Ling 566
Oct 5, 2010

Context-Free Grammar

Overview

- Failed attempts

- Formal definition of CFG

- Constituency, ambiguity, constituency tests

- Central claims of CFG

- Order independence

- Weaknesses of CFG

- If time: Work through Chapter 2, Problem 1

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 (CFG)

Chomsky Hierarchy

Type 0 Languages

Context-Sensitive Languages

Context-Free Languages

Regular Languages
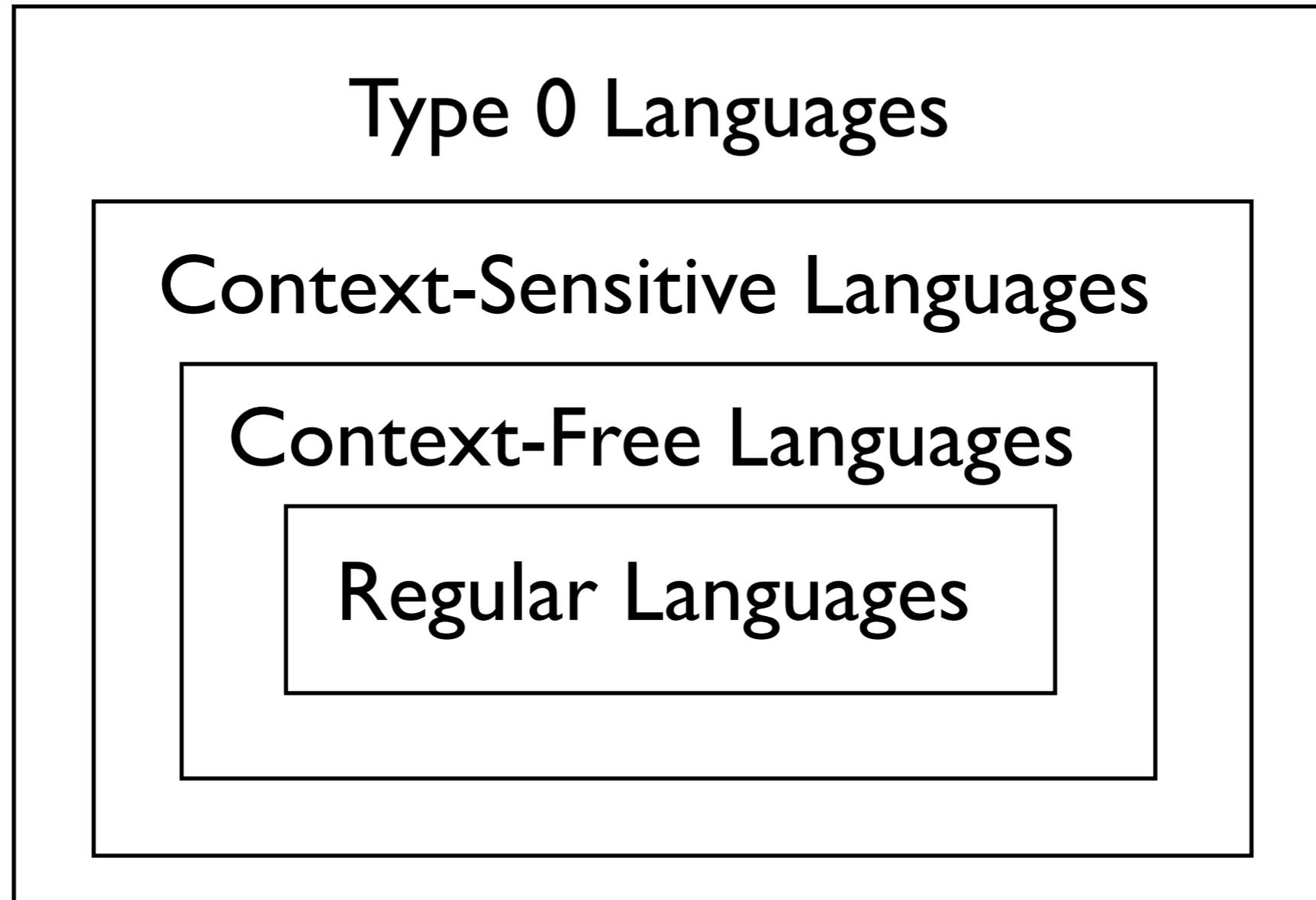
Context-Free Grammar

