

# Ling 566

## Oct 15, 2020

### Semantics

# Overview

- Some notes on the linguist's stance
- Which aspects of semantics we'll tackle
- Our formalization; Semantics Principles
- Building semantics of phrases
- Modification, coordination
- Structural ambiguity
- Reading questions

## The Linguist's Stance: Building a precise model

- Some of our statements are statements about how the model works:

“*[prep]* and *[AGR 3sing]* can't be combined because *AGR* is not a feature of the type *prep*.”
- Some of our statements are statements about how (we think) English or language in general works.

“The determiners *a* and *many* only occur with count nouns, the determiner *much* only occurs with mass nouns, and the determiner *the* occurs with either.”
- Some are statements about how we code a particular linguistic fact within the model.

“All count nouns are *[SPR < [COUNT +]>]*.”

# Semantics: Where's the Beef?

So far, our grammar has no semantic representations. We have, however, been relying on semantic intuitions in our argumentation, and discussing semantic contrasts where they line up (or don't) with syntactic ones.

Examples?

- structural ambiguity
- S/NP parallelism
- count/mass distinction
- complements vs. modifiers

Poll

# Our Slice of a World of Meanings

Aspects of meaning we won't account for

- Pragmatics
- Fine-grained lexical semantics:

The meaning of *life* is *life*', or, in our case,

$$\begin{bmatrix} \text{RELN} & \text{life} \\ \text{INST} & i \end{bmatrix}$$

# Our Slice of a World of Meanings

MODE	prop
INDEX	$s$
RESTR	$\left\langle \begin{bmatrix} \text{RELN} & \text{save} \\ \text{SIT} & s \\ \text{SAVER} & i \\ \text{SAVED} & j \end{bmatrix}, \begin{bmatrix} \text{RELN} & \text{name} \\ \text{NAME} & \text{Chris} \\ \text{NAMED} & i \end{bmatrix}, \begin{bmatrix} \text{RELN} & \text{name} \\ \text{NAME} & \text{Pat} \\ \text{NAMED} & j \end{bmatrix} \right\rangle$

“... the linguistic meaning of *Chris saved Pat* is a proposition that will be true just in case there is an actual situation that involves the saving of someone named Pat by someone named Chris.” (p. 140)

# Our Slice of a World of Meanings

What we are accounting for is the **compositionality** of sentence meaning.

- How the pieces fit together

**Semantic arguments and indices**

- How the meanings of the parts add up to the meaning of the whole.

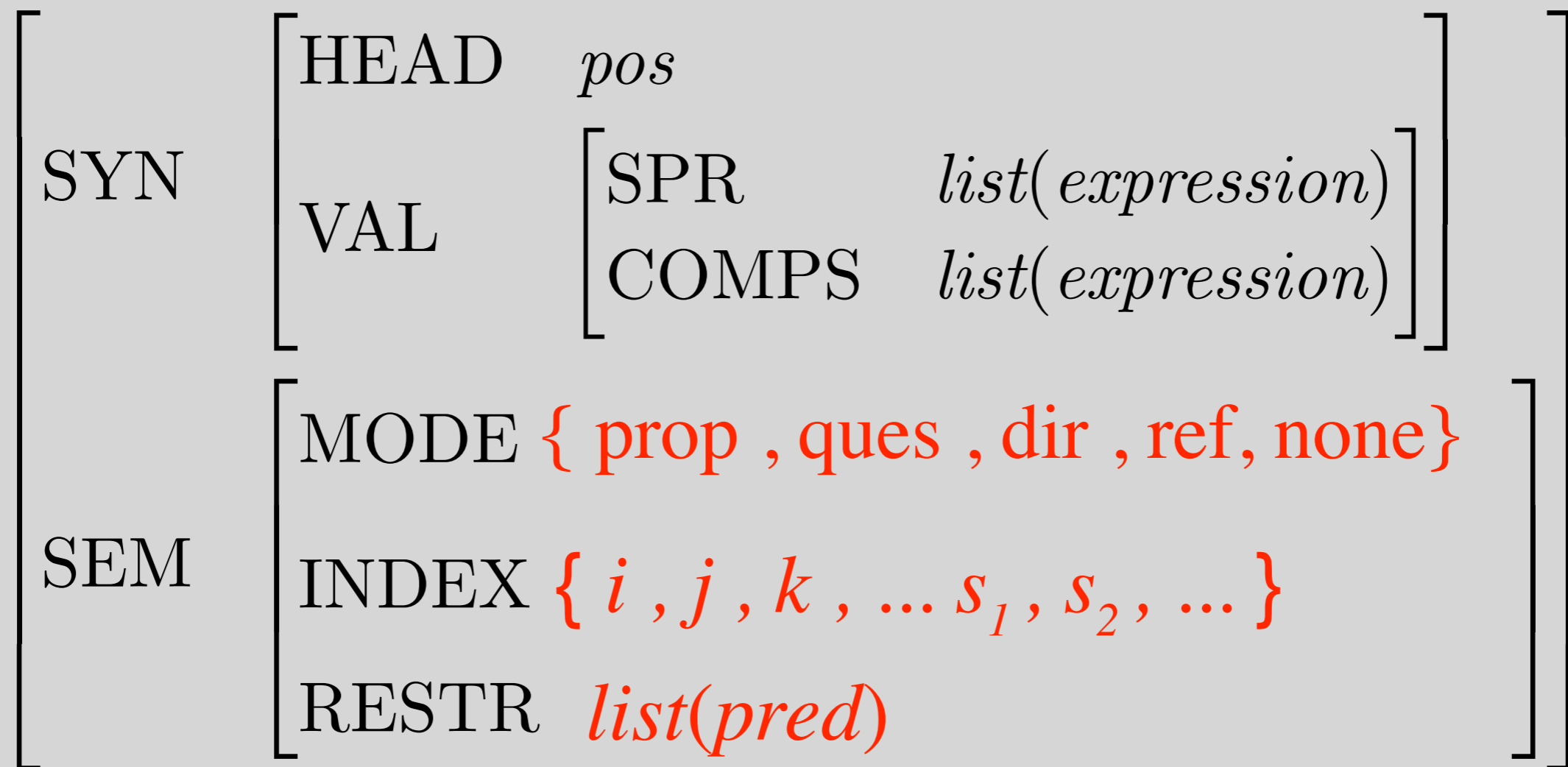
**Appending RESTR lists up the tree**

# Semantics in Constraint-Based Grammar

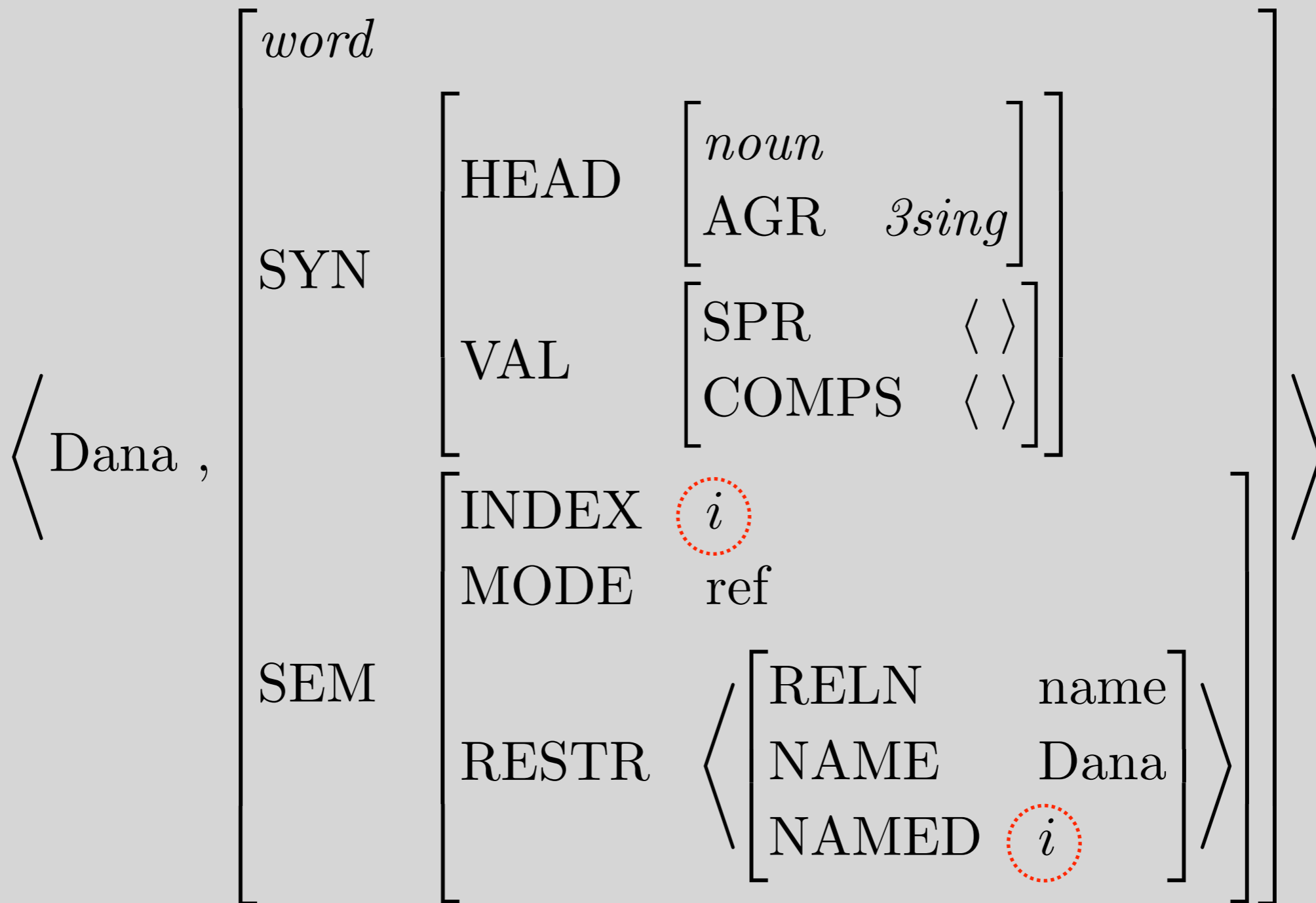
- Constraints as (generalized) truth conditions
  - proposition: what must be the case for a proposition to be true
  - directive: what must happen for a directive to be fulfilled
  - question: the kind of situation the asker is asking about
  - reference: the kind of entity the speaker is referring to
- Syntax/semantics interface: Constraints on how syntactic arguments are related to semantic ones, and on how semantic information is compiled from different parts of the sentence.



