

Ling 566

Nov 12, 2020

Non-referential NPs, Expletives, and Extraposition

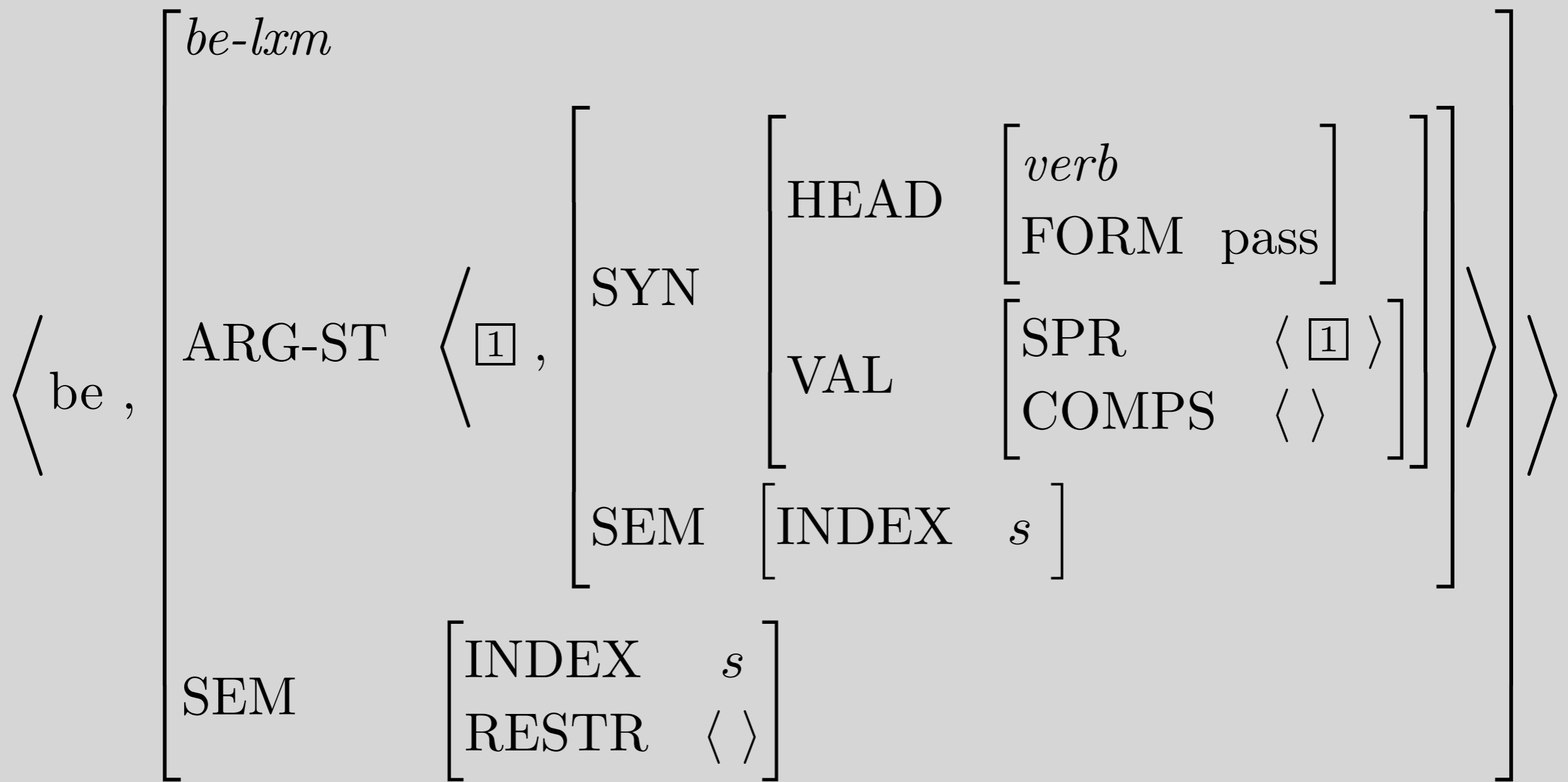
Overview

- Existentials
- Extraposition
- Idioms

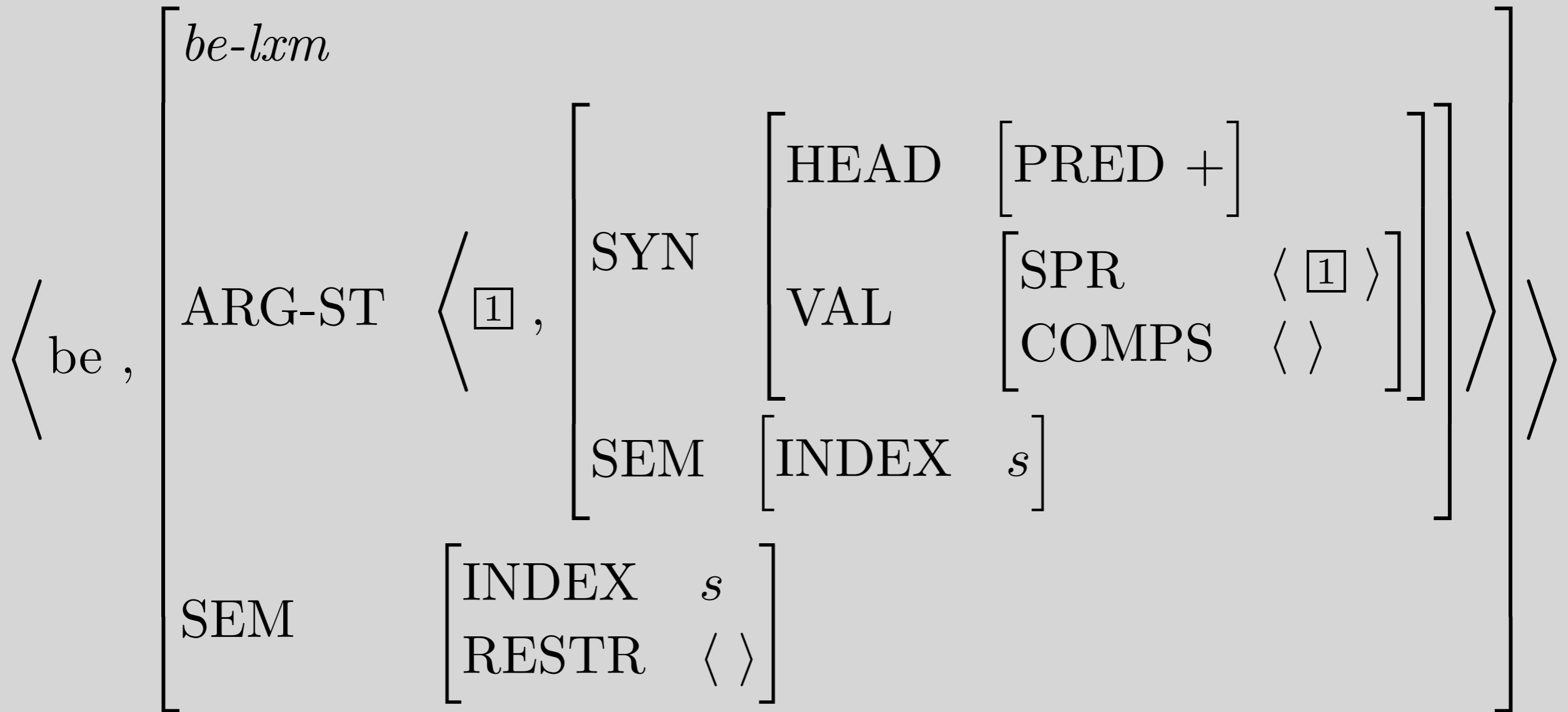
Where We Are, and Where We're Going

- Last time, we met the passive *be*.
- Passive *be* is just a special case -- that *be* generally introduces [PRED +] constituents (next slide).
- Today, we'll start with another *be*, which occurs in existential sentences starting with *there*, e.g. *There is a monster in Loch Ness*.
- Then we'll look at this use of *there*.
- Which will lead us to a more general examination of NPs that don't refer, including some uses of *it* and certain idiomatic uses of NPs.

