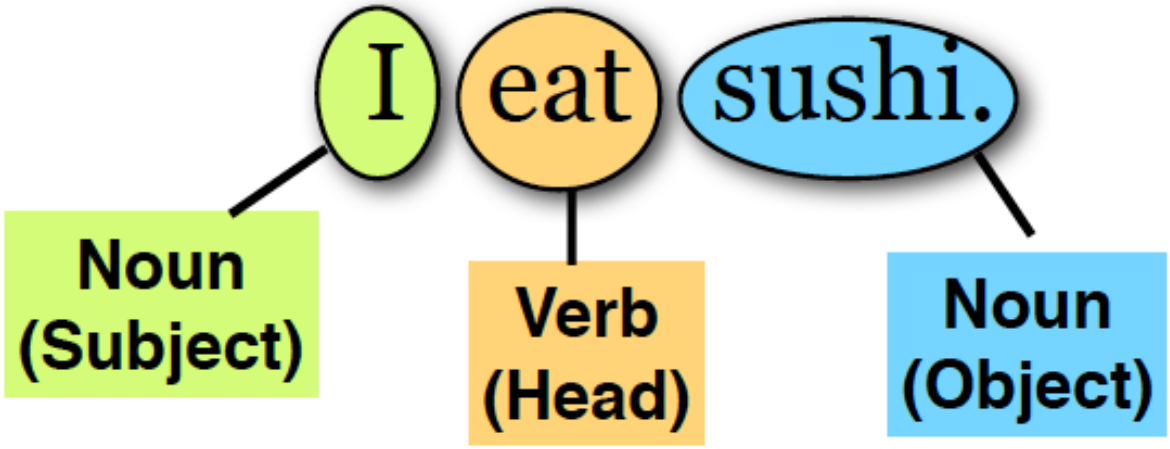
Lecture 7: Language Structure: Grammar

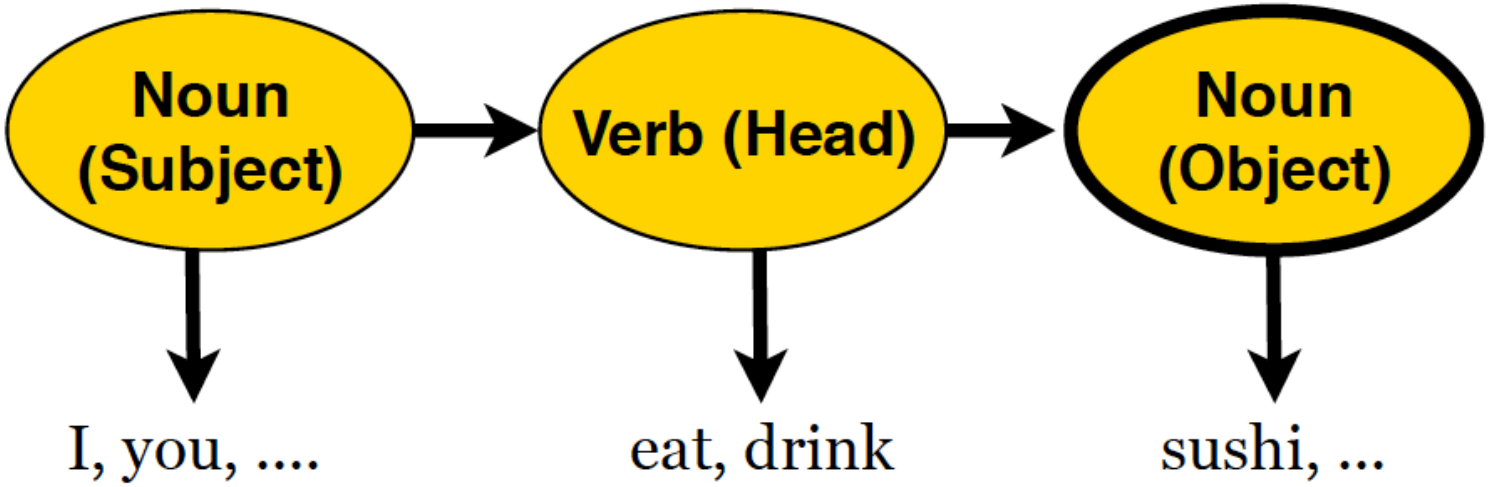
Kai-Wei Chang
CS @ UCLA
kw@kwchang.net

Couse webpage: https://uclanlp.github.io/CS269-17/

Basic sentence structure



A Markov Model



- I eat shshi; I eat meat; you eat banana…

