

# Ling 566

## Feb 20, 2019

Non-referential NPs, Expletives, and Extraposition

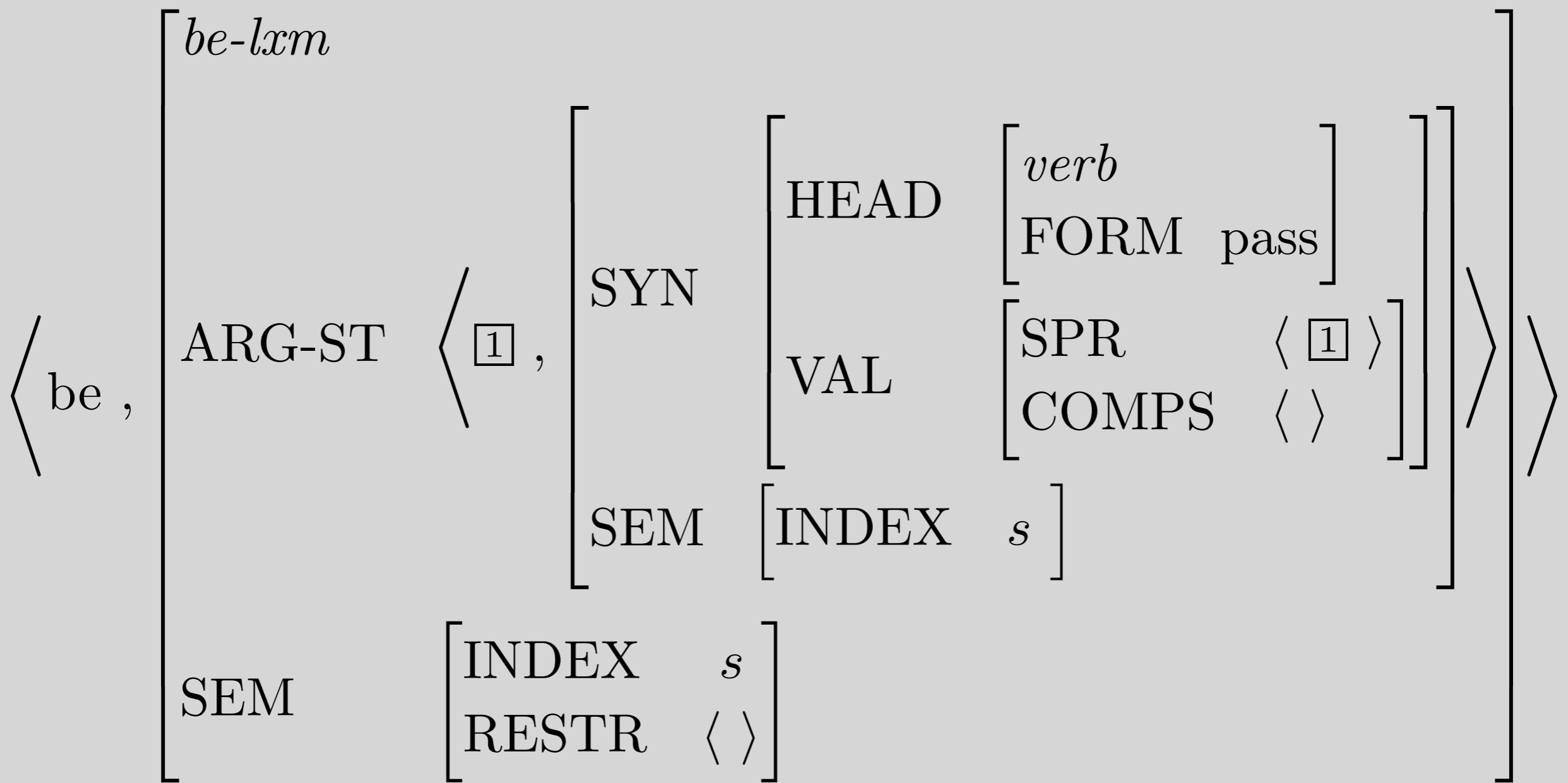
# Overview

- Existentials
- Extraposition
- Idioms

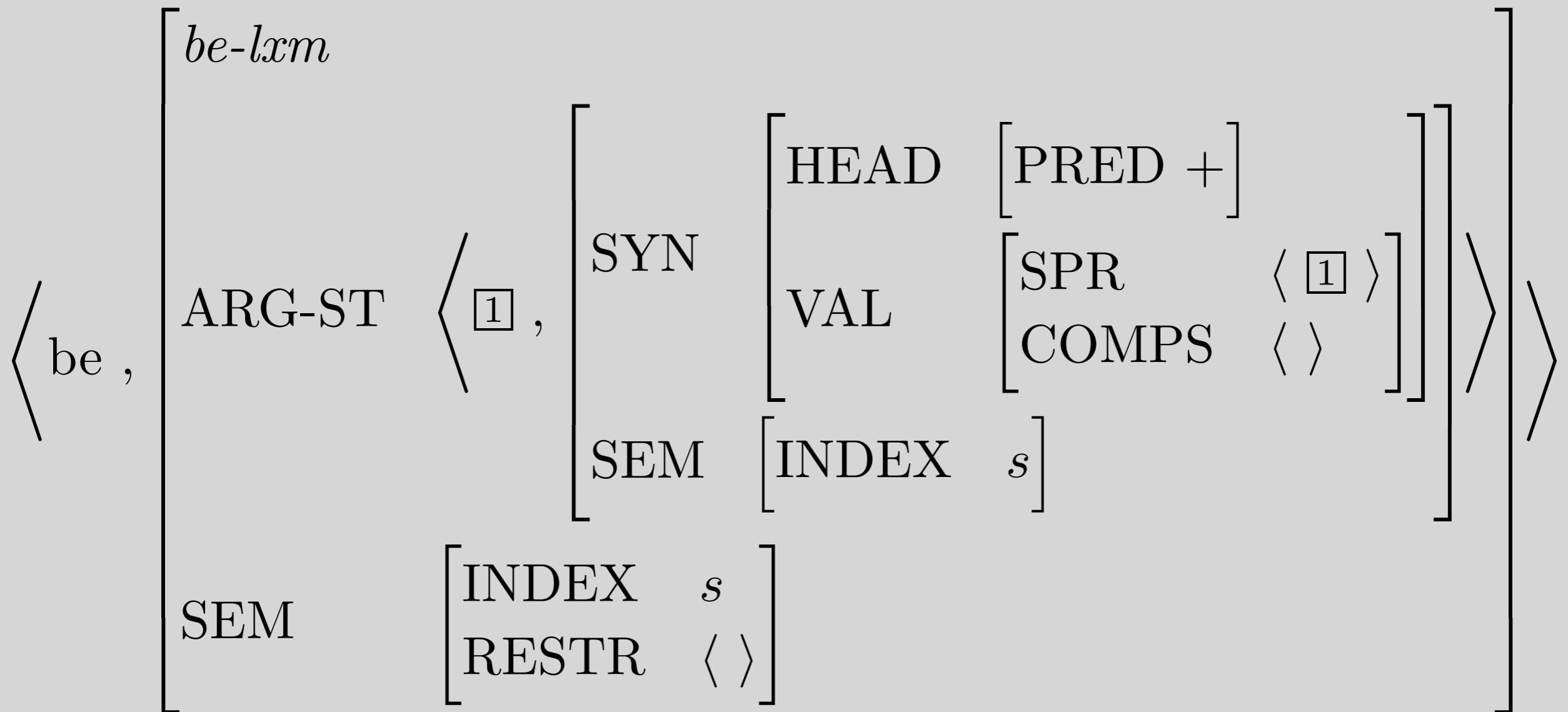
# Where We Are, and Where We're Going

- 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*



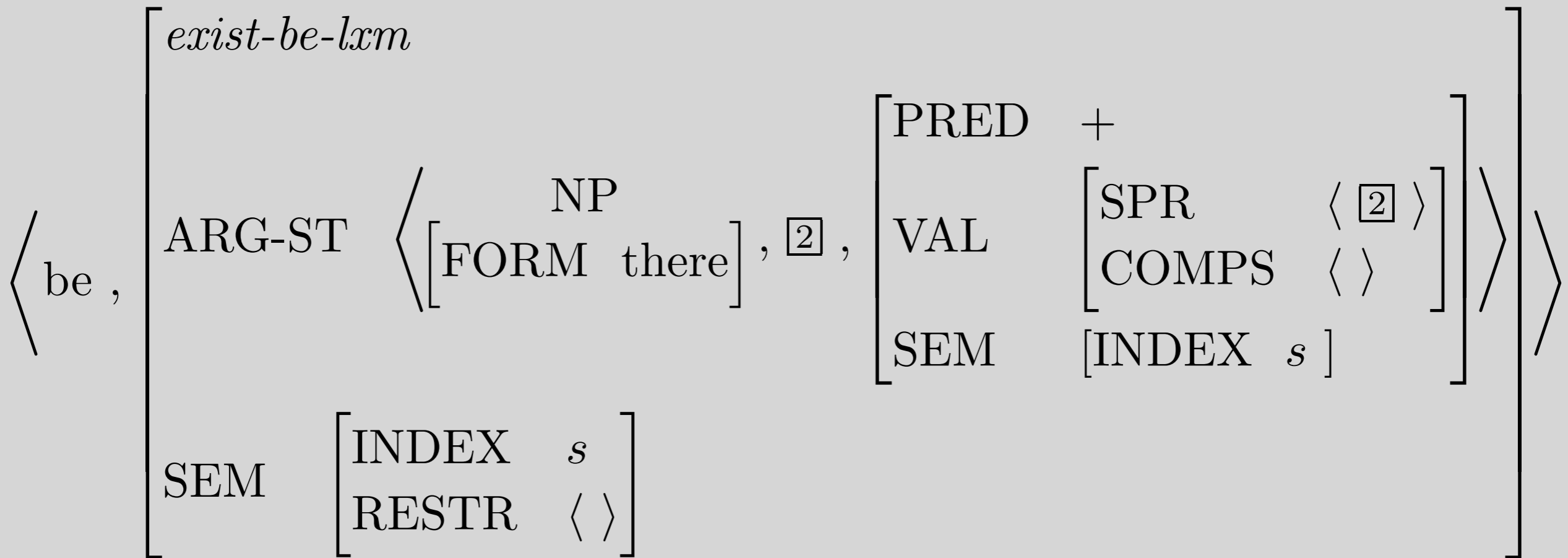
# Copula (generalized)



# Existentials

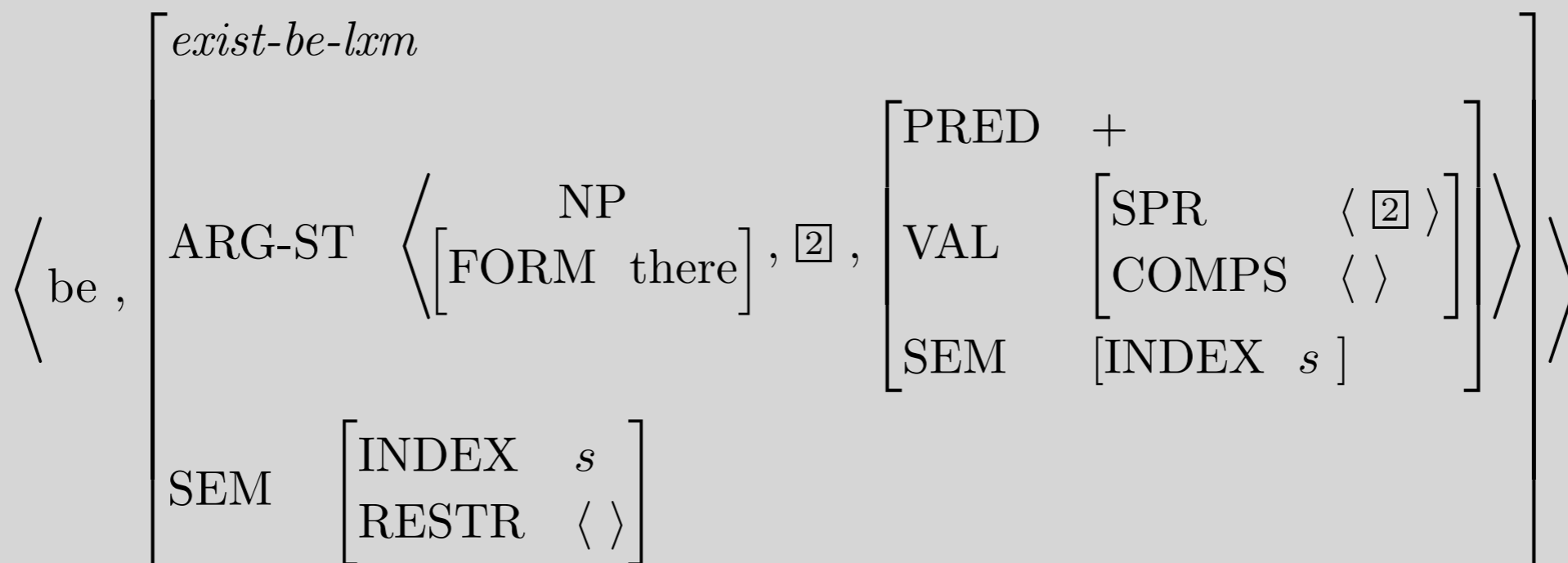
- The *be* in *There is a page missing* cannot be the same *be* that occurs in sentences like *Pat is tall* or *A cat was chased by a dog*. Why not?
- So we need a separate lexical entry for this *be*, stipulating:
  - Its SPR must be *there*
  - It takes two complements, the first an NP and the second 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*

