COMMUNICATION AND LANGUAGE

CHAPTER 22

Outline

◇ Communication

◇ Grammar

◇ Syntactic analysis

◇ Problems

Communication

"Classical" view (pre-1953):
    language consists of sentences that are true/false (cf. logic)

"Modern" view (post-1953):
    language is a form of action

Wittgenstein (1953) **Philosophical Investigations**
Austin (1962) **How to Do Things with Words**
Searle (1969) **Speech Acts**

Why utter?

Communication

"Classical" view (pre-1953):
   language consists of sentences that are true/false (cf. logic)
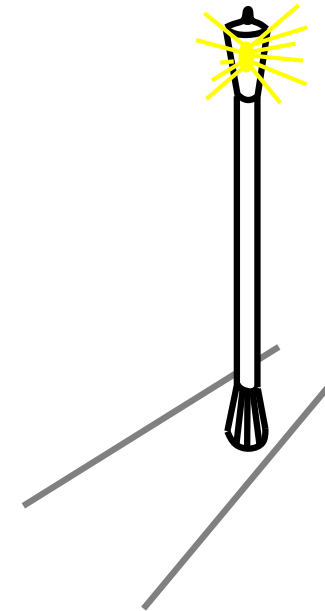
"Modern" view (post-1953):
   language is a form of action

Wittgenstein (1953) **Philosophical Investigations**
Austin (1962) **How to Do Things with Words**
Searle (1969) **Speech Acts**

Why utter?

Communication

"Classical" view (pre-1953):
   language consists of sentences that are true/false (cf. logic)

"Modern" view (post-1953):
   language is a form of action

Wittgenstein (1953) **Philosophical Investigations**
Austin (1962) **How to Do Things with Words**
Searle (1969) **Speech Acts**

Why utter?

Communication

"Classical" view (pre-1953):
  language consists of sentences that are true/false (cf. logic)
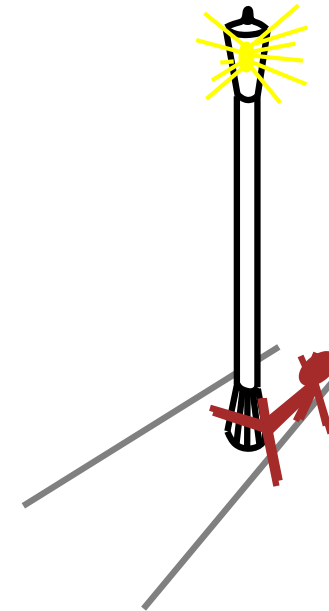
"Modern" view (post-1953):
  language is a form of action

Wittgenstein (1953) **Philosophical Investigations**
Austin (1962) **How to Do Things with Words**
Searle (1969) **Speech Acts**

Why utter?

To change the actions of other agents
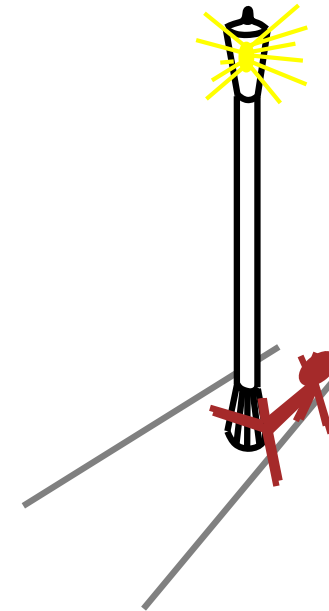
Speech acts


```
SITUATION

   Speaker  ──▶  Utterance  ──▶  Hearer
```

Speech acts achieve the speaker's goals:

Inform	"There's a pit in front of you"
Query	"Can you see the gold?"
Command	"Pick it up"
Promise	"I'll share the gold with you"
Acknowledge	"OK"

Speech act planning requires knowledge of
- Situation
- Semantic and syntactic conventions
- Hearer's goals, knowledge base, and rationality

Stages in communication (informing)

Intention   S wants to inform H that P

Generation   S selects words W to express P in context C

Synthesis   S utters words W

Perception   H perceives W' in context C'

Analysis   H infers possible meanings $P_1, \ldots P_n$

Disambiguation H infers intended meaning P_i

Incorporation   H incorporates P_i into KB

How could this go wrong?