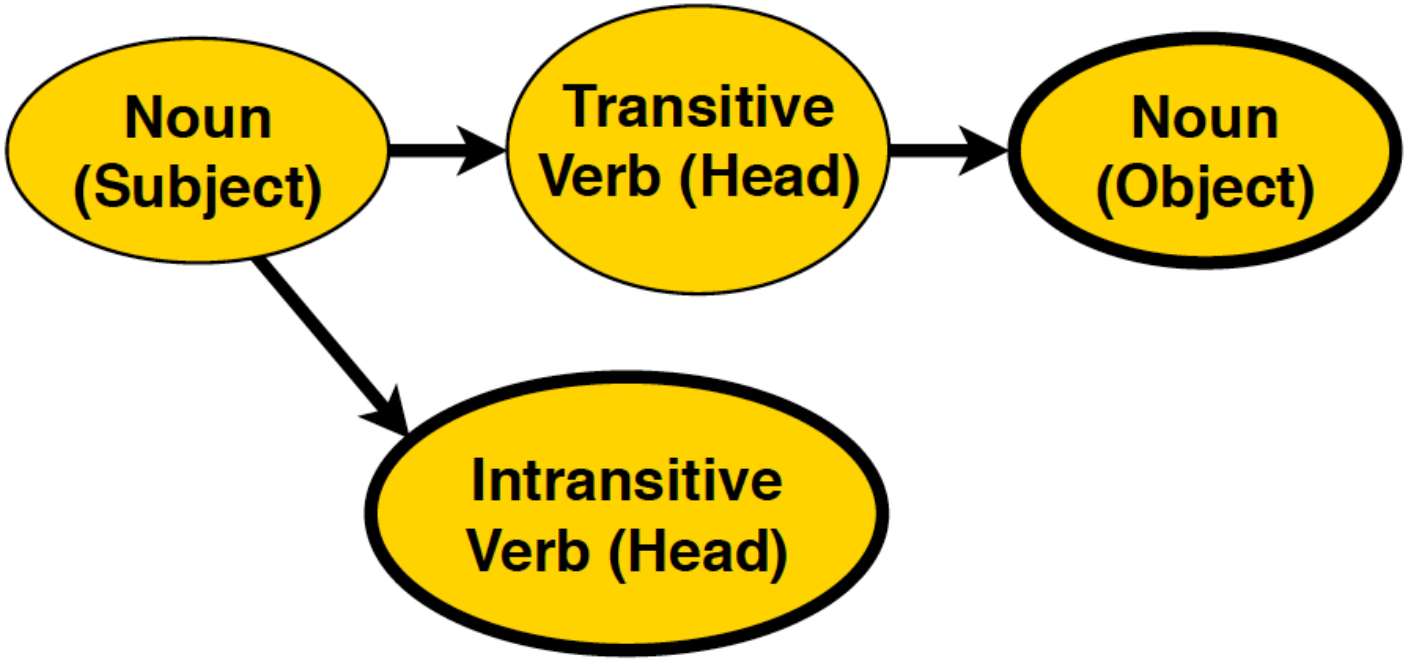
- Great, it covers many sentences

Words take different arguments

- [Good] I eat sushi
- [Bad] I run sushi
- [Bad] I give sushi

- Intransitive verbs (sleep): no object
- Transitive verbs (eat): take one direct object
- Ditransitive verbs (give): take an additional indirect object.

