Ling 566
Nov 4, 2025

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*

$$\left\langle \text{be} , \left[\begin{array}{l} \textit{be-lxm} \\ \text{ARG-ST} \left\langle \boxed{1} , \left[\begin{array}{l} \text{SYN} \left[\begin{array}{ll} \text{HEAD} & \left[\begin{array}{ll} \textit{verb} & \\ \text{FORM} & \text{pass} \end{array} \right] \\ \text{VAL} & \left[\begin{array}{ll} \text{SPR} & \left\langle \boxed{1} \right\rangle \\ \text{COMPS} & \left\langle \right\rangle \end{array} \right] \end{array} \right] \\ \text{SEM} \left[\text{INDEX} \quad s \right] \end{array} \right] \right\rangle \\ \text{SEM} \left[\begin{array}{ll} \text{INDEX} & s \\ \text{RESTR} & \left\langle \right\rangle \end{array} \right] \end{array} \right] \right\rangle$$

Copula (generalized)

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

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*

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

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?

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

The Entry for Existential *there*

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

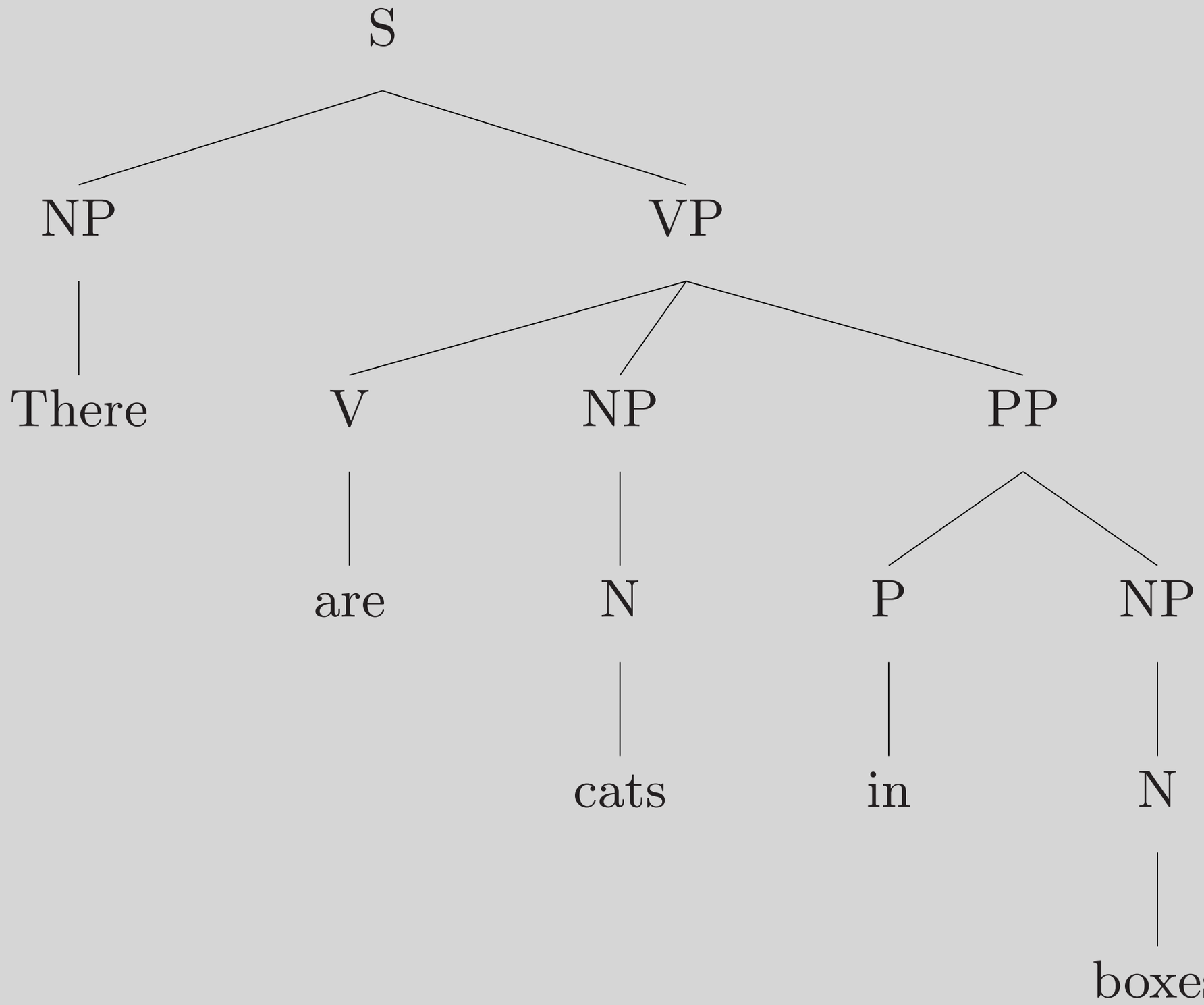
Questions About Existential *there*

- 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?

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

Sample tree for existential *be*



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 correct*

- 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 correct follows*)
- A rule to account for the relationship between pairs like *That you are correct follows* and *It follows that you are correct*

The Entry for Dummy *it*

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

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*?

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

A New Type of Lexeme: Complementizers

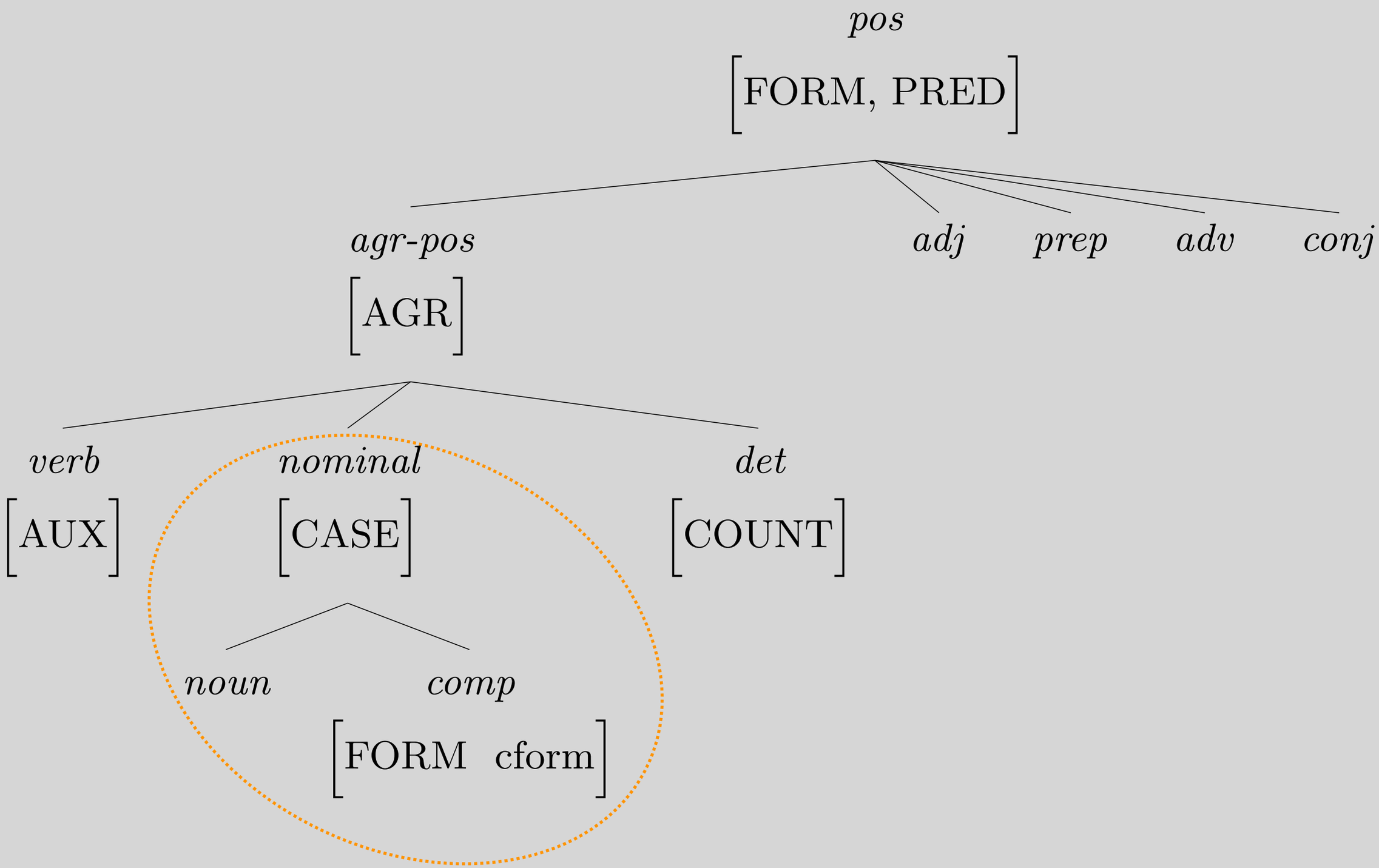
$$comp\text{-}lxm : \begin{bmatrix} \text{SYN} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} comp \\ \text{AGR} & 3sing \end{bmatrix} \\ \text{VAL} & \begin{bmatrix} \text{SPR} & \langle\ \rangle \end{bmatrix} \end{bmatrix} \\ \text{ARG-ST} & \left\langle \begin{bmatrix} \text{S} \\ \text{INDEX} &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

