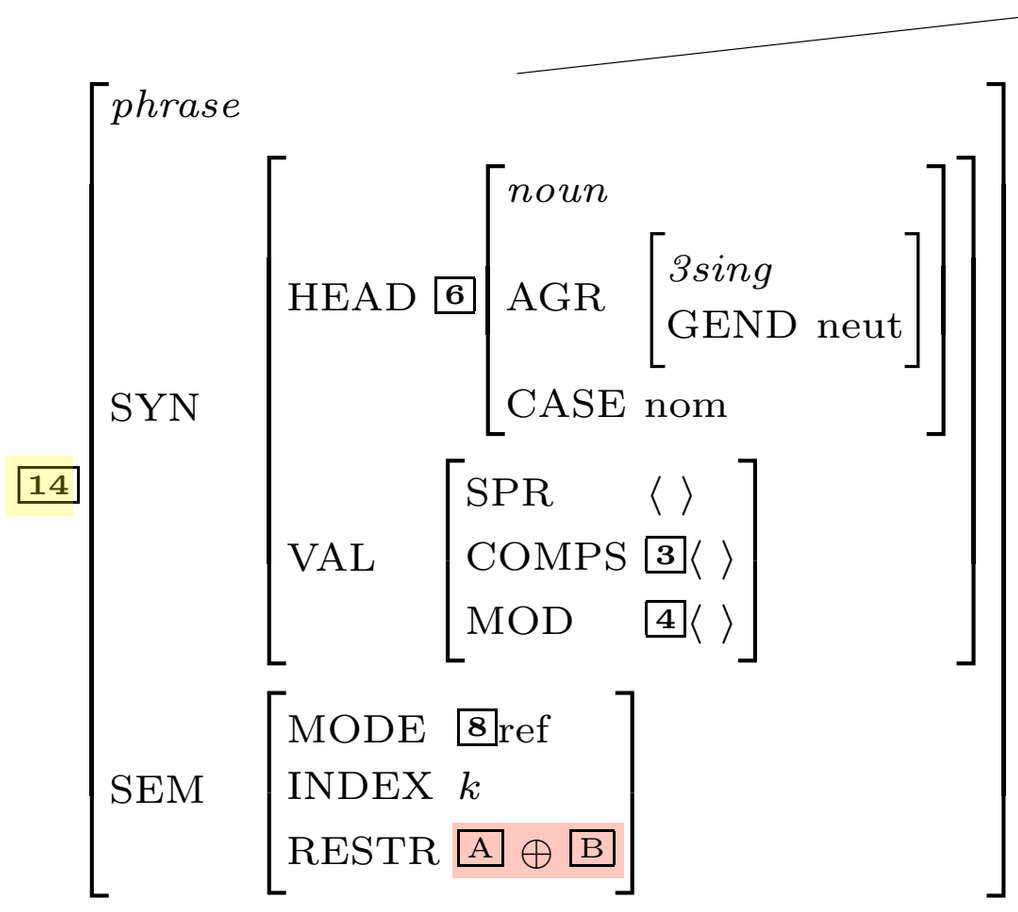
Lexical Entry for *slept*

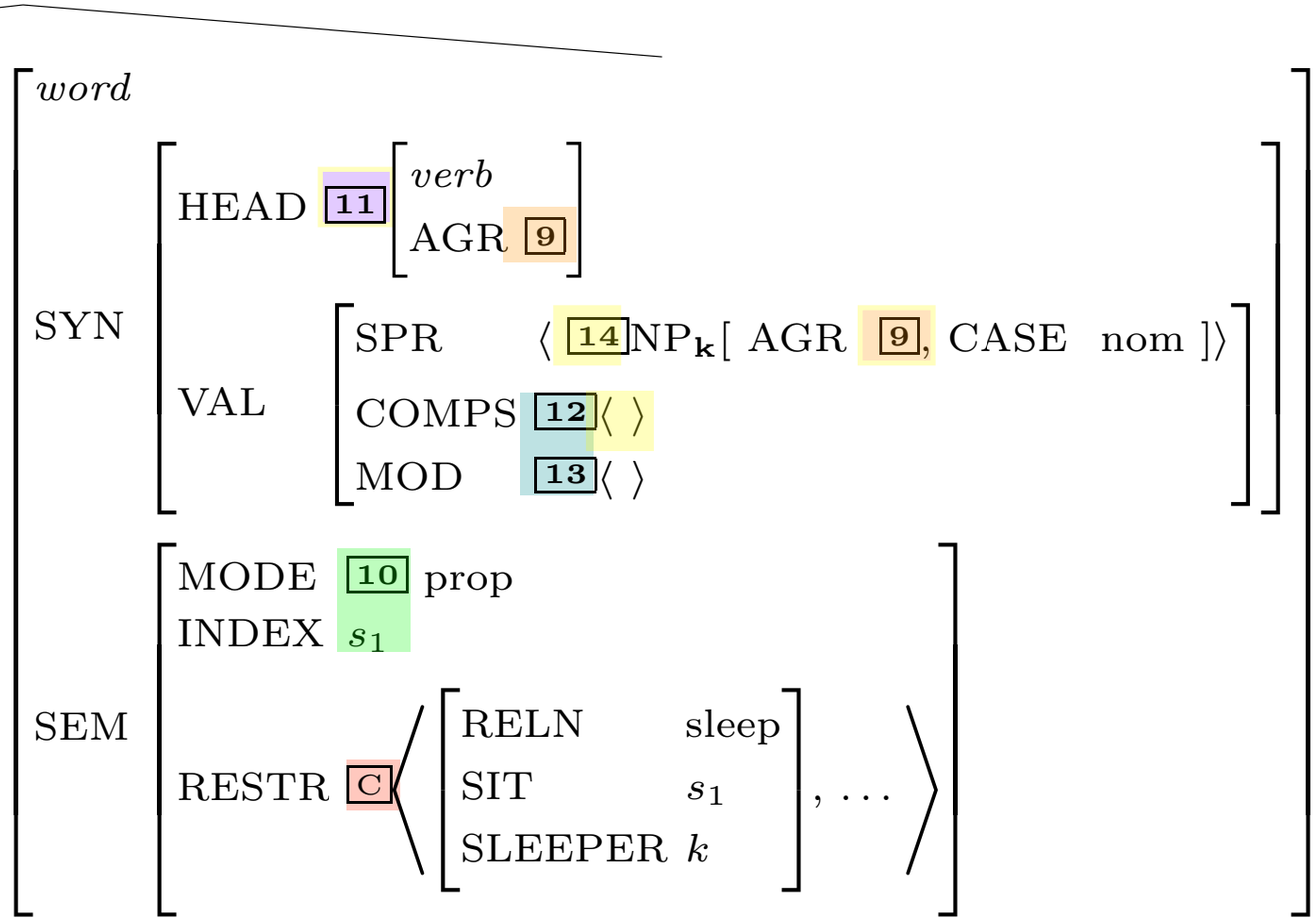
$$\left\langle \text{slept}, \left[\begin{array}{l} \textit{word} \\ \text{SYN} \left[\begin{array}{ll} \text{HEAD} & \textit{verb} \\ \text{VAL} & \left[\begin{array}{ll} \text{SPR} & \left\langle \left[\begin{array}{ll} \text{AGR} & \boxed{9} \\ \text{CASE} & \text{nom} \end{array}\right] \right\rangle_{\text{NP}_m}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

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?