A better model

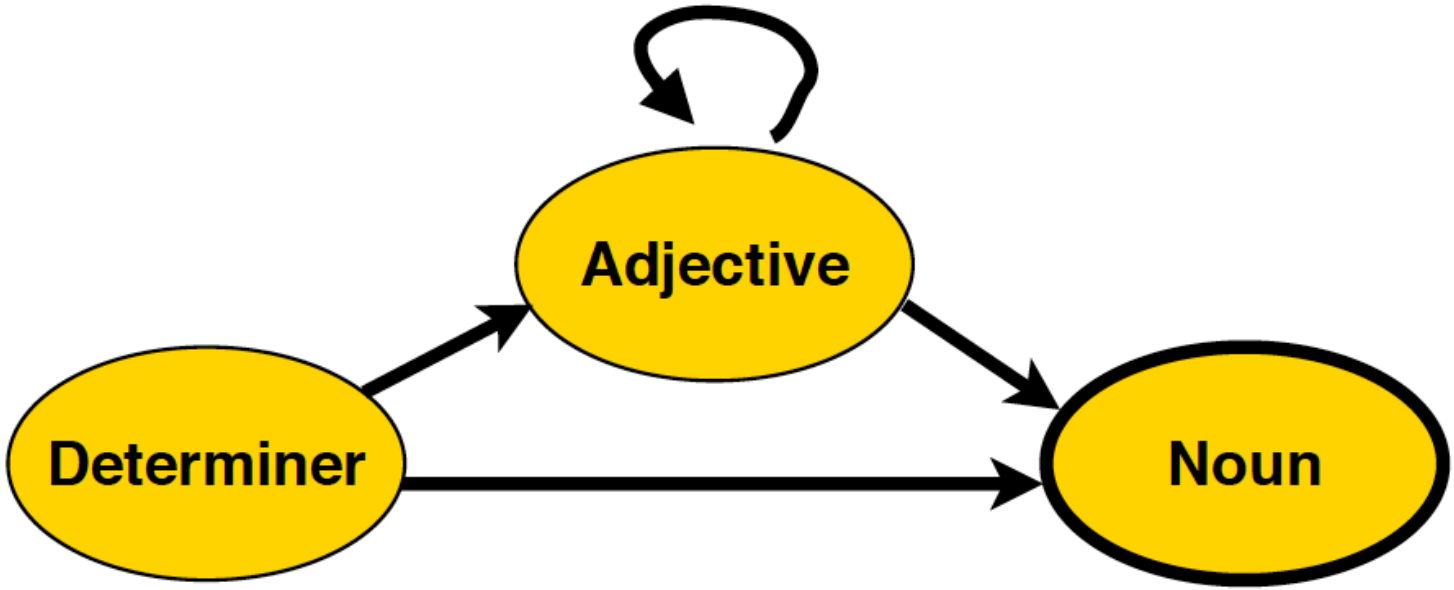


Language is recursive

the ball
the big ball
the big, red ball
the big, red, heavy ball
....

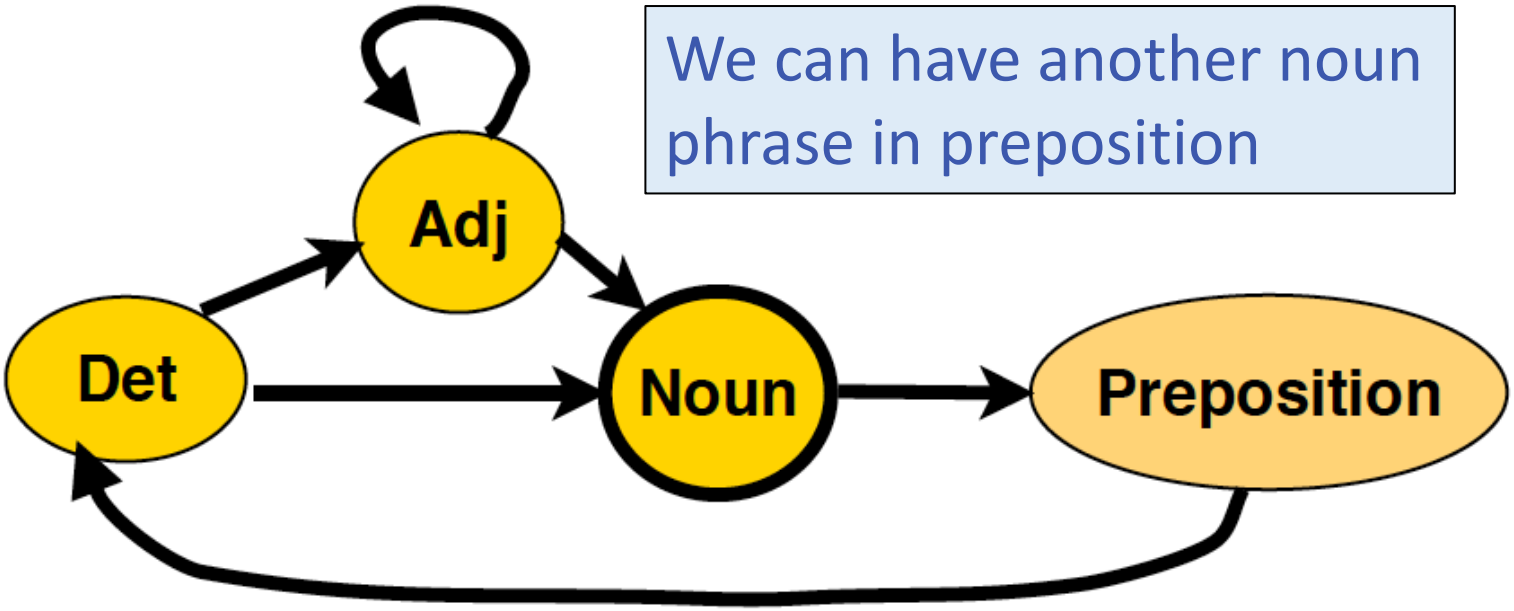
Adjectives can modify nouns.
We can have unlimited modifiers (in theory)