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?

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

The Type *comp*



The Lexical Entry for Complementizer *that*

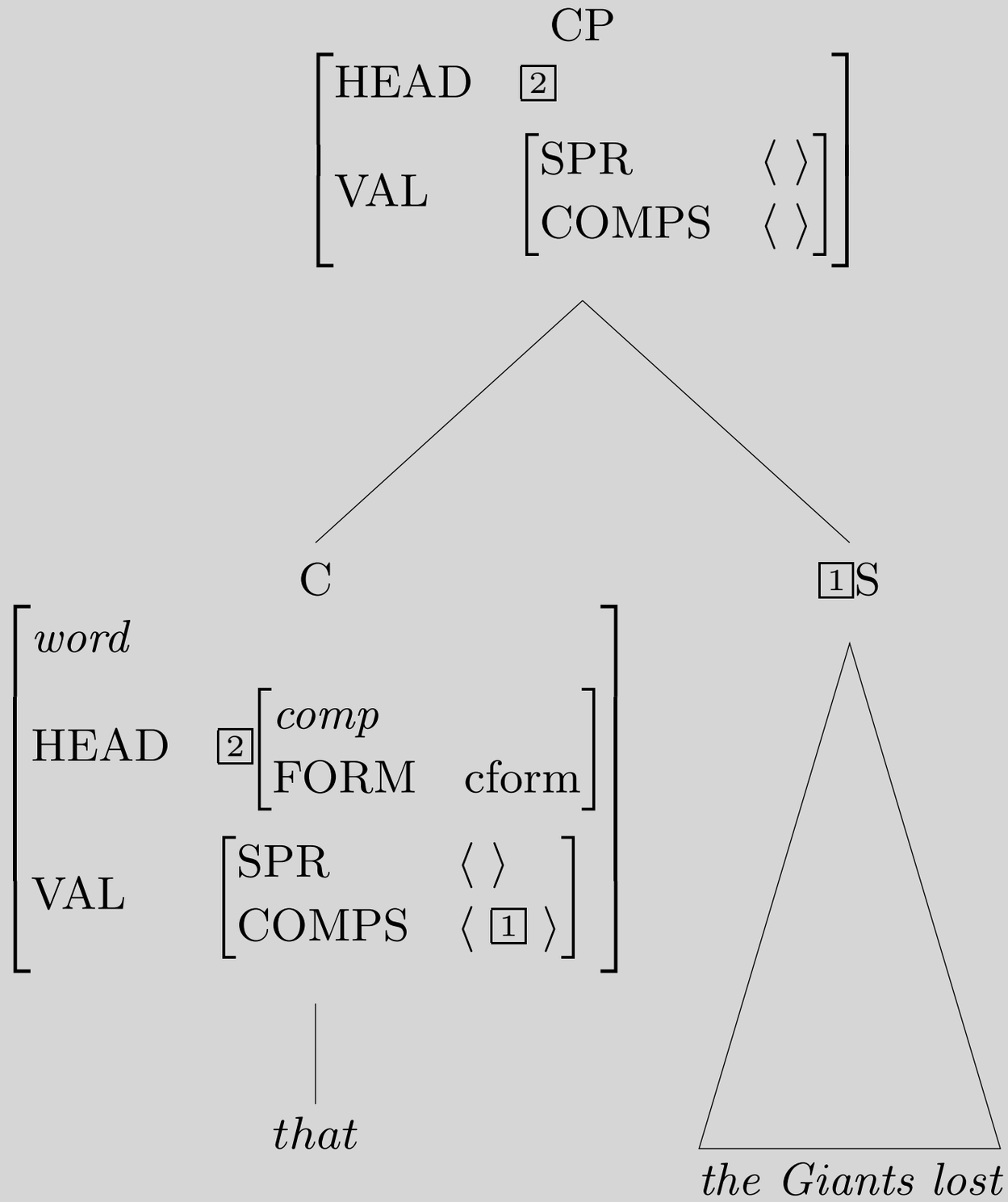
$$\left\langle \text{that} \ , \begin{bmatrix} \textit{comp-lxm} & \\ \text{ARG-ST} & \left\langle \begin{bmatrix} \text{FORM fin} \end{bmatrix} \right\rangle \\ \text{SEM} & \begin{bmatrix} \text{MODE prop} \end{bmatrix} \end{bmatrix} \right\rangle$$

