Ling 566
Sept 26, 2024

Introduction, organization, first attempts at a theory of grammar

© 2003 CSLI Publications

Overview

- Syllabus
- Prescriptive/descriptive grammar; Competence/performance
- Some history
- Why study syntax?
- Two theories that won't work
- Start on CFG

© 2003 CSLI Publications

But first…

- Building Evacuation information
- Evacuate the building using the nearest exit (or alternate if nearest exit is blocked).
- Do not use elevators.
- Take personal belongings (keys, purses, wallets, etc.).
- Secure any hazardous materials or equipment before leaving.
- Follow directions given by evacuation wardens.
- Go to Evacuation Assembly Point (EAP) designated in your building's evacuation plan and on building emergency evacuation signs. ==> Denny Yard
- Assist people with disabilities.
- https://www.washington.edu/uwem/plans-and-procedures/uw-emergency-procedures/

© 2003 CSLI Publications

But first...

- COVID-19
  - Masks are encouraged indoors
  - Please stay home if you are unwell, and attend online if you are able
  - You are welcome to attend online at any point, for any reason
  - https://www.washington.edu/coronavirus/

© 2003 CSLI Publications

Syllabus: Course requirements

http://courses.washington.edu/ling566

© 2003 CSLI Publications

The winning strategy

- Work together: make study groups
- Homework: Discuss as much as you want, write up your own answers
- Exams: No discussion
- Post to Canvas discussions
- Read the book before class (and after again, if necessary)
- Ask questions ... early and often!

© 2003 CSLI Publications

Resources

- Glossary at back of textbook
- Bender 2013 (“100 things”)
- Grammar summaries and Appendix A
- Answers to exercises at back of book
- Canvas, study groups, office hours...

© 2003 CSLI Publications

Self introductions

- Name
- Pronouns
- A part of speech that speaks to you right now

© 2003 CSLI Publications

Why study syntax?

0

Nobody has responded yet.

Hang tight! Responses are coming in.

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

Title logo

Why take this class?

0

Want to research syntax

0

Help me build better language technology

0

Help me study languages better

0

Want to get a better sense of "how language works"

0

Required for CLMS

0

I really REALLY love feature structures

0

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

Two Conceptions of Grammar

PRESCRIPTIVE

- Rules against certain usages. Few if any rules for what **is** allowed
- Proscribed forms generally in use
- Explicitly normative enterprise

DESCRIPTIVE

- Rules characterizing what people do say
- Goal to characterize all and only what speakers find acceptable
- Tries to be scientific

© 2003 CSLI Publications

Uses of Grammar

PRESCRIPTIVE

- Identify speaker's socioeconomic class & education level
- Identify level of formality of a particular usage

DESCRIPTIVE

- Understand how people produce & understand language
- Identify similarities & differences across languages
- Development of language technologies

© 2003 CSLI Publications

Prescriptive grammar

- Examples of silly prescriptive rules?
- Examples of useful prescriptive rules?
- Compling applications which might need to encode prescriptive rules?

© 2003 CSLI Publications

Fill in the blanks:
he/his, *they/their*, or something else?

1. Everyone insisted that ___ record was unblemished.
2. Everyone drives ___ own car to work.
3. Everyone was happy because ___ passed the test.
4. Everyone left the room, didn't ___?
5. Everyone left early. ___ seemed happy to get home.

© 2003 CSLI Publications

Title logo

For which sentences do he/his sound better than they/their?

0

1. Everyone insisted that ___ record was unblemished.
2. Everyone drives ___ own car to work.
3. Everyone was happy because ___ passed the test.
4. Everyone left the room, didn't ___?
5. Everyone left early. ___ seemed happy to get home.

All 0

Some 0

None 0

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

Descriptive Grammar: an example

F--- yourself!

Go f--- yourself!

F--- you!

*Go f--- you!

- Who taught you this?
- How did you learn it?

© 2003 CSLI Publications

Kinds of Things We'll Worry About

- Where to use reflexives (e.g. *myself)* vs. ordinary pronouns (*I* or *me*)
- Agreement (e.g. *We sing* vs. **We sings*)
- Word order (e.g. **Sing we*)
- Case (e.g. **Us sing*)
- Coordinate conjunction (e.g. *We sing and dance*)
- How to form questions, imperatives, negatives…
- …and much more

© 2003 CSLI Publications

Competence *vs*. Performance

- The Distinction
  - Competence - knowledge of language
  - Performance - how the knowledge is used
- Examples

  That Sandy left bothered me.

  That that Sandy left bothered me bothered Kim

  That that that Sandy left bothered me bothered Kim bothered Bo

  The horse raced past the barn fell

© 2003 CSLI Publications

Competence *v.* Performance

You are what you eat

You are what what you eat eats, too

You are what what what you eat eats eats, too

© 2003 CSLI Publications

Acceptability *vs*. grammaticality

- A sentence is **acceptable** if native speakers say it sounds good.
- A sentence is **grammatical** (with respect to a particular grammar) if the grammar licenses it.
- Linguists are sometimes sloppy about the difference.

© 2003 CSLI Publications

Some History

- Writings on grammar go back at least 3000 years
- Until 200 years ago, almost all of it was prescriptive
- Until ~60 years ago, most linguistic work concerned sound systems (phonology), word structure (morphology), and the historical relationships among languages

© 2003 CSLI Publications

The Generative Revolution

- Noam Chomsky's work in the 1950s radically changed linguistics, making syntax central.
- Chomsky has been the dominant figure in linguistics ever since.
- The theory we will develop is in the tradition started by Chomsky, but diverges from his work in many ways.

© 2003 CSLI Publications

Main Tenets of Generative Grammar

- Grammars should be formulated precisely and explicitly.
- Languages are infinite, so grammars must be tested against invented data, not just attested examples.
- The theory of grammar is a theory of human linguistic abilities.