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

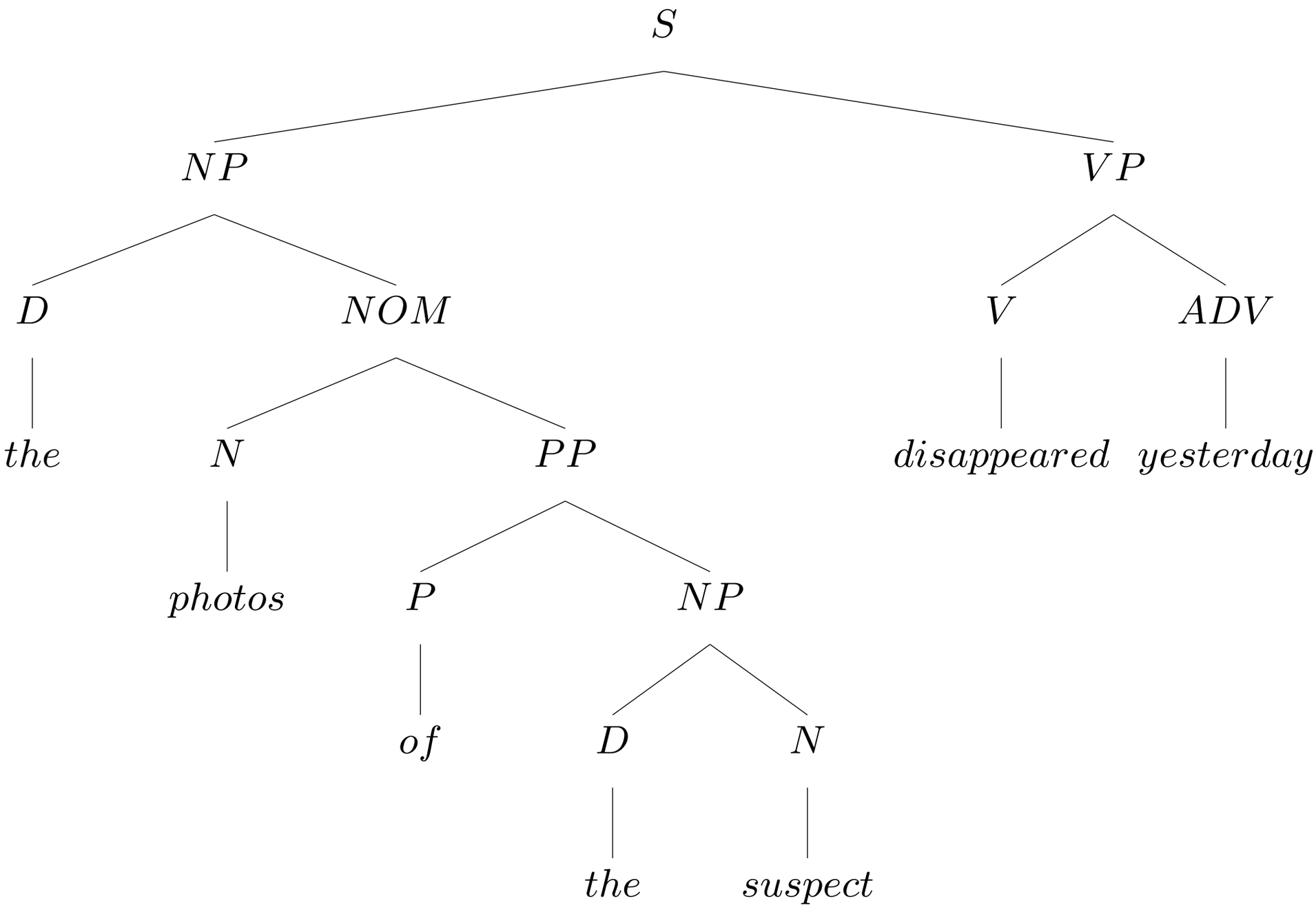
$$\boxed{14}\begin{bmatrix} \textit{phrase} & & \\ \text{SYN} & \text{HEAD } \boxed{6} & \begin{bmatrix} \textit{noun} & \\ \text{AGR} & \begin{bmatrix} \textit{3sing} \\ \text{GEND}\ \text{neut} \end{bmatrix} \\ \text{CASE}\ \text{nom} & \end{bmatrix} \\ & \text{VAL} & \begin{bmatrix} \text{SPR} & \langle\,\rangle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