…and with inherited information filled in

$$\left\langle \text{that} \;,\; \left[\begin{array}{ll} \textit{comp-lxm} & \\ \text{SYN} & \left[\begin{array}{ll} \text{HEAD} & \left[\begin{array}{ll} \textit{comp} & \\ \text{FORM} & \text{cform} \\ \text{AGR} & \textit{3sing} \end{array}\right] \\ \text{VAL} & \left[\text{SPR} \quad \langle\,\rangle\right] \end{array}\right] \\ \text{ARG-ST} & \left\langle \overset{\text{S}}{\left[\begin{array}{ll} \text{FORM} & \text{fin} \\ \text{INDEX} & s \end{array}\right]} \right\rangle \\ \text{SEM} & \left[\begin{array}{ll} \text{MODE} & \text{prop} \\ \text{INDEX} & s \\ \text{RESTR} & \langle\,\rangle \end{array}\right] \end{array}\right] \right\rangle$$

Question: Where did [FORM cform] come from?

Structure of a Complementizer Phrase



What is your impression of treating complementizers as the head of constituents like this one?

0

Fine: I know CPs from another class

0

New to me but seems legit

0

Not sure

0

Seems strange, since the main part is the S

0

Sample Verb with a CP Subject

$$\left\langle \text{matter} , \left[\begin{array}{l} \textit{siv-lxm} \\ \text{ARG-ST} \quad \left\langle \left[\text{SEM } [\text{INDEX } \boxed{1}] \right] \right\rangle \\ \\ \text{SEM} \quad \left[\begin{array}{ll} \text{INDEX} & s \\ \\ \text{RESTR} & \left\langle \left[\begin{array}{ll} \text{RELN} & \textbf{matter} \\ \text{SIT} & s \\ \text{MATTERING} & \boxed{1} \end{array} \right] \right\rangle \end{array} \right] \end{array} \right] \right\rangle$$

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.

