Ling 566
Nov 8, 2012

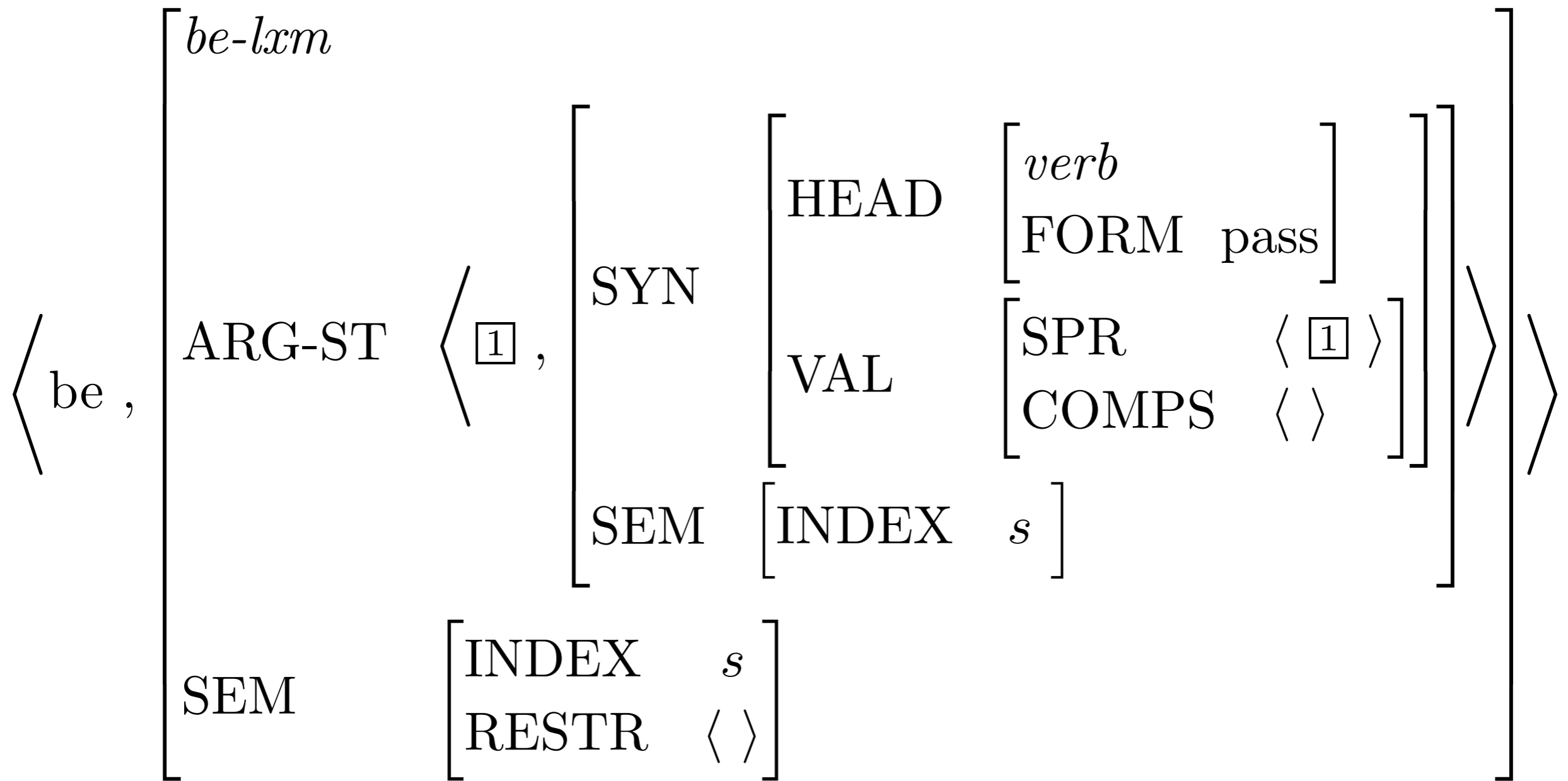## Non-referential NPs, Expletives, and Extraposition