© 2003 CSLI Publications

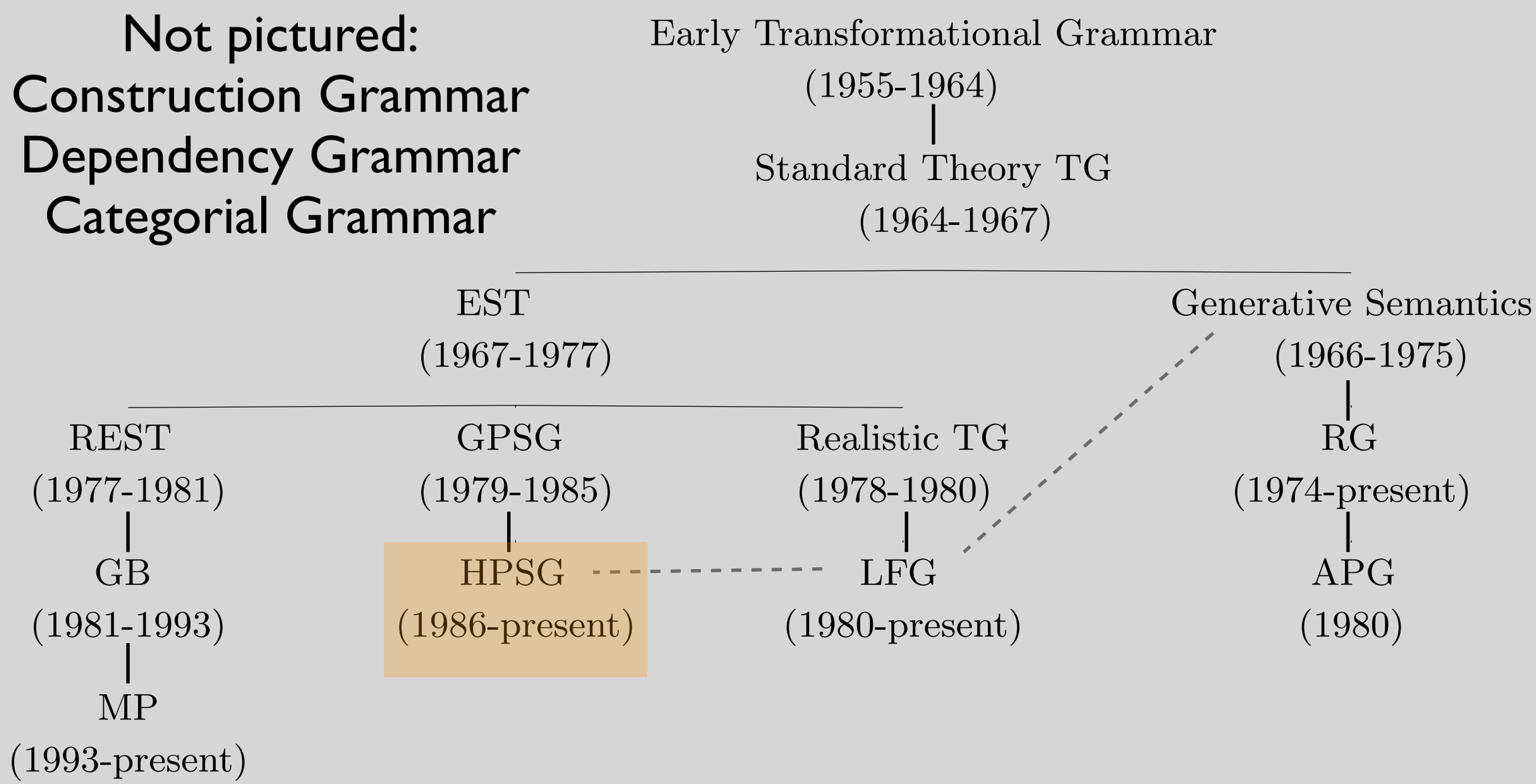
Some of Chomsky's Controversial Claims

- The superficial diversity of human languages masks their underlying similarity.
- All languages are fundamentally alike because linguistic knowledge is largely innate.
- The central problem for linguistics is explaining how children can learn language so quickly and easily.

© 2003 CSLI Publications

Partial Family Tree of Syntactic Theories

Not pictured:
Construction Grammar
Dependency Grammar
Categorial Grammar



© 2003 CSLI Publications

Breakouts: Why Study Syntax?

- Why should linguists study syntax?
- Why should computational linguists study syntax?
- Should anyone else study syntax? Why?

© 2003 CSLI Publications

Breakouts: Why Study Syntax?

- Why are you studying syntax?
- What are you looking forward to about this class?
- What are you nervous about about this class?

© 2003 CSLI Publications

Insufficient Theory #1

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

© 2003 CSLI Publications

Insufficient Theory #2: FSMs

- the noisy dogs left

D A N V

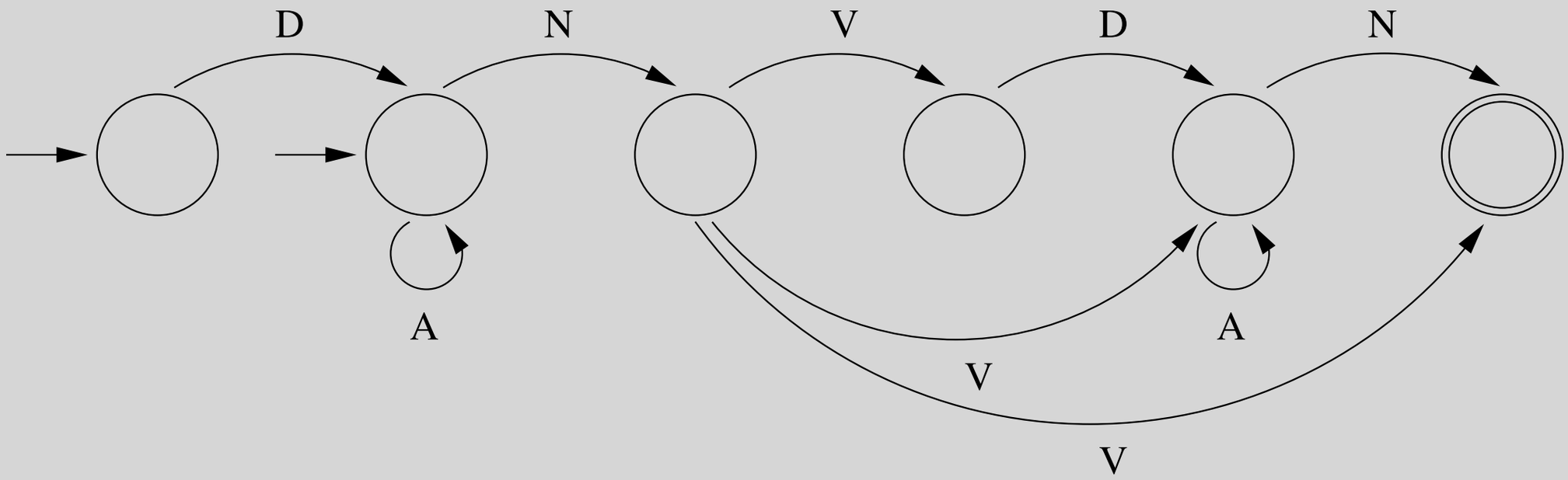
- the noisy dogs chased the innocent cats