We know how to model the simple one



Recursion can be more complex

the ball
the ball in the garden
the ball in the garden behind the house
the ball in the garden behind the house next to the school



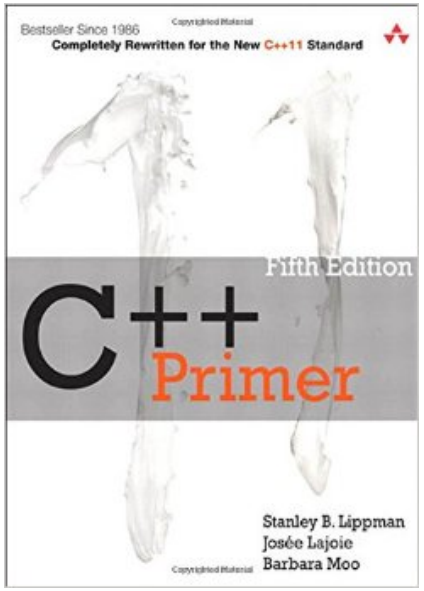
We can have another noun phrase in preposition

Syntactic parsing

- Idea: model language as a recursive generating process
  - Often use a tree structure
  - Decompose a sentence

What is grammar?

- A compact way to define and describe the structure of sentences
- Why we need grammar?
  - Number of C++ programs?



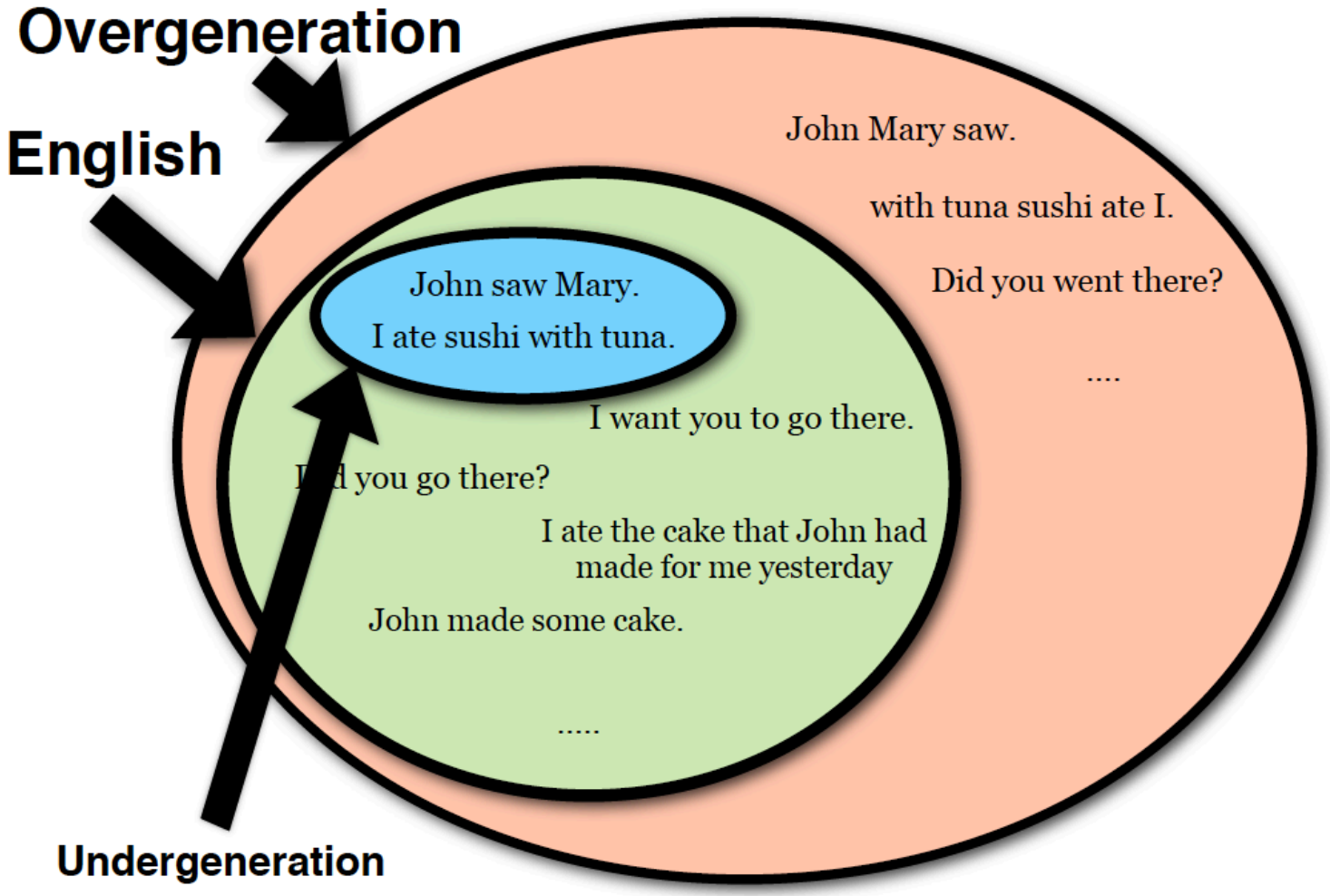
976 pages.

