Introduction to Syntax

CS 585, Fall 2018

Introduction to Natural Language Processing
http://people.cs.umass.edu/~miyyer/cs585/

Mohit Iyyer

College of Information and Computer Sciences
University of Massachusetts Amherst

(Slides and Lecture by Tu Vu)

some slides adapted from Michael Collins, Marine Carpuat, Wei Xu, and Rebecca Hwa

A Reminder!

- ## Project proposal
  - due on October 19, 2018 (*this Friday*) at 11:59 PM

- ## Midterm
  - will be held in this room on October 25, 2018 (*next Thursday*)
  - will cover text classification, word representations, language modeling, sequence labeling, and machine translation
  - will **not** cover today's lecture and next lectures
  - 20% multiple choice, 80% short answer/computational questions
  - 1-page "cheat sheet" allowed, must be hand-written

- ## Reading for the next lecture
  - JM 12

Overview

➢ An Introduction to Syntax

➢ Constituency

➢ Context-Free Grammars (CFGs)

➢ English Grammar in a Nutshell

Overview

➢ An Introduction to Syntax
- • Syntax
- • Syntax and Grammar
- • Syntax vs. Semantics
- • Syntax in NLP applications
- • Syntactic Structure

➢ Constituency

➢ Context-Free Grammars (CFGs)

➢ English Grammar in a Nutshell

Syntax

- *Sýntaxis* (setting out together or arrangement)
  - The ordering of words and how they group into phrases

Students Cook & Serve Grandparents

On Thursday, September 9, Gorman School hosted the first annual Grandparent's Day.

All Grandparents were invited to a school wide pancake breakfast. Upper grade students served as excellent chefs, as well as taking responsibility for serving the food and the clean up after-

Syntax

- *Sýntaxis* (setting out together or arrangement)
  - The ordering of words and how they group into phrases

Students Cook & Serve Grandparents

On Thursday, September 9, Gorman School hosted the first annual Grandparent's Day.

All Grandparents were invited to a school wide pancake breakfast. Upper grade students served as excellent chefs, as well as taking responsibility for serving the food and the clean up after-

- [[students][[cook and serve][grandparents]]]
- [[students][[cook][and][serve grandparents]]]

Syntax and Grammar

- Goal of syntactic theory
  - "explain how people combine words to form sentences and how children attain knowledge of sentence structure"

- Grammar
  - implicit knowledge of a native speaker
  - acquired without explicit instruction
  - minimally able to generate all and only the possible sentences of the language

Colin Phillips, "Syntax", 2003

Syntax vs. Semantics

"Colorless green ideas sleep furiously."
— Noam Chomsky
(1957)

Contrast with: "sleep green furiously ideas colorless"
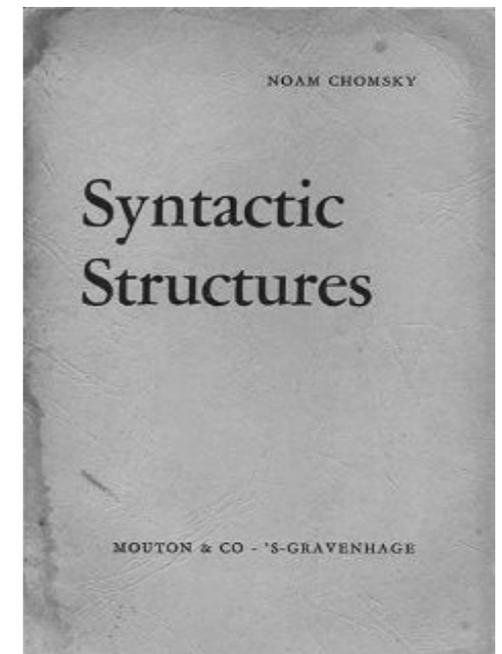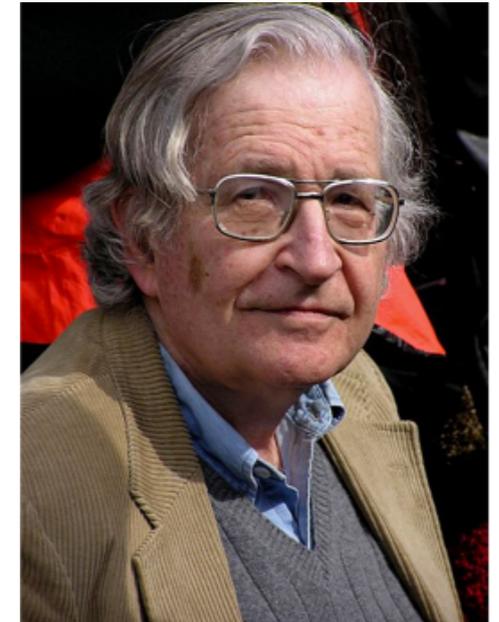
Syntax in NLP application

- Syntactic analysis is often a key component in many applications

  - Grammar checkers

  - Dialogue systems

  - Question answering

  - Information extraction

  - Machine translation

  - …

An Example: Machine Translation

- English word order is
  - *subject – verb – object*

- Japanese word order is
  - *subject – object – verb*

English: IBM bought Lotus

Japanese: *IBM Lotus bought*

English: Sources said that IBM bought Lotus yesterday

Japanese: *Sources yesterday IBM Lotus bought that said*

Another Example: Paraphrasing
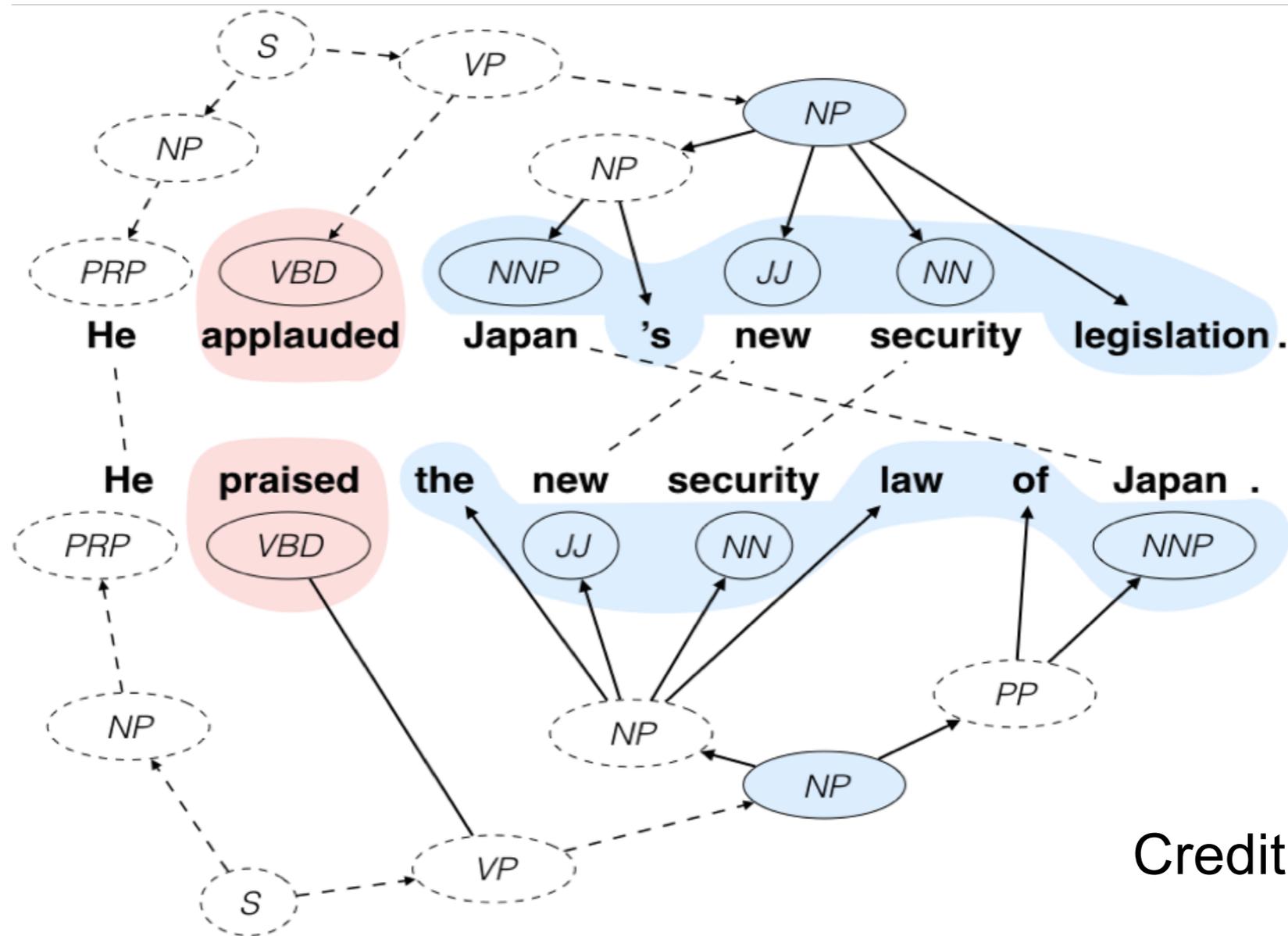
Credit: Wei Xu

Lexical	[VBD]	*applauded → praised*
Syntactic	[NP]	NNP 's JJ legislation → the JJ law of NNP

Syntactic Structure

- Constituency (phrase structure)
  - Phrase structure organizes words in nested constituents
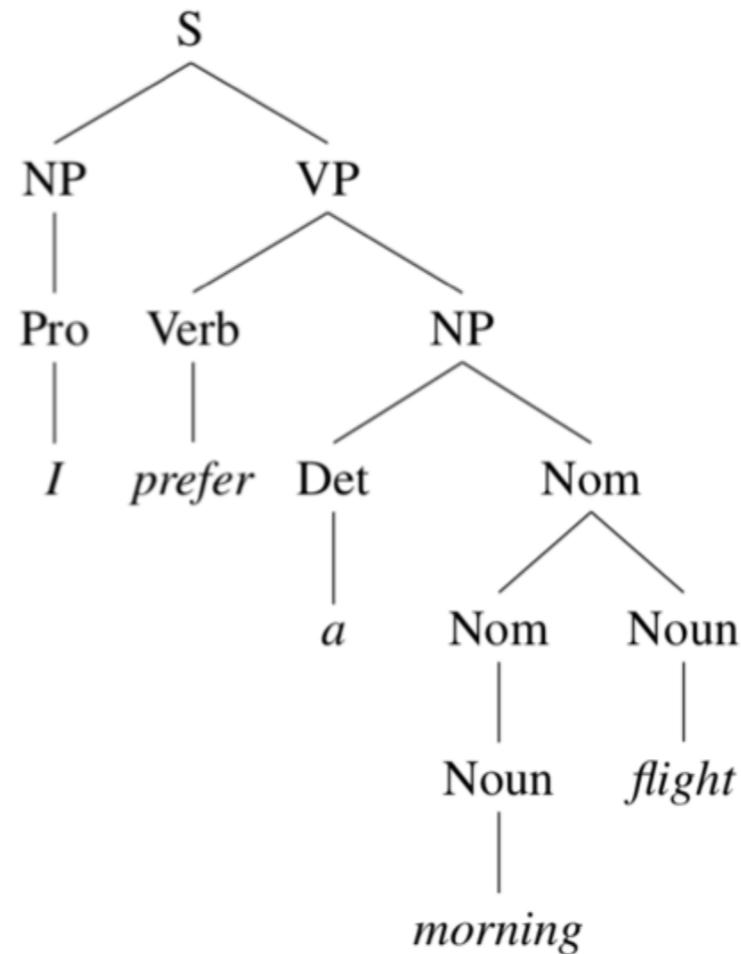
$[_S [_{NP} [_{Pro} \text{I}]] [_{VP} [_V \text{prefer}] [_{NP} [_{Det} \text{a}] [_{Nom} [_N \text{morning}] [_{Nom} [_N \text{flight}]]]]]]$

Syntactic Structure (cont.)

