



Ling 566
Nov 3, 2020
Lexical Rules

Overview

- How lexical rules fit in
- Three types of lexical rules, constraints
- Example: Plural noun lexical rule
- Advice on writing lexical rules
- Constant lexemes
- ARG-ST & ARP
- The feature FORM

Lexical Types & Lexical Rules

- Lexemes capture the similarities among *run*, *runs*, *running*, and *ran*
- The lexical type hierarchy captures the similarities among *run*, *sleep*, and *laugh*, among those and other verbs like *devour* and *hand*, and among those and other words like *book*.
- Lexical rules capture the similarities among *runs*, *sleeps*, *devours*, *hands*, ...

Poll!

Parsimony & Plausibility

- Lexical rules capture **productive** generalizations.
- There may be some ‘precompiling’ going on as well.

Three Kinds of Lexical Rules

- Inflectional: *lexeme to word*

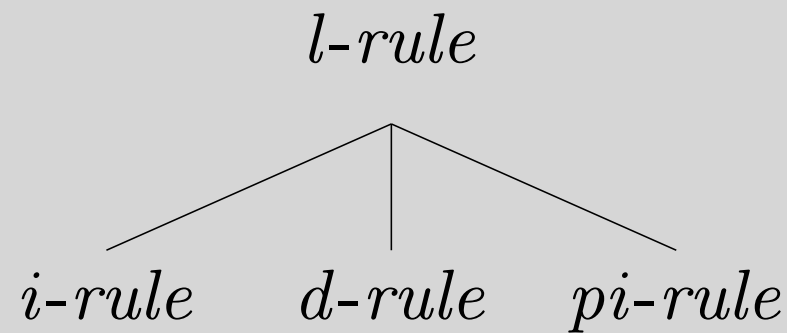
Examples?

- Derivational: *lexeme to lexeme*

Examples?

- Post-Inflectional: *word to word*
(Chapters 11, 13, 14)

Three Subtypes of *l-rule*

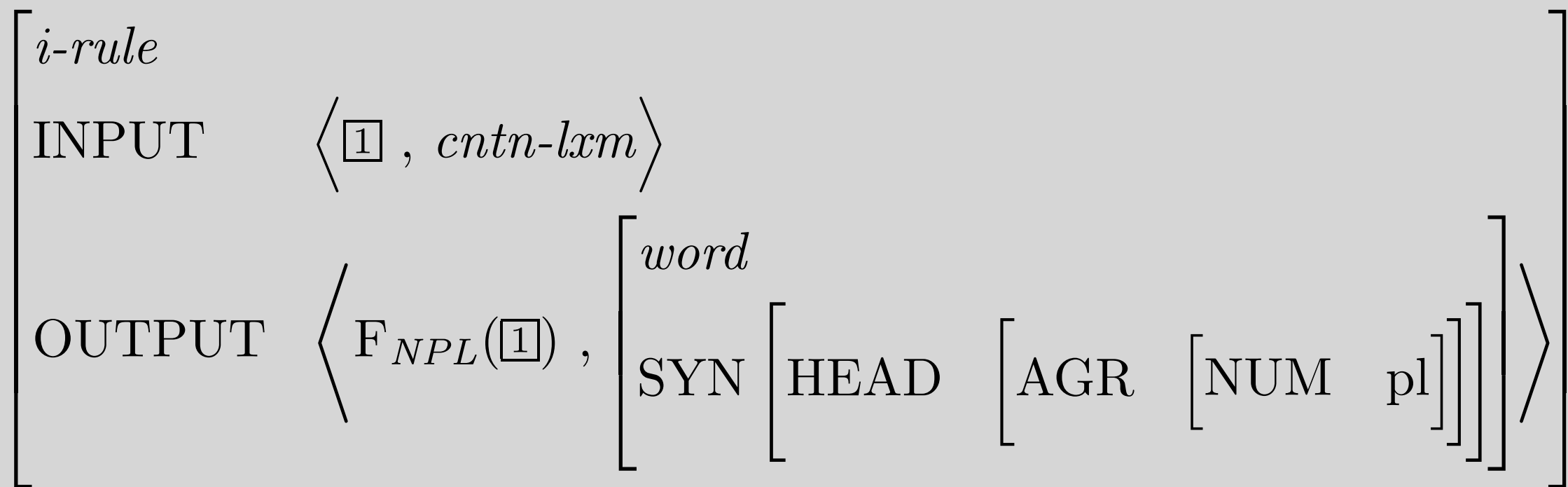


$$l\text{-rule} : \left[\begin{array}{l} \text{INPUT} \quad l\text{-sequence} \left\langle X, [\text{SEM} \quad / \boxed{2}] \right\rangle \\ \text{OUTPUT} \quad l\text{-sequence} \left\langle Y, [\text{SEM} \quad / \boxed{2}] \right\rangle \end{array} \right]$$

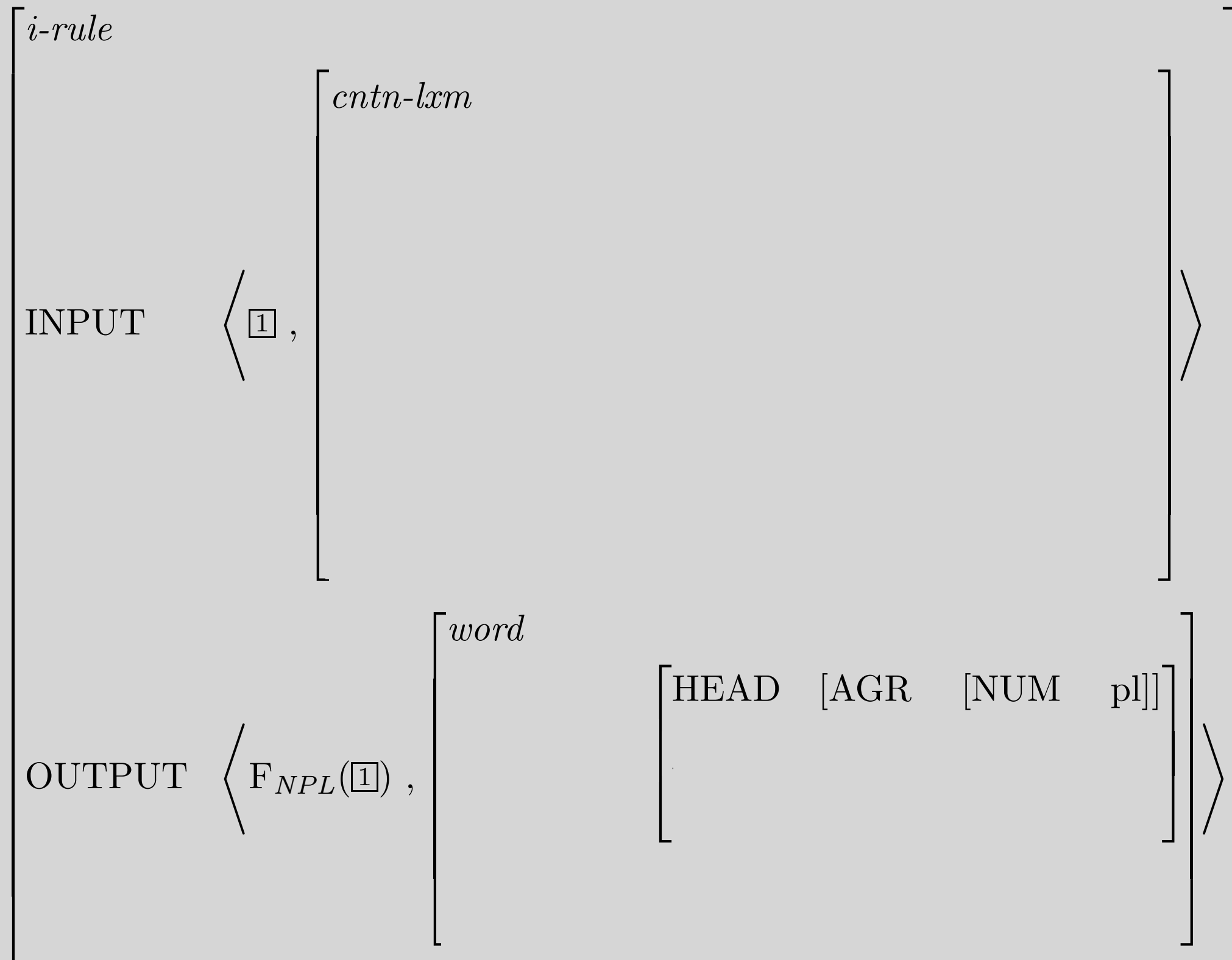
$$i\text{-rule} : \left[\begin{array}{l} \text{INPUT} \quad \left\langle X, \left[\begin{array}{l} \textit{lexeme} \\ \text{SYN} \quad \boxed{3} \\ \text{ARG-ST} \quad \boxed{A} \end{array} \right] \right\rangle \\ \text{OUTPUT} \quad \left\langle Y, \left[\begin{array}{l} \textit{word} \\ \text{SYN} \quad \boxed{3} \\ \text{ARG-ST} \quad \boxed{A} \end{array} \right] \right\rangle \end{array} \right]$$

$$d\text{-rule} : \left[\begin{array}{l} \text{INPUT} \quad \left\langle X, \left[\begin{array}{l} \textit{lexeme} \\ \text{SYN} \quad / \boxed{3} \end{array} \right] \right\rangle \\ \text{OUTPUT} \quad \left\langle Y, \left[\begin{array}{l} \textit{lexeme} \\ \text{SYN} \quad / \boxed{3} \end{array} \right] \right\rangle \end{array} \right]$$