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

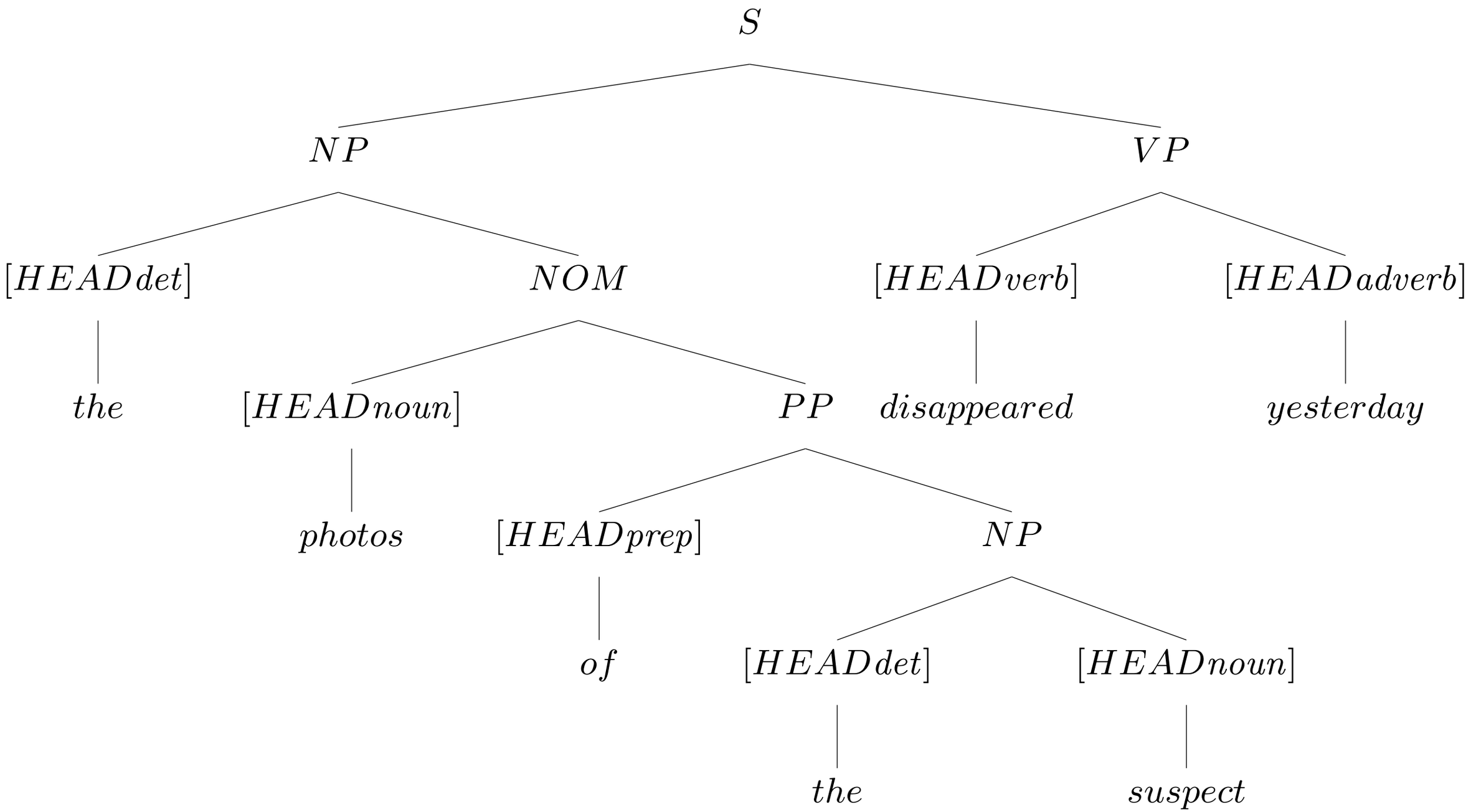
RESTR of the S node

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

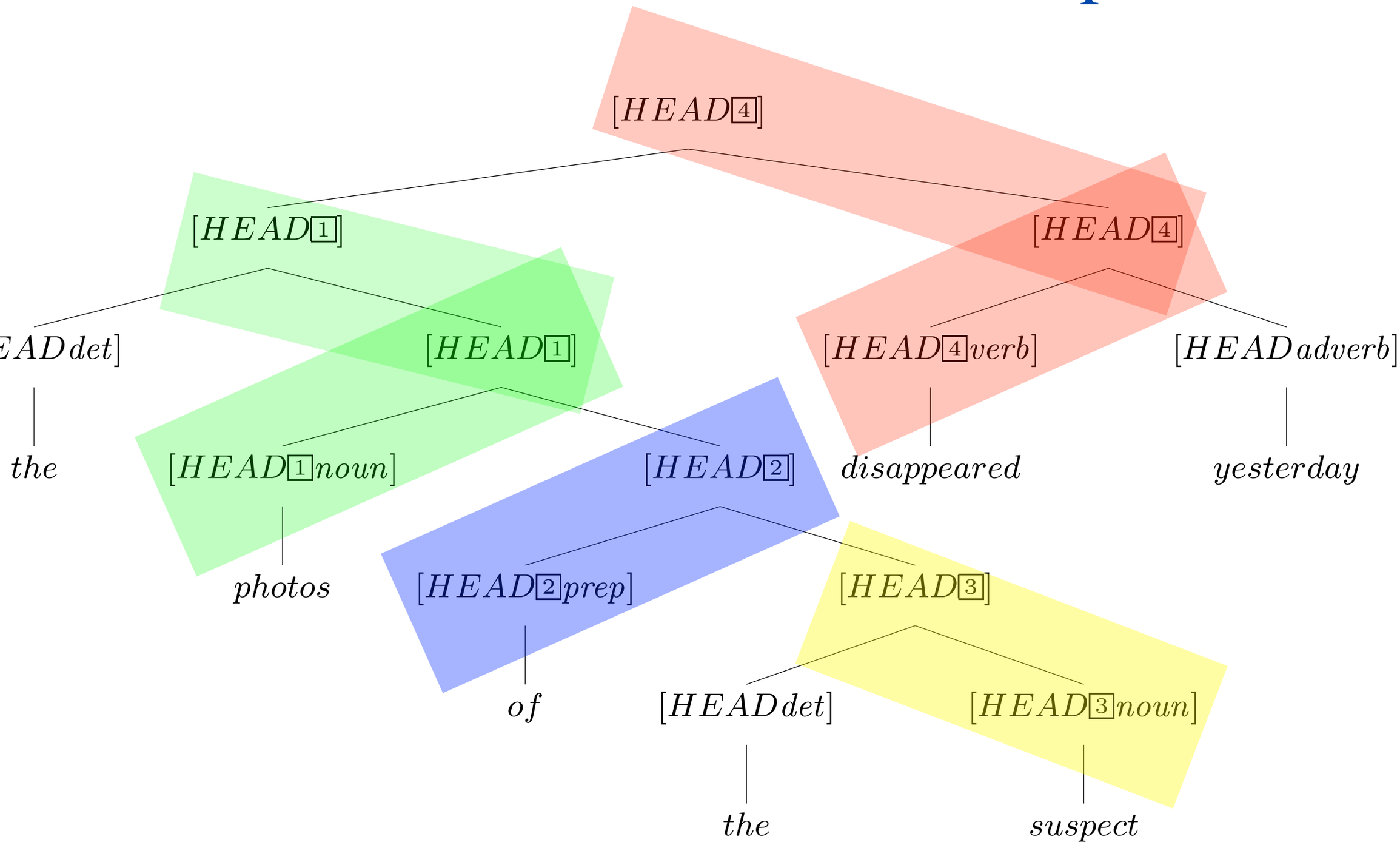
Another Example