C++ standard (2014)

ISO/IEC 14882:2014

1358 pages

Can we define a program that generates all English sentences?

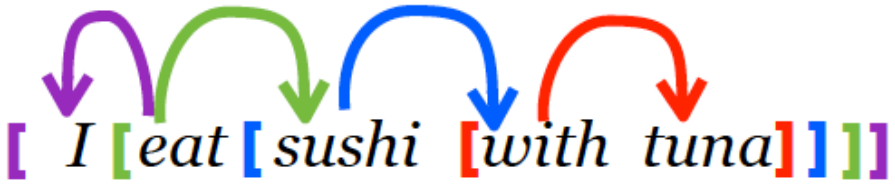


From Julia Hockenmaier, Intro to NLP

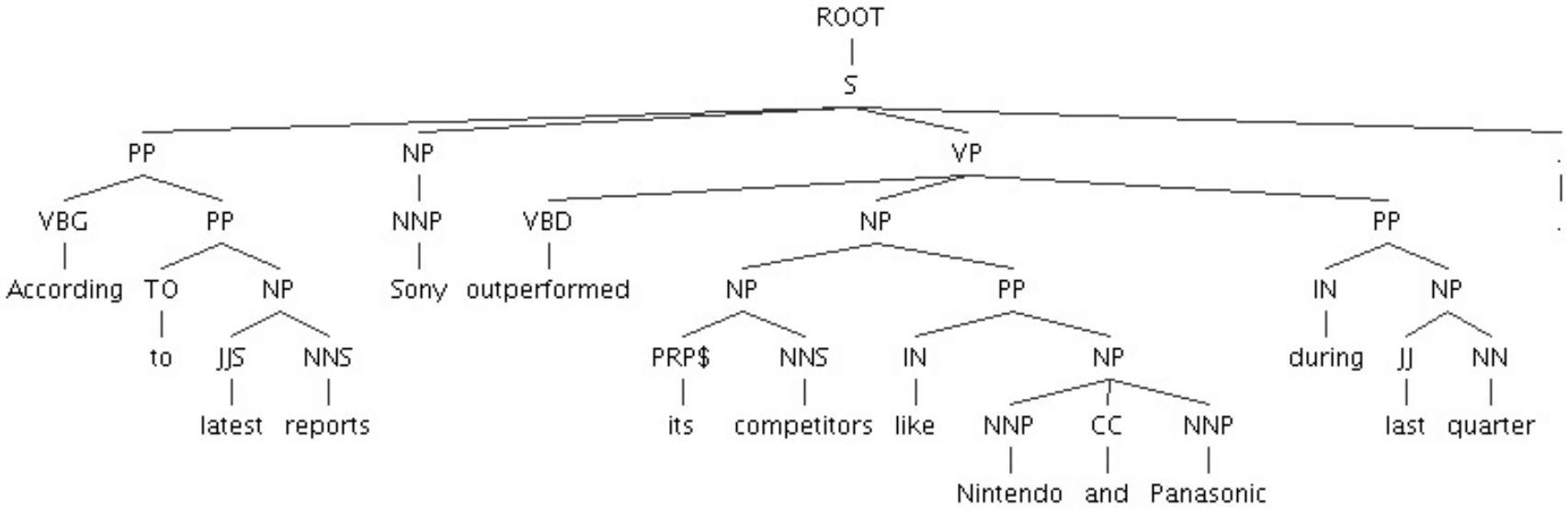
What is sentence structure

- Sentence structure is hierarchical
  - A sentence consists of phrases (or constituents)

