



Ling 566

Oct 31, 2023

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*, ...

When poll is active, respond at pollev.com/emb

Text **EMB** to **22333** once to join



Is it clear what type of regularities are

W captured by lexical types and lexical rules?
(take 3)

Not clear why we need either

Not clear what the difference is

Yes ...?

Yes

Total Results: 0

Powered by  **Poll Everywhere**

Start the presentation to see live content. For screen share software, share the entire screen. Get help at pollev.com/app

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*

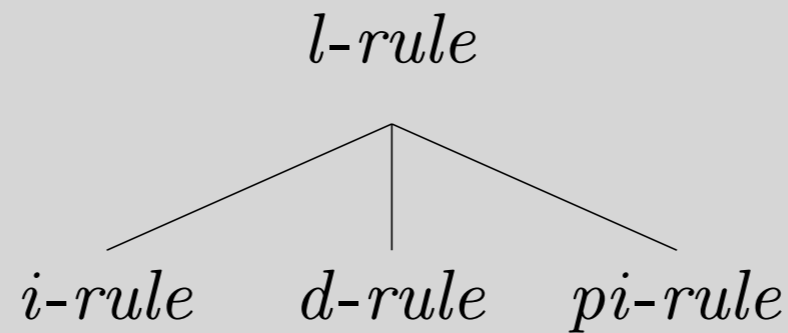
Examples?

- Derivational: *lexeme to lexeme*

Examples?

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

Three Subtypes of *l-rule*

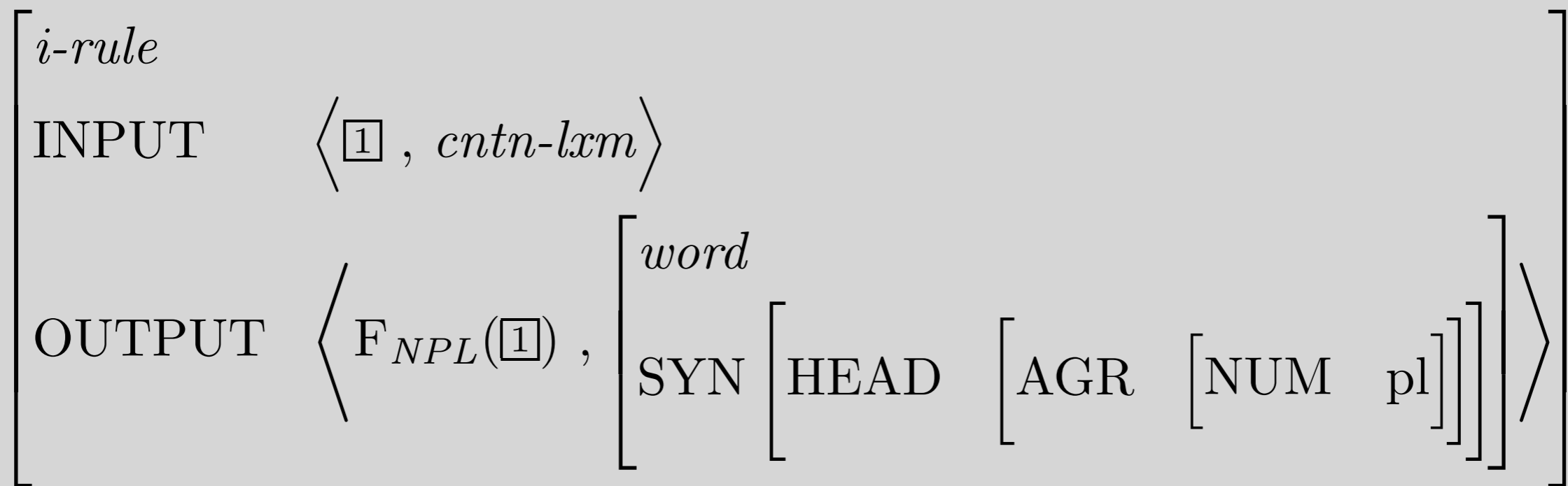


$$l\text{-rule} : \left[\begin{array}{l} \text{INPUT} \quad l\text{-sequence} \left\langle X, [\text{SEM} \quad / \boxed{2}] \right\rangle \\ \text{OUTPUT} \quad l\text{-sequence} \left\langle Y, [\text{SEM} \quad / \boxed{2}] \right\rangle \end{array} \right]$$

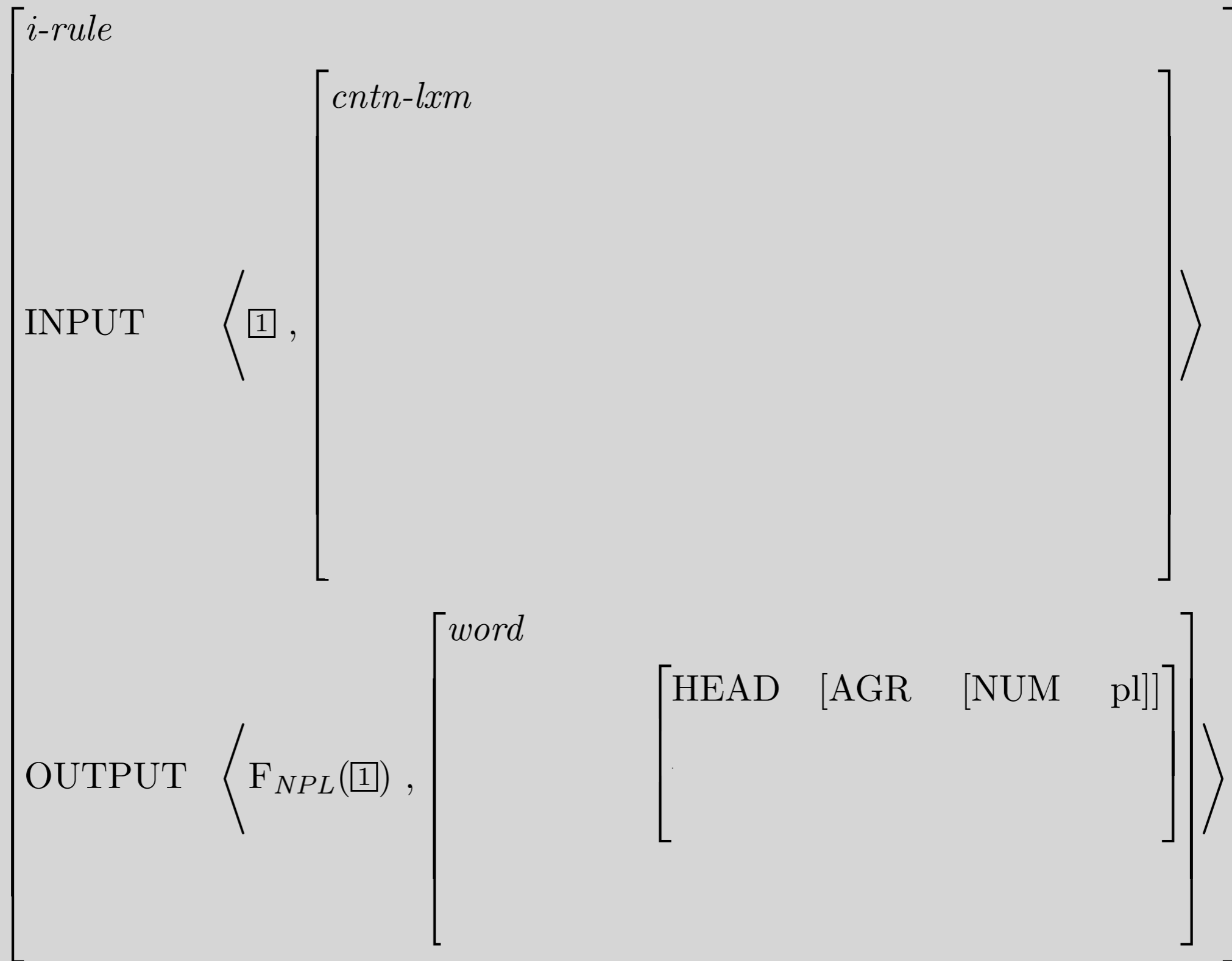
$$i\text{-rule} : \left[\begin{array}{l} \text{INPUT} \quad \left\langle X, \left[\begin{array}{l} \textit{lexeme} \\ \text{SYN} \quad \boxed{3} \\ \text{ARG-ST} \quad \boxed{A} \end{array} \right] \right\rangle \\ \text{OUTPUT} \quad \left\langle Y, \left[\begin{array}{l} \textit{word} \\ \text{SYN}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