- Dependency structure
  - Shows which words depend on (modify or are arguments of) which on other words
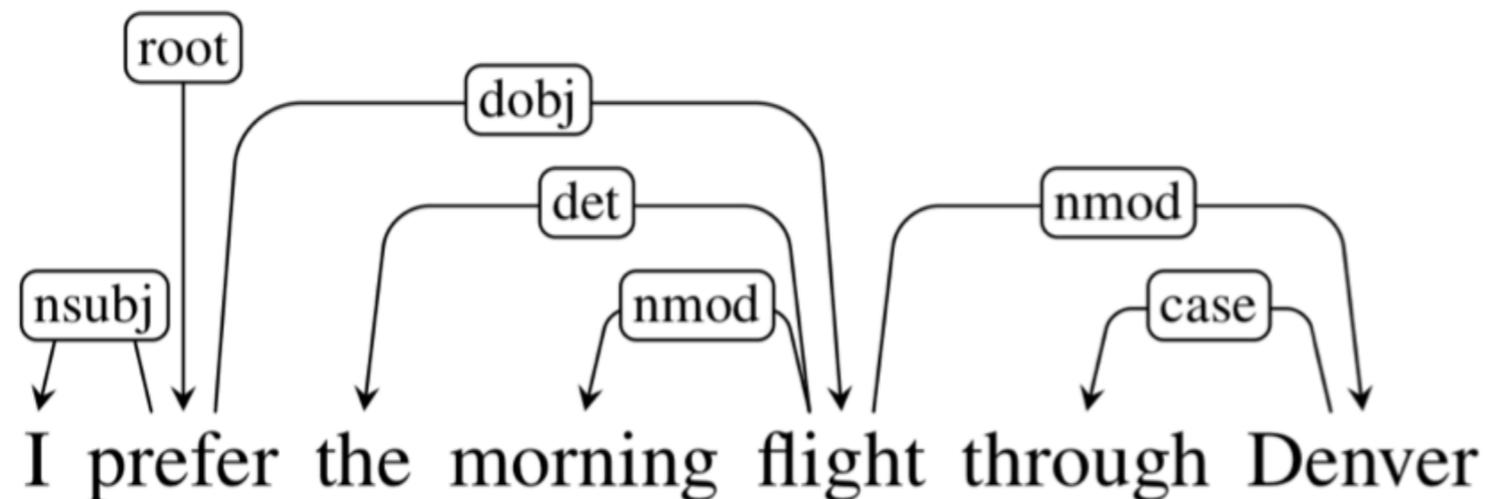
Overview

➢ An Introduction to Syntax

➢ Constituency

  • Constituency

  • Grammars and Constituency

➢ Context-Free Grammars (CFGs)

➢ English Grammar in a Nutshell

3

Constituency

- Basic idea

  - Groups of words behaving as single units, or constituents

- Constituents form coherent classes that behave similarly

  - With respect to their internal structure: e.g., at the core of a noun phrase is a noun

  - With respect to other constituents: e.g., noun phrases generally occur before verbs

Constituency (cont.)

- Examples
  - Noun phrases

Harry the Horse	a high-class spot such as Mindy's
the Broadway coppers	the reason he comes into the Hot Box
they	three parties from Brooklyn

  - Prepositional phrases

    on September seventeenth

Constituency (cont.)

- ## Examples
  - ### Noun phrases

Harry the Horse	a high-class spot such as Mindy's
the Broadway coppers	the reason he comes into the Hot Box
they	three parties from Brooklyn

  - ### Prepositional phrases

    on September seventeenth

    What evidence do we have that these groups of words behave as single units (or "form constituents")?

Constituency (cont.)

- One piece of evidence
  - They can all appear in similar syntactic environments, e.g., before a verb

> three parties from Brooklyn *arrive*…
> a high-class spot such as Mindy's *attracts*…
> the Broadway coppers *love*…
> they *sit*

Constituency (cont.)

- One piece of evidence
  - They can all appear in similar syntactic environments, e.g., before a verb

> three parties from Brooklyn *arrive*…
> a high-class spot such as Mindy's *attracts*…
> the Broadway coppers *love*…
> they *sit*

  - This is true for the entire phrase but not true of each of the individual words that make up the phrase

> *from *arrive*…   *as *attracts*…
> *the *is*…   *spot *sat*…

(*) marks fragments that are not grammatical English sentences

Constituency (cont.)

- Another piece of evidence
    - They can be placed in a number of different locations, e.g., at the beginning (preposed) or at the end (postposed) of a sentence

      On September seventeenth, I'd like to fly from Atlanta to Denver
      I'd like to fly *on September seventeenth* from Atlanta to Denver
      I'd like to fly from Atlanta to Denver *on September seventeenth*

Constituency (cont.)

- Another piece of evidence
  - They can be placed in a number of different locations, e.g., at the beginning (preposed) or at the end (postposed) of a sentence

    On September seventeenth, I'd like to fly from Atlanta to Denver

    I'd like to fly *on September seventeenth* from Atlanta to Denver

    I'd like to fly from Atlanta to Denver *on September seventeenth*

  - Again, the entire phrase can be placed differently, but the individual words that make up the phrase cannot be

    *On September, I'd like to fly <u>seventeenth</u> from Atlanta to Denver

    *<u>On</u> I'd like to fly <u>September seventeenth</u> from Atlanta to Denver

    *I'd like to fly <u>on September</u> from Atlanta to Denver <u>seventeenth</u>

Grammars and Constituency

- For a particular language:
  - What are the "right" set of constituents?
  - What rules govern how they combine?

Grammars and Constituency (cont.)

- For a particular language:
  - What are the "right" set of constituents?
  - What rules govern how they combine?

- Answer: not obvious and difficult
  - A significant part of developing a grammar involves discovering the inventory of constituents present in the language
  - That's why there are many different theories of grammar and competing analyses of the same data!

Grammars and Constituency (cont.)

- Some standard grammar formalisms:
  - Context-Free Grammar (CFG)
  - Lexical-Functional Grammar (LFG)
  - Head-Driven Phrase Structure Grammar (HPSG),
  - Tree-Adjoining Grammar (TAG),
  - Combinatory Categorial Grammar (CCG)

- While CFG emphasizes phrase-structure rules, the other approaches share the common theme of making better use of the lexicon

Overview

➢ An Introduction to Syntax

➢ Constituency

➢ Context-Free Grammars (CFGs)

- The Chomsky Hierarchy
- Context-Free Grammars (CFGs)
- Formal Definition of Context-Free Grammar
- Syntactic Parsing
- Examples of ambiguous structures

➢ English Grammar in a Nutshell

The Chomsky Hierarchy

- You've already seen one class of grammars: **regular expressions**
  - A pattern like ^[a-z][0-9]$ corresponds to a grammar which accepts (matches) some strings but not others.

- **Q**: Can regular languages define infinite languages?

The Chomsky Hierarchy (cont.)

- You've already seen one class of grammars: **regular expressions**

  - A pattern like ^[a-z][0-9]$ corresponds to a grammar which accepts (matches) some strings but not others.

- **Q**: Can regular languages define infinite languages?
  Yes, e.g. a*

The Chomsky Hierarchy (cont.)

- You've already seen one class of grammars: **regular expressions**
  - A pattern like ^[a-z][0-9]$ corresponds to a grammar which accepts (matches) some strings but not others.

- **Q**: Can regular languages define infinite languages?
  Yes, e.g. a*
- **Q**: Can regular languages define arbitrarily complex languages?

The Chomsky Hierarchy (cont.)

- You've already seen one class of grammars: **regular expressions**
  - A pattern like ^[a-z][0-9]$ corresponds to a grammar which accepts (matches) some strings but not others.

- **Q**: Can regular languages define infinite languages?

  Yes, e.g. a*

- **Q**: Can regular languages define arbitrarily complex languages?

  No. Cannot match all strings with matched parentheses or in $a^n b^n$ forms in general (recursion/arbitrary nesting).

The Chomsky Hierarchy (cont.)

- Hierarchy of classes of formal languages
  - One grammar is of greater generative power or complexity than another if it can define a language that other cannot define.
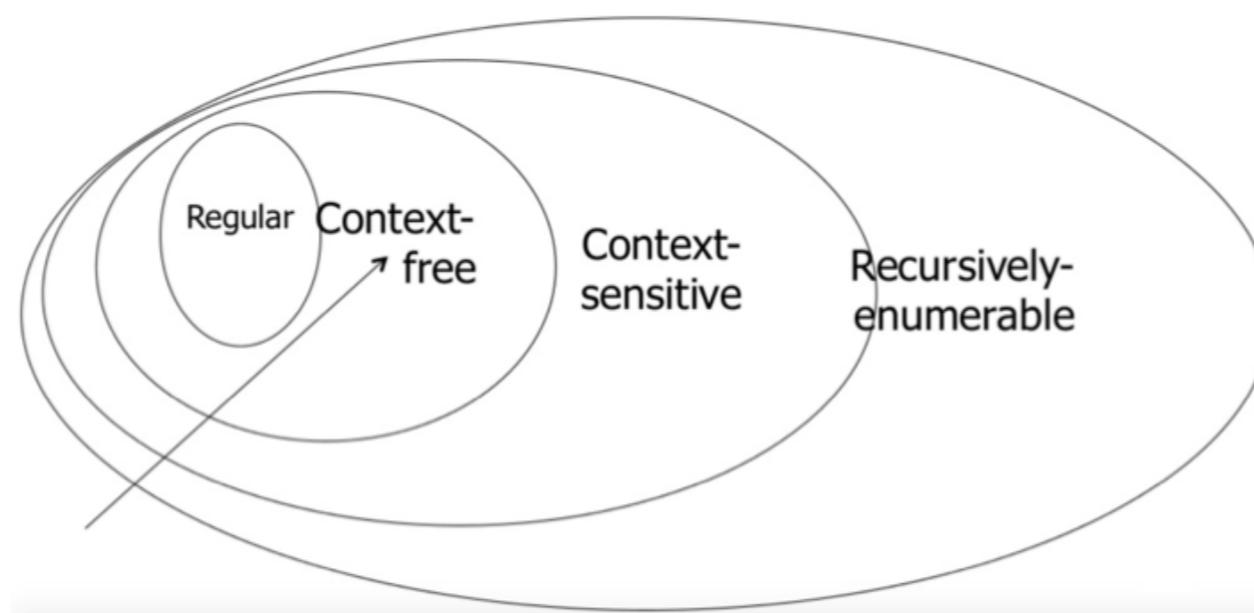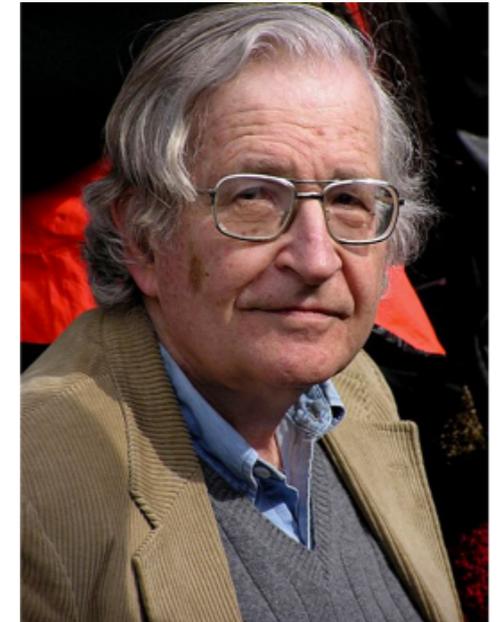
  - Context-free grammars are more powerful than regular grammars, and can account for much of the syntactic structure of English.

Context-Free Grammars

- **Context-Free Grammars** (CFGs)
    - Aka **Phrase Structure Grammars**
    - Aka **Backus-Naur Form (BNF)**
    - The most widely used formal system for modeling constituent structure in English and other natural languages
    - Good enough for most NLP applications!

    The idea of basing a grammar on constituent structure dates back to Wilhelm Wundt (1900) but was not formalized until Chomsky (1956) and, independently, Backus (1959)

- Consist of
    - Rules or productions
    - Terminals
    - Non-terminals

Context-Free Grammars (cont.)

- **Rules or productions**

  each rule can express

  - the ways that symbols of the language can be grouped and ordered together

$$NP \rightarrow Det\ Nominal$$
$$NP \rightarrow ProperNoun$$
$$Nominal \rightarrow Noun \mid Nominal\ Noun$$

- a **lexicon** of words and symbols

$$Det \rightarrow a$$
$$Det \rightarrow the$$
$$Noun \rightarrow flight$$

Context-Free Grammars (cont.)

- ## Terminals

  - Words in the language, e.g., "the", "flight"

- ## Non-terminals

  - The constituents in the language, e.g., noun phrases (NP), verb phrases (VP)

  - Express abstractions over terminals

Context-Free Grammars (cont.)