D A N V D A N

- $a^* = \{\emptyset, a, aa, aaa, aaaa, \ldots\}$
- $a^+ = \{a, aa, aaa, aaaa, \ldots\}$
- (D) A* N V ((D) A* N)

© 2003 CSLI Publications

A Finite State Machine



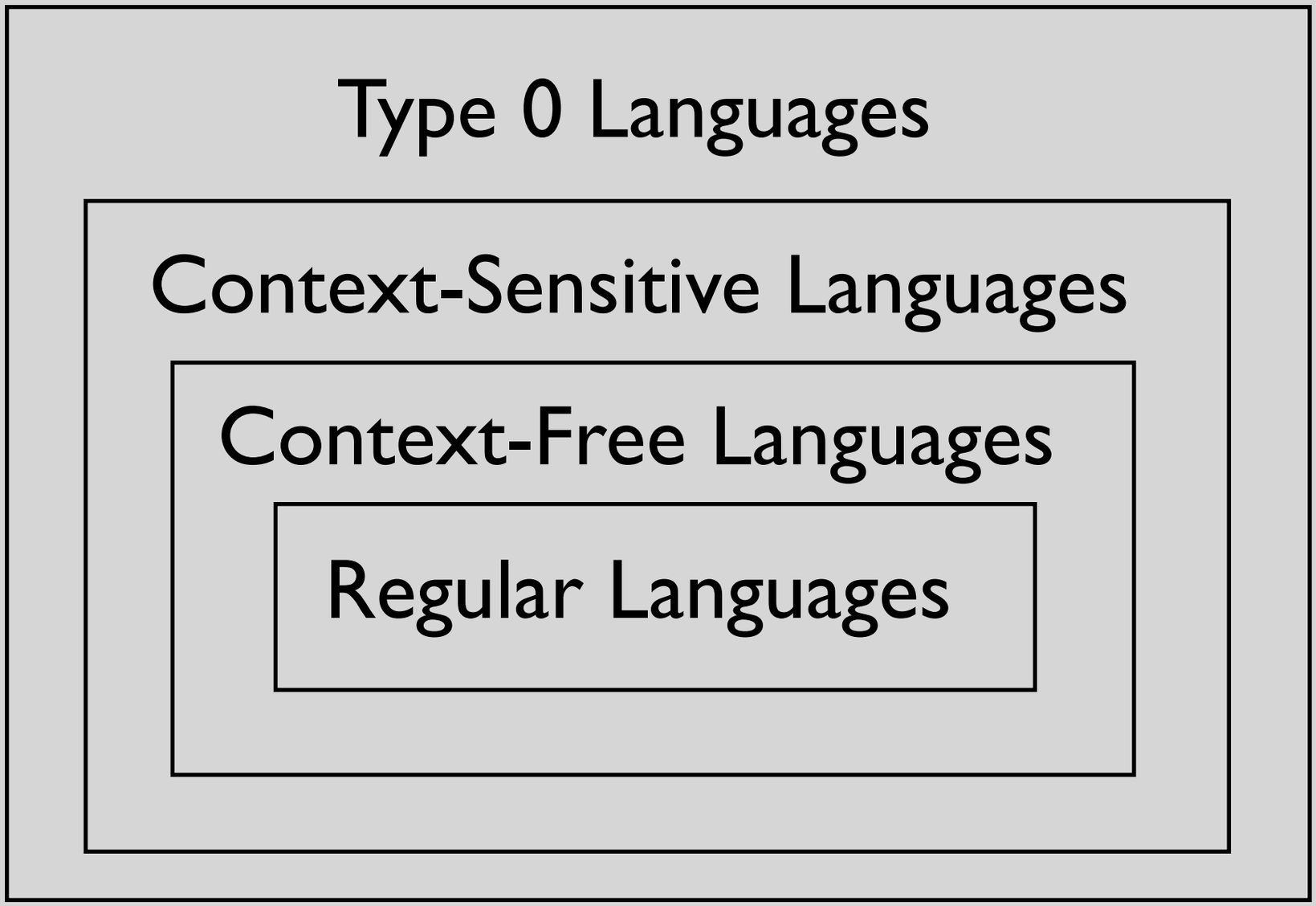
© 2003 CSLI Publications

FSMs for Grammar, cont

- Why are FSMs insufficient as a representation of natural language syntax?
- How might they be useful anyway?