⟨	there ,	<i>pron-lxm</i>	[	HEAD	[	FORM	there	]	]	⟩
		AGR			[	PER	3rd	]		
[	SEM	MODE	none	]						
		INDEX	none	]						
		RESTR	⟨ ⟩	]						

# 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?

⟨	there ,	<i>pron-lxm</i>	[	HEAD	[	FORM	there	]	]	⟩
		AGR			[	PER	3rd	]		
		SEM	[	MODE	none	]				
			[	INDEX	none	]				
			[	RESTR	⟨ ⟩	]				

# 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	$\left[ \begin{array}{l} \text{FORM } it \\ \text{AGR } 3sing \end{array} \right]$	$\rangle$
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*?

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

# A New Type of Lexeme: Complementizers

*comp-lxm* :

SYN	<table border="1"> <tr> <td>HEAD</td> <td> <table border="1"> <tr> <td><i>comp</i></td> </tr> <tr> <td>AGR</td> <td><i>3sing</i></td> </tr> </table> </td> </tr> <tr> <td>VAL</td> <td> <table border="1"> <tr> <td>SPR</td> <td><math>\langle \rangle</math></td> </tr> </table> </td> </tr> </table>	HEAD	<table border="1"> <tr> <td><i>comp</i></td> </tr> <tr> <td>AGR</td> <td><i>3sing</i></td> </tr> </table>	<i>comp</i>	AGR	<i>3sing</i>	VAL	<table border="1"> <tr> <td>SPR</td> <td><math>\langle \rangle</math></td> </tr> </table>	SPR	$\langle \rangle$
HEAD	<table border="1"> <tr> <td><i>comp</i></td> </tr> <tr> <td>AGR</td> <td><i>3sing</i></td> </tr> </table>	<i>comp</i>	AGR	<i>3sing</i>						
<i>comp</i>										
AGR	<i>3sing</i>									
VAL	<table border="1"> <tr> <td>SPR</td> <td><math>\langle \rangle</math></td> </tr> </table>	SPR	$\langle \rangle$							
SPR	$\langle \rangle$									
ARG-ST	<table border="1"> <tr> <td>S</td> </tr> <tr> <td>INDEX</td> <td><i>s</i></td> </tr> </table>	S	INDEX	<i>s</i>						
S										
INDEX	<i>s</i>									
SEM	<table border="1"> <tr> <td>INDEX</td> <td><i>s</i></td> </tr> <tr> <td>RESTR</td> <td><math>\langle \rangle</math></td> </tr> </table>	INDEX	<i>s</i>	RESTR	$\langle \rangle$					
INDEX	<i>s</i>									
RESTR	$\langle \rangle$									

