

Formal Grammars

Chapter 12 J&M'09

Recall that finite-state methods assume Natural Language sentences only involve local dependencies.

This assumption is false, and creates trivial errors:

- Machine translation

(1) This is the company that the bank bought.

- Speech synthesis

(2) Which record broke?

- Search engines (and question answering)

(3) Who would have been vice-president if Romney had won the election?

[What would happen if Obama and Romney tied? - US elections ...](#)

[www.independent.co.uk › News › World › US Elections](#)

Nov 4, 2012 – On 21 January, Mitt **Romney** takes the presidential oath **of** office. ... On **election** eve, Obama **has** a slim lead in polls ... and the Senate, likely to remain controlled by Democrats, **would** select the **vice-president**. "**If** ... Just as past presidential **election** disputes **have been** resolved with an act **of** statesmanship, ...

[What went wrong with my prediction about Mitt Romney and the ...](#)

[www.foxnews.com/.../what-went-wrong-with-my-prediction-about-...](#)

Nov 13, 2012 – I predicted that Mitt **Romney would** be our next president. ... This **election** should **have been**, 100 out **of** 100 times, a repeat **of** Reagan's landslide ... The GOP **has** to reach out to women and Latinos. ... That argument **would** have **won** the presidency. ... He is a former Libertarian **vice presidential** nominee.

[Adviser: Romney "shellshocked" by loss - CBS News](#)

[www.cbsnews.com/8301-250.../adviser-romney-shellshocked-by-loss...](#)

Nov 8, 2012 – **Romney's** pre-**election** optimism was the result **of** a few key miscalculations. ... They just couldn't believe they **had been** so wrong. ... They didn't want to **have** to withdraw their concession, like Al Gore did in 2000, and they (I, also could relate to **Vice President**, Joe Biden's behavior, in debate with Ryan.

[Live Coverage and Results - Election 2012 - NYTimes.com](#)

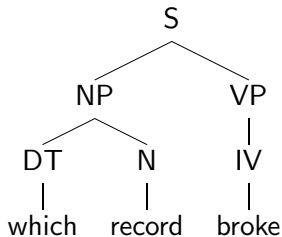
[elections.nytimes.com/2012/results/live-coverage](#)

Nov 7, 2012 – From left, Michelle Obama, President Obama, **Vice President** Joseph R. Biden and ... The candidates **have** given their speeches and gone home, and we're heading President Obama narrowly **won** Virginia over Mitt **Romney**, giving him at CHICAGO — **When** NBC projected President Obama **had been** ...

Recall the **Chomsky-Schützenberger hierarchy**

Language	Grammar	Automaton
Regular	$A \rightarrow a, A \rightarrow aB$	Finite state machine
Context-free	$A \rightarrow \gamma$	Non-deterministic pushdown automaton
Context-sensitive	$\gamma A \beta \rightarrow \alpha \gamma \beta$	Linear-bounded non-deterministic Turing machine
Recursively enumerable	$\alpha \rightarrow \beta$	Turing machine

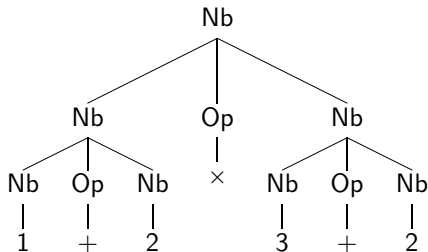
If we construct a more complete representation of the sentence, then the word tags will be more consistent with the overall structure:



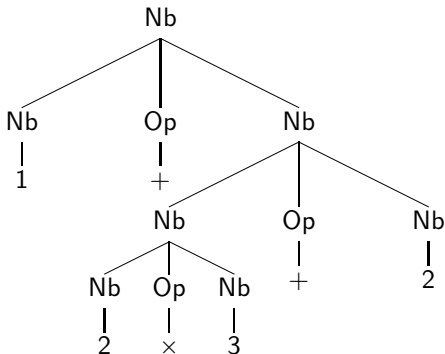
ANALOGY:

imagine modeling the language of arithmetic with n-grams or FSA:

$$(1+2) \times (3+2) = 15$$



$$1 + ((2 \times 3) + 2) = 9$$



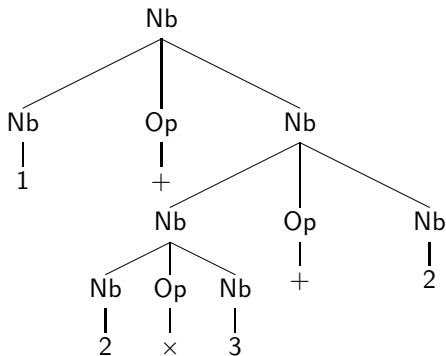
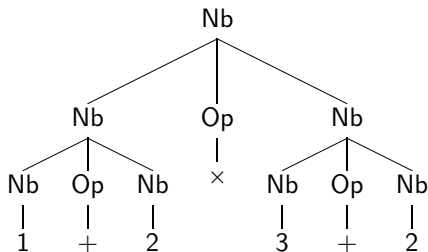
Grammars