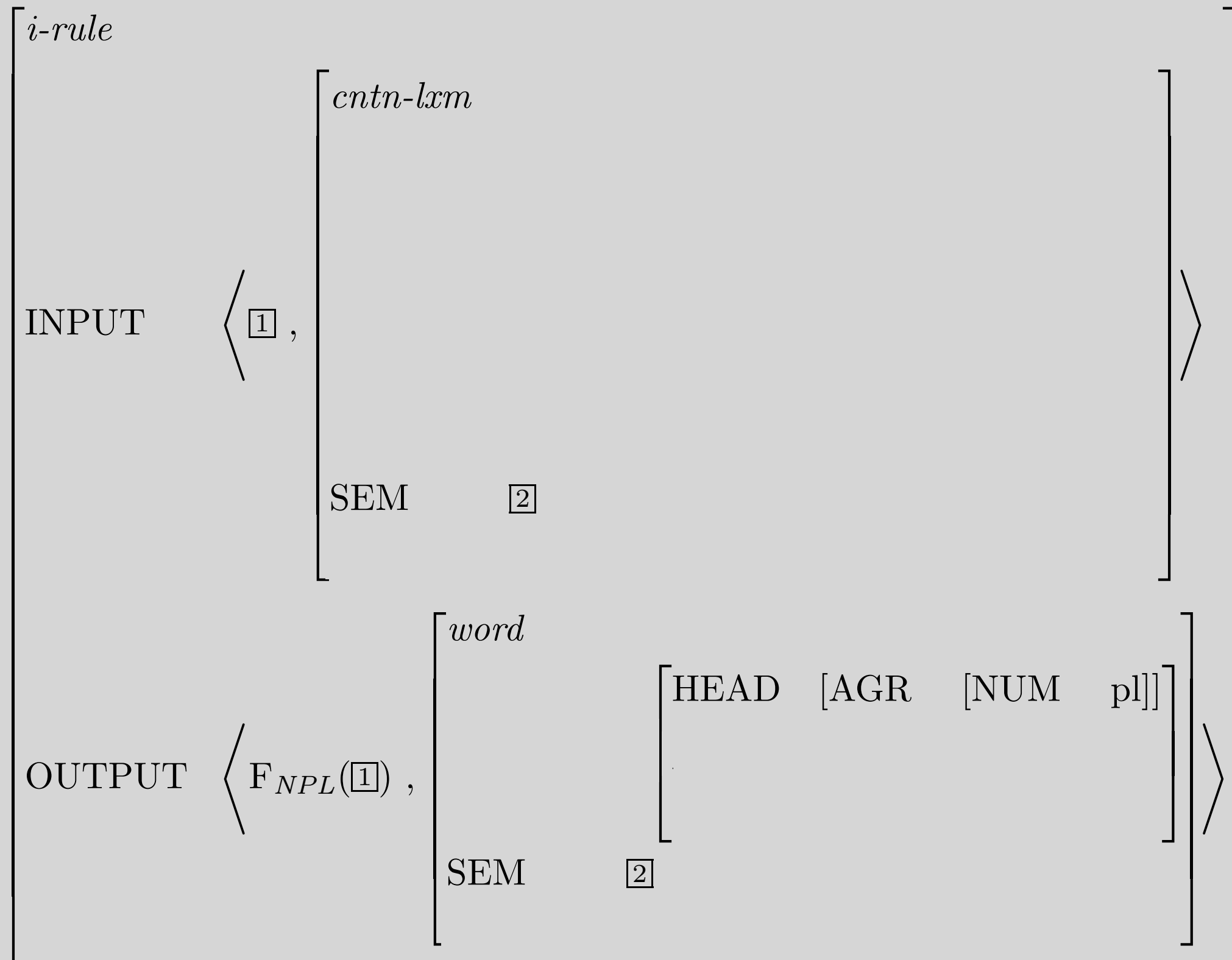
Plural Noun LR



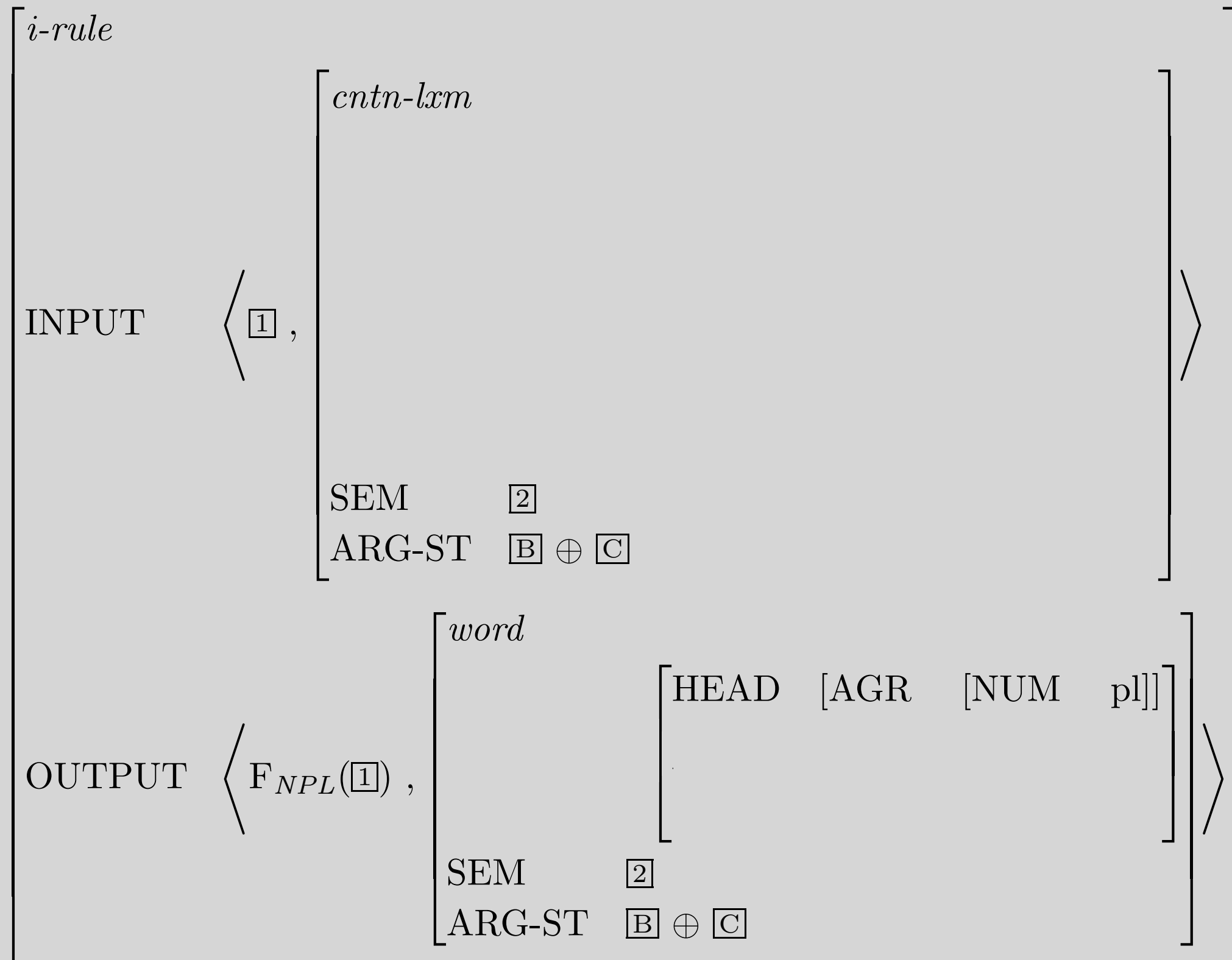
Plural Noun LR with Inherited Constraints