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

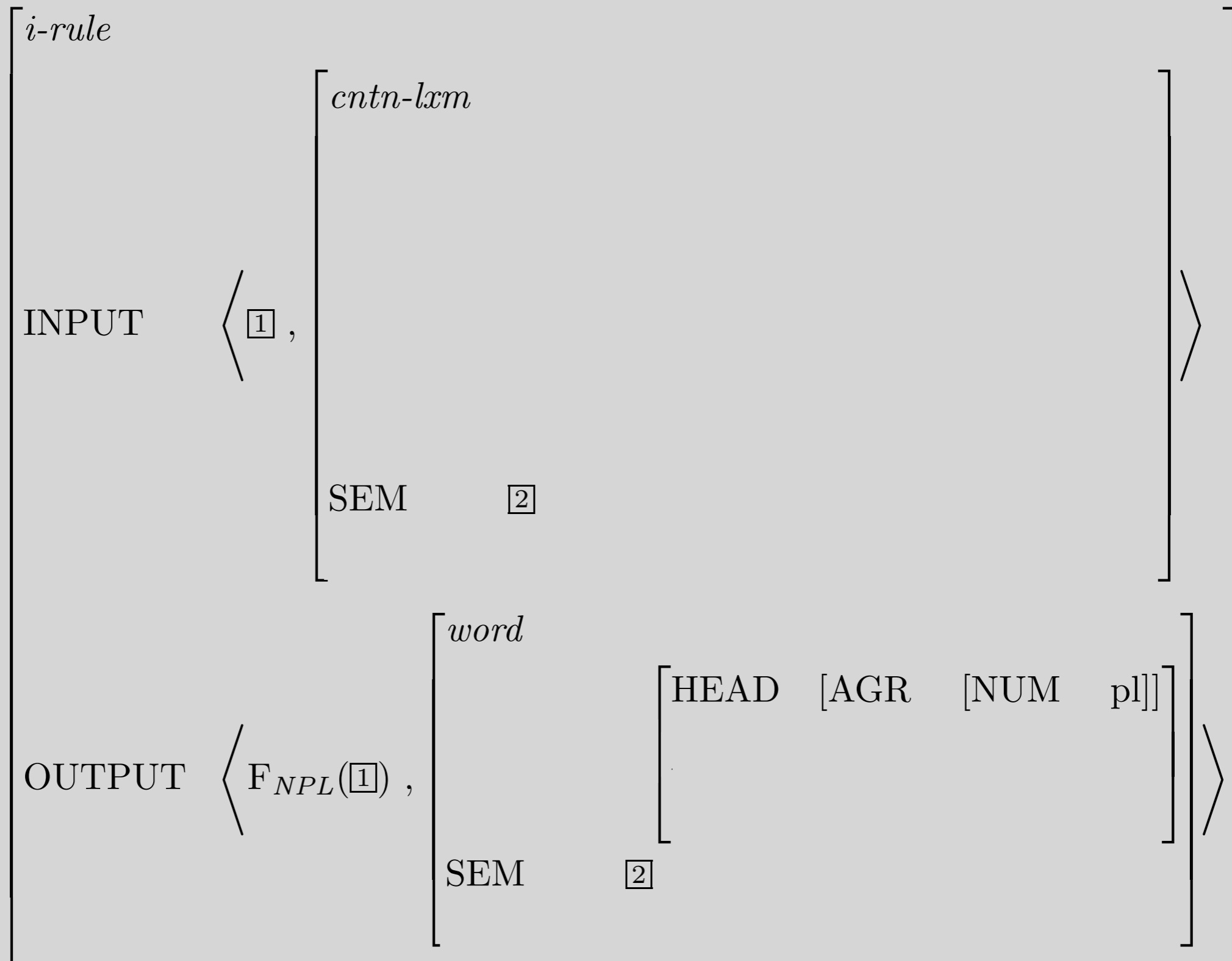
Plural Noun LR



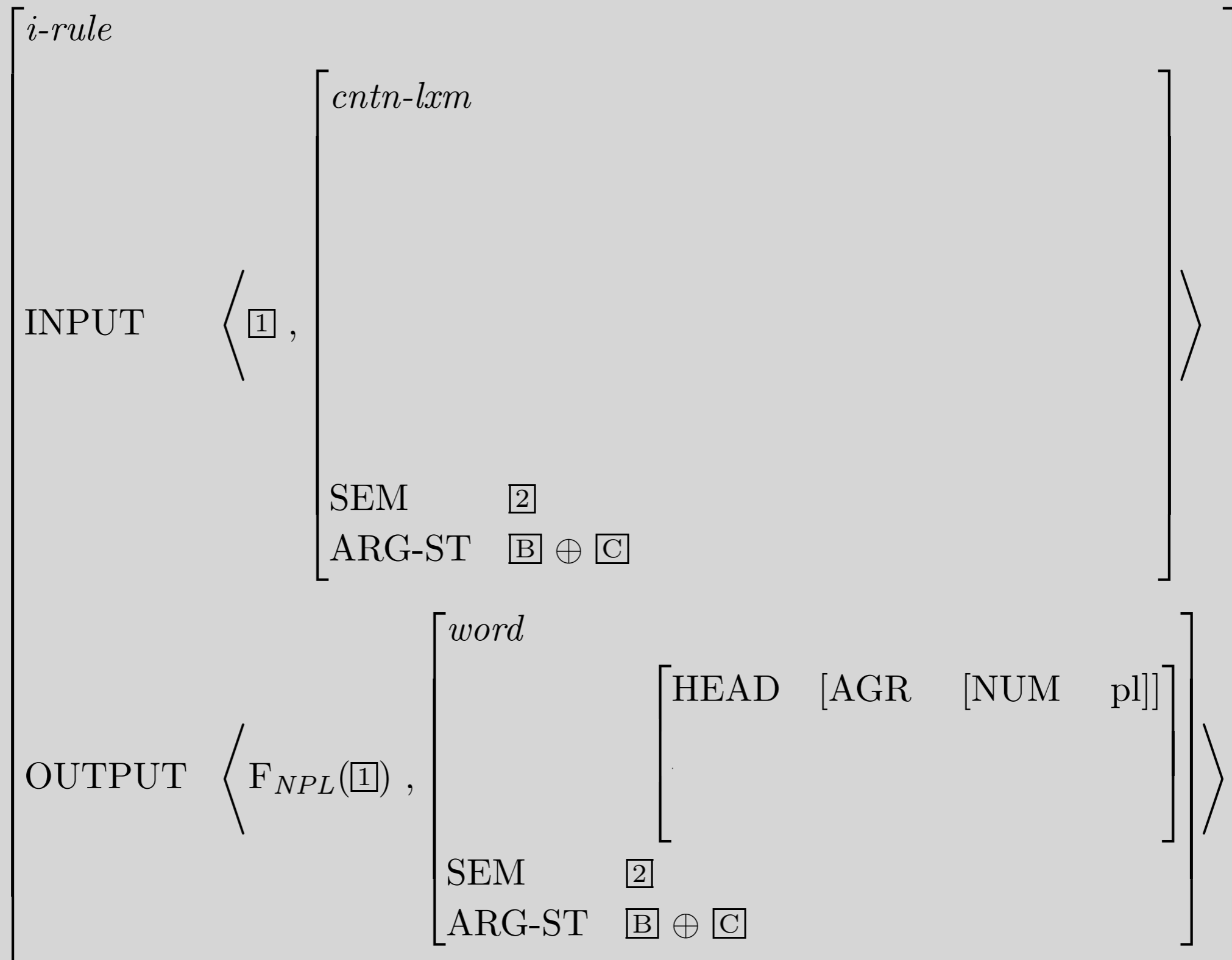
Plural Noun LR with Inherited Constraints