- A quadruple: $< C, \Sigma, P, S >$

  - 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

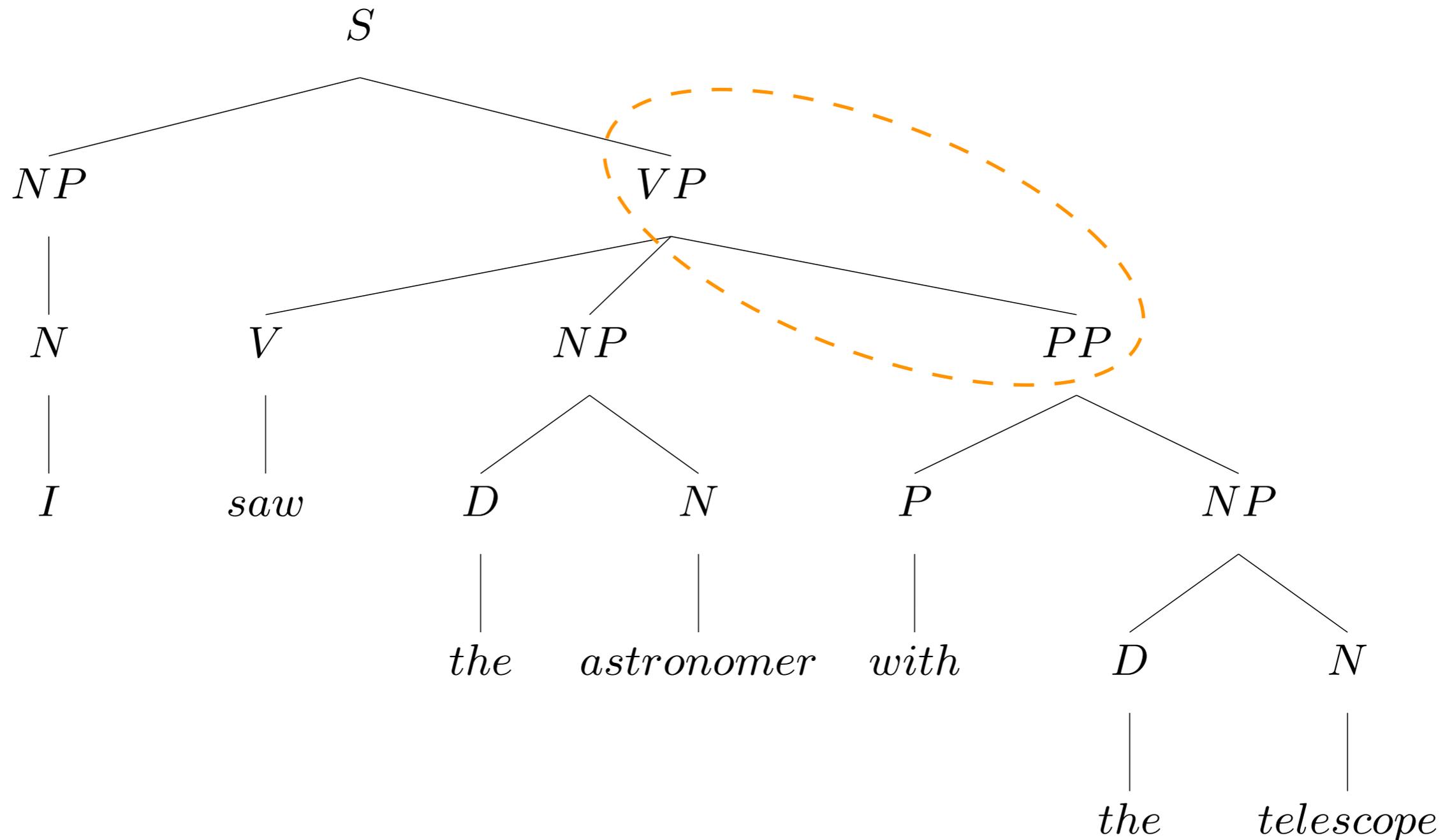N: birds, fleas, dog, hunter, I
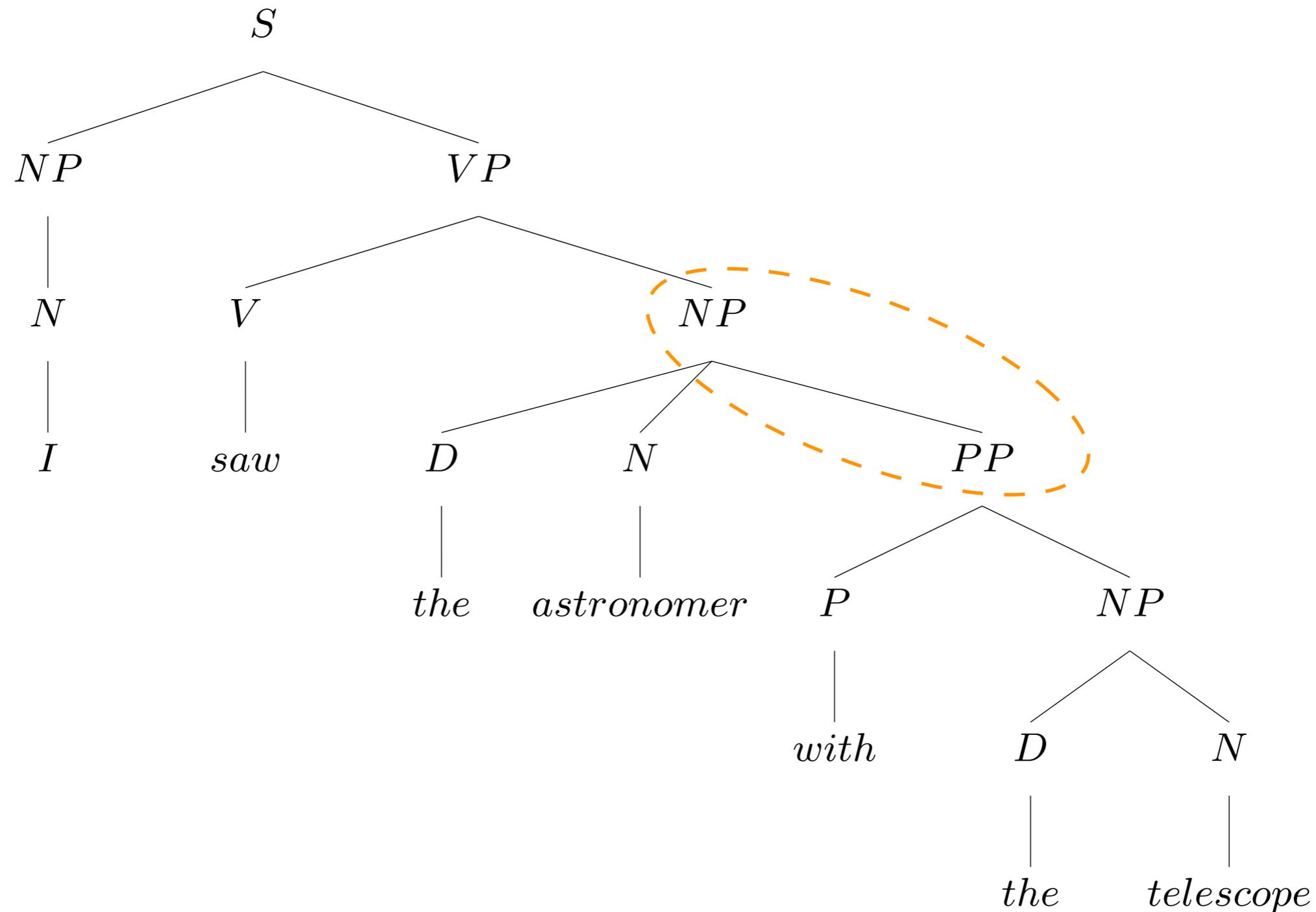
V: attack, ate, watched

