



Ling 566

Oct 4, 2022

Context-Free Grammar

Self-intro

- Please call me Emily
 - But Dr./Prof. Bender is okay
- Pronouns she/her

Overview

- Two insufficient theories
- Formal definition of CFG
- Constituency, ambiguity, constituency tests
- Central claims of CFG
- Weaknesses of CFG
- Reading questions

Insufficient Theory #1

- A grammar is simply a list of sentences.
- What's wrong with this?

An example hypothetical language (p.22)

- Some sentences go on and on
- *Sentences some go on and on
- *Some sentences go on and on and on
- Sentences some go on and on and on
- Some sentences go on and on and on and on
- *Sentences some go on and on and on and on
on

Insufficient Theory #2: FSMs

- the noisy dogs left

D A N V

- the noisy dogs chased the innocent cats

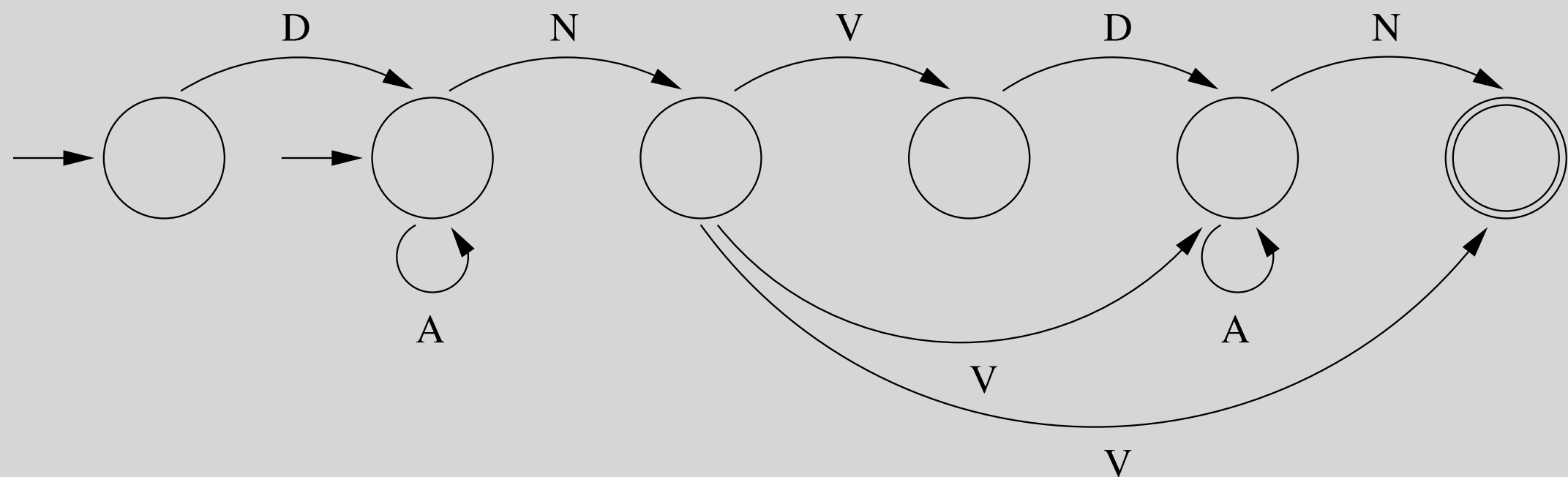
D A N V D A N

- $a^* = \{\emptyset, a, aa, aaa, aaaa, \dots\}$

- $a^+ = \{a, aa, aaa, aaaa, \dots\}$

- $(D) A^* N V ((D) A^* N)$

A Finite State Machine



Insufficient Theory #2: FSMs

- What's wrong with this?
- What can't it model?

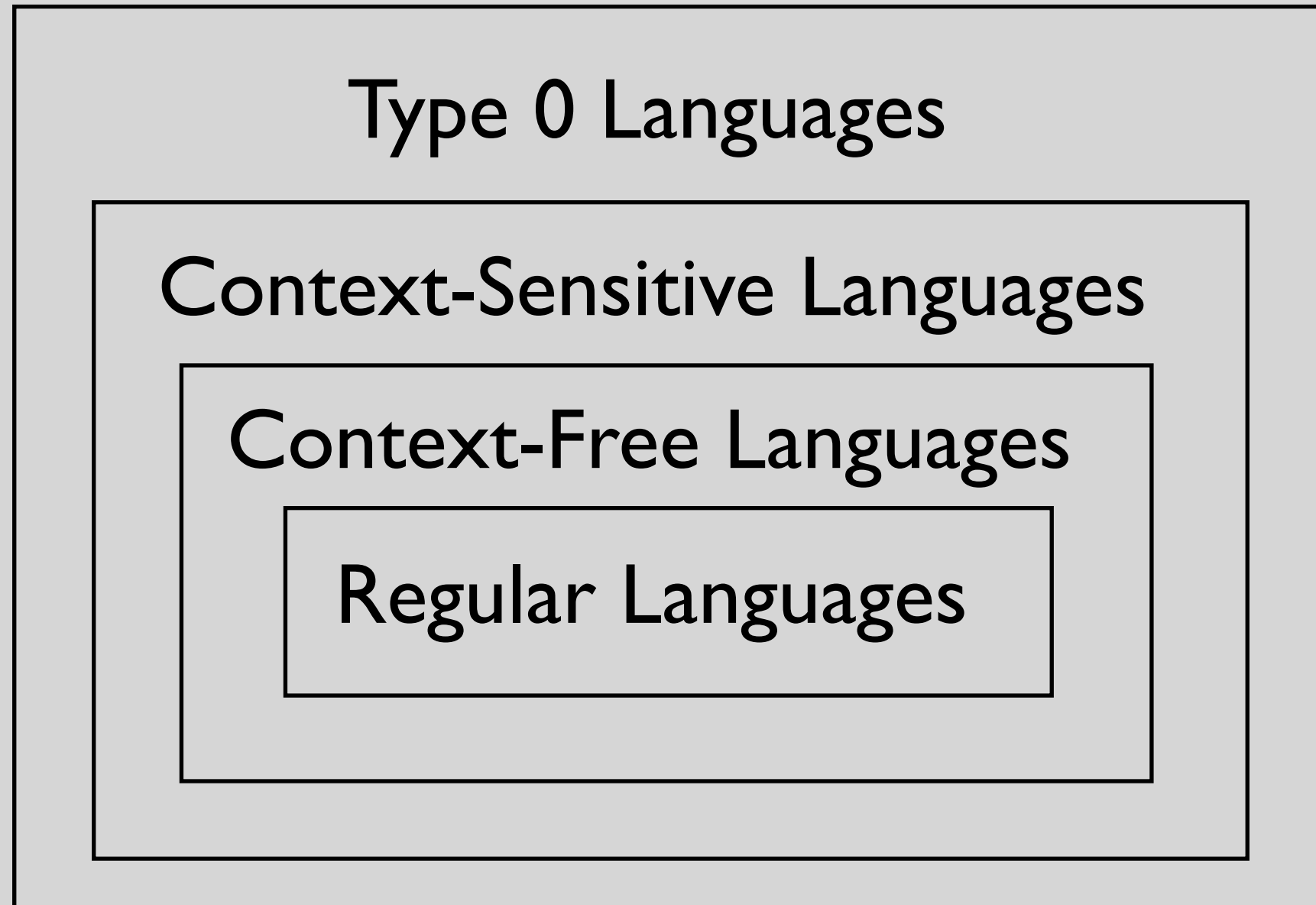
What does a theory do?