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

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

The Extraposition Lexical Rule

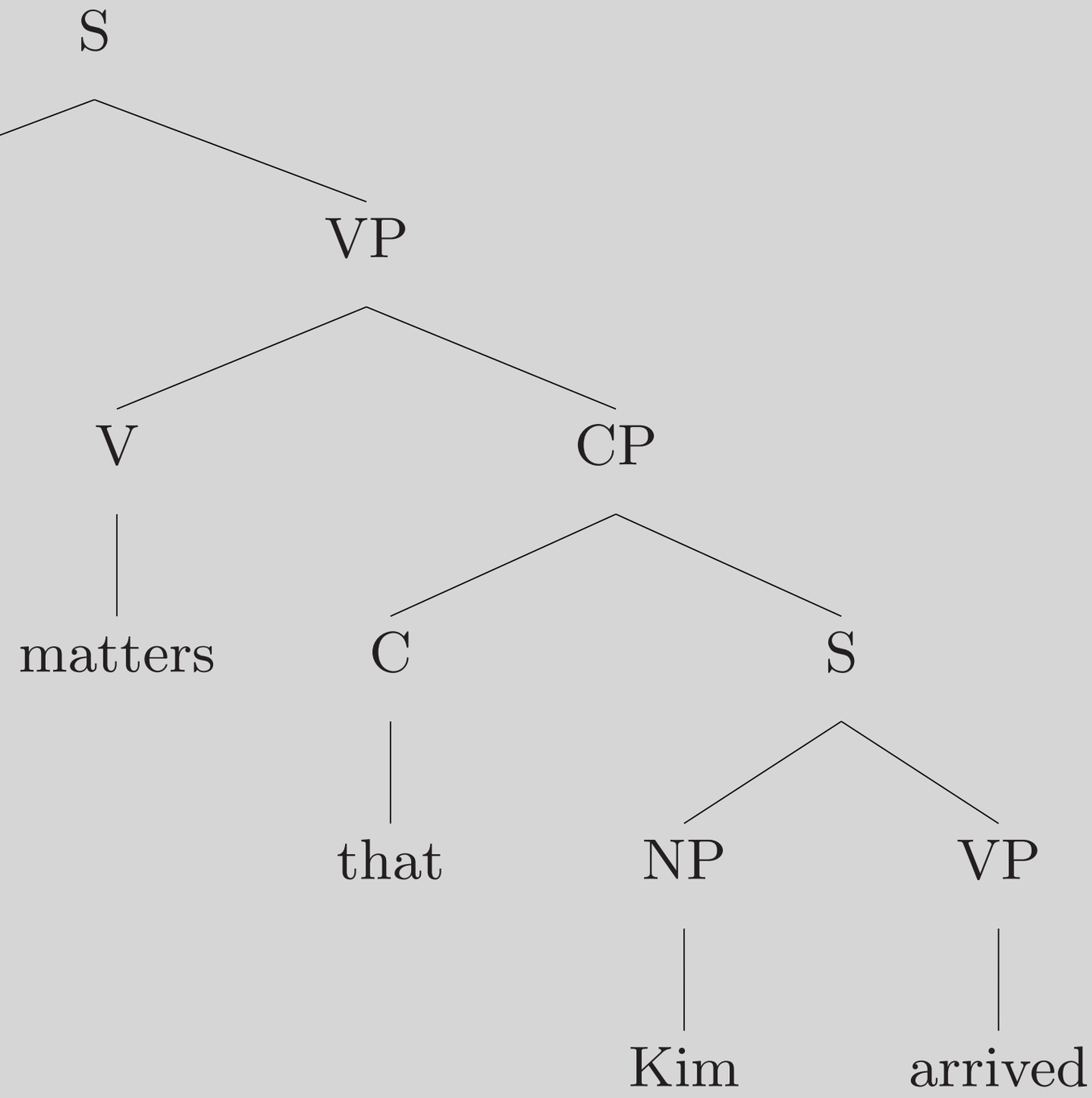
$$\left[\begin{array}{l} \textit{pi-rule} \\ \text{INPUT} \quad \left\langle \text{X} , \left[\text{SYN} \left[\text{VAL} \left[\begin{array}{ll} \text{SPR} & \langle\ \boxed{2}\text{CP}\ \rangle \\ \text{COMPS} & \boxed{\text{A}} \end{array}\right]\right]\right] \right\rangle \\ \text{OUTPUT} \quad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