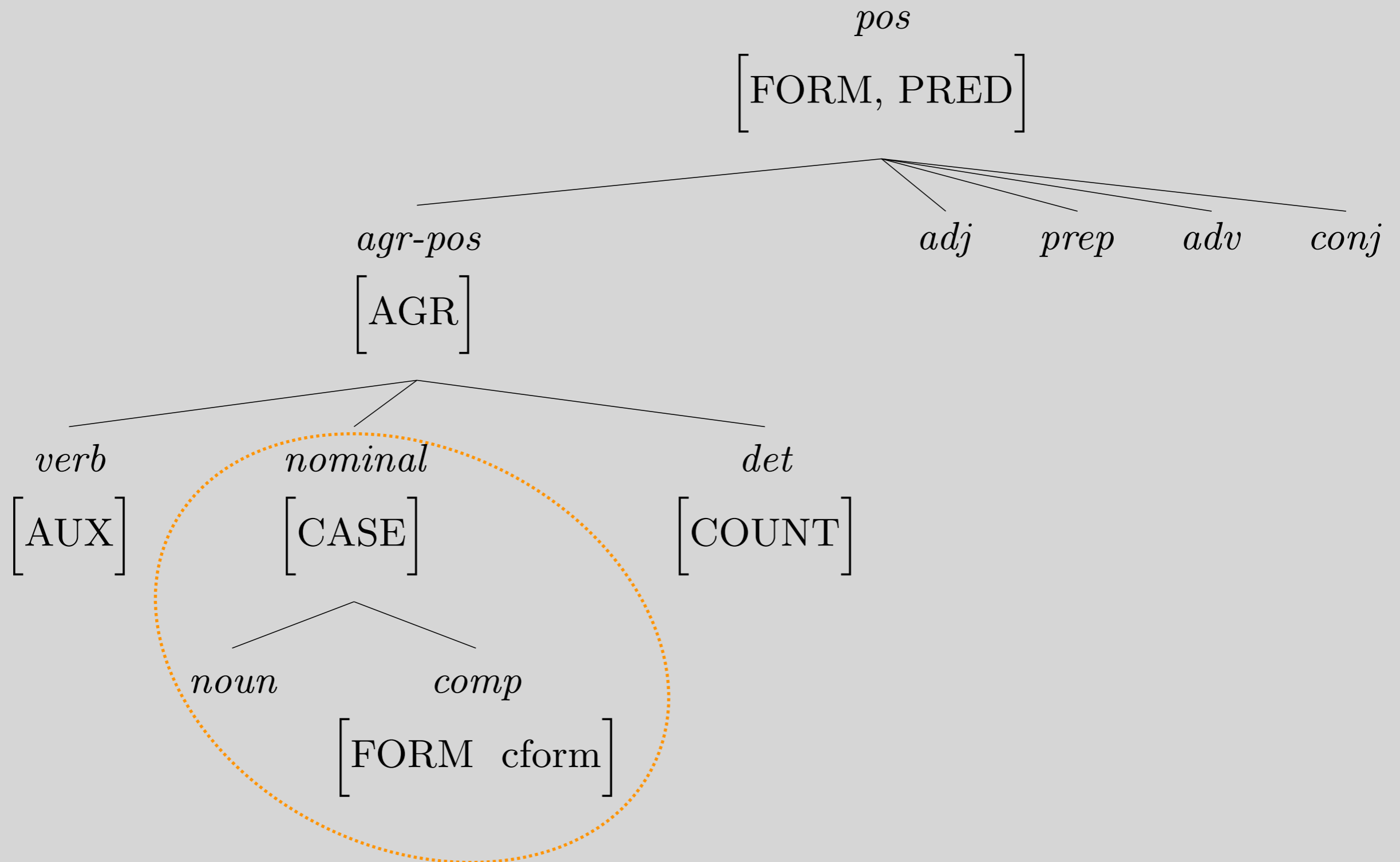
# Questions About the Type *comp-lxm*

- 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?