- A grammar with examples for each rule

Grammar Rules	Examples
$S \rightarrow NP\ VP$	I + want a morning flight
$NP \rightarrow Pronoun$	I
$\vert\ Proper\text{-}Noun$	Los Angeles
$\vert\ Det\ Nominal$	a + flight
$Nominal \rightarrow Nominal\ Noun$	morning + flight
$\vert\ Noun$	flights
$VP \rightarrow Verb$	do
$\vert\ Verb\ NP$	want + a flight
$\vert\ Verb\ NP\ PP$	leave + Boston + in the morning
$\vert\ Verb\ PP$	leaving + on Thursday
$PP \rightarrow Preposition\ NP$	from + Los Angeles

"|" indicates that a non-terminal has alternate possible expansions

Context-Free Grammars (cont.)

- A lexicon

Noun →	*flights* | *breeze* | *trip* | *morning*
Verb →	*is* | *prefer* | *like* | *need* | *want* | *fly*
Adjective →	*cheapest* | *non-stop* | *first* | *latest*
	| *other* | *direct*
Pronoun →	*me* | *I* | *you* | *it*
Proper-Noun →	*Alaska* | *Baltimore* | *Los Angeles*
	| *Chicago* | *United* | *American*
Determiner →	*the* | *a* | *an* | *this* | *these* | *that*
Preposition →	*from* | *to* | *on* | *near*
Conjunction →	*and* | *or* | *but*

"|" indicates that a non-terminal has alternate possible expansions

Context-Free Grammars (cont.)

- The form of a context-free rule $A \rightarrow \beta$
  - β is an ordered list of one or more terminals and non-terminals
  - A is a single non-terminal symbol expressing some cluster or generalization.
  - In the lexicon, β is a word and A is its lexical category, or POS

- Two view of a CFG
  - As a device for generating sentences
  - As a device for assigning a structure to a given sentence

Context-Free Grammars (cont.)

- CFG as a generator

    - We can read the rule A → β as "rewrite the symbol A on the left with string of symbols in β on the right".

An example

Rule expansions   **Rules used**

s

Grammar Rules

$S \rightarrow NP\ VP$

$NP \rightarrow Pronoun$
$\qquad |\ Proper\text{-}Noun$
$\qquad |\ Det\ Nominal$
$Nominal \rightarrow Nominal\ Noun$
$\qquad |\ Noun$

$VP \rightarrow Verb$
$\qquad |\ Verb\ NP$
$\qquad |\ Verb\ NP\ PP$
$\qquad |\ Verb\ PP$

$PP \rightarrow Preposition\ NP$

$Noun \rightarrow flight\ |\ breeze\ |\ trip\ |\ morning$
$Verb \rightarrow is\ |\ prefer\ |\ like\ |\ need\ |\ want\ |\ fly$
$Adjective \rightarrow cheapest\ |\ non\text{-}stop\ |\ first\ |\ latest$
$\qquad\qquad |\ other\ |\ direct$
$Pronoun \rightarrow me\ |\ I\ |\ you\ |\ it$
$Proper\text{-}Noun \rightarrow Alaska\ |\ Baltimore\ |\ Los\ Angeles$
$\qquad\qquad |\ Chicago\ |\ United\ |\ American$
$Determiner \rightarrow the\ |\ a\ |\ an\ |\ this\ |\ these\ |\ that$
$Preposition \rightarrow from\ |\ to\ |\ on\ |\ near$
$Conjunction \rightarrow and\ |\ or\ |\ but$

An example (cont.)

Rule expansions **Rules used**

S   S   → NP VP

Grammar Rules

→ S → NP VP

NP → Pronoun
     | Proper-Noun
     | Det Nominal
Nominal → Nominal Noun
         | Noun

VP → Verb
    | Verb NP
    | Verb NP PP
    | Verb PP

PP → Preposition NP

Noun → flight | breeze | trip | morning
Verb → is | prefer | like | need | want | fly
Adjective → cheapest | non-stop | first | latest
          | other | direct
Pronoun → me | I | you | it
Proper-Noun → Alaska | Baltimore | Los Angeles
            | Chicago | United | American
Determiner → the | a | an | this | these | that
Preposition → from | to | on | near
Conjunction → and | or | but

An example (cont.)

Rule expansions   **Rules used**

S   S   → NP VP

NP VP

Grammar Rules

$S \rightarrow NP\ VP$

$NP \rightarrow Pronoun$
$\mid Proper\text{-}Noun$
$\mid Det\ Nominal$
$Nominal \rightarrow Nominal\ Noun$
$\mid Noun$

$VP \rightarrow Verb$
$\mid Verb\ NP$
$\mid Verb\ NP\ PP$
$\mid Verb\ PP$

$PP \rightarrow Preposition\ NP$

$Noun \rightarrow flight \mid breeze \mid trip \mid morning$
$Verb \rightarrow is \mid prefer \mid like \mid need \mid want \mid fly$
$Adjective \rightarrow cheapest \mid non\text{-}stop \mid first \mid latest$
$\mid other \mid direct$
$Pronoun \rightarrow me \mid I \mid you \mid it$
$Proper\text{-}Noun \rightarrow Alaska \mid Baltimore \mid Los\ Angeles$
$\mid Chicago \mid United \mid American$
$Determiner \rightarrow the \mid a \mid an \mid this \mid these \mid that$
$Preposition \rightarrow from \mid to \mid on \mid near$
$Conjunction \rightarrow and \mid or \mid but$

An example (cont.)

Rule expansions

S

NP VP

Rules used

S → NP VP

NP → Pro

Grammar Rules

$$S \rightarrow NP\ VP$$

➡️
$$NP \rightarrow Pronoun$$
$$|\ Proper\text{-}Noun$$
$$|\ Det\ Nominal$$
$$Nominal \rightarrow Nominal\ Noun$$
$$|\ Noun$$

$$VP \rightarrow Verb$$
$$|\ Verb\ NP$$
$$|\ Verb\ NP\ PP$$
$$|\ Verb\ PP$$

$$PP \rightarrow Preposition\ NP$$

$$Noun \rightarrow flight\ |\ breeze\ |\ trip\ |\ morning$$
$$Verb \rightarrow is\ |\ prefer\ |\ like\ |\ need\ |\ want\ |\ fly$$
$$Adjective \rightarrow cheapest\ |\ non\text{-}stop\ |\ first\ |\ latest$$
$$|\ other\ |\ direct$$
$$Pronoun \rightarrow me\ |\ I\ |\ you\ |\ it$$
$$Proper\text{-}Noun \rightarrow Alaska\ |\ Baltimore\ |\ Los\ Angeles$$
$$|\ Chicago\ |\ United\ |\ American$$
$$Determiner \rightarrow the\ |\ a\ |\ an\ |\ this\ |\ these\ |\ that$$
$$Preposition \rightarrow from\ |\ to\ |\ on\ |\ near$$
$$Conjunction \rightarrow and\ |\ or\ |\ but$$

An example (cont.)

Rule expansions

S

NP VP

Pro VP

Rules used

S → NP VP

NP → Pro

Grammar Rules

$S \rightarrow NP\ VP$

$NP \rightarrow Pronoun$
$\quad\ |\ Proper\text{-}Noun$
$\quad\ |\ Det\ Nominal$
$Nominal \rightarrow Nominal\ Noun$
$\quad\ |\ Noun$

$VP \rightarrow Verb$
$\quad\ |\ Verb\ NP$
$\quad\ |\ Verb\ NP\ PP$
$\quad\ |\ Verb\ PP$

$PP \rightarrow Preposition\ NP$

$Noun \rightarrow flight\ |\ breeze\ |\ trip\ |\ morning$
$Verb \rightarrow is\ |\ prefer\ |\ like\ |\ need\ |\ want\ |\ fly$
$Adjective \rightarrow cheapest\ |\ non\text{-}stop\ |\ first\ |\ latest$
$\quad |\ other\ |\ direct$
$Pronoun \rightarrow me\ |\ I\ |\ you\ |\ it$
$Proper\text{-}Noun \rightarrow Alaska\ |\ Baltimore\ |\ Los\ Angeles$
$\quad |\ Chicago\ |\ United\ |\ American$
$Determiner \rightarrow the\ |\ a\ |\ an\ |\ this\ |\ these\ |\ that$
$Preposition \rightarrow from\ |\ to\ |\ on\ |\ near$
$Conjunction \rightarrow and\ |\ or\ |\ but$

An example (cont.)

Rule expansions

S

NP VP

Pro VP

Rules used

S   → NP VP

NP → Pro

VP → Verb NP

Grammar Rules

$$S \rightarrow NP\ VP$$

$$NP \rightarrow Pronoun$$
$$\mid Proper\text{-}Noun$$
$$\mid Det\ Nominal$$

$$Nominal \rightarrow Nominal\ Noun$$
$$\mid Noun$$

$$VP \rightarrow Verb$$
$$\mid Verb\ NP$$
$$\mid Verb\ NP\ PP$$
$$\mid Verb\ PP$$

$$PP \rightarrow Preposition\ NP$$

$$Noun \rightarrow flight \mid breeze \mid trip \mid morning$$
$$Verb \rightarrow is \mid prefer \mid like \mid need \mid want \mid fly$$
$$Adjective \rightarrow cheapest \mid non\text{-}stop \mid first \mid latest$$
$$\mid other \mid direct$$
$$Pronoun \rightarrow me \mid I \mid you \mid it$$
$$Proper\text{-}Noun \rightarrow Alaska \mid Baltimore \mid Los\ Angeles$$
$$\mid Chicago \mid United \mid American$$
$$Determiner \rightarrow the \mid a \mid an \mid this \mid these \mid that$$
$$Preposition \rightarrow from \mid to \mid on \mid near$$
$$Conjunction \rightarrow and \mid or \mid but$$

An example (cont.)

Rule expansions

S

NP VP

Pro VP

Pro Verb NP

Rules used

S → NP VP

NP → Pro

VP → Verb NP

Grammar Rules

$$S \rightarrow NP\ VP$$

$$NP \rightarrow Pronoun$$
$$| \quad Proper\text{-}Noun$$
$$| \quad Det\ Nominal$$
$$Nominal \rightarrow Nominal\ Noun$$
$$| \quad Noun$$

$$VP \rightarrow Verb$$
$$| \quad Verb\ NP$$
$$| \quad Verb\ NP\ PP$$
$$| \quad Verb\ PP$$

$$PP \rightarrow Preposition\ NP$$

$$Noun \rightarrow flight\ |\ breeze\ |\ trip\ |\ morning$$
$$Verb \rightarrow is\ |\ prefer\ |\ like\ |\ need\ |\ want\ |\ fly$$
$$Adjective \rightarrow cheapest\ |\ non\text{-}stop\ |\ first\ |\ latest$$
$$|\ other\ |\ direct$$
$$Pronoun \rightarrow me\ |\ I\ |\ you\ |\ it$$
$$Proper\text{-}Noun \rightarrow Alaska\ |\ Baltimore\ |\ Los\ Angeles$$
$$|\ Chicago\ |\ United\ |\ American$$
$$Determiner \rightarrow the\ |\ a\ |\ an\ |\ this\ |\ these\ |\ that$$
$$Preposition \rightarrow from\ |\ to\ |\ on\ |\ near$$
$$Conjunction \rightarrow and\ |\ or\ |\ but$$

An example (cont.)

Rule expansions

S

NP VP

Pro VP

Pro Verb NP

Rules used

S → NP VP

NP → Pro

VP → Verb NP

Pro → *I*

Grammar Rules

$S \rightarrow NP\ VP$

$NP \rightarrow$ *Pronoun*
$\quad\quad |$ *Proper-Noun*
$\quad\quad |$ *Det Nominal*

Nominal $\rightarrow$ *Nominal Noun*
$\quad\quad |$ *Noun*

$VP \rightarrow$ *Verb*
$\quad\quad |$ *Verb NP*
$\quad\quad |$ *Verb NP PP*
$\quad\quad |$ *Verb PP*

$PP \rightarrow$ *Preposition NP*

Noun → *flight* | *breeze* | *trip* | *morning*
Verb → *is* | *prefer* | *like* | *need* | *want* | *fly*
Adjective → *cheapest* | *non-stop* | *first* | *latest*
$\quad\quad$ | *other* | *direct*
Pronoun → *me* | *I* | *you* | *it*
Proper-Noun → *Alaska* | *Baltimore* | *Los Angeles*
$\quad\quad$ | *Chicago* | *United* | *American*
Determiner → *the* | *a* | *an* | *this* | *these* | *that*
Preposition → *from* | *to* | *on* | *near*
Conjunction → *and* | *or* | *but*

An example (cont.)