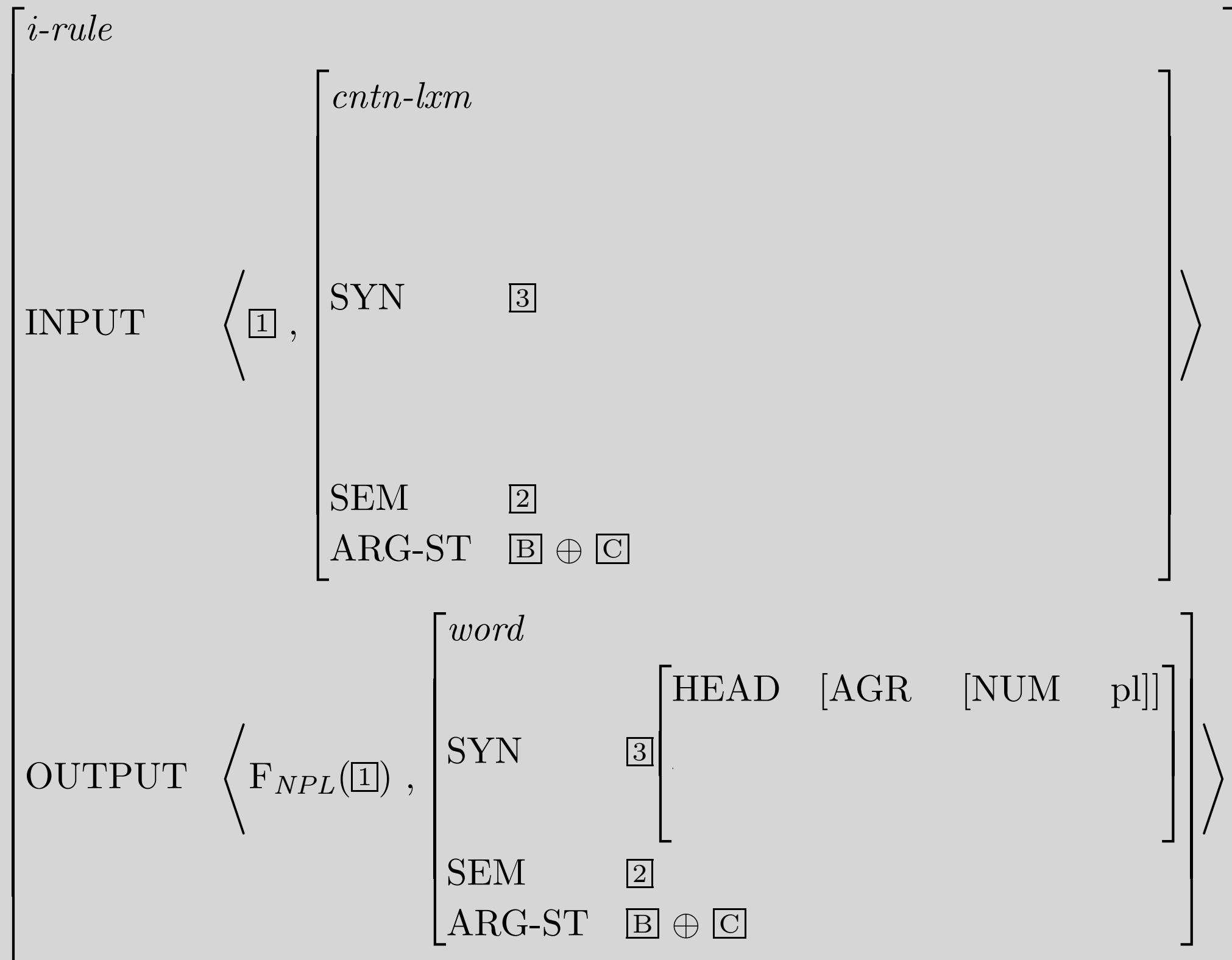
Plural Noun LR with Inherited Constraints



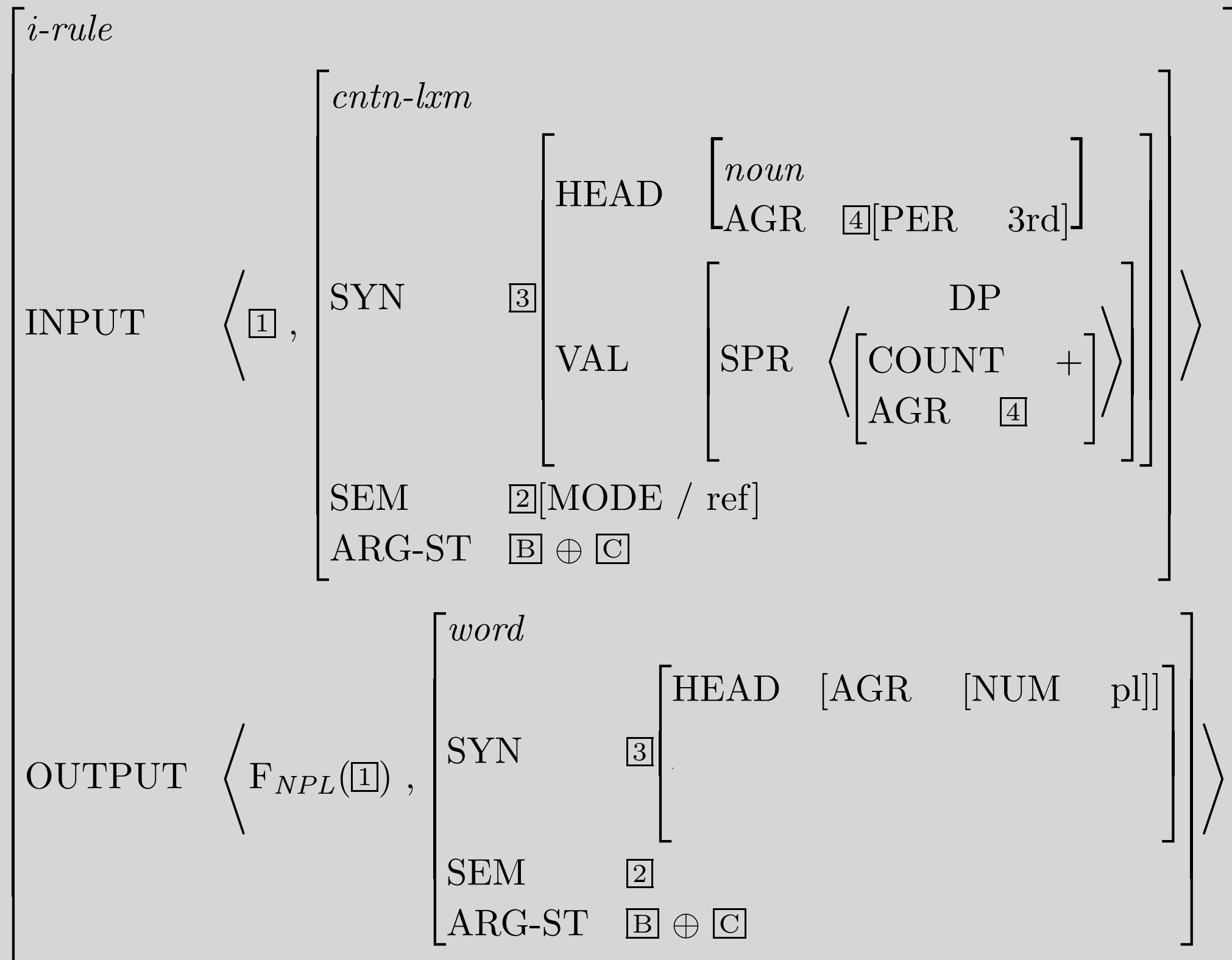
Plural Noun LR with Inherited Constraints