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→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

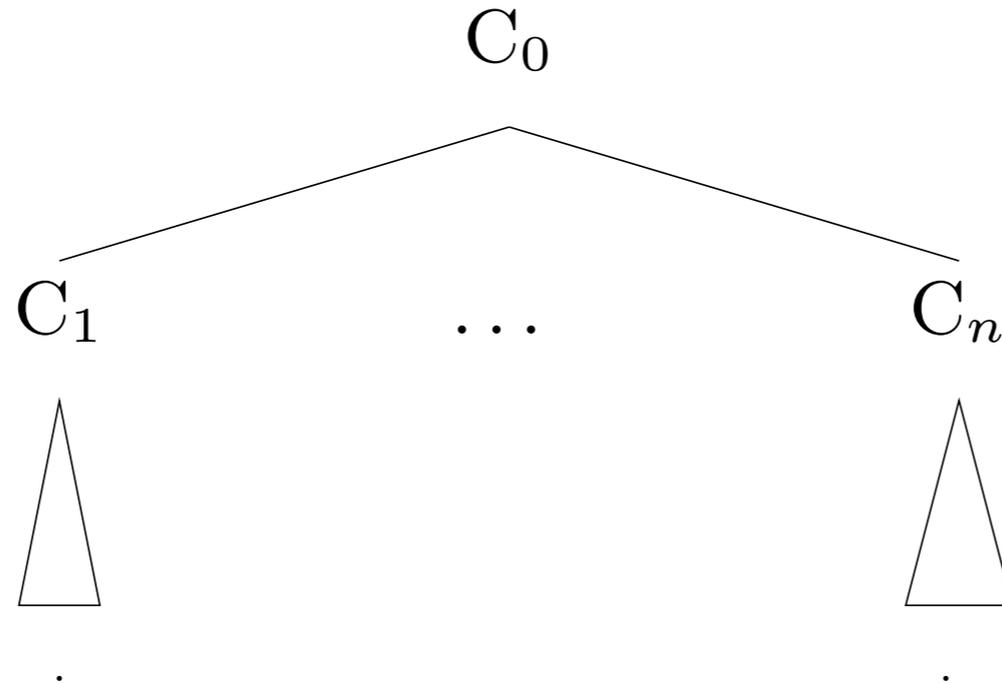S $\longrightarrow$ NP VP

NP $\longrightarrow$ D N

VP $\longrightarrow$ V NP

D: *the*

V: *chased*

N: *dog, cat*
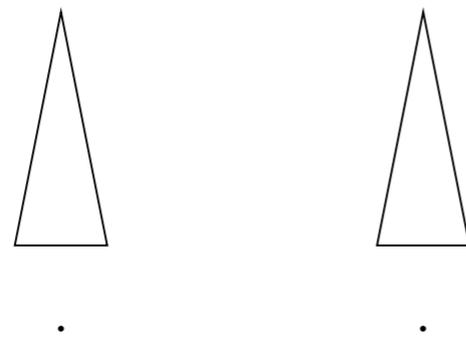
Trees and Rules

$$C_0$$

$$C_1 \quad \dots \quad C_n$$

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

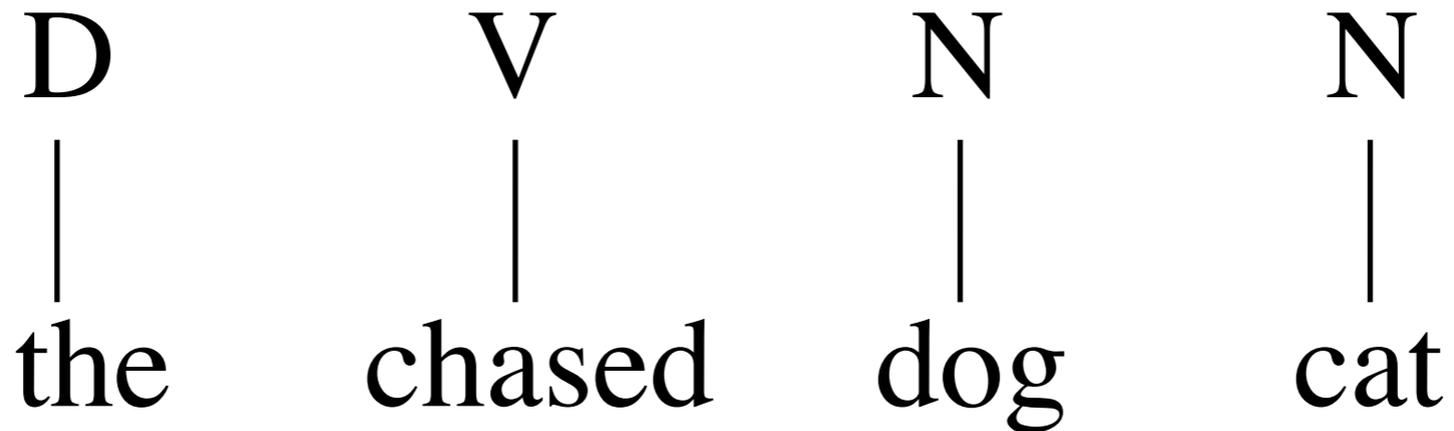$$C_n \;, \dots, C_n$$ are well-formed trees, and