Stages in communication (informing)

Intention   S wants to inform H that P
Generation   S selects words W to express P in context C
Synthesis   S utters words W

Perception   H perceives W' in context C'
Analysis   H infers possible meanings $P_1, \ldots P_n$
Disambiguation   H infers intended meaning P_i
Incorporation   H incorporates P_i into KB

How could this go wrong?
   – Insincerity (S doesn't believe P)
   – Speech wreck ignition failure
   – Ambiguous utterance
   – Differing understanding of current context ($C \neq C'$)

Grammar

Vervet monkeys, antelopes, etc. use isolated symbols for sentences
$\Rightarrow$ restricted set of communicable propositions, no generative capacity
(Chomsky (1957): **Syntactic Structures**)

Grammar specifies the compositional structure of complex messages
e.g., speech (linear), text (linear), music (two-dimensional)

A formal language is a set of strings of terminal symbols

Each string in the language can be analyzed/generated by the grammar

The grammar is a set of rewrite rules, e.g.,

$$S \rightarrow NP \; VP$$
$$Article \rightarrow \boldsymbol{the} \mid \boldsymbol{a} \mid \boldsymbol{an} \mid \ldots$$

Here S is the sentence symbol, NP, VP, and $Article$ are nonterminals

Grammar types

Regular: $nonterminal \rightarrow \boldsymbol{terminal}[nonterminal]$

$$S \rightarrow \boldsymbol{a}S$$
$$S \rightarrow \Lambda$$

Context-free: $nonterminal \rightarrow anything$

$$S \rightarrow \boldsymbol{a}S\boldsymbol{b}$$

Context-sensitive: more nonterminals on right-hand side

$$ASB \rightarrow AA\boldsymbol{a}BB$$

Recursively enumerable: no constraints

Natural languages probably context-free, parsable in real time!

Wumpus lexicon

$$Noun \rightarrow \textbf{stench} \mid \textbf{breeze} \mid \textbf{glitter} \mid \textbf{nothing}$$
$$\mid \textbf{wumpus} \mid \textbf{pit} \mid \textbf{pits} \mid \textbf{gold} \mid \textbf{east} \mid \ldots$$
$$Verb \rightarrow \textbf{is} \mid \textbf{see} \mid \textbf{smell} \mid \textbf{shoot} \mid \textbf{feel} \mid \textbf{stinks}$$
$$\mid \textbf{go} \mid \textbf{grab} \mid \textbf{carry} \mid \textbf{kill} \mid \textbf{turn} \mid \ldots$$
$$Adjective \rightarrow \textbf{right} \mid \textbf{left} \mid \textbf{east} \mid \textbf{south} \mid \textbf{back} \mid \textbf{smelly} \mid \ldots$$
$$Adverb \rightarrow \textbf{here} \mid \textbf{there} \mid \textbf{nearby} \mid \textbf{ahead}$$
$$\mid \textbf{right} \mid \textbf{left} \mid \textbf{east} \mid \textbf{south} \mid \textbf{back} \mid \ldots$$
$$Pronoun \rightarrow \textbf{me} \mid \textbf{you} \mid \textbf{I} \mid \textbf{it} \mid \ldots$$
$$Name \rightarrow \textbf{John} \mid \textbf{Mary} \mid \textbf{Boston} \mid \textbf{UCB} \mid \textbf{PAJC} \mid \ldots$$
$$Article \rightarrow \textbf{the} \mid \textbf{a} \mid \textbf{an} \mid \ldots$$
$$Preposition \rightarrow \textbf{to} \mid \textbf{in} \mid \textbf{on} \mid \textbf{near} \mid \ldots$$
$$Conjunction \rightarrow \textbf{and} \mid \textbf{or} \mid \textbf{but} \mid \ldots$$