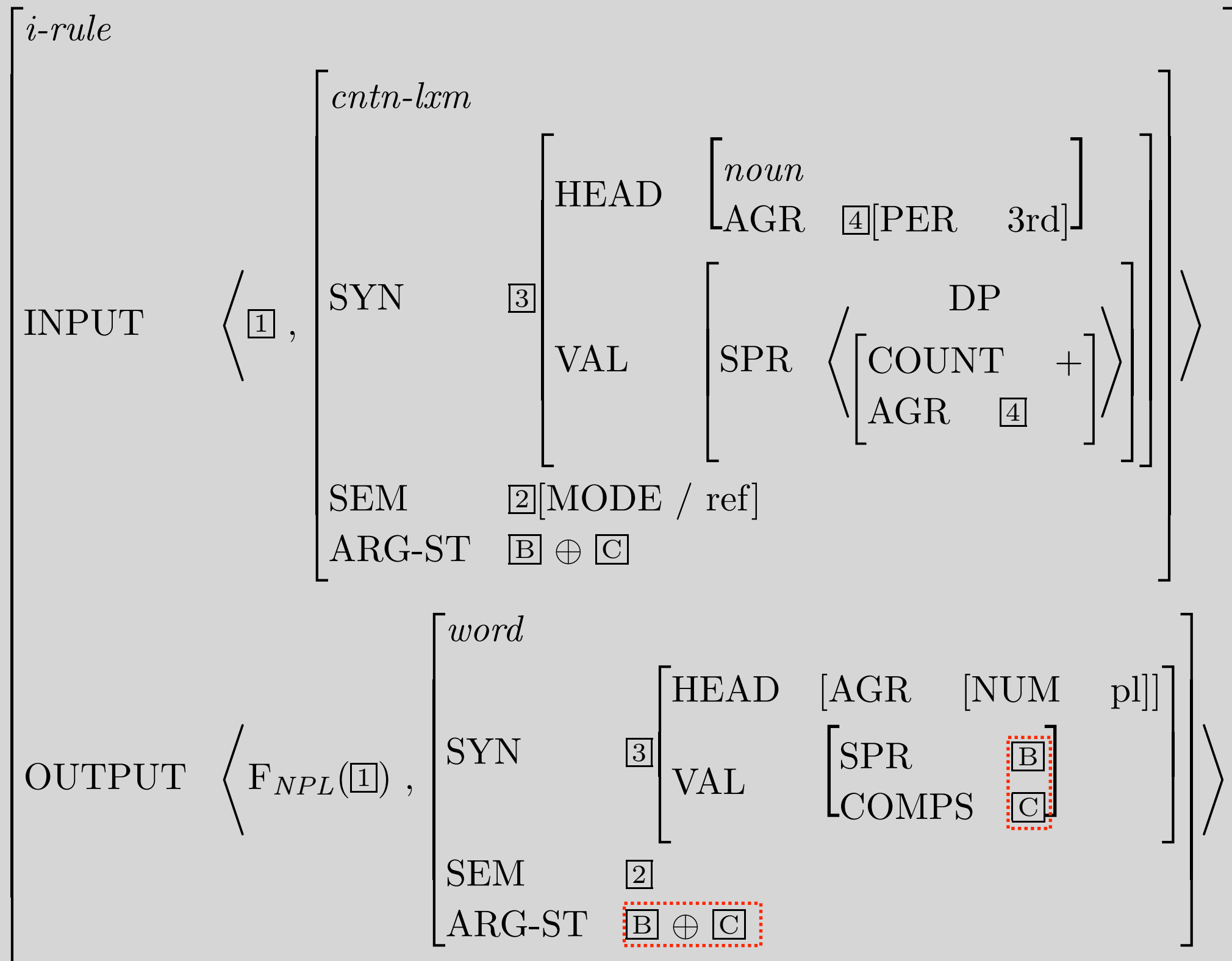
Plural Noun LR with Inherited Constraints



Plural Noun LR with Inherited Constraints



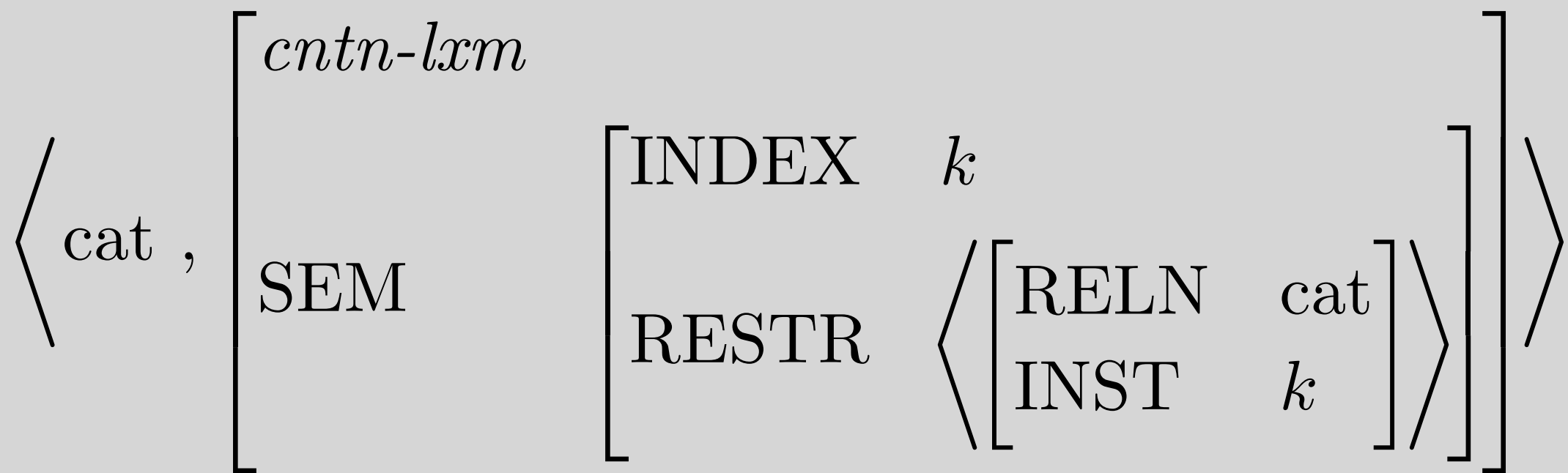
Plural Noun LR with Inherited Constraints



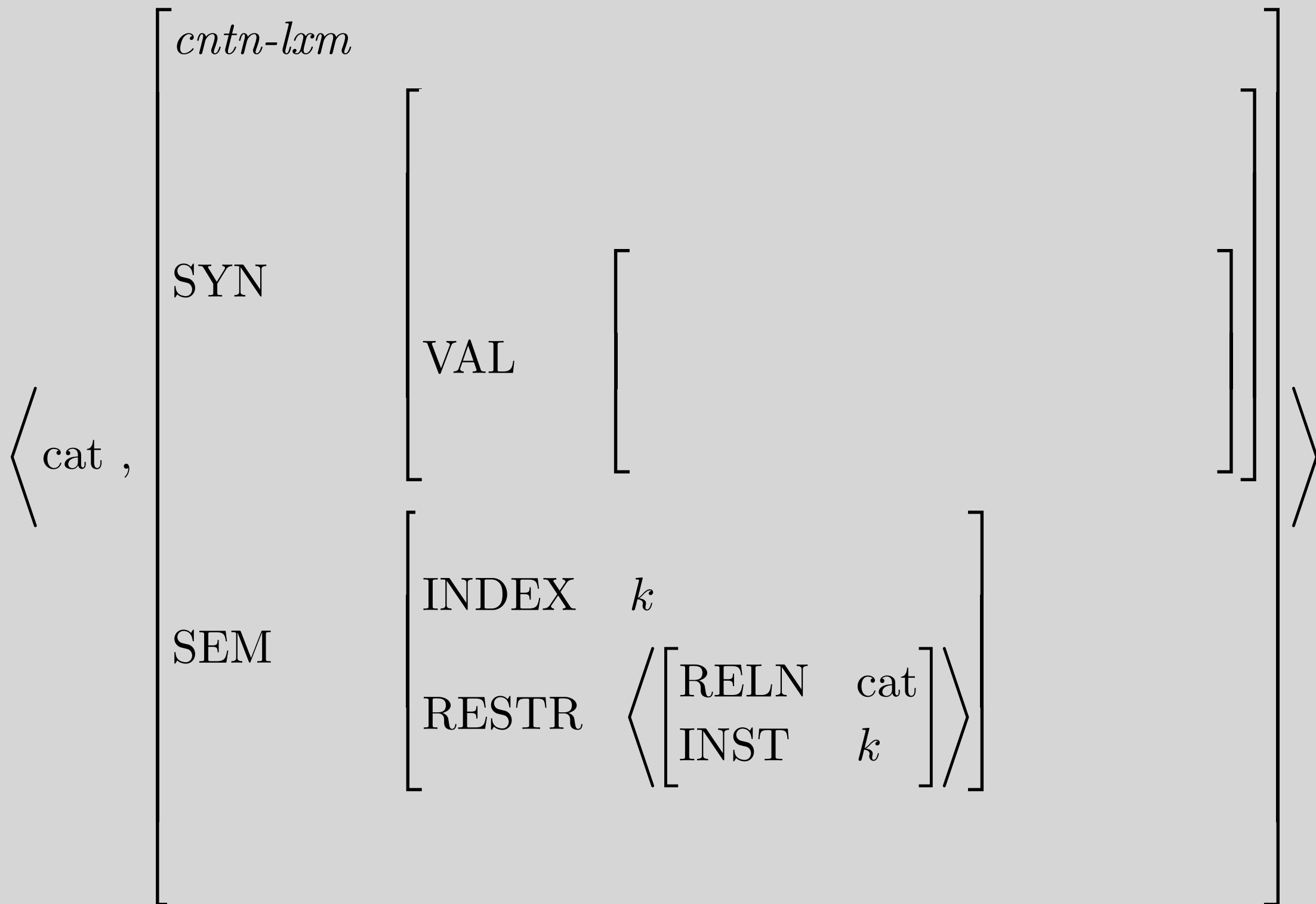
Practicalities - Applying Lexical Rules

- INPUT is a family of lexical sequences.
- OUTPUT is another family of lexical sequences.
 - ...usually a smaller family
 - ...usually a disjoint one
- The only differences between the families are those stipulated in the rule (or the rule's type).
- Similarities are handled by the constraints on *l-rule* and its subtypes.
- If we've written the LRs correctly, nothing is left underconstrained.