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



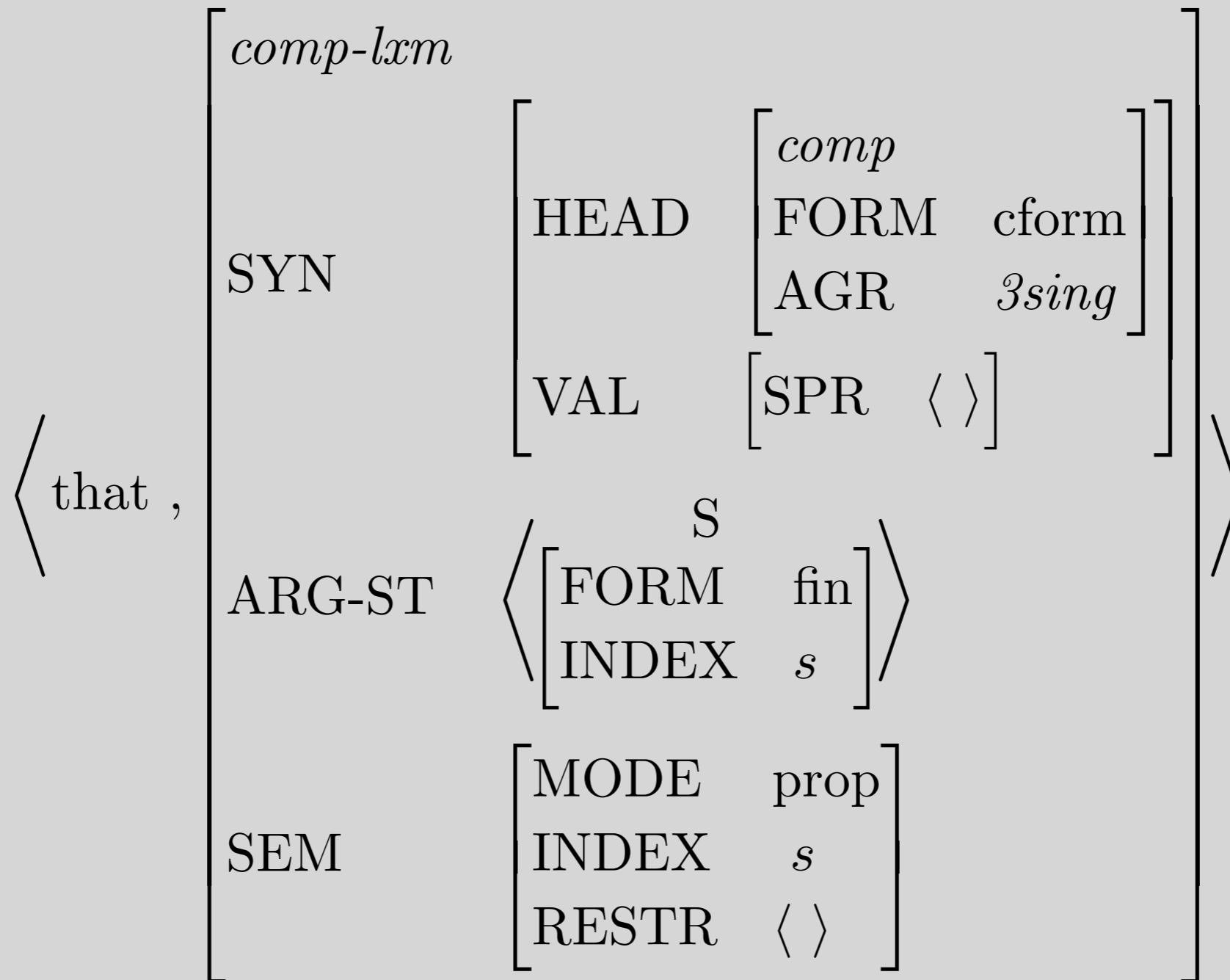
# The Type *comp*



# The Lexical Entry for Complementizer *that*

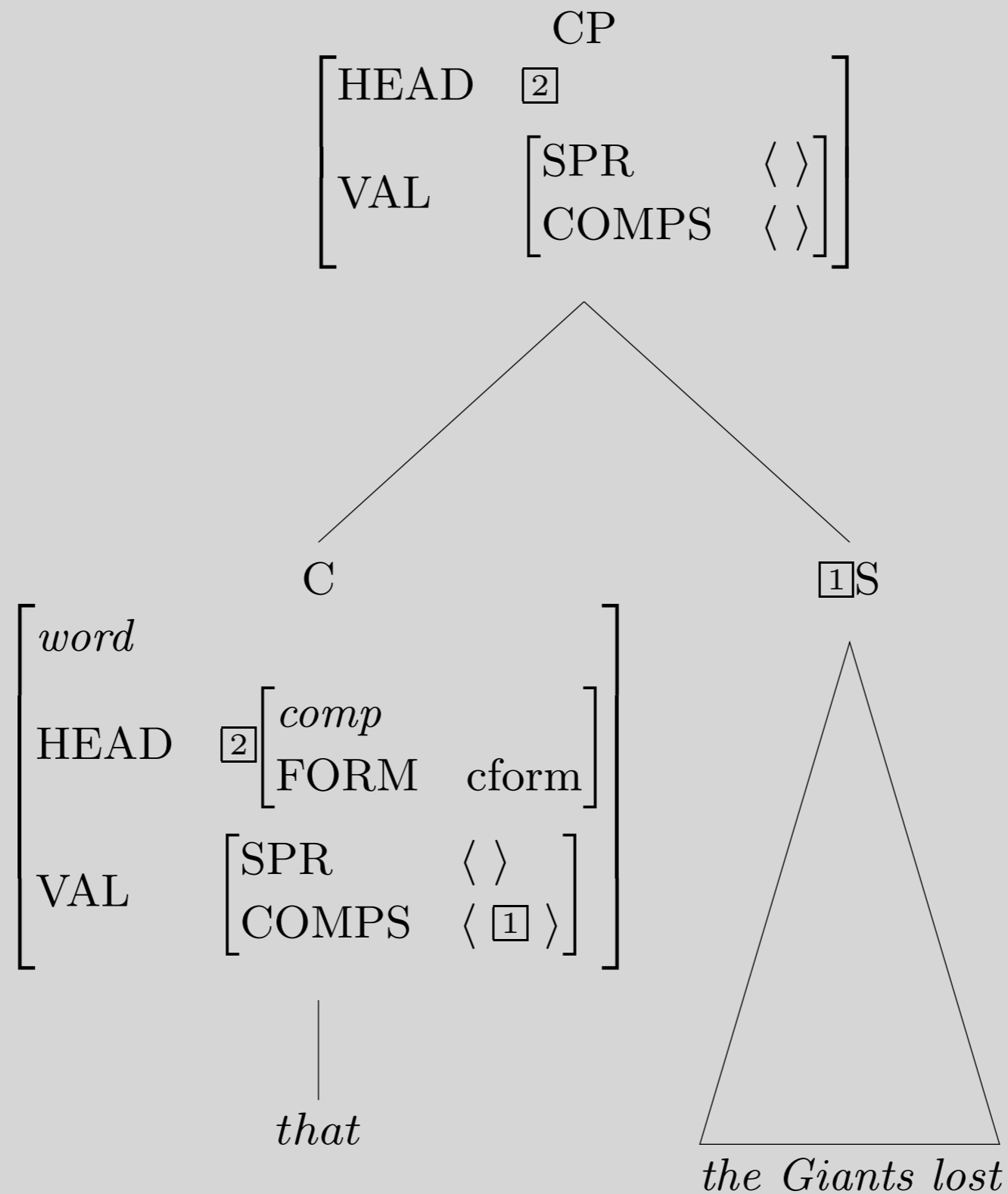
$$\left\langle \text{that} , \begin{bmatrix} \textit{comp-lxm} \\ \text{ARG-ST} \langle [\text{FORM fin}] \rangle \\ \text{SEM} \begin{bmatrix} \text{MODE prop} \end{bmatrix} \end{bmatrix} \right\rangle$$

...and with inherited information filled in

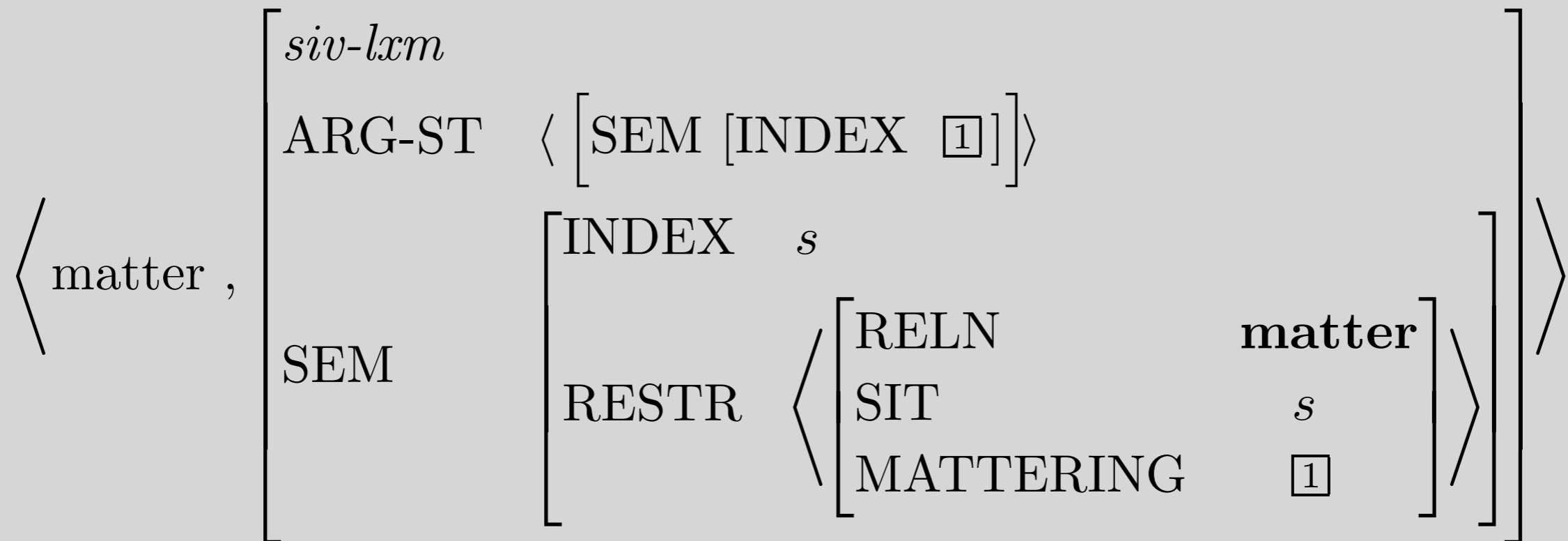


Question: Where did [FORM cform] come from?