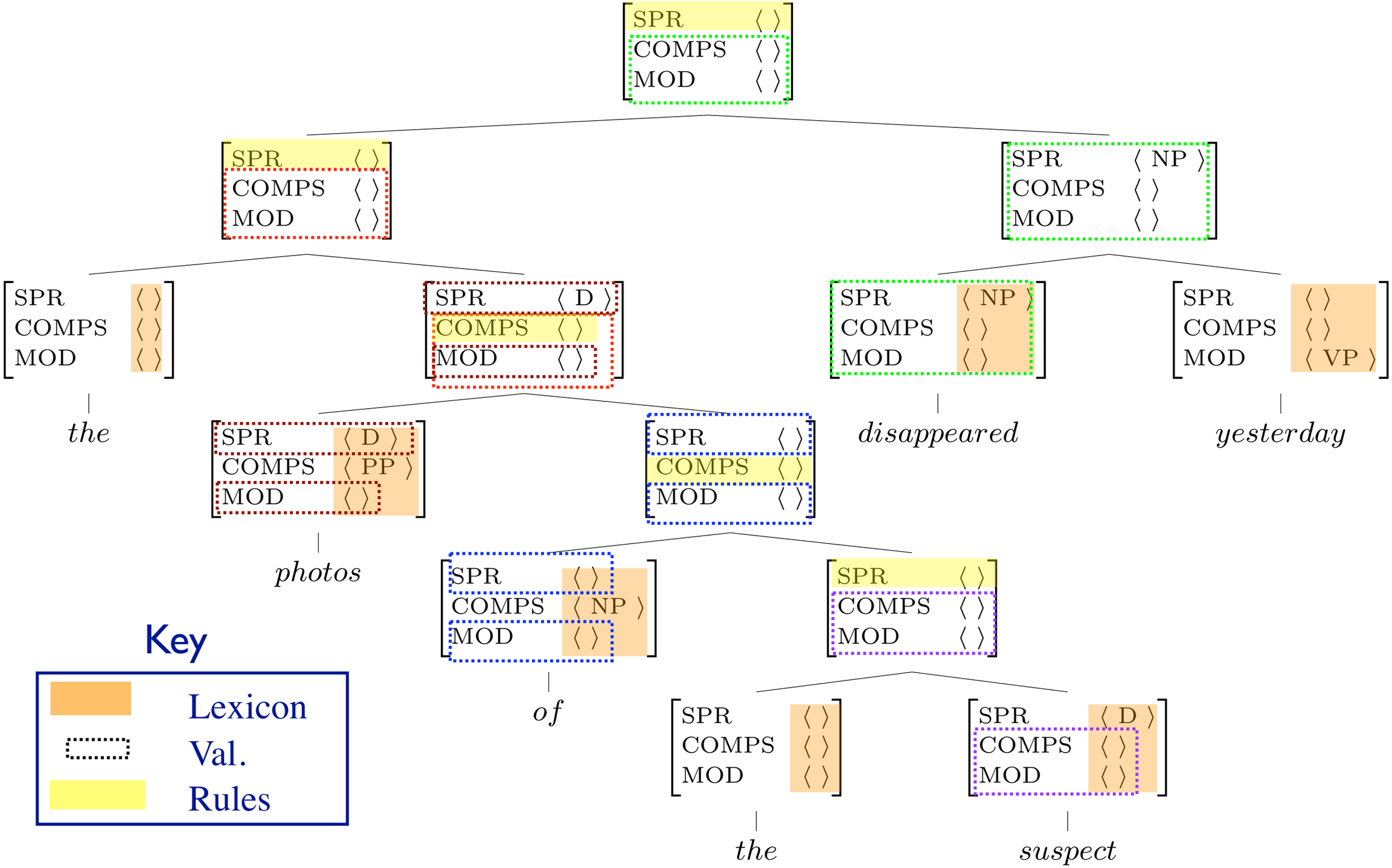
Head Features from Lexical Entries



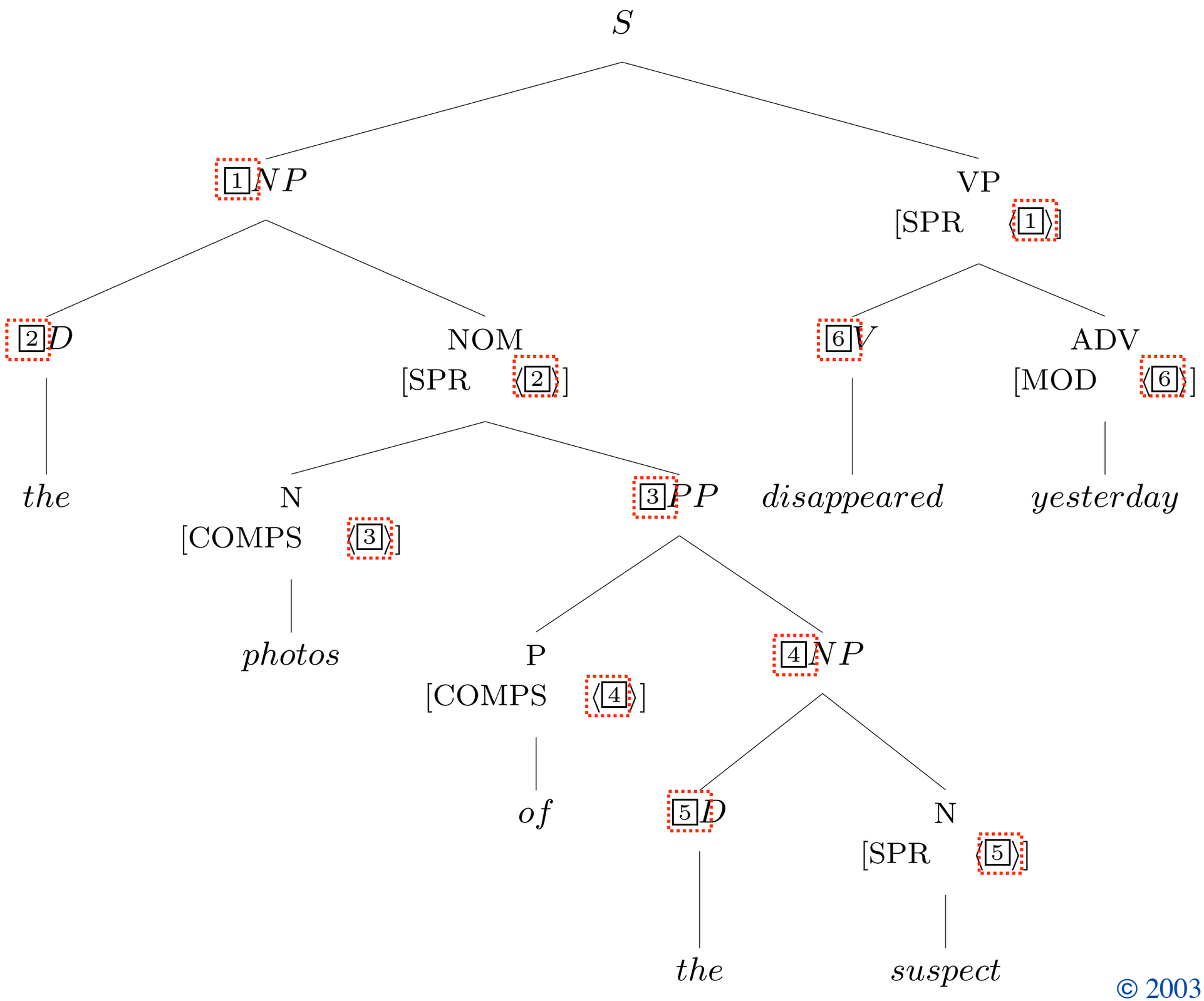
Head Features from Lexical Entries, plus HFP



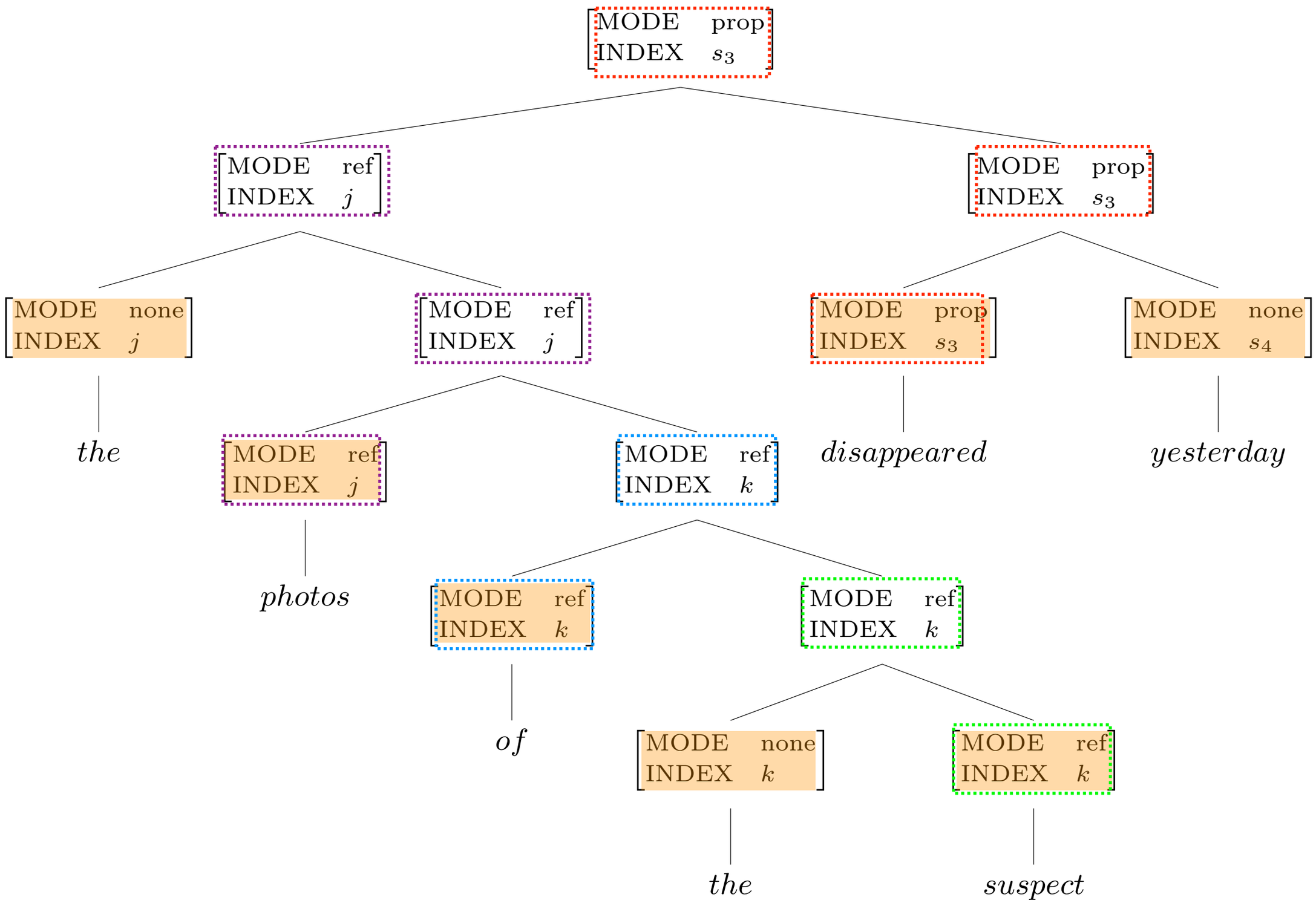
Valence Features: Lexicon, Rules, and the Valence Principle