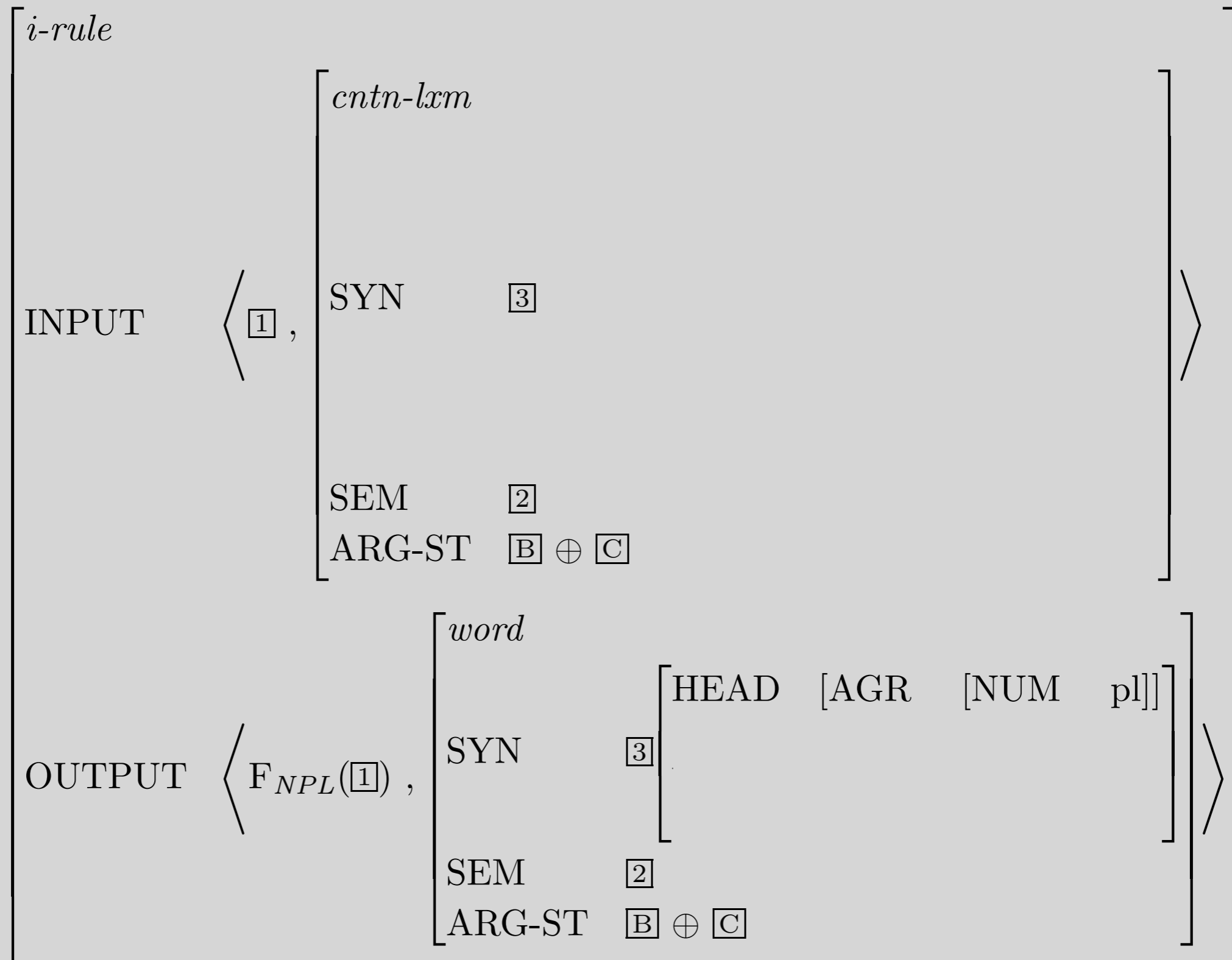
Plural Noun LR with Inherited Constraints



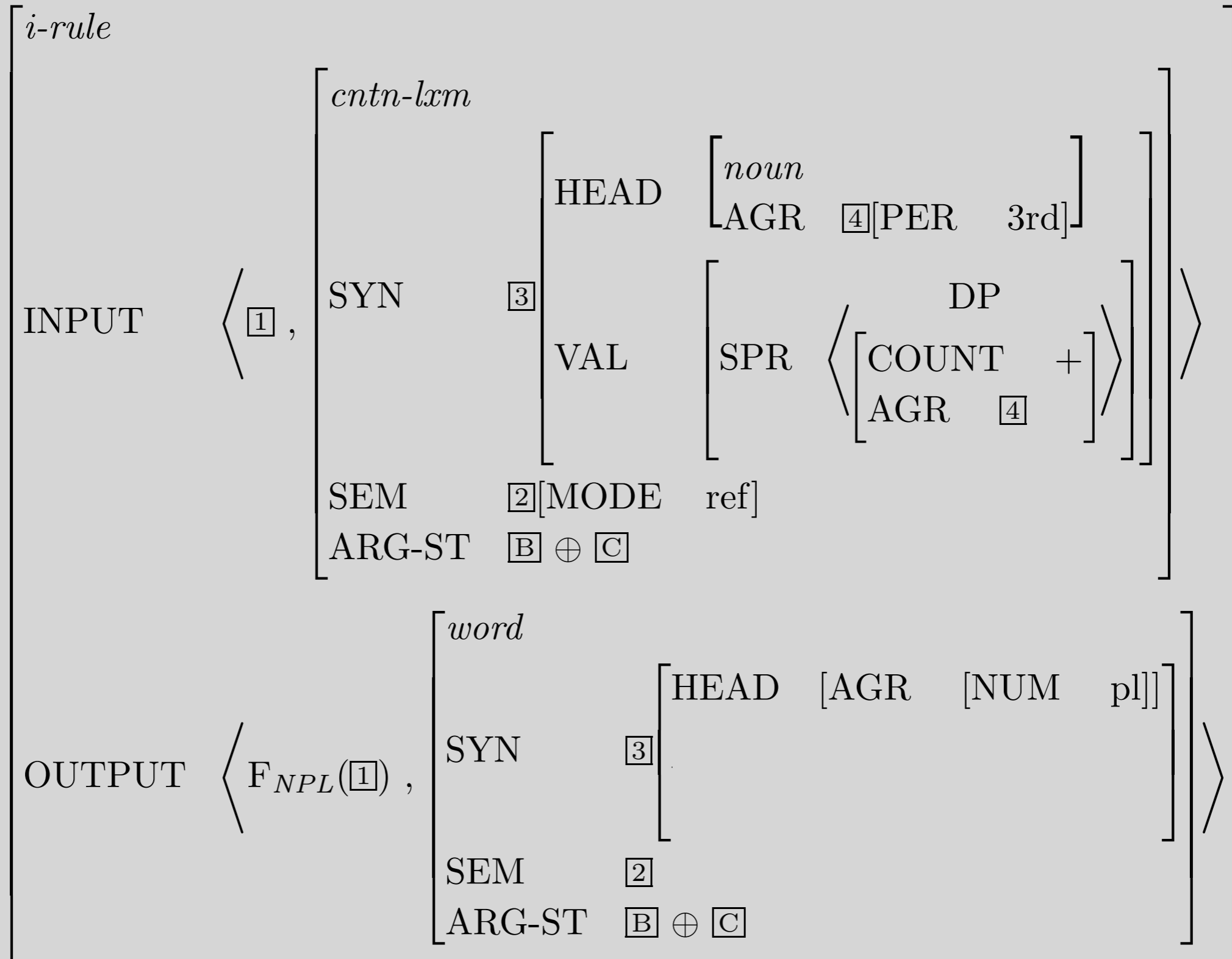
Plural Noun LR with Inherited Constraints