© 2003 CSLI Publications

Chomsky Hierarchy



© 2003 CSLI Publications

Context-Free Grammar

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

© 2003 CSLI Publications

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

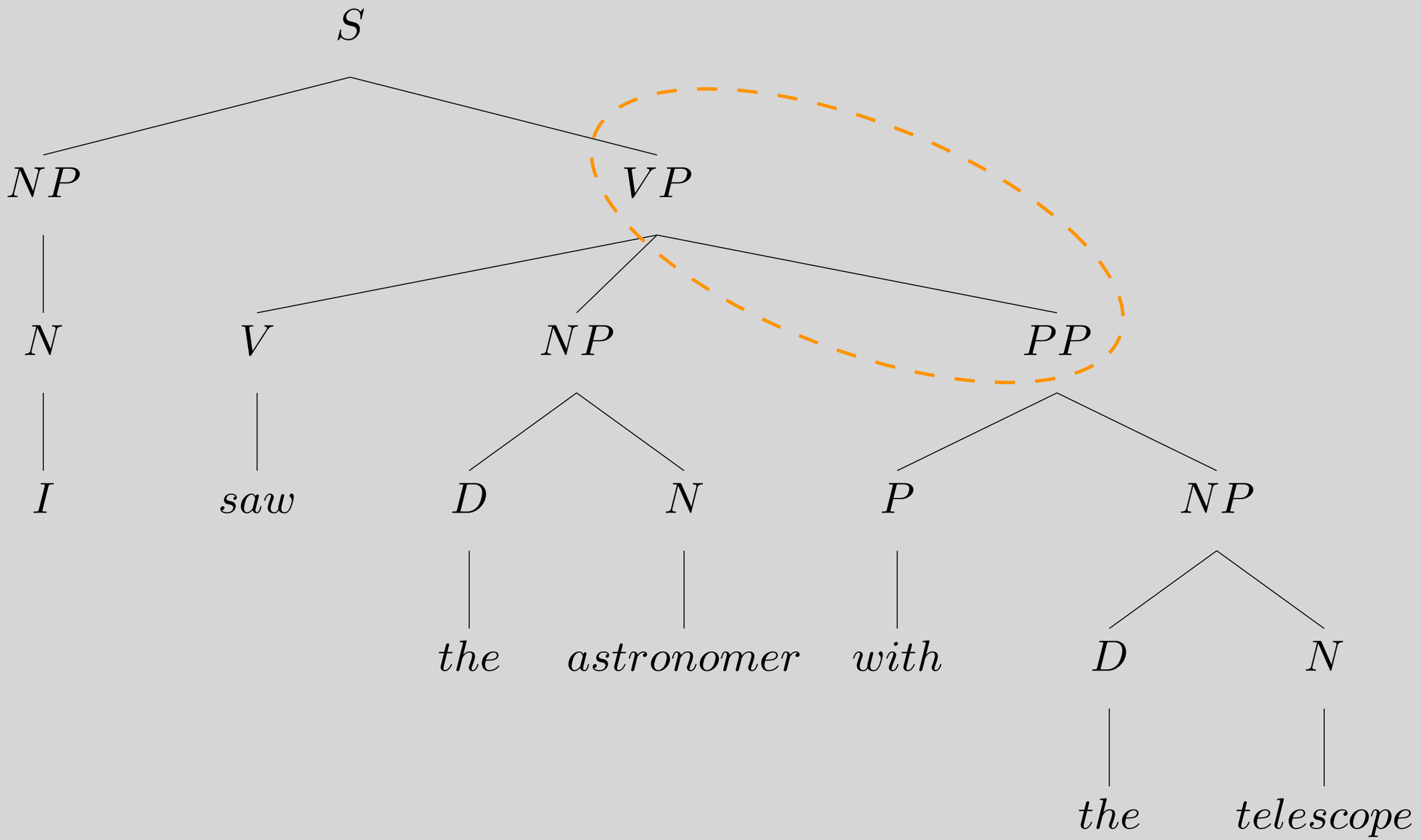
© 2003 CSLI Publications

Structural Ambiguity

I saw the astronomer with the telescope.

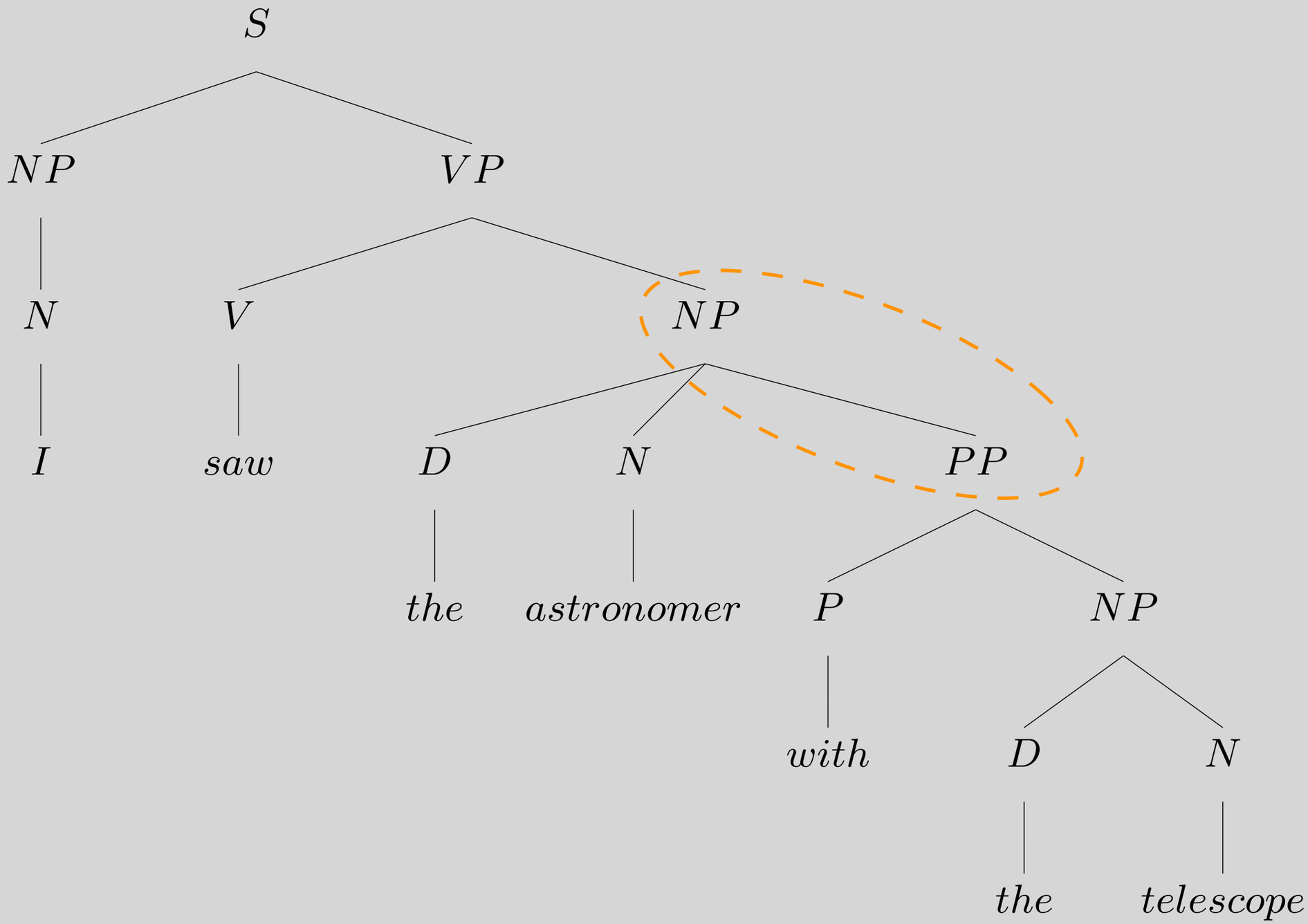
© 2003 CSLI Publications

Structure 1: PP under VP



© 2003 CSLI Publications

Structure 1: PP under NP



© 2003 CSLI Publications

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.