Overview
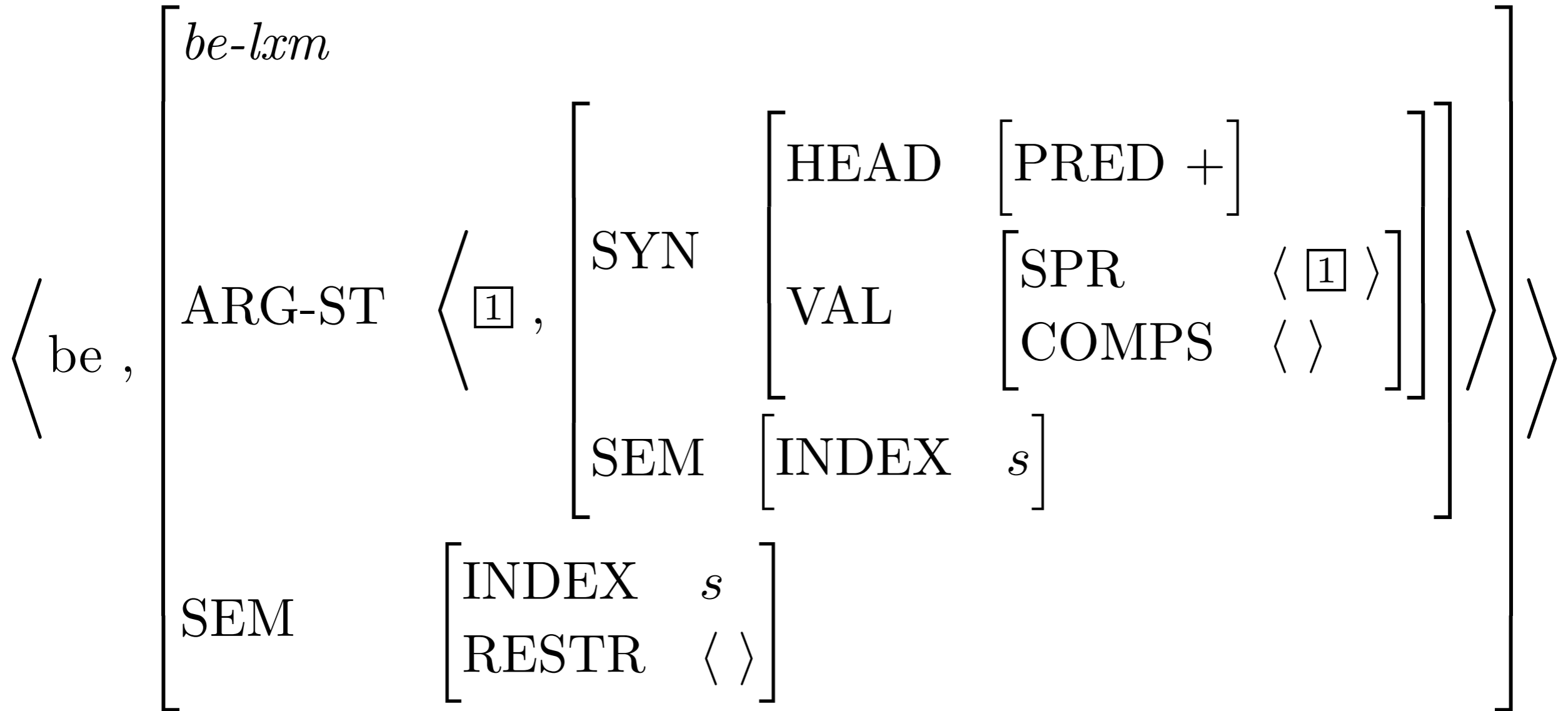
- Existentials

- 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.
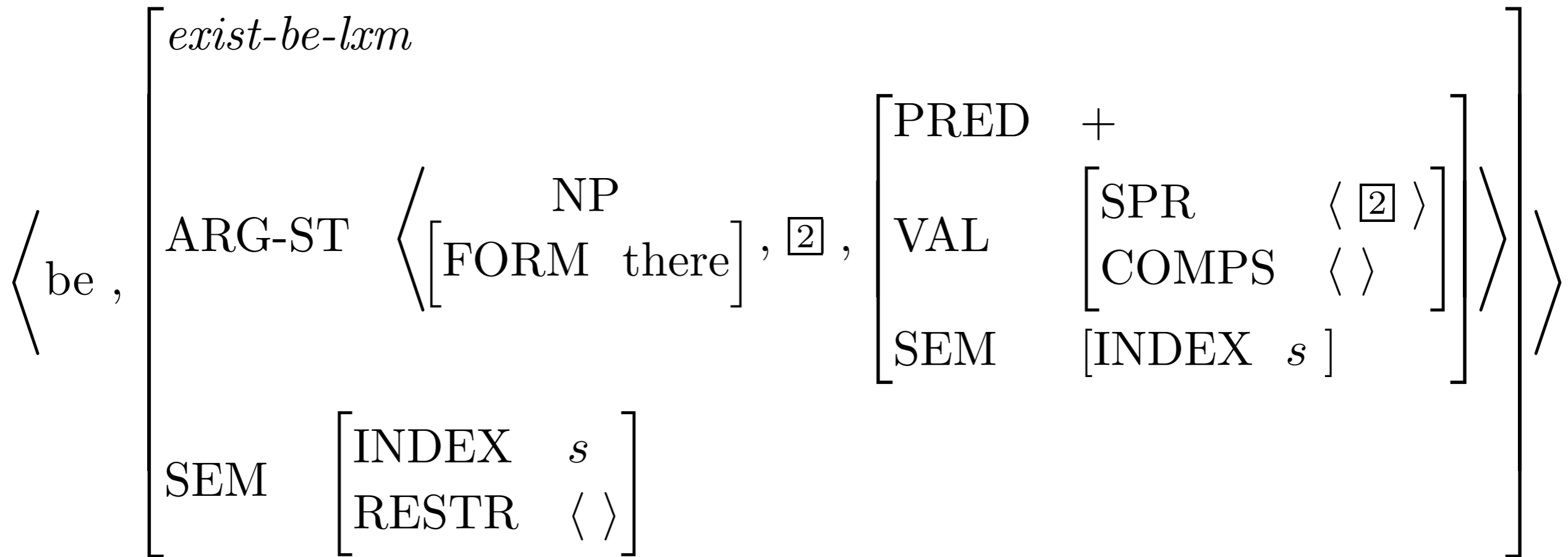
Chapter 10 entry for *be*

$$
\left\langle \text{be} ,
\begin{bmatrix}
\textit{be-lxm} \\[2ex]
\text{ARG-ST} \quad \left\langle \boxed{1} ,
\begin{bmatrix}
\text{SYN}
\begin{bmatrix}
\text{HEAD}
\begin{bmatrix}
\textit{verb} \\
\text{FORM} \quad \text{pass}
\end{bmatrix} \\[2ex]
\text{VAL}
\begin{bmatrix}
\text{SPR} \quad \langle \boxed{1} \rangle \\
\text{COMPS} \quad \langle \, \rangle
\end{bmatrix}
\end{bmatrix} \\[4ex]
\text{SEM} \begin{bmatrix} \text{INDEX} \quad s \end{bmatrix}
\end{bmatrix}
\right\rangle \\[6ex]
\text{SEM}
\begin{bmatrix}
\text{INDEX} \quad s \\
\text{RESTR} \quad \langle \, \rangle
\end{bmatrix}
\end{bmatrix}
\right\rangle
$$

Copula (generalized)

$$
\left\langle \text{be} ,\ \begin{bmatrix} \textit{be-lxm} \\[2pt] \text{ARG-ST} \quad \left\langle \boxed{1} ,\ \begin{bmatrix} \text{SYN} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{PRED} + \end{bmatrix} \\[6pt] \text{VAL} & \begin{bmatrix} \text{SPR} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \ \rangle \end{bmatrix} \end{bmatrix} \\[12pt] \text{SEM} & \begin{bmatrix} \text{INDEX} & s \end{bmatrix} \end{bmatrix} \right\rangle \\[12pt] \text{SEM} \quad \begin{bmatrix} \text{INDEX} & s \\ \text{RESTR} & \langle \ \rangle \end{bmatrix} \end{bmatrix} \right\rangle
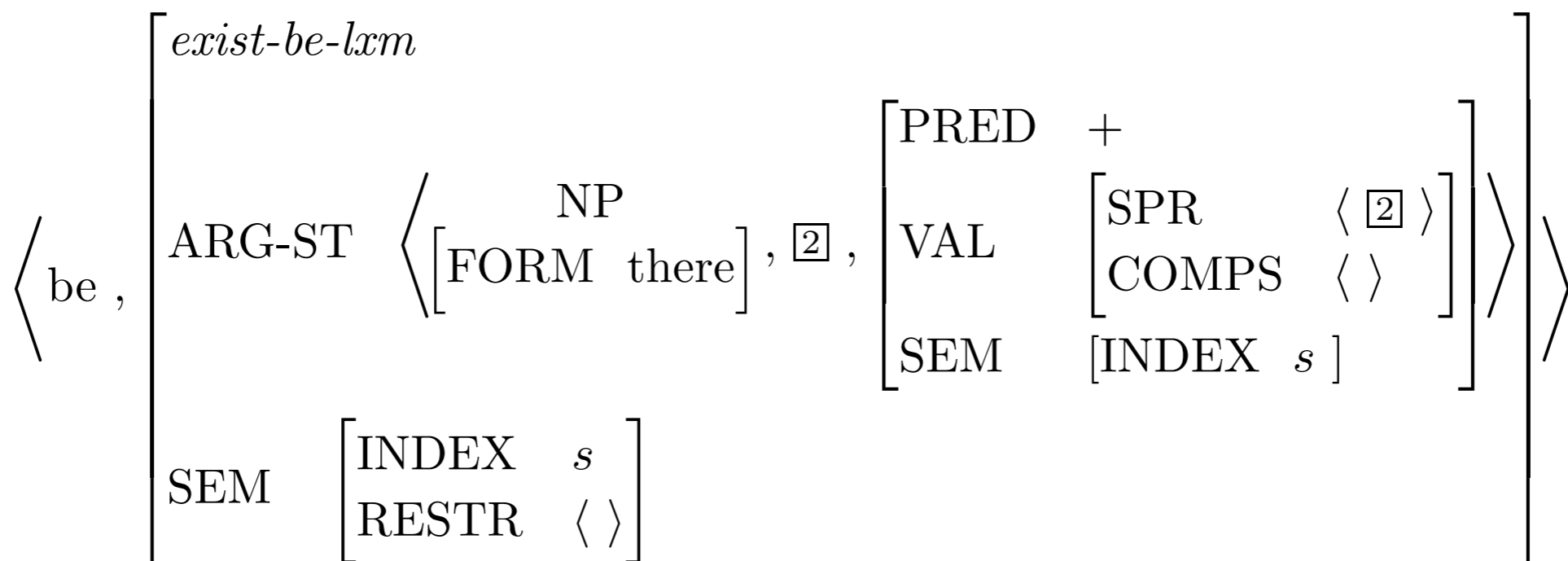$$