$$C_0 \to \; C_1 \dots C_n$$ is a grammar rule.
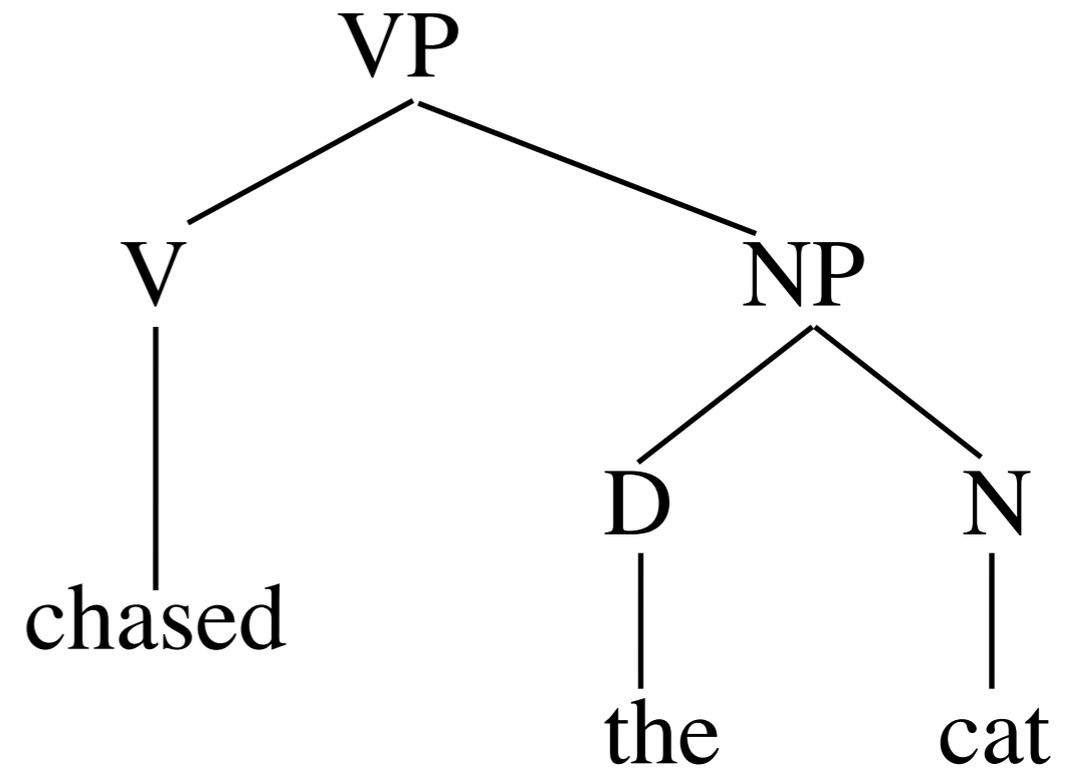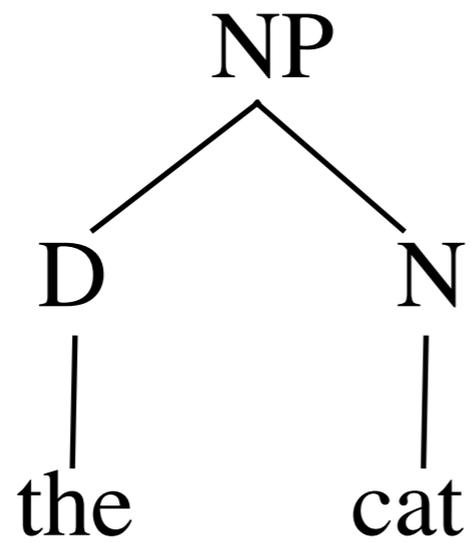