Chapter 10 entry for *be*



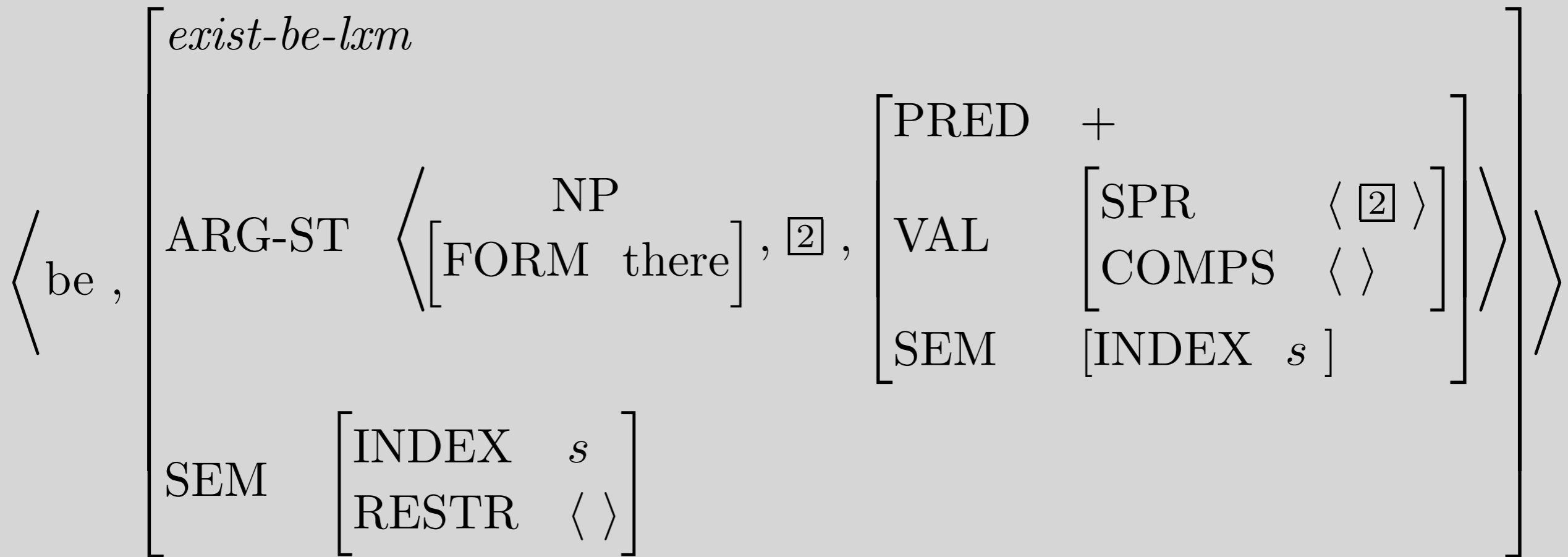
Copula (generalized)



Existentials

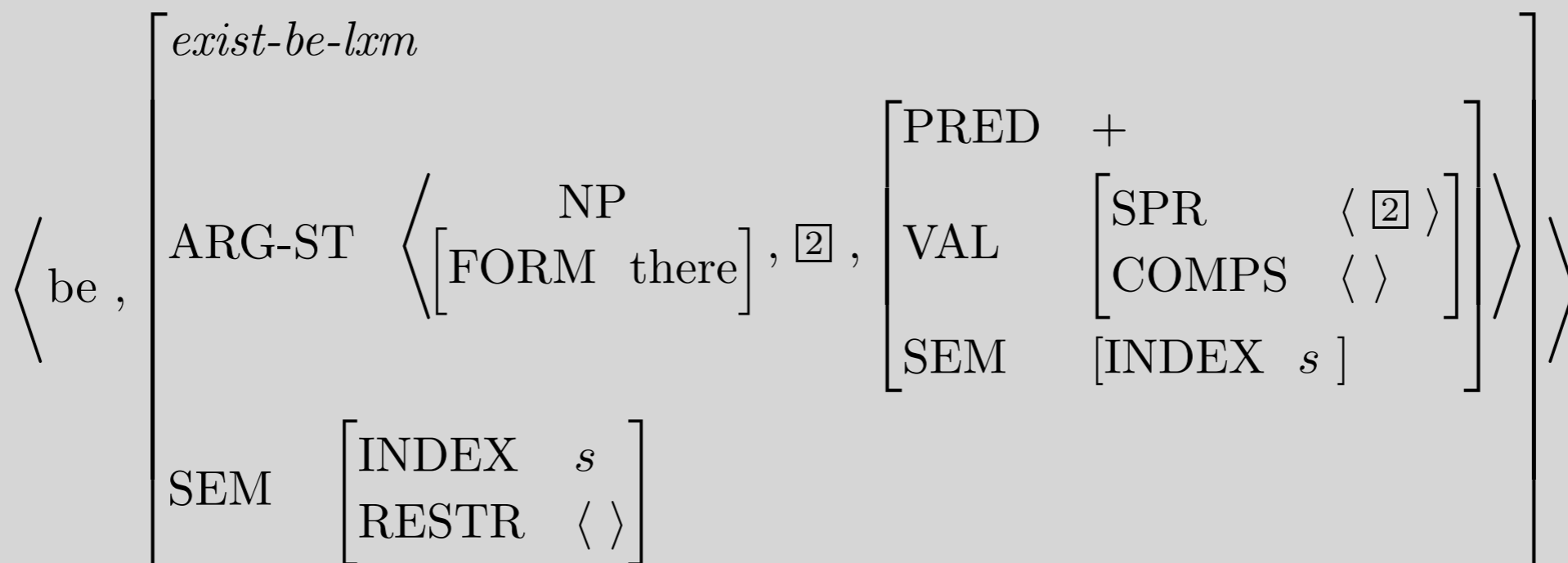
- The *be* in *There is a page missing* cannot be the same *be* that occurs in sentences like *Pat is tall* or *A cat was chased by a dog*. Why not?
- So we need a separate lexical entry for this *be*, stipulating:
 - Its SPR must be *there*
 - It takes two complements, the first an NP and the second an AP, PP, or (certain kind of) VP.
 - The semantics should capture the relation between, e.g. *There is a page missing* and *A page is missing*.