Sentence structure defines dependencies between words or phrases:



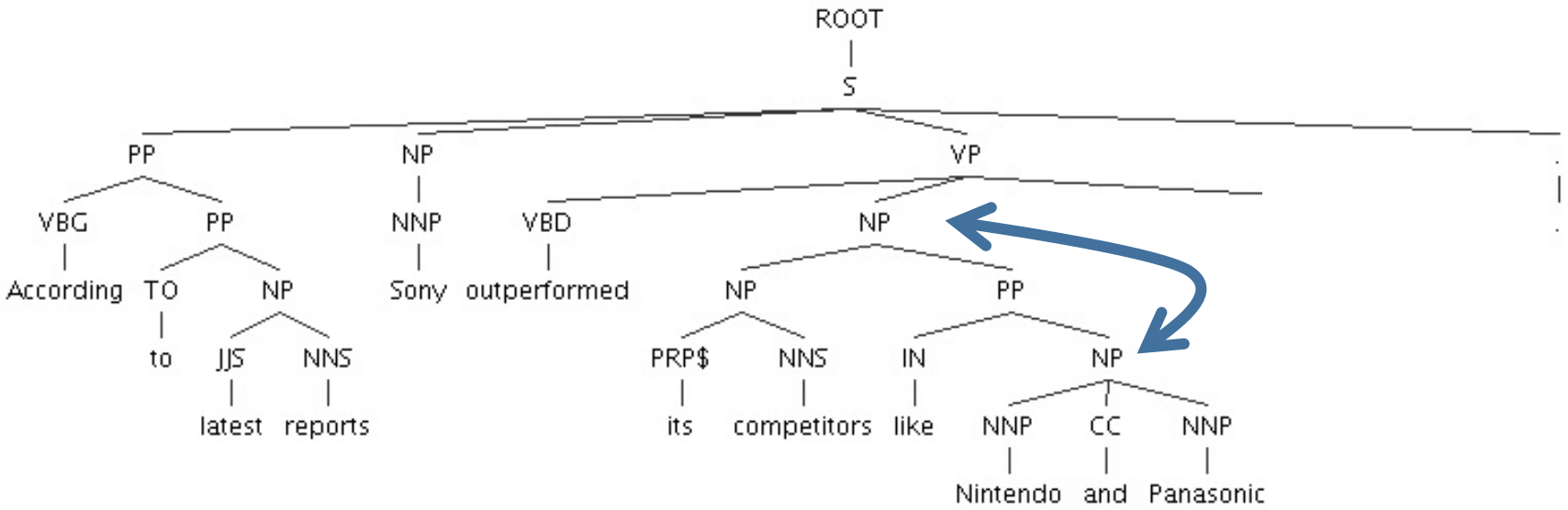
Can have complex constituents



According to latest reports Sony outperformed its competitors like Nintendo and Panasonic during last quarter.

Can have complex constituents

- Syntactically, constituents behave like simple ones



According to latest reports Sony outperformed its competitors like Nintendo and Panasonic during last quarter.

Constituency

- Groups of words that behave as a single unit or phrase
  - E.g., Noun phrases: the man, a girl with glasses
  - Prepositional phrases: with classes, on a table
  - Verb phrase: eat sushi, sleep, sleep soundly
- Phrases has a head:
  - Other parts called dependents
  - E.g, the man, a girl with glasses

Properties of constituents

- Substitution
  - He talks [in class] ⇒ He talks [there]
- It can move around in a sentence
  - He talks [in class] ⇒ [In class], he talks.
- Can be used as an answer:
  - Where does he talk? – [In class]

Types of dependencies

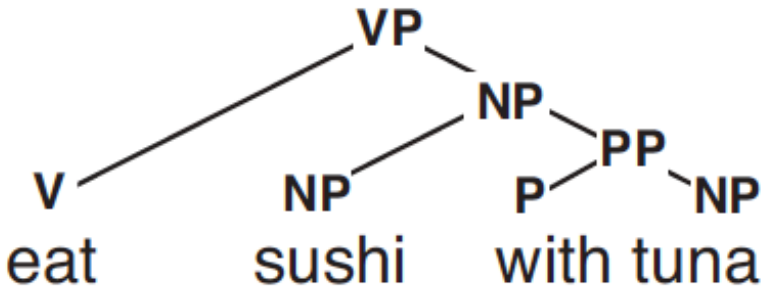
- Phrases has a head:
  - Other parts called dependents
  - E.g, the man, a girl with glasses
- Dependents can be arguments or adjuncts
- Arguments are obligatory
  - E.g., [John] likes [Mary]
- Adjuncts are optional
  - E.g., John runs [fast]
  - Adverbs, PPs, Adjectives…