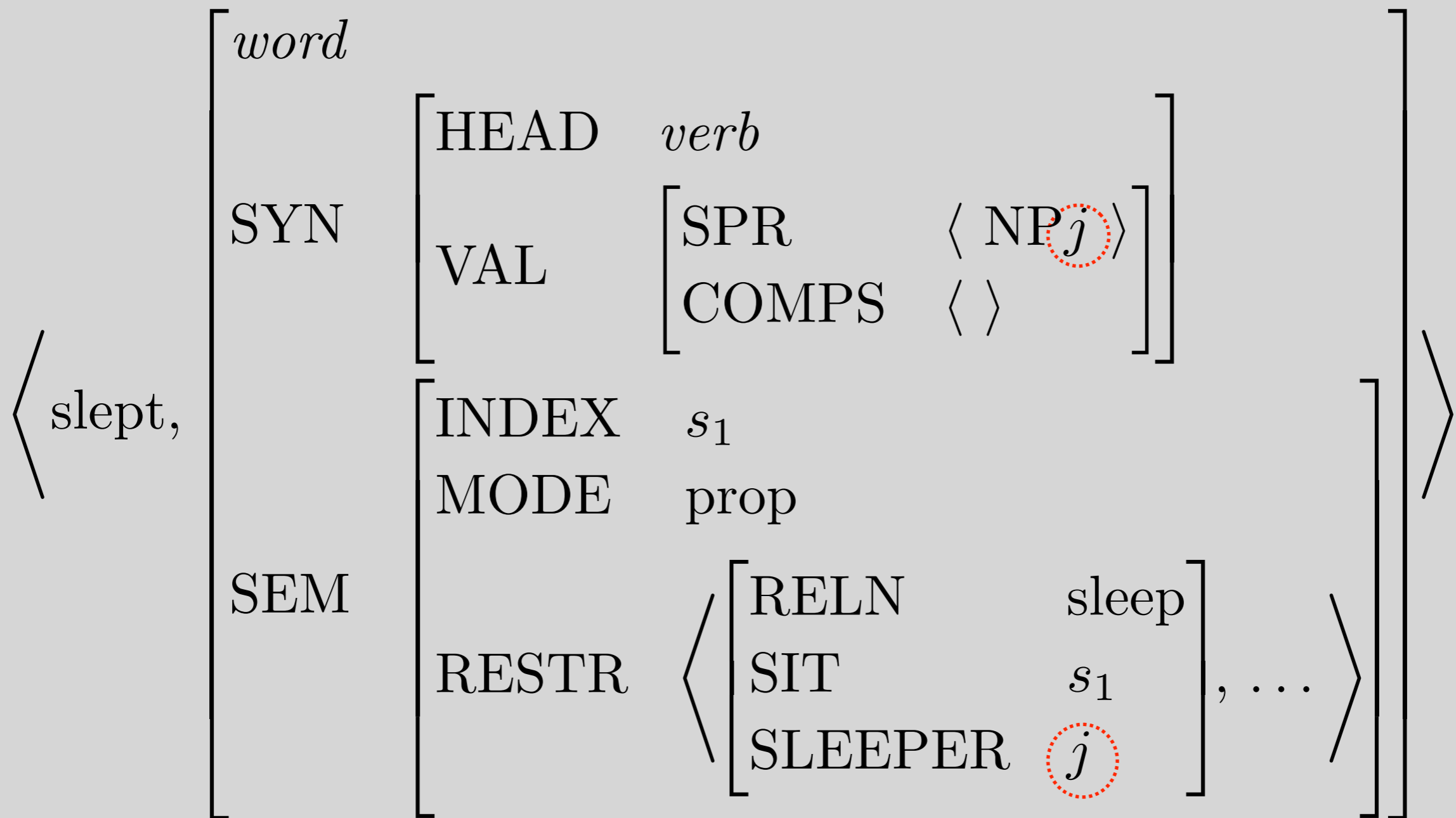
# Feature Geometry



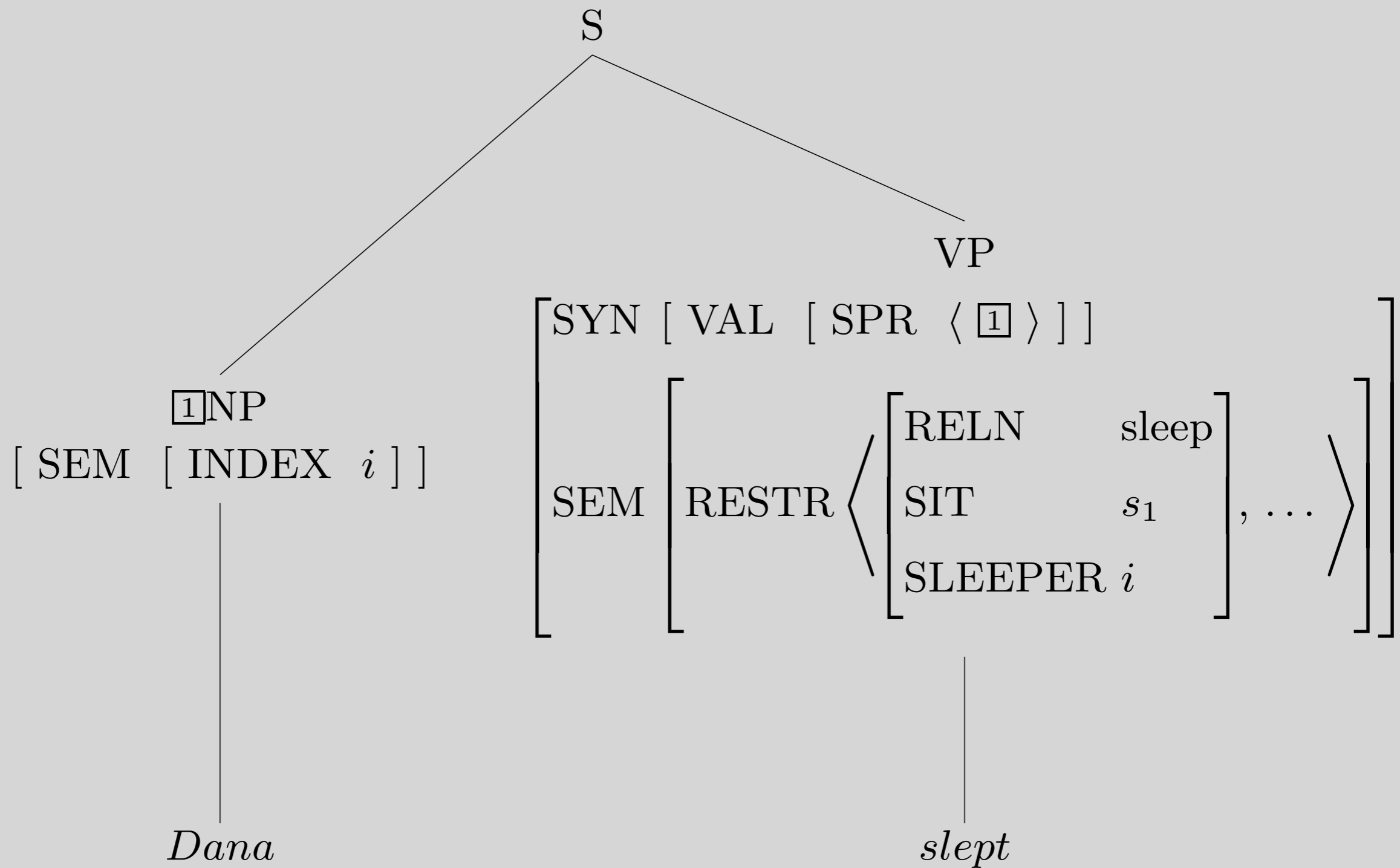
# How the Pieces Fit Together



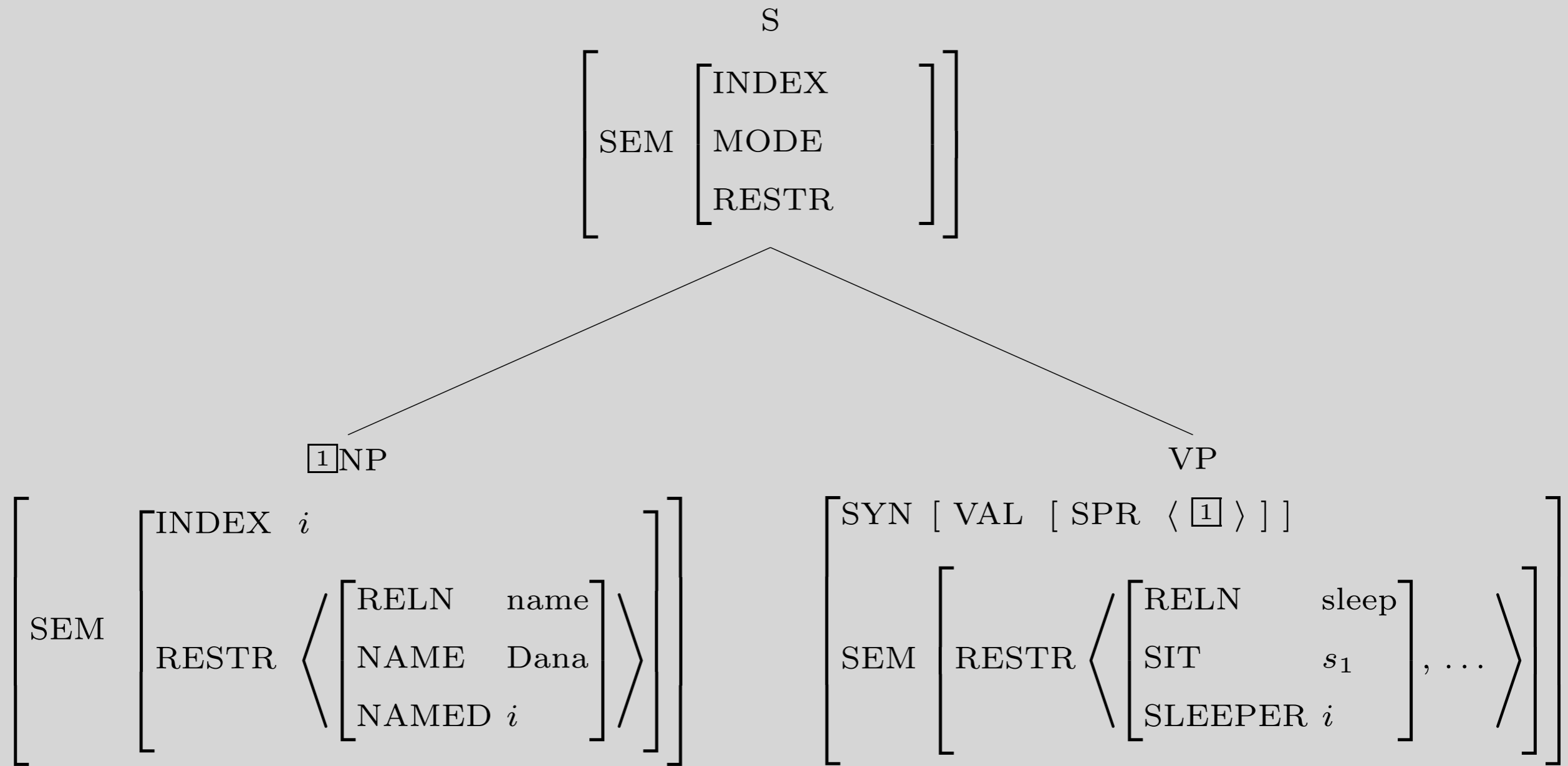
# How the Pieces Fit Together



# The Pieces Together



# A More Detailed View of the Same Tree



# To Fill in Semantics for the S-node

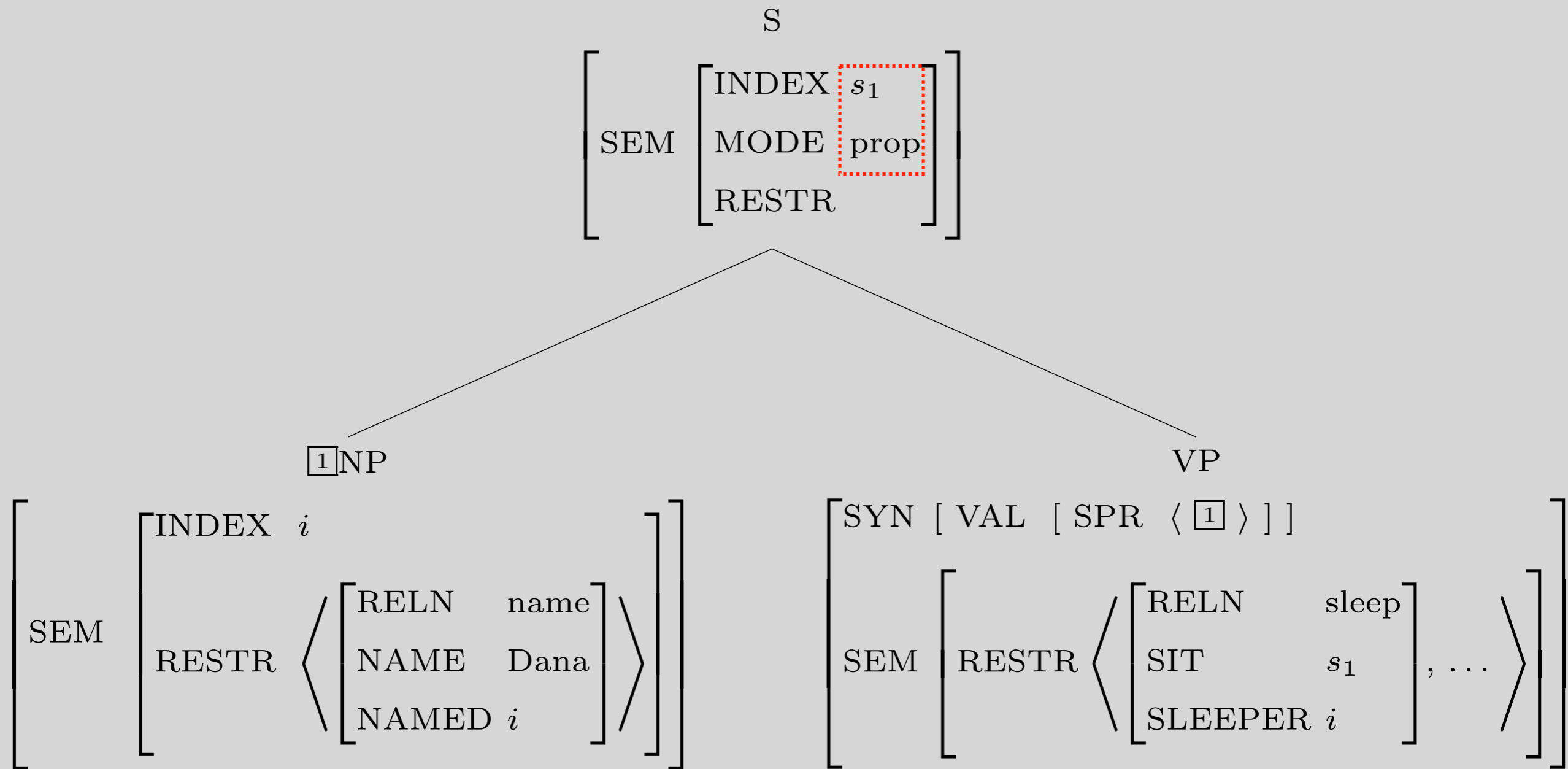
We need the Semantics Principles

- The Semantic Inheritance Principle:

In any headed phrase, the mother's **MODE** and **INDEX** are identical to those of the head daughter.

- The Semantic Compositionality Principle:

# Semantic Inheritance Illustrated



# To Fill in Semantics for the S-node

## We need the Semantics Principles

- The Semantic Inheritance Principle:

In any headed phrase, the mother's **MODE** and **INDEX** are identical to those of the head daughter.

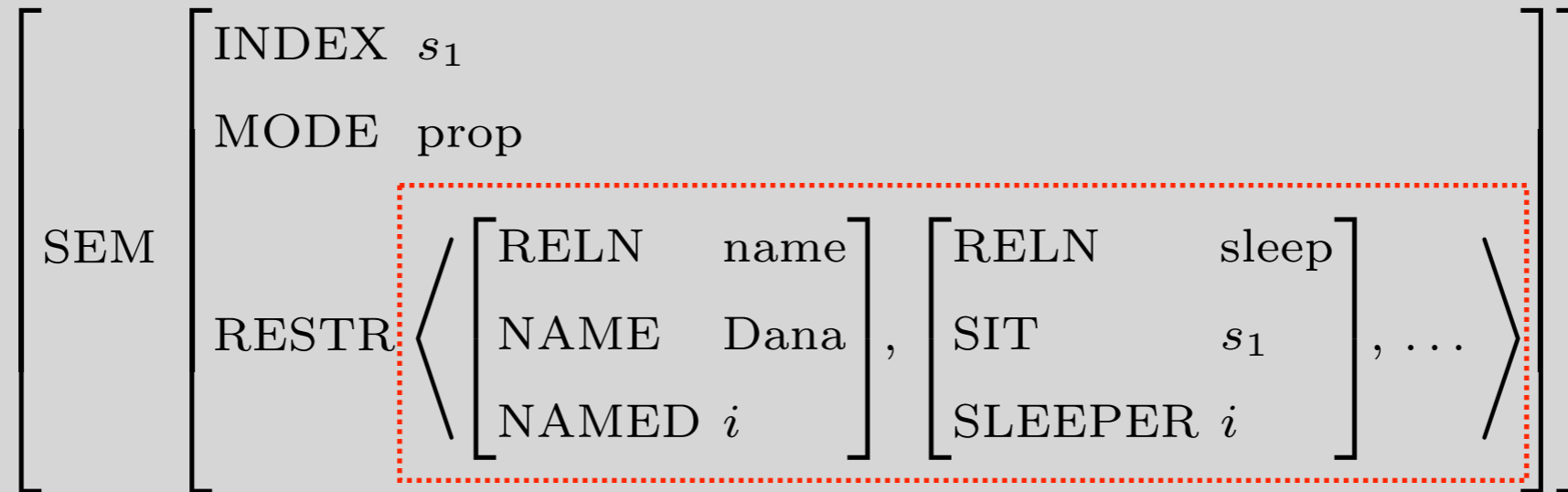
- The Semantic Compositionality Principle:

In any well-formed phrase structure, the mother's **RESTR** value is the sum of the **RESTR** values of the daughters.



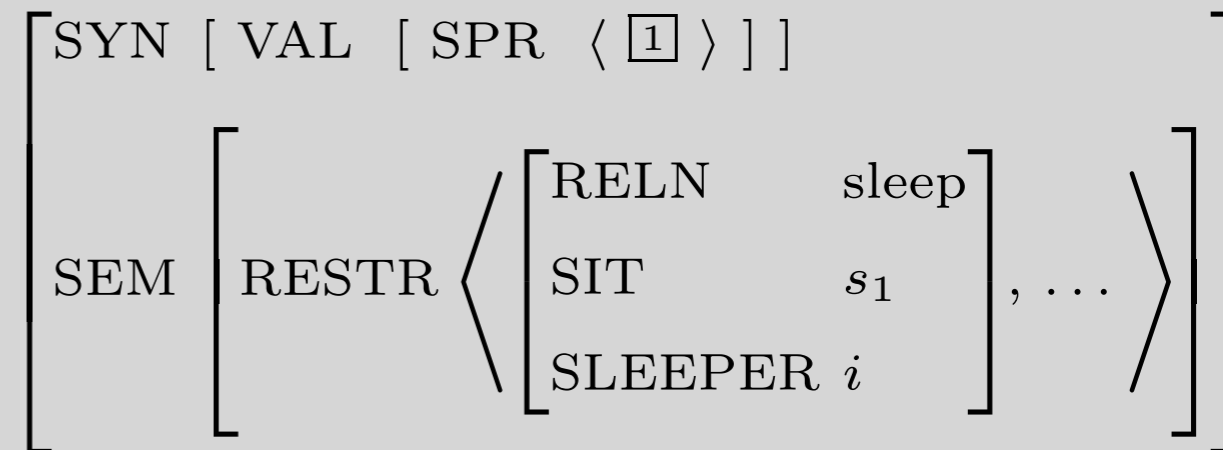
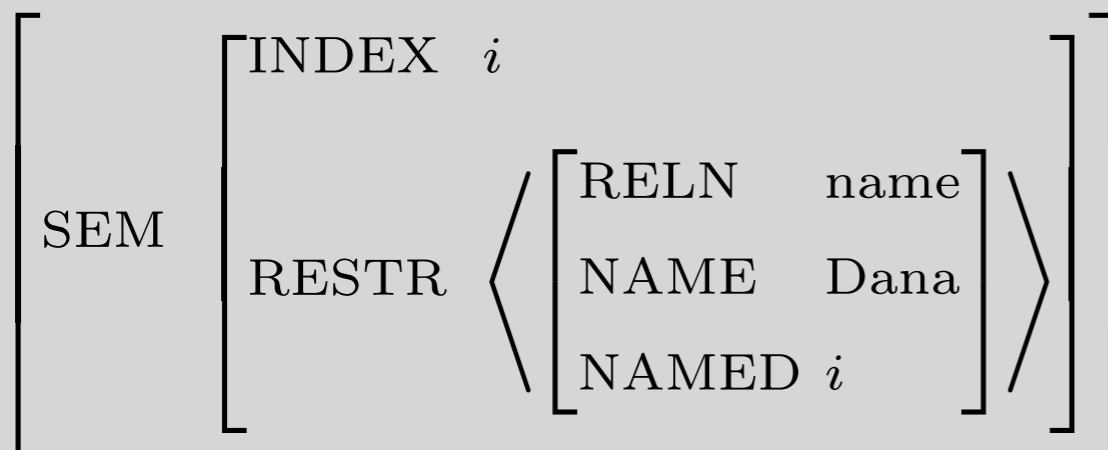
# Semantic Compositionality Illustrated