Wumpus lexicon

$$Noun \rightarrow \textbf{stench} \mid \textbf{breeze} \mid \textbf{glitter} \mid \textbf{nothing}$$
$$\mid \textbf{wumpus} \mid \textbf{pit} \mid \textbf{pits} \mid \textbf{gold} \mid \textbf{east} \mid \ldots$$
$$Verb \rightarrow \textbf{is} \mid \textbf{see} \mid \textbf{smell} \mid \textbf{shoot} \mid \textbf{feel} \mid \textbf{stinks}$$
$$\mid \textbf{go} \mid \textbf{grab} \mid \textbf{carry} \mid \textbf{kill} \mid \textbf{turn} \mid \ldots$$
$$Adjective \rightarrow \textbf{right} \mid \textbf{left} \mid \textbf{east} \mid \textbf{south} \mid \textbf{back} \mid \textbf{smelly} \mid \ldots$$
$$Adverb \rightarrow \textbf{here} \mid \textbf{there} \mid \textbf{nearby} \mid \textbf{ahead}$$
$$\mid \textbf{right} \mid \textbf{left} \mid \textbf{east} \mid \textbf{south} \mid \textbf{back} \mid \ldots$$
$$Pronoun \rightarrow \textbf{me} \mid \textbf{you} \mid \textbf{I} \mid \textbf{it} \mid \ldots$$
$$Name \rightarrow \textbf{John} \mid \textbf{Mary} \mid \textbf{Boston} \mid \textbf{UCB} \mid \textbf{PAJC} \mid \ldots$$
$$Article \rightarrow \textbf{the} \mid \textbf{a} \mid \textbf{an} \mid \ldots$$
$$Preposition \rightarrow \textbf{to} \mid \textbf{in} \mid \textbf{on} \mid \textbf{near} \mid \ldots$$
$$Conjunction \rightarrow \textbf{and} \mid \textbf{or} \mid \textbf{but} \mid \ldots$$
$$Digit \rightarrow \textbf{0} \mid \textbf{1} \mid \textbf{2} \mid \textbf{3} \mid \textbf{4} \mid \textbf{5} \mid \textbf{6} \mid \textbf{7} \mid \textbf{8} \mid \textbf{9}$$

Closed classes are small, bounded, change very slowly

Wumpus lexicon

$$Noun \rightarrow \textbf{stench} \mid \textbf{breeze} \mid \textbf{glitter} \mid \textbf{nothing}$$
$$\mid \textbf{wumpus} \mid \textbf{pit} \mid \textbf{pits} \mid \textbf{gold} \mid \textbf{east} \mid \ldots$$
$$Verb \rightarrow \textbf{is} \mid \textbf{see} \mid \textbf{smell} \mid \textbf{shoot} \mid \textbf{feel} \mid \textbf{stinks}$$
$$\mid \textbf{go} \mid \textbf{grab} \mid \textbf{carry} \mid \textbf{kill} \mid \textbf{turn} \mid \ldots$$
$$Adjective \rightarrow \textbf{right} \mid \textbf{left} \mid \textbf{east} \mid \textbf{south} \mid \textbf{back} \mid \textbf{smelly} \mid \ldots$$
$$Adverb \rightarrow \textbf{here} \mid \textbf{there} \mid \textbf{nearby} \mid \textbf{ahead}$$
$$\mid \textbf{right} \mid \textbf{left} \mid \textbf{east} \mid \textbf{south} \mid \textbf{back} \mid \ldots$$
$$Pronoun \rightarrow \textbf{me} \mid \textbf{you} \mid \textbf{I} \mid \textbf{it} \mid \textbf{\st{thou}} \mid \underline{\textbf{y'all}} \ldots$$
$$Name \rightarrow \textbf{John} \mid \textbf{Mary} \mid \textbf{Boston} \mid \textbf{UCB} \mid \textbf{PAJC} \mid \ldots$$
$$Article \rightarrow \textbf{the} \mid \textbf{a} \mid \textbf{an} \mid \ldots$$
$$Preposition \rightarrow \textbf{to} \mid \textbf{in} \mid \textbf{on} \mid \textbf{near} \mid \ldots$$
$$Conjunction \rightarrow \textbf{and} \mid \textbf{or} \mid \textbf{but} \mid \ldots$$
$$Digit \rightarrow \textbf{0} \mid \textbf{1} \mid \textbf{2} \mid \textbf{3} \mid \textbf{4} \mid \textbf{5} \mid \textbf{6} \mid \textbf{7} \mid \textbf{8} \mid \textbf{9}$$

Closed classes are small, bounded, change very slowly

Wumpus lexicon

$Noun \rightarrow$ **stench** | **breeze** | **glitter** | **nothing**
| **wumpus** | **pit** | **pits** | **gold** | **east** | ...

$Verb \rightarrow$ **is** | **see** | **smell** | **shoot** | **feel** | **stinks**
| **go** | **grab** | **carry** | **kill** | **turn** | ...

$Adjective \rightarrow$ **right** | **left** | **east** | **south** | **back** | **smelly** | ...