Rule expansions

S

NP VP

Pro VP

Pro Verb NP

I Verb NP

Rules used

S   → NP VP

NP → Pro

VP → Verb NP

Pro → *I*

Grammar Rules

S	→	NP VP
NP	→	Pronoun
	|	Proper-Noun
	|	Det Nominal
Nominal	→	Nominal Noun
	|	Noun
VP	→	Verb
	|	Verb NP
	|	Verb NP PP
	|	Verb PP
PP	→	Preposition NP

Noun → flight | breeze | trip | morning
Verb → is | prefer | like | need | want | fly
Adjective → cheapest | non-stop | first | latest
| other | direct
Pronoun → me | I | you | it
Proper-Noun → Alaska | Baltimore | Los Angeles
| Chicago | United | American
Determiner → the | a | an | this | these | that
Preposition → from | to | on | near
Conjunction → and | or | but

An example (cont.)

Rule expansions

S

NP VP

Pro VP

Pro Verb NP

I Verb NP

Rules used

S   → NP VP

NP → Pro

VP → Verb NP

Pro → *I*

Verb → prefer

Grammar Rules

$S \rightarrow NP\ VP$

$NP \rightarrow Pronoun$
$\quad |\ Proper\text{-}Noun$
$\quad |\ Det\ Nominal$

$Nominal \rightarrow Nominal\ Noun$
$\quad |\ Noun$

$VP \rightarrow Verb$
$\quad |\ Verb\ NP$
$\quad |\ Verb\ NP\ PP$
$\quad |\ Verb\ PP$

$PP \rightarrow Preposition\ NP$

$Noun \rightarrow flight\ |\ breeze\ |\ trip\ |\ morning$
$Verb \rightarrow is\ |\ prefer\ |\ like\ |\ need\ |\ want\ |\ fly$
$Adjective \rightarrow cheapest\ |\ non\text{-}stop\ |\ first\ |\ latest$
$\quad |\ other\ |\ direct$
$Pronoun \rightarrow me\ |\ I\ |\ you\ |\ it$
$Proper\text{-}Noun \rightarrow Alaska\ |\ Baltimore\ |\ Los\ Angeles$
$\quad |\ Chicago\ |\ United\ |\ American$
$Determiner \rightarrow the\ |\ a\ |\ an\ |\ this\ |\ these\ |\ that$
$Preposition \rightarrow from\ |\ to\ |\ on\ |\ near$
$Conjunction \rightarrow and\ |\ or\ |\ but$

An example (cont.)

Rule expansions

S

NP VP

Pro VP

Pro Verb NP

I Verb NP

I prefer NP

Rules used

S   → NP VP

NP → Pro

VP → Verb NP

Pro → *I*

Verb → prefer

Grammar Rules

$S \rightarrow NP\ VP$

$NP \rightarrow Pronoun$
$\quad\quad | \quad Proper\text{-}Noun$
$\quad\quad | \quad Det\ Nominal$
$Nominal \rightarrow Nominal\ Noun$
$\quad\quad | \quad Noun$

$VP \rightarrow Verb$
$\quad\quad | \quad Verb\ NP$
$\quad\quad | \quad Verb\ NP\ PP$
$\quad\quad | \quad Verb\ PP$

$PP \rightarrow Preposition\ NP$

$Noun \rightarrow flight\ |\ breeze\ |\ trip\ |\ morning$
$Verb \rightarrow is\ |\ prefer\ |\ like\ |\ need\ |\ want\ |\ fly$
$Adjective \rightarrow cheapest\ |\ non\text{-}stop\ |\ first\ |\ latest$
$\quad\quad |\ other\ |\ direct$
$Pronoun \rightarrow me\ |\ I\ |\ you\ |\ it$
$Proper\text{-}Noun \rightarrow Alaska\ |\ Baltimore\ |\ Los\ Angeles$
$\quad\quad |\ Chicago\ |\ United\ |\ American$
$Determiner \rightarrow the\ |\ a\ |\ an\ |\ this\ |\ these\ |\ that$
$Preposition \rightarrow from\ |\ to\ |\ on\ |\ near$
$Conjunction \rightarrow and\ |\ or\ |\ but$

An example (cont.)

Rule expansions

S

NP VP

Pro VP

Pro Verb NP

I Verb NP

I prefer NP

Rules used

S   → NP VP

NP → Pro

VP → Verb NP

Pro → *I*

Verb → prefer

NP → Det Nom

Grammar Rules

$S \rightarrow NP\ VP$

$NP \rightarrow Pronoun$
  $|\ Proper\text{-}Noun$
  $|\ Det\ Nominal$
$Nominal \rightarrow Nominal\ Noun$
  $|\ Noun$

$VP \rightarrow Verb$
  $|\ Verb\ NP$
  $|\ Verb\ NP\ PP$
  $|\ Verb\ PP$

$PP \rightarrow Preposition\ NP$

$Noun \rightarrow flight\ |\ breeze\ |\ trip\ |\ morning$
$Verb \rightarrow is\ |\ prefer\ |\ like\ |\ need\ |\ want\ |\ fly$
$Adjective \rightarrow cheapest\ |\ non\text{-}stop\ |\ first\ |\ latest$
  $|\ other\ |\ direct$
$Pronoun \rightarrow me\ |\ I\ |\ you\ |\ it$
$Proper\text{-}Noun \rightarrow Alaska\ |\ Baltimore\ |\ Los\ Angeles$
  $|\ Chicago\ |\ United\ |\ American$
$Determiner \rightarrow the\ |\ a\ |\ an\ |\ this\ |\ these\ |\ that$
$Preposition \rightarrow from\ |\ to\ |\ on\ |\ near$
$Conjunction \rightarrow and\ |\ or\ |\ but$

An example (cont.)

Rule expansions

S

NP VP

Pro VP

Pro Verb NP

I Verb NP

I prefer NP

I prefer Det Nom

Rules used

S   → NP VP

NP → Pro

VP → Verb NP

Pro → *I*

Verb → prefer

NP → Det Nom

Grammar Rules

S	→	NP VP
NP	→	Pronoun
	|	Proper-Noun
	|	Det Nominal
Nominal	→	Nominal Noun
	|	Noun
VP	→	Verb
	|	Verb NP
	|	Verb NP PP
	|	Verb PP
PP	→	Preposition NP

Noun → *flight* | *breeze* | *trip* | *morning*
Verb → *is* | *prefer* | *like* | *need* | *want* | *fly*
Adjective → *cheapest* | *non-stop* | *first* | *latest*
   | *other* | *direct*
Pronoun → *me* | *I* | *you* | *it*
Proper-Noun → *Alaska* | *Baltimore* | *Los Angeles*
   | *Chicago* | *United* | *American*
Determiner → *the* | *a* | *an* | *this* | *these* | *that*
Preposition → *from* | *to* | *on* | *near*
Conjunction → *and* | *or* | *but*

An example (cont.)

Rule expansions

S

NP VP

Pro VP

Pro Verb NP

I Verb NP

I prefer NP

I prefer Det Nom

Rules used

S   → NP VP

NP → Pro

VP → Verb NP

Pro → *I*

Verb → prefer

NP → Det Nom

Det → *a*

Grammar Rules

$S \rightarrow NP\ VP$

$NP \rightarrow Pronoun$
$| \quad Proper\text{-}Noun$
$| \quad Det\ Nominal$
$Nominal \rightarrow Nominal\ Noun$
$| \quad Noun$

$VP \rightarrow Verb$
$| \quad Verb\ NP$
$| \quad Verb\ NP\ PP$
$| \quad Verb\ PP$

$PP \rightarrow Preposition\ NP$

$Noun \rightarrow flight \mid breeze \mid trip \mid morning$
$Verb \rightarrow is \mid prefer \mid like \mid need \mid want \mid fly$
$Adjective \rightarrow cheapest \mid non\text{-}stop \mid first \mid latest$
$\mid other \mid direct$
$Pronoun \rightarrow me \mid I \mid you \mid it$
$Proper\text{-}Noun \rightarrow Alaska \mid Baltimore \mid Los\ Angeles$
$\mid Chicago \mid United \mid American$
$Determiner \rightarrow the \mid a \mid an \mid this \mid these \mid that$
$Preposition \rightarrow from \mid to \mid on \mid near$
$Conjunction \rightarrow and \mid or \mid but$

An example (cont.)

Rule expansions

S

NP VP

Pro VP

Pro Verb NP

I Verb NP

I prefer NP

I prefer Det Nom

I prefer a Nom

Rules used

S   → NP VP

NP → Pro

VP → Verb NP

Pro → *I*

Verb → prefer

NP → Det Nom

Det → *a*

Grammar Rules

$S \rightarrow NP\ VP$

$NP \rightarrow Pronoun$
$\quad | \quad Proper\text{-}Noun$
$\quad | \quad Det\ Nominal$
$Nominal \rightarrow Nominal\ Noun$
$\quad | \quad Noun$

$VP \rightarrow Verb$
$\quad | \quad Verb\ NP$
$\quad | \quad Verb\ NP\ PP$
$\quad | \quad Verb\ PP$

$PP \rightarrow Preposition\ NP$

$Noun \rightarrow flight\ |\ breeze\ |\ trip\ |\ morning$
$Verb \rightarrow is\ |\ prefer\ |\ like\ |\ need\ |\ want\ |\ fly$
$Adjective \rightarrow cheapest\ |\ non\text{-}stop\ |\ first\ |\ latest$
$\quad |\ other\ |\ direct$
$Pronoun \rightarrow me\ |\ I\ |\ you\ |\ it$
$Proper\text{-}Noun \rightarrow Alaska\ |\ Baltimore\ |\ Los\ Angeles$
$\quad |\ Chicago\ |\ United\ |\ American$
$Determiner \rightarrow the\ |\ a\ |\ an\ |\ this\ |\ these\ |\ that$
$Preposition \rightarrow from\ |\ to\ |\ on\ |\ near$
$Conjunction \rightarrow and\ |\ or\ |\ but$

An example (cont.)

Rule expansions

S

NP VP

Pro VP

Pro Verb NP

I Verb NP

I prefer NP

I prefer Det Nom

I prefer a Nom

Rules used

S   → NP VP

NP → Pro

VP → Verb NP

Pro → *I*

Verb → prefer

NP → Det Nom

Det → *a*

Nom → Nom Noun

Grammar Rules

$S \rightarrow NP\ VP$

$NP \rightarrow$ *Pronoun*
| *Proper-Noun*
| *Det Nominal*

→ *Nominal* → *Nominal Noun*
| *Noun*

$VP \rightarrow$ *Verb*
| *Verb NP*
| *Verb NP PP*
| *Verb PP*

$PP \rightarrow$ *Preposition NP*

Noun → *flight* | *breeze* | *trip* | *morning*
Verb → *is* | *prefer* | *like* | *need* | *want* | *fly*
Adjective → *cheapest* | *non-stop* | *first* | *latest*
| *other* | *direct*
Pronoun → *me* | *I* | *you* | *it*
Proper-Noun → *Alaska* | *Baltimore* | *Los Angeles*
| *Chicago* | *United* | *American*
Determiner → *the* | *a* | *an* | *this* | *these* | *that*
Preposition → *from* | *to* | *on* | *near*
Conjunction → *and* | *or* | *but*

An example (cont.)

Rule expansions

S

NP VP

Pro VP

Pro Verb NP

I Verb NP

I prefer NP

I prefer Det Nom

I prefer a Nom

I prefer a Nom Noun

Rules used

S   → NP VP

NP → Pro

VP → Verb NP

Pro → *I*

Verb → prefer

NP → Det Nom

Det → *a*

Nom → Nom Noun

Grammar Rules

$$S \rightarrow NP\ VP$$

$$NP \rightarrow Pronoun$$
$$| \quad Proper\text{-}Noun$$
$$| \quad Det\ Nominal$$
$$Nominal \rightarrow Nominal\ Noun$$
$$| \quad Noun$$

$$VP \rightarrow Verb$$
$$| \quad Verb\ NP$$
$$| \quad Verb\ NP\ PP$$
$$| \quad Verb\ PP$$

$$PP \rightarrow Preposition\ NP$$

$$Noun \rightarrow flight\ |\ breeze\ |\ trip\ |\ morning$$
$$Verb \rightarrow is\ |\ prefer\ |\ like\ |\ need\ |\ want\ |\ fly$$
$$Adjective \rightarrow cheapest\ |\ non\text{-}stop\ |\ first\ |\ latest$$
$$|\ other\ |\ direct$$
$$Pronoun \rightarrow me\ |\ I\ |\ you\ |\ it$$
$$Proper\text{-}Noun \rightarrow Alaska\ |\ Baltimore\ |\ Los\ Angeles$$
$$|\ Chicago\ |\ United\ |\ American$$
$$Determiner \rightarrow the\ |\ a\ |\ an\ |\ this\ |\ these\ |\ that$$
$$Preposition \rightarrow from\ |\ to\ |\ on\ |\ near$$
$$Conjunction \rightarrow and\ |\ or\ |\ but$$

An example (cont.)