Lexical Entry for the Existential *be*



Questions About the Existential *be*

- What type of constituent is the third argument?
- Why is the third argument [PRED +]?
- Why is the second argument tagged as identical to the SPR of the third argument?
- What is the contribution of this *be* to the semantics of the sentences it occurs in?
- Can all [PRED +] predicates appear as the third argument in existentials?



The Entry for Existential *there*

| | | | | |
|-------------------|------------|-------------------------|------------|--|
| \langle there , | $pron-lxm$ | | | |
| | SYN | HEAD | FORM there | |
| | | AGR | PER 3rd | |
| | SEM | MODE none | | |
| | | INDEX none | | |
| | | RESTR $\langle \rangle$ | | |

Questions About Existential *there*

- Why do we call it a pronoun?
- Why don't we give it a value for NUM?
- What does this entry claim is *there*'s contribution to the semantics of the sentences it appears in?
Is this a correct claim?

| | | | | | | | | | |
|-------------------|-----|-----------------|-----|-------|-------------------|------|-------|---|---|
| \langle there , | [| <i>pron-lxm</i> | [| HEAD | [| FORM | there |] |] |
| | SYN | [| AGR | [| PER | 3rd |] |] | |
| | [| SEM | [| MODE | none |] | | |] |
| | | | [| INDEX | none |] | | | |
| | | | [| RESTR | $\langle \rangle$ |] | | | |

Other NPs that don't seem to refer

- *It sucks that the Rockies lost the series.*
- *It is raining.*
- *Andy took **advantage** of the opportunity.*
- *Lou kicked **the bucket**.*

What we need to deal with examples like *It follows that you are wrong*

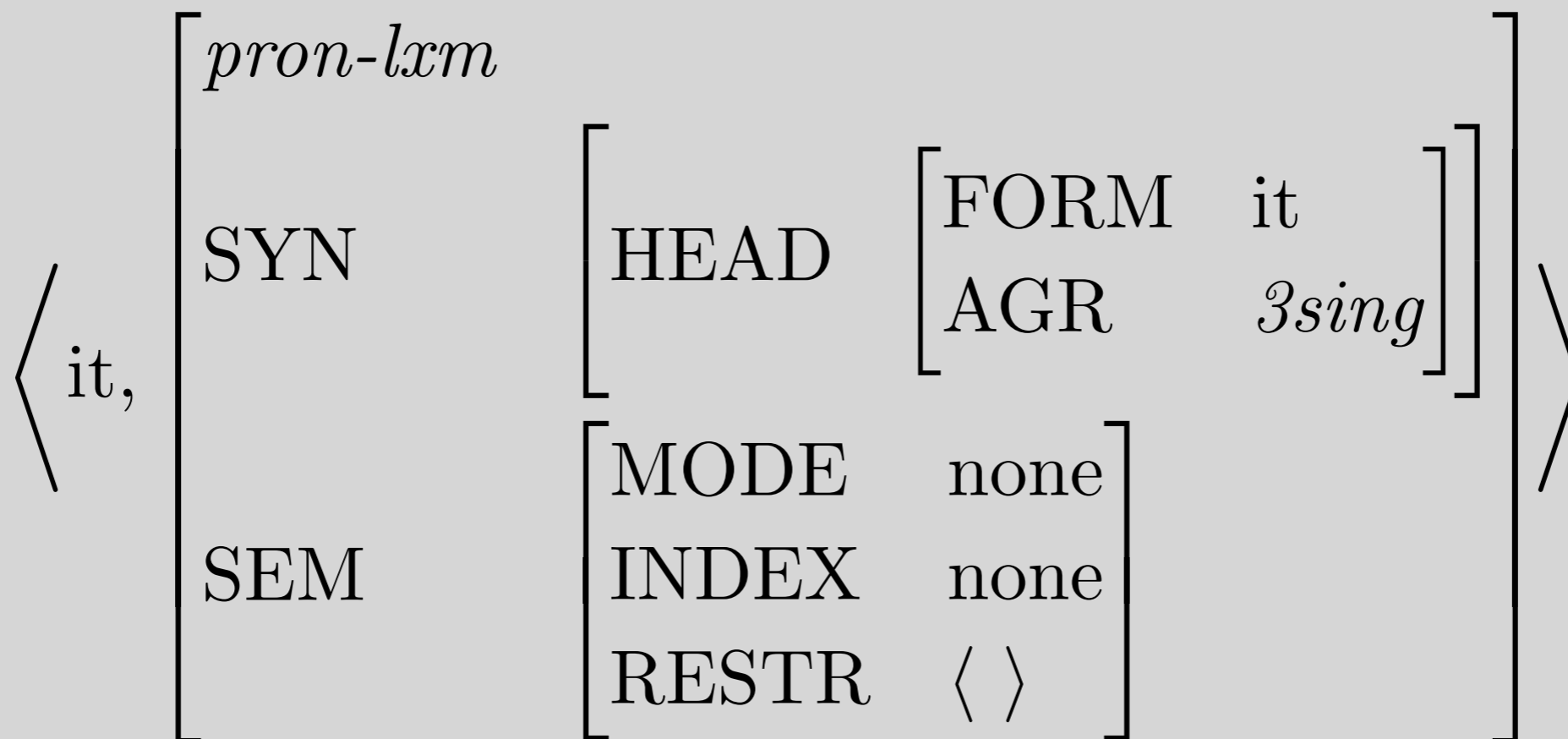
- A lexical entry for this dummy *it*
- An analysis of this use of *that*
- Entries for verbs that take clausal subjects
(as in *That you are wrong follows*)
- A rule to account for the relationship
between pairs like *That you are wrong
follows* and *It follows that you are wrong*

The Entry for Dummy *it*

| | | | |
|--------------------------------|-----------------|-------------------|------------------------------------------------------------------------------------|
| \langle <i>it,</i> \rangle | <i>pron-lxm</i> | | |
| | SYN | HEAD | $\left[\begin{array}{l} \text{FORM } it \\ \text{AGR } 3sing \end{array} \right]$ |
| SEM | MODE | none | |
| | INDEX | none | |
| | RESTR | $\langle \rangle$ | |

Questions About Dummy *it*

- How does it differ from the entry for dummy *there*? Why do they differ in this way?
- Is this the only entry for *it*?