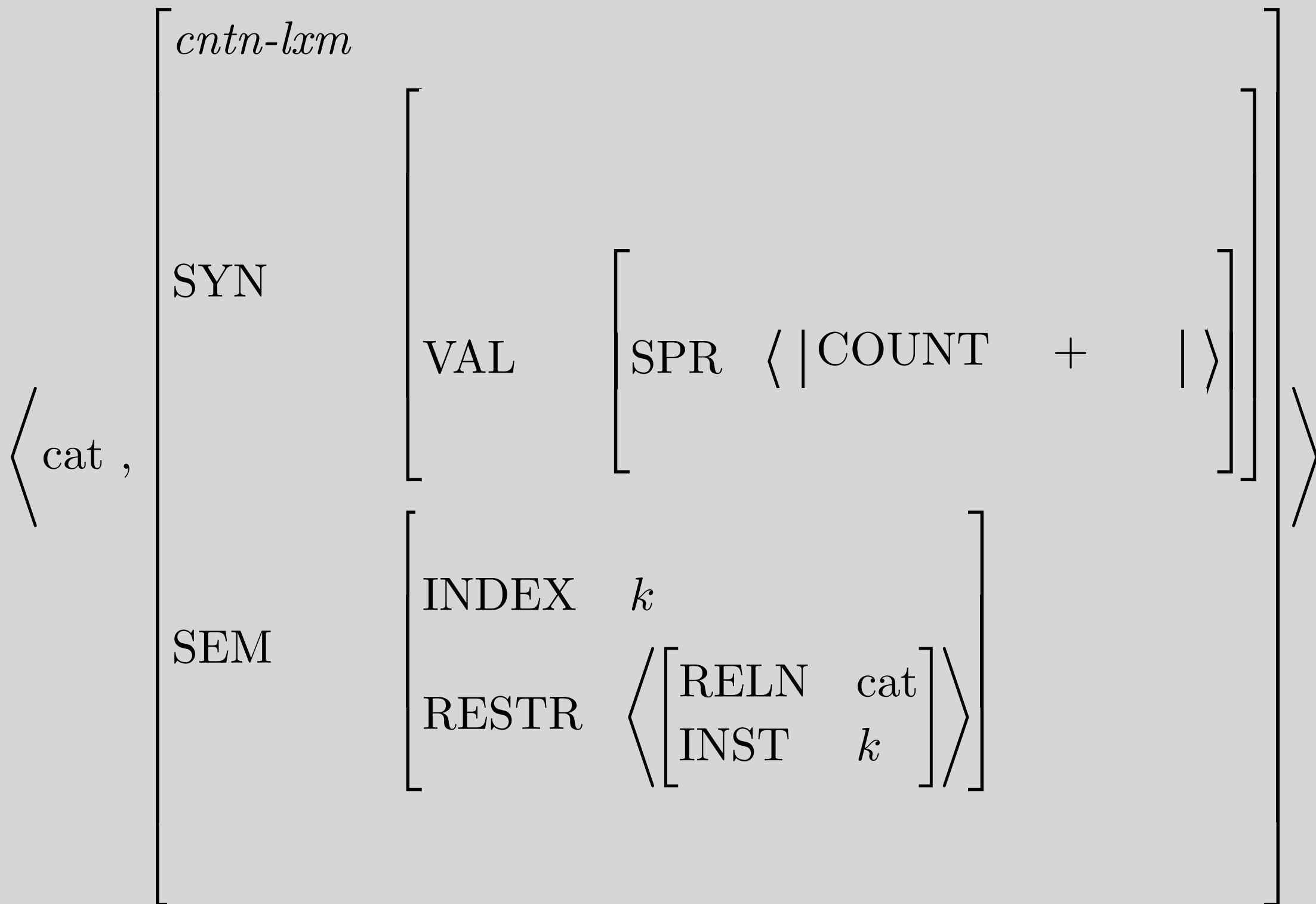
Example: Lexical Entry for *cat*



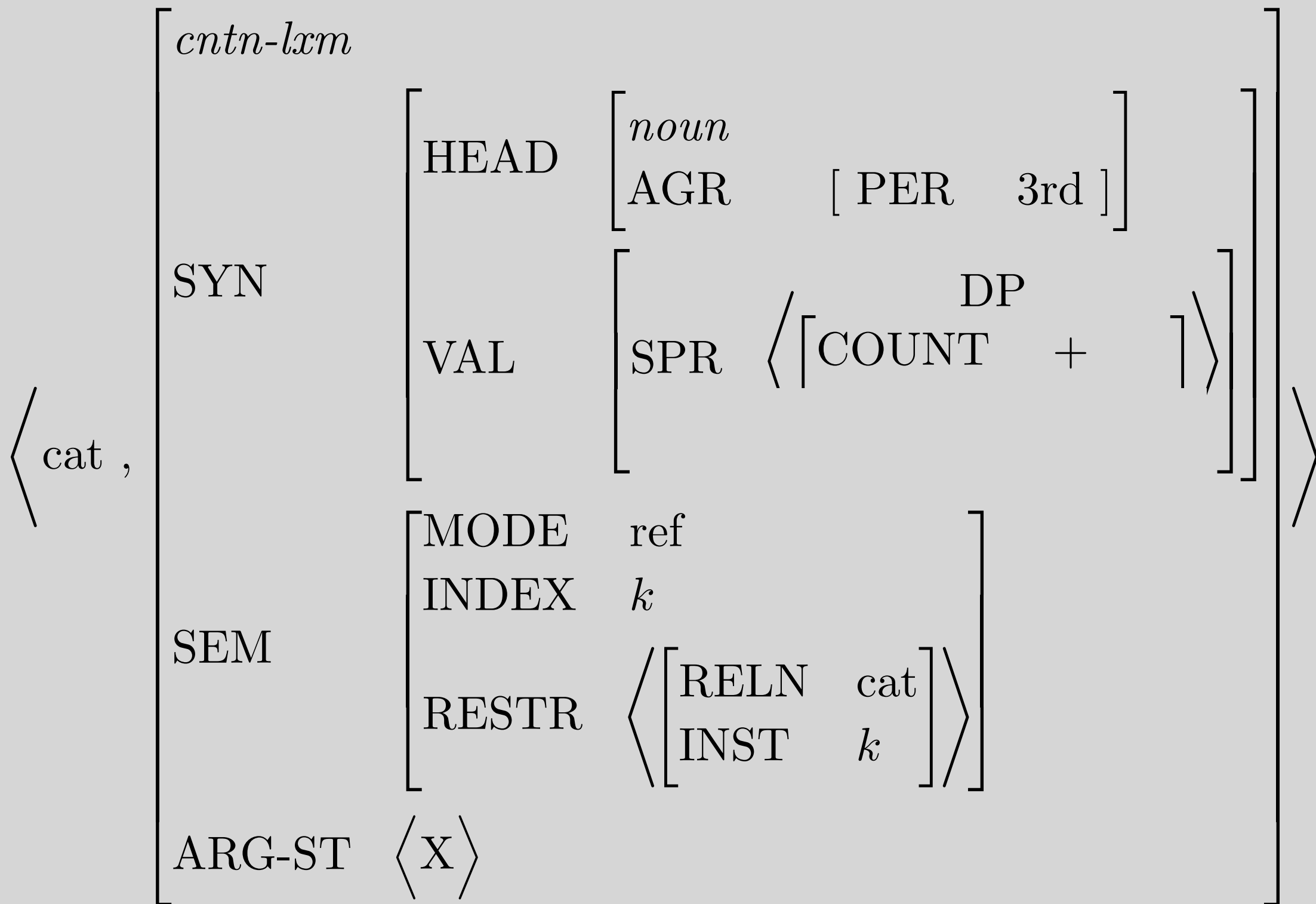
Example: *cat*, with inheritance



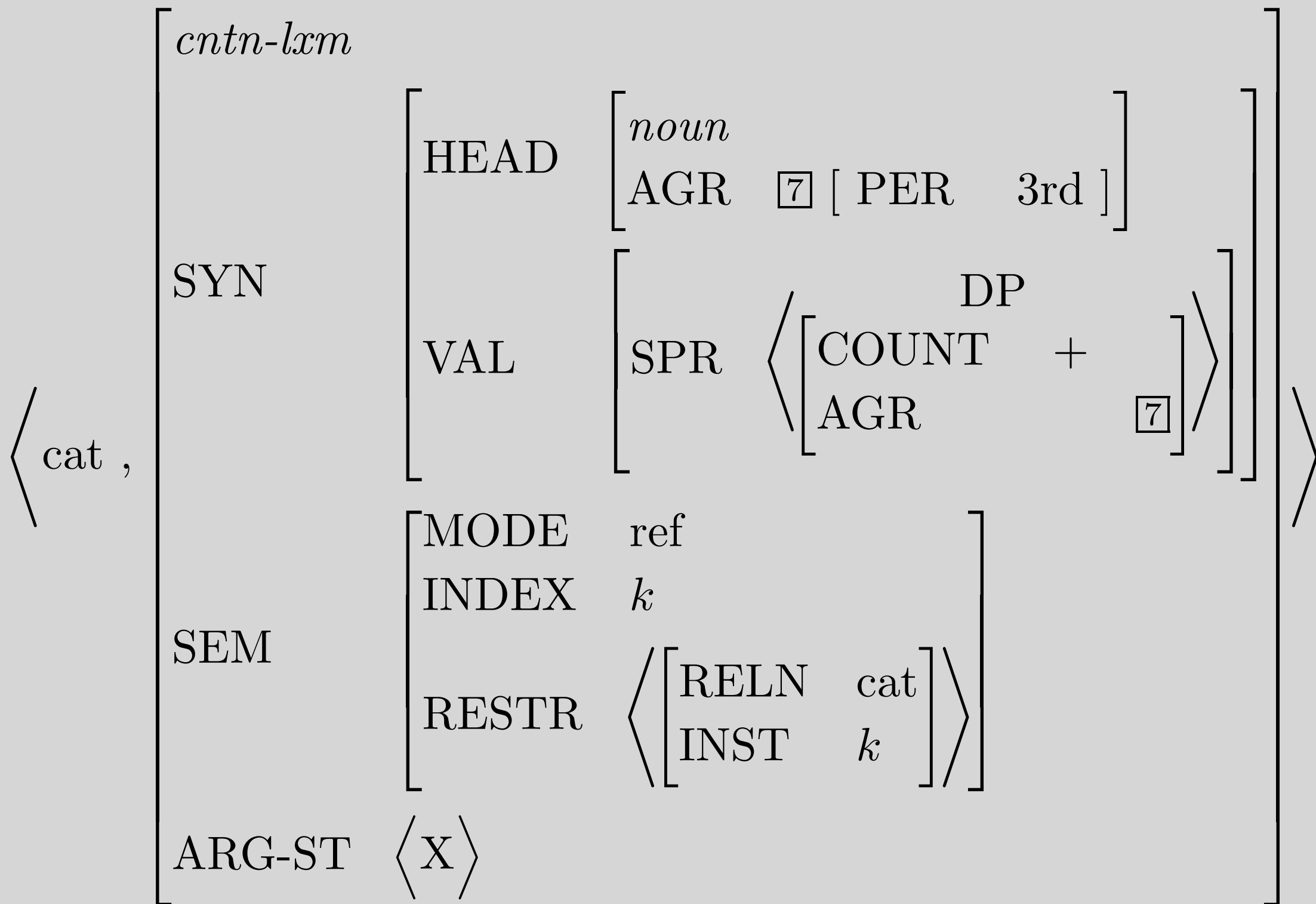
Example: *cat*, with inheritance



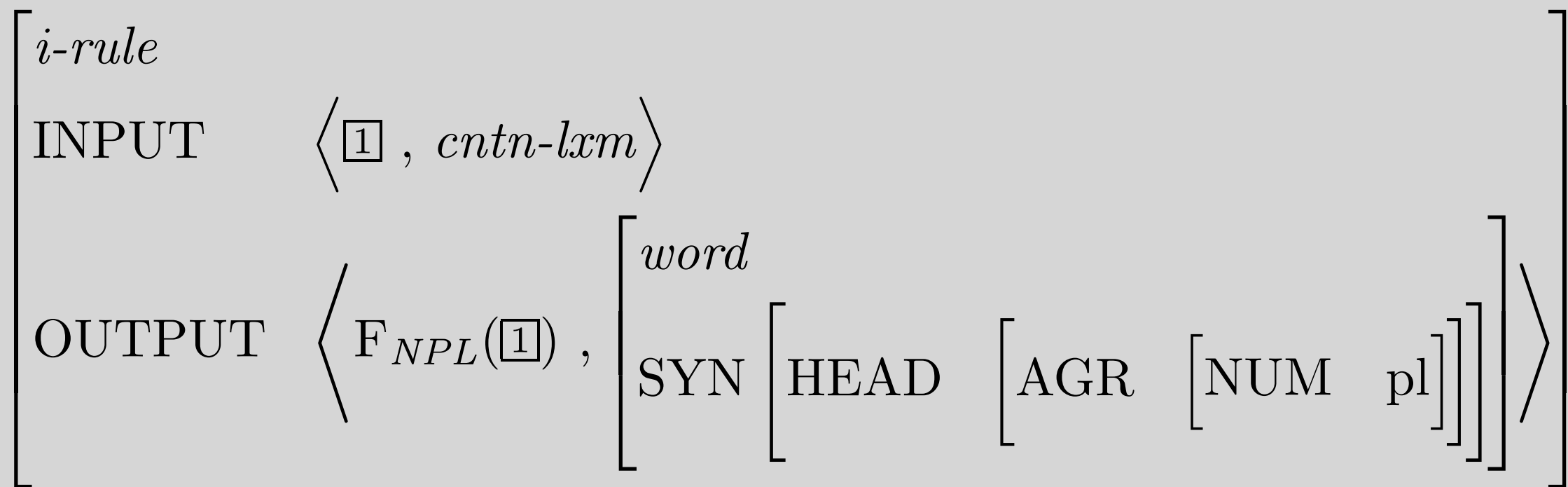
Example: *cat*, with inheritance



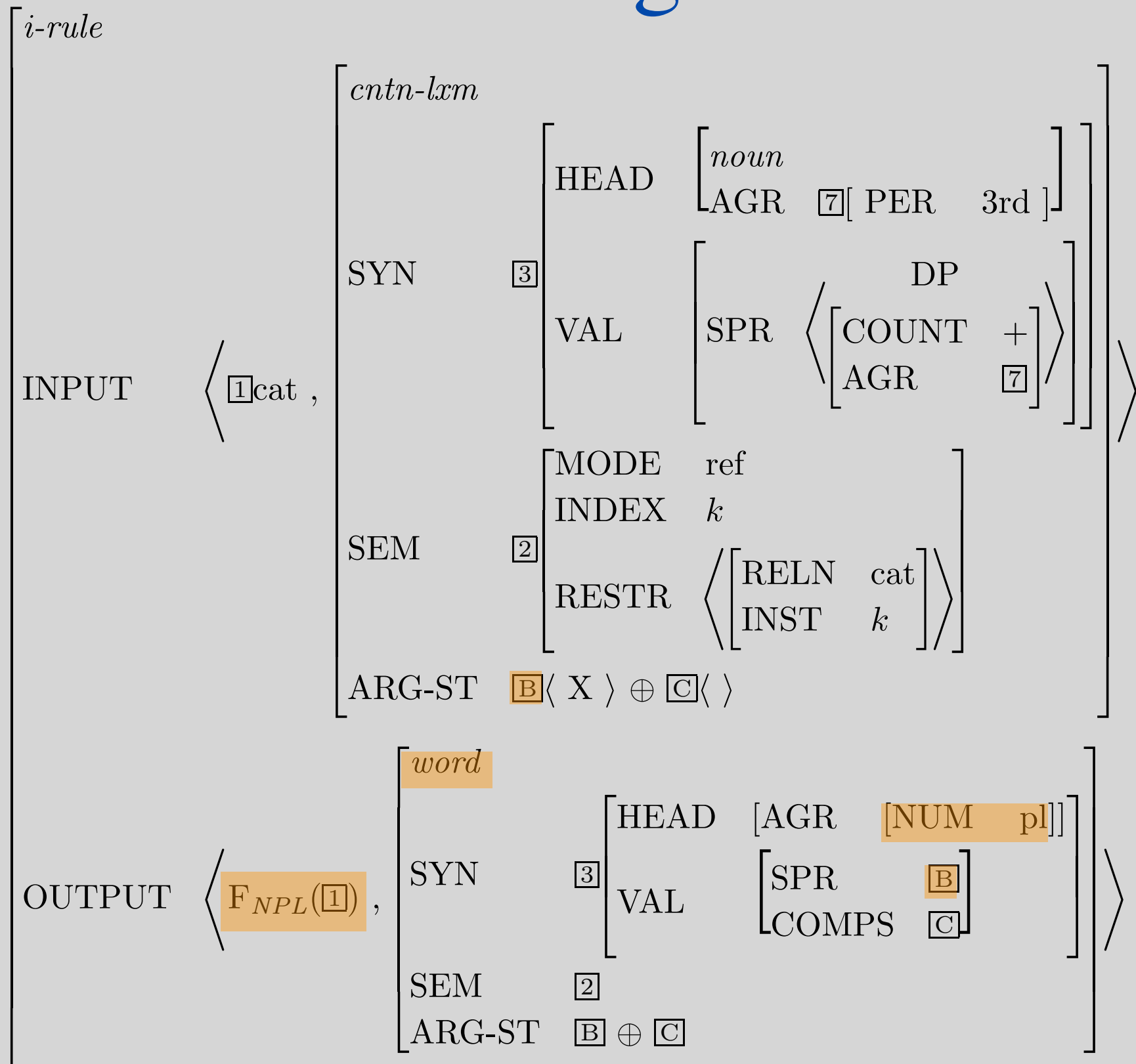
Example: *cat*, with inheritance



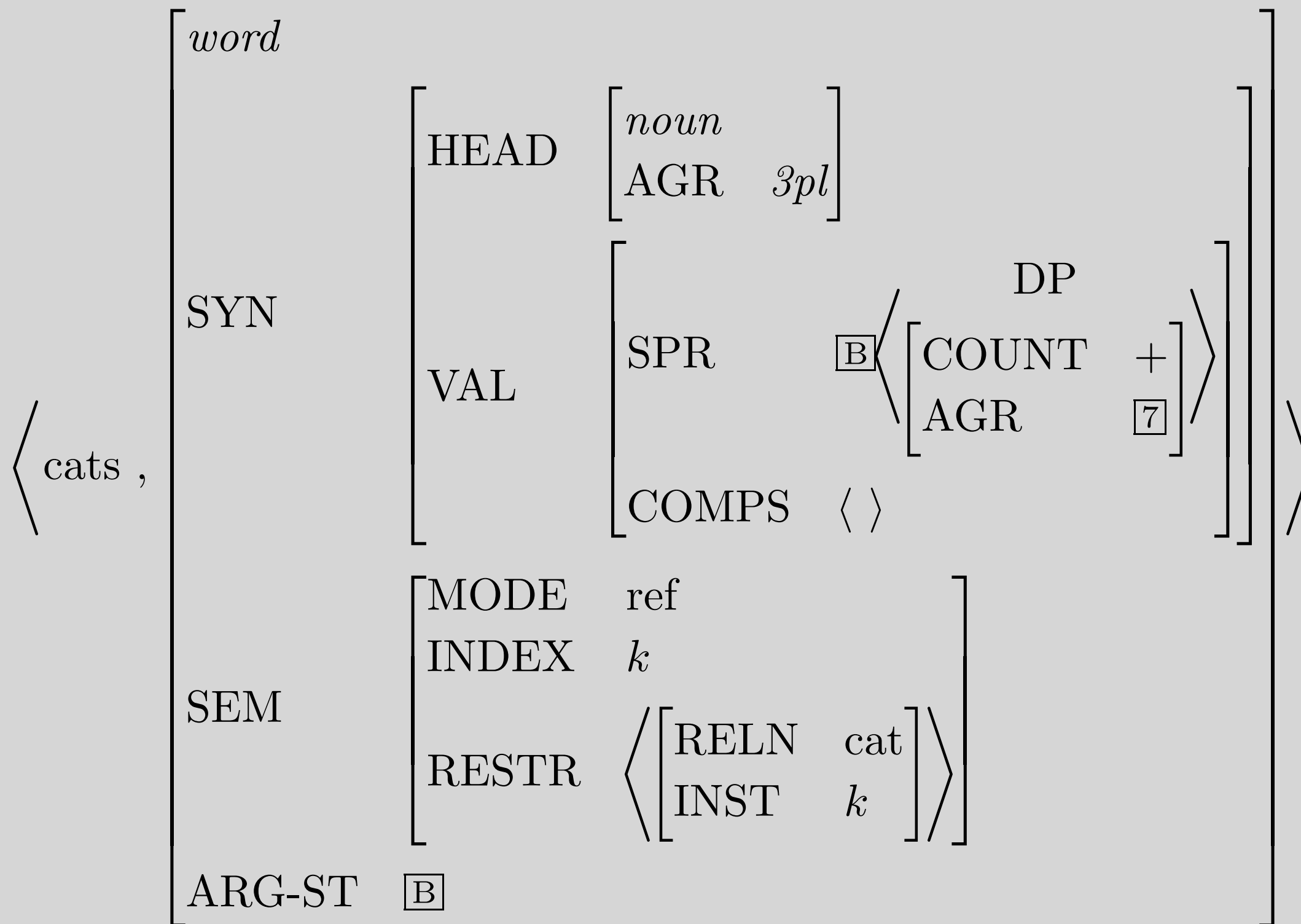
Plural Noun LR



Licensing *cats*



cats: The (family of) Lexical Sequence_(s)



Practicalities -- Writing Lexical Rules

- Determine the type of the LR.
- Determine the class of possible inputs.
- Determine what should change.
 - If INPUT and OUTPUT values are identified (by default or otherwise) and only OUTPUT value is mentioned, then...
information is added.

(Lexical sequences incompatible with that value are not possible inputs)
 - If INPUT and OUTPUT values are identified by default, but different values are given on the INPUT and OUTPUT of the rule, then...
information is changed.
 - If INPUT and OUTPUT values are identified by an inviolable constraint, but different values are given on the INPUT and OUTPUT of the rule, then...
there is no well-formed output

Constant lexemes

- What kinds of words are constant lexemes in our grammar?
- Why do we need a rule for these words?
