

Ling 566  
Oct 29, 2013  
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*, ...

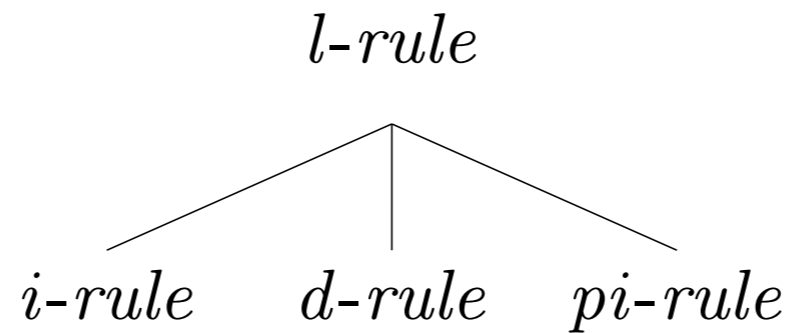
# Parsimony & Plausibility

- 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*

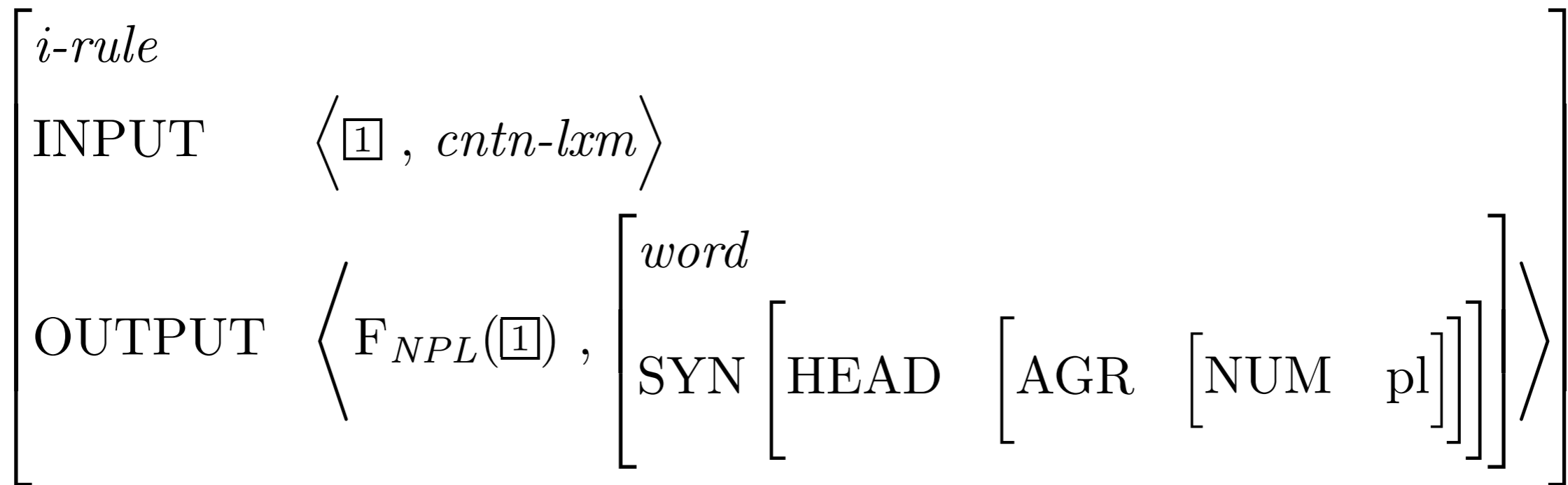


$$l\text{-rule} : \left[ \begin{array}{l} \text{INPUT} \quad l\text{-sequence} \left\langle X, [ \text{SEM} \quad / \boxed{2} ] \right\rangle \\ \text{OUTPUT} \quad l\text{-sequence} \left\langle Y, [ \text{SEM} \quad / \boxed{2} ] \right\rangle \end{array} \right]$$

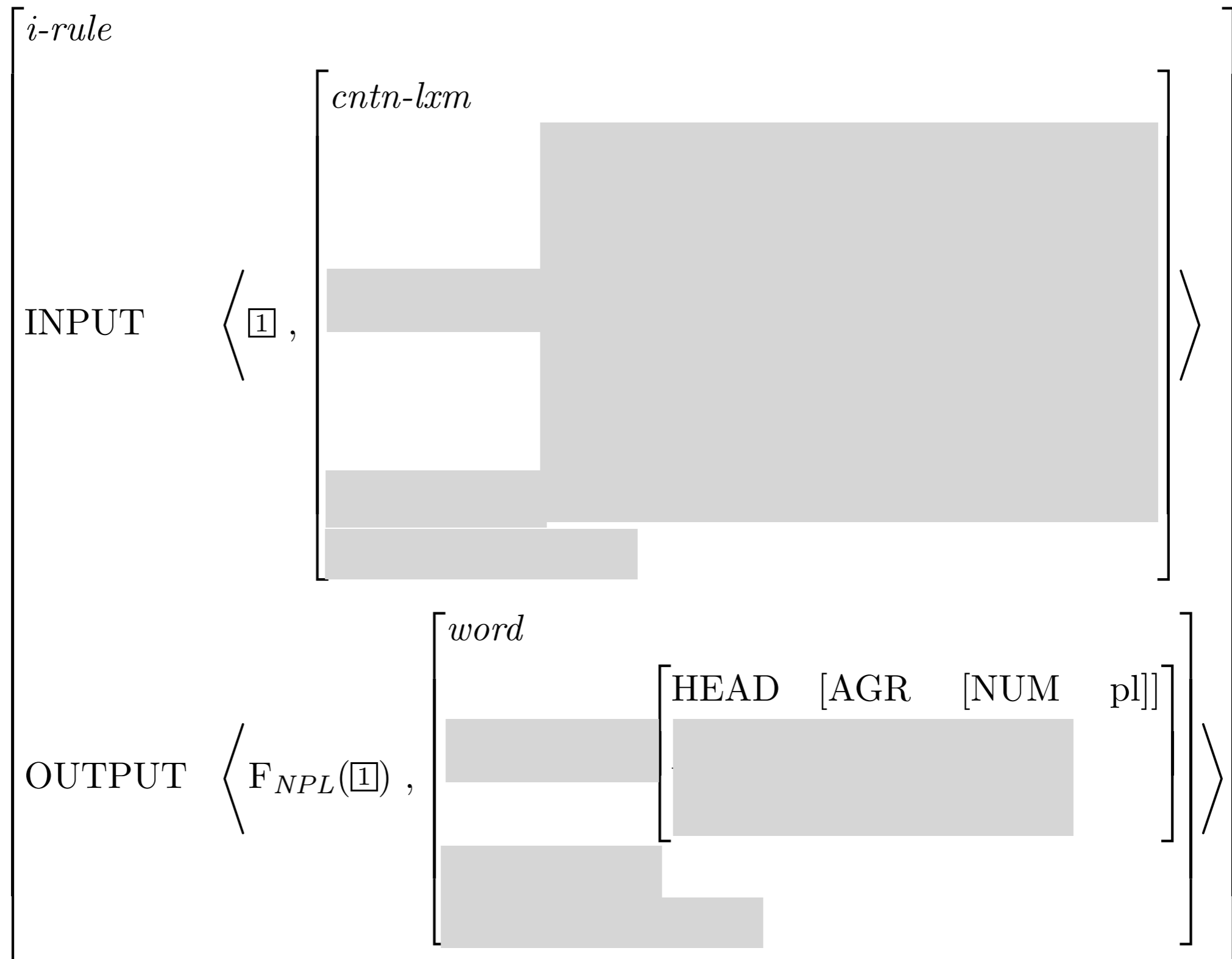
$$i\text{-rule} : \left[ \begin{array}{l} \text{INPUT} \quad \left\langle X, \left[ \begin{array}{l} \textit{lexeme} \\ \text{SYN} \quad \boxed{3} \\ \text{ARG-ST} \quad \boxed{A} \end{array} \right] \right\rangle \\ \text{OUTPUT} \quad \left\langle Y, \left[ \begin{array}{l} \textit{word} \\ \text{SYN}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