A New Type of Lexeme: Complementizers

comp-lxm :

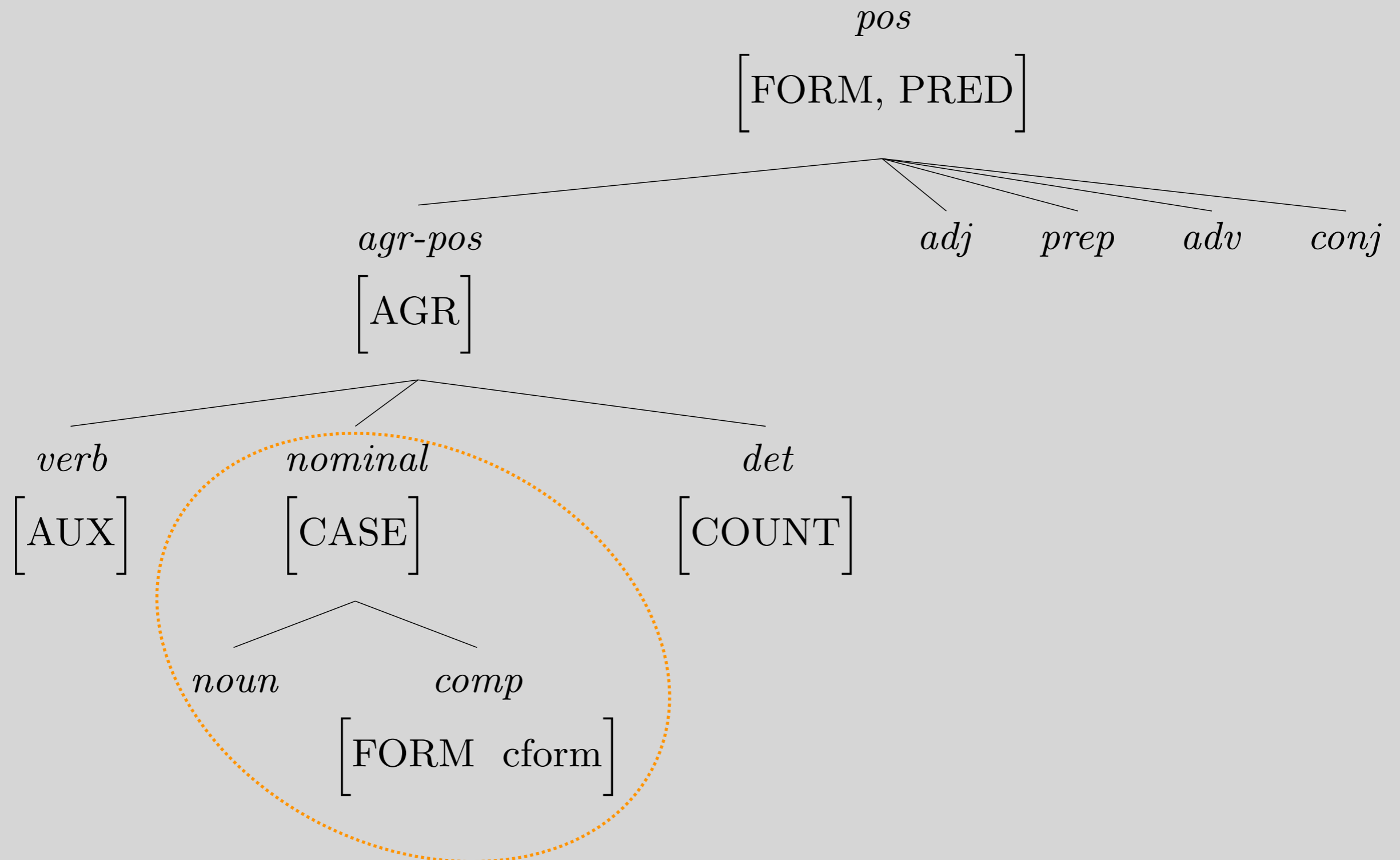
| | | | | | | | | | | | |
|-------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|--------------|-----|--------------|-----|----------------------------------------------------------------------------------------------------------------------------------|-----|-----|
| SYN | <table border="0"> <tr> <td style="vertical-align: middle;">HEAD</td> <td style="vertical-align: middle;"> <table border="0"> <tr> <td style="vertical-align: middle;"><i>comp</i></td> <td></td> </tr> <tr> <td style="vertical-align: middle;">AGR</td> <td style="vertical-align: middle;"><i>3sing</i></td> </tr> </table> </td> </tr> <tr> <td style="vertical-align: middle;">VAL</td> <td style="vertical-align: middle;"> <table border="0"> <tr> <td style="vertical-align: middle;">SPR</td> <td style="vertical-align: middle;">⟨ ⟩</td> </tr> </table> </td> </tr> </table> | HEAD | <table border="0"> <tr> <td style="vertical-align: middle;"><i>comp</i></td> <td></td> </tr> <tr> <td style="vertical-align: middle;">AGR</td> <td style="vertical-align: middle;"><i>3sing</i></td> </tr> </table> | <i>comp</i> | | AGR | <i>3sing</i> | VAL | <table border="0"> <tr> <td style="vertical-align: middle;">SPR</td> <td style="vertical-align: middle;">⟨ ⟩</td> </tr> </table> | SPR | ⟨ ⟩ |
| HEAD | <table border="0"> <tr> <td style="vertical-align: middle;"><i>comp</i></td> <td></td> </tr> <tr> <td style="vertical-align: middle;">AGR</td> <td style="vertical-align: middle;"><i>3sing</i></td> </tr> </table> | <i>comp</i> | | AGR | <i>3sing</i> | | | | | | |
| <i>comp</i> | | | | | | | | | | | |
| AGR | <i>3sing</i> | | | | | | | | | | |
| VAL | <table border="0"> <tr> <td style="vertical-align: middle;">SPR</td> <td style="vertical-align: middle;">⟨ ⟩</td> </tr> </table> | SPR | ⟨ ⟩ | | | | | | | | |
| SPR | ⟨ ⟩ | | | | | | | | | | |
| ARG-ST | <table border="0"> <tr> <td style="vertical-align: middle;">S</td> <td></td> </tr> <tr> <td style="vertical-align: middle;">INDEX</td> <td style="vertical-align: middle;"><i>s</i></td> </tr> </table> | S | | INDEX | <i>s</i> | | | | | | |
| S | | | | | | | | | | | |
| INDEX | <i>s</i> | | | | | | | | | | |
| SEM | <table border="0"> <tr> <td style="vertical-align: middle;">INDEX</td> <td style="vertical-align: middle;"><i>s</i></td> </tr> <tr> <td style="vertical-align: middle;">RESTR</td> <td style="vertical-align: middle;">⟨ ⟩</td> </tr> </table> | INDEX | <i>s</i> | RESTR | ⟨ ⟩ | | | | | | |
| INDEX | <i>s</i> | | | | | | | | | | |
| RESTR | ⟨ ⟩ | | | | | | | | | | |

Questions About the Type *comp-lxm*

- Why does it stipulate values for both SPR and ARG-ST?
- Why is its INDEX value the same as its argument's?
- What is its semantic contribution?

$$\textit{comp-lxm} : \left[\begin{array}{l} \text{SYN} \\ \text{ARG-ST} \\ \text{SEM} \end{array} \left[\begin{array}{l} \text{HEAD} \left[\begin{array}{l} \textit{comp} \\ \text{AGR} \quad \textit{3sing} \end{array} \right] \\ \text{VAL} \left[\begin{array}{l} \text{SPR} \quad \langle \rangle \end{array} \right] \\ \text{INDEX} \quad \textit{s} \\ \text{RESTR} \quad \langle \rangle \end{array} \right] \left[\begin{array}{l} \text{S} \\ \text{INDEX} \quad \textit{s} \end{array} \right] \end{array} \right]$$