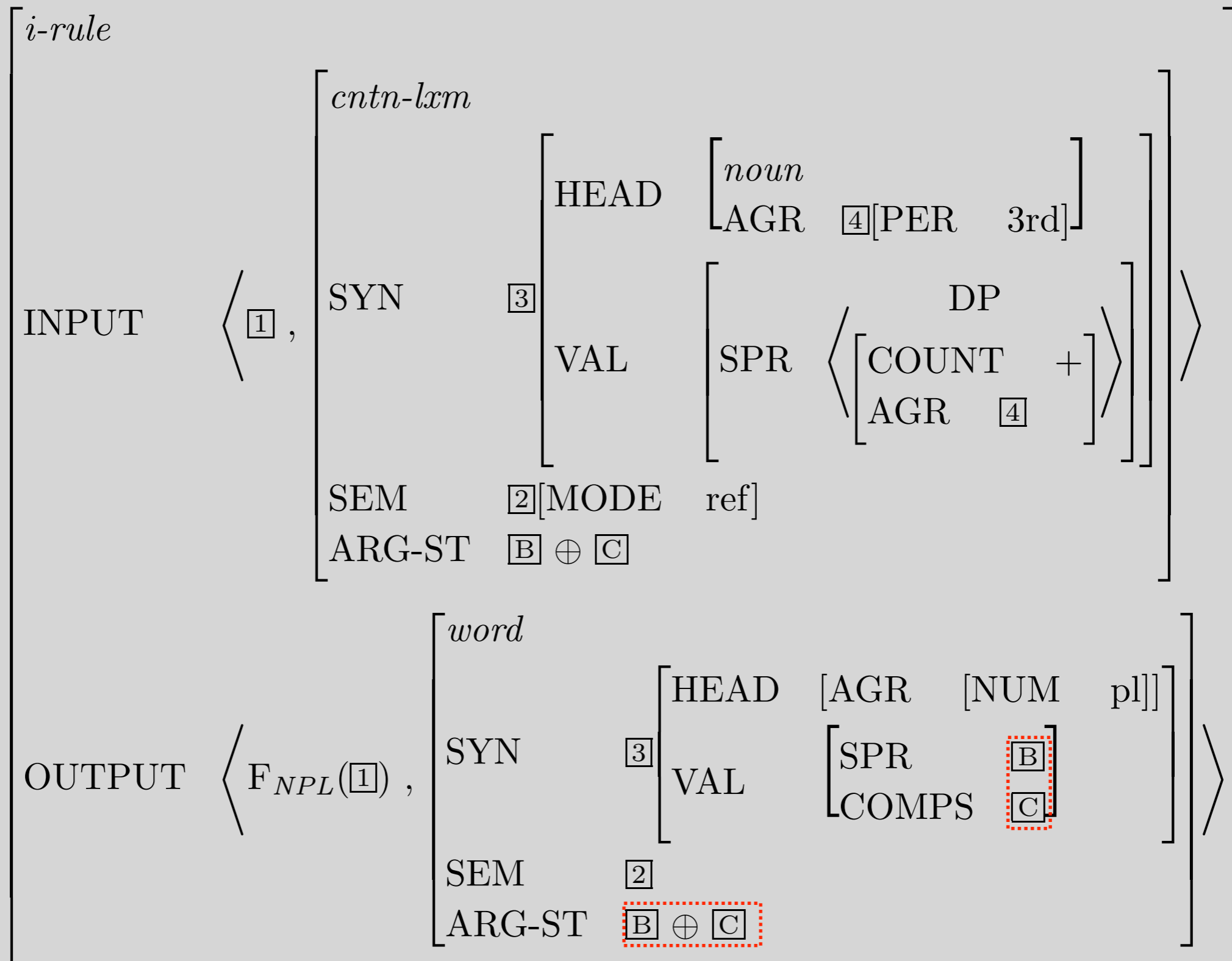
Plural Noun LR with Inherited Constraints



Plural Noun LR with Inherited Constraints



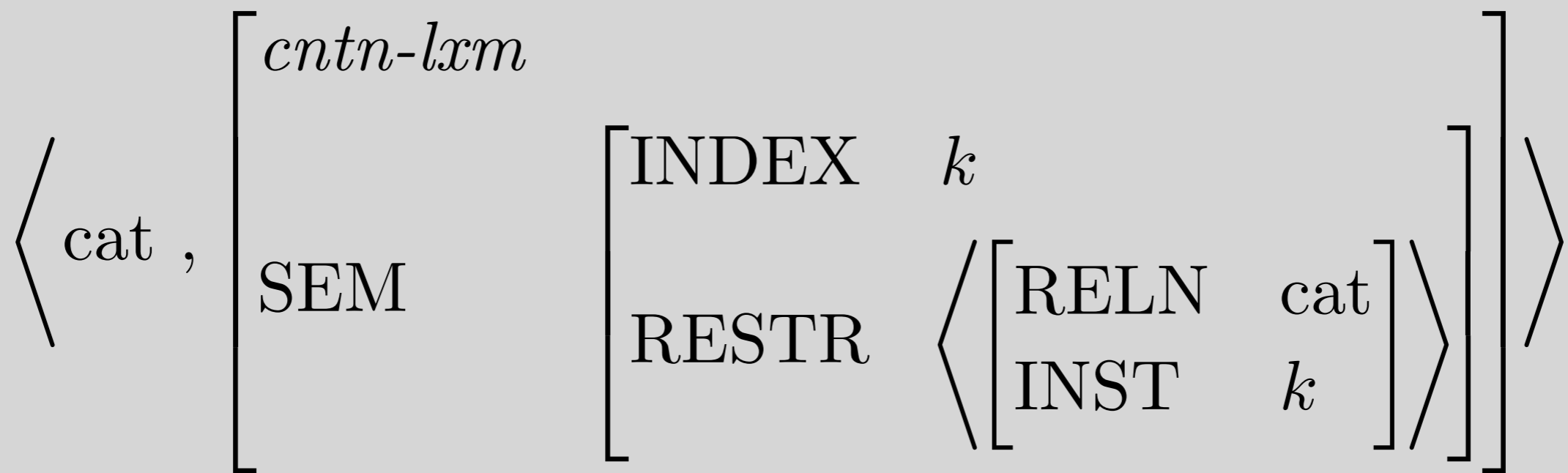
Plural Noun LR with Inherited Constraints