- Monolingual
 - Model grammaticality/acceptability
 - Model relationships between sentences (internal structure)
- Multilingual
 - Model relationships between languages
 - Capture generalizations about possible languages

Summary

- Grammars as lists of sentences:
 - Runs afoul of creativity of language
- Grammars as finite-state machines:
 - No representation of structural ambiguity
 - Misses generalizations about structure
 - (Not formally powerful enough)
- Next attempt: Context-free grammar

Chomsky Hierarchy



Context-Free Grammar

- A quadruple: $\langle C, \Sigma, P, S \rangle$
 - C : set of categories
 - Σ : set of terminals (vocabulary)
 - P : set of rewrite rules $\alpha \rightarrow \beta_1, \beta_2, \dots, \beta_n$
 - S in C : start symbol
 - For each rule $\alpha \rightarrow \beta_1, \beta_2, \dots, \beta_n \in P$
 $\alpha \in C$; $\beta_i \in C \cup \Sigma$; $1 \leq i \leq n$

A Toy Grammar

RULES

$S \longrightarrow NP VP$

$NP \longrightarrow (D) A^* N PP^*$

$VP \longrightarrow V (NP) (PP)$

$PP \longrightarrow P NP$

LEXICON

D: the, some

A: big, brown, old

N: birds, fleas, dog, hunter, I

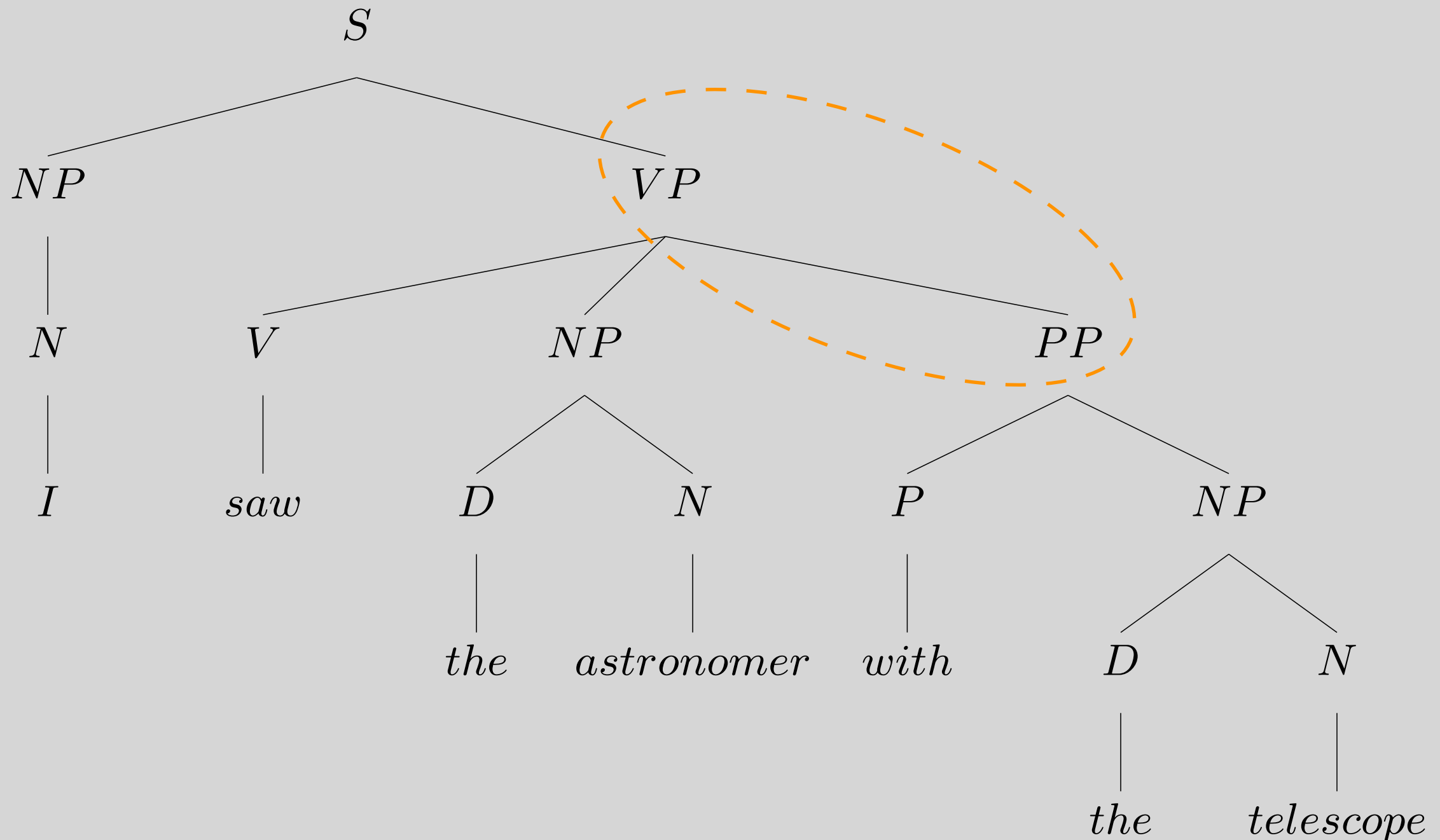
V: attack, ate, watched

P: for, beside, with

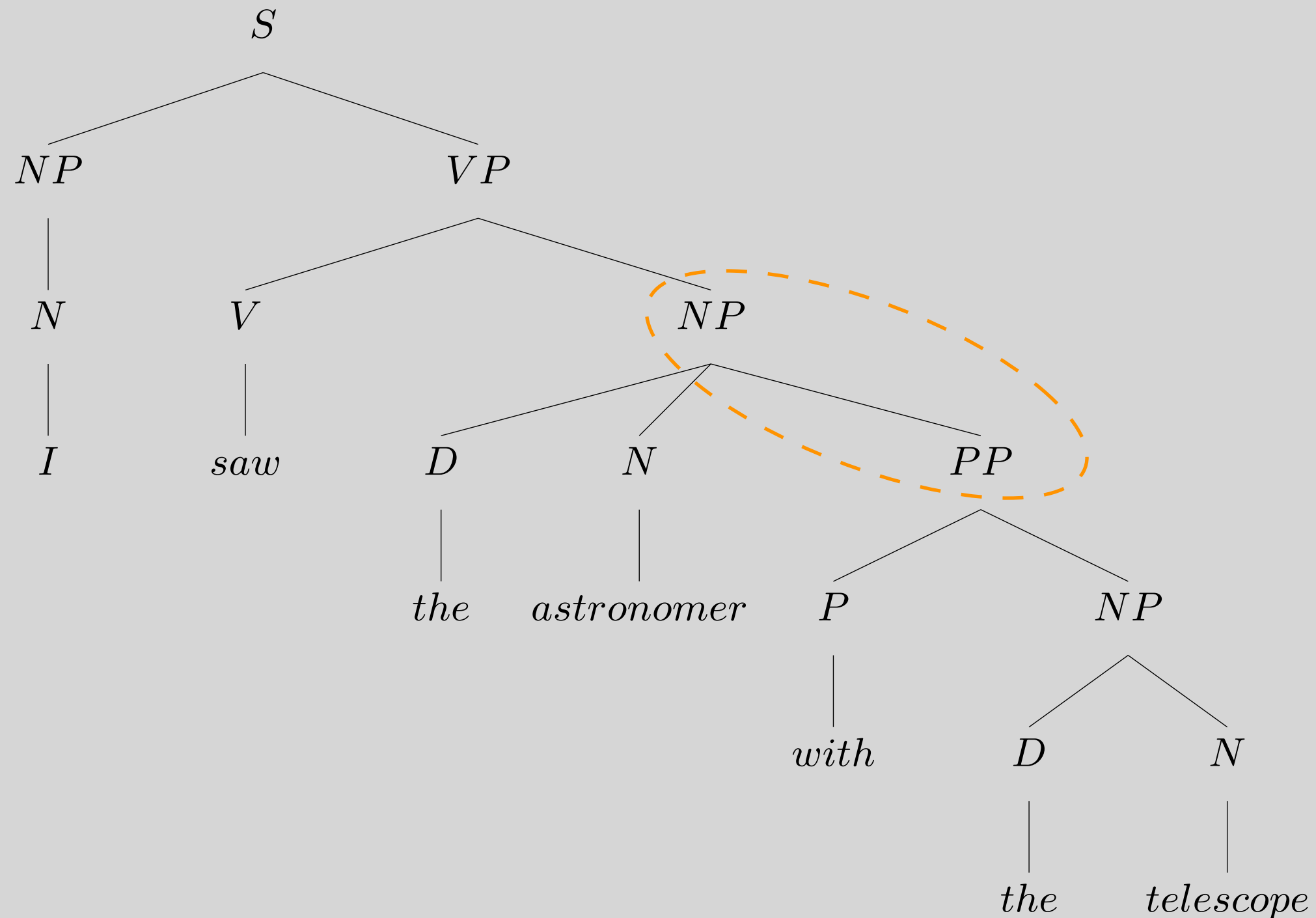
Structural Ambiguity

I saw the astronomer with the telescope.

Structure 1: PP under VP



Structure 1: PP under NP



Constituents

- How do constituents help us? (What's the point?)
- What aspect of the grammar determines which words will be modeled as a constituent?
- How do we tell which words to group together into a constituent?
- What does the model claim or predict by grouping words together into a constituent?

Constituency Tests

- Recurrent Patterns

The quick brown fox with the bushy tail jumped over the lazy brown dog with one ear.

- Coordination