© 2003 CSLI Publications

General Types of Constituency Tests

- Distributional
- Intonational
- Semantic
- Psycholinguistic

... but they don't always agree.

© 2003 CSLI Publications

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.

© 2003 CSLI Publications

- 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.

© 2003 CSLI Publications

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.*

© 2003 CSLI Publications

A Trivial CFG

S $\longrightarrow$ NP VP

NP $\longrightarrow$ D N

VP $\longrightarrow$ V NP

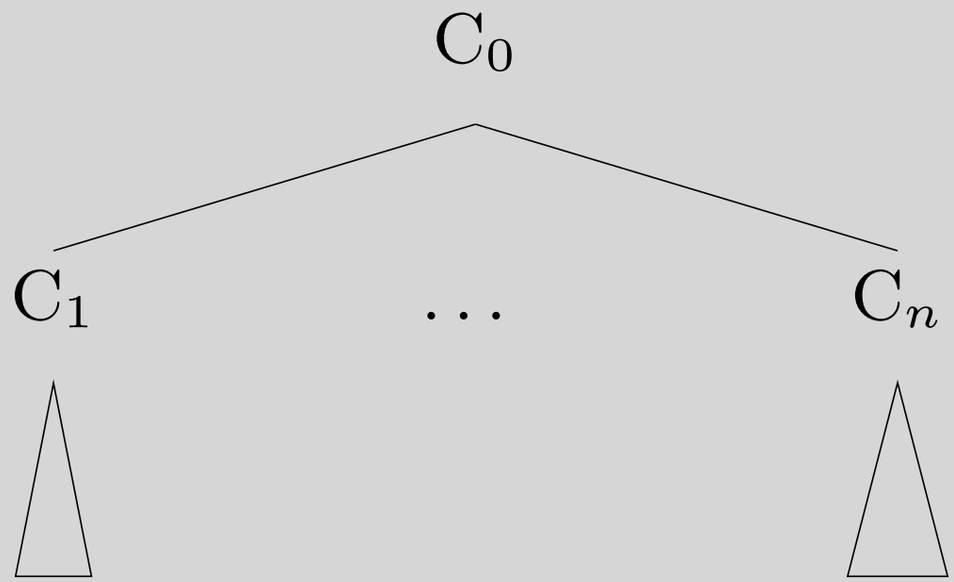
D: *the*

V: *chased*

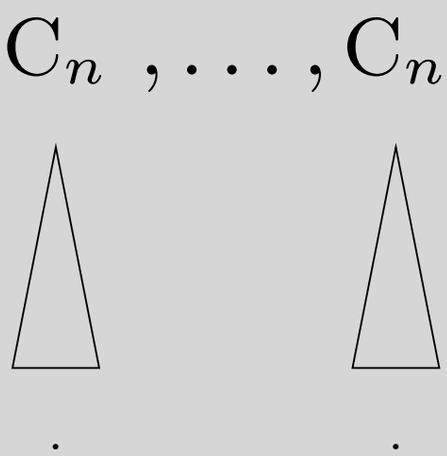
N: *dog, cat*

© 2003 CSLI Publications

Trees and Rules



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



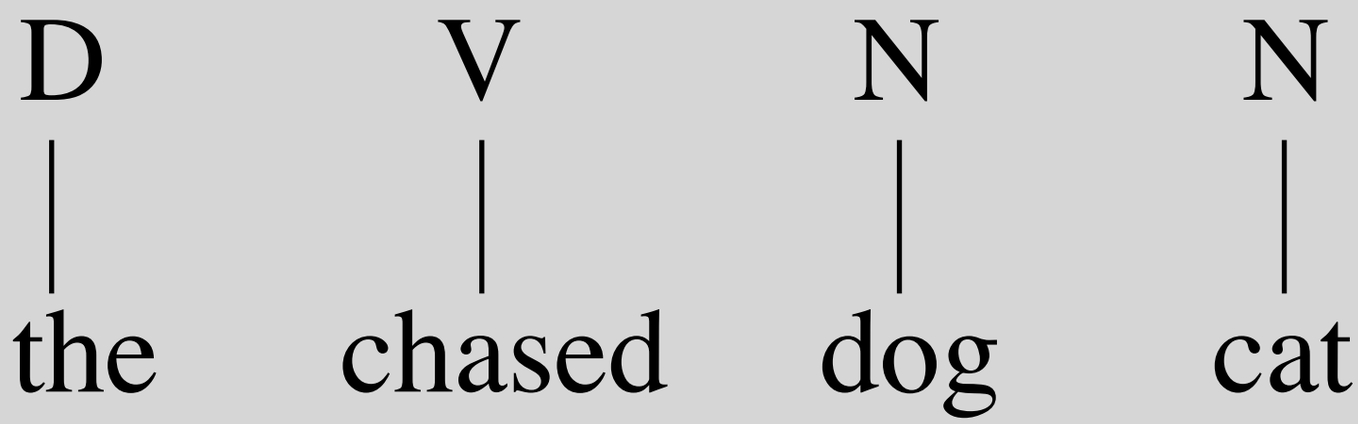
are well-formed trees, and

$C_0 \rightarrow C_1 \ldots Cn$ is a grammar rule.