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*

$$
\left\langle \text{be} \, , \,
\begin{bmatrix}
\textit{exist-be-lxm} \\[4pt]
\text{ARG-ST} \quad
\left\langle
\begin{bmatrix} \text{NP} \\ \text{FORM} \quad \text{there} \end{bmatrix} , \; \boxed{2} \, , \;
\begin{bmatrix}
\text{PRED} \quad + \\
\text{VAL} \quad
\begin{bmatrix} \text{SPR} \quad \langle \, \boxed{2} \, \rangle \\ \text{COMPS} \quad \langle \, \rangle \end{bmatrix} \\
\text{SEM} \quad [\text{INDEX} \; s \,]
\end{bmatrix}
\right\rangle \\[20pt]
\text{SEM} \quad
\begin{bmatrix} \text{INDEX} \quad s \\ \text{RESTR} \quad \langle \, \rangle \end{bmatrix}
\end{bmatrix}
\right\rangle
$$

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?
- How do we rule out **There was a greyhound a good runner*?

$$
\left\langle \text{be} , \begin{bmatrix} \textit{exist-be-lxm} \\[4pt] \text{ARG-ST} \quad \left\langle \begin{bmatrix} \text{NP} \\ \text{FORM} \quad \text{there} \end{bmatrix}, \; \boxed{2} , \; \begin{bmatrix} \text{PRED} & + \\ \text{VAL} & \begin{bmatrix} \text{SPR} & \langle \boxed{2} \rangle \\ \text{COMPS} & \langle \, \rangle \end{bmatrix} \\ \text{SEM} & [\text{INDEX} \; s] \end{bmatrix} \right\rangle \\[4pt] \text{SEM} \quad \begin{bmatrix} \text{INDEX} & s \\ \text{RESTR} & \langle \, \rangle \end{bmatrix} \end{bmatrix} \right\rangle
$$

The Entry for Existential *there*

$$
\left\langle \text{there} \,, \begin{bmatrix} \textit{pron-lxm} \\[4pt] \text{SYN} \quad \begin{bmatrix} \text{HEAD} \quad \begin{bmatrix} \text{FORM} \quad \text{there} \\ \text{AGR} \quad \begin{bmatrix} \text{PER} \quad \text{3rd} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\[20pt] \text{SEM} \quad \begin{bmatrix} \text{MODE} \quad \text{none} \\ \text{INDEX} \quad \text{none} \\ \text{RESTR} \quad \langle\, \rangle \end{bmatrix} \end{bmatrix} \right\rangle
$$

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?

$$\left\langle \text{there}, \begin{bmatrix} \textit{pron-lxm} \\ \\ \text{SYN} \quad \begin{bmatrix} \text{HEAD} \quad \begin{bmatrix} \text{FORM} & \text{there} \\ \text{AGR} & [\text{PER} \quad \text{3rd}] \end{bmatrix} \end{bmatrix} \\ \\ \text{SEM} \quad \begin{bmatrix} \text{MODE} & \text{none} \\ \text{INDEX} & \text{none} \\ \text{RESTR} & \langle \, \rangle \end{bmatrix} \end{bmatrix} \right\rangle$$

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*

$$\left\langle \text{it,} \begin{bmatrix} \textit{pron-lxm} \\[2ex] \text{SYN} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{FORM} & \text{it} \\ \text{AGR} & \textit{3sing} \end{bmatrix} \end{bmatrix} \\[4ex] \text{SEM} & \begin{bmatrix} \text{MODE} & \text{none} \\ \text{INDEX} & \text{none} \\ \text{RESTR} & \langle\,\rangle \end{bmatrix} \end{bmatrix} \right\rangle$$

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*?

$$\left\langle \text{it,} \begin{bmatrix} pron\text{-}lxm \\ \\ \text{SYN} \quad \begin{bmatrix} \text{HEAD} \quad \begin{bmatrix} \text{FORM} & \text{it} \\ \text{AGR} & 3sing \end{bmatrix} \end{bmatrix} \\ \\ \text{SEM} \quad \begin{bmatrix} \text{MODE} & none \\ \text{INDEX} & none \\ \text{RESTR} & \langle \ \rangle \end{bmatrix} \end{bmatrix} \right\rangle$$

A New Type of Lexeme: Complementizers

$$
\textit{comp-lxm} : \begin{bmatrix} \text{SYN} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{comp} \\ \text{AGR} & \textit{3sing} \end{bmatrix} \\ \text{VAL} & \begin{bmatrix} \text{SPR} & \langle\ \rangle \end{bmatrix} \end{bmatrix} \\ \text{ARG-ST} & \left\langle \begin{matrix} \text{S} \\ \begin{bmatrix} \text{INDEX} & s \end{bmatrix} \end{matrix} \right\rangle \\ \text{SEM} & \begin{bmatrix} \text{INDEX} & s \\ \text{RESTR} & \langle\ \rangle \end{bmatrix} \end{bmatrix}
$$

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?

$$
\textit{comp-lxm} : \begin{bmatrix} \text{SYN} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{comp} \\ \text{AGR} & \textit{3sing} \end{bmatrix} \\ \text{VAL} & \begin{bmatrix} \text{SPR} & \langle \, \rangle \end{bmatrix} \end{bmatrix} \\ \text{ARG-ST} & \left\langle \begin{array}{c} \text{S} \\ \begin{bmatrix} \text{INDEX} & s \end{bmatrix} \end{array} \right\rangle \\ \text{SEM} & \begin{bmatrix} \text{INDEX} & s \\ \text{RESTR} & \langle \, \rangle \end{bmatrix} \end{bmatrix}
$$