The quick brown fox with the bushy tail and the lazy brown dog with one ear are friends.

- Sentence-initial position

The election of 2000, everyone will remember for a long time.

- Cleft sentences

It was a book about syntax they were reading.

General Types of Constituency Tests

- Distributional
- Intonational
- Semantic
- Psycholinguistic

... but they don't always agree.

Central claims implicit in CFG formalism:

1. Parts of sentences (larger than single words) are linguistically significant units, i.e. phrases play a role in determining meaning, pronunciation, and/or the acceptability of sentences.
2. Phrases are contiguous portions of a sentence (no discontinuous constituents).
3. Two phrases are either disjoint or one fully contains the other (no partially overlapping constituents).
4. What a phrase can consist of depends only on what kind of a phrase it is (that is, the label on its top node), not on what appears around it.

- Claims 1-3 characterize what is called ‘phrase structure grammar’
- Claim 4 (that the internal structure of a phrase depends only on what type of phrase it is, not on where it appears) is what makes it ‘context-free’.
- There is another kind of phrase structure grammar called ‘context-sensitive grammar’ (CSG) that gives up 4. That is, it allows the applicability of a grammar rule to depend on what is in the neighboring environment. So rules can have the form $A \rightarrow X$, in the context of Y_Z .

Possible Counterexamples

- To Claim 2 (no discontinuous constituents):

A technician arrived who could solve the problem.

- To Claim 3 (no overlapping constituents):

*I read *what* was written about me.*

- To Claim 4 (context independence):

- *He arrives this morning.*
- **He arrive this morning.*
- **They arrives this morning.*
- *They arrive this morning.*