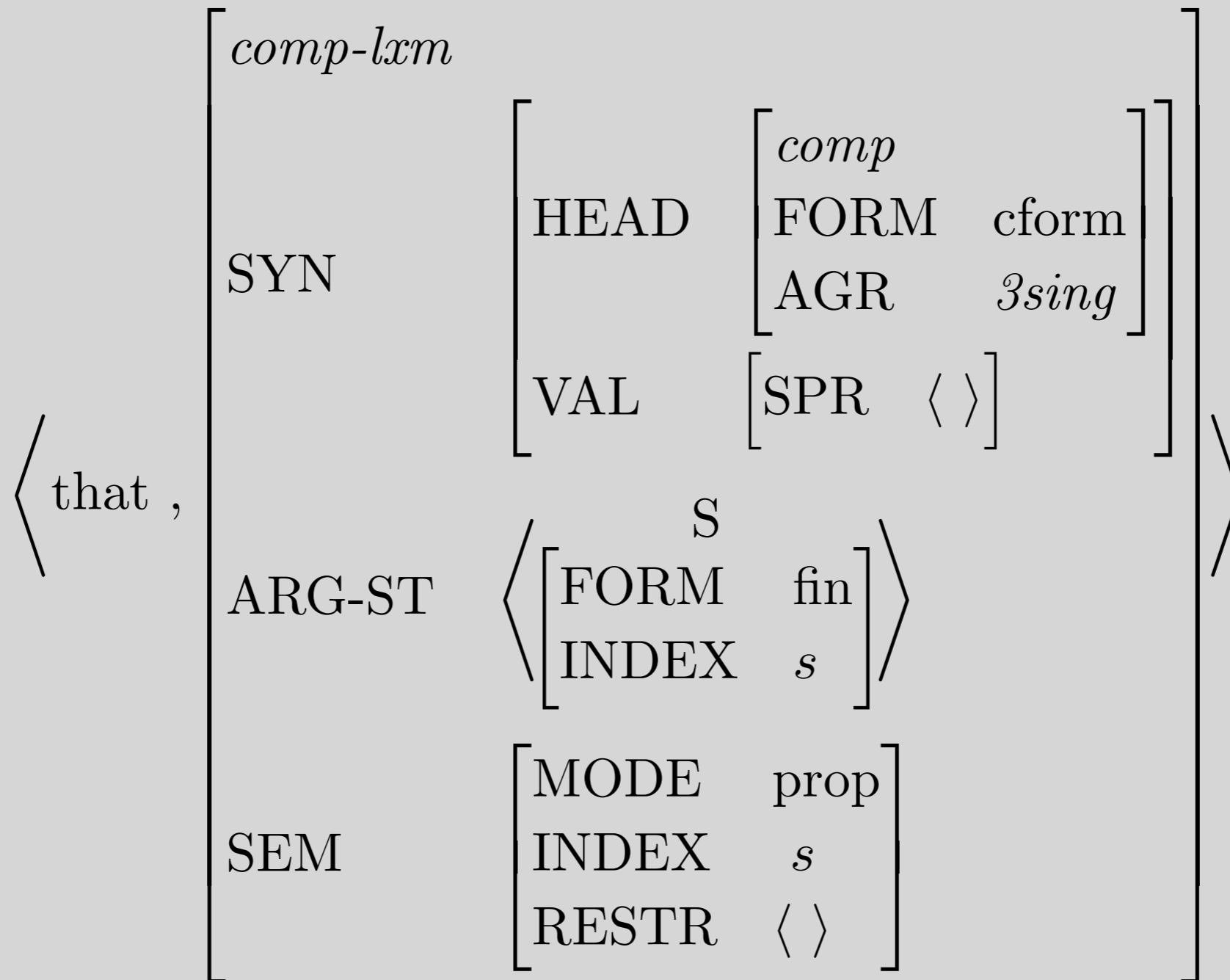
The Type *comp*



The Lexical Entry for Complementizer *that*

$$\left\langle \text{that} , \begin{bmatrix} \textit{comp-lxm} \\ \text{ARG-ST} < [\text{FORM fin}] > \\ \text{SEM} & [\text{MODE prop}] \end{bmatrix} \right\rangle$$

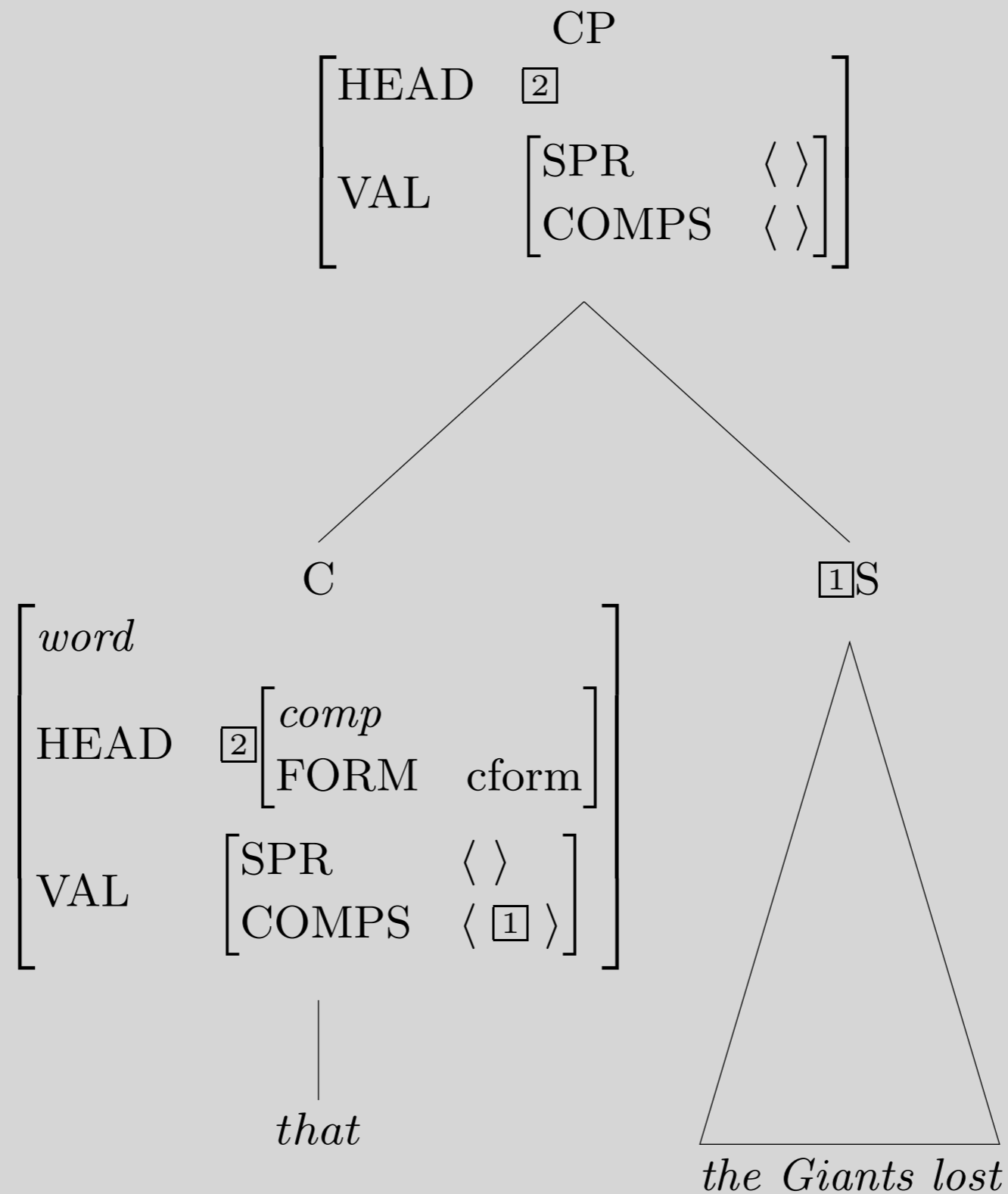
...and with inherited information filled in