Lexicon:

Nb \rightarrow 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9

Op \rightarrow + | - | \times | =

Syntax: Nb \rightarrow Nb Op Nb



NLs have a bigger lexicon, more rules, and no visible brackets.

- (4) a. [French [history teacher]]
b. [[French history] teacher]
- (5) a. [The reporter [said [that the elephant died yesterday]]].
b. [The reporter [said [that the elephant died] yesterday]].
- (6) a. The room contained [[noisy children] and animals].
b. The room contained [noisy [children and animals]].
- (7) a. Put the block in the box.
b. Put the block in the box on the table.
c. Put the block in the box on the table in the hallway.
d. Put the block in the box on the table in the hallway outside the bedroom.
e. Put the block in the box on the table in the hallway outside the bedroom by the pool.

In many languages, word order is relatively free.

Again, Russian exemplifies:

- (8) a. Anna uvidela Natashu
 'Anna saw Natasha'
 - b. Anna Natashu uvidela.
 - c. Natashu Anna uvidela.
 - d. Natashu uvidela Anna.
 - e. Uvidela Natashu Anna.
 - f. Uvidela Anna Natashu.
- (9) a. Natasha uvidela Annu
 'Natasha saw Anna'
 - b. Annu Natasha uvidela.
 - c. Natasha Annu uvidela.
 - d. Natasha uvidela Annu.
 - e. Uvidela Natasha Annu.
 - f. Uvidela Annu Natasha.

Subject, Verb and Object order in a sample of 1228 languages:

- 497 are SOV (e.g. like Japanese – Japan)
- 435 are SVO (e.g. like Mandarin – China)
- 85 are VSO (e.g. like Irish – Ireland)
- 26 are VOS (e.g. like Nias – Indonesia)
- 9 are OVS (e.g. like Hixkaryana – Brazil)
- 4 are OSV (e.g. like Jaupés-Japurá – Brazil)
- 172 are mixed (e.g. Syrian Arabic; both SVO/VSO)

Haspelmath et al. 2005 The World Atlas of Language Structures, Oxford Press

Word order map of the [world's languages](#)

There are many grammar-based systems (some with online demos):

Lisp:

- [ERG](#)
- [KRG](#)
- [Heart of Gold](#)
- [Delph-In](#)

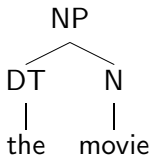
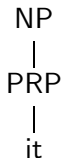
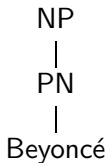
Java:

- [Stanford parser](#)
- [CCG](#)

Easy access to lots of grammars and parsers in Java via:

- StanfordNLP's [parser](#)
- OpenNLP's [parser](#)
- OpenCCG's [parser](#)
- NLTK's [parsers](#)

Three examples of basic types of NP:



Minigrammar

PRP \rightarrow *we* | *she* | *he* | *her* | ...

DT \rightarrow *the* | *each* | *a* | ...

N \rightarrow *movie* | *person* | *cat* | ...

PN \rightarrow *Tom* | *Beyoncé* | ...

NP \rightarrow PRP

NP \rightarrow PN

NP \rightarrow DT N

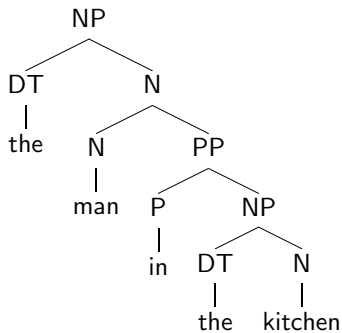
More rules

$N \rightarrow N \text{ PP}$

$\text{PP} \rightarrow P \text{ NP}$

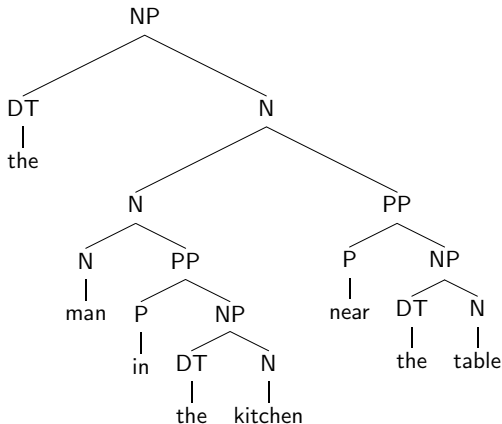
$P \rightarrow \textit{with} \mid \textit{on} \mid \textit{in} \mid \dots$