S

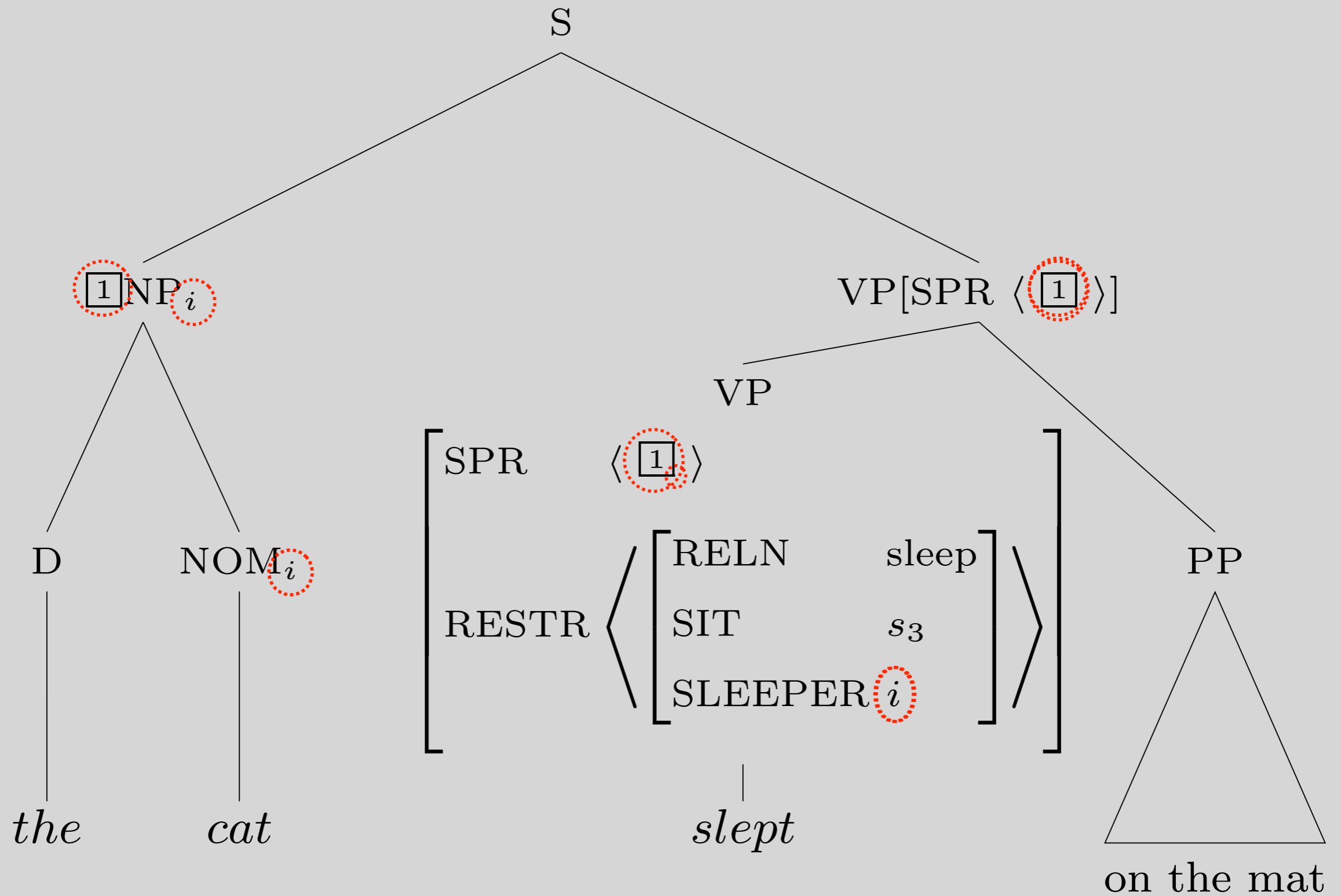


$\boxed{1}$ NP

VP

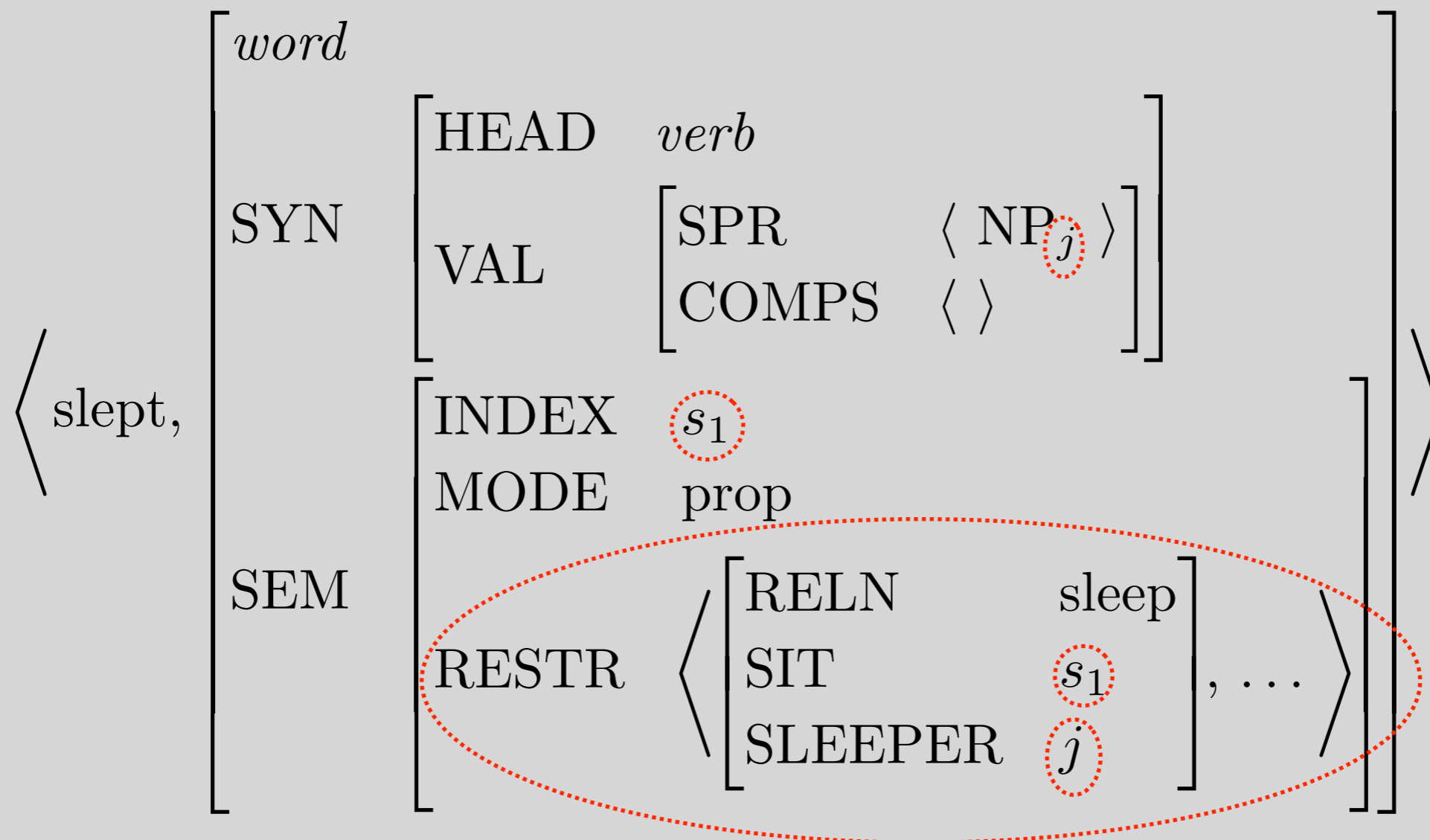


# What Identifies Indices?



# Summary: Words ...

- contribute predications
- ‘expose’ one index in those predications, for use by words or phrases
- relate syntactic arguments to semantic arguments



# Summary: Grammar Rules ...

- identify feature structures (including the INDEX value) across daughters

## Head Specifier Rule

$$\left[ \begin{array}{l} \textit{phrase} \\ \text{SYN} \left[ \text{VAL} \left[ \text{SPR} \langle \rangle \right] \right] \end{array} \right] \rightarrow \boxed{1} \mathbf{H} \left[ \text{SYN} \left[ \text{VAL} \left[ \begin{array}{l} \text{SPR} \langle \boxed{1} \rangle \\ \text{COMPS} \langle \rangle \end{array} \right] \right] \right]$$

## Head Complement Rule

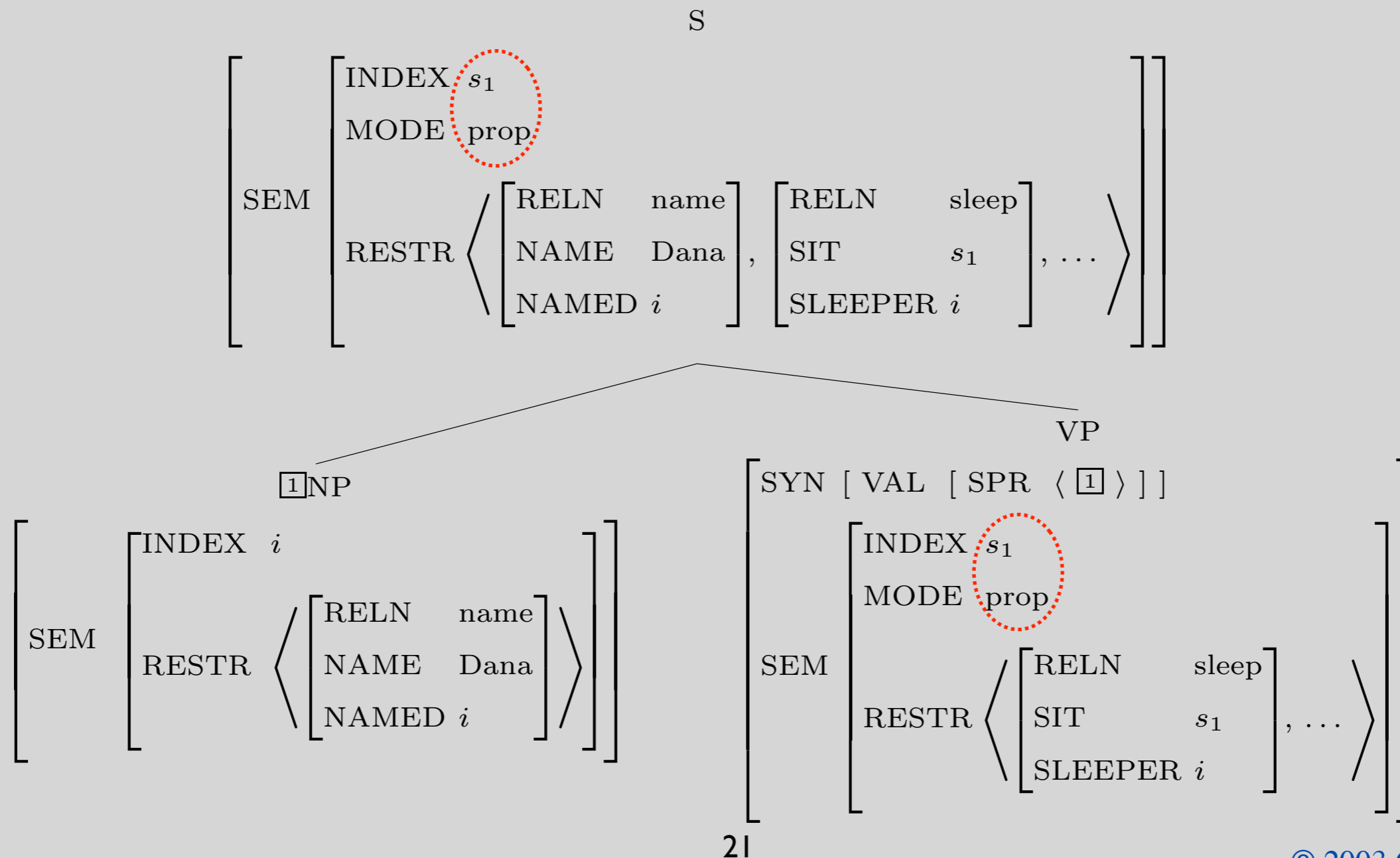
$$\left[ \begin{array}{l} \textit{phrase} \\ \text{SYN} \left[ \text{VAL} \left[ \text{COMPS} \langle \rangle \right] \right] \end{array} \right] \rightarrow \mathbf{H} \left[ \begin{array}{l} \textit{word} \\ \text{SYN} \left[ \text{VAL} \left[ \text{COMPS} \langle \boxed{1}, \dots, \boxed{n} \rangle \right] \right] \end{array} \right] \boxed{1} \dots \boxed{n}$$