The Type *comp*

pos

$$\begin{bmatrix} \text{FORM, PRED} \end{bmatrix}$$

agr-pos

$$\begin{bmatrix} \text{AGR} \end{bmatrix}$$

adj   *prep*   *adv*   *conj*

verb

$$\begin{bmatrix} \text{AUX} \end{bmatrix}$$

nominal

$$\begin{bmatrix} \text{CASE} \end{bmatrix}$$

det

$$\begin{bmatrix} \text{COUNT} \end{bmatrix}$$

noun   *comp*

$$\begin{bmatrix} \text{FORM} & \text{cform} \end{bmatrix}$$
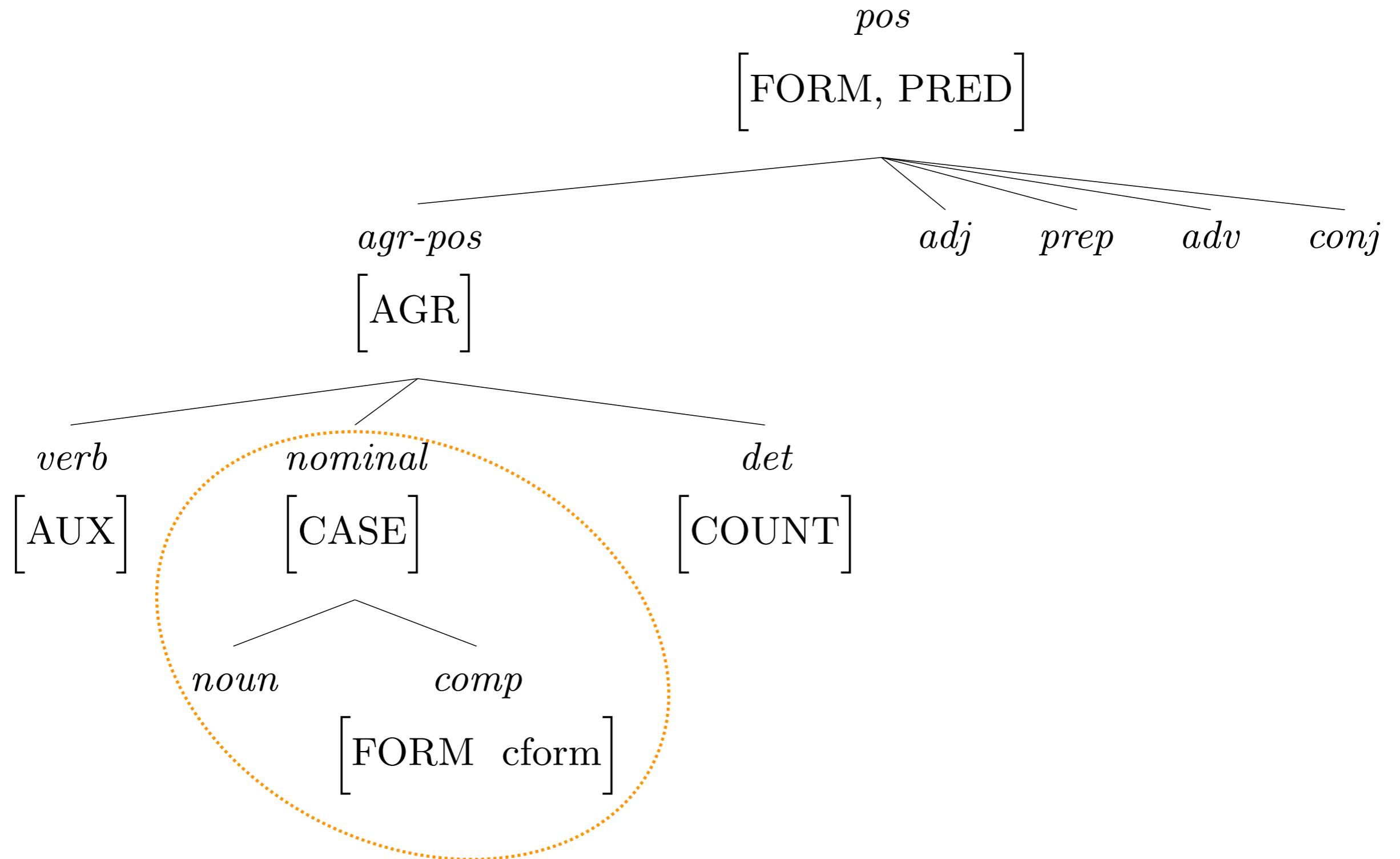
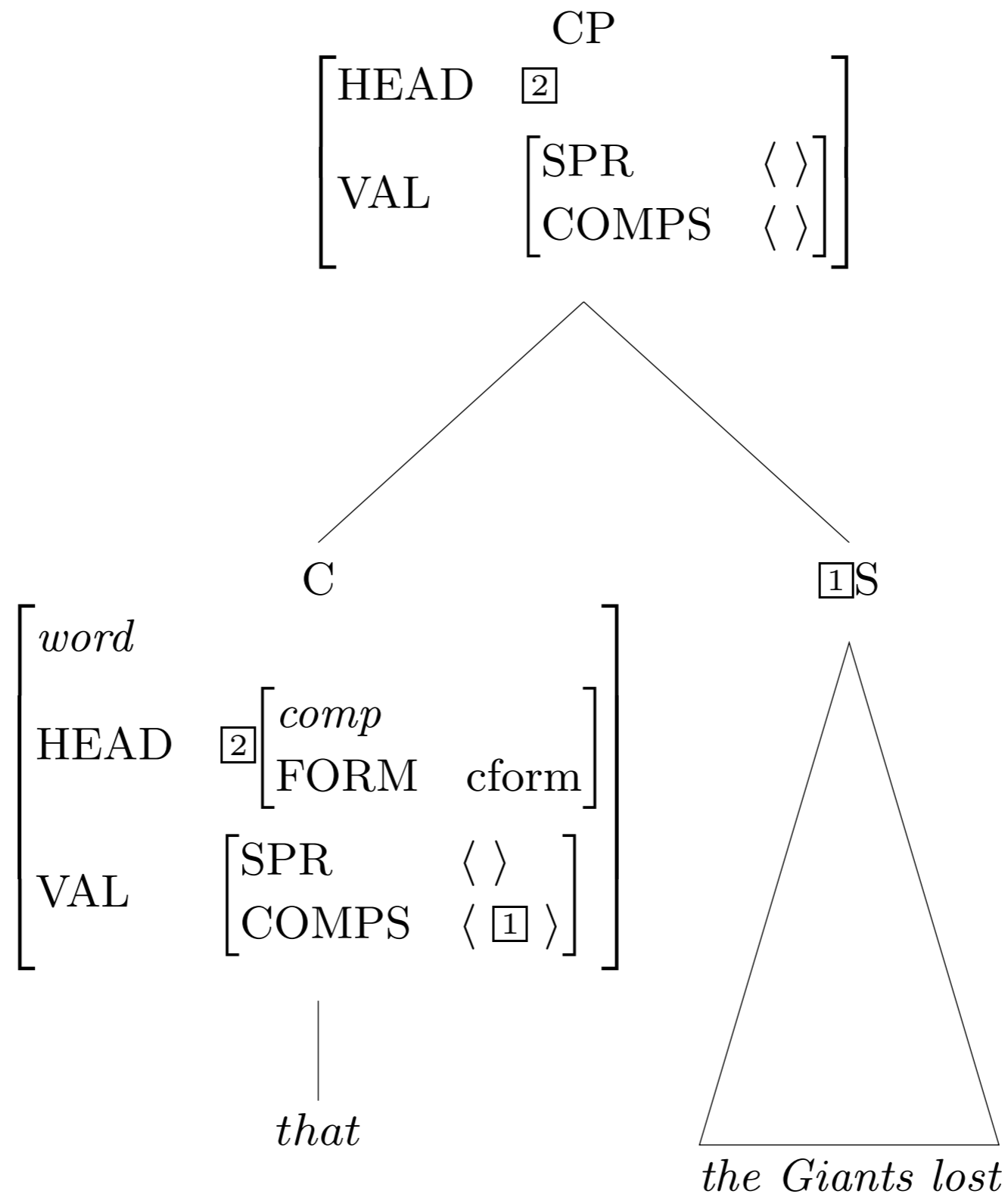# The Lexical Entry for Complementizer *that*