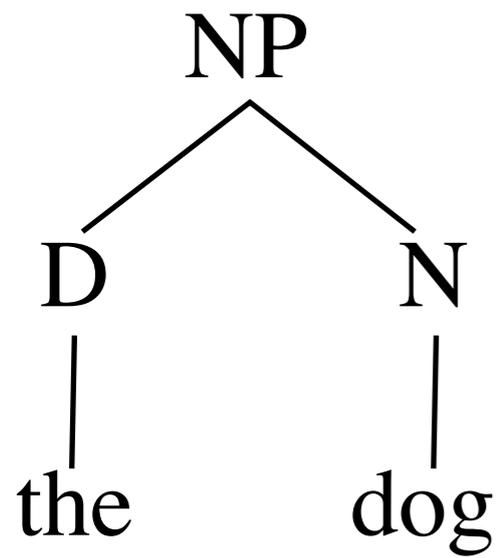
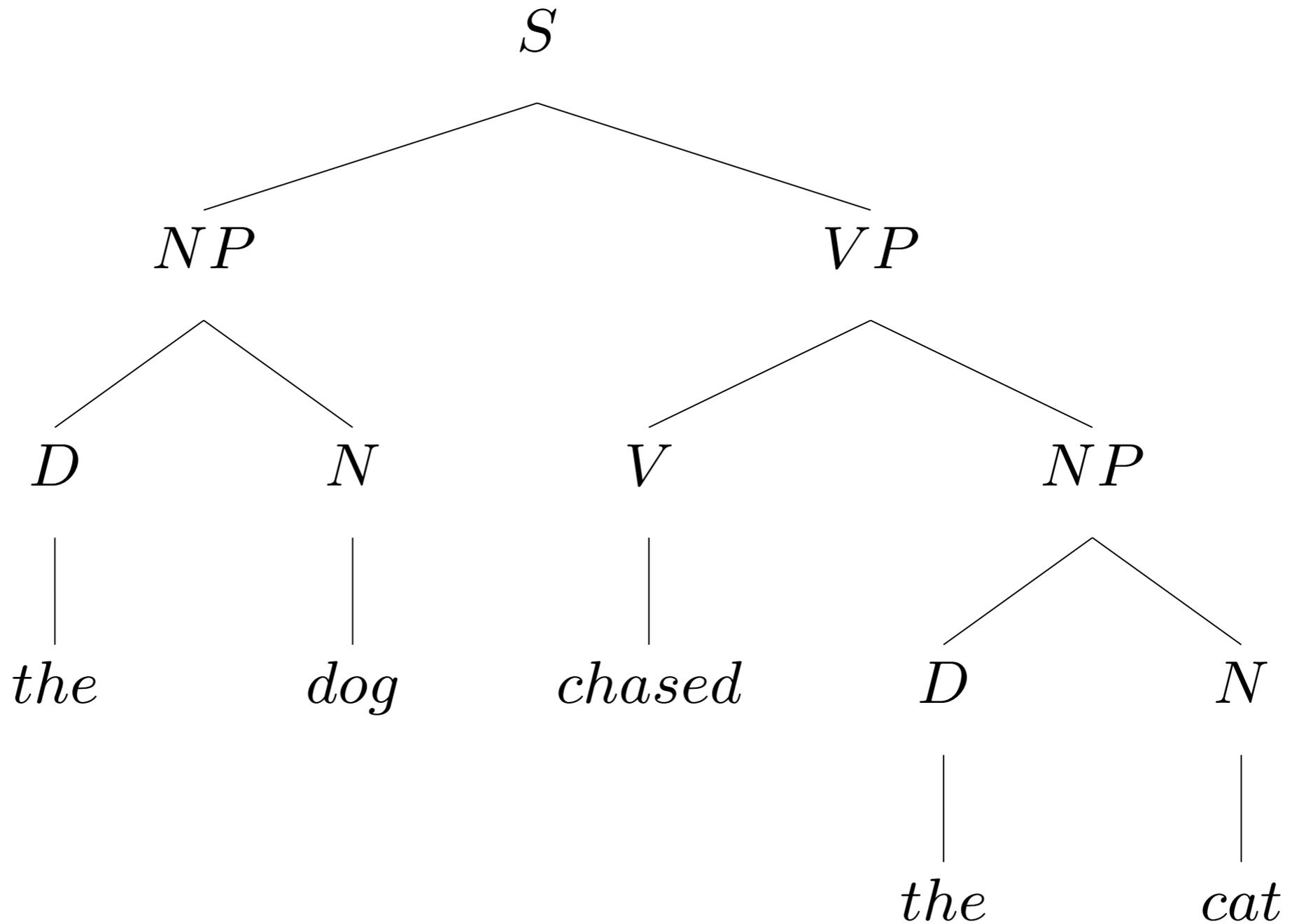# Bottom-up Tree Construction