Rule expansions

S

NP VP

Pro VP

Pro Verb NP

I Verb NP

I prefer NP

I prefer Det Nom

I prefer a Nom

I prefer a Nom Noun

Rules used

S   → NP VP

NP → Pro

VP → Verb NP

Pro → *I*

Verb → prefer

NP → Det Nom

Det → *a*

Nom → Nom Noun

Nom → Noun

Grammar Rules

$S \rightarrow NP\ VP$

$NP \rightarrow Pronoun$
$\quad | \quad Proper\text{-}Noun$
$\quad | \quad Det\ Nominal$

➡ $Nominal \rightarrow Nominal\ Noun$
$\quad | \quad Noun$

$VP \rightarrow Verb$
$\quad | \quad Verb\ NP$
$\quad | \quad Verb\ NP\ PP$
$\quad | \quad Verb\ PP$

$PP \rightarrow Preposition\ NP$

$Noun \rightarrow flight\ |\ breeze\ |\ trip\ |\ morning$
$Verb \rightarrow is\ |\ prefer\ |\ like\ |\ need\ |\ want\ |\ fly$
$Adjective \rightarrow cheapest\ |\ non\text{-}stop\ |\ first\ |\ latest$
$\quad |\ other\ |\ direct$
$Pronoun \rightarrow me\ |\ I\ |\ you\ |\ it$
$Proper\text{-}Noun \rightarrow Alaska\ |\ Baltimore\ |\ Los\ Angeles$
$\quad |\ Chicago\ |\ United\ |\ American$
$Determiner \rightarrow the\ |\ a\ |\ an\ |\ this\ |\ these\ |\ that$
$Preposition \rightarrow from\ |\ to\ |\ on\ |\ near$
$Conjunction \rightarrow and\ |\ or\ |\ but$

An example (cont.)

Rule expansions

S

NP VP

Pro VP

Pro Verb NP

I Verb NP

I prefer NP

I prefer Det Nom

I prefer a Nom

I prefer a Nom Noun

I prefer a Noun Noun

Rules used

S   → NP VP

NP → Pro

VP → Verb NP

Pro → *I*

Verb → prefer

NP → Det Nom

Det → *a*

Nom → Nom Noun

Nom → Noun

Grammar Rules		
S	→	NP VP
NP	→	Pronoun
	|	Proper-Noun
	|	Det Nominal
Nominal	→	Nominal Noun
	|	Noun
VP	→	Verb
	|	Verb NP
	|	Verb NP PP
	|	Verb PP
PP	→	Preposition NP

Noun → *flight* | *breeze* | *trip* | *morning*
Verb → *is* | *prefer* | *like* | *need* | *want* | *fly*
Adjective → *cheapest* | *non-stop* | *first* | *latest*
          | *other* | *direct*
Pronoun → *me* | *I* | *you* | *it*
Proper-Noun → *Alaska* | *Baltimore* | *Los Angeles*
            | *Chicago* | *United* | *American*
Determiner → *the* | *a* | *an* | *this* | *these* | *that*
Preposition → *from* | *to* | *on* | *near*
Conjunction → *and* | *or* | *but*

An example (cont.)

Rule expansions

S

NP VP

Pro VP

Pro Verb NP

I Verb NP

I prefer NP

I prefer Det Nom

I prefer a Nom

I prefer a Nom Noun

I prefer a Noun Noun

Rules used

S → NP VP

NP → Pro

VP → Verb NP

Pro → *I*

Verb → prefer

NP → Det Nom

Det → *a*

Nom → Nom Noun

Nom → Noun

Noun → *flight*

Grammar Rules

S	→	NP VP
NP	→	Pronoun
	|	Proper-Noun
	|	Det Nominal
Nominal	→	Nominal Noun
	|	Noun
VP	→	Verb
	|	Verb NP
	|	Verb NP PP
	|	Verb PP
PP	→	Preposition NP

Noun → flight | breeze | trip | morning
Verb → is | prefer | like | need | want | fly
Adjective → cheapest | non-stop | first | latest
| other | direct
Pronoun → me | I | you | it
Proper-Noun → Alaska | Baltimore | Los Angeles
| Chicago | United | American
Determiner → the | a | an | this | these | that
Preposition → from | to | on | near
Conjunction → and | or | but

An example (cont.)

Rule expansions

S

NP VP

Pro VP

Pro Verb NP

I Verb NP

I prefer NP

I prefer Det Nom

I prefer a Nom

I prefer a Nom Noun

I prefer a Noun Noun

I prefer a Noun *flight*

Rules used

S   → NP VP

NP → Pro

VP → Verb NP

Pro → *I*

Verb → prefer

NP → Det Nom

Det → *a*

Nom → Nom Noun

Nom → Noun

Noun → *flight*

Grammar Rules

$S \rightarrow NP\ VP$

$NP \rightarrow Pronoun$
$\quad\ \ |\ Proper\text{-}Noun$
$\quad\ \ |\ Det\ Nominal$
$Nominal \rightarrow Nominal\ Noun$
$\quad\ \ |\ Noun$

$VP \rightarrow Verb$
$\quad\ \ |\ Verb\ NP$
$\quad\ \ |\ Verb\ NP\ PP$
$\quad\ \ |\ Verb\ PP$

$PP \rightarrow Preposition\ NP$

$Noun \rightarrow flight\ |\ breeze\ |\ trip\ |\ morning$
$Verb \rightarrow is\ |\ prefer\ |\ like\ |\ need\ |\ want\ |\ fly$
$Adjective \rightarrow cheapest\ |\ non\text{-}stop\ |\ first\ |\ latest$
$\quad\ \ |\ other\ |\ direct$
$Pronoun \rightarrow me\ |\ I\ |\ you\ |\ it$
$Proper\text{-}Noun \rightarrow Alaska\ |\ Baltimore\ |\ Los\ Angeles$
$\quad\ \ |\ Chicago\ |\ United\ |\ American$
$Determiner \rightarrow the\ |\ a\ |\ an\ |\ this\ |\ these\ |\ that$
$Preposition \rightarrow from\ |\ to\ |\ on\ |\ near$
$Conjunction \rightarrow and\ |\ or\ |\ but$

An example (cont.)

Rule expansions

S

NP VP

Pro VP

Pro Verb NP

I Verb NP

I prefer NP

I prefer Det Nom

I prefer a Nom

I prefer a Nom Noun

I prefer a Noun Noun

I prefer a Noun *flight*

Rules used

S   → NP VP

NP → Pro

VP → Verb NP

Pro → *I*

Verb → prefer

NP → Det Nom

Det → *a*

Nom → Nom Noun

Nom → Noun

Noun → *flight*

Noun → *morning*

Grammar Rules

$S \rightarrow NP\ VP$

$NP \rightarrow Pronoun$
$\quad\ \ |\ Proper\text{-}Noun$
$\quad\ \ |\ Det\ Nominal$
$Nominal \rightarrow Nominal\ Noun$
$\quad\ \ |\ Noun$

$VP \rightarrow Verb$
$\quad\ \ |\ Verb\ NP$
$\quad\ \ |\ Verb\ NP\ PP$
$\quad\ \ |\ Verb\ PP$

$PP \rightarrow Preposition\ NP$

$Noun \rightarrow flight\ |\ breeze\ |\ trip\ |\ morning$
$Verb \rightarrow is\ |\ prefer\ |\ like\ |\ need\ |\ want\ |\ fly$
$Adjective \rightarrow cheapest\ |\ non\text{-}stop\ |\ first\ |\ latest$
$\quad\ \ |\ other\ |\ direct$
$Pronoun \rightarrow me\ |\ I\ |\ you\ |\ it$
$Proper\text{-}Noun \rightarrow Alaska\ |\ Baltimore\ |\ Los\ Angeles$
$\quad\ \ |\ Chicago\ |\ United\ |\ American$
$Determiner \rightarrow the\ |\ a\ |\ an\ |\ this\ |\ these\ |\ that$
$Preposition \rightarrow from\ |\ to\ |\ on\ |\ near$
$Conjunction \rightarrow and\ |\ or\ |\ but$

An example (cont.)

Rule expansions

S

NP VP

Pro VP

Pro Verb NP

I Verb NP

I prefer NP

I prefer Det Nom

I prefer a Nom

I prefer a Nom Noun

I prefer a Noun Noun

I prefer a Noun *flight*

I prefer a morning flight

Rules used

S   → NP VP

NP → Pro

VP → Verb NP

Pro → *I*

Verb → prefer

NP → Det Nom

Det → *a*

Nom → Nom Noun

Nom → Noun

Noun → *flight*

Noun → *morning*

Grammar Rules

$S \rightarrow NP\ VP$

$NP \rightarrow Pronoun$
$\quad |\quad Proper\text{-}Noun$
$\quad |\quad Det\ Nominal$
$Nominal \rightarrow Nominal\ Noun$
$\quad |\quad Noun$

$VP \rightarrow Verb$
$\quad |\quad Verb\ NP$
$\quad |\quad Verb\ NP\ PP$
$\quad |\quad Verb\ PP$

$PP \rightarrow Preposition\ NP$

$Noun \rightarrow flight \mid breeze \mid trip \mid morning$
$Verb \rightarrow is \mid prefer \mid like \mid need \mid want \mid fly$
$Adjective \rightarrow cheapest \mid non\text{-}stop \mid first \mid latest$
$\quad \mid other \mid direct$
$Pronoun \rightarrow me \mid I \mid you \mid it$
$Proper\text{-}Noun \rightarrow Alaska \mid Baltimore \mid Los\ Angeles$
$\quad \mid Chicago \mid United \mid American$
$Determiner \rightarrow the \mid a \mid an \mid this \mid these \mid that$
$Preposition \rightarrow from \mid to \mid on \mid near$
$Conjunction \rightarrow and \mid or \mid but$

An example (cont.)

- **Some Terminologies**
    - Each grammar must have one designated **start symbol**, S
    - We say the string "I prefer a morning flight" can be derived from S and the sequence of rule expansions is called a **derivation** of the string
    - A CFG $\mathcal{L}$ can be used to generate a set of strings. This set of strings is called the **formal language** defined by $\mathcal{L}$
    - Sentences that **can** be derived by $\mathcal{L}$ are called **grammatical** sentences in the formal language defined by $\mathcal{L}$
    - Sentences that **cannot** be derived by $\mathcal{L}$ are called **ungrammatical** sentences in the formal language defined by $\mathcal{L}$

Sequence of

rule expansions

S

NP VP

Pro VP

Pro Verb NP

I Verb NP

I prefer NP

I prefer Det Nom

I prefer a Nom

I prefer a Nom Noun

I prefer a Noun Noun

I prefer a Noun *flight*

I prefer a morning flight

An example (cont.)

- We can represent the derivation by a parse tree or in bracketed notation
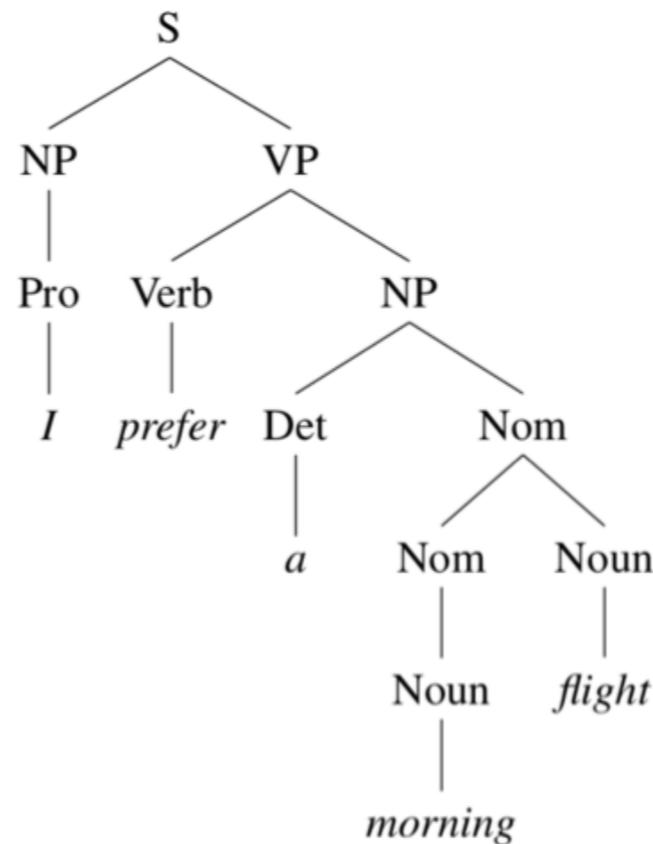
$[_S [_{NP} [_{Pro} \text{I}]] [_{VP} [_V \text{prefer}] [_{NP} [_{Det} \text{a}] [_{Nom} [_N \text{morning}] [_{Nom} [_N \text{flight}]]]]]]$

Q: What information is conveyed by a parse tree?