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

# Plural Noun LR

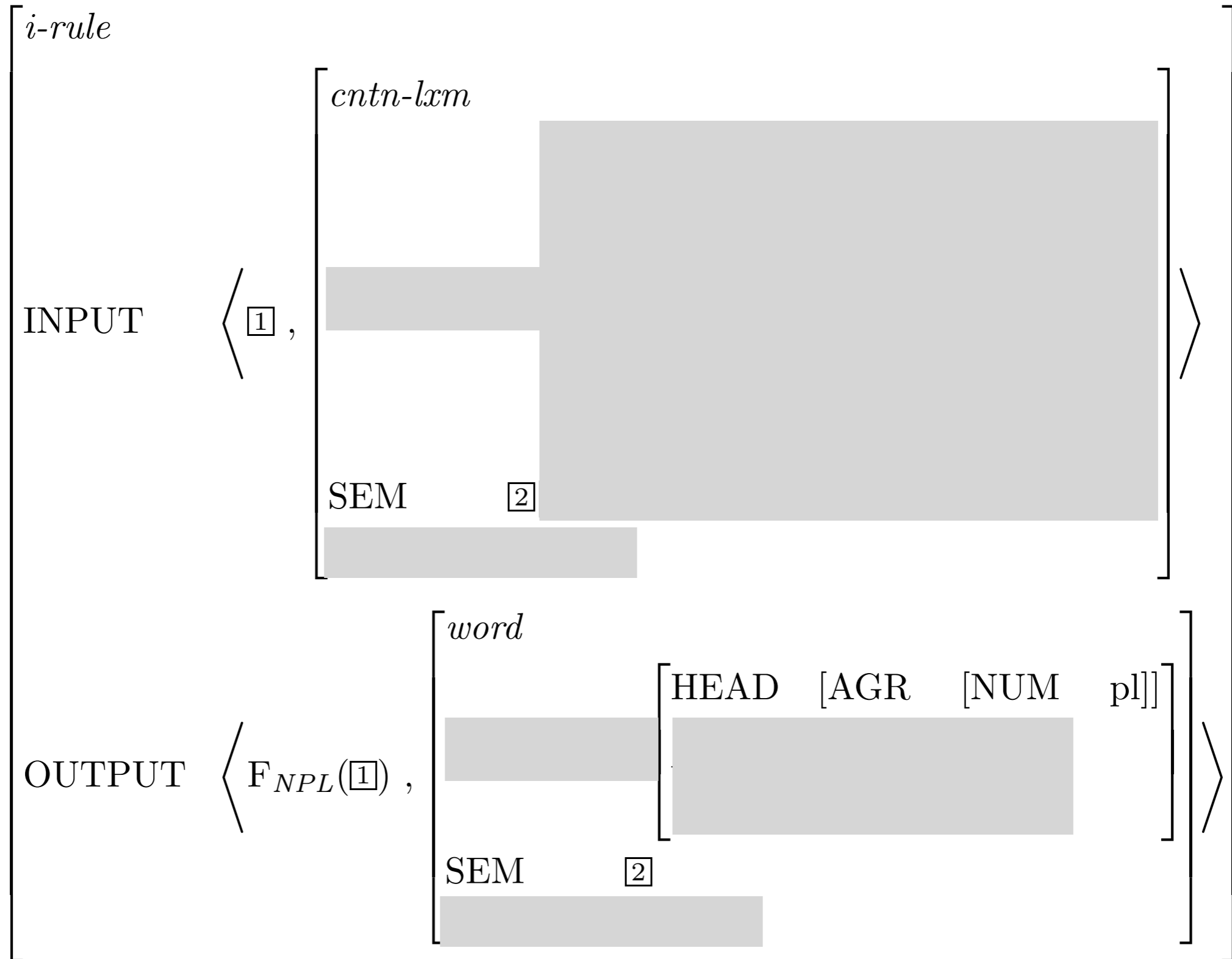


# Plural Noun LR with Inherited Constraints

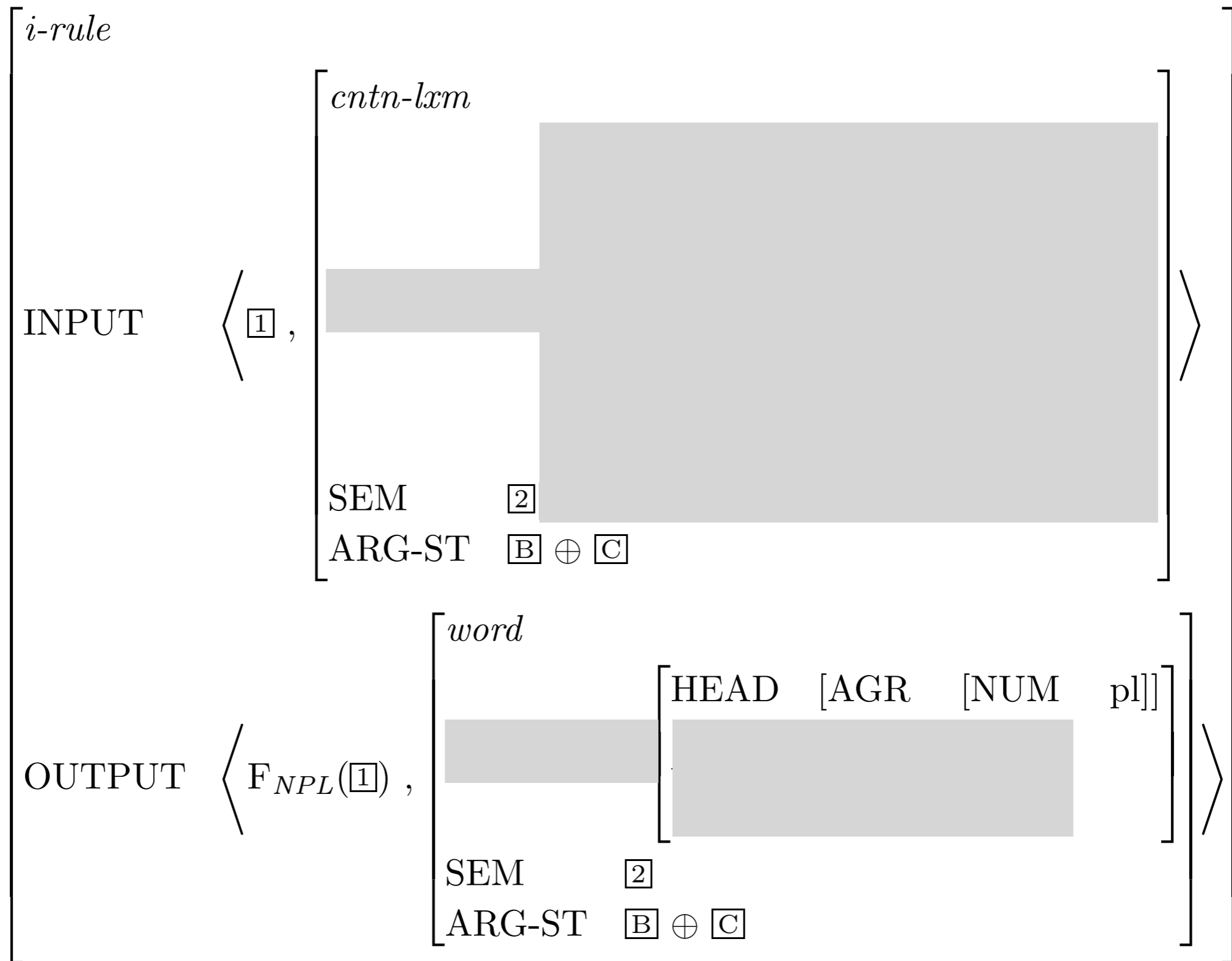