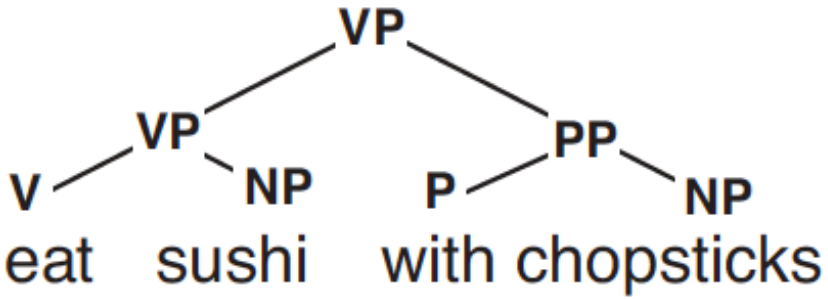
All arguments have to be present and cannot be occupied multiple times

Can be an arbitrary number of adjuncts

How to represent the structure

Phrase structure trees

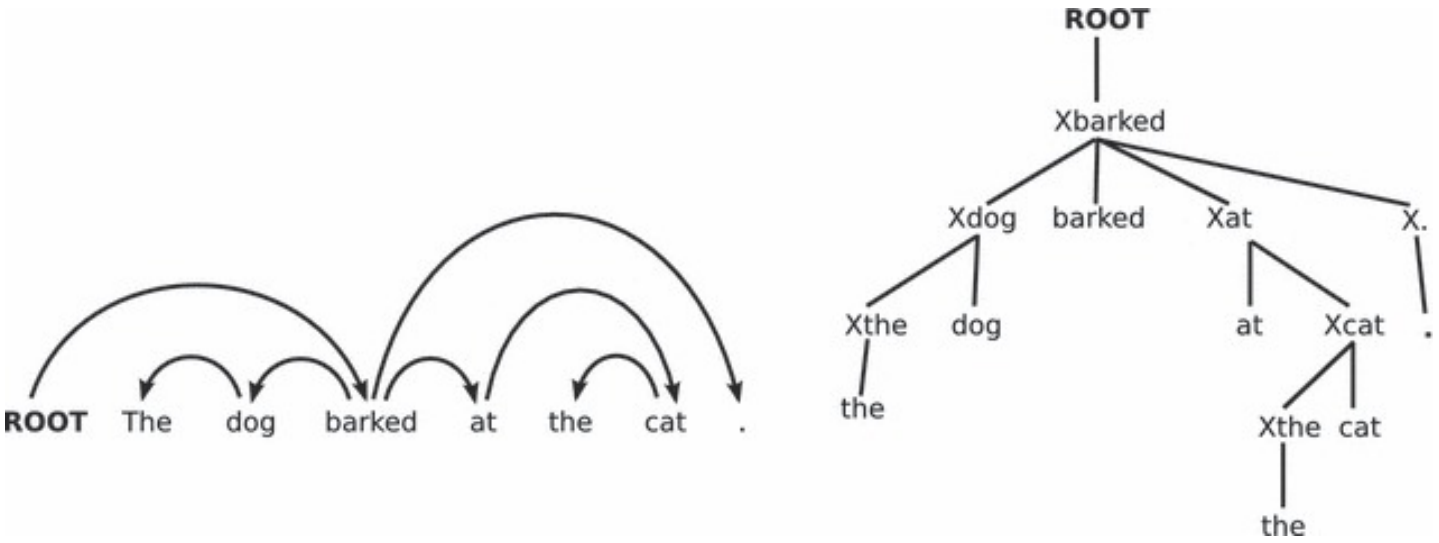




Dependency trees

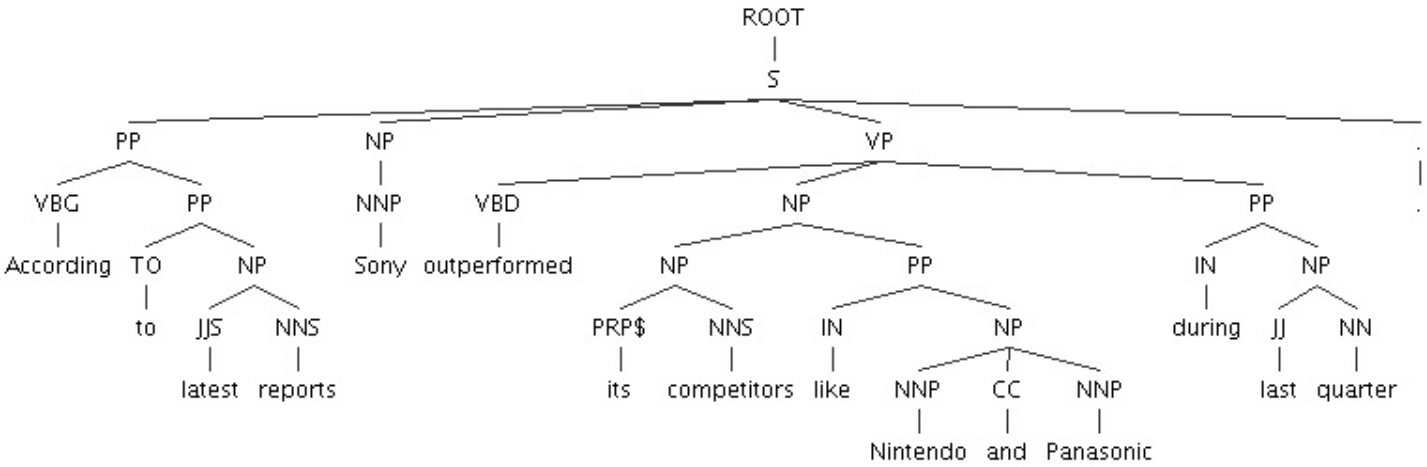
Dependency Trees

- Dependency grammar describe the structure of sentences as a graph (tree)
  - Nodes represent words
  - Edges represent dependencies



Phrases structure trees

- Can be modeled by Context-free grammars



According to latest reports Sony outperformed its competitors like Nintendo and Panasonic during last quarter.

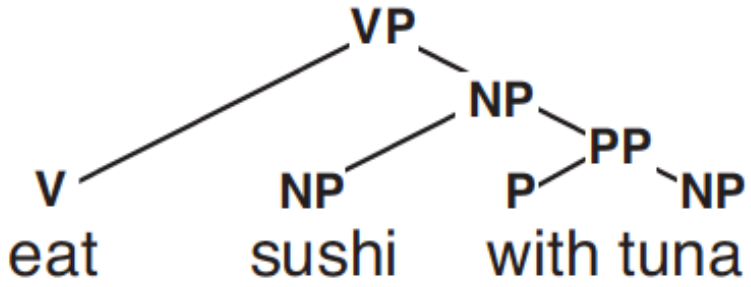
Context-free grammars

DT → {the, a}
N → {ball, garden, house, sushi }
P → {in, behind, with}
NP → **DT N**
NP → **NP PP**
PP → **P NP**

N: noun
P: preposition
NP: "noun phrase"
PP: "prepositional phrase"

Parse tree defined by CFG

N → {sushi, tuna}
P → {with}
V → {eat}
NP → **N**
NP → **NP PP**
PP → **P NP**
VP → **V NP**



Generate sentences by CFG

The mouse ate the corn.
The mouse that the snake ate ate the corn.
The mouse that the snake that the hawk ate ate ate the corn.
....

These sentences are all grammatical.
They can be generated by a CFG:

S	→	NP	VP
NP	→	NP	RelClause
RelClause	→	that	NP ate

Example: Noun Phrases

Simple NPs:

[He] sleeps. **(pronoun)**
[John] sleeps. **(proper name)**
[A student] sleeps. **(determiner + noun)**

Complex NPs:

[A tall student] sleeps. **(det + adj + noun)**
[The student in the back] sleeps. **(NP + PP)**
[The student who likes MTV] sleeps. **(NP + Relative Clause)**

Example: verb phrase

He [eats].
He [eats sushi].
He [gives John sushi].
He [eats sushi with chopsticks].

VP → V
VP → V NP
VP → V NP PP
VP → VP PP

V → *{eats, sleeps gives,...}*

Sentences

[He eats sushi].
[Sometimes, he eats sushi].
[In Japan, he eats sushi].

S → NP VP
S → AdvP S
S → PP S

He says [he eats sushi].
VP → V_{comp} S
V_{comp} → {says, think, believes}