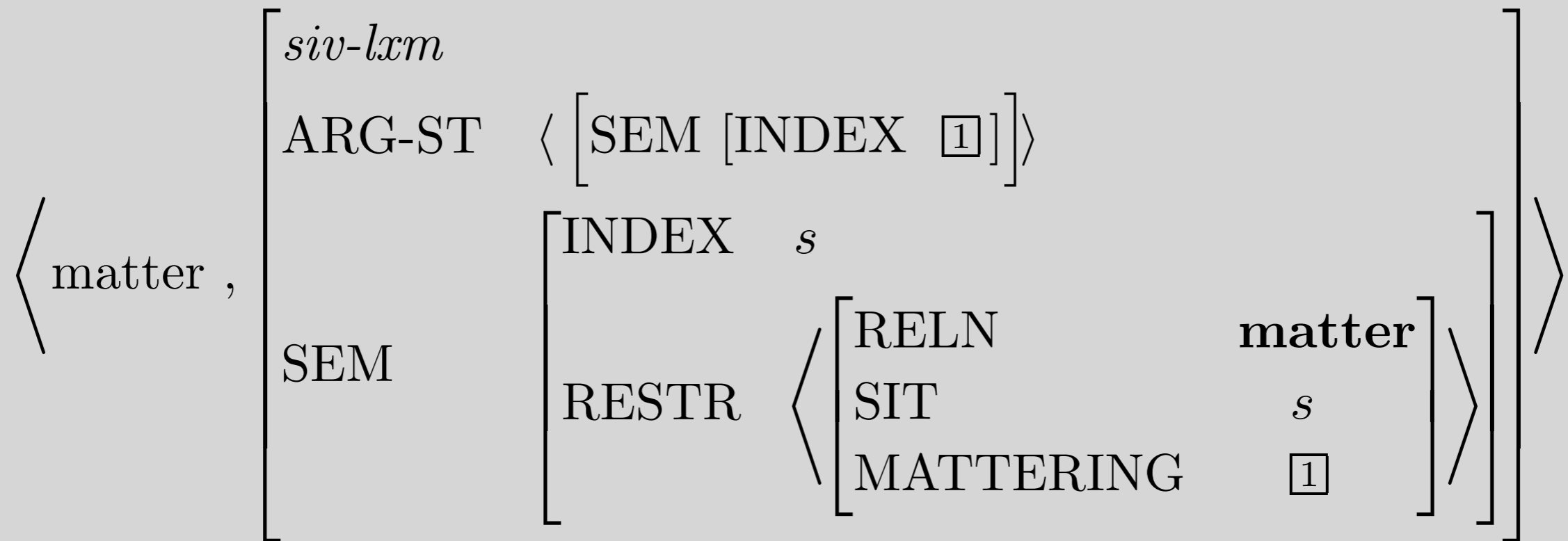
Question: Where did [FORM cform] come from?

Structure of a Complementizer Phrase

Poll!



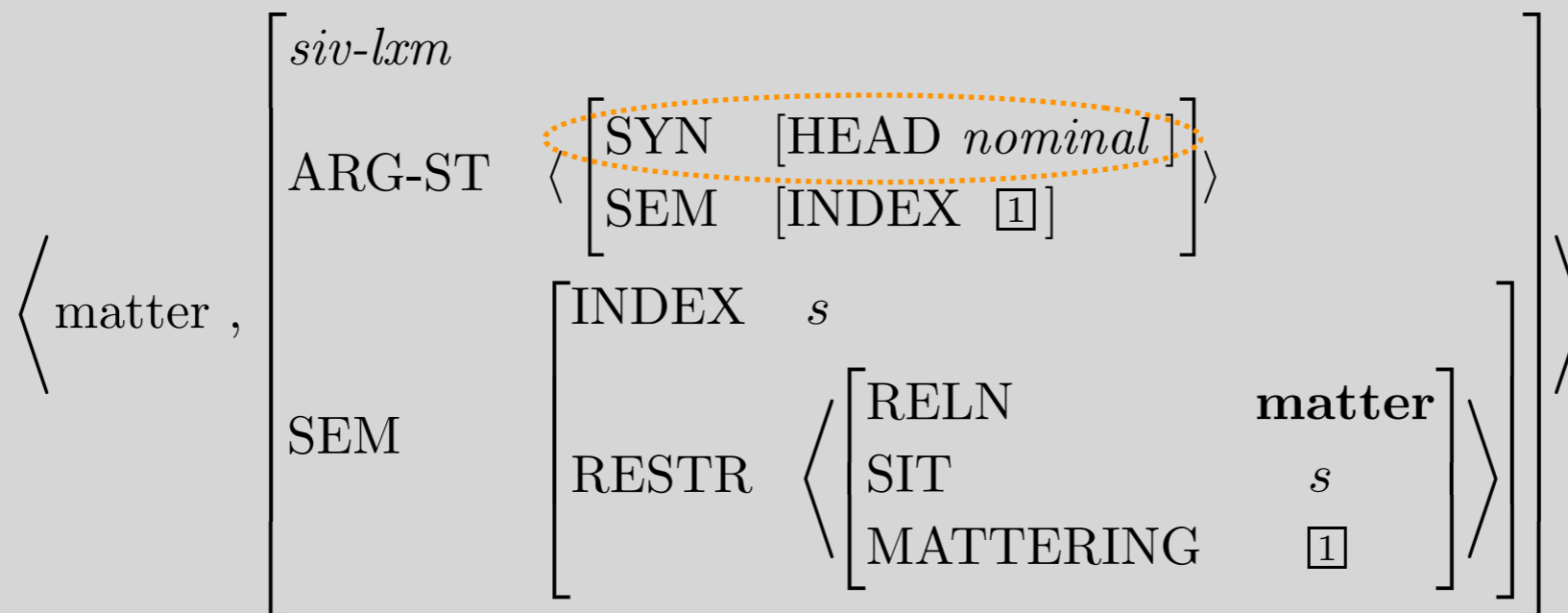
Sample Verb with a CP Subject



Note: the only constraint on the first argument is semantic

A Problem

- We constrained the subject of *matter* only semantically. However...
 - CP and S are semantically identical, but we get:
That Bush won matters vs. **Bush won matters*
 - Argument-marking PPs are semantically identical to their object NPs, but we get:
The election mattered vs. **Of the election mattered*
- So we need to add a syntactic constraint.