# Structure of a Complementizer Phrase



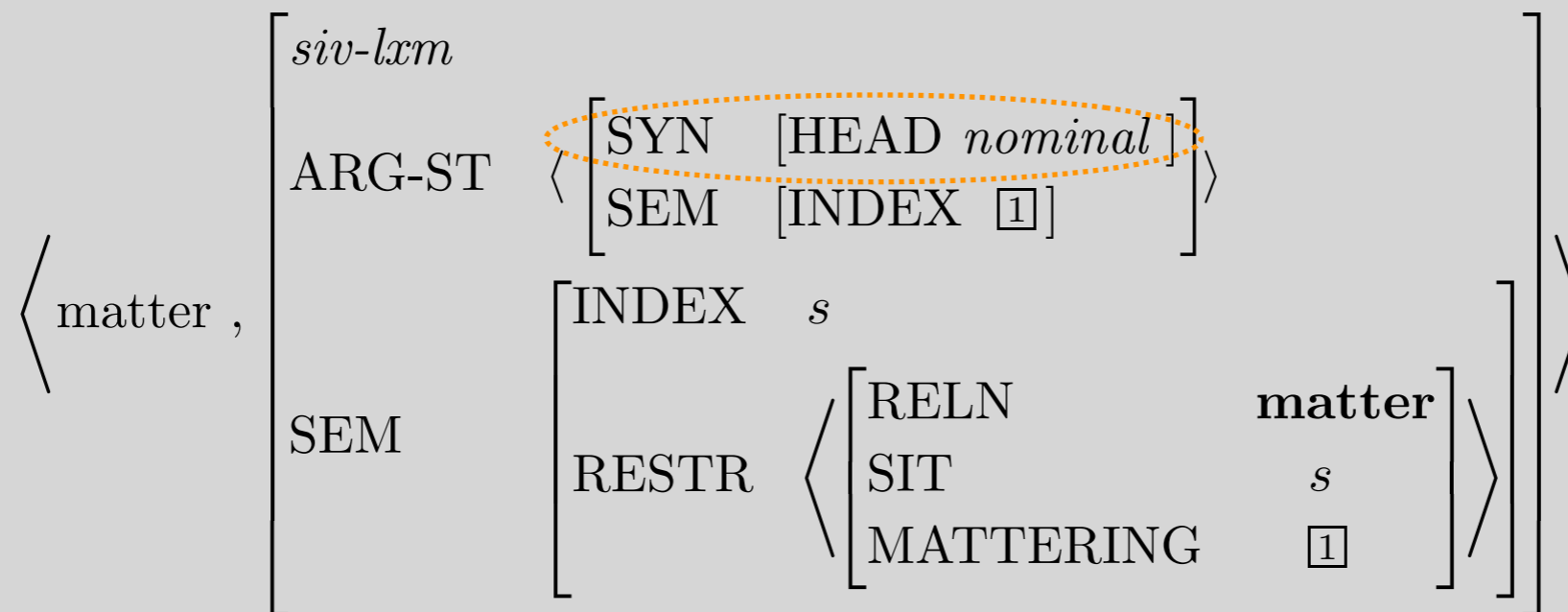
# Sample Verb with a CP Subject



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*.

# The Extraposition Lexical Rule

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

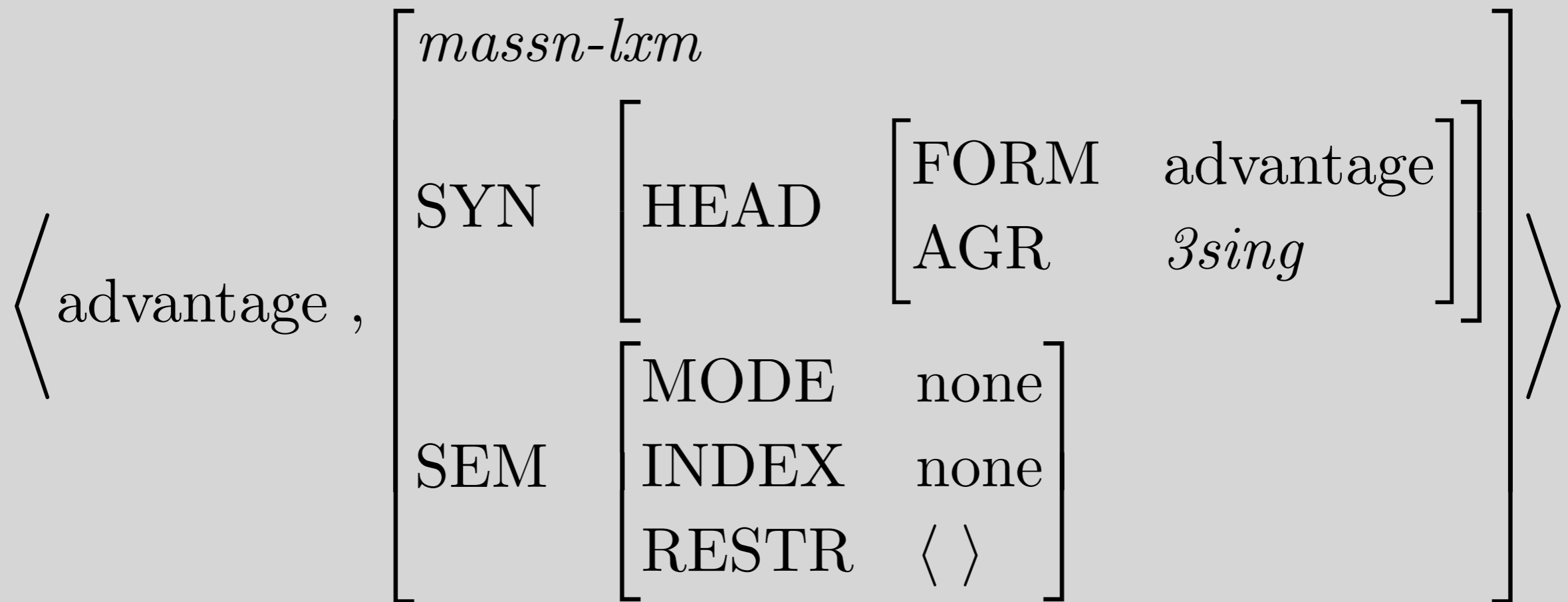
- Why is the type *pi-rule*?
- Why doesn't it say anything about the semantics?
- Why is the COMPS on INPUT  $\boxed{A}$ , not  $\langle \ \ \rangle$ ?