$Adverb \rightarrow$ **here** | **there** | **nearby** | **ahead**
| **right** | **left** | **east** | **south** | **back** | ...

$Pronoun \rightarrow$ **me** | **you** | **I** | **it** | ...

$Name \rightarrow$ **John** | **Mary** | **Boston** | **UCB** | **PAJC** | ...

$Article \rightarrow$ **the** | **a** | **an** | ...

$Preposition \rightarrow$ **to** | **in** | **on** | **near** | ...

$Conjunction \rightarrow$ **and** | **or** | **but** | ...

Open classes are large, unbounded, change very fast

Wumpus lexicon

$$Noun \rightarrow \textbf{stench} \mid \textbf{breeze} \mid \textbf{glitter} \mid \textbf{nothing}$$
$$\mid \textbf{wumpus} \mid \textbf{pit} \mid \textbf{pits} \mid \textbf{gold} \mid \textbf{east} \mid \ldots$$
$$Verb \rightarrow \textbf{is} \mid \textbf{see} \mid \textbf{smell} \mid \textbf{shoot} \mid \textbf{feel} \mid \textbf{stinks}$$
$$\mid \textbf{go} \mid \textbf{grab} \mid \textbf{carry} \mid \textbf{kill} \mid \textbf{turn} \mid \underline{google} \ldots$$
$$Adjective \rightarrow \textbf{right} \mid \textbf{left} \mid \textbf{east} \mid \textbf{south} \mid \textbf{back} \mid \textbf{smelly} \mid \ldots$$
$$Adverb \rightarrow \textbf{here} \mid \textbf{there} \mid \textbf{nearby} \mid \textbf{ahead}$$
$$\mid \textbf{right} \mid \textbf{left} \mid \textbf{east} \mid \textbf{south} \mid \textbf{back} \mid \ldots$$
$$Pronoun \rightarrow \textbf{me} \mid \textbf{you} \mid \textbf{I} \mid \textbf{it} \mid \ldots$$
$$Name \rightarrow \textbf{John} \mid \textbf{Mary} \mid \textbf{Boston} \mid \textbf{UCB} \mid \textbf{PAJC} \mid \underline{Google} \ldots$$
$$Article \rightarrow \textbf{the} \mid \textbf{a} \mid \textbf{an} \mid \ldots$$
$$Preposition \rightarrow \textbf{to} \mid \textbf{in} \mid \textbf{on} \mid \textbf{near} \mid \ldots$$
$$Conjunction \rightarrow \textbf{and} \mid \textbf{or} \mid \textbf{but} \mid \ldots$$

Open classes are large, unbounded, change very fast

Wumpus grammar

$S \rightarrow NP\ VP$ I + feel a breeze

$|\ S\ Conjunction\ S$ I feel a breeze + and + I smell a wumpus

$NP \rightarrow Pronoun$ I

$|\ Noun$ pits

$|\ Article\ Noun$ the + wumpus

$|\ Digit\ Digit$ 3 4

$|\ NP\ PP$ the wumpus + to the east

$|\ NP\ RelClause$ the wumpus + that is smelly

$VP \rightarrow Verb$ stinks

$|\ VP\ NP$ feel + a breeze

$|\ VP\ Adjective$ is + smelly

$|\ VP\ PP$ turn + to the east

$|\ VP\ Adverb$ go + ahead

$PP \rightarrow Preposition\ NP$ to + the east

$RelClause \rightarrow \mathbf{that}\ VP$ that + is smelly

Grammaticality judgements

Formal language L_1 may differ from natural language L_2
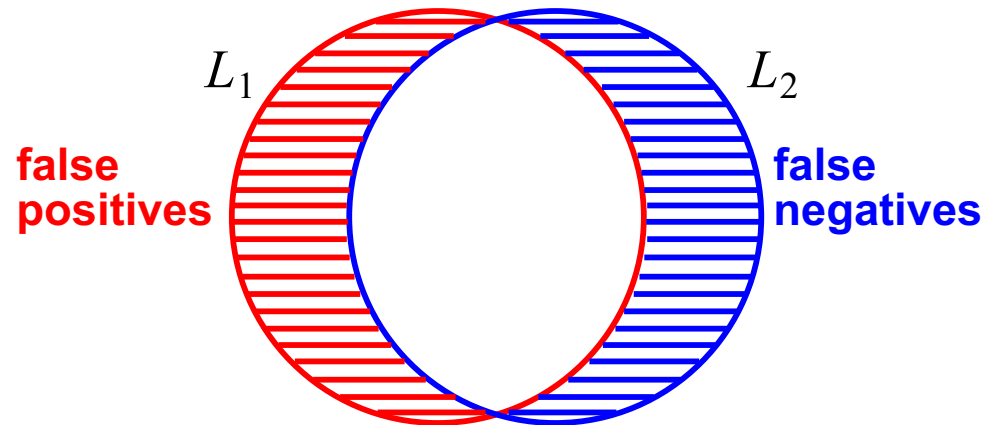
Adjusting L_1 to agree with L_2 is a learning problem!