## Head Modifier Rule

$$[\textit{phrase}] \rightarrow \mathbf{H} \boxed{1} \left[ \text{SYN} \left[ \text{COMPS} \langle \rangle \right] \left[ \text{SYN} \left[ \text{VAL} \left[ \begin{array}{l} \text{COMPS} \langle \rangle \\ \text{MOD} \langle \boxed{1} \rangle \end{array} \right] \right] \right] \right]$$

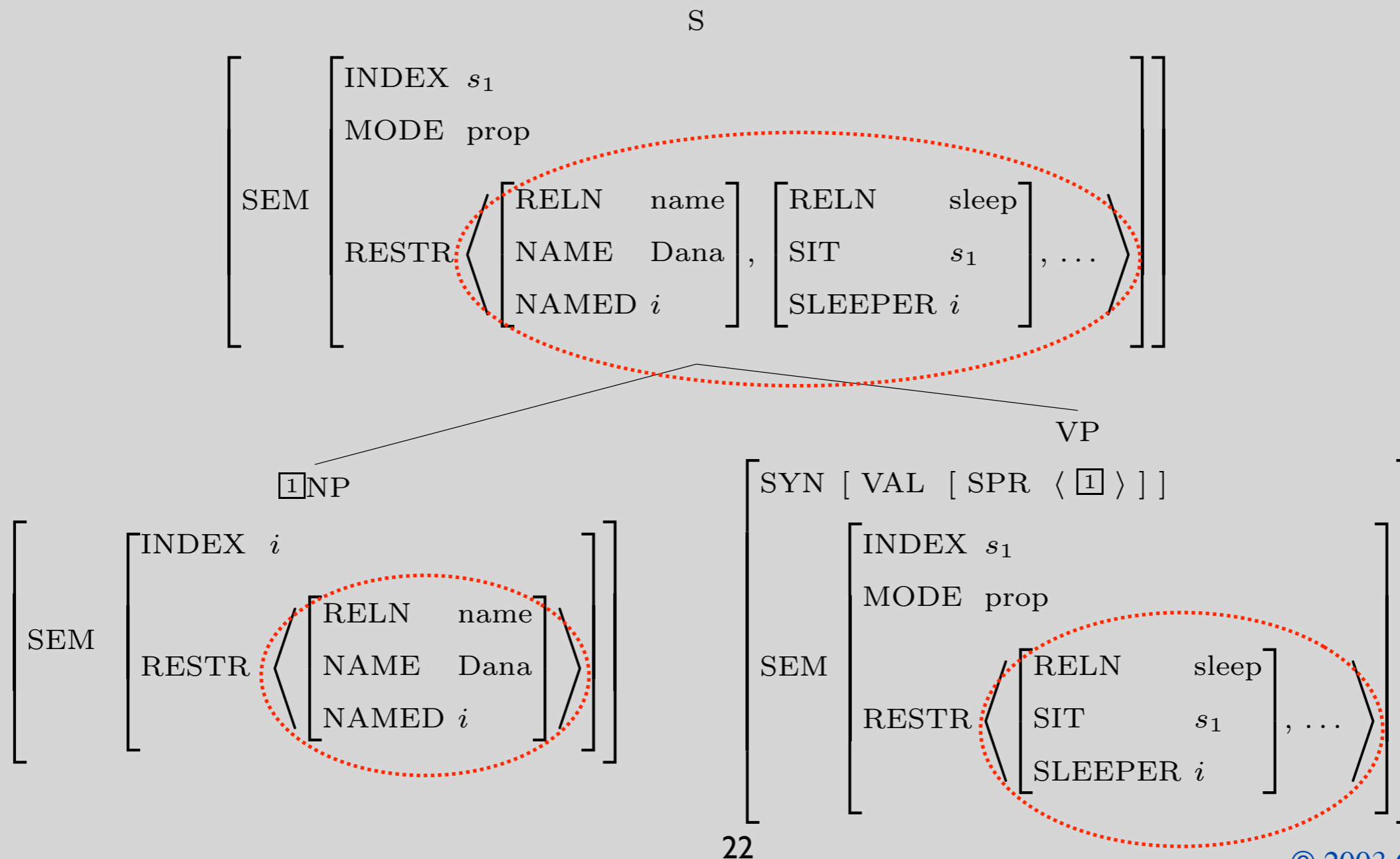
# Summary: Grammar Rules ...

- identify feature structures (including the INDEX value) across daughters
- license trees which are subject to the semantic principles
  - SIP 'passes up' MODE and INDEX from head daughter



# Summary: Grammar Rules ...

- identify feature structures (including the INDEX value) across daughters
- license trees which are subject to the semantic principles
  - SIP 'passes up' MODE and INDEX from head daughter
  - SCP: 'gathers up' predications (RESTR list) from all daughters



# Other Aspects of Semantics

- Tense, Quantification (only touched on here)
- Modification
- Coordination
- Structural Ambiguity

# Evolution of a Phrase Structure Rule

Ch. 2: NOM --> NOM PP  
 VP --> VP PP

Ch. 3: 
$$\left[ \begin{array}{l} \textit{phrase} \\ \text{VAL} \left[ \begin{array}{l} \text{COMPS itr} \\ \text{SPR } - \end{array} \right] \end{array} \right] \rightarrow \mathbf{H} \left[ \begin{array}{l} \textit{phrase} \\ \text{VAL} \left[ \begin{array}{l} \text{SPR } - \end{array} \right] \end{array} \right] \text{PP}$$

Ch. 4: 
$$[\textit{phrase}] \rightarrow \mathbf{H} \left[ \text{VAL} \left[ \text{COMPS} \langle \rangle \right] \right] \text{PP}$$

Ch. 5: 
$$[\textit{phrase}] \rightarrow \mathbf{H}[\boxed{1}] \left[ \text{SYN} \left[ \text{VAL} \left[ \text{COMPS} \langle \rangle \right] \right] \right] \left[ \text{SYN} \left[ \text{VAL} \left[ \begin{array}{l} \text{COMPS} \langle \rangle \\ \text{MOD} \langle \boxed{1} \rangle \end{array} \right] \right] \right]$$

Ch. 5 (abbreviated): 
$$[\textit{phrase}] \rightarrow \mathbf{H}[\boxed{1}] \left[ \text{COMPS} \langle \rangle \right] \left[ \begin{array}{l} \text{COMPS} \langle \rangle \\ \text{MOD} \langle \boxed{1} \rangle \end{array} \right]$$



# Evolution of Another Phrase Structure Rule

Ch. 2:  $X \dashrightarrow X^+ \text{ CONJ } X$

Ch. 3:  $\boxed{1} \rightarrow \boxed{1}^+ \begin{bmatrix} \textit{word} \\ \text{HEAD} \textit{ conj} \end{bmatrix} \boxed{1}$

Ch. 4:  $\begin{bmatrix} \text{VAL} \boxed{1} \end{bmatrix} \rightarrow \begin{bmatrix} \text{VAL} \boxed{1} \end{bmatrix}^+ \begin{bmatrix} \textit{word} \\ \text{HEAD} \textit{ conj} \end{bmatrix} \begin{bmatrix} \text{VAL} \boxed{1} \end{bmatrix}$

Ch. 5:  $\begin{bmatrix} \text{SYN} \text{ [VAL } \boxed{0}] \\ \text{SEM} \text{ [IND } s_0] \end{bmatrix} \rightarrow$   
 $\begin{bmatrix} \text{SYN} \text{ [VAL } \boxed{0}] \\ \text{SEM} \text{ [IND } s_1] \end{bmatrix} \cdots \begin{bmatrix} \text{SYN} \text{ [VAL } \boxed{0}] \\ \text{SEM} \text{ [IND } s_{n-1}] \end{bmatrix} \begin{bmatrix} \text{SYN} \text{ [HEAD } \textit{ conj}] \\ \text{SEM} \begin{bmatrix} \text{IND } s_0 \\ \text{RESTR} \langle \text{[ARGS } \langle s_1 \dots s_n \rangle \rangle \end{bmatrix} \end{bmatrix} \begin{bmatrix} \text{SYN} \text{ [VAL } \boxed{0}] \\ \text{SEM} \text{ [IND } s_n] \end{bmatrix}$

Ch. 5 (abbreviated):

$$\begin{bmatrix} \text{VAL} \boxed{0} \\ \text{IND } s_0 \end{bmatrix} \rightarrow \begin{bmatrix} \text{VAL} \boxed{0} \\ \text{IND } s_1 \end{bmatrix} \cdots \begin{bmatrix} \text{VAL} \boxed{0} \\ \text{IND } s_{n-1} \end{bmatrix} \begin{bmatrix} \text{HEAD } \textit{ conj} \\ \text{IND } s_0 \\ \text{RESTR} \langle \text{[ARGS } \langle s_1 \dots s_n \rangle \rangle \end{bmatrix} \begin{bmatrix} \text{VAL} \boxed{0} \\ \text{IND } s_n \end{bmatrix}$$