A Trivial CFG

$S \rightarrow NP VP$

$NP \rightarrow D N$

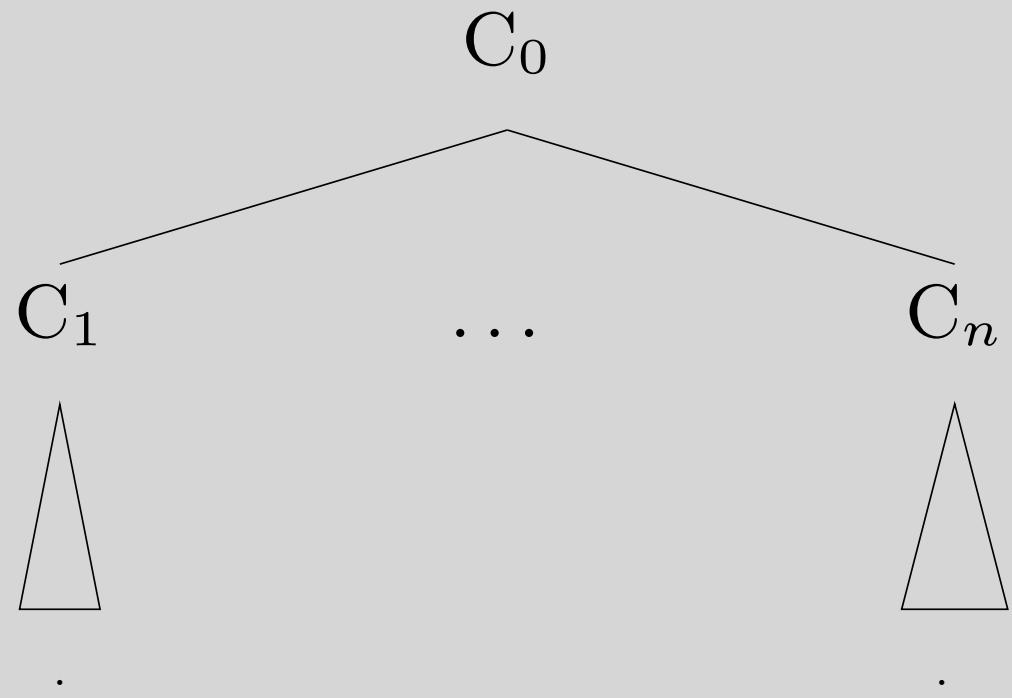
$VP \rightarrow V NP$

D: *the*

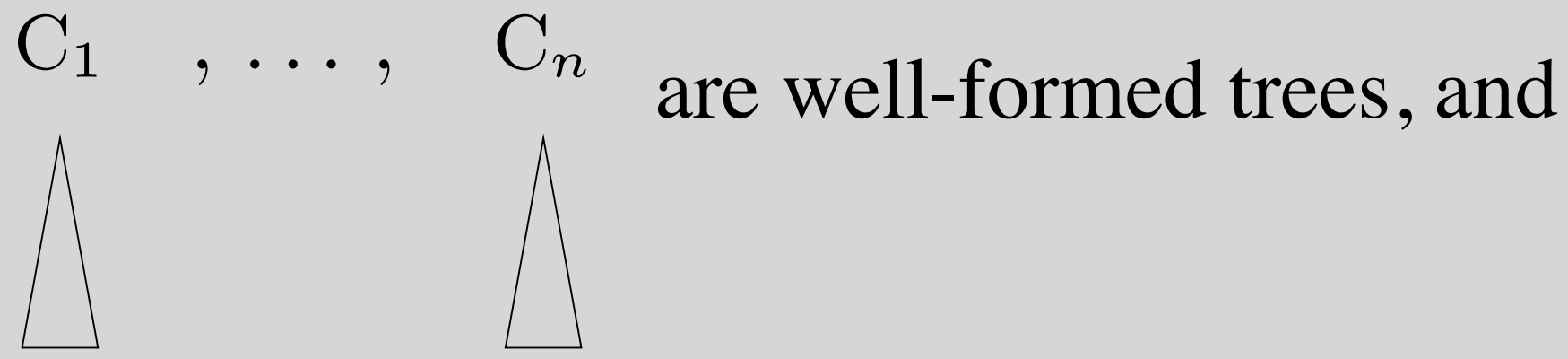
V: *chased*

N: *dog, cat*

Trees and Rules



is a well-formed nonlexical tree if (and only if)



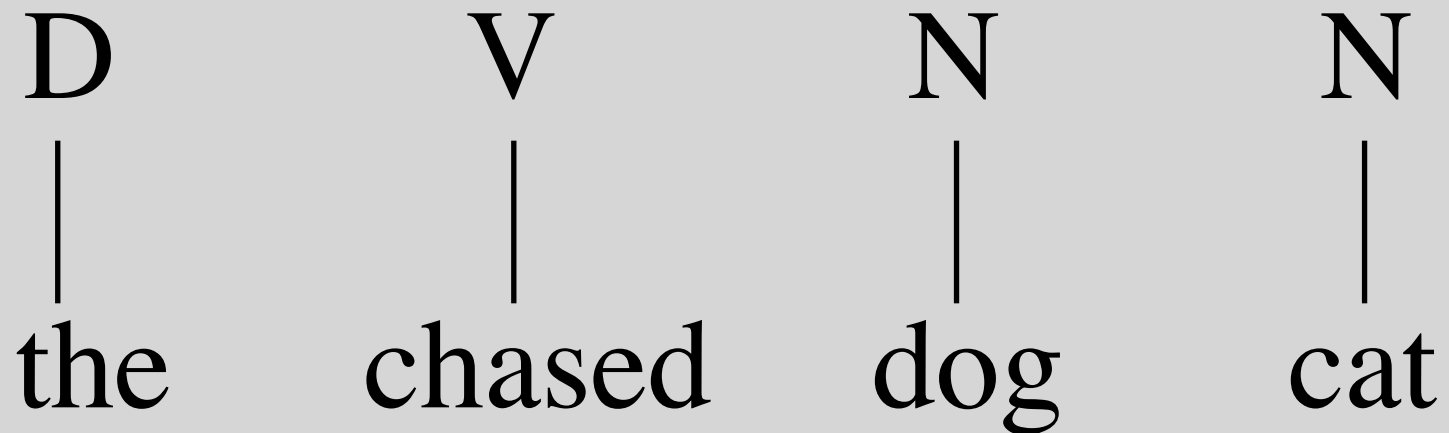
$C_0 \rightarrow C_1 \dots C_n$ is a grammar rule.