© 2003 CSLI Publications

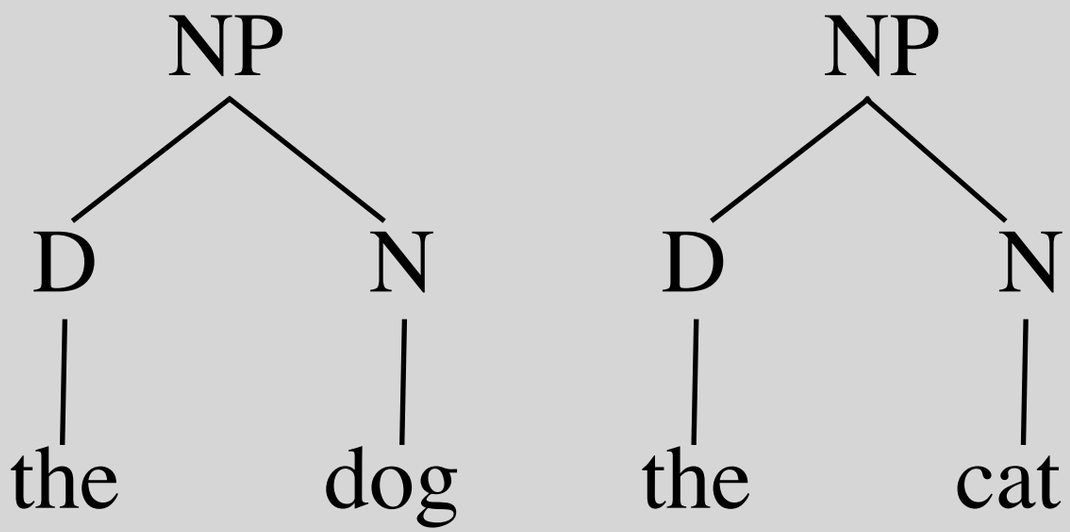
Bottom-up Tree Construction

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

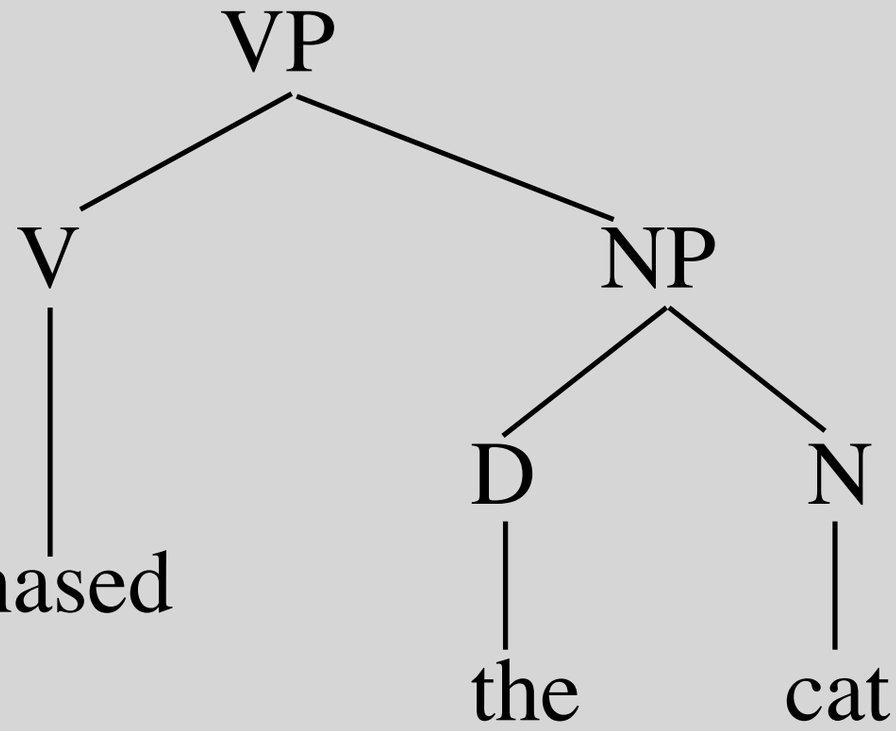


© 2003 CSLI Publications

NP ⟶ D N

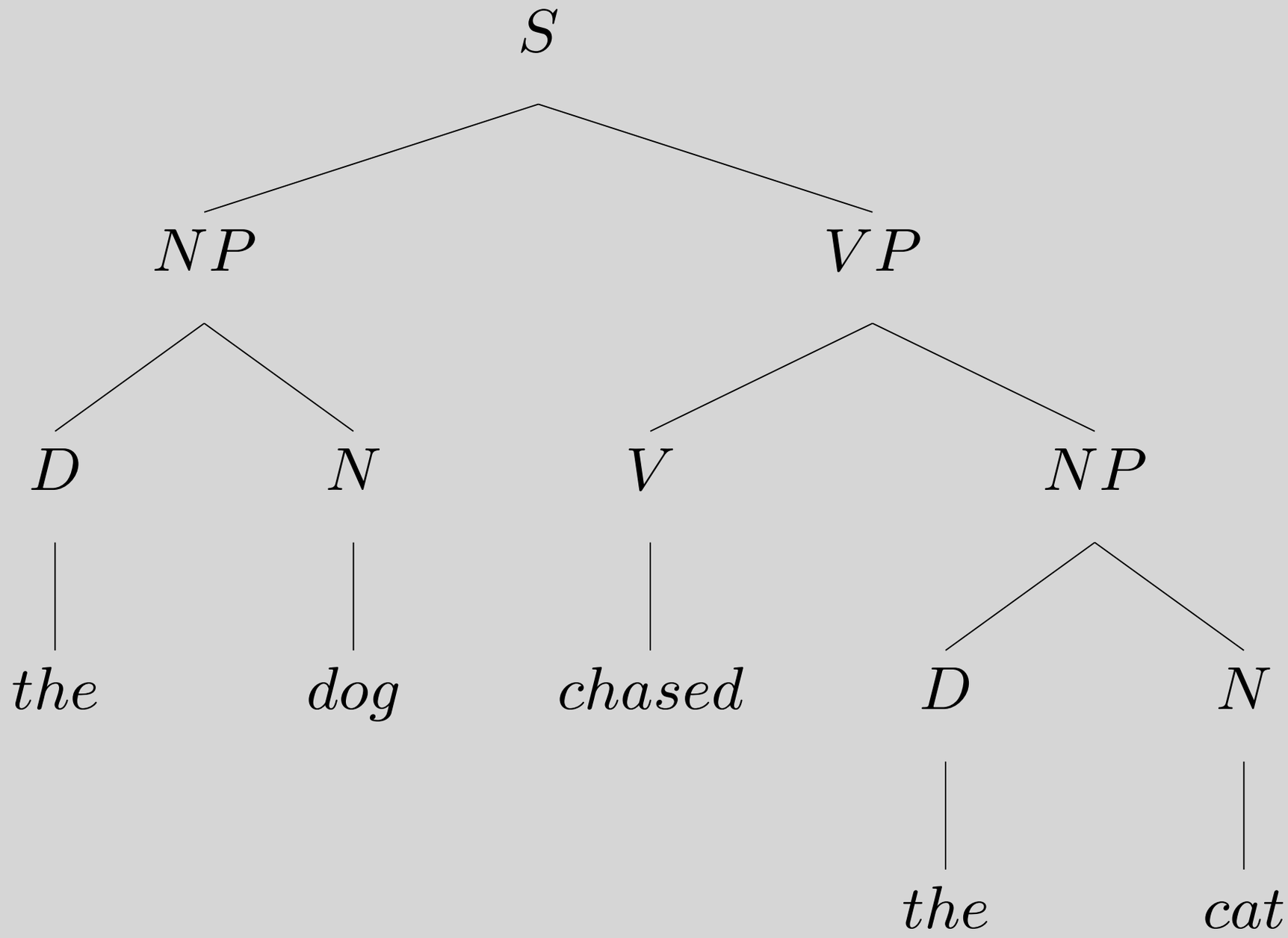


VP ⟶ V NP



© 2003 CSLI Publications