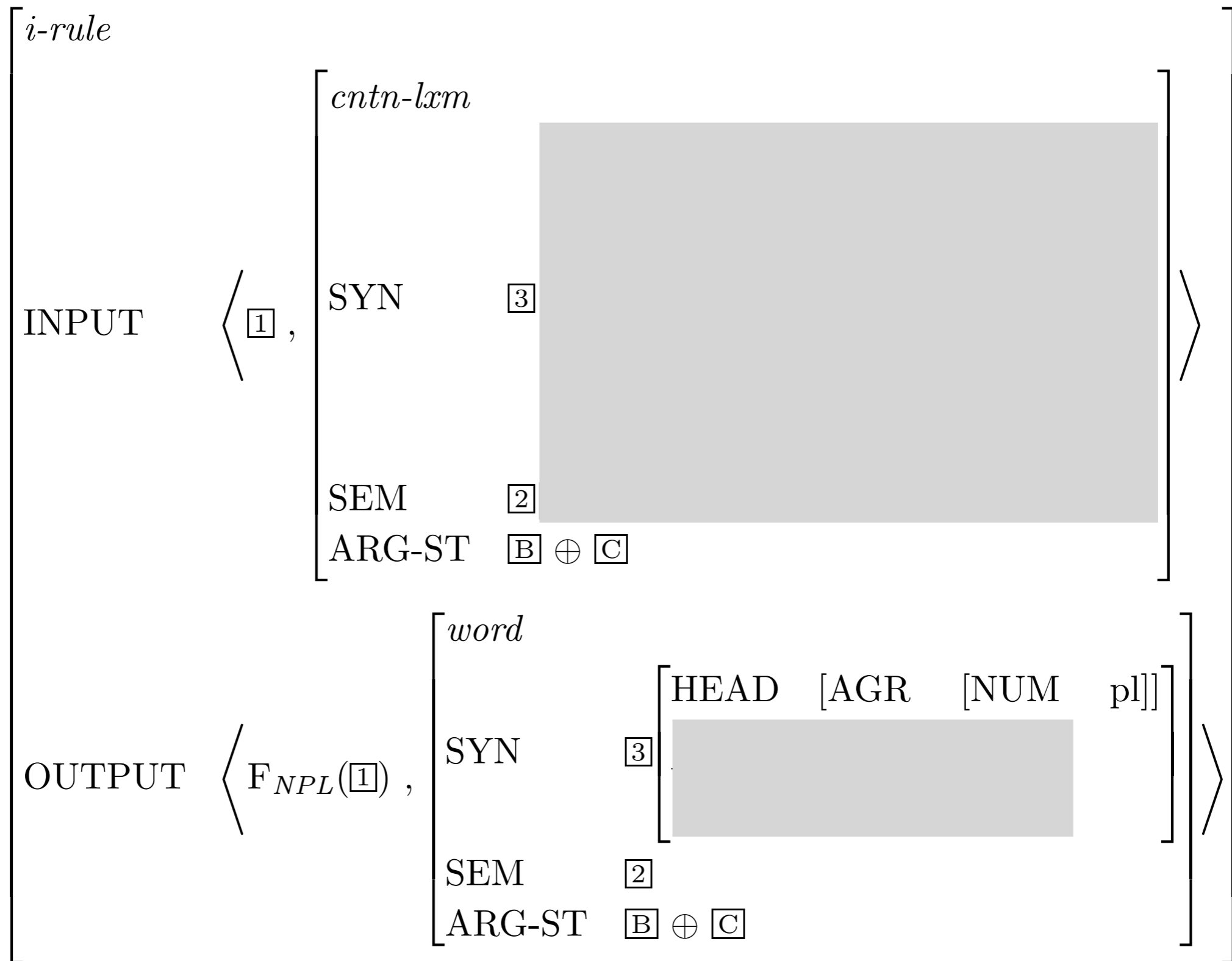
# Plural Noun LR with Inherited Constraints



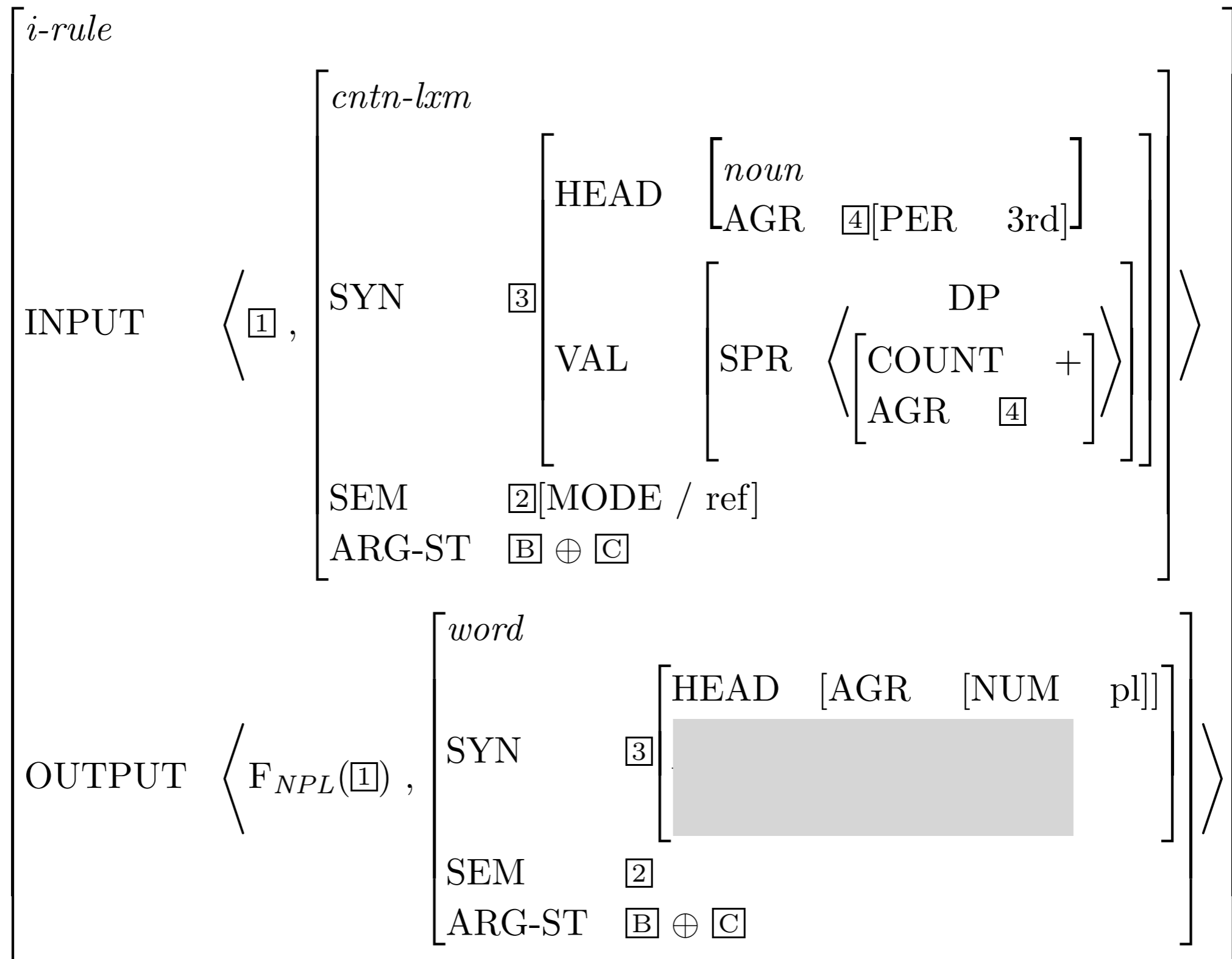
# Plural Noun LR with Inherited Constraints