- What would be an alternative analysis?

Constant Lexeme LR

<i>i-rule</i>	
INPUT	$\langle \boxed{1}, \textit{const-lxm} \rangle$
OUTPUT	$\left[\text{FIRST } \boxed{1} \right]$

- What keeps this from applying to, say, verb lexemes?
- Why is this an *i-rule*?

Poll!

ARG-ST & ARP

- Given the ARP, what do we need to specify about the valence properties of words?
- Why isn't the ARP a constraint on the type *lexeme*?

The Feature FORM

- Different inflected forms of verbs show up in different syntactic environments. Examples?
- These different forms are syntactically distinguished by the feature FORM, as assigned by lexical rules.
- FORM is also useful in our analyses of coordination and PP selection.

How do we rule these out?

- *Kim eat pizza.
- *Kim seems to eats pizza.
- *Dana helped Leslie [pack and moved].
- *Kim relies for Sandy.
- *Dana walked and Kim.

Overview

- How lexical rules fit in
- Three types of lexical rules, constraints
- Example: Plural noun lexical rule
- Advice on writing lexical rules
- Constant lexemes
- ARG-ST & ARP
- The feature FORM
- Reading Questions

Reading Questions

- lexeme
- lexical entry
- lexical rule
- lexical rule instantiation
- lexical sequence
- word structure

Reading Questions

- Where are the morphological functions like FFAST defined? Are they lexeme-specific, or are there more generic rules for them too?
- Our functions in the lexical rules still seem like black boxes, and we take the rules' word for it that the output will be correct, regardless of whether the verb is regular/irregular, or strange differences in plural for words like "moose" and "goose". Where in our grammar is the most intuitive and easiest way to handle all the exceptions to these inflectional functions?
- Is there any standard around the different forms of F? I saw F-er FFAST etc in the reading but it all seems arbitrary (although self-explanatory) to me.
- For languages that have morphological processes such as reduplication and partial reduplication, do we need to formulate different lexical rules to capture those processes?

Reading Questions

- I am also confused about why we can use an inflectional rule to generate the past tense, yet we need a derivational rule for the past participle. Does it have something to do with finite vs. non-finite and inflectional rules only generating words?
- Chapter 8.8 introduces a few derivational rules. Since there are lots of derivational morphemes, I am wondering do we have to formulate a d-rule for any derivational morpheme? For example, should civil -> civilize -> civilization all be generated by some d-rules so that they are derivable from civil?