# Extraposition with Verbs whose COMPS Lists are Nonempty

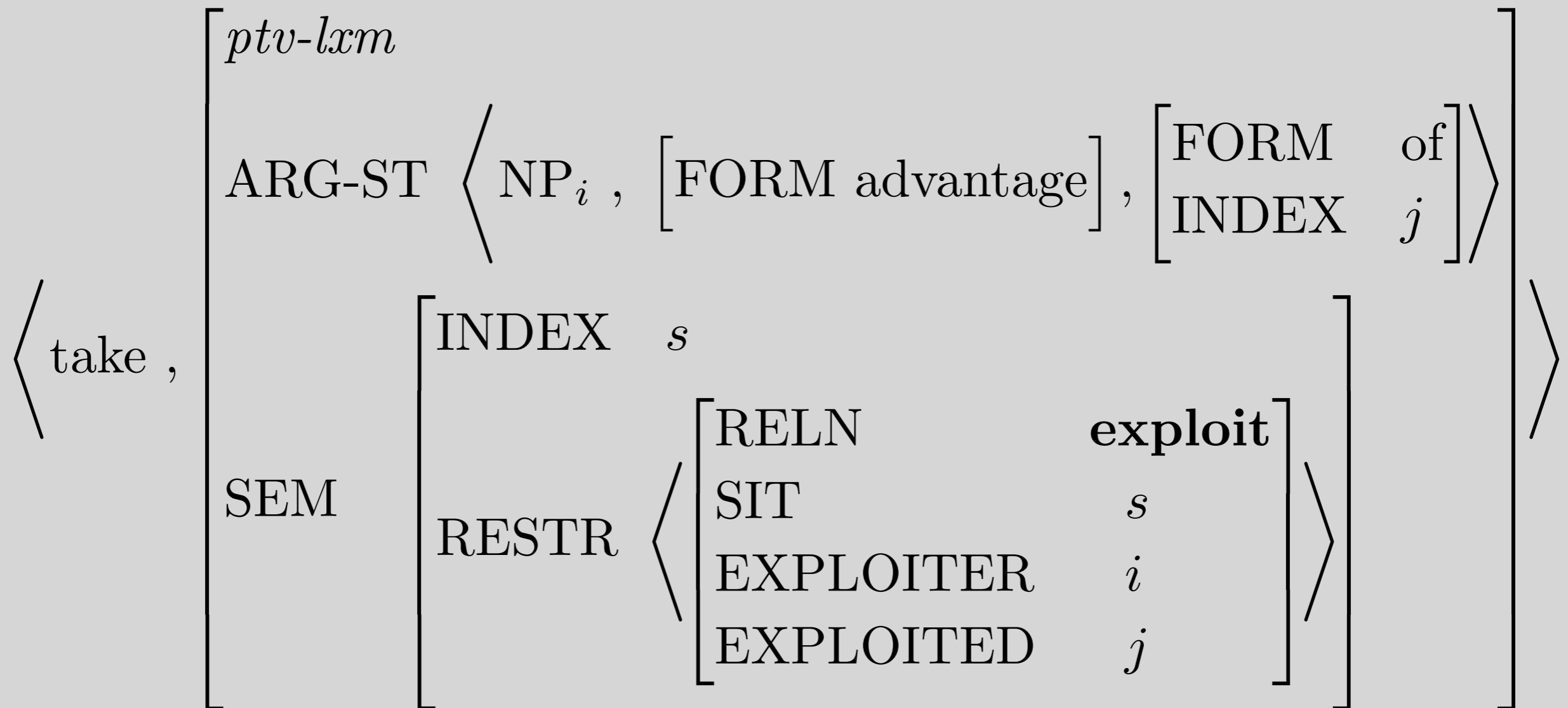
- *It worries me that war is imminent.*
- *It occurred to Pat that Chris knew the answer.*
- *It endeared you to Andy that you wore a funny hat.*



# Another Nonreferential Noun



# The Verb that Selects *advantage*



# 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

- In the p. 335 example *The book is under the table*, why does *under* take *book* as its SPR? That structure doesn't make sense when I consider constructions like *The book under the table is heavy*: by the Head-Specifier rule, that would mean *book under the table* is headed by the preposition *under* and thus is a PP and not an NP, which seems wrong.

# Reading Questions

- In fn 2, why does *a scholar* have to take MODE prop in (i)? Isn't it also possible to structure the semantics of *is* as an equivalence relation so that we just need to pick as the referent a scholar such that Pat is the scholar?
- Considering 11.7 on Page 350, why PRED is a feature necessary for adj, prep, adv and conj?

# Reading Questions

- Why are complementizers [ AGR *3sing* ]?

# Reading Questions

- So CPs can act as both the specifiers and the complements of verbs:
  - (a) That the dog was barking annoyed Sandy.
  - (b) Sandy hated that the dog was barking.
- but not every slot in a sentence that could take an NP can take a CP:
  - (c) \*That the dog was barking hated Sandy.
  - (d) \*Sandy annoyed that the dog was barking.
- Are (c) and (d) truly ungrammatical, or just problematic in a "colorless green ideas" way? Should these restrictions be specified in the lexical entries for the verbs?



# Reading Questions

- In the case where complementizer can be omitted, do we create a new rule for it like imperatives?
- Why would omitting *that* in (a) be fine, but omitting *that* in (b) leads to a weird-sounding sentence?

(a) I believed that the sky is blue.

(b) That Cleveland lost the series sucks.

# Reading Questions

- For That-Clauses, how can we distinguish the clauses in which *that* can be omitted and those in which *that* can not from the lexical information of CP?
- So comp is given CASE so it can be underspecified and therefore appear in places that request case, not because it is actually marked for case, right? This is said on page 341. Are there languages where it would be actually marked? It seems like it might be a good candidate for something like this based on distribution, no?