Formal Definition of Context-Free Grammar

- A context-free grammar G is defined by four parameters: *N, Σ, R, S*

N a set of **non-terminal symbols** (or **variables**)

Σ a set of **terminal symbols** (disjoint from N)

R a set of **rules** or productions, each of the form $A \rightarrow \beta$,

     where A is a non-terminal,

     β is a string of symbols from the infinite set of strings $(\Sigma \cup N)*$

S a designated **start symbol** and a member of N

- The Kleene star means "zero or more occurrences of the immediately previous character or regular expression"

Syntactic Parsing

- The problem of mapping from a sentence (a string of words) to its parse tree

Boeing is located in Seattle.
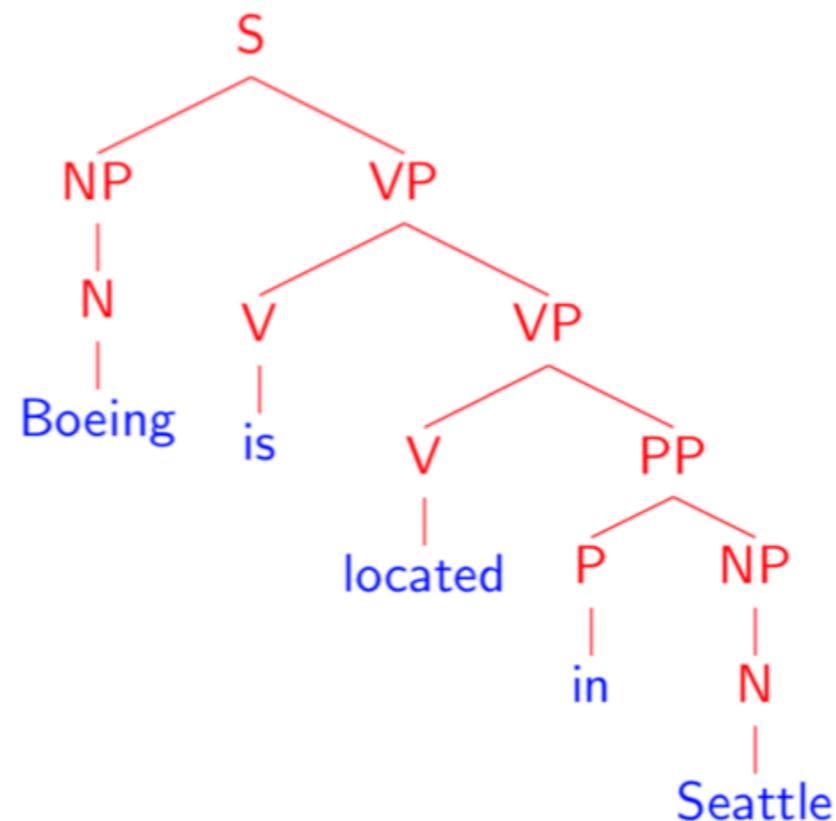
Syntactic Parsing (cont.)

- The problem with parsing: ambiguity
  - *Each string in the language defined by a CFG may have more than one derivation ("ambiguity")*

INPUT:

She announced a program to promote safety in trucks and vans

⇓

POSSIBLE OUTPUTS:

And there are more…
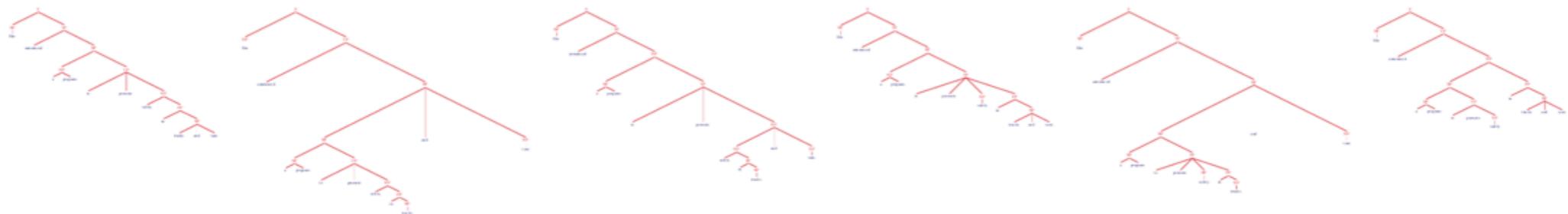
Syntactic Parsing (cont.)

- Sources of ambiguity

  - Lexical ambiguity, e.g., multiple word senses, multiple parts-of-speech

  - Structural ambiguity

Examples of ambiguous structures (cont.)

- Example 1:

    "I saw her duck with a telescope"

Examples of ambiguous structures (cont.)

- Example 1:

  "I saw her duck with a telescope"

  - Part-of-Speech ambiguity

    NN → duck

    Vi → duck

Examples of ambiguous structures (cont.)

- Example 1:

  "I saw her duck with a telescope"
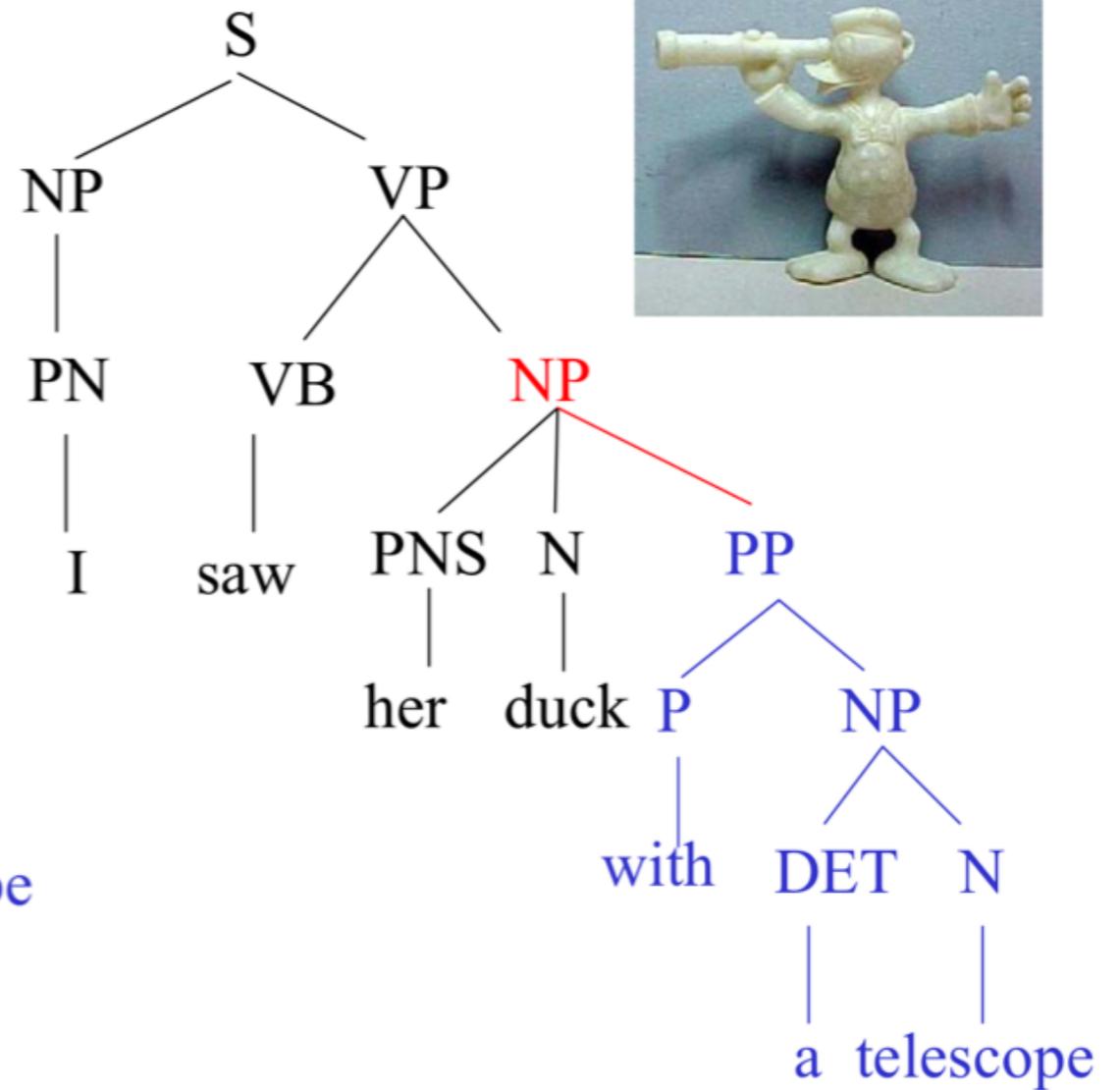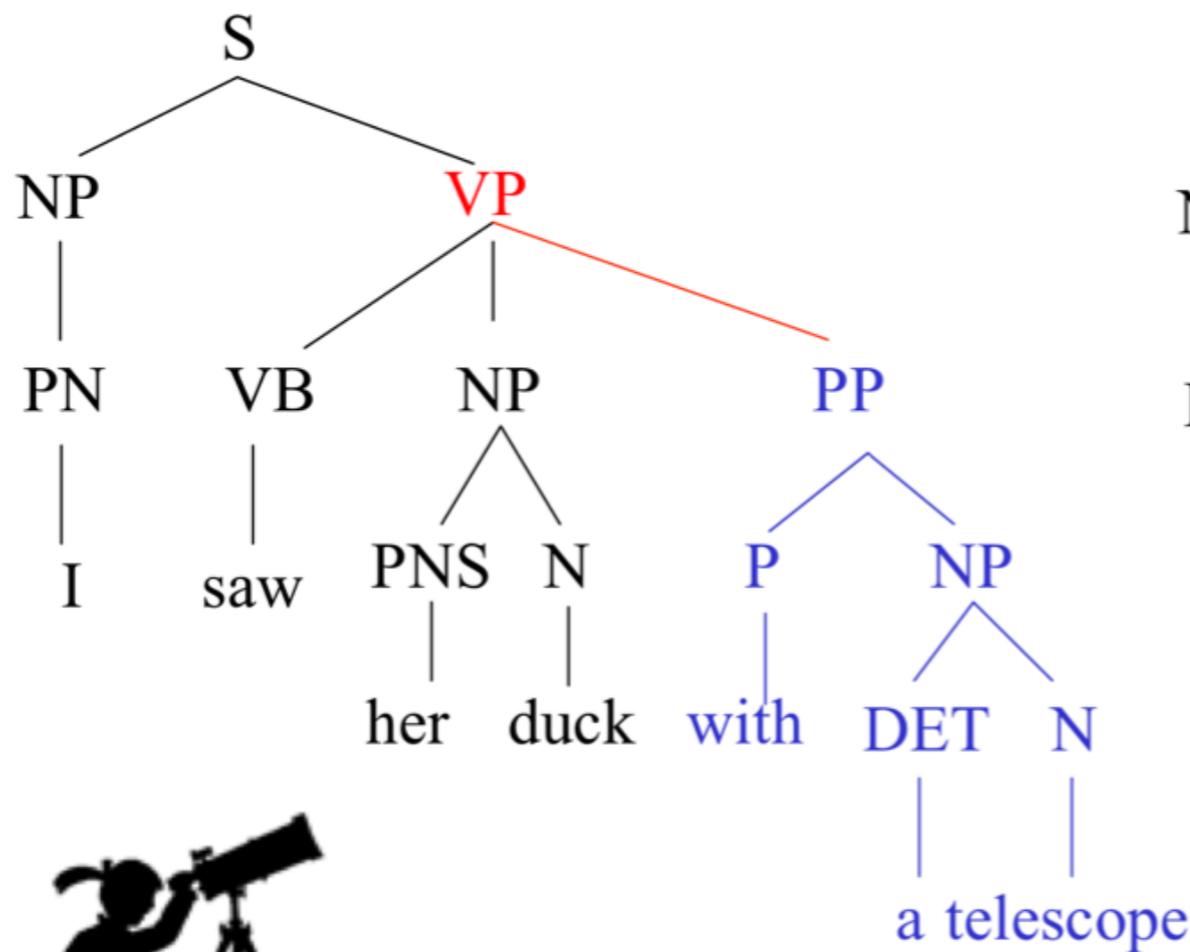
Examples of ambiguous structures (cont.)