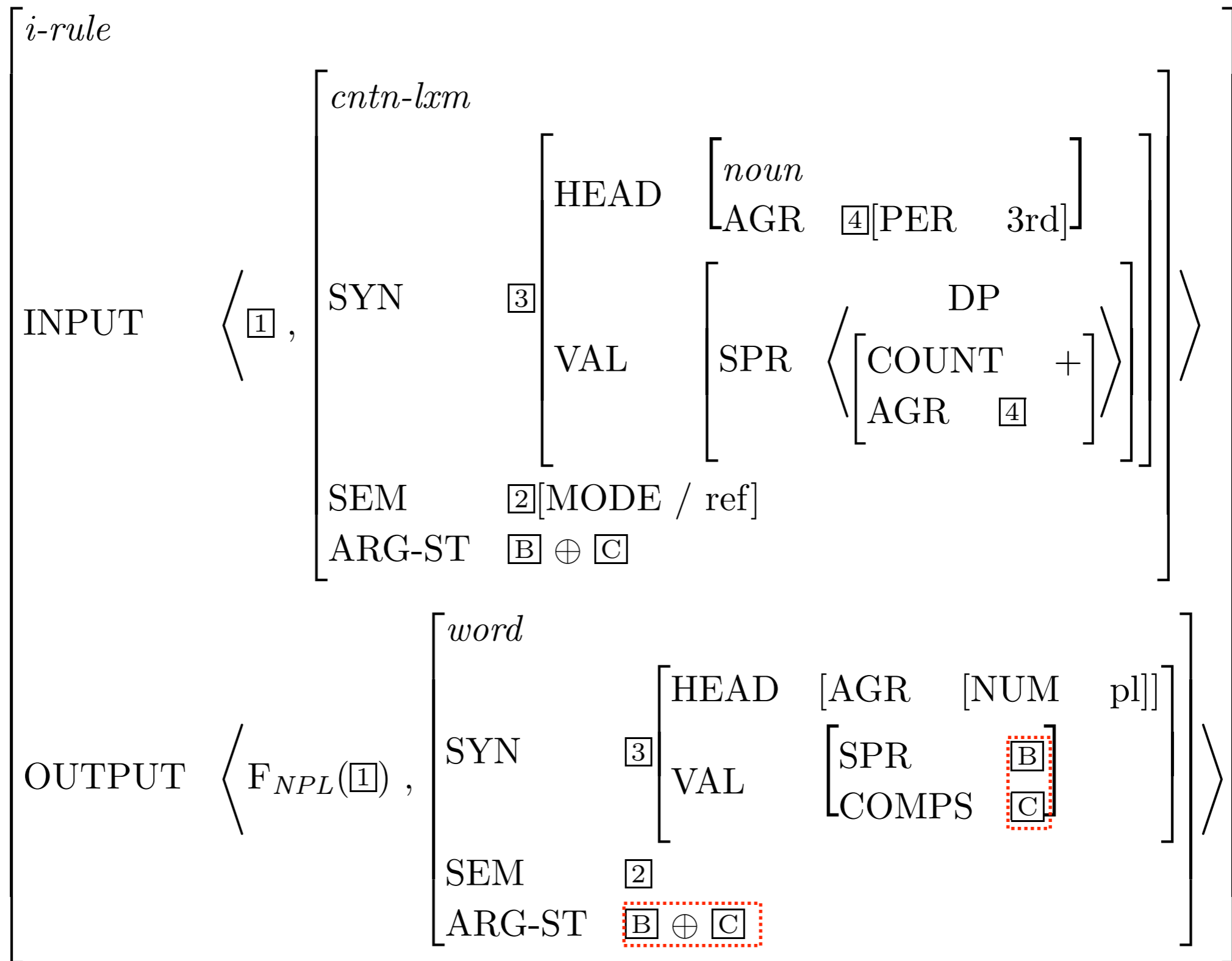
# Plural Noun LR with Inherited Constraints



# Plural Noun LR with Inherited Constraints



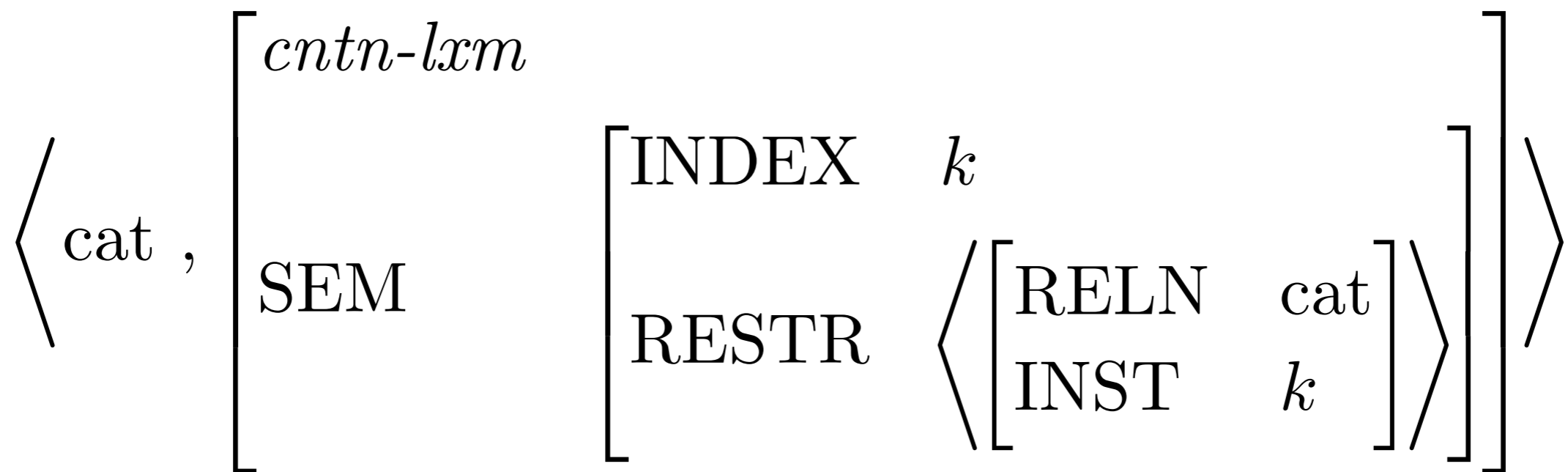
# Plural Noun LR with Inherited Constraints



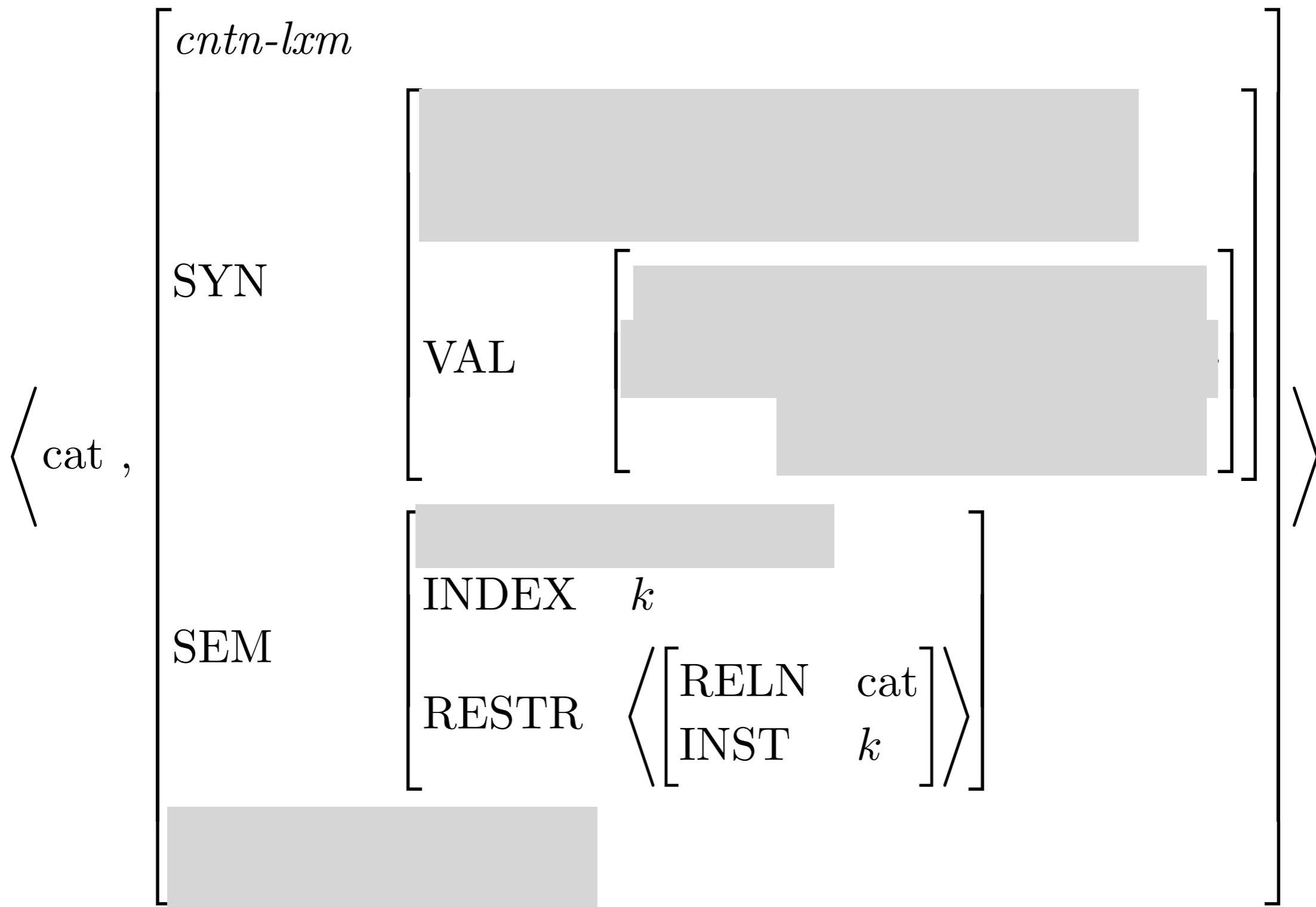
# Practicalities - Applying Lexical Rules

- 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 LR's correctly, nothing is left underconstrained.