# Combining Constraints and Coordination

## Coordination Rule

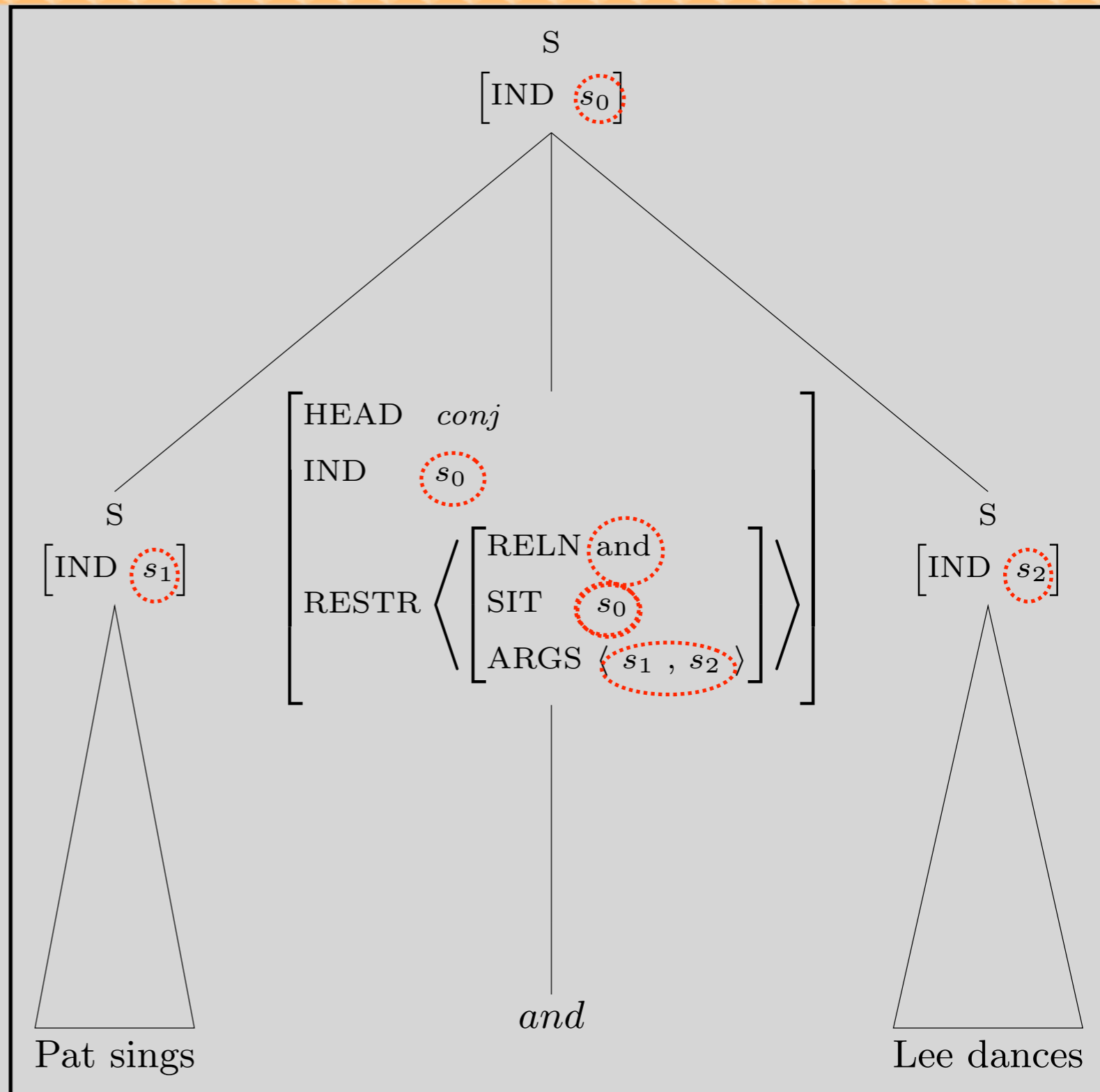
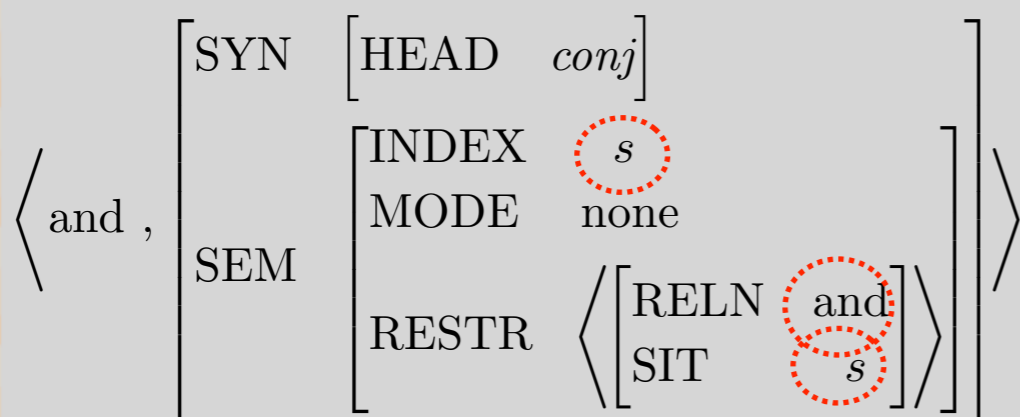
$$\begin{bmatrix} \text{VAL} & \boxed{0} \\ \text{IND} & s_0 \end{bmatrix} \rightarrow \begin{bmatrix} \text{VAL} & \boxed{0} \\ \text{IND} & s_1 \end{bmatrix} \cdots \begin{bmatrix} \text{VAL} & \boxed{0} \\ \text{IND} & s_{n-1} \end{bmatrix} \begin{bmatrix} \text{HEAD} & conj \\ \text{IND} & s_0 \\ \text{RESTR} & \langle \text{ARGS} \langle s_1 \dots s_n \rangle \rangle \end{bmatrix} \begin{bmatrix} \text{VAL} & \boxed{0} \\ \text{IND} & s_n \end{bmatrix}$$

## Lexical Entry for a Conjunction

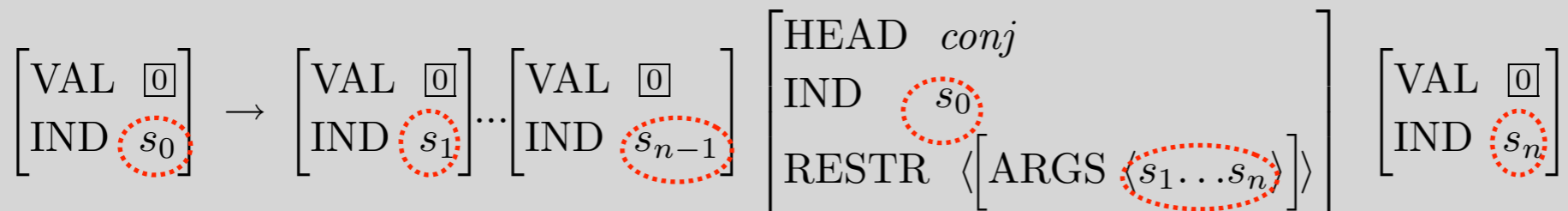
$$\left\langle \text{and} , \begin{bmatrix} \text{SEM} \begin{bmatrix} \text{SYN} & \begin{bmatrix} \text{HEAD} & conj \end{bmatrix} \\ \text{INDEX} & s \\ \text{MODE} & none \\ \text{RESTR} & \left\langle \begin{bmatrix} \text{RELN} & and \\ \text{SIT} & s \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \right\rangle$$