Bottom-up Tree Construction

D: *the*

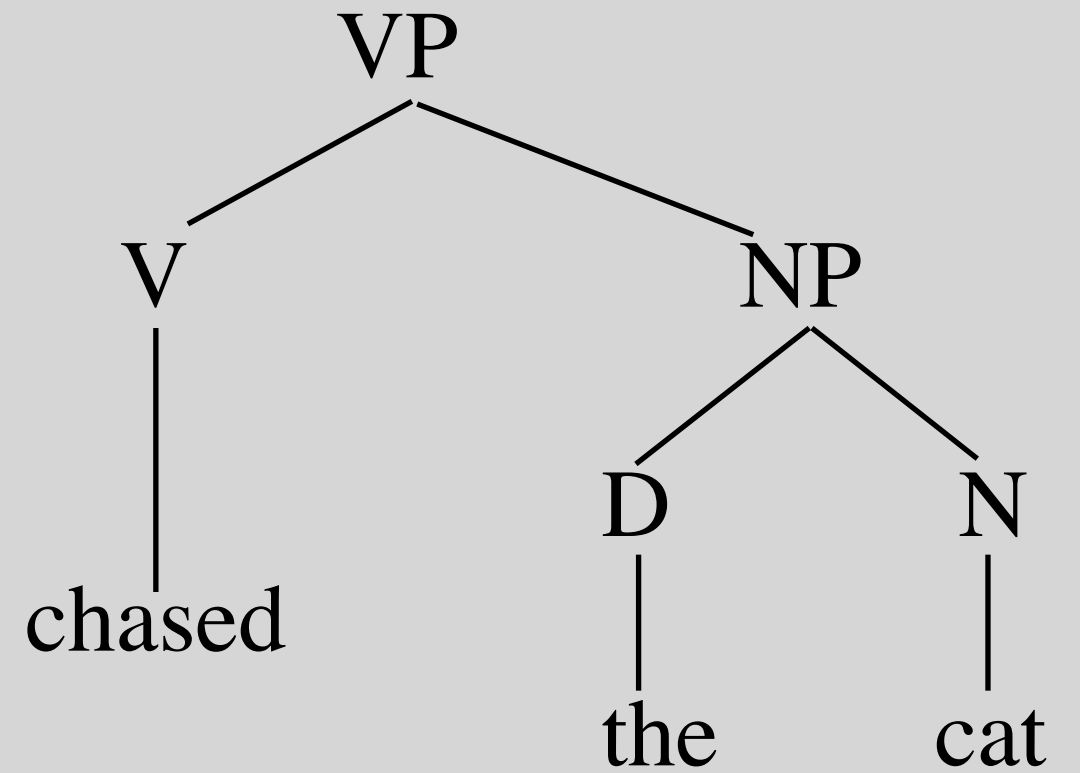
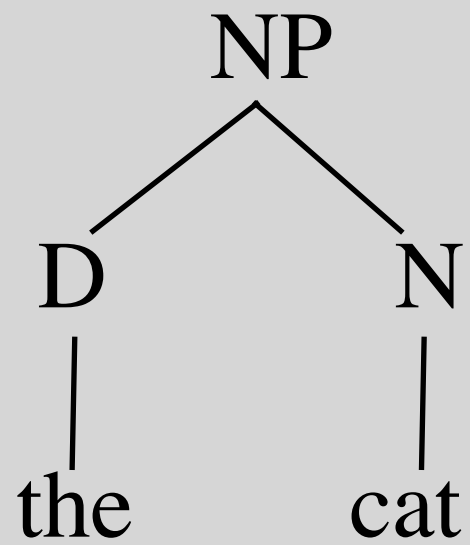
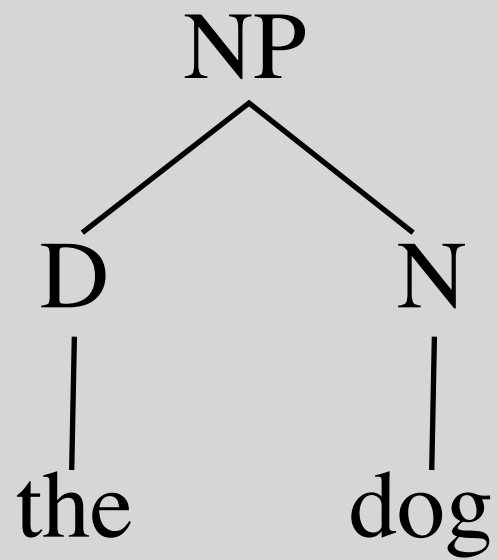
V: *chased*