# Example: Lexical Entry for *cat*

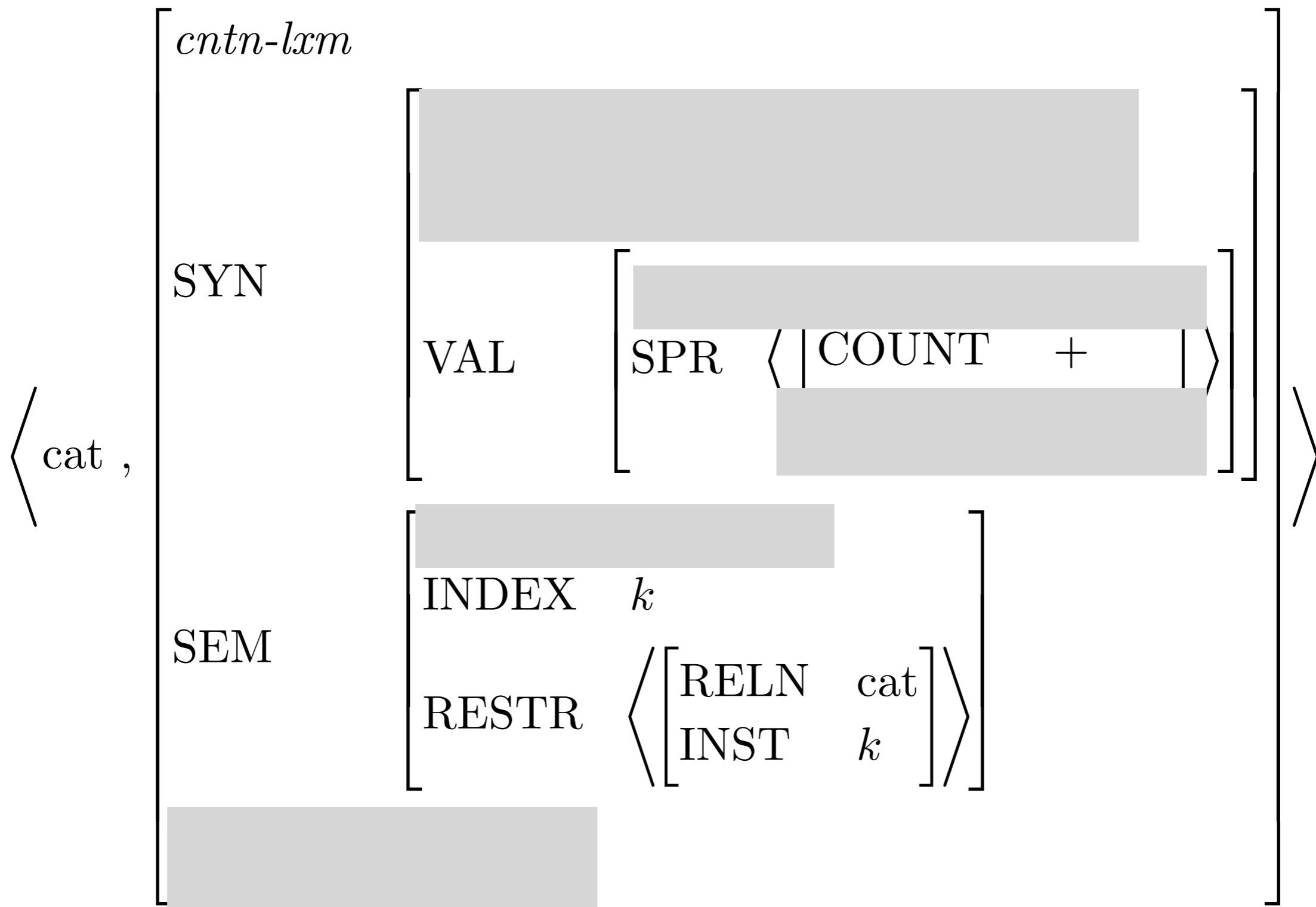


# Example: *cat*, with inheritance

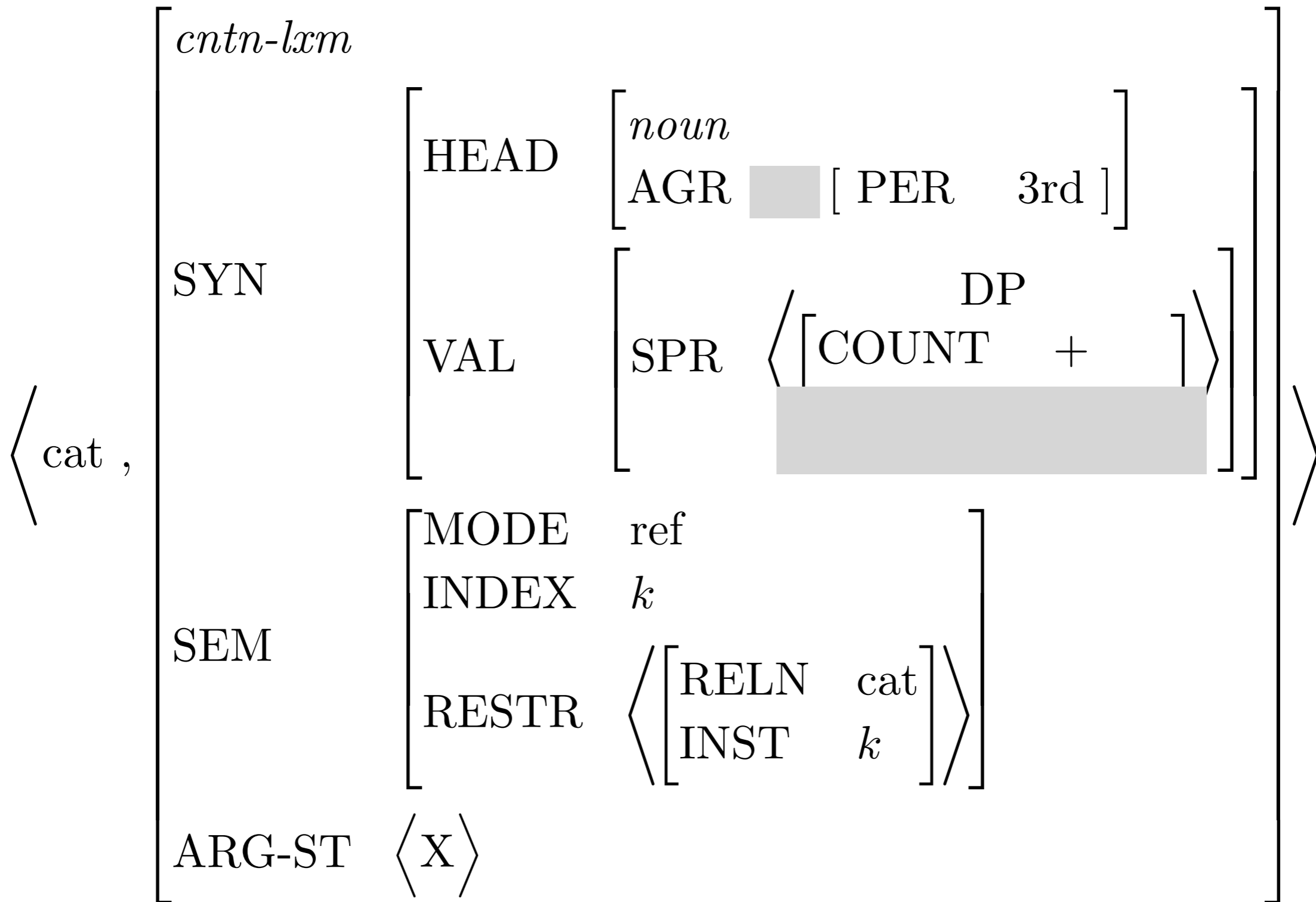




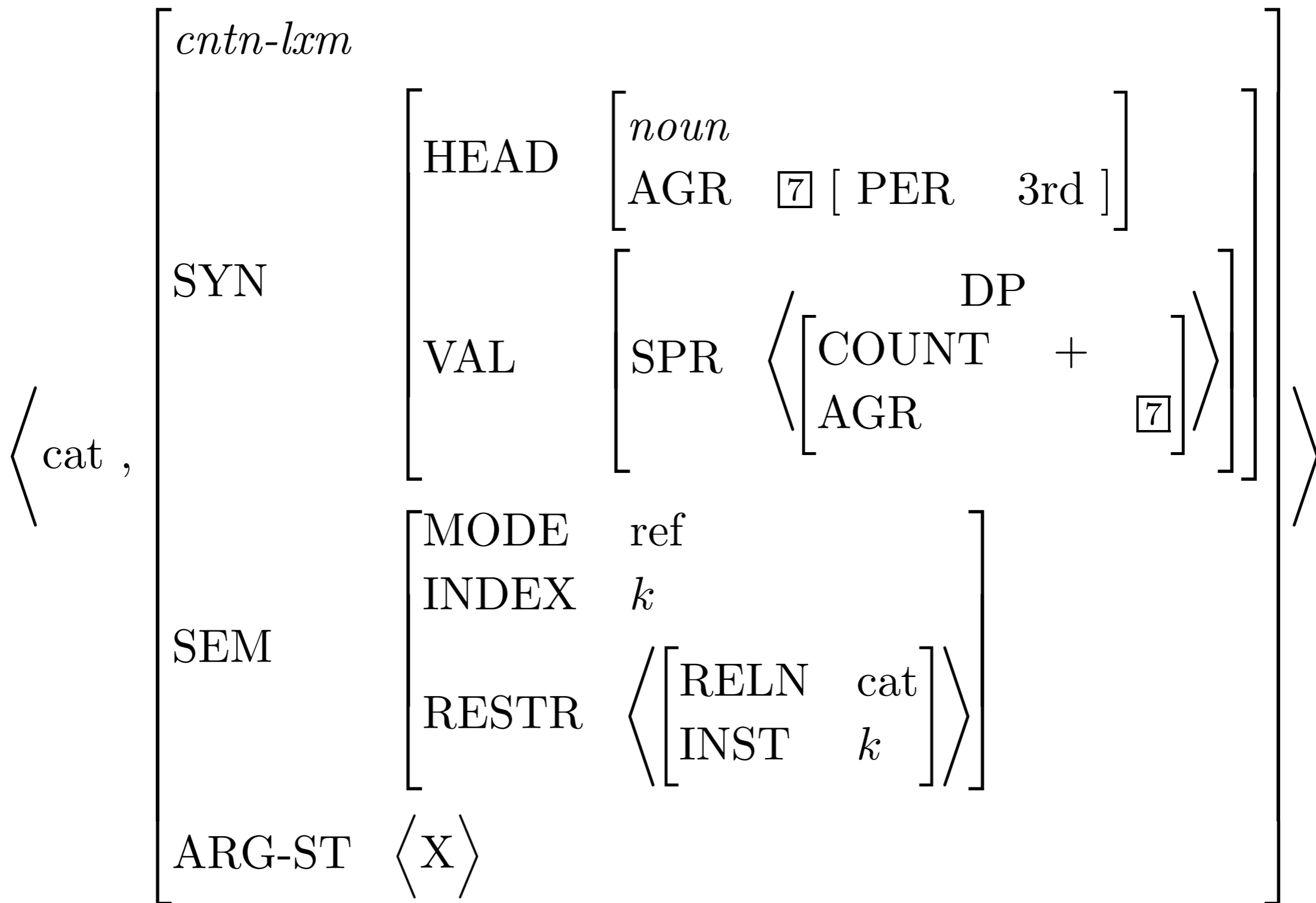
# Example: *cat*, with inheritance



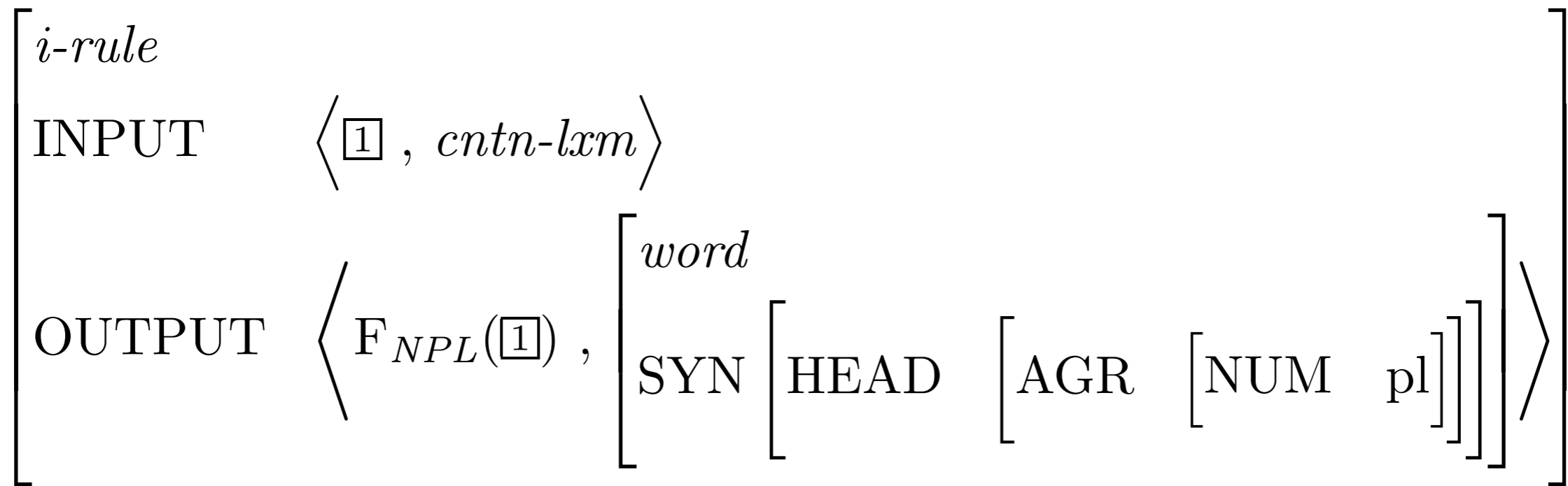
# Example: *cat*, with inheritance



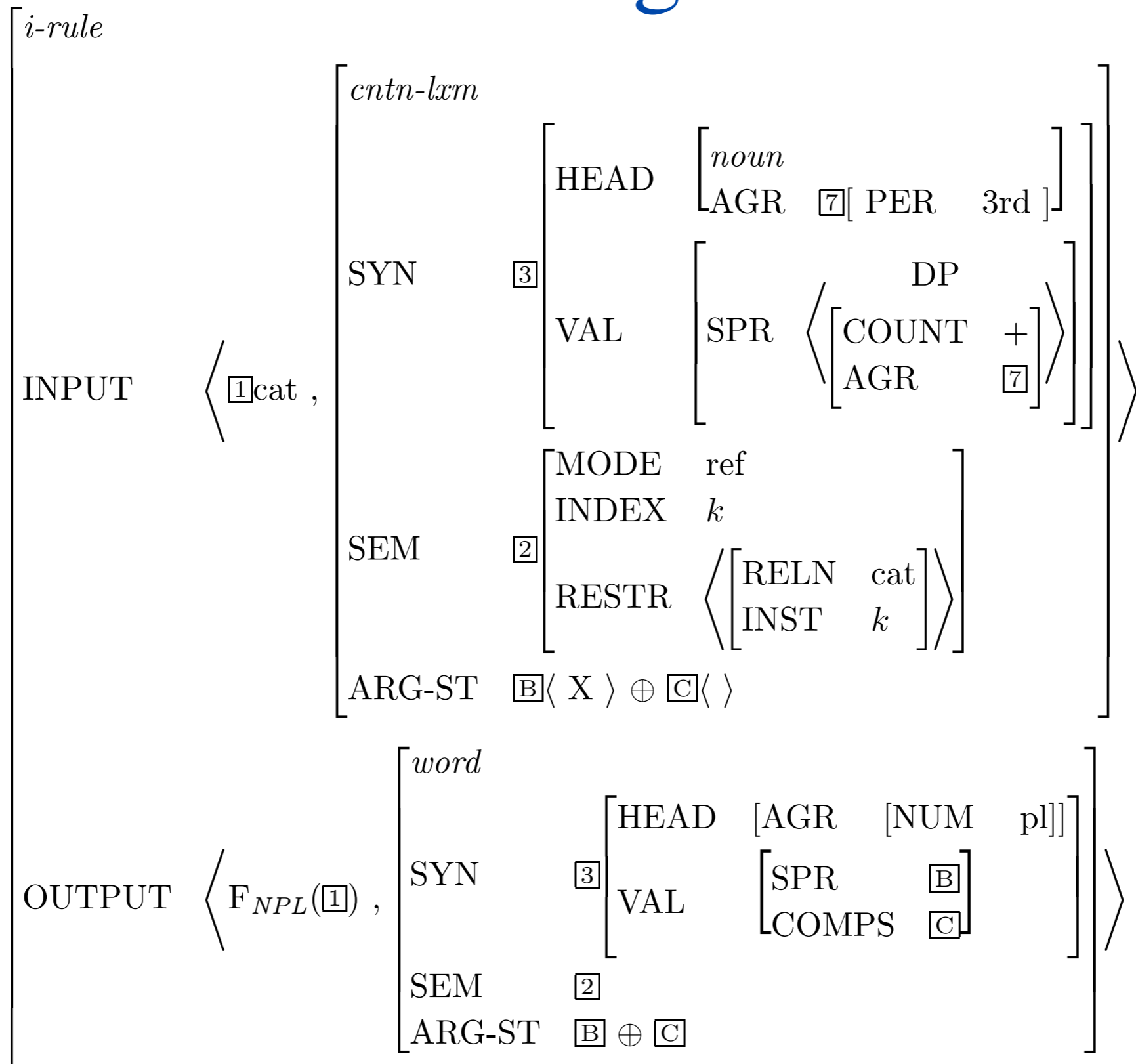
# Example: *cat*, with inheritance



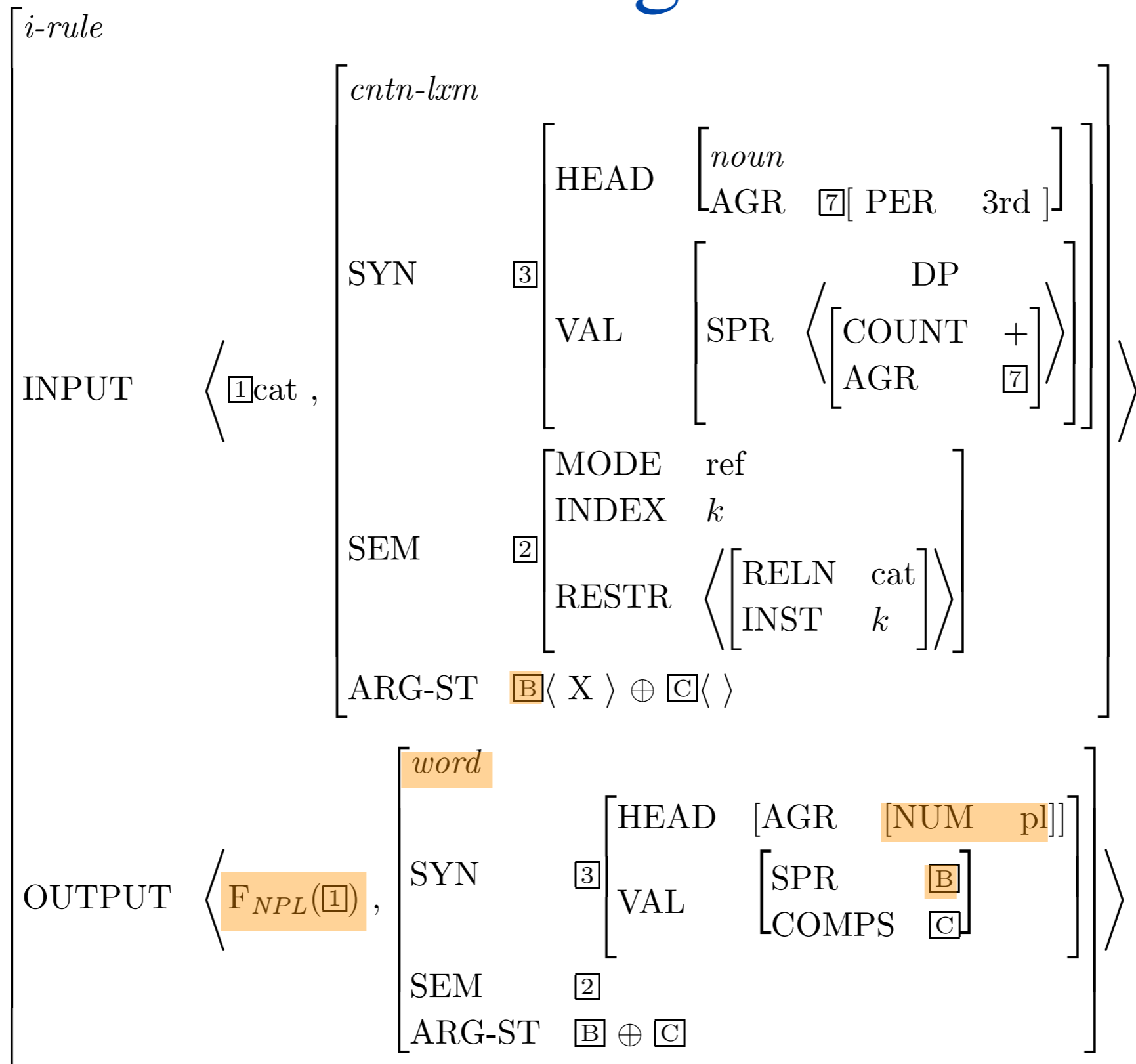
# Plural Noun LR



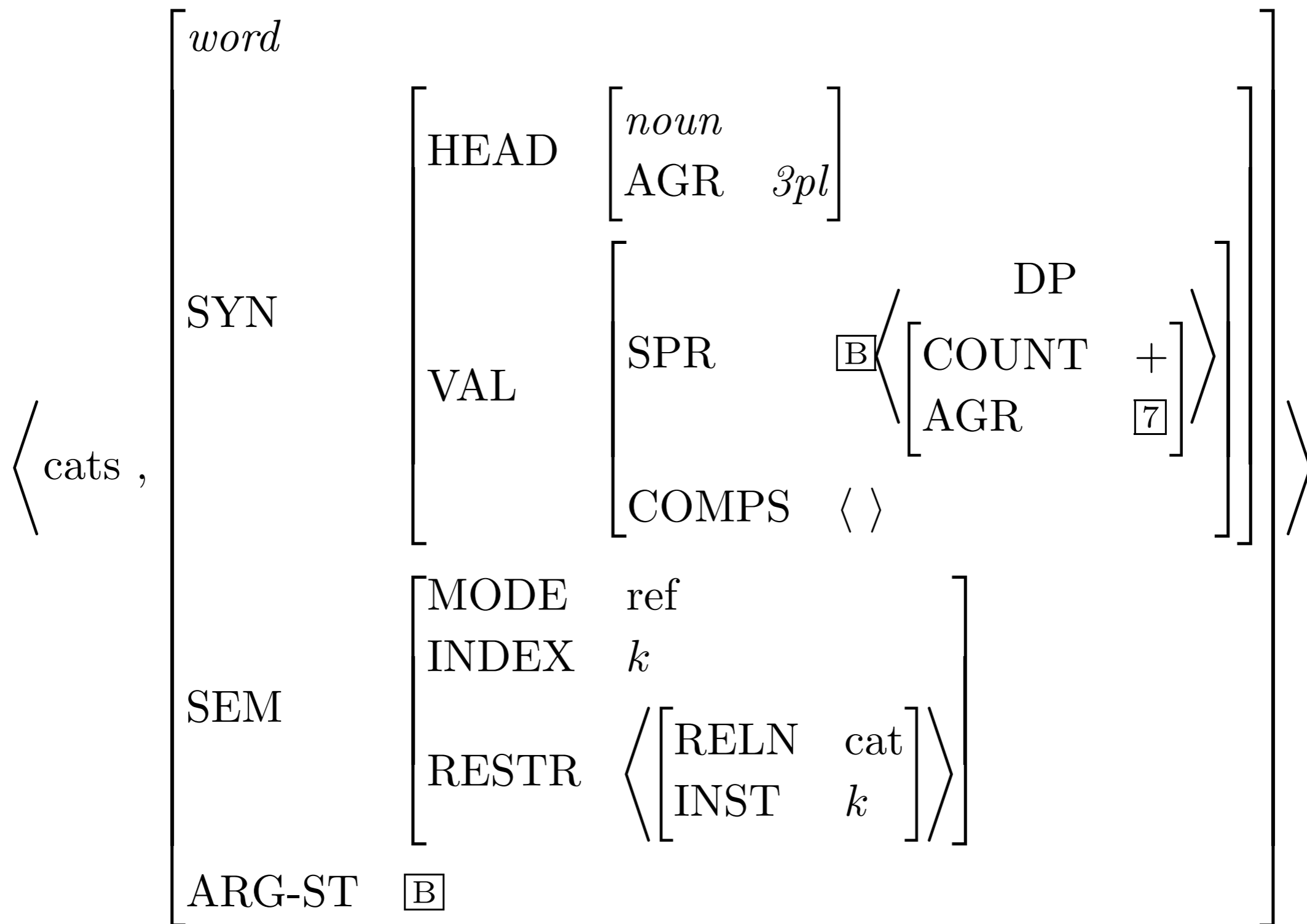
# Licensing *cats*



# Licensing *cats*



# *cats*: The Lexical Sequence



# 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

$$\left[ \begin{array}{l} i\text{-rule} \\ \text{INPUT} \quad \langle \boxed{1}, \text{const-lxm} \rangle \\ \text{OUTPUT} \quad \left[ \text{FIRST} \quad \boxed{1} \right] \end{array} \right]$$

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

# 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.

# What rules these out?

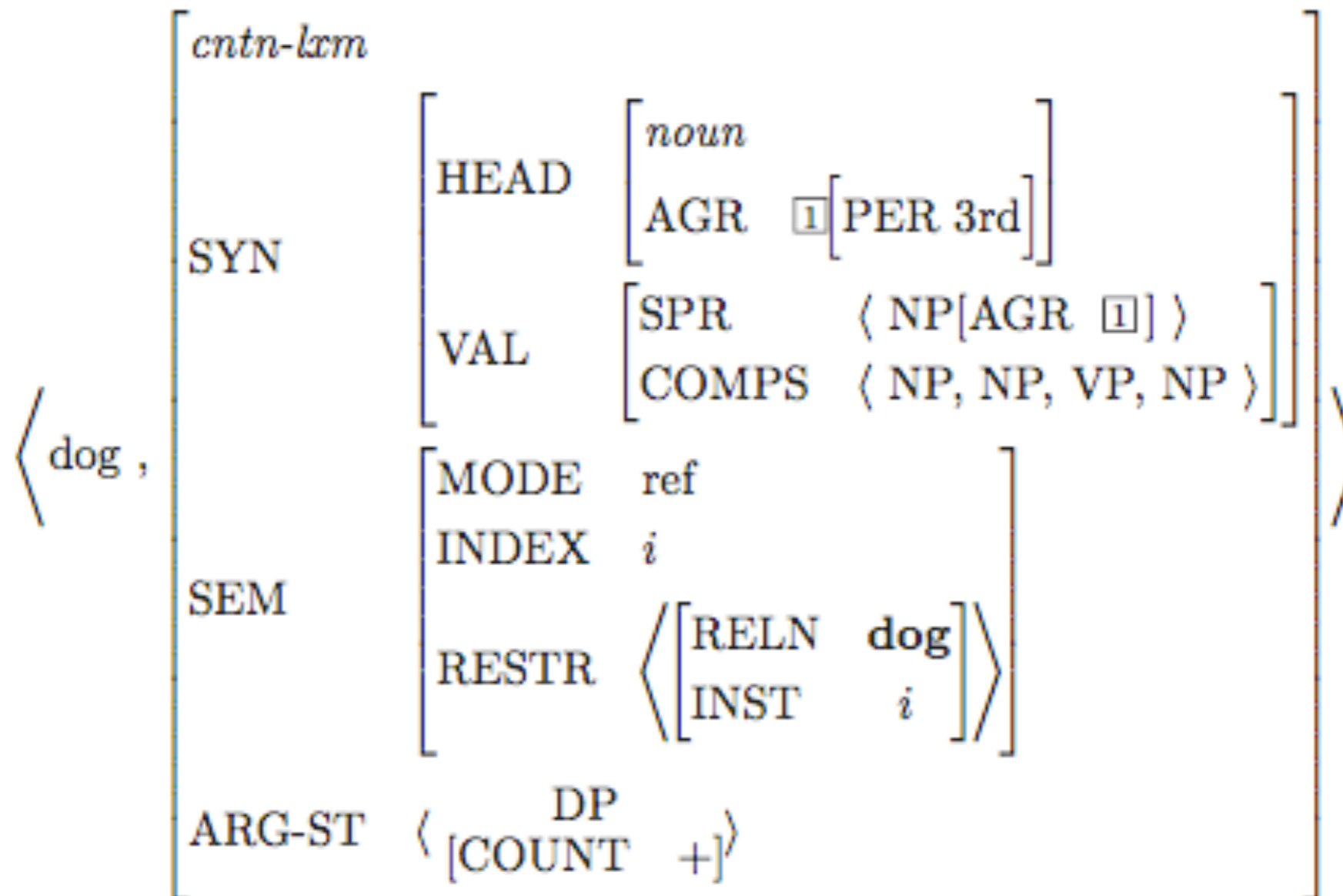
- \*Kim eat pizza.
- \*Kim seems to eats pizza.
- \*Dana helped Leslie [pack and moved].
- \*Kim relies for Sandy.