# Combining Constraints and Coordination

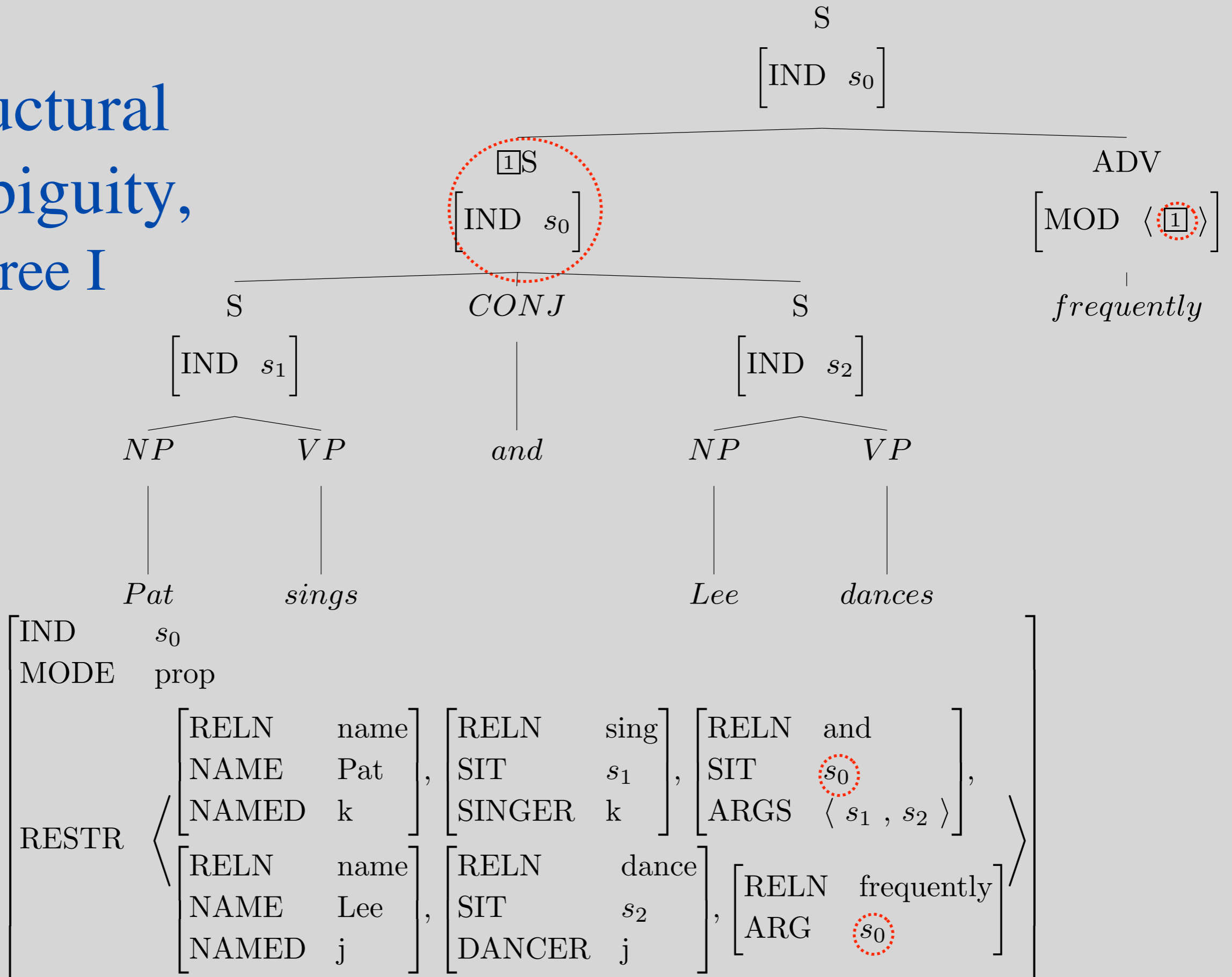
## Lexical Entry for *and*



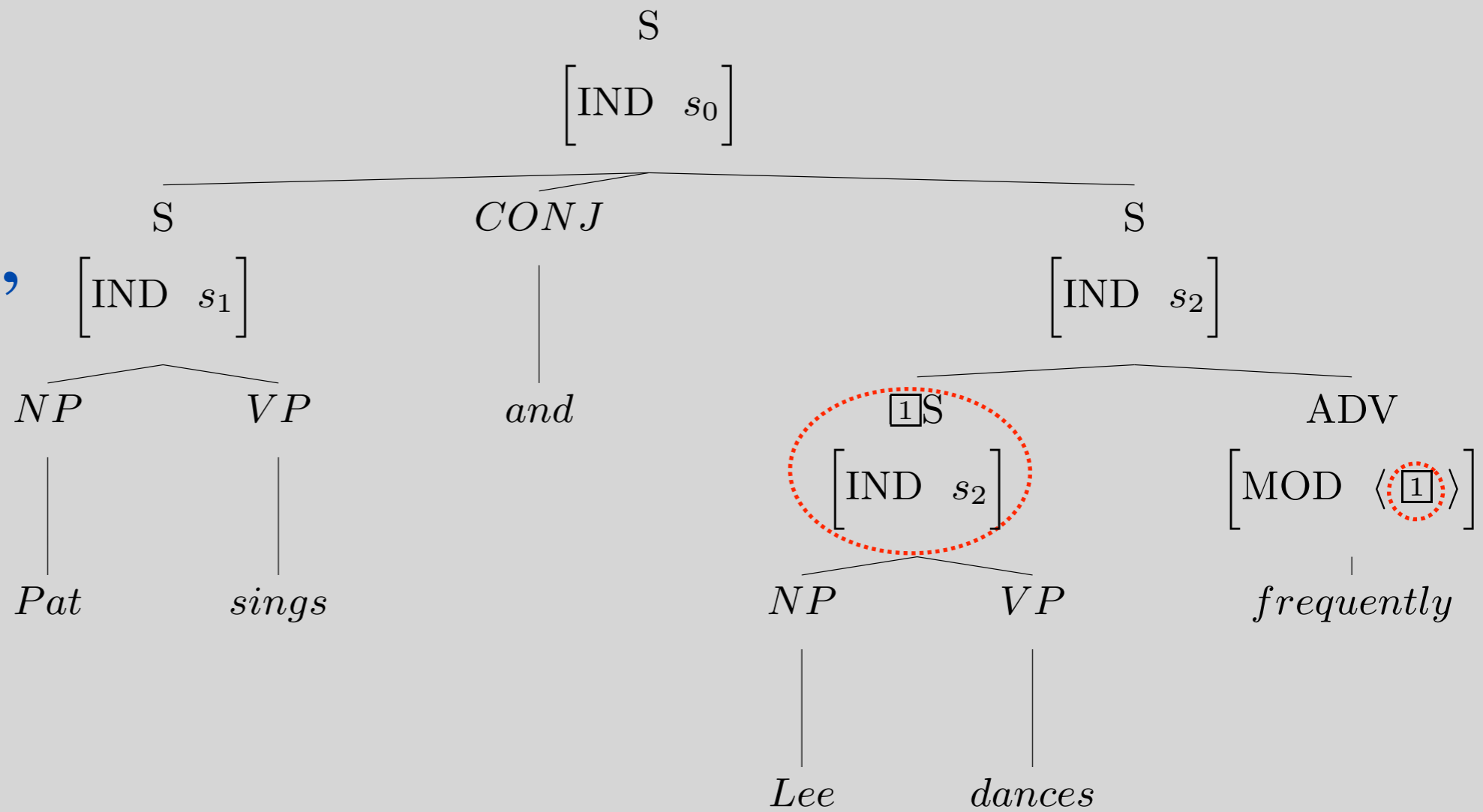
## Coordination Rule



# Structural Ambiguity, Tree I



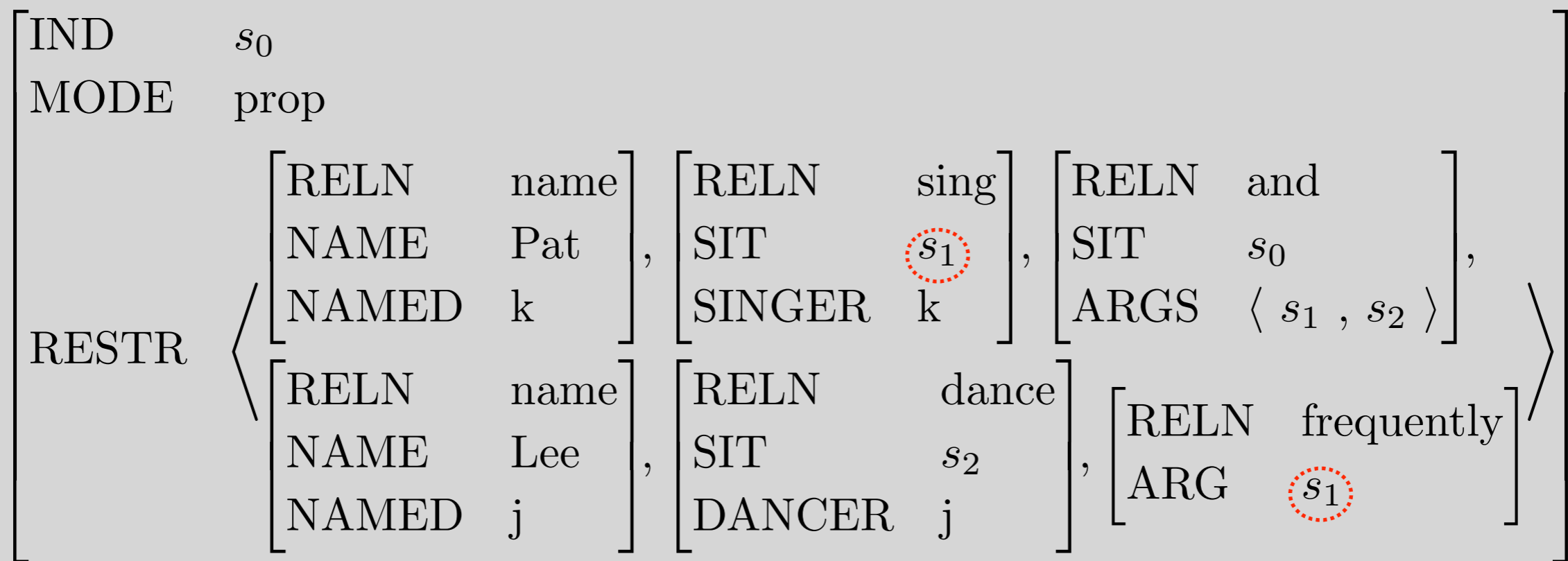
# Structural Ambiguity, Tree II



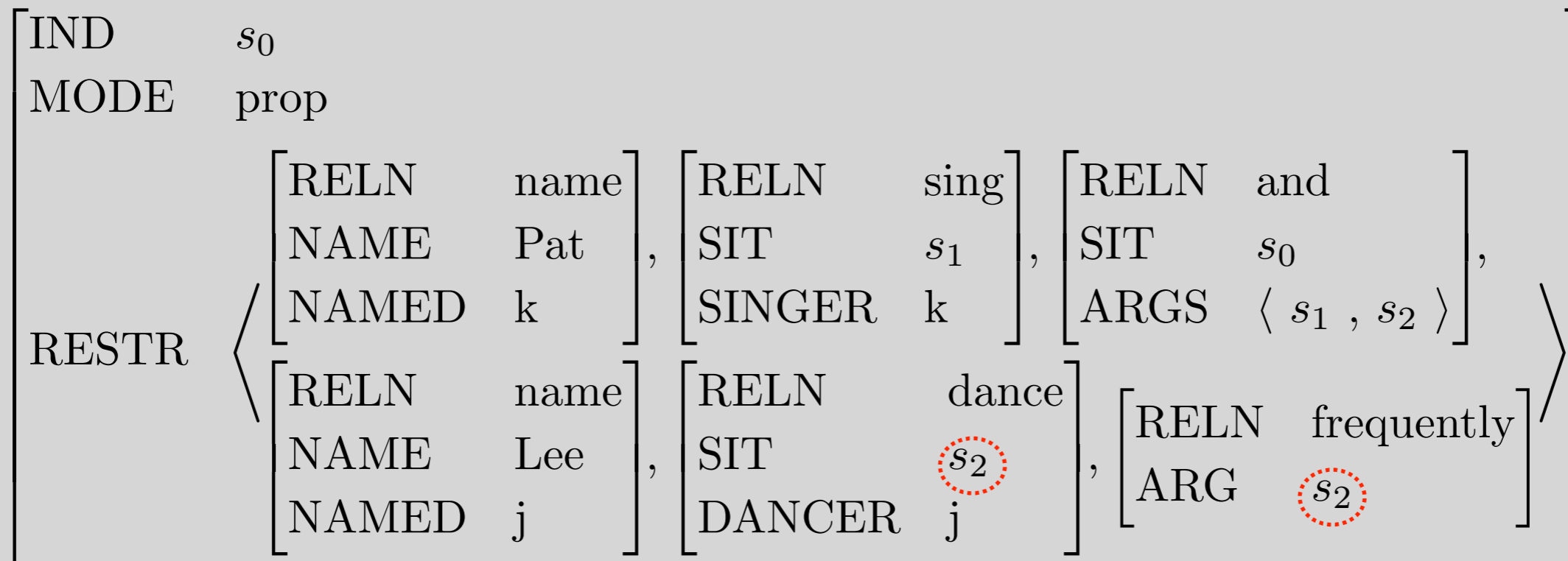
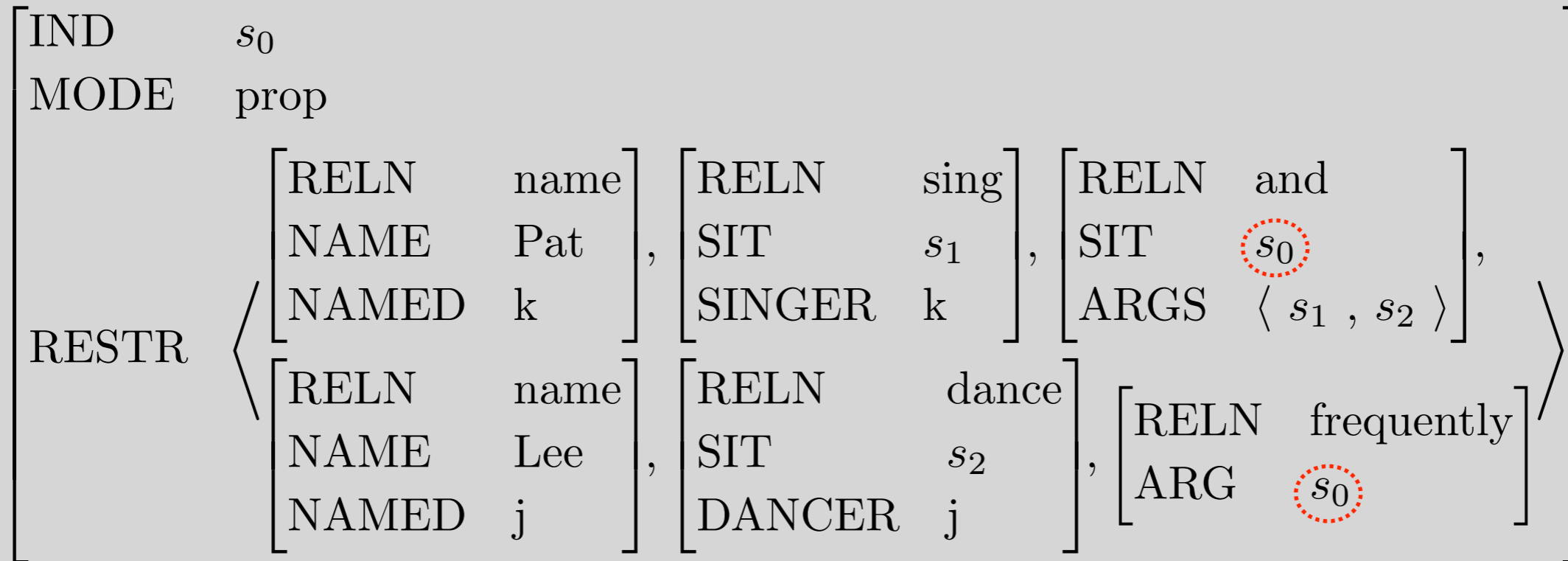
IND	$s_0$				
MODE	prop				
RESTR	$\left\langle \begin{array}{l} \left[ \begin{array}{ll} \text{RELN} & \text{name} \\ \text{NAME} & \text{Pat} \\ \text{NAMED} & \text{k} \end{array} \right], \left[ \begin{array}{ll} \text{RELN} & \text{sing} \\ \text{SIT} & s_1 \\ \text{SINGER} & \text{k} \end{array} \right], \left[ \begin{array}{ll} \text{RELN} & \text{and} \\ \text{SIT} & s_0 \\ \text{ARGS} & \langle s_1, s_2 \rangle \end{array} \right], \\ \left[ \begin{array}{ll} \text{RELN} & \text{name} \\ \text{NAME} & \text{Lee} \\ \text{NAMED} & \text{j} \end{array} \right], \left[ \begin{array}{ll} \text{RELN} & \text{dance} \\ \text{SIT} & s_2 \\ \text{DANCER} & \text{j} \end{array} \right], \left[ \begin{array}{ll} \text{RELN} & \text{frequently} \\ \text{ARG} & s_2 \end{array} \right] \end{array} \right\rangle$				