- Example 2:

  "I drove down the road in the car"

- Example 2:
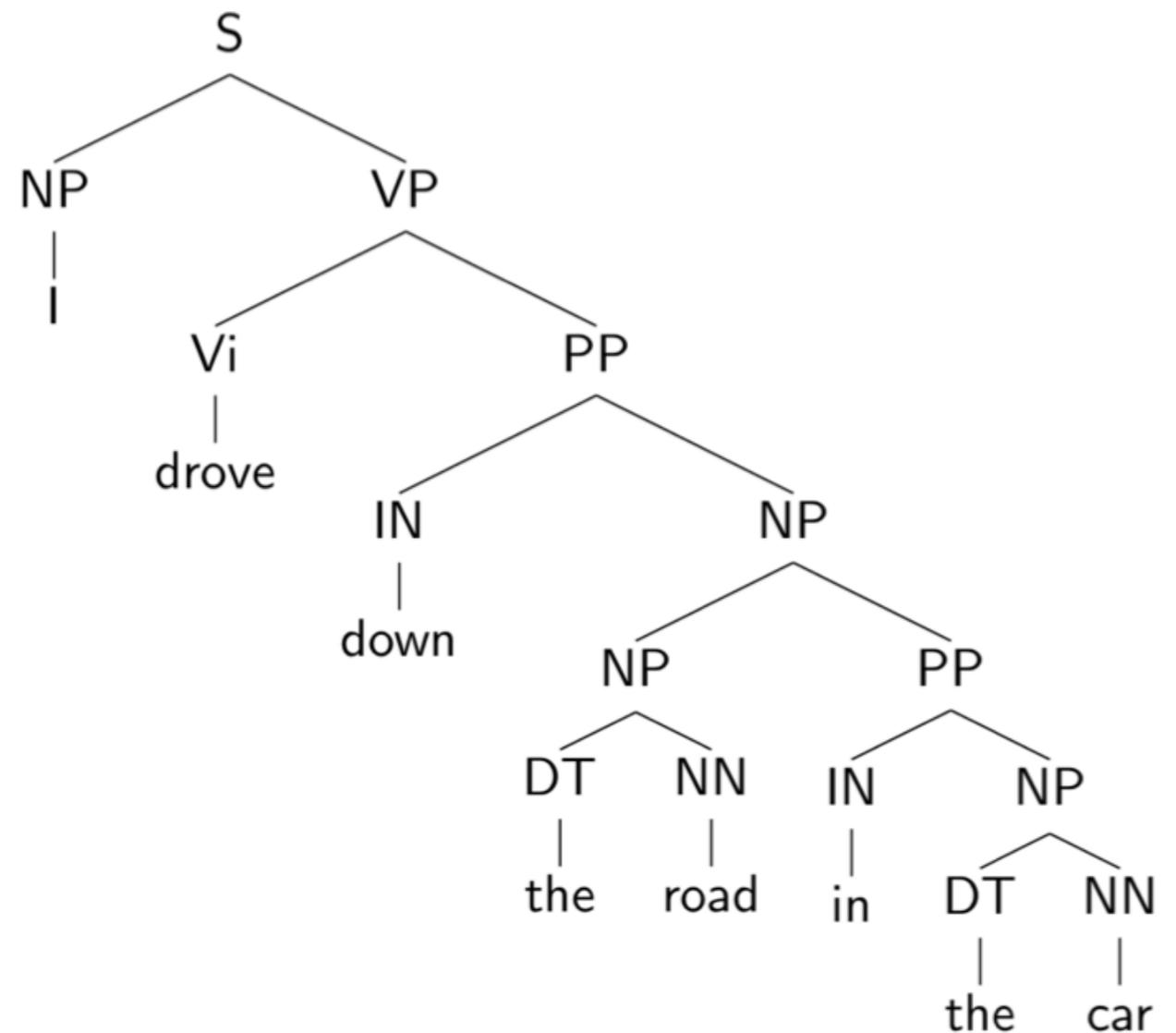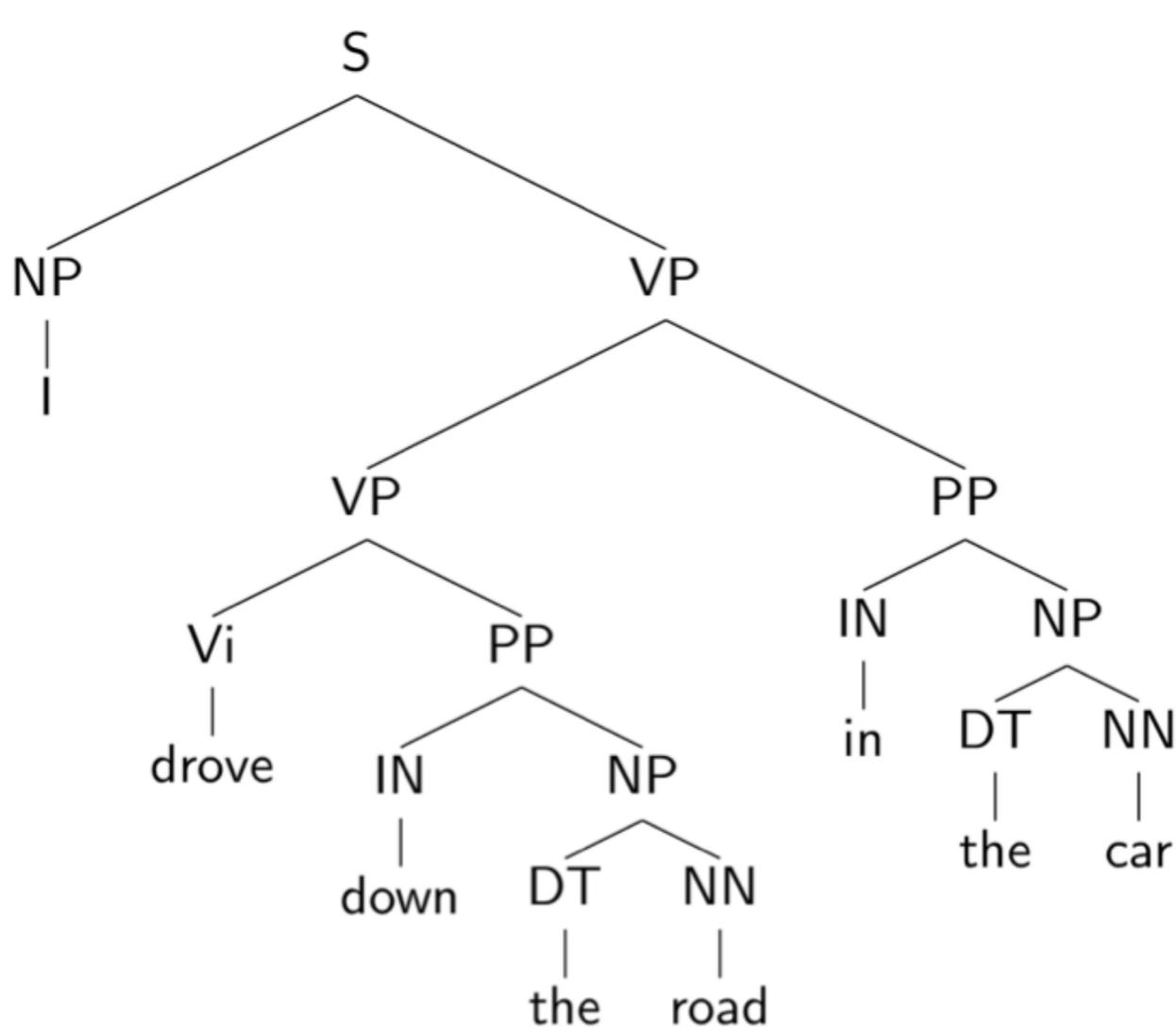
# Examples of ambiguous structures (cont.)

- Example 3:

  "the fast car mechanic"

Examples of ambiguous structures (cont.)

- Example 3:
  - Noun premodifiers
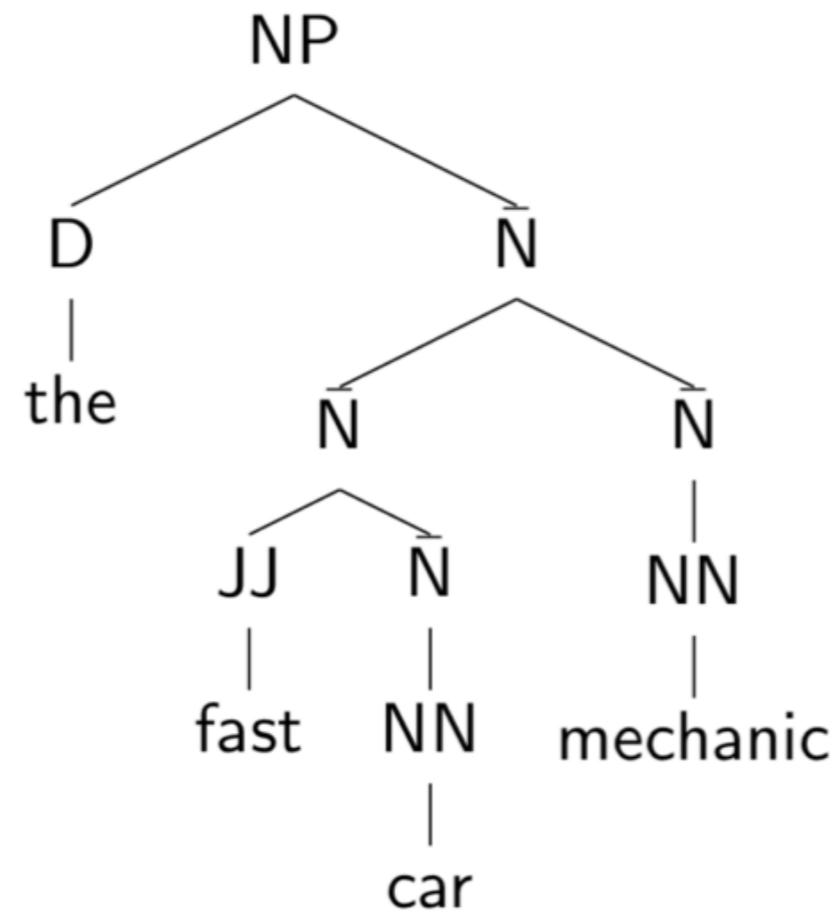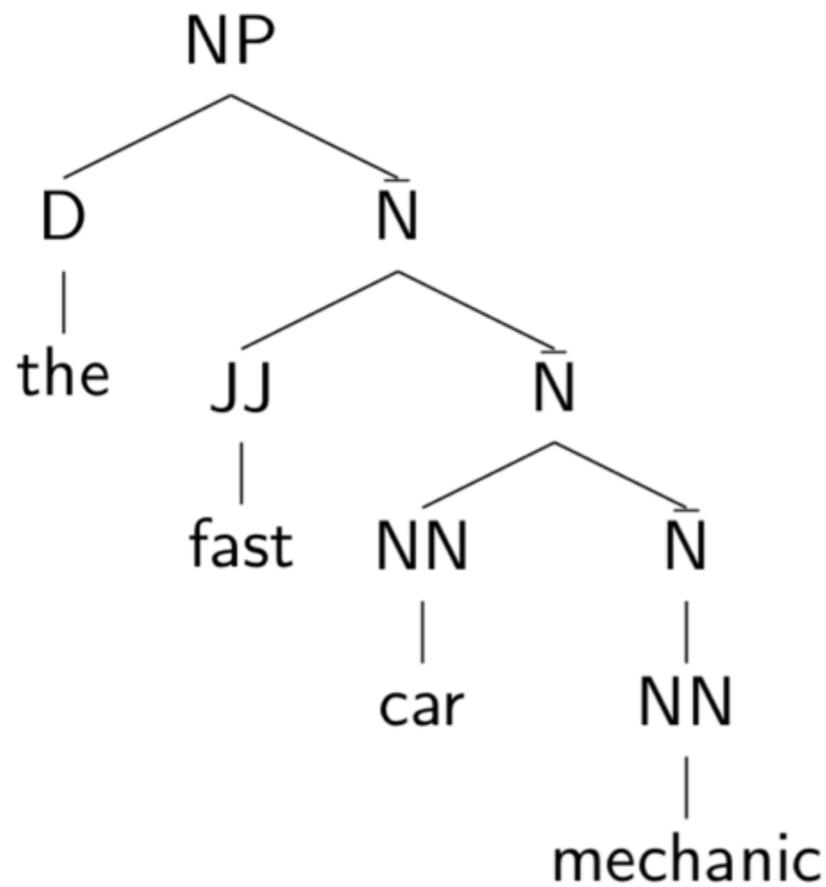
Overview

➢ An Introduction to Syntax

➢ Constituency

➢ Context-Free Grammars (CFGs)

➢ English Grammar in a Nutshell

- Some Grammar Rules

- Treebanks

Some Grammar Rules

- Sentence-level Constructions

    - Declaratives

   - Imperatives

    - Yes-no questions

    - Wh-questions

Some Grammar Rules (cont.)