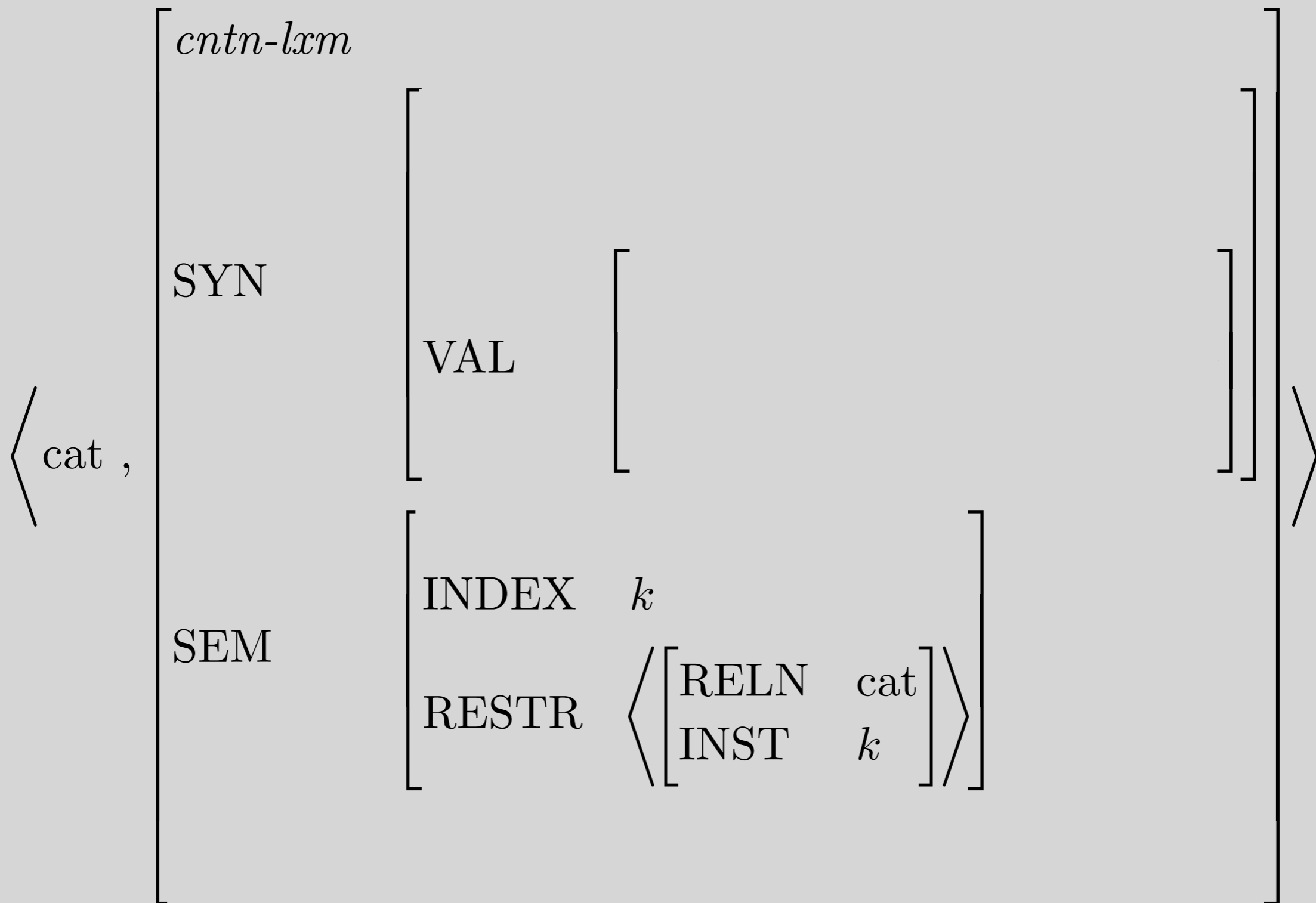
Practicalities - Applying Lexical Rules

- 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.