- Why is the type *pi-rule*?
- Why doesn't it say anything about the semantics?
- Why is the COMPS on INPUT $\boxed{\text{A}}$, 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.*

Sample tree with extraposition



Another Nonreferential Noun

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

The Verb that Selects *advantage*

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

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

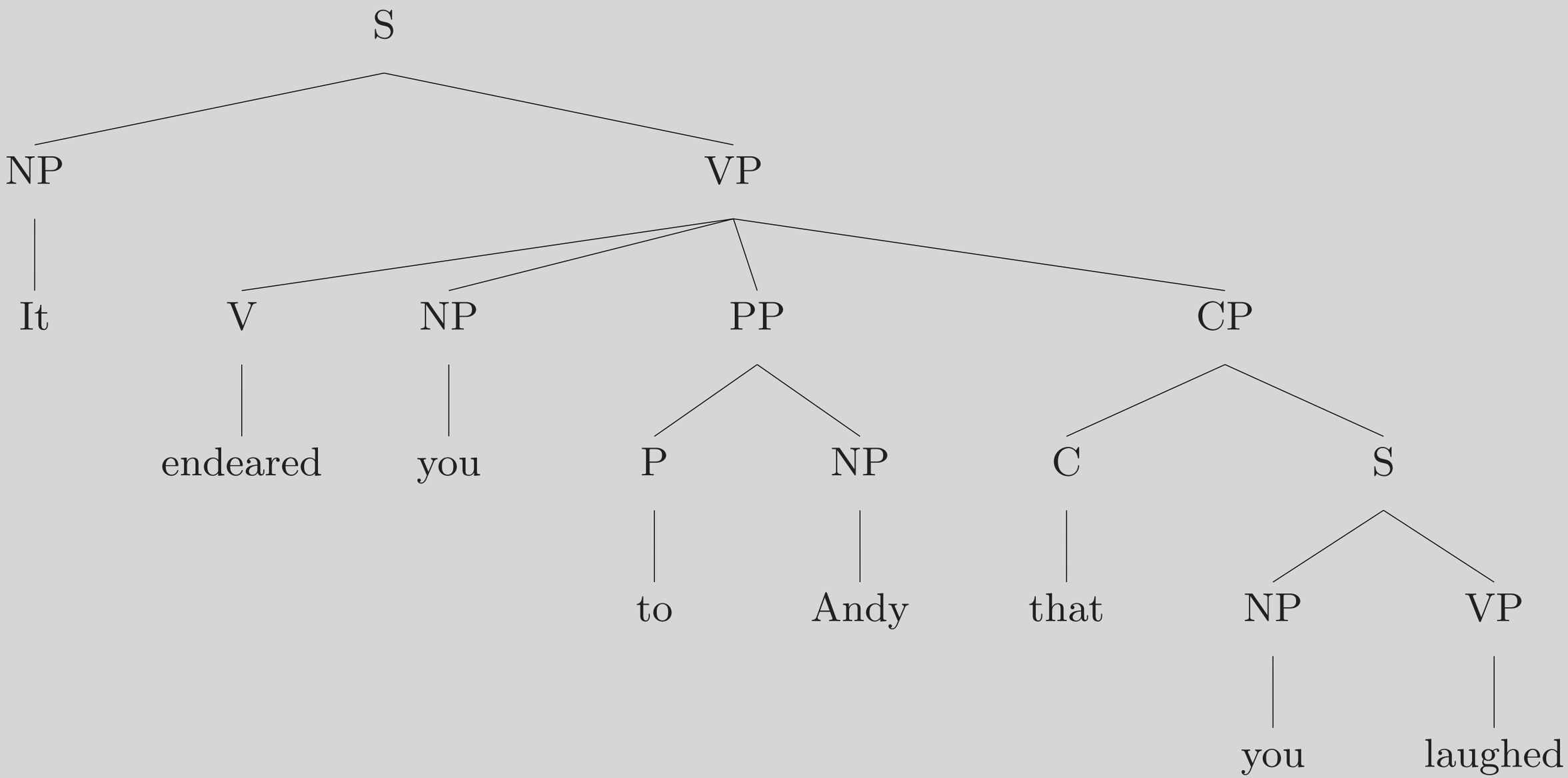
- Existentials (*there*, *be*)
- Extraposition (*that*, *it*, *LR*)
- Idioms