# Reading Questions

- How could we analyze: *It was thought that the campaign worked*
- Why does *It sucks that the Giant lost* sound better than *That the Giants lost sucks*?
- Why do we need post-inflectional rules?  
What makes d-rules and i-rules insufficient?  
Why can't the extraposition lexical rule just be a d-rule?

# Reading Questions

- Do we analyze the “weather ‘it’ “ in the same way as the extraposition “it”? It’s definitely hard to argue that it’s referring to anything (maybe nature as a whole?? ) but it doesn’t intuitively seem to be quite the same thing as the purely vacuous “placeholder” it in extraposition.

# Reading Questions

- *A unicorn is in the garden* can be paraphrased as *There is a unicorn in the garden*, but according to fn 5, *A vase is blue* can not be paraphrased as *There is a vase blue*. How can we differentiate the usage of *blue* and *in the garden* in the *there* construction sentence?

# Reading Questions

- What about sentences like *There is a unicorn?*
- What's going on with *There lived a woman with a love of dogs?*
- *There exists a unicorn in the garden* Should this also have the same RESTR as *There is a unicorn in the garden?*

# Reading Questions

- *Here is the book you asked for.* Is this use of *here* different from the existential *there*?
- Does our current existential *there* work for predications like *there is a unicorn* where it's just claiming existence? There's not really a sentential paraphrase of *there is a unicorn* or *there are unicorns*. It seems like the paraphrase would just be *unicorns*, which normally is not a full sentence.

# Reading Questions

- p.336 fn describes a "so-called definiteness restriction" on the NP following the word *is*. It is suggested that there is a semantic distinction which could invalidate the two examples:
- i) ?\*There is each unicorn in the garden. (ii) ?There was the felon elected to the city council.
- Why are we not concerned about representing this phenomenon despite claiming to represent semantics within our grammatical formalism? This phenomenon does seem to interact with the syntax despite being a distinction in meaningfulness of sentence.



# Reading Questions

- How do we account for flexible idioms?
  - I spilled the beans on our project. → The beans were spilled on our project.
  - The old man kicked the bucket. → #The bucket was kicked (by the old man).
  - They spilled the surprising beans.
- What are some restrictions on using idioms? Even changing the tenses sometimes sounds a little weird to me? Like "He had kicked the bucket."

# Reading Questions

- The use of FORM to identify and select for the correct components of certain idioms seems really clunky to me. Is there not a more elegant way to do this? (There are so many idioms out there, it seems like our FORM value is quickly getting very saturated.)

# Reading Questions

- Let's say we're encoding this grammar computationally, how would we effectively indicate the semantics of an idiom in our grammar in a way that a computer could make sense of the true meaning instead of the literal one?
- What about nonverbal idioms, such as *piece of cake* or *under the weather*? How does/should our grammar handle these?

# Reading Questions

- I hadn't previously realised that some idioms allow for the passive form while others do not. Is there any specific reason why *kick the bucket* cannot be used in the passive form?

# Reading Questions

- Does that mean we are considering these verbs to have a different sense when they are used in idioms that when they aren't? Because while I can understand that analysis for idiom's like *kick the bucket* or *take the cake* I don't really see it as much for verbs like *take* in *take advantage*. It seems like in this idiom, the verb is being used with the same sense as it is in any other case. And for the verb *keep*, there is also the phrase *keep watch*, that means something similar but a different argument on the verb. Would this also be analyzed the same way? How do we distinguish between an idiomatic meaning that requires this method of analysis and just an alternative meaning that can use a typical verbal analysis?

# Reading Questions

- Constraining keep's complements to [FORM tabs] and [FORM on] seems oddly specific (48.a, p.348). I want to understand the motivation for inventing new FORM values (and corresponding lexemes to instantiate those FORM values) instead of following the suggested treatment of treating *kick the bucket* e.g. string-as-lexical-entry e.g. (51) p.495  
<kick,the,bucket> ?

# Reading Questions

- In a real world grammar how clean (or cluttered) are the actual FORM values?
- Regarding idioms, I was convinced by the "can we?" part, but it seems like the "should we?" wasn't addressed. Why are idioms properly part of a grammar? If the rest of our semantics are happy to leave synonymy up to external semantic interpretation, it seems odd that *kicked the bucket* should be exactly semantically equivalent to *died*. If we must account for idioms, why not other kinds of poetic and pragmatic language? Where is the boundary?



# Reading Questions

- I may end up blurring the line between collocations/formulaic language/idioms here, but where do we draw the line when deciding if something is an idiom or not? I understand that semantic opacity is important, but I feel like there is a finer distinction to be made. How would we deal with the following examples?
  - phrasal verbs like *put up with*
  - *make amends*
  - *I beg your pardon*



# Reading Questions

- In conceptual metaphor theory there is emphasis on the salience of the source domain on the target domain in metaphors such as *I've had my ups and down in life* and *spend time*, and this is often also relevant for idiomatic expressions based on metaphors. Out of curiosity, in HPSG is there a way to represent in the features that some idioms simultaneously carry the shadow of the source domain sense alongside the target?

# Reading Questions

- How common are these "semantically empty" words across languages? Example (6) shows us where they don't appear in other languages but due to the English focus of the text (and my lack of broader linguistic knowledge) I'm left wondering how often they'll show up across languages and how we'd handle the FORM values when they do.

# Reading Questions

- Do versions of dummy it/there occur in most languages? Also, for pro-drop languages, would pronouns technically be "dummies" and/or optionally semantically empty?

# Reading Questions

- The solutions to the grammatical phenomena raised in this chapter seem kind of "hacky" (in that we have new FORM values and lexical entries to deal with a single idiom or the existential be) but necessary to deal with all of English syntax. In other words, are other languages, in general, more regular than English, or do most other languages also need language specific hacks to make the grammar work?

# Reading Questions

- Where do we get data in order to make judgements about what the grammar needs to handle. Idioms can certainly be quite variable across dialects/regions. How do we decide what gets captured? Do we cast as wide a net as possible, and does this get unruly? Is there a widely accepted list of idiomatic FORM values?