Reading Questions

- Hmm, I'm wondering how agglutinative languages would look like with our grammar now (probably not well since we only have a few morphological things happening so far)?

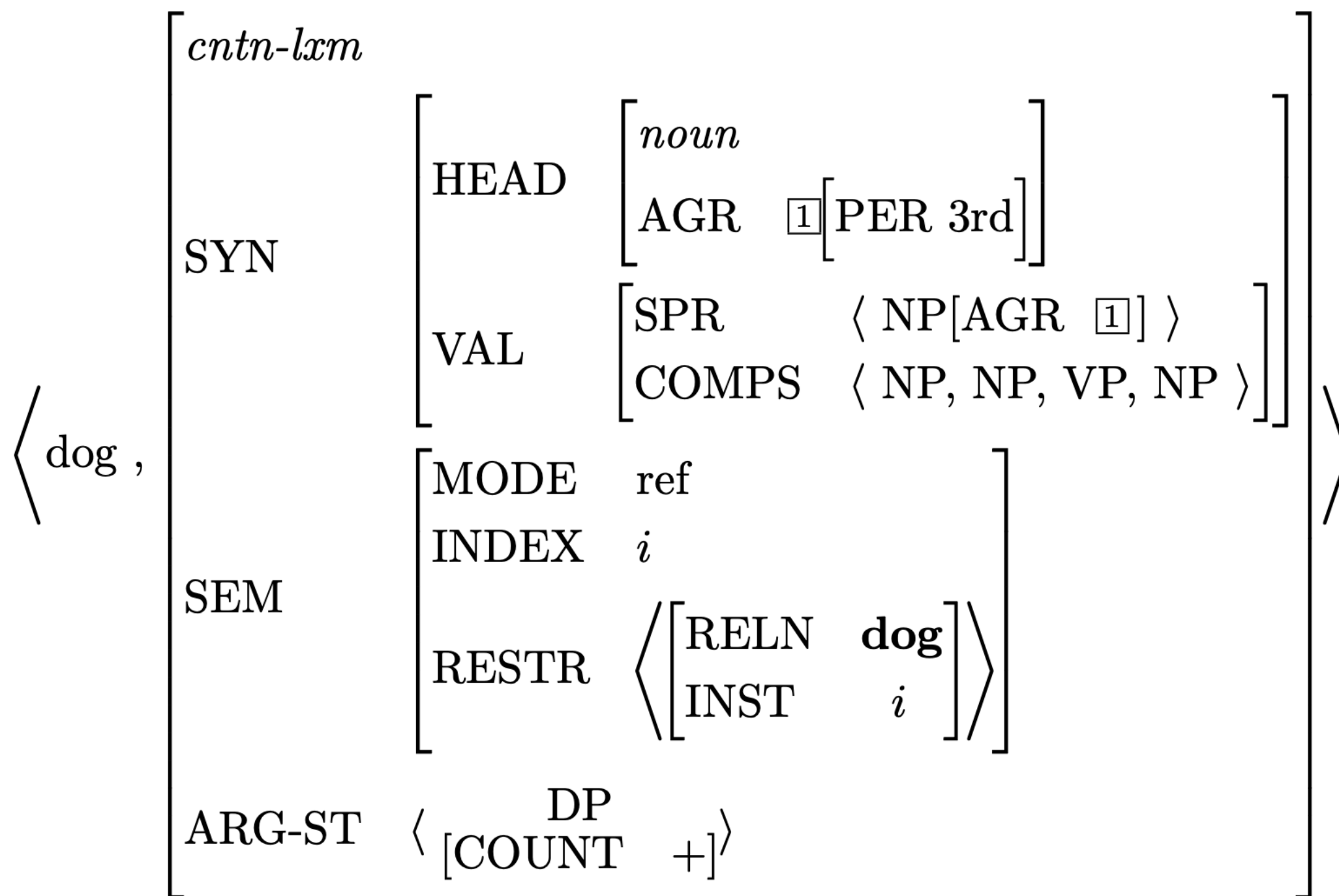
Reading Questions

- Can a feature structure directly above a word in a tree be a lexeme type or must they always be word type?
- Why are lexical rule instantiations not considered to be models of words or sentences even though they are fully specified features structures? Especially since the OUTPUT is inherited from the defeasible constraint of the INPUT?
- The reading mentions that lexical rule instantiations are objects that satisfy lexical rules, but the term "object" was a little vague to me. What does object mean in this case/what would be an example of a valid object?

Reading Questions

- It makes sense to me why our inflectional rules allow for the construction of well-formed feature structures of words, but I don't fully understand how they disallow entries like (74) in section 8.7.4. Specifically, I don't understand, "the SYN and ARG-ST values of the INPUT and the OUTPUT are identified, which means the INPUT will always, as a side-effect, also obey the ARP". I see the ARG-ST value of the output in (74), but where is the ARG-ST value of the input identified? I don't see how the ARP is followed by default.

(74) A lexical sequence that doesn't give rise to any words



Poll!

Reading Questions

- One of the recent assignments for my LING 570 class made use of finite-state transducers in which an input becomes exchanged for an output (e.g. every instance of 'aa' yields a 'b'). Do INPUT and OUTPUT work the same way, or at least similarly?

Reading Questions

- Does the '...' stand in for any tense (past/present/future)?
- The ...s in SEM are for tense, but the ...s in ARG-ST? Are those for complements?
- On page 247, the textbook discussed different kinds of FORM features, which the fin stands for both present or past tense, and on page 258, the past-tense verb lexical rule suggests using the function $F_{\{past\}}$ "for the morphological relation between verbal lexemes and their past tense forms". Is this mean that when we are draw trees, we need to put past tense like slept to something like $F_{\{past\}}(\text{Sleep})$, since we do not introduce tense in SYN?

Poll!

Reading Questions

- Why in the imperative rule, does the VP have FORM base - the rule basically assumes the VP has an unrealized 2nd person NP, shouldn't that make the FORM fin instead, i.e. intuitively, when you say "Eat rice" to someone, it's basically saying "(you) eat rice".
- In (76) there in F-er's ARG-ST there is a PP with FORM of. I might have missed it, but I'm not sure what of stands for. Does it mean the literal "PP -> of NP"?

Reading Questions

- Why do we differentiate between FORM psp and FORM pass in English?
- Additionally, if implementing this feature programmatically, how would one be able to differentiate between these two FORMs and other FORMs that are similar on the surface level but represent different types of entities? Is this is accounted for in the lexical rules?

Reading Questions

- page 248 mentions the idea of a "stand-alone" sentence. One thing I imagine this is useful for is distinguishing between Ss that can be embedded, but not sound grammatical when they're not embedded, vs. those that do sound fine alone in this sense. But I'm intrigued by the idea of a stand-alone sentence in the first place. Is "a complete thought" an unproblematic criterion? And how often do we run into problems when parsing strings as candidates for stand-alone sentences?
- On p. 248 it says that Ss are specified as fin since only these can be stand-alone sentences. Are utterances that are answers to questions ("Eating breakfast" in response to "What's Lisa doing?" for example) not considered sentences for our grammar?

Reading Questions

- With the introduction of FORM for verbs, I'm still unclear how we handle what I guess are traditionally called "verbal nouns" (like gerunds, infinitives) that syntactically appear where NPs would. Seeing VPs appear as the complement as other VPs is nothing new for our grammar, but can VPs appear as specifiers (examples: "To live is to suffer", "Painting requires practice")? Or in this case have they undergone a lexical rule to effectively become nouns? I am especially curious about infinitives as my first instinct is to treat them as a (noun? verb?) phrase with a prep SPR and a FORM base verb HEAD.
- In my previous syntax classes, we had other labels for FORM such as gerund, preterite, etc. What is the reason for the different types of labels in syntax? And how are the ones we're using different/similar?

Reading Questions

- I always love footnotes, and I was fascinated by footnote 34 (p. 258), which is about the creation of specialized lexical rules for *be*. What do these specialized rules look like? What are some other cases where we'd need a solution like that?

Poll!

Reading Questions

- Are the i-rule and d-rule lexical rules applicable to all / most languages? It seems to me, since the i-rule invariably requires the SYN values of the INPUT and OUTPUT to be identical, but the d-rule only defeasibly requires this, that
 - 1) the i-rule tells us how to alter certain lexemes into lexically-aligned specific words (e.g. drives from drive), and
 - 2) the d-rule tells us how to generate new words from lexemes that aren't lexically-aligned (e.g. *driver* from *drive*).
- Looking at Arabic, with the triliteral root system, it makes me wonder if we should treat the root "k-t-b" as a base form verb and apply i-rules and d-rules similarly as we do in English, or if a different system of rules is more appropriate?