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

…and with inherited information filled in

$$
\left\langle \text{that}, \begin{bmatrix} \textit{comp-lxm} \\ \text{SYN} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{comp} \\ \text{FORM} & \text{cform} \\ \text{AGR} & \textit{3sing} \end{bmatrix} \\ \text{VAL} & \begin{bmatrix} \text{SPR} & \langle \ \rangle \end{bmatrix} \end{bmatrix} \\ \text{ARG-ST} & \left\langle \begin{bmatrix} \text{S} \\ \text{FORM} & \text{fin} \\ \text{INDEX} & s \end{bmatrix} \right\rangle \\ \text{SEM} & \begin{bmatrix} \text{MODE} & \text{prop} \\ \text{INDEX} & s \\ \text{RESTR} & \langle \ \rangle \end{bmatrix} \end{bmatrix} \right\rangle
$$
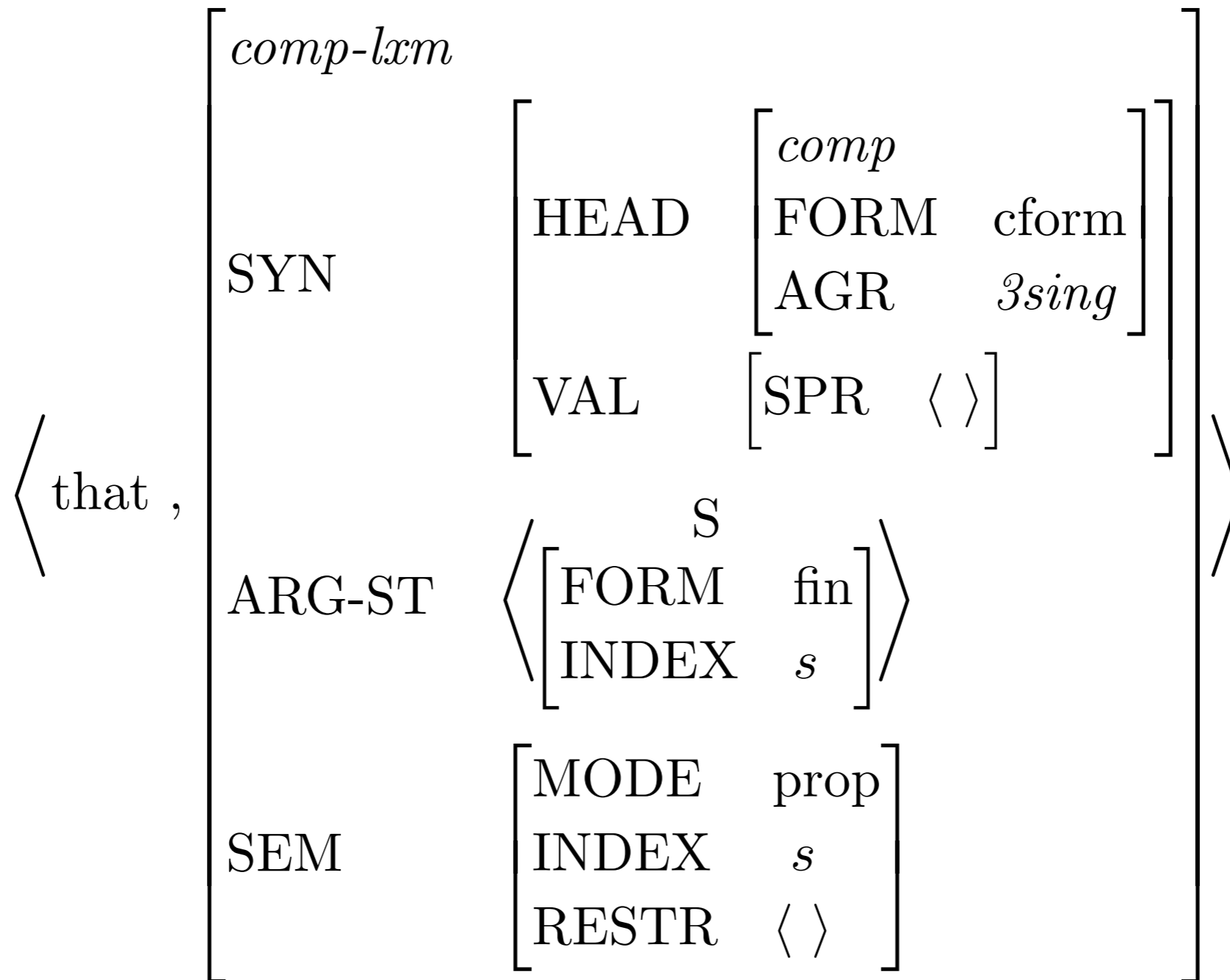
Question: Where did [FORM cform] come from?

Structure of a Complementizer Phrase

CP

$$\begin{bmatrix} \text{HEAD} & \boxed{2} \\ \text{VAL} & \begin{bmatrix} \text{SPR} & \langle\,\rangle \\ \text{COMPS} & \langle\,\rangle \end{bmatrix} \end{bmatrix}$$

C

$$\begin{bmatrix} \textit{word} \\ \text{HEAD} & \boxed{2}\begin{bmatrix} \textit{comp} \\ \text{FORM} & \text{cform} \end{bmatrix} \\ \text{VAL} & \begin{bmatrix} \text{SPR} & \langle\,\rangle \\ \text{COMPS} & \langle\,\boxed{1}\,\rangle \end{bmatrix} \end{bmatrix}$$