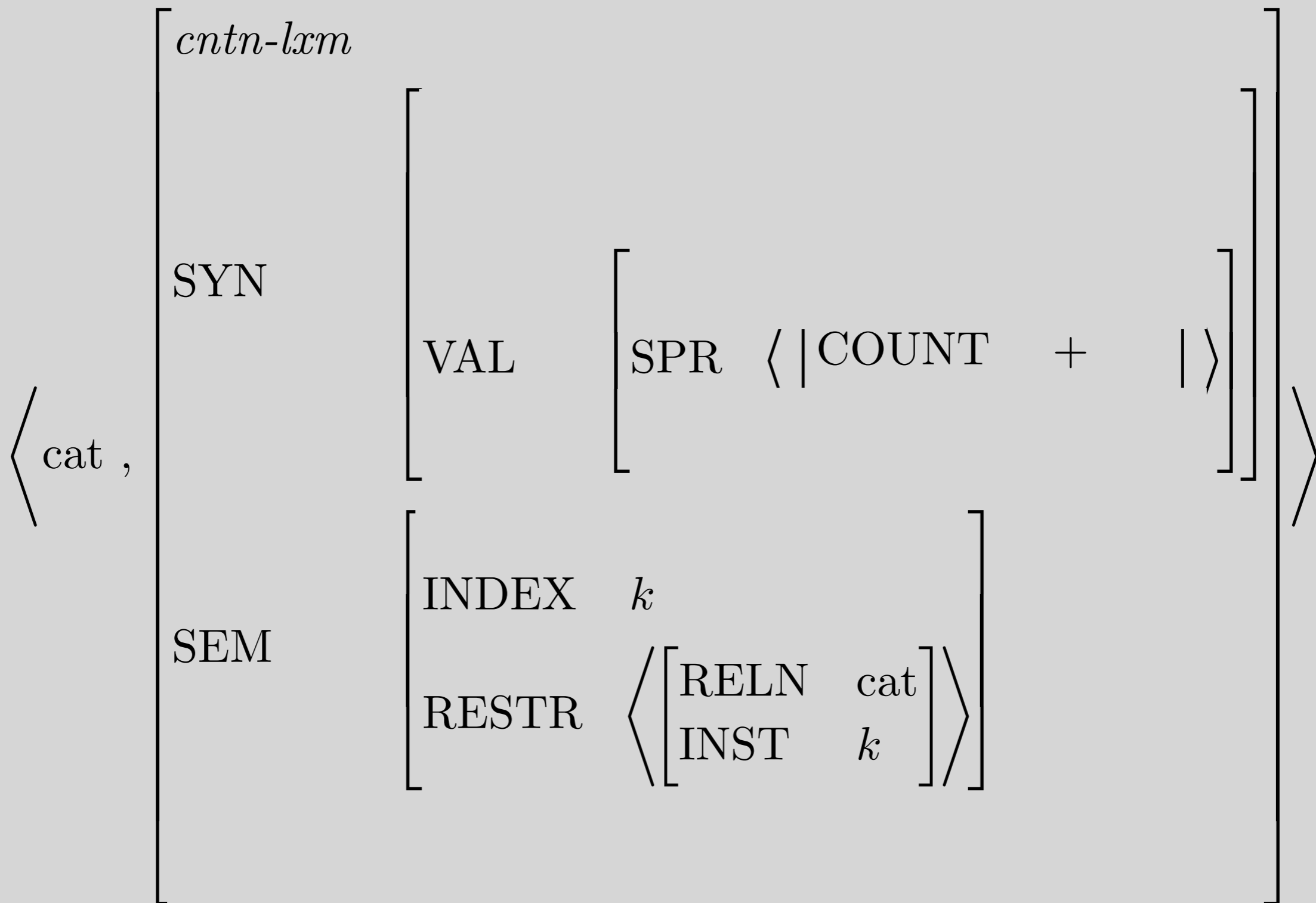
Example: Lexical Entry for *cat*



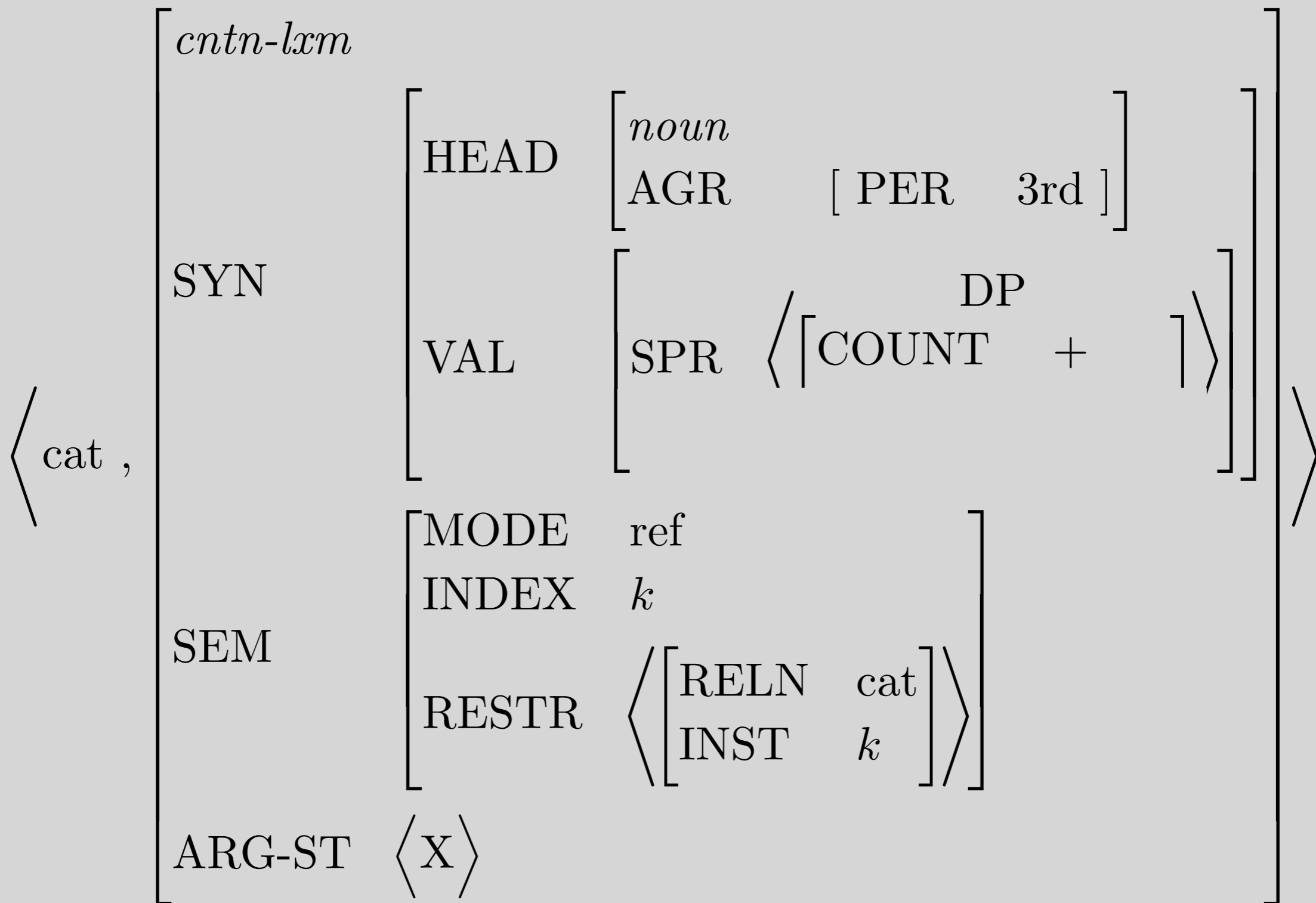
Example: *cat*, with inheritance



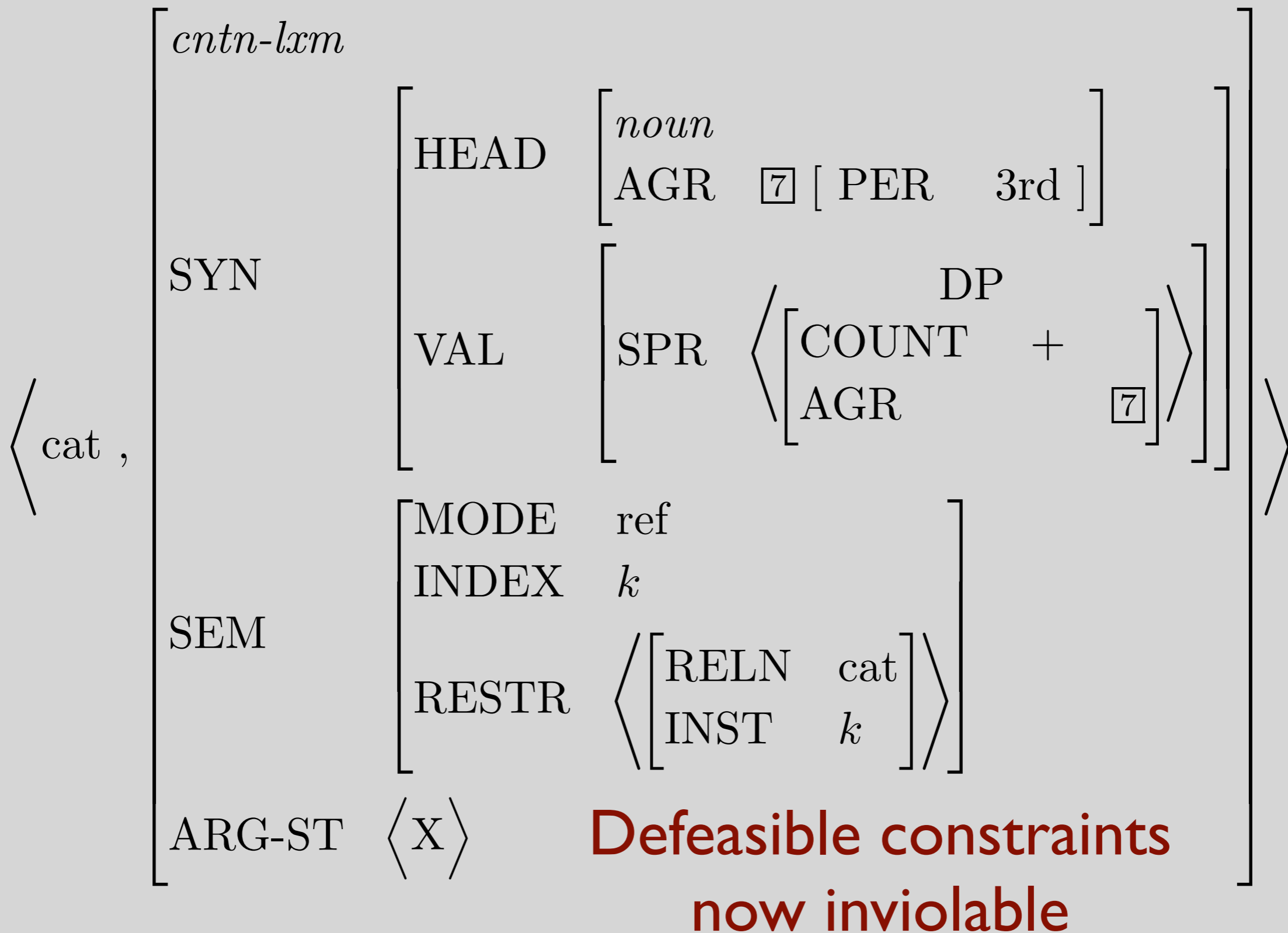
Example: *cat*, with inheritance



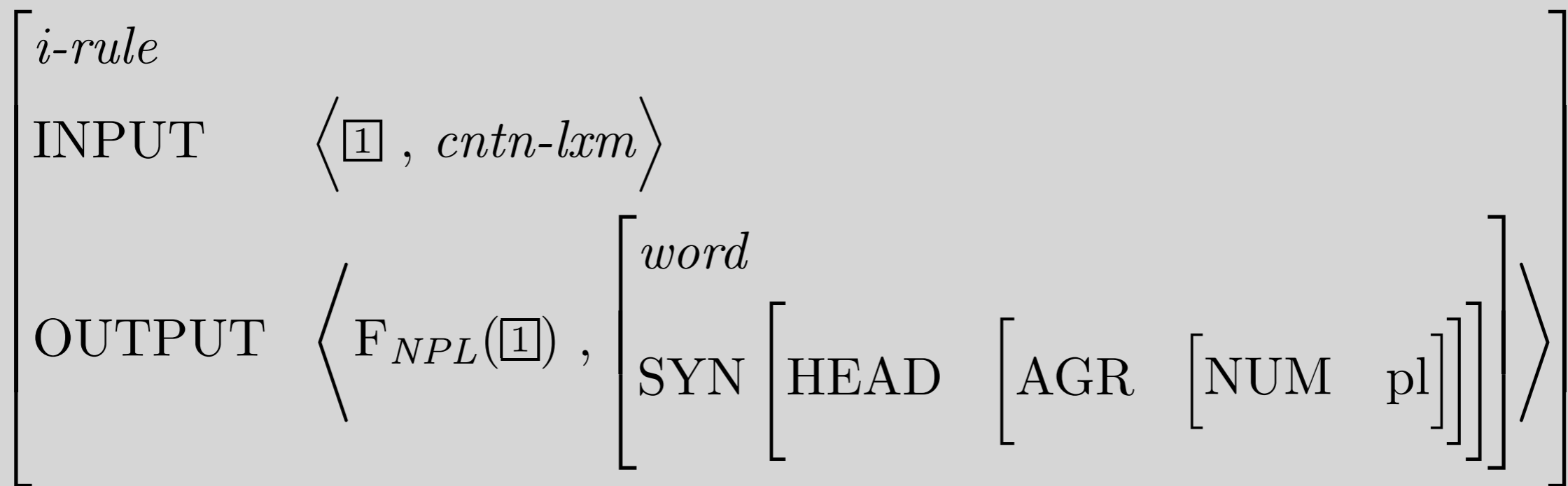
Example: *cat*, with inheritance



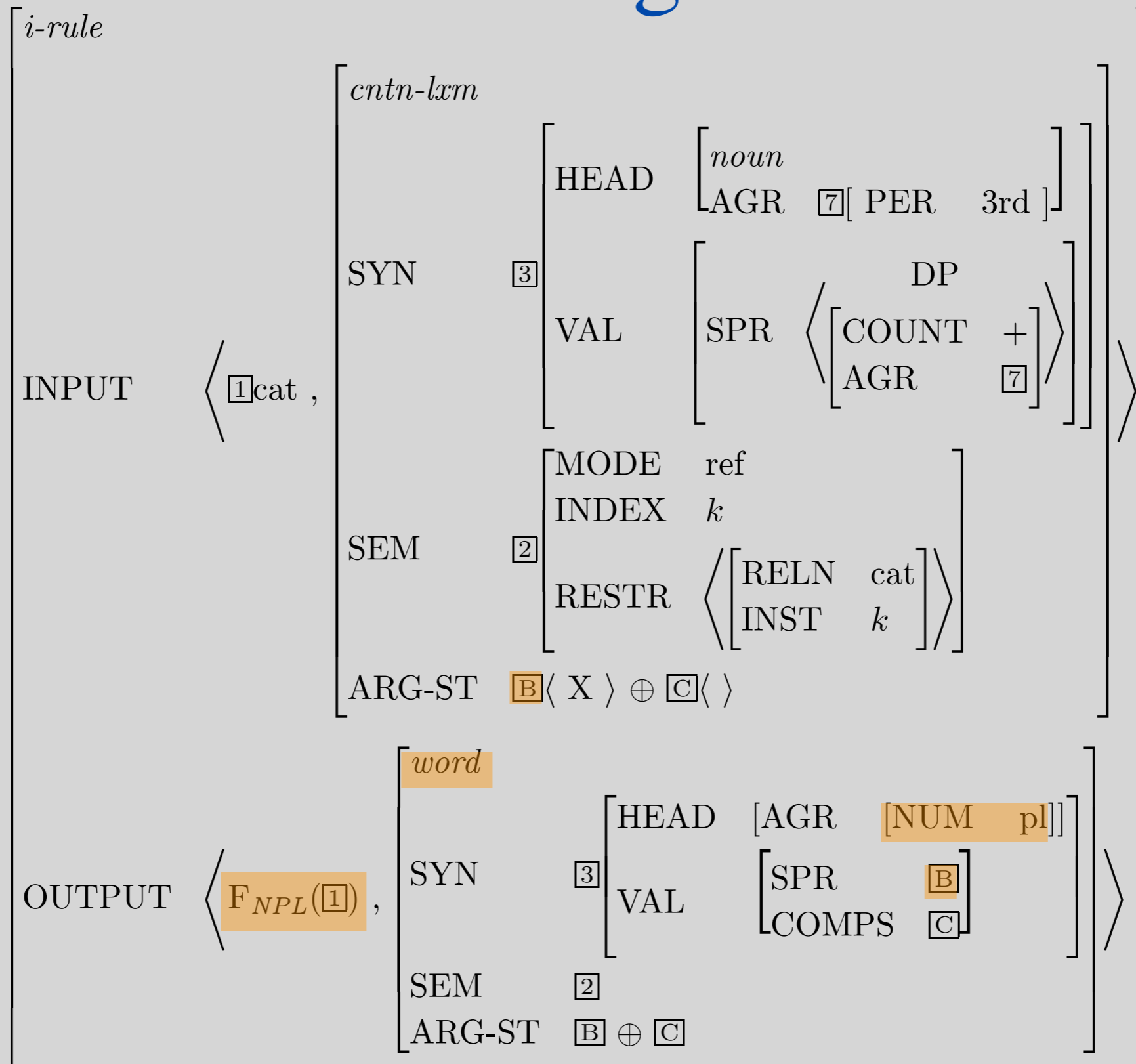
Example: *cat*, with inheritance



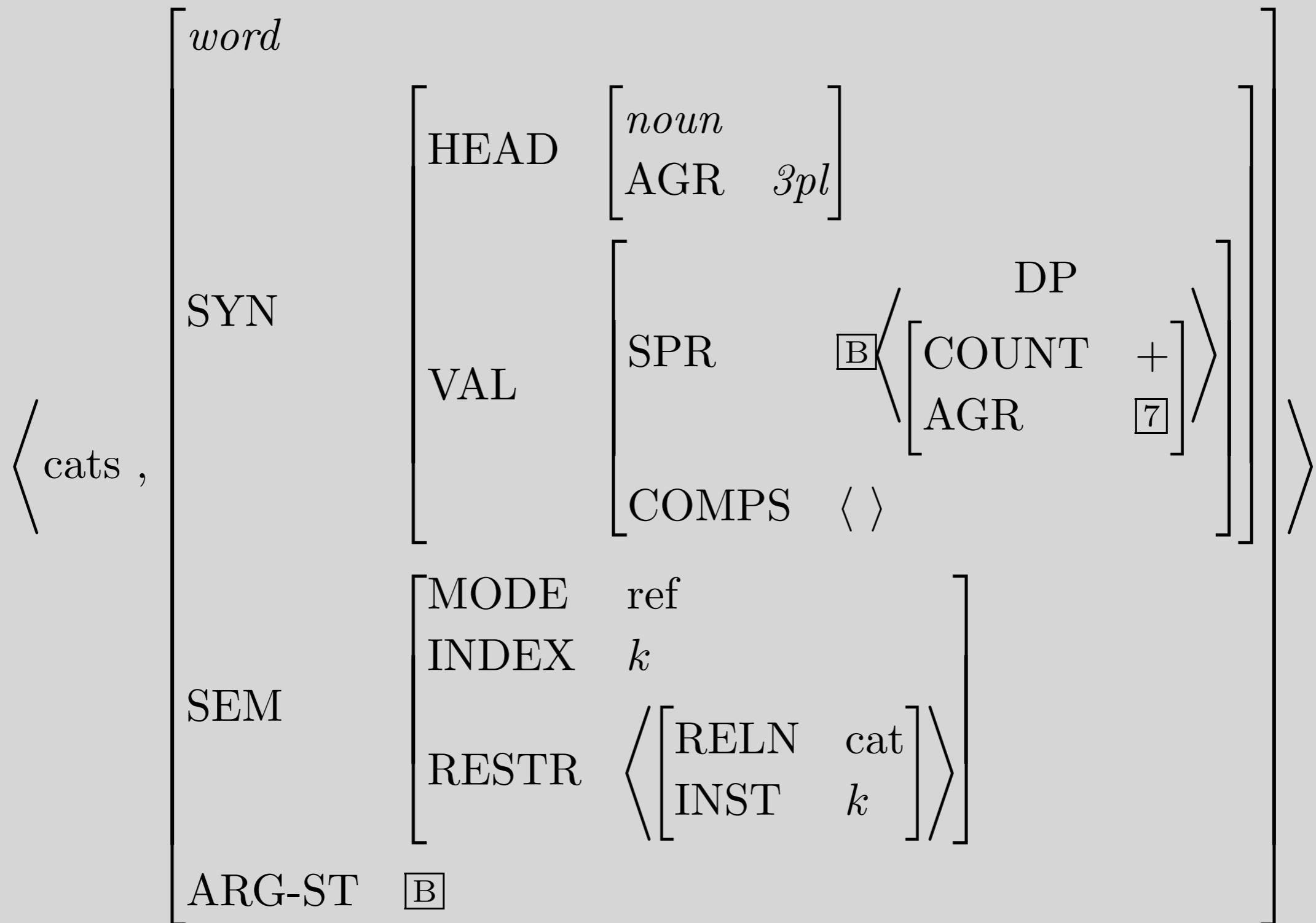
Plural Noun LR



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	$\langle \boxed{1}, \text{const-lxm} \rangle$
OUTPUT	$\left[\text{FIRST } \boxed{1} \right]$

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

Constant Lexeme LR

<i>i-rule</i>	
INPUT	$\langle \boxed{1}, \text{const-lexm} \rangle$
OUTPUT	$\langle \boxed{1}, [] \rangle$

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

When poll is active, respond at pollev.com/emb

Text **EMB** to **22333** once to join



Is it clear what type of regularities are

W captured by lexical types and lexical rules?
(take 4)

Not clear why we need either

Not clear what the difference is

Yes ...?

Yes

Total Results: 0

Powered by  **Poll Everywhere**

Start the presentation to see live content. For screen share software, share the entire screen. Get help at pollev.com/app

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

RQs: Morphophonology

- Are morphological functions (such as Fnpl) considered part of our grammar? And, how are these morphological functions implemented in our computer programs?
- Regarding common noun inflection, I'm wondering how these rules apply to irregular nouns in English. Are there any consistent patterns or rules for nouns like 'child' and 'children', or 'man' and 'men', or are they simply exceptions that we have to memorize

RQs: FORM

- Why isn't future tense included on the FORM list in (48)? How is the future tense analyzed?