$N \rightarrow \textit{man} \mid \textit{cat} \mid \textit{rock} \mid \dots$



- (10) a. I know the man in the kitchen.
b. I know the man in the kitchen, near the table.
c. I know the man in the kitchen, near the table, in the blue shirt.
d. I know the man in the kitchen, near the table, in the blue shirt, with the mustache.

This ' $N \rightarrow N PP$ ' rule can apply to its own output:



In this case, each PP is imposing constraints on the same noun:
... 'the man who is in the kitchen, who is near the table'

But the added PPs need not be about the same N:

- (11) a. The winner [of the race]
b. The winner [of the race [on TV]]
c. The winner [of the race [on TV [in the restaurant]]]
d. The winner [of the race [on TV [in the bar [near my house]]]]
e.

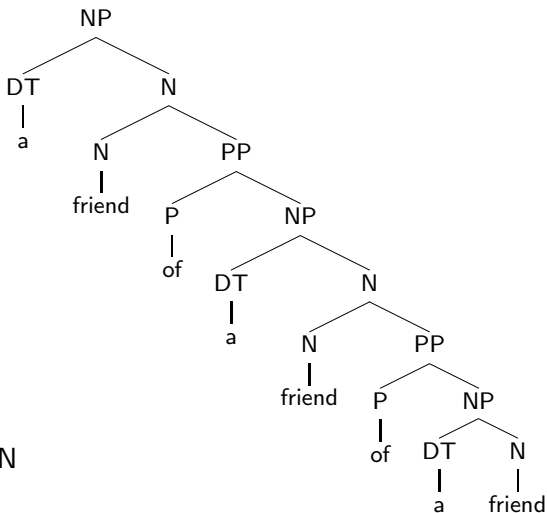
(12) I met a friend [of a friend [of a friend [of a friend [of Sue]]]].

This is already predicted by our rules

$NP \rightarrow DT\ N$

$N \rightarrow N\ PP$

$PP \rightarrow P\ NP$



$NP \rightarrow DT\ N$

$N \rightarrow N\ PP$

$PP \rightarrow P\ NP$

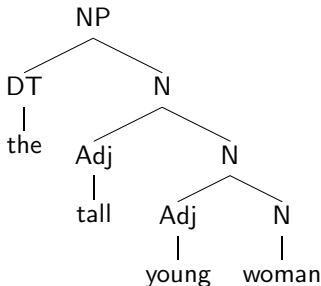
NPs optionally contain adjectives

- (13) a. The tall woman
b. The tall young woman
c. The tall young obnoxious woman
d. The tall young obnoxious Italian woman
e. ...

Two new rules:

$N \rightarrow \text{Adj } N$

$\text{Adj} \rightarrow \text{tall} \mid \text{short} \mid \text{green} \mid \dots$



Verb classes

- **Intransitive verbs** (no complements)

VP \rightarrow IV

IV \rightarrow *died* | *smiled* | *fainted* | *blushed* | ...

(14) a. Tom fainted.

b. *Tom fainted it

- **Transitive verbs** (one direct object complement)

VP \rightarrow TV NP

TV: \rightarrow *devour* | *print* | *hit* | *drive* | ...

(15) a. Tom devoured the burger.

b. *Tom devoured.

- **Ditransitive verbs** (two complements)

VP \rightarrow DTV NP NP

DTV \rightarrow *gave* | *sent* | *offer* | *supply* | ...

(16) a. Tom gave Sue the burger.

b. *Tom gave Sue.

Verbs combine with whatever complements they need in order to form **verb phrases** (VPs):

$$\text{VP} = \left\{ \begin{array}{l} \text{laughed} \\ \text{opened the door.} \\ \text{tossed him a book.} \\ \dots \end{array} \right\}$$

VPs are mobile constituents:

- (17) a. I wanted to sleep, and I did [sleep].
b. I wanted to sleep, and [sleep] I did.
c. I wanted to read a book, and I WILL [read a book].
d. I wanted to read a book, and [read a book] I WILL.
e.*I wanted to read a book, and [read] I WILL a book.