$\boxed{1}$S

that

the Giants lost

Sample Verb with a CP Subject

$$
\left\langle \text{matter} \;,\; \begin{bmatrix} \textit{siv-lxm} \\[6pt] \text{ARG-ST} \quad \left\langle \begin{bmatrix} \text{SEM } [\text{INDEX} \quad \boxed{1}] \end{bmatrix} \right\rangle \\[12pt] \text{SEM} \quad \begin{bmatrix} \text{INDEX} \quad s \\[12pt] \text{RESTR} \quad \left\langle \begin{bmatrix} \text{RELN} & \textbf{matter} \\ \text{SIT} & s \\ \text{MATTERING} & \boxed{1} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \right\rangle
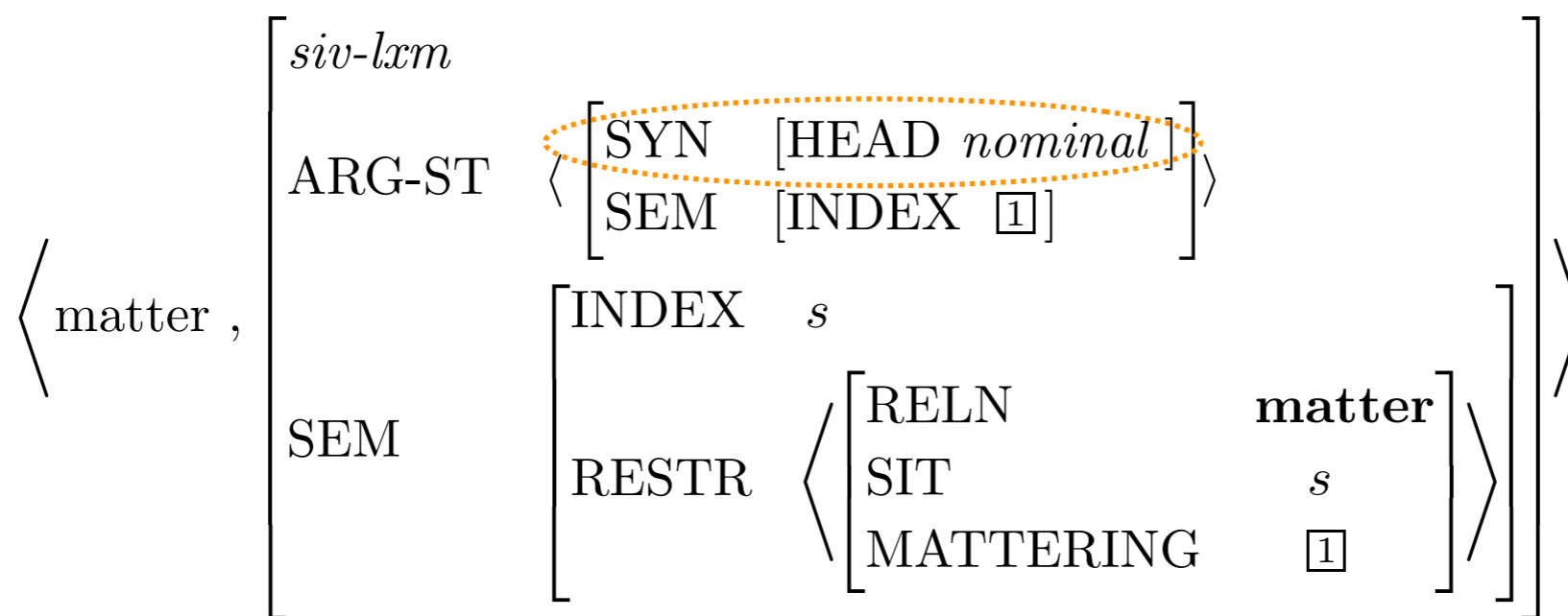$$
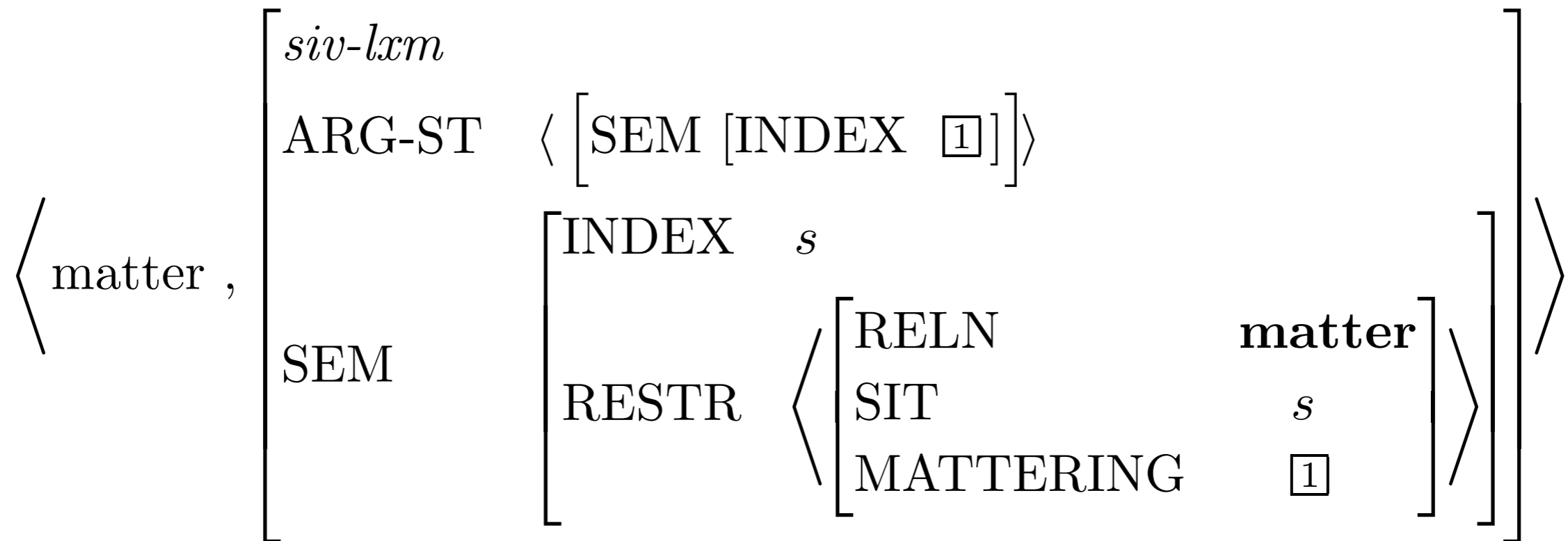
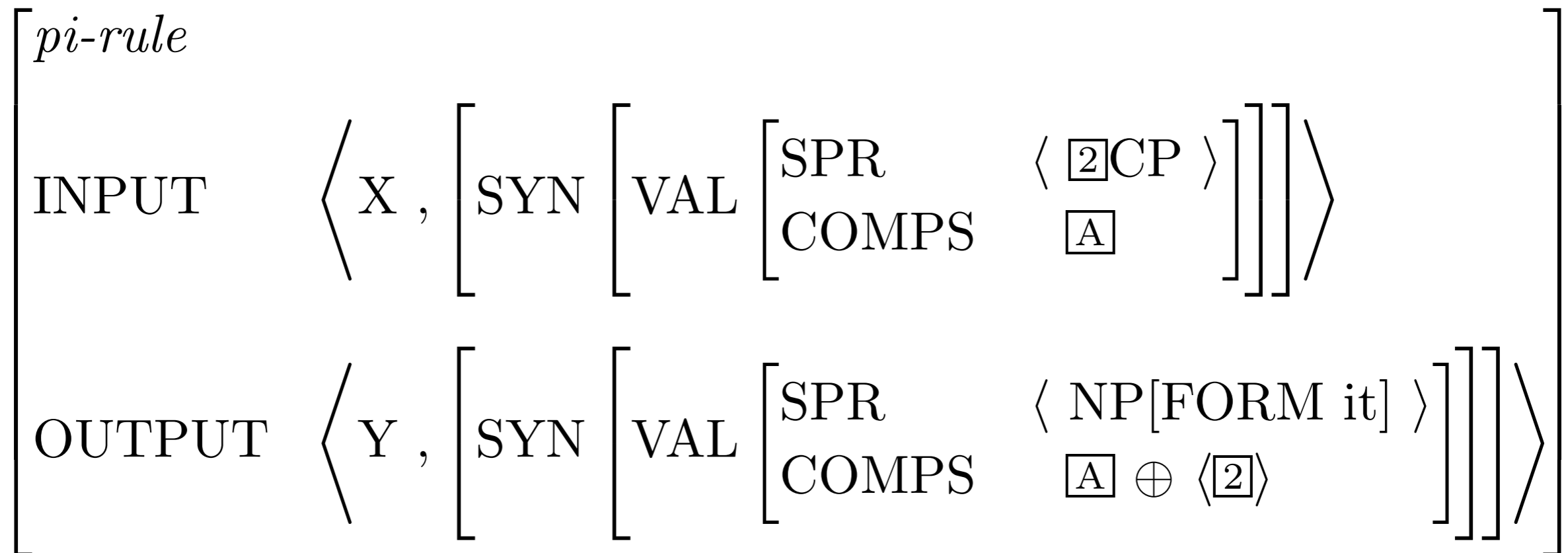Note: the only constraint on the first argument is semantic

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.

$$
\left\langle \text{matter},
\begin{bmatrix}
\textit{siv-lxm} \\[4pt]
\text{ARG-ST} \quad \left\langle
\begin{bmatrix}
\text{SYN} & [\text{HEAD } \textit{nominal}] \\
\text{SEM} & [\text{INDEX } \boxed{1}]
\end{bmatrix}
\right\rangle \\[10pt]
\text{SEM} \quad
\begin{bmatrix}
\text{INDEX} & s \\[4pt]
\text{RESTR} & \left\langle
\begin{bmatrix}
\text{RELN} & \textbf{matter} \\
\text{SIT} & s \\
\text{MATTERING} & \boxed{1}
\end{bmatrix}
\right\rangle
\end{bmatrix}
\end{bmatrix}
\right\rangle
$$