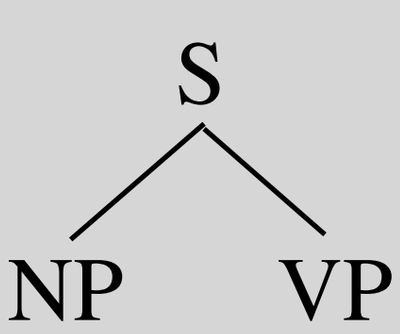
S ⟶ NP VP



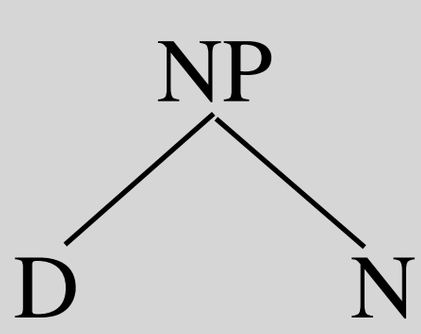
© 2003 CSLI Publications

Top-down Tree Construction

$S \longrightarrow NP\ \ VP$

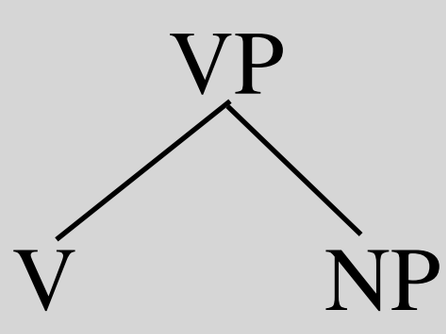


$NP \longrightarrow D\ \ N$

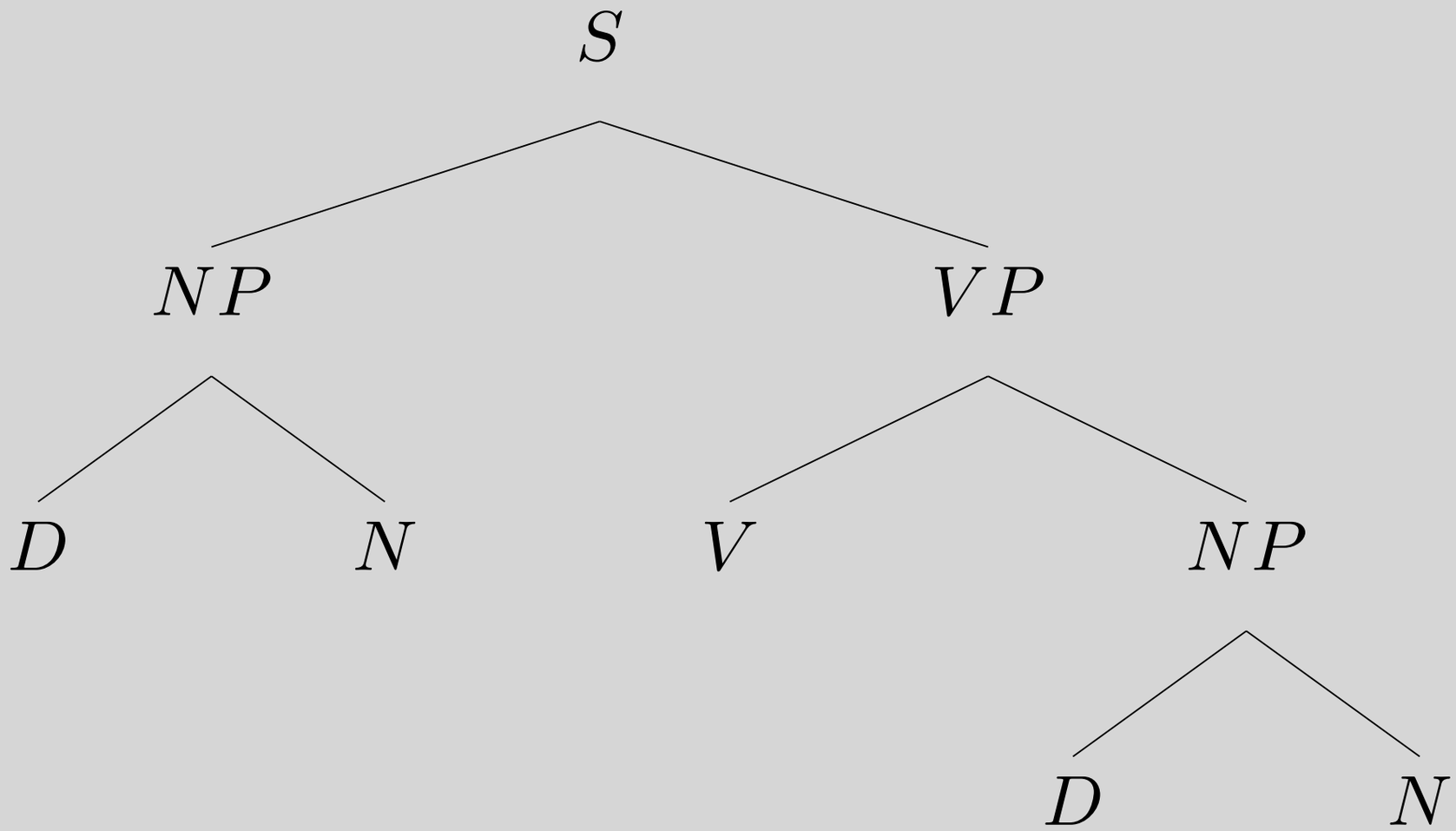


(twice)