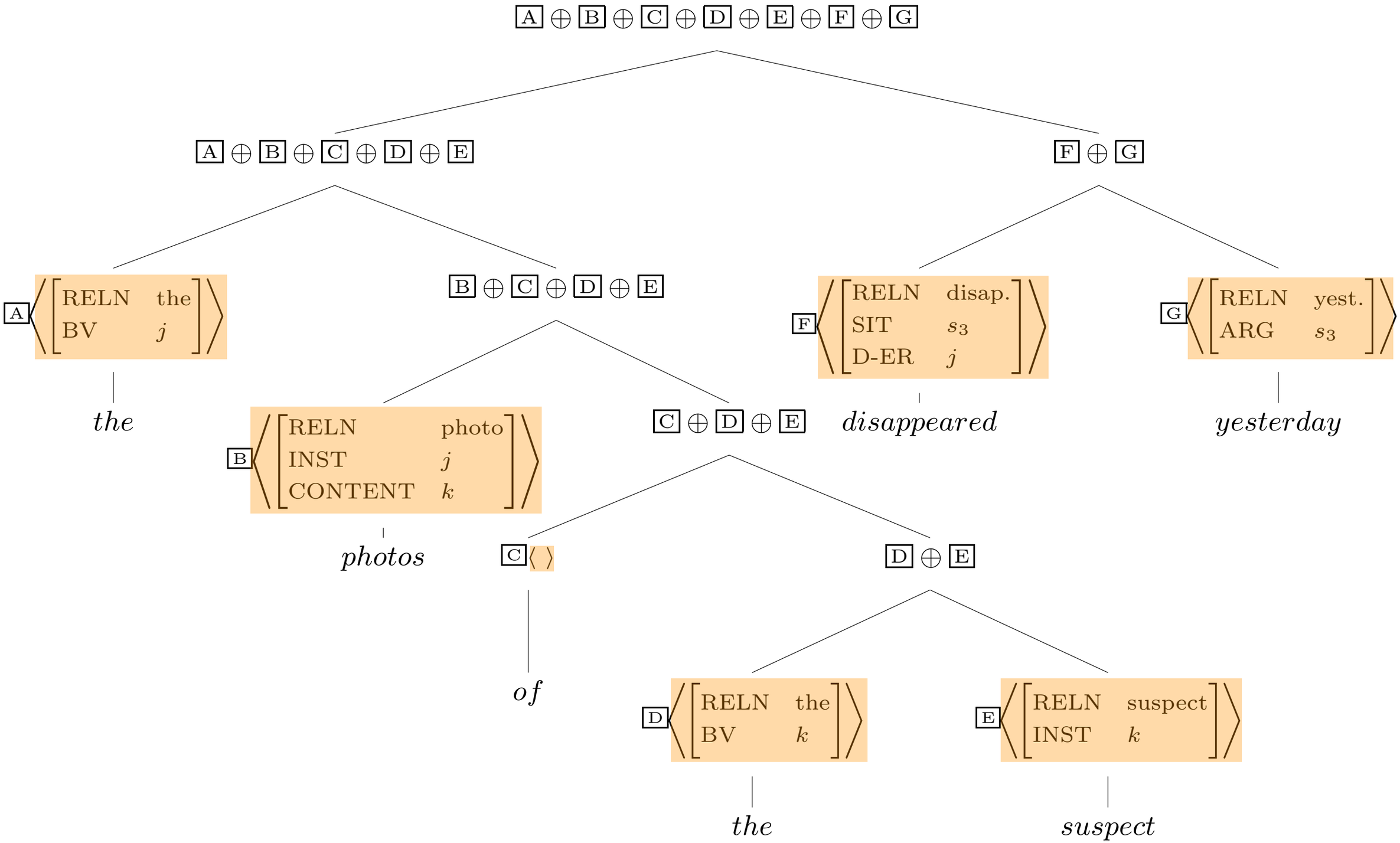
Required Identities: Grammar Rules



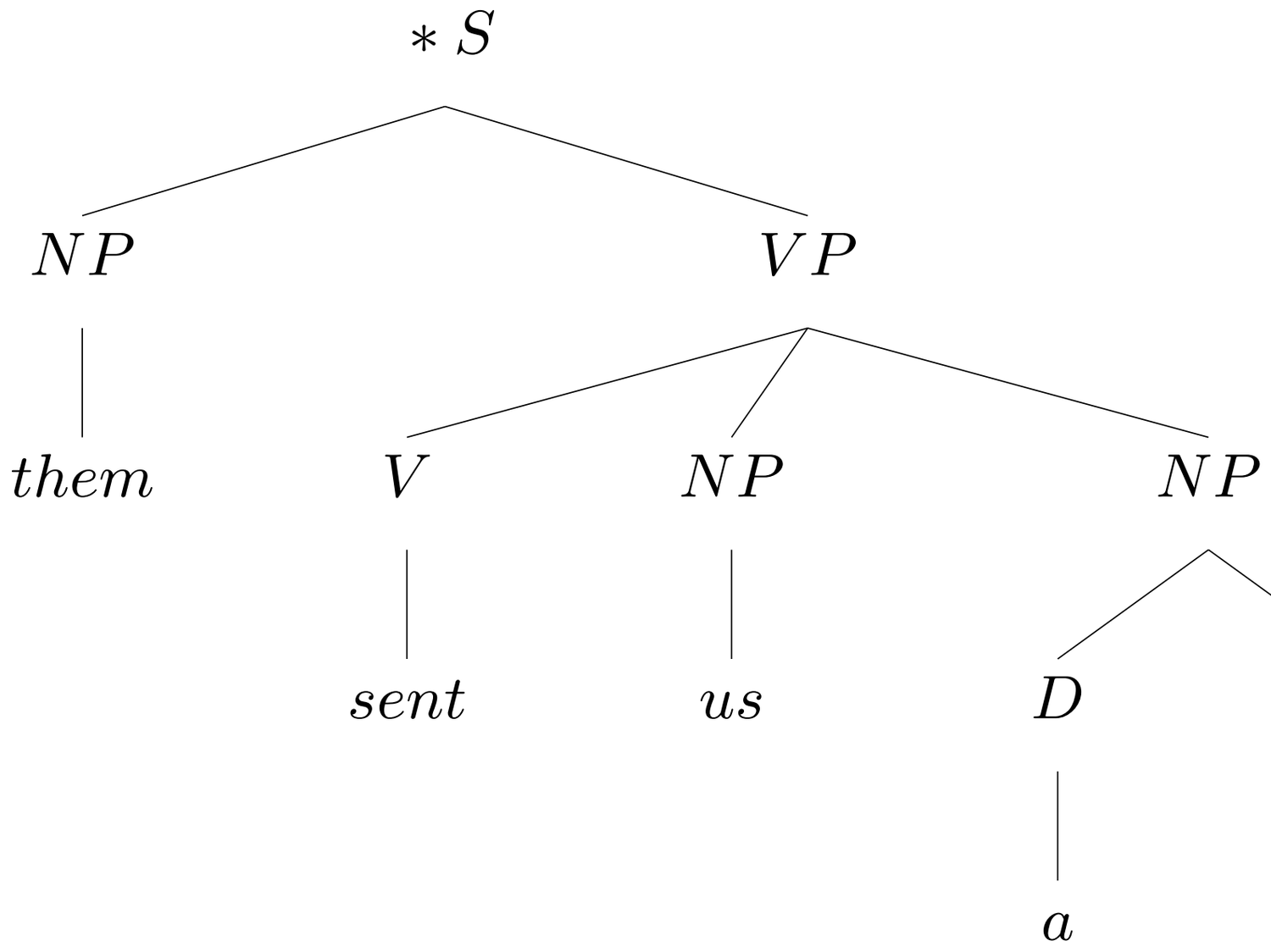
Two Semantic Features: the Lexicon & SIP