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

The Extraposition Lexical Rule

$$
\begin{bmatrix}
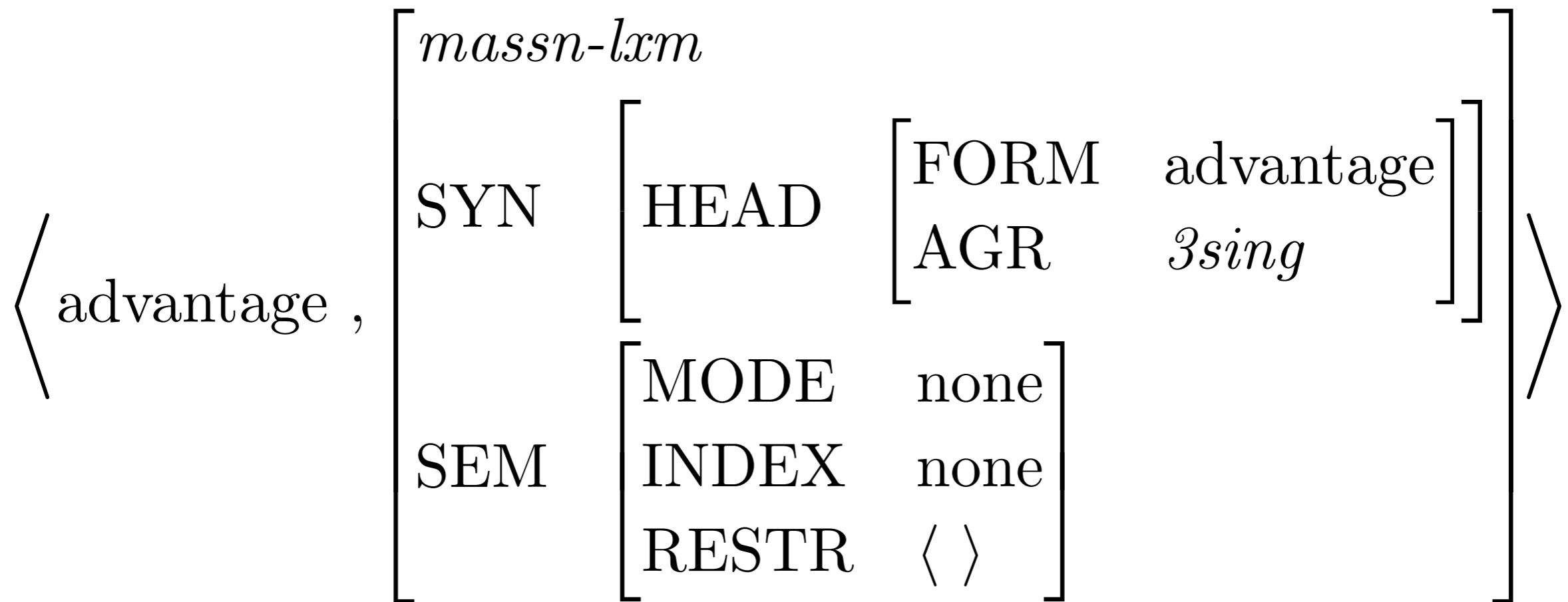pi\text{-}rule \\
\\
\text{INPUT} \quad \left\langle \text{X} , \begin{bmatrix} \text{SYN} \begin{bmatrix} \text{VAL} \begin{bmatrix} \text{SPR} & \langle\, \boxed{2}\text{CP}\, \rangle \\ \text{COMPS} & \boxed{A} \end{bmatrix} \end{bmatrix} \end{bmatrix} \right\rangle \\
\\
\text{OUTPUT} \quad \left\langle \text{Y} , \begin{bmatrix} \text{SYN} \begin{bmatrix} \text{VAL} \begin{bmatrix} \text{SPR} & \langle\, \text{NP[FORM it]}\, \rangle \\ \text{COMPS} & \boxed{A} \oplus \langle\boxed{2}\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \right\rangle
\end{bmatrix}
$$

- Why is the type *pi-rule*?
- Why doesn't it say anything about the semantics?
- Why is the COMPS value $\boxed{A}$, not < >?
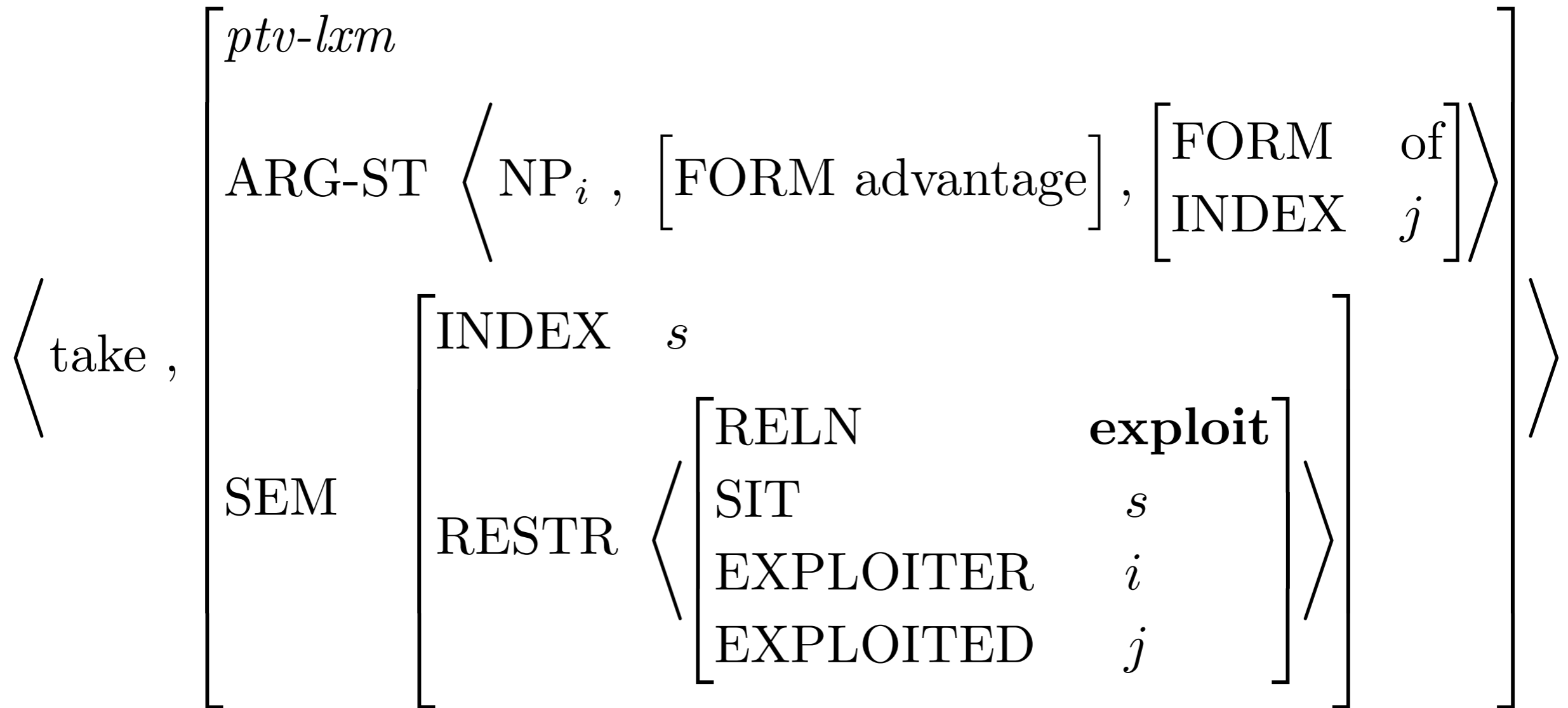
Extraposition with Verbs whose COMPS Lists are Nonempty

- *It worries <u>me</u> that war is imminent.*

- *It occurred <u>to Pat</u> that Chris knew the answer.*

- *It endeared <u>you</u> <u>to Andy</u> that you wore a funny hat.*

Another Nonreferential Noun

$$\left\langle \text{advantage} \, , \begin{bmatrix} \textit{massn-lxm} \\[4pt] \text{SYN} \begin{bmatrix} \text{HEAD} \begin{bmatrix} \text{FORM} & \text{advantage} \\ \text{AGR} & \textit{3sing} \end{bmatrix} \end{bmatrix} \\[20pt] \text{SEM} \begin{bmatrix} \text{MODE} & \text{none} \\ \text{INDEX} & \text{none} \\ \text{RESTR} & \langle \, \rangle \end{bmatrix} \end{bmatrix} \right\rangle$$

The Verb that Selects *advantage*

$$
\left\langle \text{take} \ , \
\begin{bmatrix}
\textit{ptv-lxm} \\[2ex]
\text{ARG-ST} \ \left\langle \text{NP}_i \ , \ \begin{bmatrix} \text{FORM advantage} \end{bmatrix} , \ \begin{bmatrix} \text{FORM} & \text{of} \\ \text{INDEX} & j \end{bmatrix} \right\rangle \\[3ex]
\text{SEM} \ \begin{bmatrix}
\text{INDEX} & s \\[1ex]
\text{RESTR} \ \left\langle \begin{bmatrix}
\text{RELN} & \textbf{exploit} \\
\text{SIT} & s \\
\text{EXPLOITER} & i \\
\text{EXPLOITED} & j
\end{bmatrix} \right\rangle
\end{bmatrix}
\end{bmatrix}
\right\rangle
$$

Our analyses of idioms and passives interact...