- Declaratives
  - Form: **S → NP VP**   I prefer a morning flight
  - Have a great number of uses

- Imperatives
  - Form: **S → VP**   Show the lowest fare
  - Used for commands and suggestions

Some Grammar Rules (cont.)

- Yes-no questions
  - Form: **S → Aux NP VP**   Do any of these flights have stops?
  - Often used to ask questions

- Wh-questions
  - Wh-subject-questions
    - Form: **S → Wh-NP VP**   What airlines fly from Burbank to Denver?
    - Identical to the declarative structure, except that the first NP contains some wh-word
  - Wh-no-subject-questions
    - Form: **S → Wh-NP Aux NP VP**

      What flights do you have from Burbank to Tacoma Washington?

Some Grammar Rules (cont.)

- Clauses and Sentences
    - The **S** rules are intended to account for entire sentences that stand alone as fundamental units of discourse

    - **S** can also occur on the right-hand side of grammar rules and can be embedded within larger sentences

    - The **S** rules are some sense *complete* (i.e., forming a complete thought)*.* They correspond to the notion of **clause**.

Some Grammar Rules (cont.)

- ## Noun Phrases
  - Can be complicated
    - Determiners
    - Pre-modifiers
    - Post-modifiers

"head" = central, most critical part of the NP
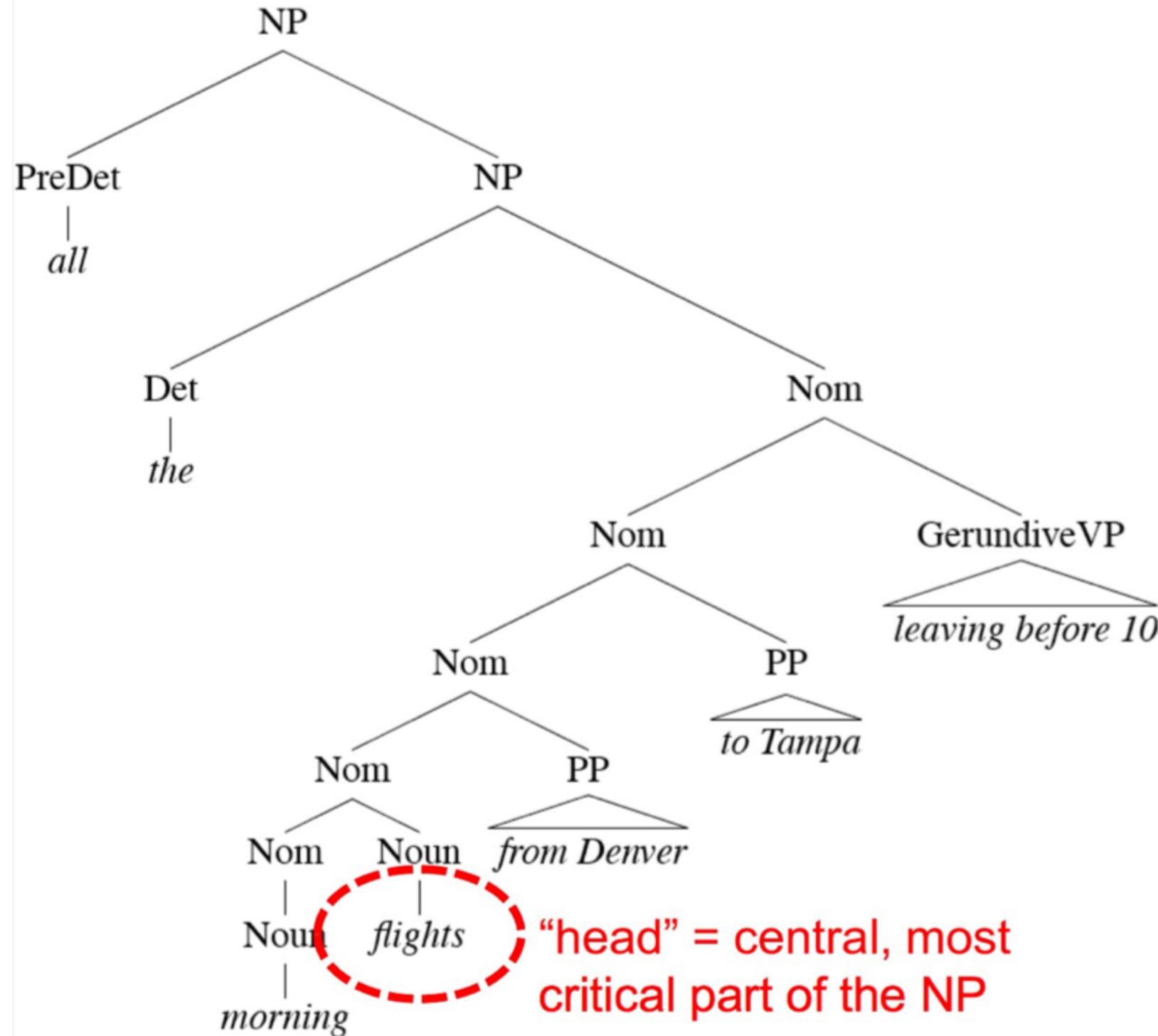
Some Grammar Rules (cont.)

- Determiners
  - Noun phrases can begin with determiners …
  - Determiners can be
    - simple lexical items a, the, this, those, any, some, etc.
    - simple possessives John's car
    - complex recursive versions of that

      John's sister's husband's son's car

Some Grammar Rules (cont.)

- Premodifiers
  - Come before the head
  - Examples
    - Cardinal numbers   one, two, three
    - Ordinal numbers   first, next, other
    - Quantifiers   many, (a) few, several
    - Adjectives   first-class, longest, non-stop
  - Ordering constraints

    three large cars vs. large three cars

Some Grammar Rules (cont.)

- Postmodifiers
  - Come after the head
  - Three kinds
    - Prepositional phrases   from Seattle
    - Non-finite clauses   arriving before noon
    - Relative clauses   that serve breakfast
  - Similar recursive rules to handle these:
    - Nominal → Nominal PP
    - Nominal → Nominal GerundVP
    - Nominal → Nominal RelClause

Some Grammar Rules (cont.)

- Agreement Issues
  - Agreement: constraints that hold among various constituents
  - For example, subjects must agree with their verbs on person and number:

    I am cold.   You are cold.   He is cold.

    * I are cold.   * You is cold.   * He am cold.
  - Requires separate productions for each combination in CFG:
    - S → NP1stPersonSing VP1stPersonSing
    - S → NP2ndPersonSing VP2ndPersonSing
    - NP1stPersonSing → …
    - VP1stPersonSing → …
    - NP2ndPersonSing → …
    - VP2ndPersonSing → …

Some Grammar Rules (cont.)

- Other agreement Issues
  - Pronouns have case (e.g. nominative, accusative) that must agree with their syntactic position.

    I gave him the book.   * I gave he the book.

    He gave me the book.   * Him gave me the book.

  - Many languages have gender agreement.

    Los Angeles   * Las Angeles

    Las Vegas   * Los Vegas

Some Grammar Rules (cont.)

- Verb Phrases

  - English verb phrases consists of

    – Head verb

    – Zero or more following constituents (called arguments)

  - Sample rules

    – VP → Verb   disappear

    – VP → Verb NP   prefer a morning flight

    – VP → Verb NP PP   leave Boston in the morning

    – VP → Verb PP   leave on Thursday

Some Grammar Rules (cont.)

- Subcategorization Issues
  - Specific verbs take some types of arguments but not others
    - Transitive verb: "found" requires a direct object

    John found the ring.   * John found.

    - Intransitive verb: "disappeared" cannot take one

    John disappeared.   * John disappeared the ring.

    - "gave" takes both a direct and indirect object

    John gave Mary the ring.   * John gave Mary.   *John gave the ring

    - "want" takes an NP, or non-finite VP or S

    John wants a car.   John wants to buy a car.

   John wants Mary to take the ring.   * John wants.

Some Grammar Rules (cont.)

- Subcategorization frames
    - Specify the range of argument types that a given verb can take.

Frame	Verb	Example
$\emptyset$	eat, sleep	I ate
NP	prefer, find, leave	Find [$_{NP}$ the flight from Pittsburgh to Boston]
NP NP	show, give	Show [$_{NP}$ me] [$_{NP}$ airlines with flights from Pittsburgh]
$PP_{from}\ PP_{to}$	fly, travel	I would like to fly [$_{PP}$ from Boston] [$_{PP}$ to Philadelphia]
$NP\ PP_{with}$	help, load	Can you help [$_{NP}$ me] [$_{PP}$ with a flight]
VPto	prefer, want, need	I would prefer [$_{VPto}$ to go by United airlines]
VPbrst	can, would, might	I can [$_{VPbrst}$ go from Boston]
S	mean	Does this mean [$_{S}$ AA has a hub in Boston]

Treebanks

- Data for parsing experiments
  - Penn WSJ Treebank = 50,000 sentences with associated trees
  - Usual set-up: 40,000 training sentences, 2400 test sentences
  - Example tree
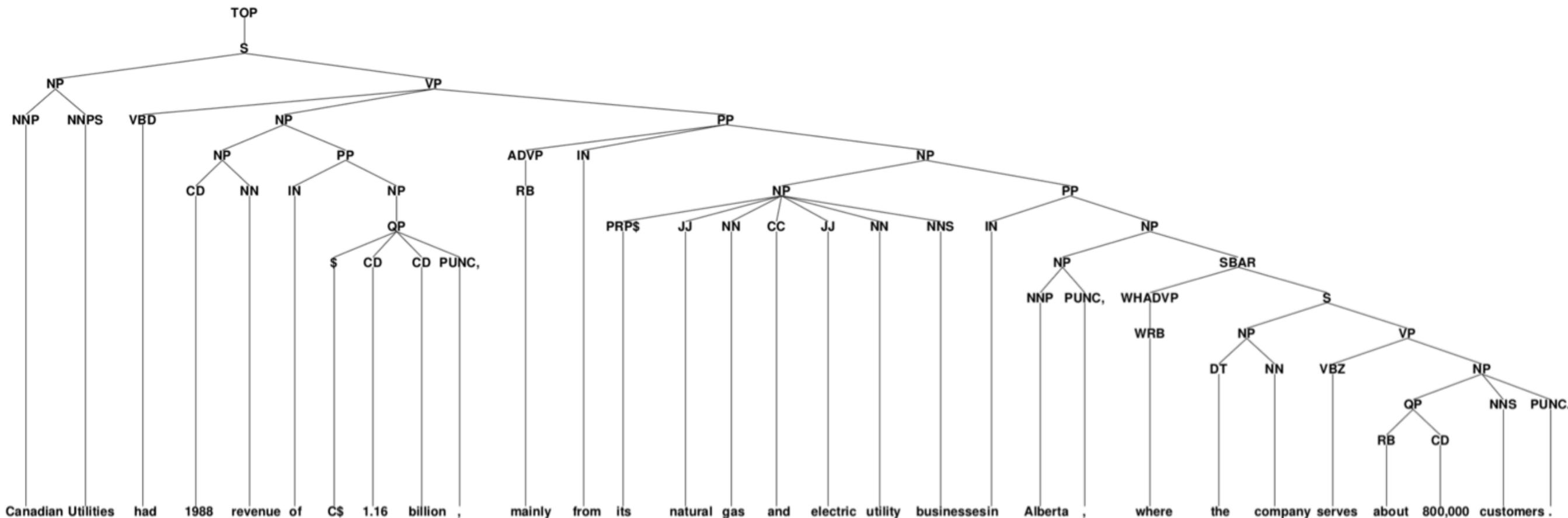
Treebanks (cont.)

- ## Penn Treebank

  - Treebanks implicitly define a grammar for the language

  - Penn Treebank has 4500 different rules for VPs, including…

    - VP → BD PP

    - VP → VBD PP PP

    - VP → VBD PP PP PP

    - VP → VBD PP PP PP PP

exercise!