RESTR Values and the SCP

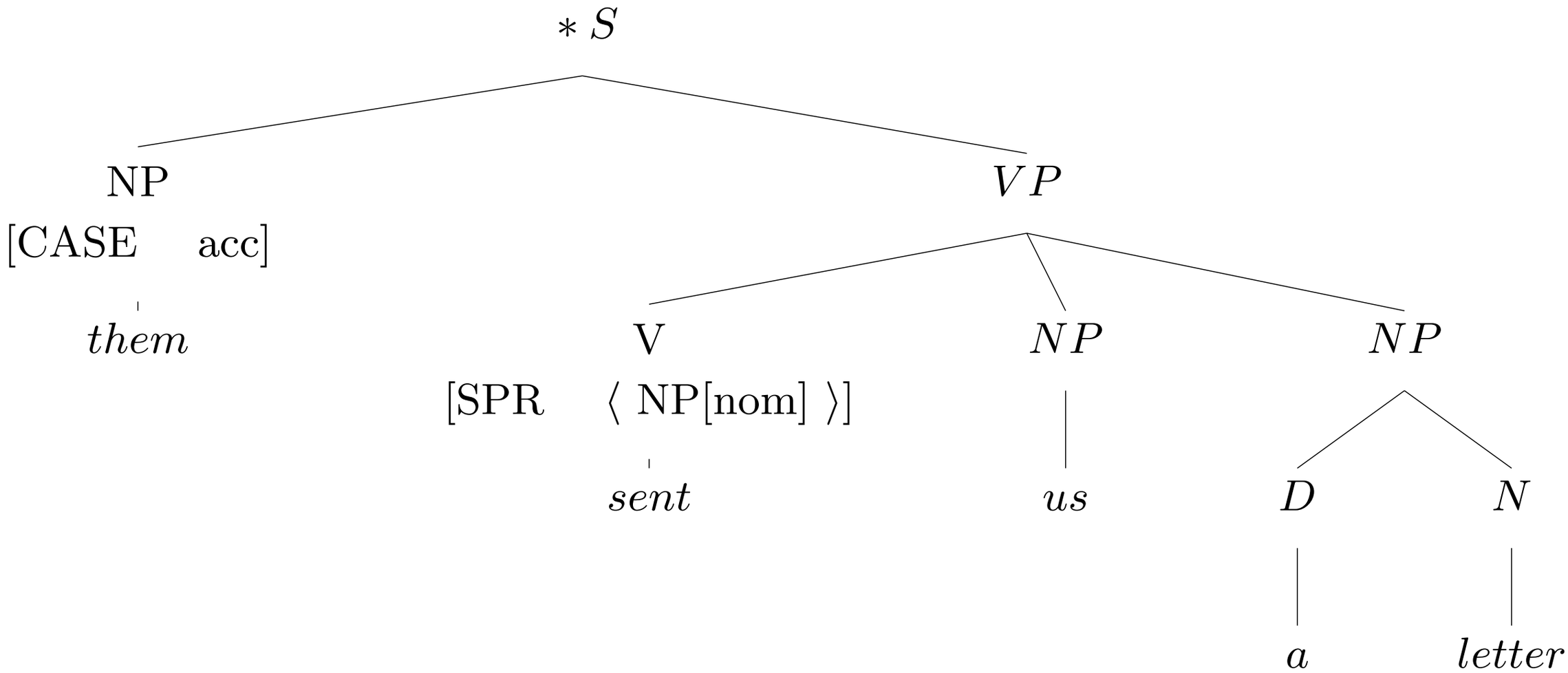


An Ungrammatical Example



What's wrong with this sentence?

An Ungrammatical Example

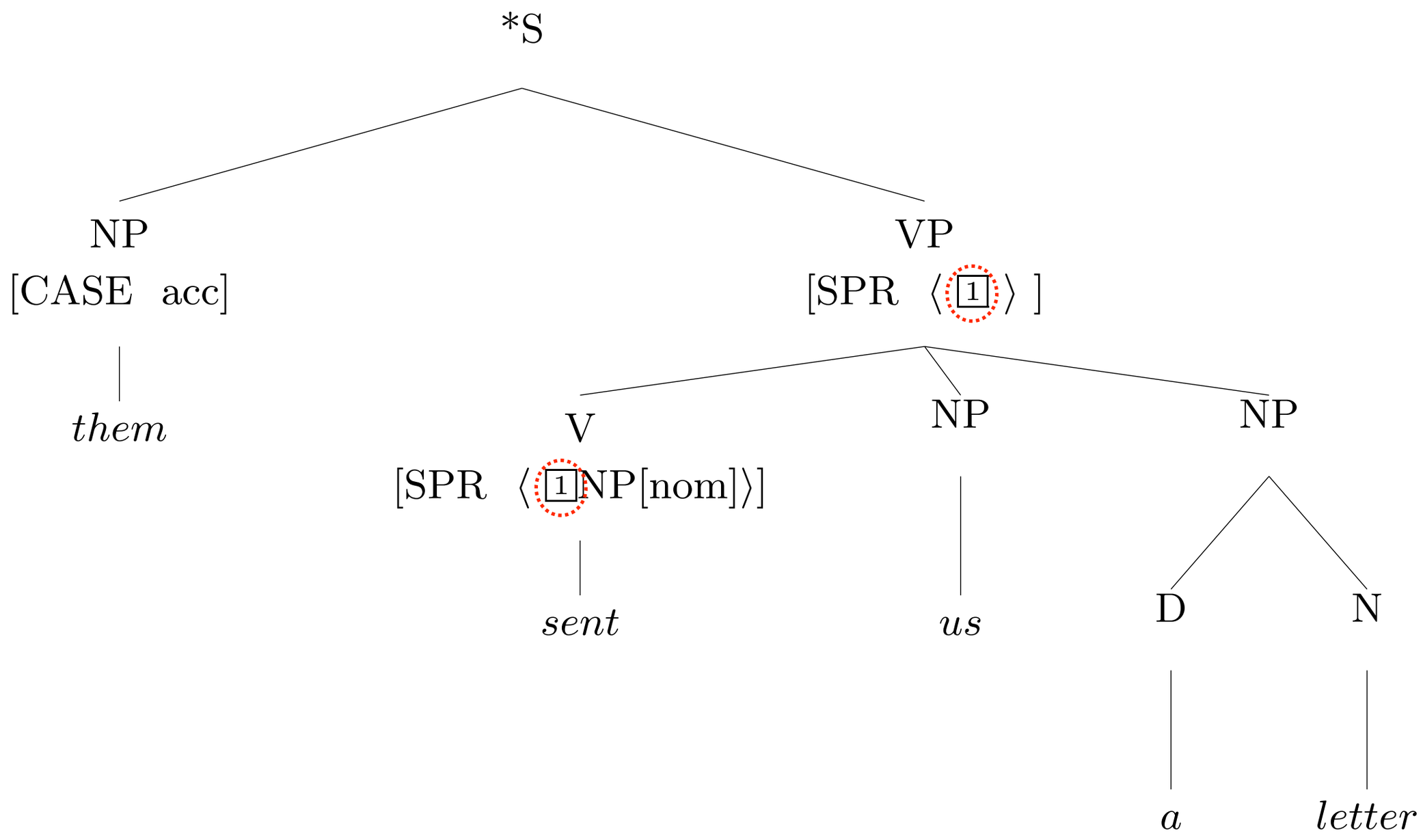


What's wrong with this sentence?