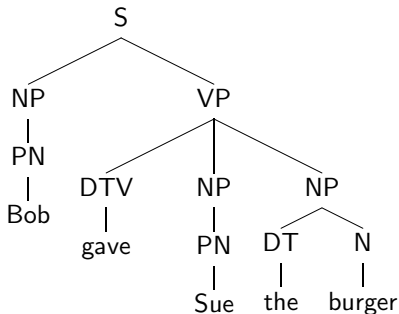
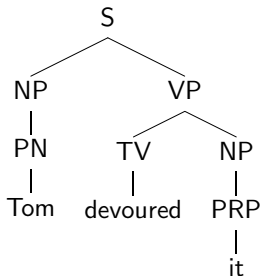
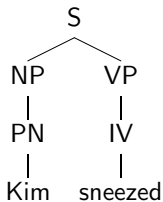
VPs can conjoin with other VPs to create another VP:

- (18) You can't [[drive a car] and [drink whiskey]].

Adding a NP to the left of a VP creates a sentence:

$S \rightarrow NP VP$

For example:

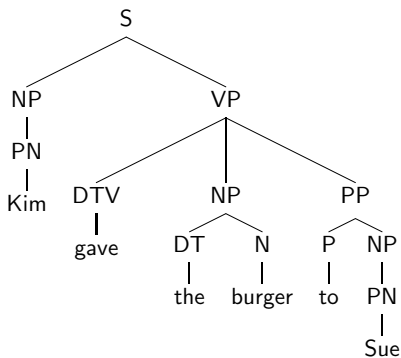
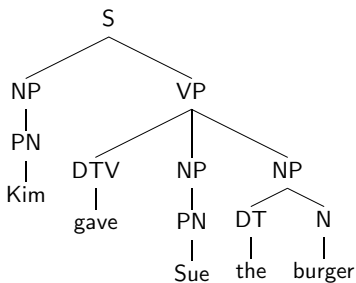


Ditransitive verbs can be used as Prepositional Dative verbs:

- (19) a. Kim gave Sue the burger.
b. Kim gave the burger to Sue.

We can capture this fact with an extra rule:

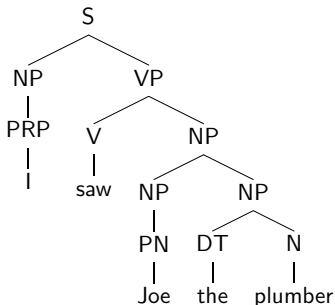
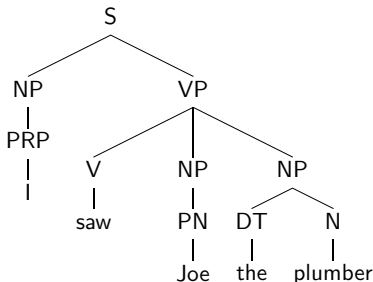
$VP \rightarrow DTV \ NP \ PP$



Fine-grained distinctions like TV / DTV are useful

- they reduce incorrect parses;

(20) I saw Joe the plumber.



- they are crucial for sentence generation:

(21) a.*Saw I Joe the plumber.

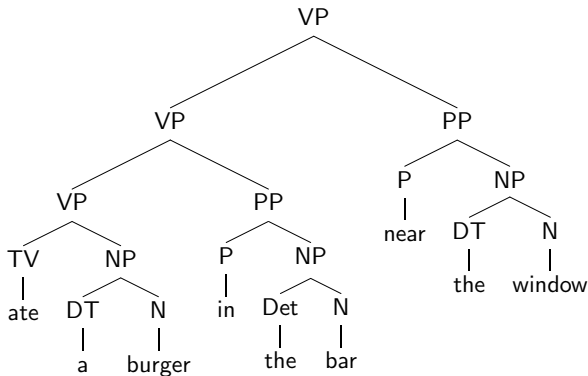
b.*Her saw the plumber Joe.

VPs can be 'modified' by PPs

- (22) a. Tom ate a burger.
b. Tom ate a burger in the bar.
c. Tom ate a burger in the bar, near the window.
d. Tom ate a burger in the bar, near the window, at noon.
e. ...