# Question About Structural Ambiguity

Why isn't this a possible semantic representation for the string *Pat sings and Lee dances frequently*?



# Semantic Compositionality



# Overview

- Some notes on the linguist's stance
- Which aspects of semantics we'll tackle
- Our formalization; Semantics Principles
- Building semantics of phrases
- Modification, coordination
- Structural ambiguity
- Next time: How the grammar works



# But first!

- More polls :)

# Reading Questions

- Why do we even bother including semantics at all? It feels like everything that we're doing can still be handled by the lexicon. Is this for handling concepts like reference in more complicated sentences?
- It feels like this adds a considerable amount of complexity, and that an alternative framework using exclusively the lexicon for these tasks would be possible. Even the final Head-Modifier Rule seems to not address SEM directly. Is this an example of grounding the model more with what we observe in real language rather than keeping it as simple as possible?

# Reading Questions

- Why is the MOD feature under the syntax feature structure instead of semantics?

# Reading Questions

- Is the main motivation for introducing semantics at this point to address syntactic changes due to mode (questions/imperatives/etc) and to identify indices for coordination?
- Now that we are adding a semantic analysis to our grammar, it seems like we will evaluate false sentence as ungrammatical on the basis that the semantic constraints are not met. Surely this is not desirable in a syntactic analysis of language.

# Reading Questions

- How would we be able to account for things like sarcasm or questions that aren't meant to be answered (more like statements) with MODE? Like for example, if I were to say, *Are you kidding me?* I'm saying this as an expression of my displeasure rather than asking for an answer. Would this still be labeled as "question" for MODE?
- In *Pat aches*), how does *aches* receive a "prop" MODE? I know that it says that the mother and head daughter need to agree on the INDEX and MODE features, but this feels nonsensical to me wrt MODE because something like being a proposition seems like a property of the sentence construction, not any given constituent?

# Reading Questions

- What is a situation? It didn't feel like we got a satisfying definition.
- What's the point of the INDEX feature?
- Why can it "take an unlimited number of different values"?
- How is the INDEX feature different from the ARG feature?

# Reading Questions

- Regarding RELN attributes - can these be anything? Are they just labels that are meant to reference other attributes within an expression's other Semantic restrictions?
- For predications, is there a standard way to know that you have included all of the features that are semantically relevant for a certain word? Also, are there any naming conventions when labeling the features of the predication?

# Reading Questions

- It seems that RELN values can be pretty much anything-- verbs, nouns, adjectives, prepositions. Why is that? If their value can be so many different things, then wouldn't predication be very difficult to model in feature structures, or wouldn't it cause ambiguity?
- Are there words that share the same RELN other than **name**? For example, could the words *throw* and *toss* share the same RELN? Does sharing a RELN make two words synonyms? Can connotations also be represented in predications? Or do "synonyms" with different connotations have different RELNs?



# Reading Questions

- For the two predictions for “Chris” and “Pat”, why can’t they be written as something like:

$$\begin{bmatrix} \text{RELN} & \text{Chris/Pat} \\ \text{SIT} & s \\ \text{INST} & i \end{bmatrix}$$

- because it seems to me the value for RELN is just the individual word in a sentence. And we use INST for a single argument situation. Or for proper nouns specifically, we further generalize them as “name” and use NAME + NAMED instead of INST

# Reading Questions

- What's going on with quantifiers?
- What are BV, QRESTR and QSCOPE?
- How does scope underspecification work and why do we want it?

- (123) a.  $a_D(X, Y) \iff X \cap Y \neq \emptyset$   
 b.  $every_D(X, Y) \iff X \subseteq Y$

Other quantifiers can be defined as relations between sets as well, as illustrated in (125) which gives representations for the quantifiers used in (124):

- (124) a. More than two dogs barked.  
 b. Neither dog barked.  
 c. More than half of the dogs barked.  
 d. Only dogs bark.  
 e. All but one of the dogs barked.
- (125) a.  $more\ than\ 2_D(X, Y) \iff |X \cap Y| \geq 2$   
 b.  $neither_D(X, Y) \iff |X| = 2 \text{ and } X \cap Y = \emptyset$   
 c.  $more\ than\ half_D(X, Y) \iff |X \cap Y| > \frac{|X|}{2}$   
 d.  $only_D(X, Y) \iff X \cap Y \neq \emptyset \text{ and } (D \setminus X) \cap Y = \emptyset$   
 e.  $all\ but\ one_D(X, Y) \iff |X \cap Y| = |X| - 1 \text{ and } |X \cap (D \setminus Y)| = 1$

$$(50) \left[ \text{RESTR} \left\langle \begin{array}{l} \left[ \begin{array}{ll} \text{RELN} & \text{exist} \\ \text{BV} & i \\ \text{QRESTR} & \boxed{1} \\ \text{QSCOPE} & \boxed{2} \end{array} \right], \boxed{1} \left[ \begin{array}{ll} \text{RELN} & \text{dog} \\ \text{INST} & i \end{array} \right], \boxed{2} \left[ \begin{array}{ll} \text{RELN} & \text{all} \\ \text{BV} & j \\ \text{QRESTR} & \boxed{3} \\ \text{QSCOPE} & \boxed{4} \end{array} \right], \\ \boxed{3} \left[ \begin{array}{ll} \text{RELN} & \text{family} \\ \text{INST} & j \end{array} \right], \boxed{4} \left[ \begin{array}{ll} \text{RELN} & \text{save} \\ \text{SAVER} & i \\ \text{SAVED} & j \end{array} \right] \end{array} \right\rangle \right]$$

$$(51) \left[ \text{RESTR} \left\langle \begin{array}{l} \left[ \begin{array}{ll} \text{RELN} & \text{exist} \\ \text{BV} & i \\ \text{QRESTR} & \boxed{1} \\ \text{QSCOPE} & \boxed{4} \end{array} \right], \boxed{1} \left[ \begin{array}{ll} \text{RELN} & \text{dog} \\ \text{INST} & i \end{array} \right], \left[ \begin{array}{ll} \text{RELN} & \text{all} \\ \text{BV} & j \\ \text{QRESTR} & \boxed{3} \\ \text{QSCOPE} & \boxed{2} \end{array} \right], \\ \boxed{3} \left[ \begin{array}{ll} \text{RELN} & \text{family} \\ \text{INST} & j \end{array} \right], \boxed{4} \left[ \begin{array}{ll} \text{RELN} & \text{save} \\ \text{SAVER} & i \\ \text{SAVED} & j \end{array} \right] \end{array} \right\rangle \right]$$

# Reading Questions

- On page 154, the text talks about how determiner identifies its own index with the value of  $BV(i)$ , which it turn is the noun of the NP, *a dog*. So I understand we want to link the determiner and the noun in each of their SEM features as well as SYN, but why is it that in the SYN features we make the noun point to its determiner in SPR, but in the SEM features it looks like the determiner is pointing instead to its noun in the BV?

# Reading Questions

- I am wondering how useful the SEM feature is in real-world applications. This is because the chapter outlines many ways that object relationships can be described by the SEM feature (including some forms of predicate logic), which makes me wonder whether this can be useful in extracting triples for knowledge graphs/ontology-based systems.

# Reading Questions

- Are these semantic features what are used to aid in machine translation? For instance, in Spanish, the verbs change to an imperative form when someone is making a command unlike in English.
- I'm curious about whether semantic analysis with feature structures for each word is related to sentiment analysis--I don't know much about sentiment analysis, but I could imagine words having a feature structure but with features that indicate different feelings or intentions instead.

# Reading Questions

- If there is nobody to consume language, does it have any meaning?

Bender Emily M. and Alexander Koller.

2020. Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data. *ACL*

2020 [.bib] [audiopaper] [ACL 2020 presentation]

[slides]