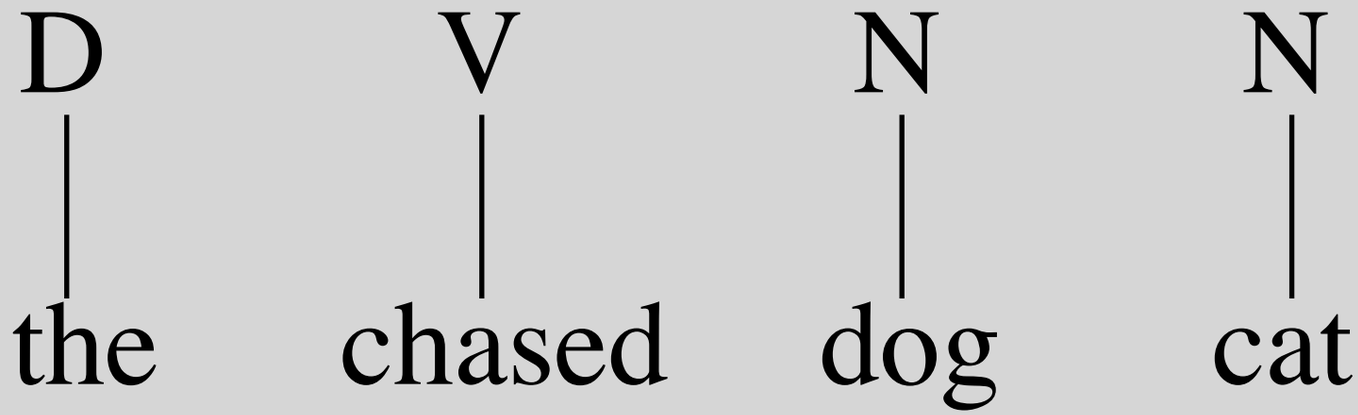
$VP \longrightarrow V\ \ NP$



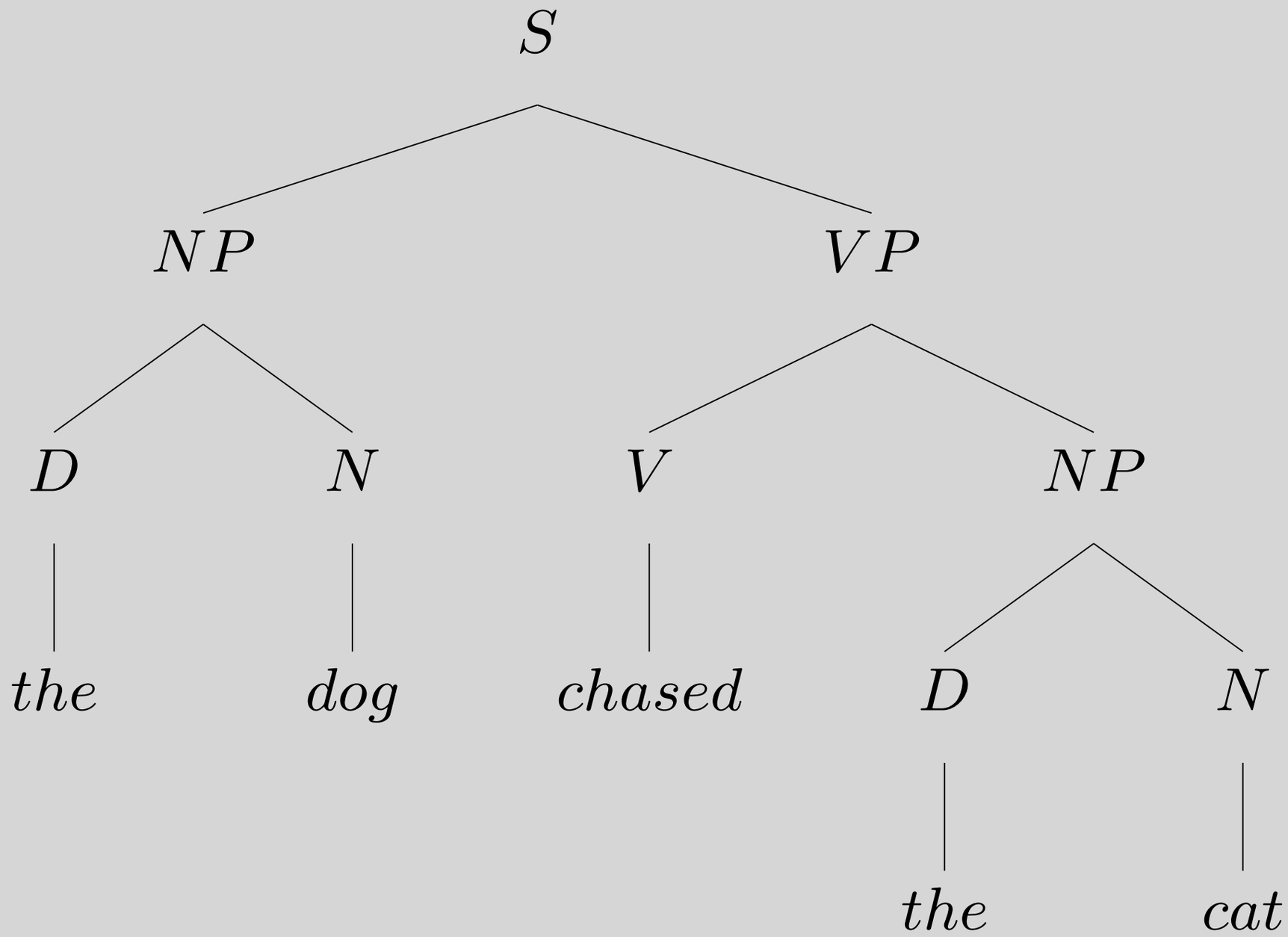
© 2003 CSLI Publications



© 2003 CSLI Publications



© 2003 CSLI Publications



© 2003 CSLI Publications

Weaknesses of CFG (atomic node labels)

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

$$\text{VP} \rightarrow \text{P NP}$$
$$\text{NP} \rightarrow \text{VP S}$$

- 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.

© 2003 CSLI Publications

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.

© 2003 CSLI Publications

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.

© 2003 CSLI Publications

Overview

- Syllabus
- Prescriptive/descriptive grammar; Competence/performance
- Some history
- Why study syntax?
- Two theories that won’t work
- Start on CFG

© 2003 CSLI Publications