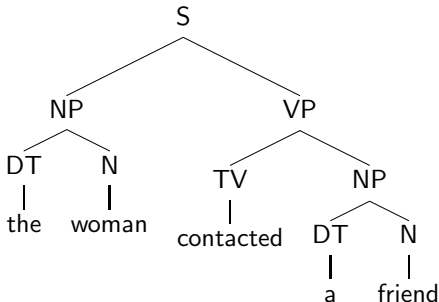
New rule:

VP \rightarrow VP PP



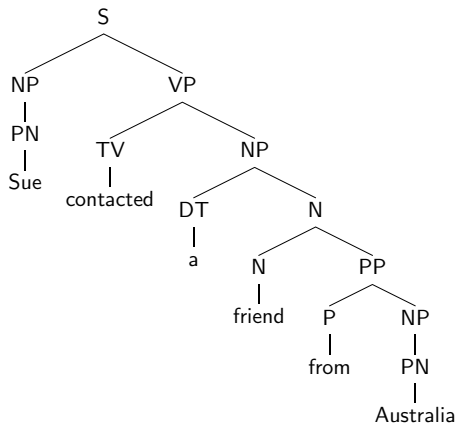
Our phrasal grammar rules so far:

- 1 NP \rightarrow DT N
- 2 N \rightarrow N PP
- 3 N \rightarrow Adj N
- 4 NP \rightarrow PN
- 5 NP \rightarrow PRP
- 6 PP \rightarrow P NP
- 7 S \rightarrow NP VP
- 8 VP \rightarrow IV
- 9 VP \rightarrow TV NP
- 10 VP \rightarrow DTV NP NP
- 11 VP \rightarrow DTV NP PP
- 12 VP \rightarrow VP PP

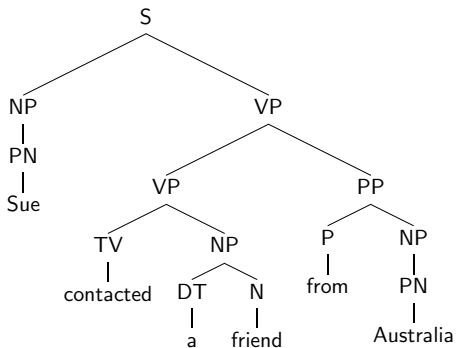


Where can '*from Australia*' go?

$N \rightarrow N PP$



$VP \rightarrow VP PP$



(23) Sue contacted a friend from Australia.

Interpretation 1 (I1): involves a friend from Australia

Interpretation 2 (I2): involves Sue being in Australia

This ambiguity vanishes in the **subject position**:

(24) A friend from Australia contacted Sue. (I1)

And in **it-cleft** sentences:

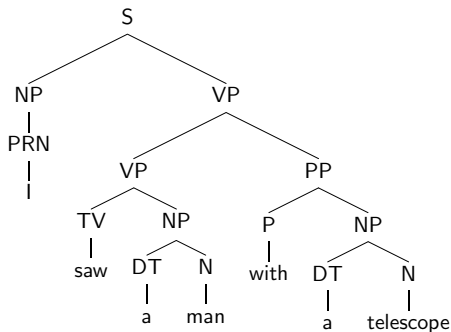
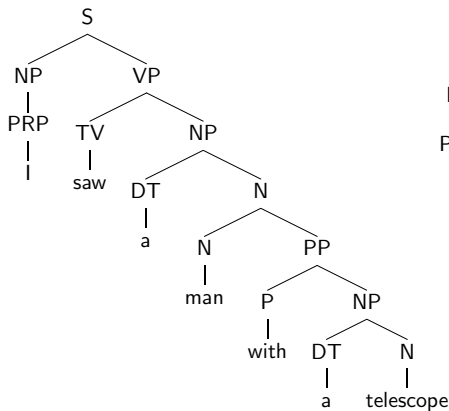
(25) a. It was a friend from Australia that Sue contacted. (I1)

b. It was from Australia that Sue contacted a friend. (I2)

c. It was a friend that Sue contacted from Australia. (I2)

The same set of rules explains the ambiguity of (26).

(26) I saw a man with a telescope.