- \*Dana walked and Kim.

# 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

- Why doesn't this lexical sequence give rise to any words?



# Reading Questions

- Why is the ARP a constraint on type word?
- If the ARP applies to the INPUT of i-rules via identity with the SYN of the OUTPUT, why do the lexical entries for various words not give concrete values for COMPS?



# Reading Questions

- Example 49 on page 247 has  $S \rightarrow NP VP$  and then  $VP \rightarrow V VP$ . Why not eliminate the latter rule, name  $V$  as auxiliary, and make it a feature inside  $VP$ ?
- If *may* and *like* both say [FORM base], how can *Kim may like Sandy be* [FORM fin]?
- What is the head of an  $S$  with an auxiliary in it?

# Reading Questions

- So FORM as defined in the book is a feature in the HEAD matrix (i.e. a feature of pos). But lexemes also have HEAD feature. What FORM values do we give to lexemes?
- Re: (48) on p. 247: What motivates defining pass and psp as distinct values of the feature FORM, while both the "eat" in the directive "Eat rice!" and in "Andy tried to eat rice" have the same base value? Off the top of my head, I can't think of my values where the inflection of a passivized verb would be different from that of a past participle in English.

# Reading Questions

- What is a lexical sequence?
- What are the morphological functions, and how are they supposed to work?
- What do X and Y mean in the definition of l-rule?

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

# Reading Questions

- How do the lexical rules fit in in the course of building trees?
- Why are the lexical rules but not the phrase structure rules organized into a hierarchy?
- Could we recast the PSR as having INPUTs and OUTPUTs too?

# Reading Questions

- For the lexical rule feature structures, there are nodes specified as INPUT, which appears to be a more general set of constraints, and OUTPUT, which is said to be the instantiation of those. Why do we need a system like this? We seemed to be doing fine before, and I had no problem understanding theoretical templates when only given the basic categorical groups (the ones that would correspond to OUTPUT). It just seems redundant to me. Is there a real purpose to this that gives us something we didn't have before?

# Reading Questions

- Do the inflectional rules apply to all lexemes of the appropriate types? How do we account for count nouns that only have plural forms, e.g. jeans, glasses, goods, congratulations, (the) police?

# Reading Questions

- What does the Past Tense Verb LR do for us, if we're not modeling tense?
- Why can't the INDEX of the INPUT and OUTPUT of the Agent Nominalization LR be identified?

# Reading Questions

- With the i-rule constraint stated in (60), the entire SYN block must be the same between the input and the output in inflections, including the AGR in the HEAD. But with the Singular and Plural Noun Lexical Rules, the AGR value of the output has NUM sg and NUM pl respectively, which means that the AGR value of the input must be the same. Doesn't this result in there being distinct lexeme structures for each inflection - say, one for dog, [... [NUM sg] and one for dog, [...[NUM pl]? I thought the point of lexemes was that there would only be one lexeme that covers (or could be converted into) all the possible inflections.