D: *the*
V: *chased*
N: *dog, cat*

D	V	N	N
the	chased	dog	cat

NP ⟶ D N   VP ⟶ V NP

$$S \longrightarrow NP \; VP$$

```
                              S
                    ┌─────────┴─────────┐
                   NP                    VP
                ┌───┴───┐           ┌─────┴─────┐
                D       N           V           NP
                │       │           │        ┌───┴───┐
               the     dog       chased      D       N
                                             │       │
                                            the     cat
```

Top-down Tree Construction

S ⟶ NP VP   NP ⟶ D N   VP ⟶ V NP

(twice)

D   V   N   N
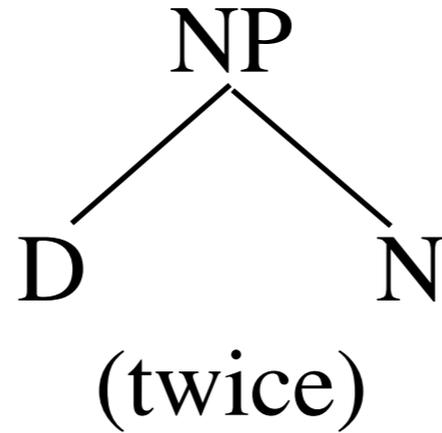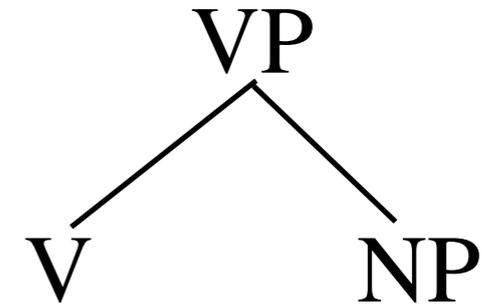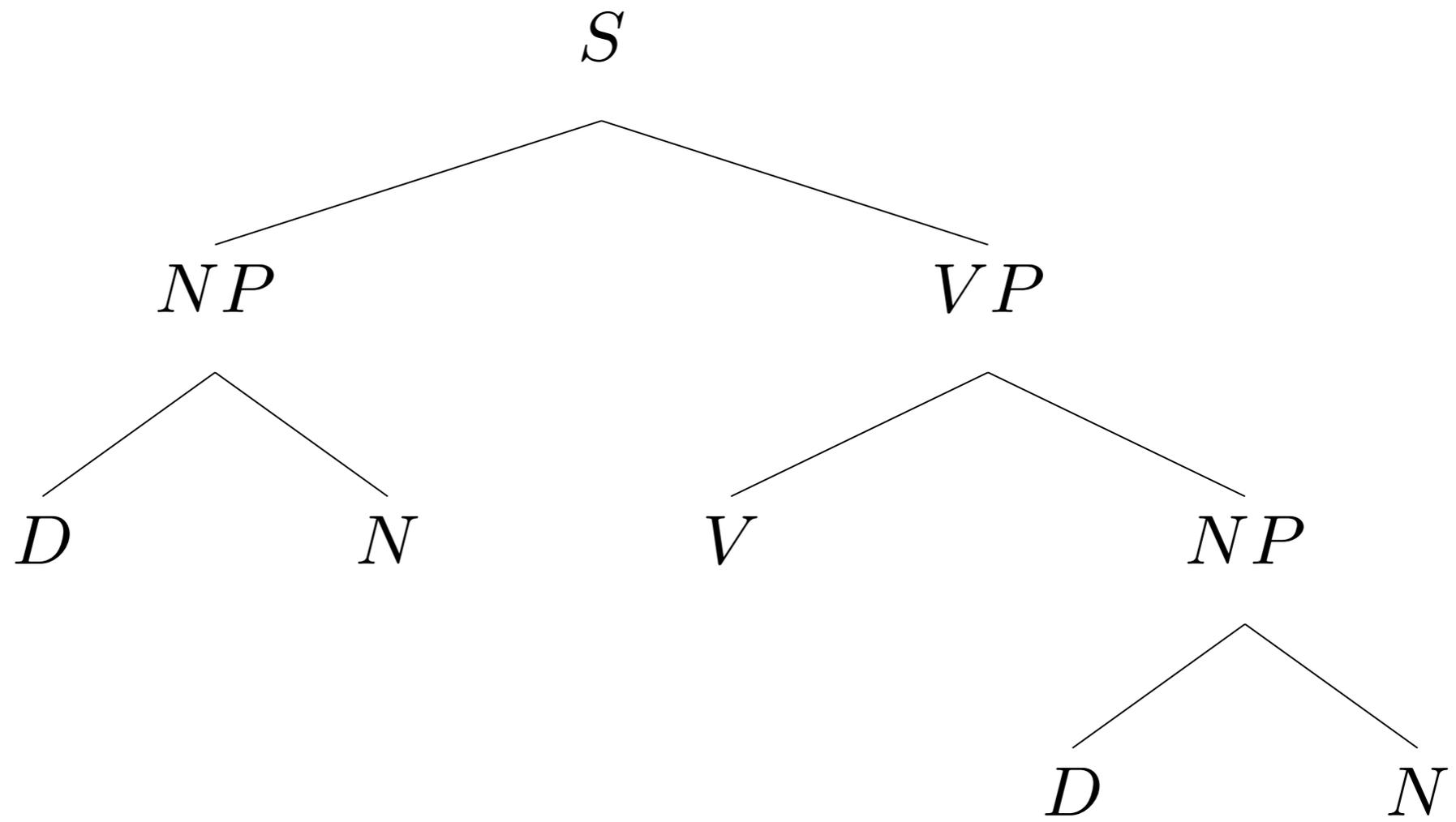
the chased dog cat
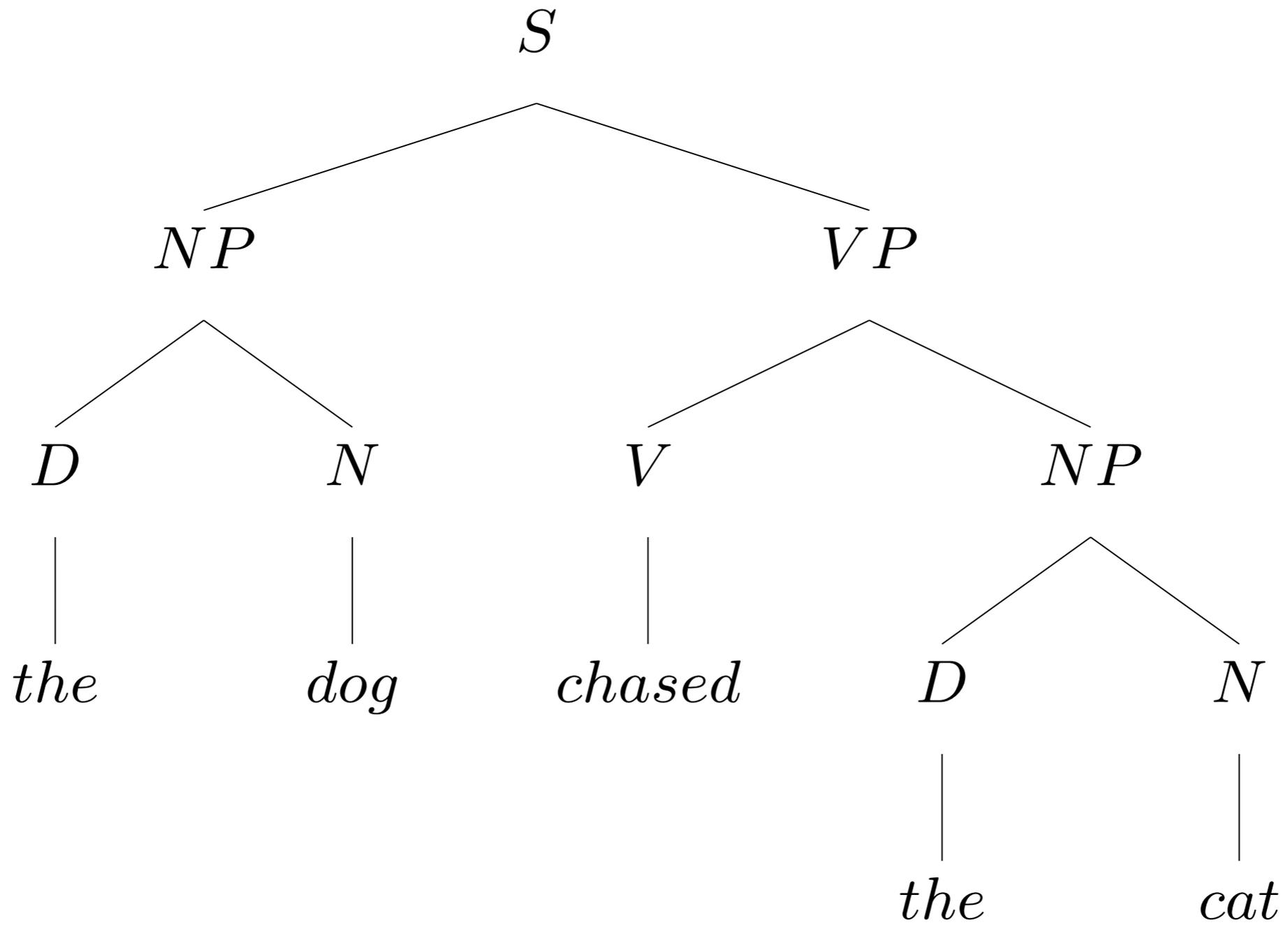
Weaknesses of CFG (atomic node labels)

- 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.

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.

Chapter 2, Problem 1

S → NP VP   NOM → NOM PP

NP → (D) NOM   VP → VP PP

VP → V (NP) (NP)   PP → P NP

NOM → N   X → X$^+$ CONJ X

D: a, the

V: admired, disappeared, put, relied

N: cat, dog, hat, man, woman, roof

P: in, on, with

CONJ: and, or

Chapter 2, Problem 1

- Well-formed English sentence unambiguous according to this grammar

- Well-formed English sentence ambiguous according to this grammar

- Well-formed English sentence not licensed by this grammar

- String licensed by this grammar that is not a well-formed English sentence

- How many strings does this grammar license?

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 $f(\mathrm{SG})$:

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

- Let r be the regular language $wa^*b^*xc^*d^*y$

- $f(\mathrm{SG}) \cap r = wa^m b^n x c^m d^n y$

- $wa^m b^n x c^m d^n y$ is not context free.

- But context free languages are closed under intersection.

- $\therefore f(\mathrm{SG})$ (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 SW 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.

Overview

- Failed attempts

- Formal definition of CFG

- Constituency, ambiguity, constituency tests

- Central claims of CFG

- Order independence

- Weaknesses of CFG

- Next time: Feature structures