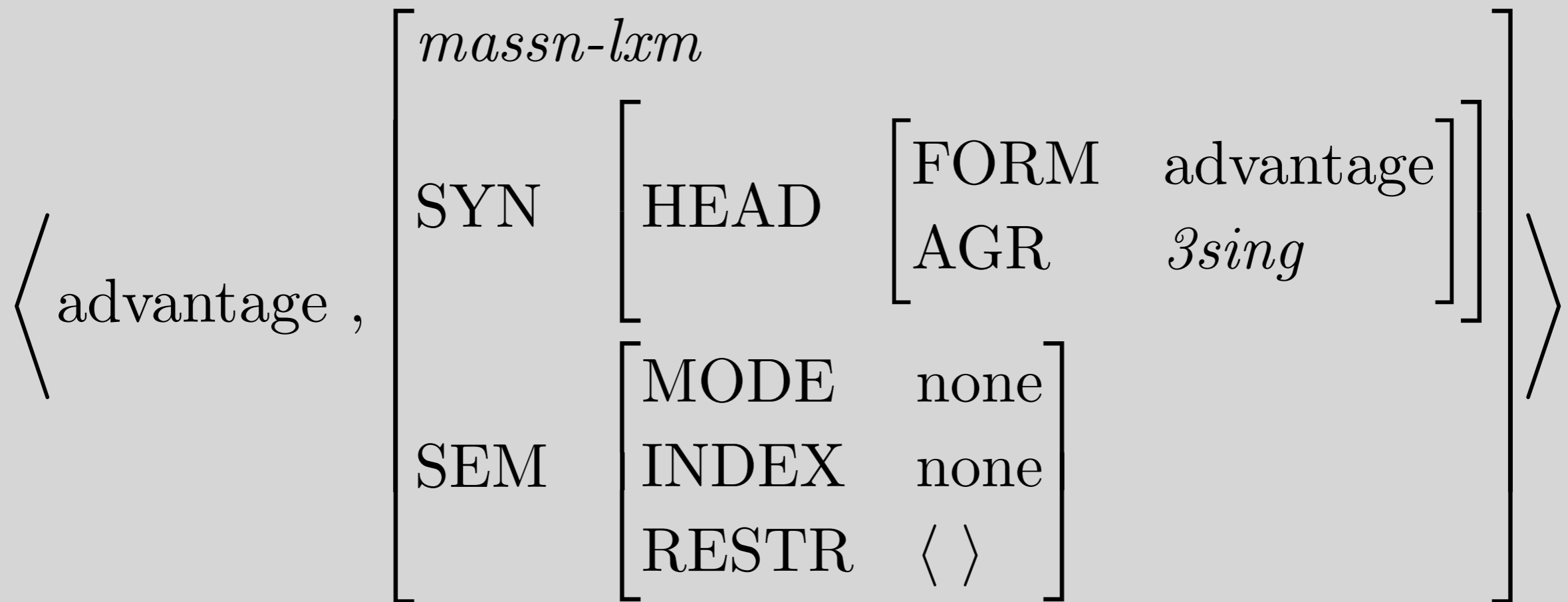


- S and PP subjects are generally impossible, so this constraint belongs on *verb-lxm*.

Extraposition with Verbs whose COMPS Lists are Nonempty

- *It worries me that war is imminent.*
- *It occurred to Pat that Chris knew the answer.*
- *It endeared you to Andy that you wore a funny hat.*

Another Nonreferential Noun



Our analyses of idioms and passives interact...

- We generate

Advantage was taken of the situation by many people.

Tabs are kept on online activists.

- But not:

Many people were taken advantage of.

- Why not?

Overview

- Existentials (*there, be*)
- Extraposition (*that, it, LR*)
- Idioms

Reading Questions

- How can we tell if a word has no referential potential and therefore no referential index?
- What are the deciding factors when determining that a rule needs to be a *pi-rule*?
- Why complementizers (that, whether, if, etc.) are AGR 3sing, and why can't the AGR value be underspecified?

Reading Questions

- I am curious to know what we would be violating if we modify our *be-lxm* to allow an "optional" third argument, instead of creating a separate entry for existential be?