* the gold grab the wumpus
* I smell the wumpus the gold
  I give the wumpus the gold
* I donate the wumpus the gold

Intersubjective agreement somewhat reliable, independent of semantics!
Real grammars 10–500 pages, insufficient even for "proper" English

Parse trees

Exhibit the grammatical structure of a sentence

I   **shoot**   **the**   **wumpus**

Parse trees

Exhibit the grammatical structure of a sentence

Pronoun　　　**Verb**　　　**Article**　　　**Noun**

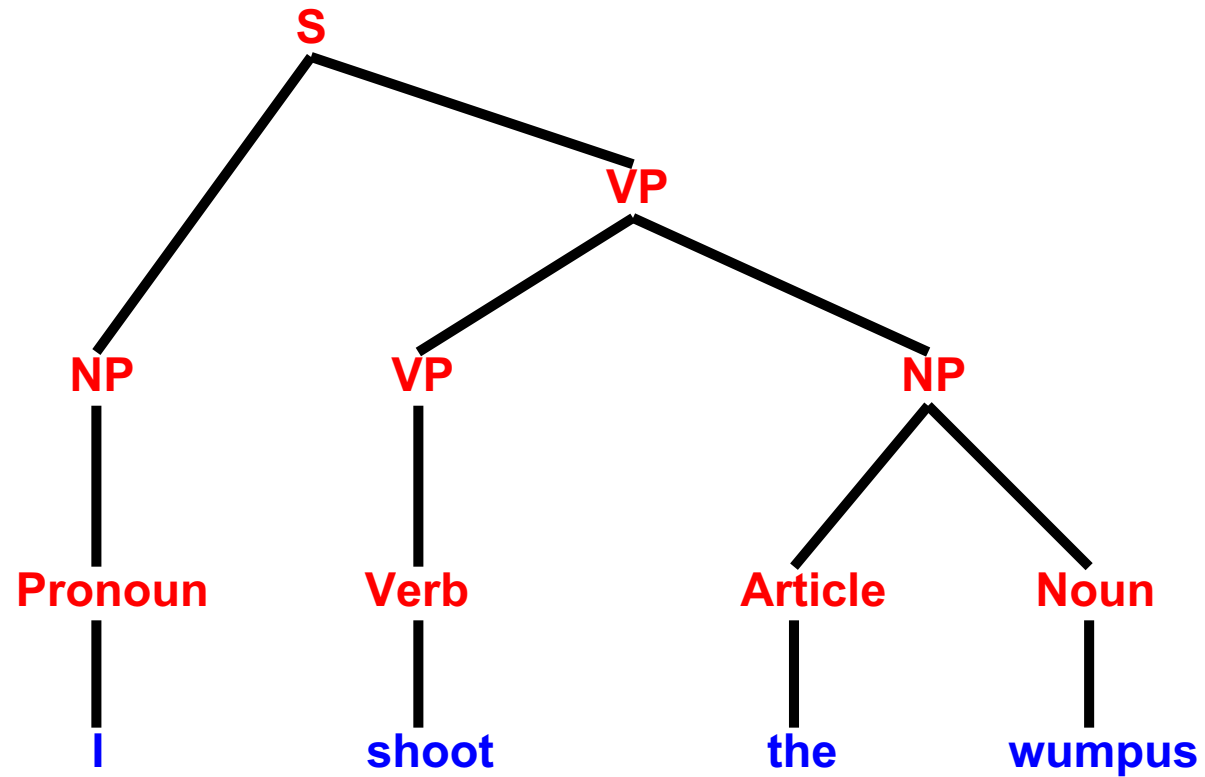I　　　shoot　　　the　　　wumpus

Parse trees

Exhibit the grammatical structure of a sentence
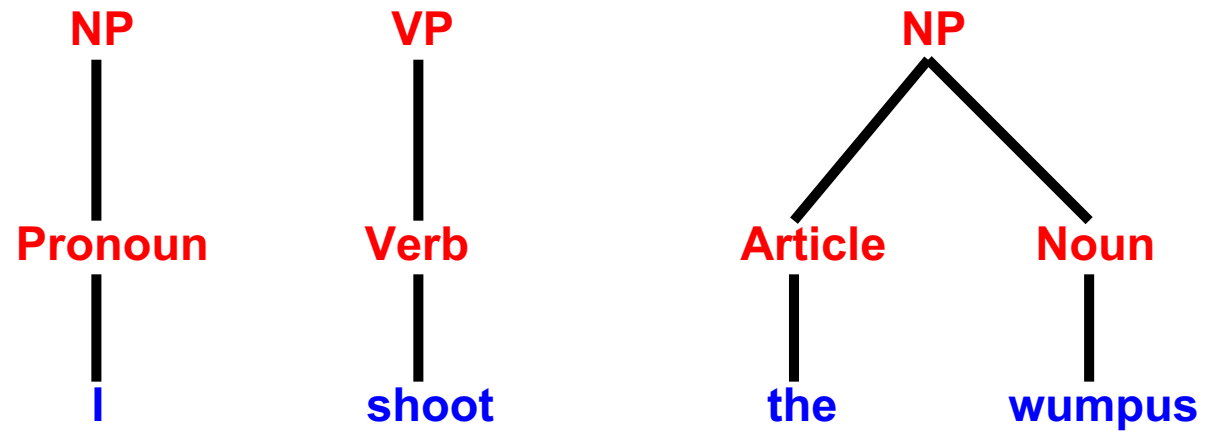
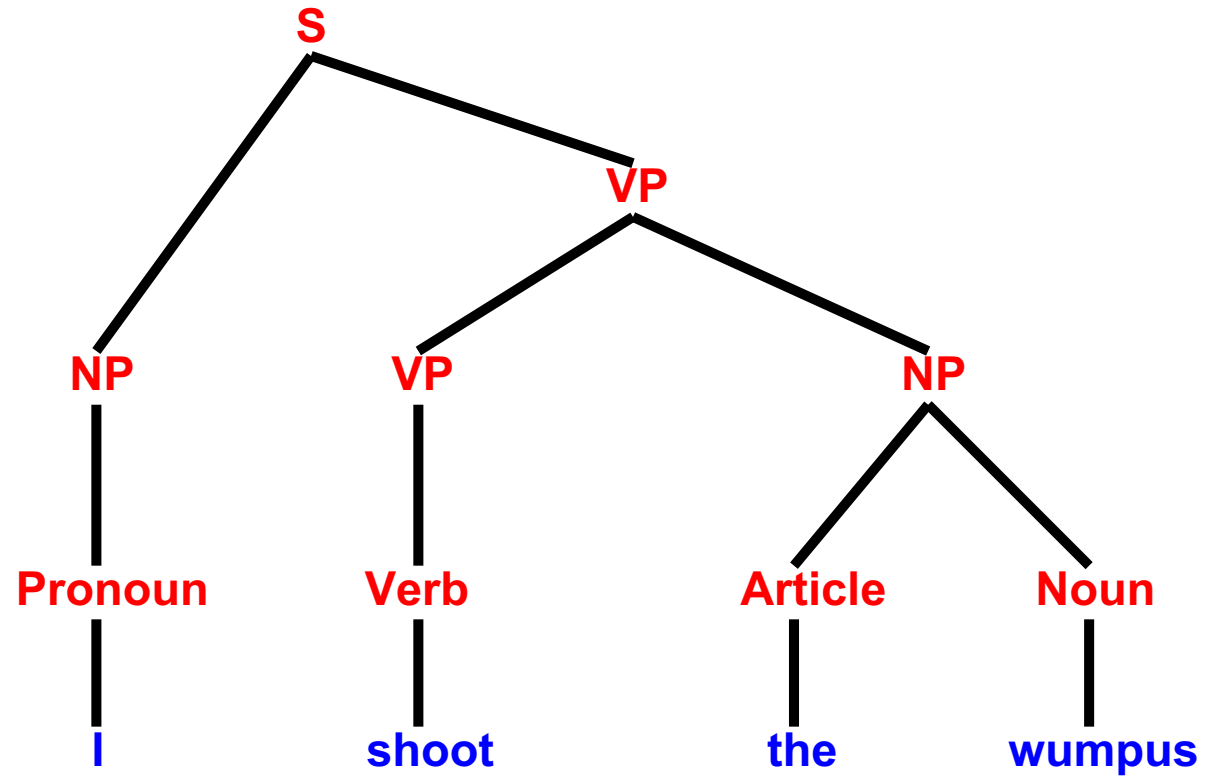# Parse trees