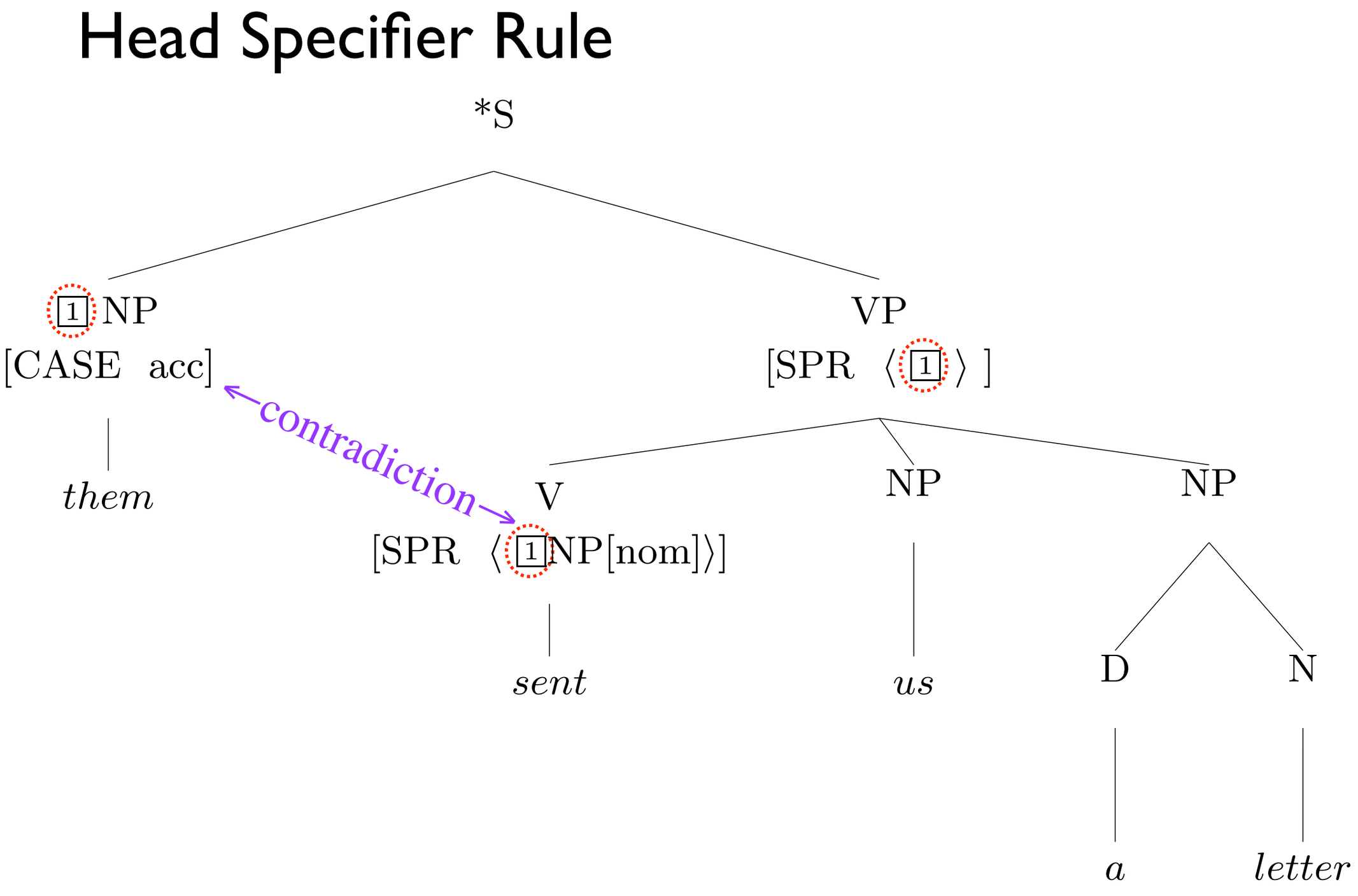
So what?

An Ungrammatical Example

The Valence Principle



An Ungrammatical Example



Exercise in Critical Thinking

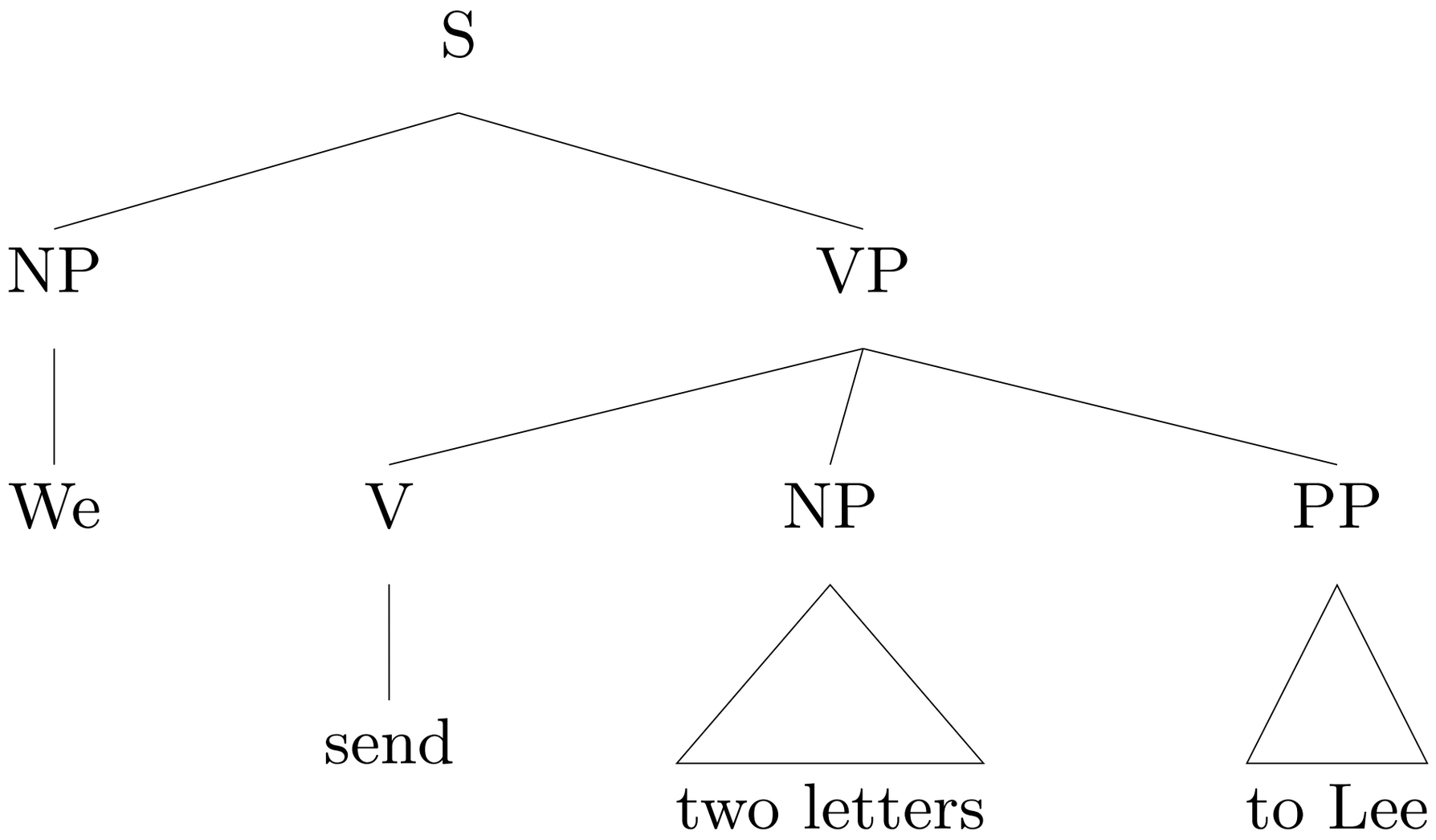
- Our grammar has come a long way since Ch 2, as we've added ways of representing different kinds of information:
  - generalizations across categories
  - semantics
  - particular linguistic phenomena: valence, agreement, modification
- What else might we add? What facts about language are as yet unrepresented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