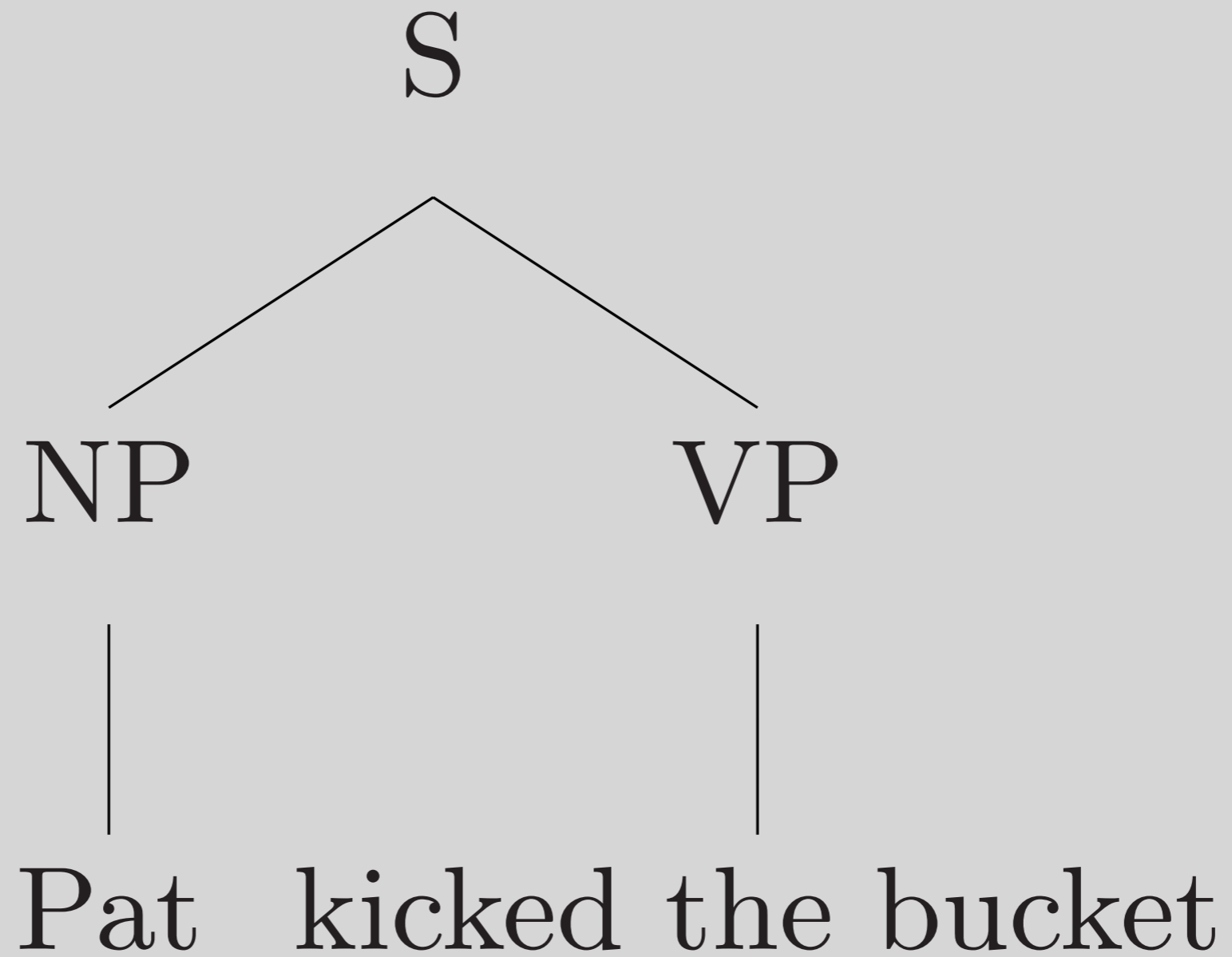
Reading Questions

- When it says that we can make [PRED -] a constraint of verb-lxm, does this mean it is a defeasible constraint? If not, why can't we make it a defeasible constraint? Also, why do inflectional rules not affect the PRED specification? I'm unclear on this generalization.
- I'm surprised that PRED is listed directly under pos in the type hierarchy, along side FORM rather than somewhere more specific. The difference to me is that FORM seems to be applicable to all parts of speech, while at least in this chapter, we're mainly shown how PRED is used with verbs and adjectives, and I can't imagine how could be for conjunctions or determiners. Are there more uses of PRED that we will learn?

Reading Questions

- I'm also curious about why this grammar would try to capture idiomatic expressions at all. Now that there is precedence for giving a single lexical entry for an entire idiom like "kick the bucket", couldn't we do that for any other tricky phrases that come up?
- The approach in 11.5 to idioms does not seem scalable. Due to the specificity of this kind of phrase, though, is there no way around hard-coding idioms like "kick the bucket"?
- How would the lexical entry for 'kick the bucket' in example (51) be used in creating a tree structure for a phrase like 'Pat kicked the bucket'?

POLL!



Reading Questions

- I'm not entirely convinced by the "kick the bucket" argument, either. I can imagine a sentence where one might infix an adjective, like "Well John, he kicked the old bucket a few months back," which isn't possible with this hard-coding of individual words inside the idiom. (Though I guess we could treat this as some kind of infixation process like f***-insertion inside words?)
- If "kick the bucket" can be hardcoded, why we need to have separate entries for other idioms such as "keep tabs on"?
- By hardcoding the idioms, will sentence like "Carrie kicked the bucket angrily" be interpreted as "Carrie died angrily" since "kick the bucket" may have priority in methods such as Maximum match theory. Should we produce to separate syntax reading instead?

Reading Questions

- If we allow multiple entries for an identical string in our lexicon, how do we resolve the lexical entries of such strings in the implementation of HPSG? I can imagine the contexts (i.e. other resolved entries) can help to make the decision but does it mean that the HPSG parser will have to process lexical entries in a specific order?

Reading Questions

- In 11.5, is there a general rule which types of idioms can have passive forms and which cannot (beyond entering all constituting words into a lexical entry)?
- The section on idioms sounds very interesting, but how does it scale as the language might evolve? Also, how does the feature structure encode subtle semantics like euphemism (e.g. *pass away* vs. *die*)?

Reading Questions

- I'm curious how the analysis of idioms in this chapter would treat the idiom *keep an eye on*. Assuming it has a passive form, what would the FORM values be for the NP *an eye* as well as *an* and *eye* individually? This seemed straightforward in the example *keep tabs on* when the NP *tabs* was a single word, but I'm unsure of how to handle a case with an NP with more than one word.

Reading Questions

- I was curious about the idiom, *put a lid on* (something), this cannot be expressed in a passive form, but is also a transitive verb, *he put a lid on the complainers*, would this be identical to the *kick the bucket* entry with another RESTR value in the list? If this were in the imperative form, *put a lid on it!*, would we try to mirror the imperative form values discussed in earlier chapters, or would this differ?

Reading Questions

- The reading mentions that it can be argued that selectional restrictions for verbs are semantic in nature. What are the arguments for this view? Does this mean that the constraint would only have to be expressed in the semantics, and not in the syntax?

Reading Questions

- When discussing that not all transitive verbs take that-clause complements, it was mentioned that there are verbs that can take CP complements but not NPs. And "it might plausibly be argued that these selectional restrictions are semantic in nature, so that this constraint need not be specified in their ARG-ST values." I expected the footnote to expand a bit on this, but instead it just mentioned further complications. I'm interested in how we can have semantic restrictions to resolve what complements a word can take. My thinking is that a lexical entry for a word like hope would have a RESTR that has a SIT feature and not a INDEX feature, so hope isn't compatible with NPs because the SIP would mean the NP would have an index instead of a situation, which would be incompatible with hope's RESTR needing to be matched to another situation.

Reading Questions

- While in a previous chapter there were a lot of terms with "lexical" in them, this time there's a lot of "semantic" terms. Semantic embedding, semantic complex, semantic analysis, along with familiar terms such as semantic structures, and semantic index. Could we review these the way we did with all the "lexical" terms?

Reading Questions

- Footnote 4 says that "our use of FORM values may seem somewhat promiscuous. In actual practice, however, we believe that the number of words entering into such morphologically-sensitive co-occurrence relations in any language is quite manageable." Who might it seem promiscuous to, and what would those people prefer we do? Are the FORM values on idiom verbs also part of this potential criticism? Are idioms approximately as closed of a class as the word 'be'?

Reading Questions

- I'm still not sure I understand why "FORM" is used outside of verbs. It feels natural to me for verbs, but when we created "nform," "aform," and others, it kinda just felt like a hack because we had chosen to put in on pos instead of just verbs in order to have only one coordination rule. I think the consequences of that decision are why the use of FORM values may feel "promiscuous" to some readers. Were any other alternatives investigated? FORM is probably the feature that feels the least intuitive to me except when used for verbs.

Reading Questions

- In some cases could we ever see the FORM value for the verb of an idiom overlap with itself? I.e. "turn over a new leaf" vs "turn the tables", vs "turn the tide of the battle" do all of these have FORM turn and just have separate lexical entries with different SEM values?

Reading Questions

- This chapter seems to be taking a very specific and idiosyncratic aspect of English and finding ways to design our lexical rules such that they account for it. Are there procedures for approaching phenomena that might not easily fit into the grammar, both in English and across languages?