N: *dog, cat*

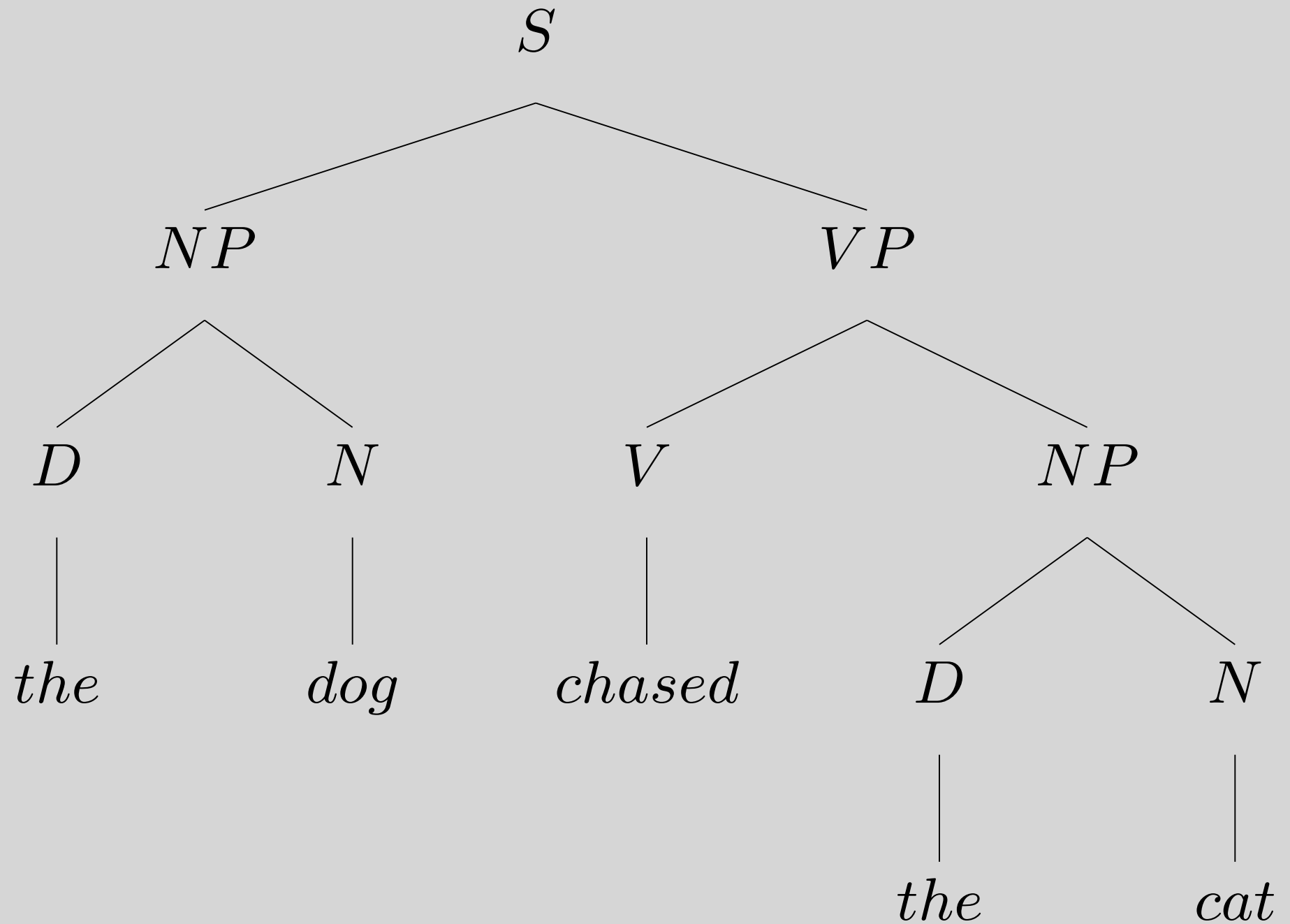


$NP \rightarrow D N$

$VP \rightarrow V NP$

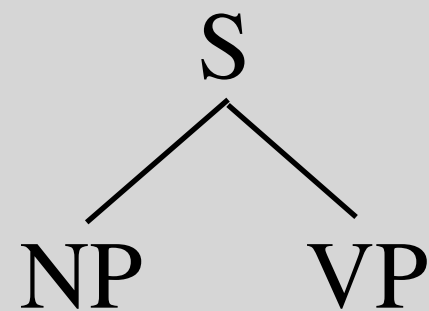


$S \rightarrow NP VP$

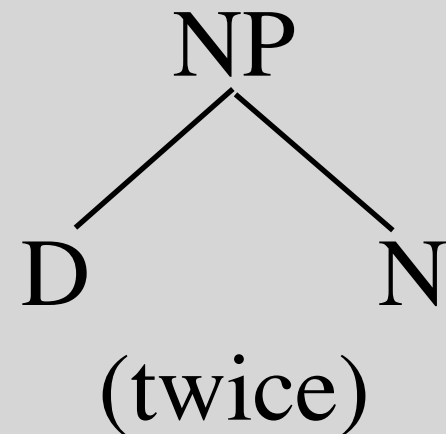


Top-down Tree Construction

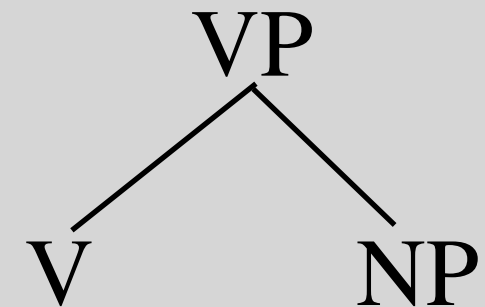
$S \longrightarrow NP VP$

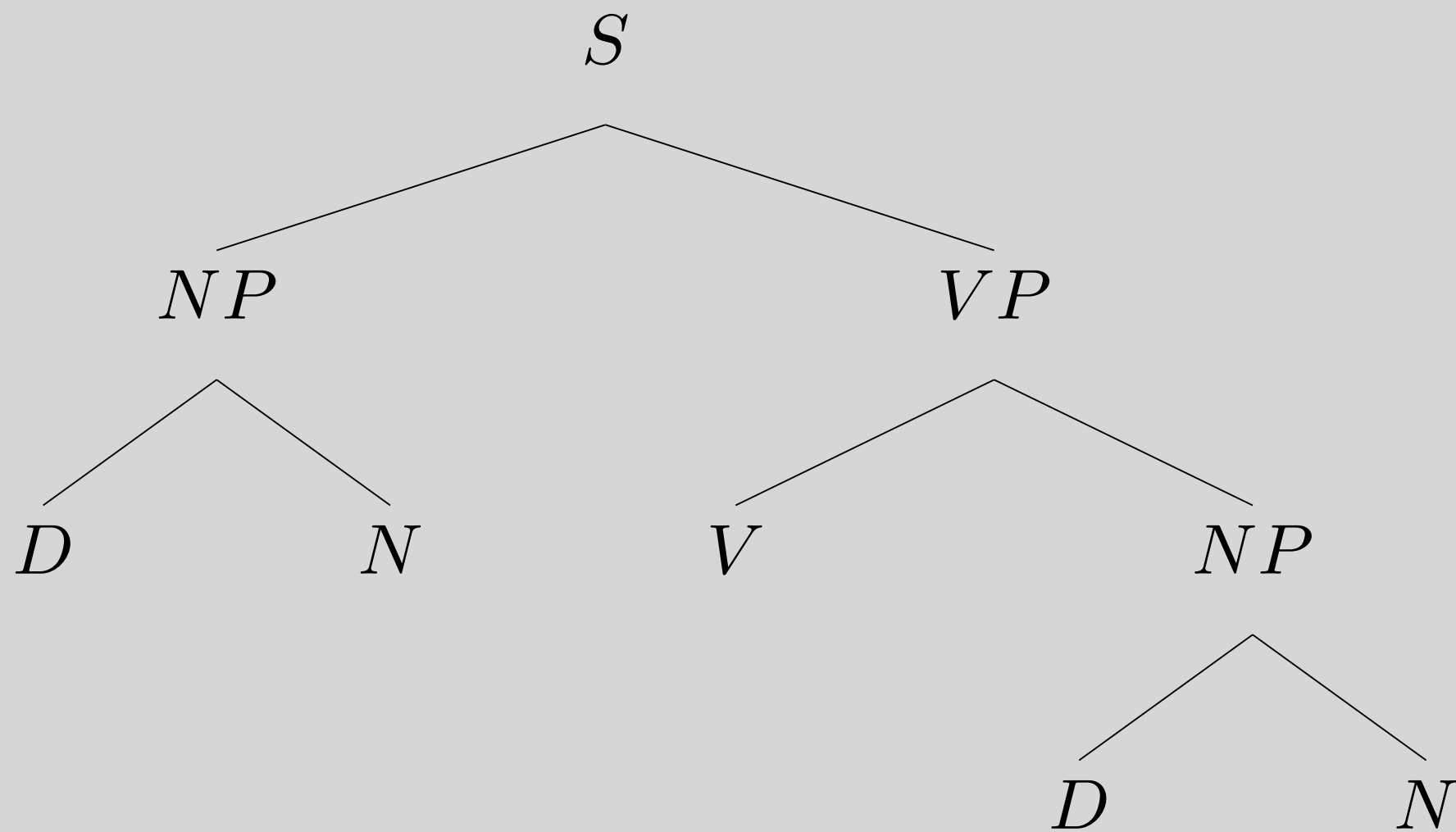


$NP \longrightarrow D N$



$VP \longrightarrow V NP$



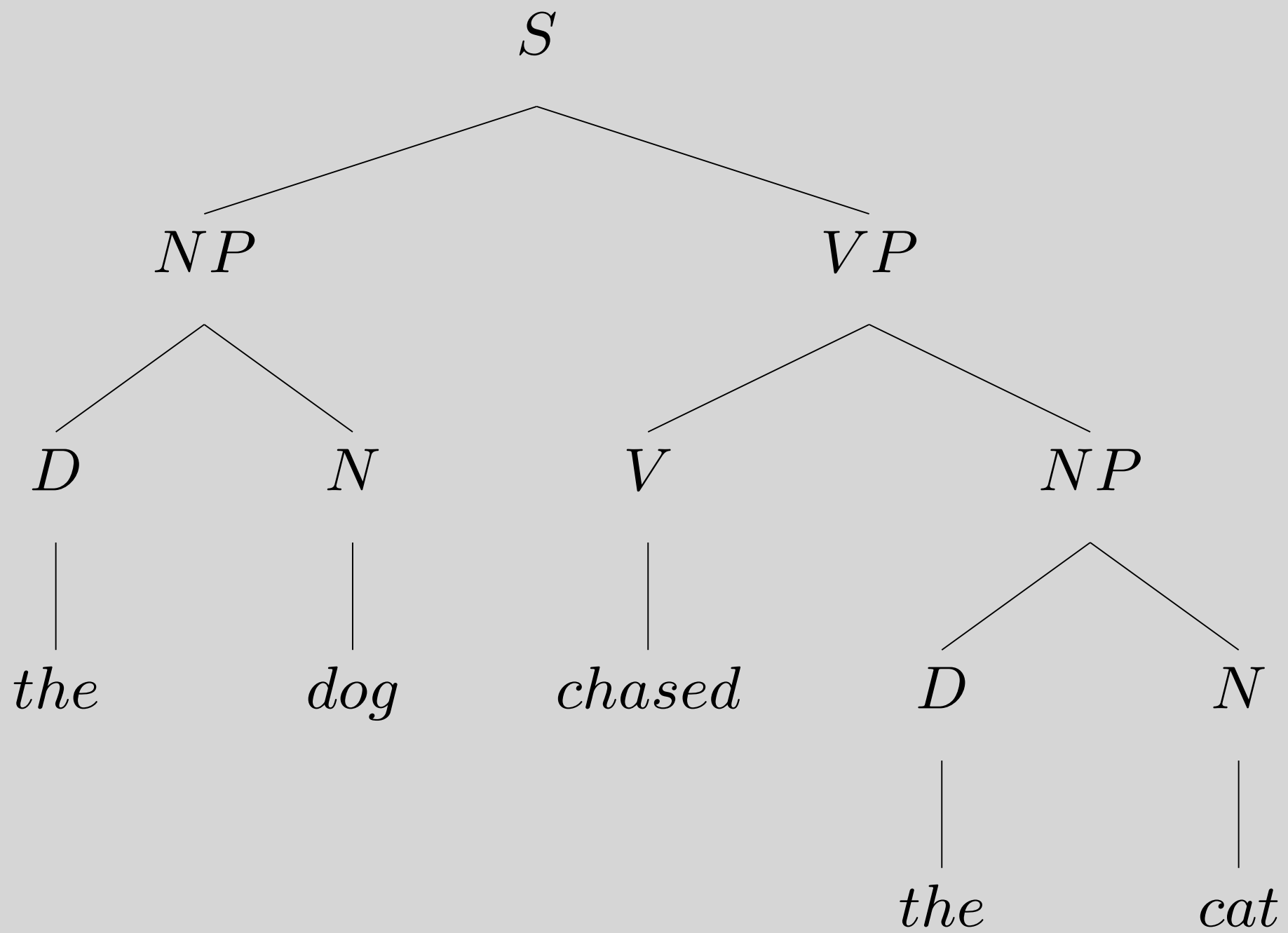


D
|
the

V
|
chased

N
|
dog

N
|
cat



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

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



W For parsing, which style feels most intuitive to you?

Bottom up

Top down

Left to right

Something else

None of the above

Weaknesses of CFG (w/atomic node labels)

- It doesn't tell us what constitutes a linguistically natural rule

$$VP \rightarrow P NP$$
$$NP \rightarrow VP S$$

- Rules get very cumbersome once we try to deal with things like agreement and transitivity.
- It has been argued that certain languages (notably Swiss German and Bambara) contain constructions that are provably beyond the descriptive capacity of CFG.

Agreement & Transitivity

S	→	NP-SG VP-SG	VP-SG	→	IV-SG
S	→	NP-PL VP-PL	VP-PL	→	IV-PL
NP-SG	→	(D) NOM-SG	VP-SG	→	TV-SG NP
NP-PL	→	(D) NOM-PL	VP-PL	→	TV-PL NP
NOM-SG	→	NOM-SG PP	VP-SG	→	DTV-SG NP NP
NOM-PL	→	NOM-PL PP	VP-PL	→	DTV-PL NP NP
NOM-SG	→	N-SG	VP-SG	→	CCV-SG S
NOM-PL	→	N-PL	VP-PL	→	CCV-PL S
NP	→	NP-SG	VP-SG	→	VP-SG PP
NP	→	NP-PL	VP-PL	→	VP-PL PP
	

Shieber 1985

- Swiss German example:

... mer d'chind em Hans es huus lönd hälfe aastriiche
... we the children-ACC Hans-DAT the hous-ACC let help paint
... we let the children help Hans paint the house

- Cross-serial dependency:

- *let* governs case on *children*

- *help* governs case on *Hans*

- *paint* governs case on *house*

Shieber 1985

- Define a new language $f(\text{SG})$:

$$\begin{array}{llll} f(\text{d'chind}) & = & a & f(\text{Jan säit das mer}) & = & w \\ f(\text{em Hans}) & = & b & f(\text{es huus}) & = & x \\ f(\text{lönde}) & = & c & f(\text{aastriiche}) & = & y \\ f(\text{hälfe}) & = & d & f([\text{other}]) & = & z \end{array}$$