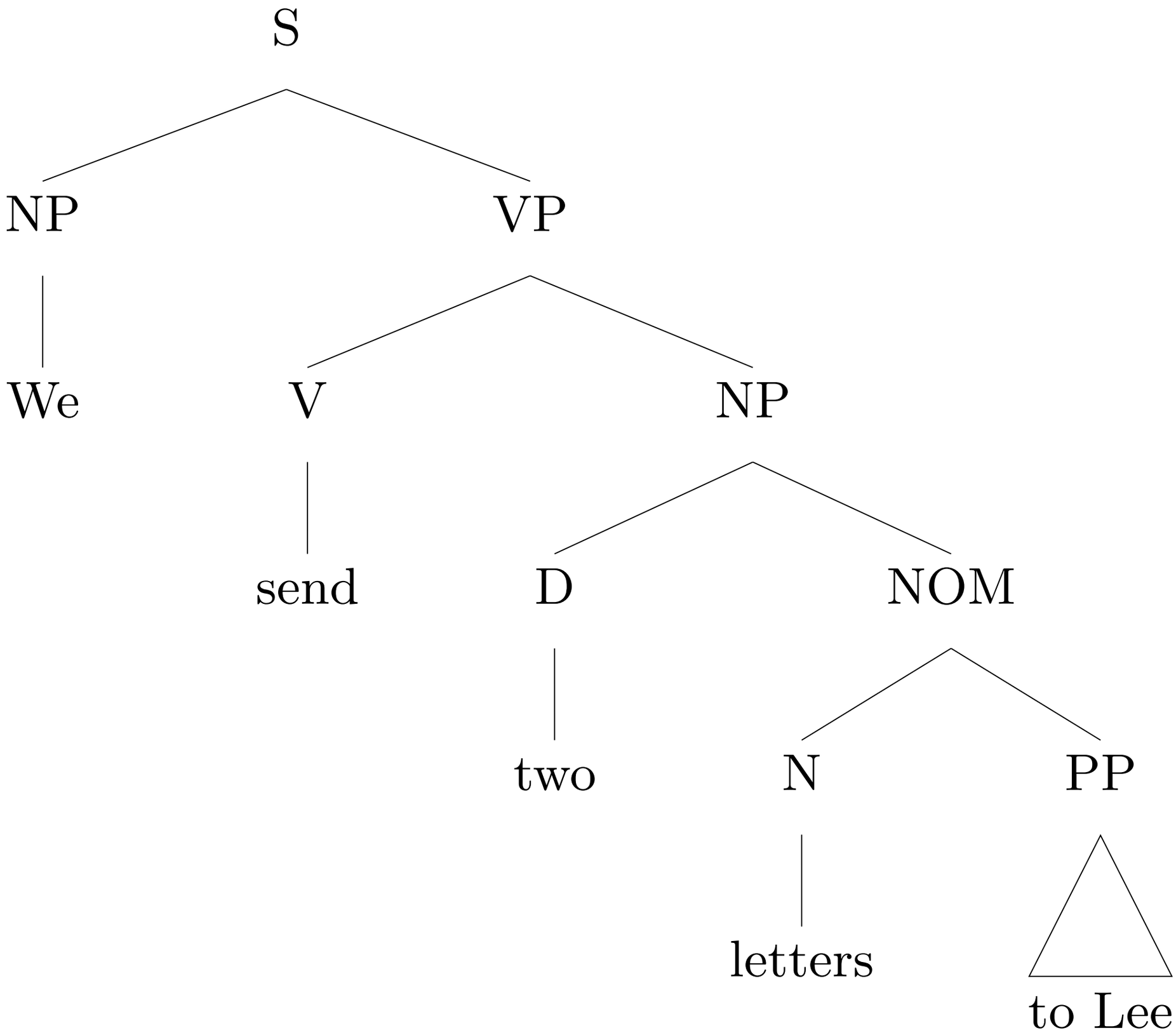
Overview

- 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
- Next time: Catch up & review

Reading Questions

- In what way does the actual meaning of the two structures assigned to this sentence differ?
  - *We sent two letters to Lee*.
- Are they really both grammatical?





$$\begin{bmatrix}\text{RELN} & \textbf{group} \\ \text{INST} & i\end{bmatrix}, \begin{bmatrix}\text{RELN} & \textbf{speaker} \\ \text{INST} & l\end{bmatrix}, \begin{bmatrix}\text{RELN} & \textbf{member} \\ \text{SET} & i \\ \text{ELEMENT} & l\end{bmatrix},$$

$$\begin{bmatrix}\text{RELN} & \textbf{send} \\ \text{SIT} & s_7 \\ \text{SENDER} & i \\ \\ \text{SENDEE} & j \\ \text{SENT} & k\end{bmatrix}, \begin{bmatrix}\text{RELN} & \textbf{two} \\ \text{BV} & k\end{bmatrix}, \begin{bmatrix}\text{RELN} & \textbf{letter} \\ \text{INST} & k \\ \text{ADDRESSEE} & m\end{bmatrix},$$

$$\begin{bmatrix}\text{RELN} & \textbf{name} \\ \text{NAME} & \text{Lee} \\ \text{NAMED} & j\end{bmatrix}$$

$$\begin{bmatrix} \text{RELN} & \textbf{group} \\ \text{INST} & i \end{bmatrix}, \begin{bmatrix} \text{RELN} & \textbf{speaker} \\ \text{INST} & l \end{bmatrix}, \begin{bmatrix} \text{RELN} & \textbf{member} \\ \text{SET} & i \\ \text{ELEMENT} & l \end{bmatrix},$$

$$\begin{bmatrix} \text{RELN} & \textbf{send} \\ \text{SIT} & s_7 \\ \text{SENDER} & i \\ \text{SENDEE} & j \\ \text{SENT} & k \end{bmatrix}, \begin{bmatrix} \text{RELN} & \textbf{two} \\ \text{BV} & k \end{bmatrix}, \begin{bmatrix} \text{RELN} & \textbf{letter} \\ \text{INST} & k \\ \text{ADDRESSEE} & m \end{bmatrix},$$

$$\begin{bmatrix} \text{RELN} & \textbf{name} \\ \text{NAME} & \text{Lee} \\ \text{NAMED} & m \end{bmatrix}$$

Reading Questions

- How do we know what features to put into a *predication*?
- Would *letters* as in letters of the alphabet have the same lexical entry as *letters* like what's usually sent in the mail?
- How do we represent the difference in meaning between *send* and *sent*?

Reading Questions

- How do we get enough different INDEX values for a whole dictionary?
- Why sometimes *s* and sometimes s_n, and not *t*, *u*, *v*?
- How can *to* be semantically empty and still have a meaningful INDEX value?
- How can the head of a phrase be semantically empty?
- Why does *letter* share its INDEX with it's SPR?

Reading Questions

- Does set of well-formed structures correspond exactly to the set of well-formed English sentences?
- Do we have to understand the squiggly bits?
- Why bother formalizing?
- Don’t these feature structures get ridiculously large?

Reading Questions

- Does English have dative case?
- Is it redundant to have a feature CASE for English given that we mostly use prepositions to mark 'case'?
- English nouns (other than pronouns) are underspecified for CASE. How do we figure out their particular CASE values when they are used in a tree?

Reading Questions

- Is position alone enough to tell whether something is SPR or COMPS?
- Will this approach work for morphologically complex languages as well?
- What ever happened to NOM?
- Is it worth memorizing the rules now?
- Why aren't we using NumP?

Reading Questions

- Is top-down or bottom-up more efficient in actual processing?
- How can we possibly do "simultaneous satisfaction" of all constraints?
- What are the best practices for writing trees going bottom-up (order of things to put in)?

Reading Questions

- Does not having to realize semantic roles mean we can license semantically weird sentences?
- Can we build a grammar that works with more than one sentence at a time? (I.e., paragraphs)