- We generate

  Advantage was taken of the situation by many people.

  Tabs are kept on foreign students.

- But not:

  Many people were taken advantage of.

- Why not?

Overview

- Existentials (*there*, *be*)

- Extraposition (*that*, *it*, *LR*)

- Idioms

Reading Questions

- Is PRED a cop-out? Is there a deeper analysis of what's predicative and what's not?

- How do we handle empty copula examples like (6a,b)? A lexical rule that somehow makes the copula null?

    b. A   magyar   zászló piros-fehér-zőld.
      the Hungarian flag   red-white-green
      'The Hungarian flag is red, white, and green.'
      (Hungarian)

Reading Questions

- For (32) (revised definition of tv-lxm), should [HEAD nominal] be defeasible?

$$(32) \quad \begin{bmatrix} \text{ARG-ST} & \left\langle \text{X}, \begin{bmatrix} \text{HEAD} & nominal \\ \text{VAL} & \begin{bmatrix} \text{SPR} & \langle\ \rangle \\ \text{COMPS} & \langle\ \rangle \end{bmatrix} \end{bmatrix}, \dots \right\rangle \end{bmatrix}$$

Reading Questions

- In *That Sandy smokes matters*, what's the SPR of *matters* --- the CP or the NP? What is the specifier for *Sandy smokes* - the C?

- So, when there is no C present in, for instance,

  We think we're going to go skiing this weekend,

  There would still be a complementizer with a lexical entry, but the name would be empty, right?

Reading Questions

- "that" - Would a sentence like "That doesn't make sense." be counted as a complement phrase? What is the type for "that" in this case? Could "this" also be used to denote complement phrases? "Pat runs." -> "Pat runs this way."

Reading Questions

- Also, in the idioms section there's a reference to only tv-lxm being able to passivize. What do we do with examples like, *He was thought to be the cookie monster*?

Reading Questions

- Why does comp need CASE? If a verb is assigning nom case to its subject, is that affected? In the example, *We forgot that we needed invitations*, wouldn't *needed* assign nom case to *we*?

Reading Questions

- What does the pi-rule constraint on p. 346 tell us? It looks like it doesn't say anything, but I assume it does.

$$
\textit{pi-rule} : \begin{bmatrix} \text{INPUT} & \left\langle / \boxed{0} , \begin{bmatrix} \textit{word} \\ \text{SYN} \begin{bmatrix} \text{HEAD} & / \boxed{1} \\ \text{VAL} & \begin{bmatrix} \text{MOD} & \boxed{A} \end{bmatrix} \end{bmatrix} \end{bmatrix} \right\rangle \\ \\ \text{OUTPUT} & \left\langle / \boxed{0} , \begin{bmatrix} \textit{word} \\ \text{SYN} \begin{bmatrix} \text{HEAD} & / \boxed{1} \\ \text{VAL} & \begin{bmatrix} \text{MOD} & \boxed{A} \end{bmatrix} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}
$$

Reading Questions

- Can you explain the reasons that the Extraposition Lexical Rule is not formulated in terms of the ARG-ST the way the Passive Lexical Rule is?

Reading Questions

- For some reason, saying existential "it" doesn't have an INDEX feels wrong.

  (i) The Giants lost. It sucks.
  (ii) That the Giants lost sucks.
  (iii) It sucks that the Giants lost.

- To me it seems that the INDEX of "it" is the situation of "the Giants lost" and that it would be easier/more general to have a rule that allows it to refer to the situation and underspcify the FORM of CP (to avoid the Binding Theory problems that might arrive)...? I mean, we have sentences like "She sang very well, that woman in the red dress." (Though perhaps that's more colloquial than formal.)

Reading Questions

- Is the extraposition lexical rule or something like it involved in these examples?

I will join you if you come.

If you come, I will join you.

Reading Questions

- Is extraposition involved in it-clefts?

  It is Kim who read the book.

- If so, does 'who' here have the same function as 'that'?

Reading Questions

- Also, are things like the dummy *it* and *there* universals? Or widespread?

- How does pro-drop relate to dummy pronouns?

Reading Questions

- How do we handle agreement in existentials?

  There is a unicorn in the garden.

  There are unicorns in the garden.

Reading Questions

- There are at least a few other verbs that can also take an NP[FORM there] specifier:

  There remain several questions to be answered.
  There exist at least a few other verbs that can take 'there' as a specifier.
  There persist certain problems with this assumption.

- We formulated a new lexeme type for be, exist-be-lxm, in order to handle its co-occurrence with there. Should the same be done for remain, exist and persist? Or are they type exist-be-lxm?

Reading Questions

- It seems like having one lexical entry for be is another, maybe or maybe not better, option. My question is whether there is a good reason to create a second entry instead of having a lexical rule that transforms the one entry for be. Would also help with those other verbs...

Reading Questions

- The approach to idioms doesn't seem very elegant or scalable. Is FORM really the right way to go?

- The analyses in this chapter seem less general than the previous ones. Do the phenomena of passivization and/or existentials get reanalyzed somewhat for greater generality of the type hierarchy and FORM values?

Reading Questions

- Also, it seems to me like our list of potential values for type FORM is becoming quite large, almost like a second lexicon of sorts. Does this become problematic? At what point do we decide if something should be posited as a new FORM value (i.e. how do we stop ourselves from over-filling FORM with so many potential values, perhaps when we don't need to?)?

Reading Questions

- Lastly, is the solution for keeping kick the bucket away from the passive going to work for all the other idioms? By this I mean are we going to be able to use one of those two solutions for all other idioms, either multi-word lxm or lexical entries with very restrictive FORM values.

Reading Questions

- In the sentence 'The FBI kept close tabs on Sandy' how does the semantics of 'carefully observe' get attributed to the verb phrase?

- Would the treatment of phrasal verbs be the same as what is outlined for idioms? Or are phrasal verbs considered idioms in the is grammar?

Reading Questions

- Where do we draw the line for what counts as an idiom?

- How do people learn idioms? How can computers learn them?

- Would the multi-word entries also work for things like "Bank of America" and "New York Stock Exchange"?

Reading Questions

- When a change needs to be made to cover a new type of construction, how do you decide which part of the grammar to modify: the lexeme constraints, the lexical entries, or lexical rules? For example, section 11.4.1 adds a constraint on the type *tv-lxm* to allow for *that*-clause complements, then adds additional constraints to the lexical entries of all the verbs of that type for which it does not apply. What is the motivation behind this decision, versus adding constraints to the lexical entries for which *that*-clause complements do apply, or perhaps making a subtype?