- Let r be the regular language $wa^*b^*xc^*d^*y$
- $f(\text{SG}) \cap r = wa^mb^nc^md^ny$
- $wa^mb^nc^md^ny$ is not context free.
- But context free languages are closed under intersection. w/reg languages
- $\therefore f(\text{SG})$ (and by extension Swiss German) must not be context free.

Strongly/weakly CF

- A language is *weakly* context-free if the set of strings in the language can be generated by a CFG.
- A language is *strongly* context-free if the CFG furthermore assigns the correct structures to the strings.
- Shieber's argument is that SG is not *weakly* context-free and *a fortiori* not *strongly* context-free.
- Bresnan et al (1983) had already argued that Dutch is *strongly* not context-free, but the argument was dependent on linguistic analyses.

On the other hand....

- It's a simple formalism that can generate infinite languages and assign linguistically plausible structures to them.
- Linguistic constructions that are beyond the descriptive power of CFG are rare.
- It's computationally tractable and techniques for processing CFGs are well understood.

So.....

- CFG has been the starting point for most types of generative grammar.
- The theory we develop in this course is an extension of CFG.

Overview

- Two insufficient theories
- Formal definition of CFG
- Constituency, ambiguity, constituency tests
- Central claims of CFG
- Weaknesses of CFG
- Reading questions

Reading Questions

- Why do we need NOM?
- For the Prepositional Phrases rule (no. 20) why would we break $\text{NOM} \rightarrow \text{N}$ and $\text{NOM} \rightarrow \text{NOM PP}$ separate? Wouldn't it just be $\text{NOM} \rightarrow \text{NOM N PP}$? And Is the difference between NP and NOM adjective before a noun? Also, I notice adjective A is gone from one rule $\text{NP} \rightarrow (\text{D}) \text{A}^* \text{N PP}^*$ to the other $\text{NP} \rightarrow (\text{D}) \text{NOM}$. Why is that?