Which *that*?

- *that*: complementizer, demonstrative pronoun, demonstrative determiner, relative pronoun
- The teacher said that that that that that student wrote was superfluous.

Practice tree

- It endeared you to Andy that you laughed



RQs: So many *be*s

- When do we need a separate lexical entry for a same-looking lexeme? The chapter says we need another entry for *be* because we want it to take three elements in the ARG-ST list. So if ARG-ST list changes, we need a new entry? Can't we just incorporate in the original lexical entry of *be* an optional third element?

RQs: Idioms

- From the section about idioms, I was still a little confused why idioms like *kick the bucket* and *kept tabs on* had different analyses. I understand that tabs is semantically empty because it isn't providing actual meaning but without *tabs* there, the *kept* makes no sense on its own so the *tabs* is still required to get the full idiomatic meaning. So why do we not want to analyze *kept tabs* in the same way as *kick the bucket* as a multi word verb?

RQs: Idioms

- Does the multi-word idiom *kick the bucket* allow for transformations for l-rules? Because obviously we can say *kicked the buckets*, *kicks the bucket* etc.

- If we take the rule that the morphological functions happen to the first element in the list of our idiom, how do we deal with agreement further on in the list (for example, *I had butterflies in my stomach* vs *He had butterflies in his stomach*)?

RQs: Idioms

- What determines whether something is a true idiom (non-compositional) or just a highly conventional phrase? Would introducing a new FORM value for each idiom blur the line between idiomatic meaning and ordinary lexical semantics?
- How does the ERG handle idioms? Especially newer idioms like *no cap* (means to accuse you of lying or are you serious)? Do they needed to be manually added?

RQs: Idioms

- But what about idioms like *The early bird catches the worm*? Would "fixed idiomatic sentences" just have set ARG-ST values that state that they can only take words with FORM values related to the idiom?

RQs: FORM

- Is there a list of all the form values that are we are currently using at this point? The form variable is a newer feature introduced so I am still having some trouble remembering all of the abbreviations and what they mean.
- What is the general rule of thumb on when we want to posit more FORM values vs just use the ones that have already been defined? As we can posit more with each idiom we can think of, it seems FORM could have any number of potential values.

RQs: *pi-rule*

- Why is the Extraposition Lexical Rule formulated in terms of SPR and COMPS, rather than ARG-ST? What's the difference?
- Is there a specific reason that the ELR is written in terms of SPR/COMPS and not of ARG-ST?

RQs: CP subjects

- "Alternatively, 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." Could you please explain this?

RQs: Lexical preferences

- Extraposition moves CP from SPR to final COMPS position in the lexicon. Are there verbs that resist extraposition (and some prefer)? If so, should such resistance or preference be stipulated in the lexical entries?
- Some verbs seem to accept both, like *it surprised me that she left* vs. *that she left surprised me* although the former sounds more natural. *She leaving surprised me* should be grammatical too (or *her leaving*? I have seen instances of both).

RQs: Other complementizers

- The other examples of complementizers given on p. 343 are *whether*, *if*, and *for*. Are complementizers different from other ways of creating a subordinate clause, such as relative pronouns?

RQs: Existential *there*

- I can follow the analysis that has "dummy" *there*, but I also wonder if it would be possible to analyze the existential *there* as an adverb with topicalization/emphasis.
- Why is *there* person 3rd? why does it have AGR at all?