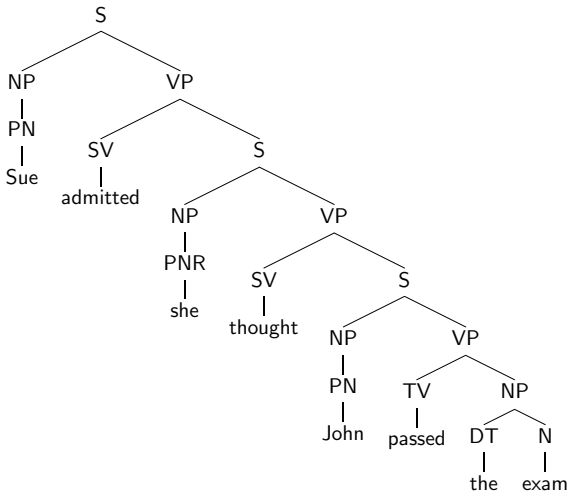


Indirect discourse structures

VP \rightarrow SV S

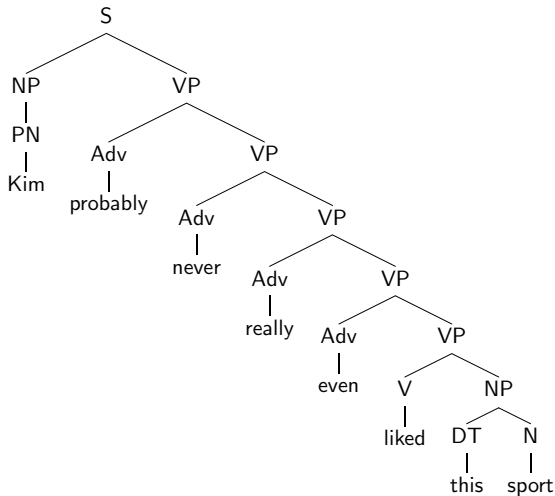
SV \rightarrow *said* | *thought* | *feared* | *denied* | *shouted* | *wished* | ...

(27) I think [Sue admitted [she thought [John passed the exam]]].



Adverbial adjunction

$VP \rightarrow ADV VP$

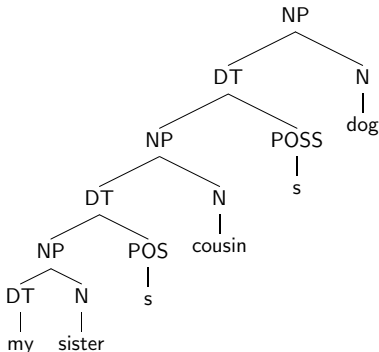


Genitive structures

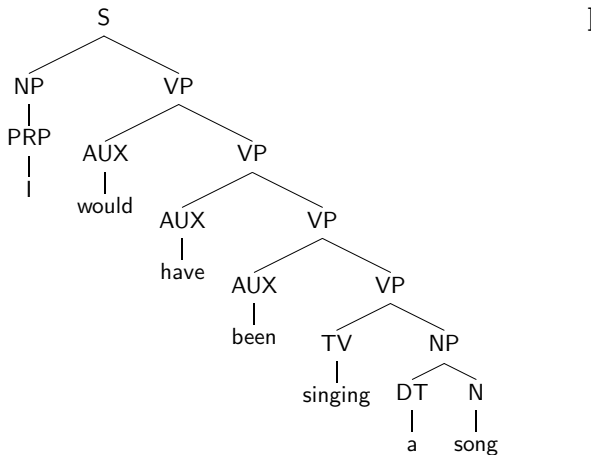
DT → NP POS

POS → 's

- (28) a. [[My father]'s tortoise] died.
b. [[[[My father]'s father]'s father]'s tortoise] died.
c. [[[[[My teacher]'s sister]'s best friend]'s cousin]'s dog] died.



Other sources of recursion:



$VP \rightarrow AUX VP$

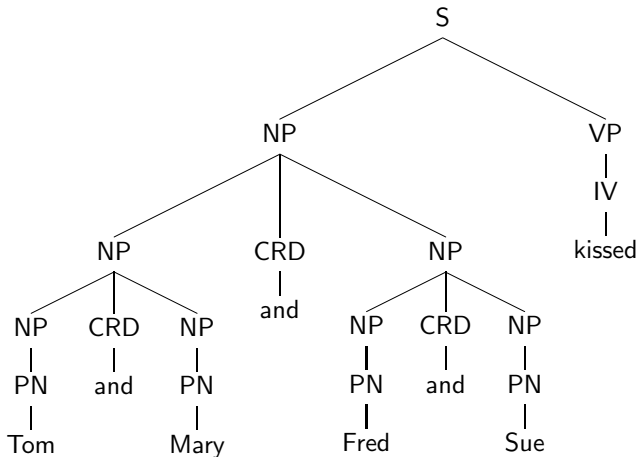
More sources of recursion:

$X \rightarrow X_1 \dots X_{n-1} \text{ CRD } X_n$

$\text{CRD} \rightarrow \text{and} \mid \text{or} \mid \text{but}$

- (29) a. [[John] and [the neighbor]] caught the cat.
b. [[The car and the truck]] collided.
- (30) a. The trips [[to France] and [to Iceland]] were great.
b. Fred relied either [[on Sue] or [on Kim]].
- (31) a. I bought the [[big] and [yellow]] book.
b. Robin is both [[tall] and [slim]].
- (32) a. Kim [[washed the car] and [cleaned the garage]].
b. You can't [[drink booze] and [drive a car]].
- (33) a. [[Tom went to NY] and [Mary stayed in California]].
b. He said [[Kim was friendly] and [Bob was annoying]].

Embedding coordination:



(34) John met with either [[John and Mary] or [Tom and Sue]].

How big is the English Grammar?

Nobody knows. Still an open problem.

Active formal research started in the 50s.

Modern large-scale computational grammars:

- Hand-written feature-based grammar
ERG: total of 266 rules (morphology + syntax)
- Corpora-induced grammars: 17,500 rules (Penn Treebank)
Over 17k (J&M'09:§12.4.2)

Do we really need type 2 (context-free) grammars?

Yes: some structures have the form $a^n b^n$

- (35) a. The movie was *Inception*.
b. The movie everyone loved was *Inception*.
c. The movie everyone I interviewed loved was *Inception*.

Do we need type 1 (context-sensitive) grammars?

Yes: some structures have the form $a^n b^m c^n d^m$

- Swiss German (Shieber 1985)
- Dutch (Bresnan & al. 1982)
- Bambara (Culy 1985)
- ... etc. (Walther 1999, Beesley & Karttunen 2000).

Dutch example: ... $noun_1 noun_2^m \dots verb_1^n verb_2^m \dots$

- (36) ... dat ik₁ Henk₂ haar₃ de nijlpaarden₃ zag₁ helpen₂ voeren₃
... that I Henk her the hippos saw help feed
'... that I saw Henk help her feed the hippos'

Some formalisms ‘push’ the grammar into the lexicon

A **categorial grammar** consists of:

① Categories

① Basic Categories: S, NP, PP, ...

② Complex Categories:

If X and Y are a categories, then so are X/Y and $X \backslash Y$.

② Two functional rules

$(X/Y) Y \Rightarrow X$

(Forward composition ($>$))

$Y (X \backslash Y) \Rightarrow X$

(Backward composition ($<$))

Example:

(37) Tom sneezed.

$$\frac{\frac{Tom}{NP} \quad \frac{sneezed}{(S \backslash NP)}}{S} <$$

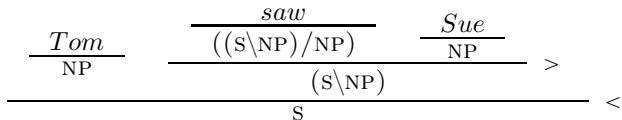
(38) Kim saw Sandy.

$$\frac{\frac{Kim}{NP} \quad \frac{\frac{saw}{((S \backslash NP) / NP)} \quad \frac{Sandy}{NP}}{(S \backslash NP)}}{S} <$$

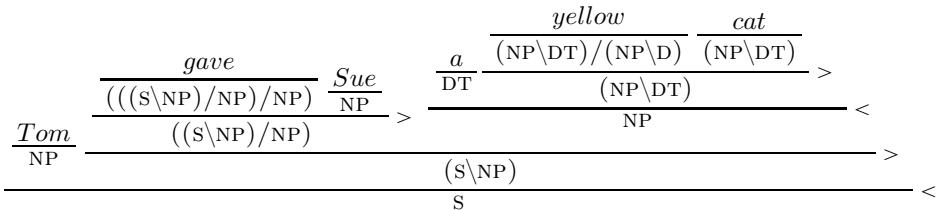
(39) Kim sneezed yesterday.

$$\frac{\frac{Kim}{NP} \quad \frac{\frac{sneezed}{(S \backslash NP)} \quad \frac{yesterday}{(S \backslash NP) \backslash (S \backslash NP)}}{(S \backslash NP)}}{S} <$$

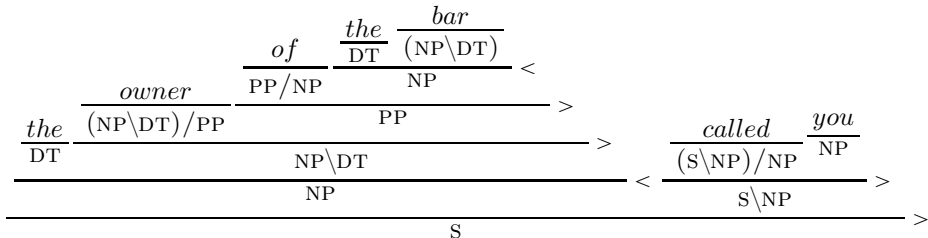
(40) Tom saw Sue.