Exhibit the grammatical structure of a sentence
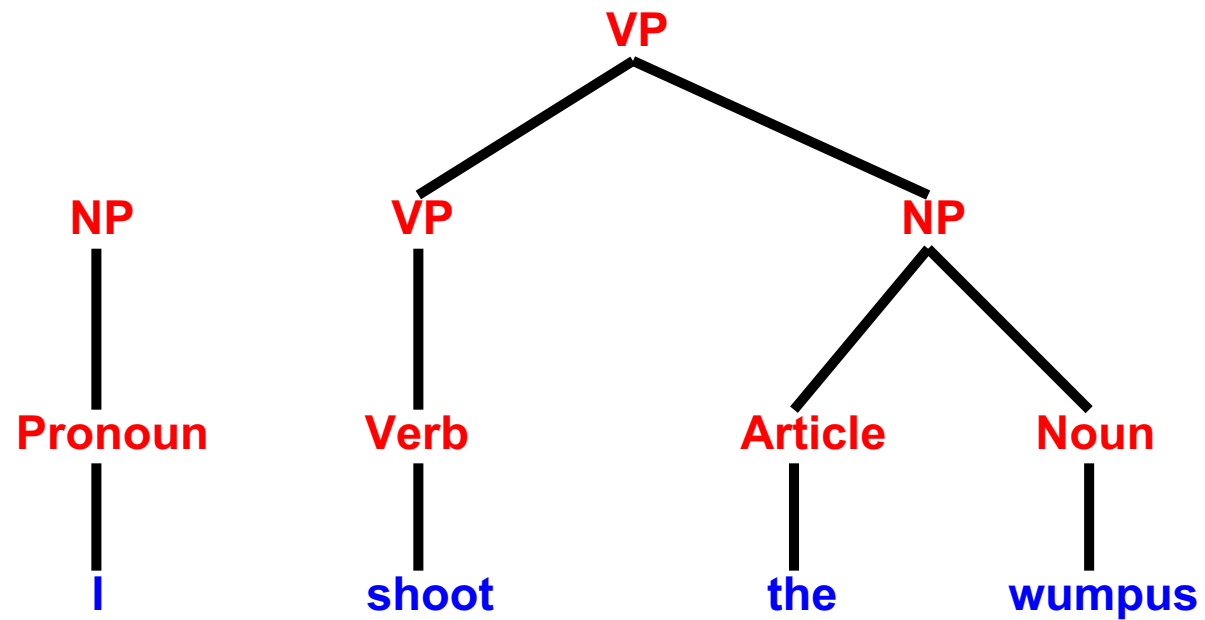
Parse trees

Exhibit the grammatical structure of a sentence

Parse trees

Exhibit the grammatical structure of a sentence

Efficient CFG algorithms (e.g., chart parsing, Section 22.3) $O(n^3)$

Syntax in NLP

Most view syntactic structure as an essential step towards meaning;
  "Mary hit John" $\neq$ "John hit Mary"

Nonetheless, ungrammatical sentence may be understood:

Syntax in NLP

Most view syntactic structure as an essential step towards meaning;
"Mary hit John" $\neq$ "John hit Mary"

Nonetheless, ungrammatical sentence may be understood:

"Georgie give Georgie breakfast dinosaur!! Dinosaur brush teeth!!!"

Syntax in NLP

Most view syntactic structure as an essential step towards meaning;

"Mary hit John" $\neq$ "John hit Mary"

Nonetheless, ungrammatical sentence may be understood:

"Georgie give Georgie breakfast to dinosaur!! Need teeth brush!!!"

Not all grammatical sentences are easy to understand:

"Wouldn't the sentence 'I want to put a hyphen between the words Fish and And and And and Chips in my Fish-And-Chips sign' have been clearer if quotation marks had been placed before Fish, and between Fish and and, and and and And, and And and and, and and and And, and And and and, and and and Chips, as well as after Chips?"