- Why are both past and present verbs FORM fin if they still get separate rules? Will the distinction between past and present tense verbs preserved somewhere in a word's feature structure?

RQs: FORM

- Does this treatment of PPs (like PP[FORM of]) have anything to do with the predication vs. arg-marking element of PP's that we've seen so far?
- Do we specify FORM in lexical entries?
- Is FORM a meaningful feature for parts of speech that only have one FORM, or is it just a convenience for referencing part of speech during coordination? Are there multiple FORMS for certain parts of speech in other languages that we just happen to lack in English, or do some coordinatable parts of speech not take multiple forms in any language?

RQs: FORM

- I am curious about how the FORM feature works in languages with a more complex conjugation system, such as Spanish, and how it works together with the AGR to create specific conjugations. Would Spanish have 14 forms, the same number as its tenses?
- Does the distinction between tense, aspect and mood have any specific role, or do all these come together to create the different forms?

RQs: ARG-ST/ARP

- From the first paragraph of page 260, we know that the SYN and ARG-ST values of the INPUT and the OUTPUT are identified, which means that the INPUT will always, as a side-effect, also obey the ARP. However, on the other hand, our ARP only applies "officially" to the word type. So why don't we explicitly extend the ARP to the lexeme type given this side-effect?

RQs: *i-rules*

- How come the HEAD values of the INPUT and OUTPUT are completely identified in (67)? Wouldn't this assign [NUM pl] to the INPUT as well? I'm confused because I was expecting the INPUT to have an underspecified or singular NUM value.

RQs: *i-rules*