(41) Tom gave Sue a yellow cat.



(42) The owner of the bar called you.



Coordination

and : $(X \backslash X) / X$

(43) Sue and Tom

$$\begin{array}{c}
 \frac{\frac{Sue}{NP} \quad \frac{\frac{and}{(NP \backslash NP) / NP} \quad \frac{Tom}{NP}}{NP \backslash NP}}{NP} \quad > <
 \end{array}$$

(44) Tom challenged and defeated Robin.

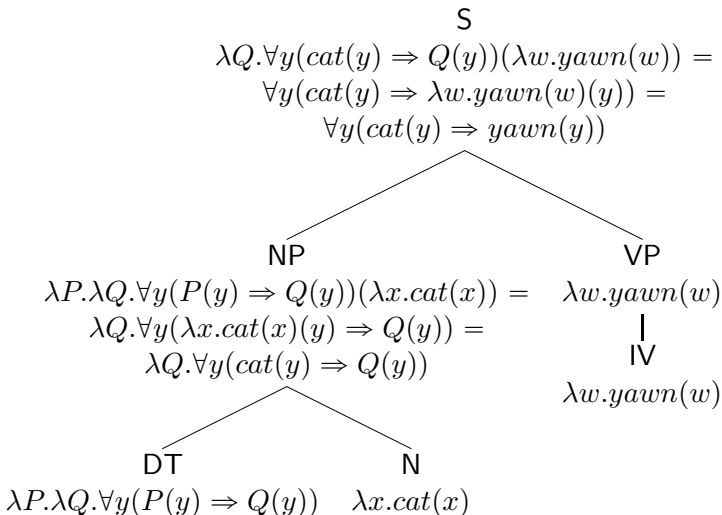
$$\begin{array}{c}
 \frac{\frac{Tom}{NP} \quad \frac{\frac{\frac{played}{(S \backslash NP) / NP} \quad \frac{\frac{\frac{and}{(((S \backslash NP) / NP) \backslash ((S \backslash NP) / NP)) / ((S \backslash NP) / NP)}{((S \backslash NP) / NP) \backslash ((S \backslash NP) / NP)}}{(S \backslash NP) / NP} \quad \frac{\frac{defeated}{(S \backslash NP) / NP}}{NP}}{S \backslash NP}}{S} \quad > <
 \end{array}$$

One of the great advantages of having access to syntactic structure is that this gives us some access to semantic structure.

- (45) a. $cat : N ; \lambda x.cat(x)$
b. $every : DT ; \lambda P.\lambda Q.\forall y(P(y) \Rightarrow Q(y))$
c. $a : DT ; \lambda P.\lambda Q.\exists y(P(y) \wedge Q(y))$
d. $yawn : IV ; \lambda w.yawn(w)$

- $\forall x(\phi)$ translates as 'for all values of x , ϕ is true'
- $\exists x(\phi)$ translates as 'for some value of x , ϕ is true'
- $P(x)$ translates as the value of x has property P
- $\phi \wedge \psi$ translates as ' ϕ and ψ are both true'
- $\phi \Rightarrow \psi$ translates as 'if ϕ is true then ψ is true too'
- $\lambda x.\phi$ translates as 'replace x 's in ϕ with a given argument'

- (46) a. $NP;\phi(\psi) \rightarrow DT;\phi \quad N;\psi$
b. $VP;\phi \rightarrow IV;\phi$
c. $S;\phi(\psi) \rightarrow NP;\phi \quad VP;\psi$



More words and rules:

- ① *Robin* : PN ; $\lambda P.P(\textit{robin})$
- ② *chased* : TV ; $\lambda R.\lambda z.R(\lambda w.\textit{chase}(z, w))$
- ① NP; $\phi \rightarrow$ PN; ϕ
- ② VP; $\phi(\psi) \rightarrow$ TV; ϕ NP; ψ