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

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



Which rules license "no painting by A or drawing by B was displayed"? (Together with PP -> P NP & coord rule)

NP -> NP PP

NP -> D NOM; NOM -> NOM PP; NOM -> N

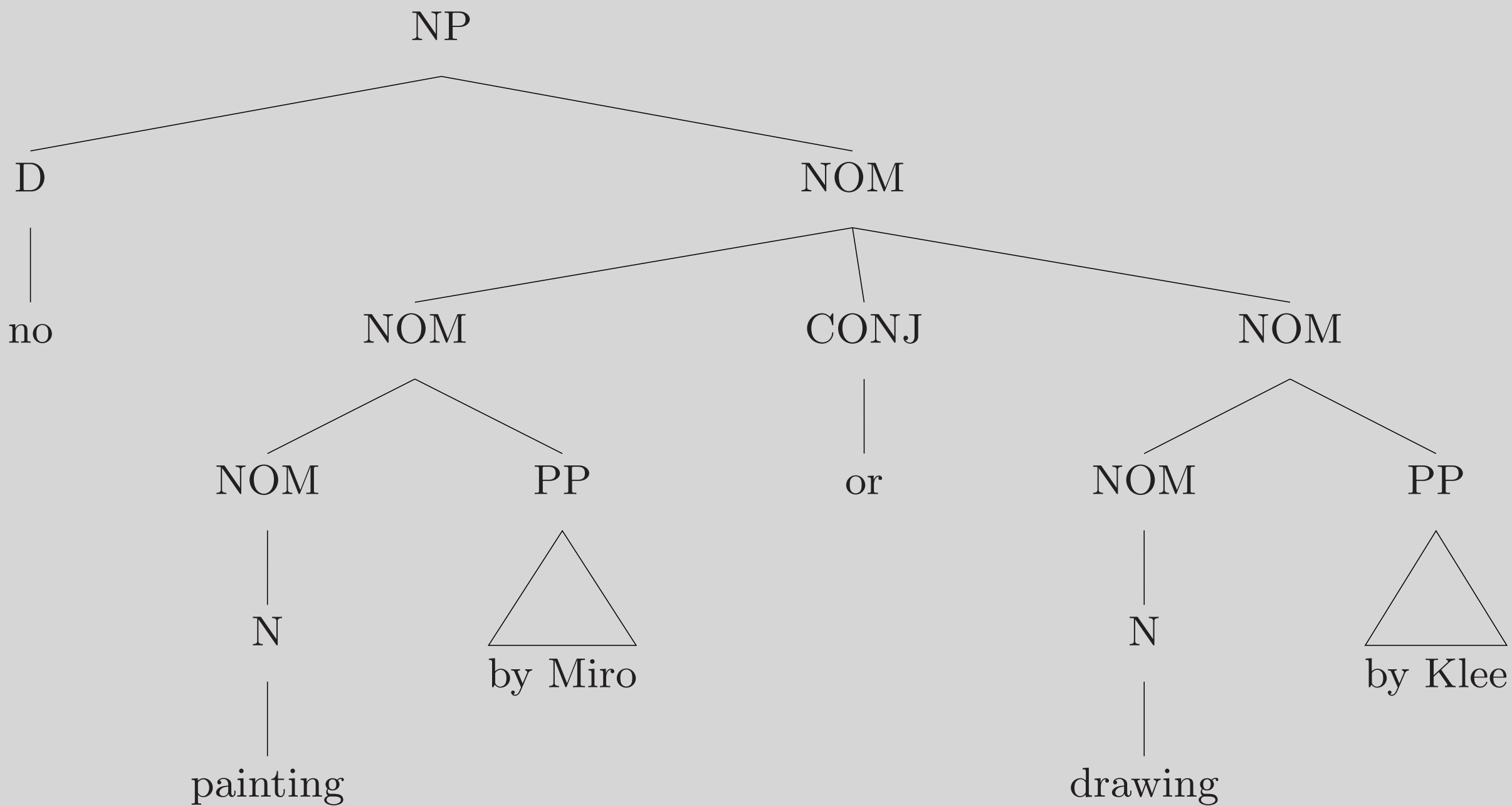
NP -> (D) N (PP)

None of the above

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



Reading Questions

- Is there a reason to prefer ternary (or quaternary etc.) trees over binary trees or vice versa?
- There is an explanation of the Kleene star (p. 24) which is demonstrated in (11) NP > (D) N PP* but then completely dropped and not appearing on the trees nor in the rules. Is there a reason for that?

Reading Questions

- The example in 2.7.2 involves transitivity and the fact V requires intransitive, transitive, and ditransitive subcategories. However, English seems to have many verbs which could fit two or more of these subcategories such as *sing* (*The bird sings.* or *The performer sings a song.*). Are the subcategories reserved for verbs which can only be placed in one or would we put a verb like *sing* in multiple subcategories? Does the lexicon allow for categories with subcategories to still be used? In other words, would we classify *sing* as V in the lexicon while *laugh* and *die* would be intransitive, *show* and *give* would be ditransitive, etc.?

Reading Questions

- With regards to the constituency diagnostics, do different theories offer predictions for what properties of a phrase might cause it to fail certain diagnostics, but pass others? Are there any predictions that can be made about the differing behavior of phrases that pass many different diagnostics such as VPs or NPs versus ones that don't pass as many such as NOM?

Reading Questions

- How could we reasonably represent sentence inversion if examples like "Dark is the night" are abundant in a language? Should we alter the sequence of the rules accordingly, or adopt a rule in the original order + some stress marker? (Or what are the potential advantages and limitations)

Reading Questions

- How do constituency diagnostics work for languages other than English? Can we use similar tests or are tests handcrafted for each language? Can they be more/less reliable than the ones for English?

Reading Questions

- Why NP and not DP?
- I don't really understand why CFGs can't capture headedness - is it just because the non lexical category names are mnemonic? Or is it that headedness introduces additional behavior that can't be captured by a CFG?

Reading Questions

- How do syntacticians and HPSG account for language evolution? As posted above, 2.7.2 32(b) is "ungrammatical" to the authors but I have seen sentences like that on the Internet as perfectly acceptable to certain subcultures.

Reading Questions

- I am curious as to why we will be learning a "nontransformational grammar" rather than a "transformational grammar" like we learned in the Ling 46x series at UW. What benefit, especially as it has to do with computational linguistics and the implementation of grammars in a system, does a "nontransformational grammar" have over a "transformational" one? It seems to me that using a "nontransformational grammar" might come off as English-centric, but maybe this is a benefit when implementing a grammar for English exclusively?

Reading Questions

- Are there languages whose syntax is more difficult to model by HPSG? How would HPSG model bilingual/sentence-internal code-switching?

Reading Questions

- Can you provide some examples or further explanation of grammars that are not direction and process neutral? e.g. How does a top-down vs. bottom-up perspective change our approach to interpreting a statement or judging its acceptability?
- What do direction neutrality and process neutrality mean in Section 2.5?

Reading Questions

- In the examples intended to demonstrate semantic ambiguity, I'm curious to know how much of that ambiguity is a bi-product of the fact that we're primarily dealing with written, not oral, representations of sentences? When I think about saying a potentially ambiguous sentence out loud, I tend to change my pause lengths and intonation to disambiguate the sentence. E.g. the way I say *I forgot how good beer tastes* changes quite drastically when I change the intended meaning. Does this sort of disambiguation have a place in syntax, or is it purely attributed to other, non-syntactic, fields of study?