- The footnote 28 on page 253 says that "if an input were specified as [NUM pl], then it would fail to undergo the lexical rule." Is it because that a lexeme should have its AGR value unspecified? I don't quite understand the following sentence though: "there could be no relation between the input lexical sequence and any output lexical sequence that satisfied the constraint specified in (62)". However, if an input were specified as [NUM pl], why are we using the Singular Noun Lexical Rule (since we have already known it is plural)?

RQs: *i-rules*

- When apply the *i-rule*, the SYN should be identical between the input and output, does this mean the lexeme should have same FORM value with the word? so for the same word, they come from different lexeme?
- Why does the output of the Past-tense lexical rule specify CASE on it the first argument on the ARG-ST? I am wondering because I thought that inflectional rules do not effect changes in ARG-ST.

RQs: *d-rules*

- Is there any limitation on zero derivation instances? How does one determine whether a verbal lexeme converted from noun lexeme is acceptable?

RQs: *l-rules* + trees

- I'm curious about what the relation between the lexical/derivational rules and our previously examined grammar rules/principles is. Are lexical/derivational rules just a way to conveniently convey common features between lexemes or can they come into play within our trees to sanction specific inflections? For example, is there a way to show in our trees that a past tense inflection rule is required to sanction a sentence which would otherwise just express the need for a FORM "fin" verb (thus accepting both present or past tense)?

RQs: Hierarchy of rules

- At least part of the motivation for introducing lexical rules seems to be because they lend themselves to organization within a type hierarchy. Is there an inherent advantage to our grammar containing these hierarchies and structures besides just for organizational purposes and ease of use? Is it to mirror the trees themselves? Are these reasons enough to motivate the introduction of this entirely new type of rule?