Understanding Sentences

Informatics 1 CG: Lecture 14

Mirella Lapata

School of Informatics
University of Edinburgh
mlap@inf.ed.ac.uk

February 11, 2016

Reading:

Trevor Harley (2001). The Psychology of Language, Chapter 9

Recap: The Associationist View of Word Meaning

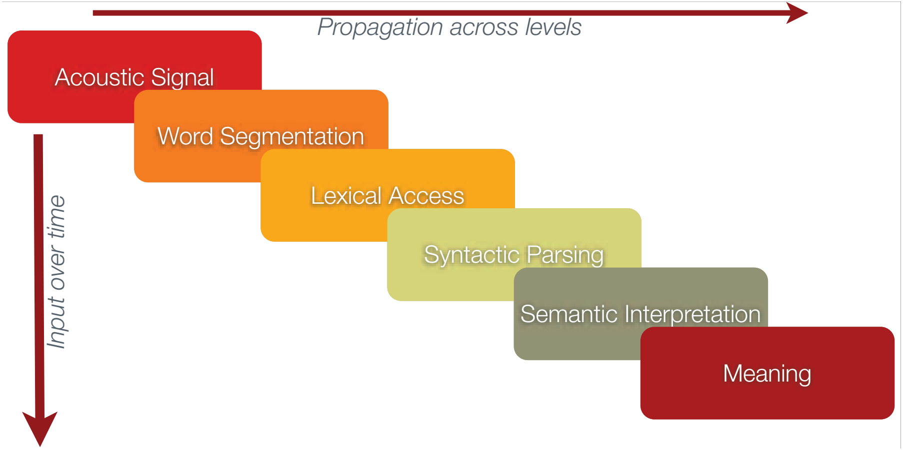
- You shall know a word by the company it keeps (Firth, 1957).
- Distributional hypothesis about word meaning: the context surrounding a given word provides information about its meaning.
- Experimental evidence indicates that the cognitive system is sensitive to distributional information.
- Construction of vector spaces.
- Latent Semantic Analysis.

Recap: The Associationist View of Word Meaning

- You shall know a word by the company it keeps (Firth, 1957).
- Distributional hypothesis about word meaning: the context surrounding a given word provides information about its meaning.
- Experimental evidence indicates that the cognitive system is sensitive to distributional information.
- Construction of vector spaces.
- Latent Semantic Analysis.

What happens after we recognize a word? How do we put together words to form meaningful sentences?

Sound to Meaning



Parsing

How do we understand a sentence? First step is to "parse" it.

- Parsing takes place unconsciously (involves finding the subjects, verbs, objects and so on).
- A parser is the mental program that analyzes sentence structure during language comprehension.
- Ultimately parsing helps us to interpret sentences during language comprehension.

1. The ghost chased the vampire.
2. The vampire chased the ghost.
3. The vampire was chased by the ghost.

Parsing

How do we understand a sentence? First step is to "parse" it.

- Parsing takes place unconsciously (involves finding the subjects, verbs, objects and so on).
- A parser is the mental program that analyzes sentence structure during language comprehension.
- Ultimately parsing helps us to interpret sentences during language comprehension.

1. The ghost chased the vampire.
2. The vampire chased the ghost.
3. The vampire was chased by the ghost.

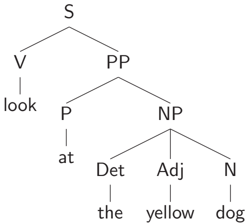
Parsing

How do we understand a sentence? First step is to “parse” it.

- Parsing takes place unconsciously (involves finding the subjects, verbs, objects and so on).
- A parser is the mental program that analyzes sentence structure during language comprehension.
- Ultimately parsing helps us to interpret sentences during language comprehension.

1. The ghost chased the vampire.
2. The vampire chased the ghost.
3. The vampire was chased by the ghost.

⇒ look at the yellow dog ⇒



Today we will look at parsing, which turns a sequence of words into a syntactic representation.

Syntactic representations make explicit how the words in a sentence relate to each other.

Phrase Structure Grammar

In order to build syntactic representations, we need a grammar. The simplest type of grammar is a context-free grammar:

Phrasal categories:
S: sentence, NP: noun phrase, VP: verb phrase

Lexical categories (aka parts of speech):
Det: determiner, N: noun, V: verb

Phrase structure rules:

S	$\rightarrow$	NP VP	Det	$\rightarrow$	the
NP	$\rightarrow$	Det N	N	$\rightarrow$	kitten
VP	$\rightarrow$	V NP	N	$\rightarrow$	dog
			V	$\rightarrow$	bit

Phrase Structure Grammar

In order to build syntactic representations, we need a grammar.
The simplest type of grammar is a context-free grammar:

Phrasal categories:
S: sentence, NP: noun phrase, VP: verb phrase

Lexical categories (aka parts of speech):
Det: determiner, N: noun, V: verb

Phrase structure rules:

$$\begin{aligned} S &\rightarrow NP\ VP \\ NP &\rightarrow Det\ N \\ VP &\rightarrow V\ NP \end{aligned}$$

$$\begin{aligned} Det &\rightarrow the \\ N &\rightarrow kitten \\ N &\rightarrow dog \\ V &\rightarrow bit \end{aligned}$$

Phrase Structure Grammar

In order to build syntactic representations, we need a grammar. The simplest type of grammar is a context-free grammar:

Phrasal categories:

S: sentence, NP: noun phrase, VP: verb phrase

Lexical categories (aka parts of speech):

Det: determiner, N: noun, V: verb

Phrase structure rules:

S	$\rightarrow$	NP VP	Det	$\rightarrow$	the
NP	$\rightarrow$	Det N	N	$\rightarrow$	kitten
VP	$\rightarrow$	V NP	N	$\rightarrow$	dog
			V	$\rightarrow$	bit

Phrase Structure Grammar

In order to build syntactic representations, we need a grammar. The simplest type of grammar is a context-free grammar:

Phrasal categories:
S: sentence, NP: noun phrase, VP: verb phrase

Lexical categories (aka parts of speech):
Det: determiner, N: noun, V: verb

Phrase structure rules:

S	$\rightarrow$	NP VP	Det	$\rightarrow$	the
NP	$\rightarrow$	Det N	N	$\rightarrow$	kitten
VP	$\rightarrow$	V NP	N	$\rightarrow$	dog
			V	$\rightarrow$	bit

Derivations and Syntax Trees

A derivation is the sequence of strings that results from applying a sequence of grammar rules:

Derivation

S $\longrightarrow$ NP VP $\longrightarrow$ NP V NP $\longrightarrow$ NP V Det N $\longrightarrow$ NP bit Det N $\longrightarrow$ NP bit Det dog $\longrightarrow$ NP bit the dog $\longrightarrow$ Det N bit the dog $\longrightarrow$ the N bit the dog $\longrightarrow$ the kitten bit the dog

Derivations and Syntax Trees

A derivation is the sequence of strings that results from applying a sequence of grammar rules:

Derivation

S $\longrightarrow$ NP VP $\longrightarrow$ NP V NP $\longrightarrow$ NP V Det N $\longrightarrow$ NP bit Det N $\longrightarrow$ NP bit Det dog $\longrightarrow$ NP bit the dog $\longrightarrow$ Det N bit the dog $\longrightarrow$ the N bit the dog $\longrightarrow$ the kitten bit the dog

Derivations and Syntax Trees

A derivation is the sequence of strings that results from applying a sequence of grammar rules:

Derivation

S $\longrightarrow$ NP VP $\longrightarrow$ NP V NP $\longrightarrow$ NP V Det N $\longrightarrow$ NP bit Det N $\longrightarrow$ NP bit Det dog $\longrightarrow$ NP bit the dog $\longrightarrow$ Det N bit the dog $\longrightarrow$ the N bit the dog $\longrightarrow$ the kitten bit the dog

Derivations and Syntax Trees

A derivation is the sequence of strings that results from applying a sequence of grammar rules:

Derivation

S $\rightarrow$ NP VP $\rightarrow$ NP V NP $\rightarrow$ NP V Det N $\rightarrow$ NP bit Det N $\rightarrow$ NP bit Det dog $\rightarrow$ NP bit the dog $\rightarrow$ Det N bit the dog $\rightarrow$ the N bit the dog $\rightarrow$ the kitten bit the dog

Derivations and Syntax Trees

A derivation is the sequence of strings that results from applying a sequence of grammar rules:

Derivation

S $\longrightarrow$ NP VP $\longrightarrow$ NP V NP $\longrightarrow$ NP V Det N $\longrightarrow$ NP bit Det N $\longrightarrow$ NP bit Det dog $\longrightarrow$ NP bit the dog $\longrightarrow$ Det N bit the dog $\longrightarrow$ the N bit the dog $\longrightarrow$ the kitten bit the dog

Derivations and Syntax Trees

A derivation is the sequence of strings that results from applying a sequence of grammar rules:

Derivation

S $\longrightarrow$ NP VP $\longrightarrow$ NP V NP $\longrightarrow$ NP V Det N $\longrightarrow$ NP bit Det N $\longrightarrow$ NP bit Det dog $\longrightarrow$ NP bit the dog $\longrightarrow$ Det N bit the dog $\longrightarrow$ the N bit the dog $\longrightarrow$ the kitten bit the dog

Derivations and Syntax Trees

A derivation is the sequence of strings that results from applying a sequence of grammar rules:

Derivation

S $\longrightarrow$ NP VP $\longrightarrow$ NP V NP $\longrightarrow$ NP V Det N $\longrightarrow$ NP bit Det N $\longrightarrow$ NP bit Det dog $\longrightarrow$ NP bit the dog $\longrightarrow$ Det N bit the dog $\longrightarrow$ the N bit the dog $\longrightarrow$ the kitten bit the dog

Derivations and Syntax Trees

A derivation is the sequence of strings that results from applying a sequence of grammar rules:

Derivation

S → NP VP → NP V NP → NP V Det N → NP bit Det N → NP bit Det dog → NP bit the dog → Det N bit the dog → the N bit the dog → the kitten bit the dog

Derivations and Syntax Trees

A derivation is the sequence of strings that results from applying a sequence of grammar rules:

Derivation

S $\longrightarrow$ NP VP $\longrightarrow$ NP V NP $\longrightarrow$ NP V Det N $\longrightarrow$ NP bit Det N $\longrightarrow$ NP bit Det dog $\longrightarrow$ NP bit the dog $\longrightarrow$ Det N bit the dog $\longrightarrow$ the N bit the dog $\longrightarrow$ the kitten bit the dog

Derivations and Syntax Trees

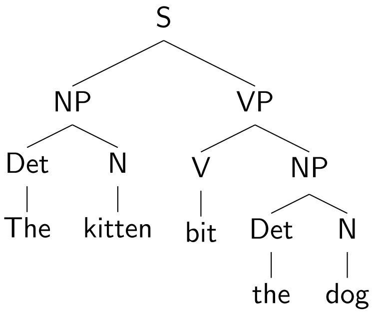
A derivation is the sequence of strings that results from applying a sequence of grammar rules:

Derivation

S $\rightarrow$ NP VP $\rightarrow$ NP V NP $\rightarrow$ NP V Det N $\rightarrow$ NP bit Det N $\rightarrow$ NP bit Det dog $\rightarrow$ NP bit the dog $\rightarrow$ Det N bit the dog $\rightarrow$ the N bit the dog $\rightarrow$ the kitten bit the dog

Derivations and Syntax Trees

Derivations are represented as syntax trees:



Syntactic ambiguity: a sentence can have multiple syntax trees.
These correspond to different interpretations.

Human Sentence Processing

A parser takes a sentence and computes a syntax tree for it, given a grammar. This is a prerequisite for assigning an interpretation to the sentence.

The cognitive device that performs syntactic parsing is called human sentence processing mechanism (HSPM).

Parsing is incremental: the HSPM builds structures word by word as the input arrives.

But what if more than one structure is compatible with the input:

- at the current point but not later: local ambiguity;
- for the input overall: global ambiguity.

Global Ambiguity

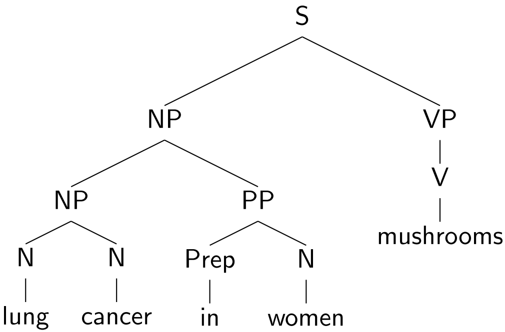
Given a grammar, strings that have more than one complete syntax tree (parse) are said to have global structural ambiguity.

Examples

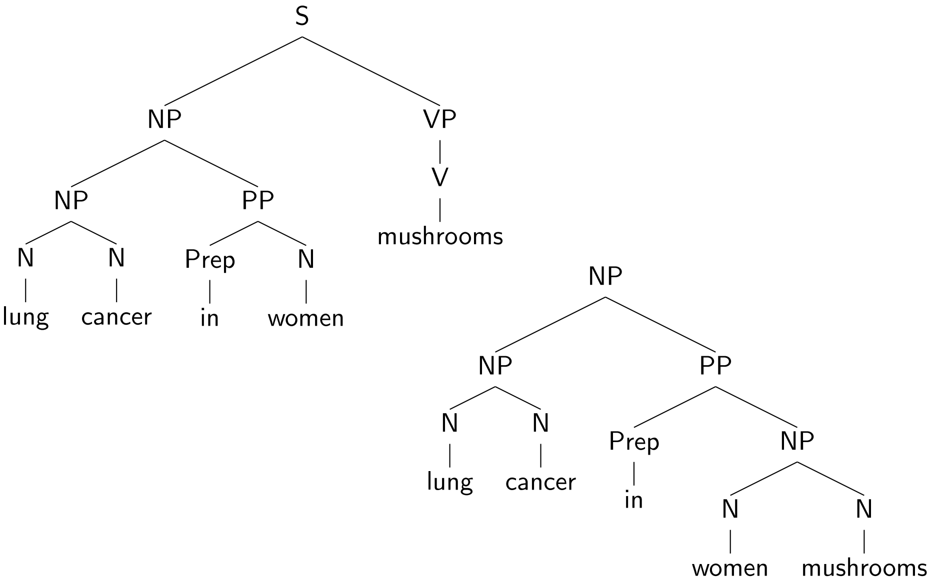
1. She sat on the chair covered in dust.
2. I put the book on the table in the kitchen.
3. Lung cancer in women mushrooms.

Examples from `http://www.fun-with-words.com/ambiguous_garden_path.html`

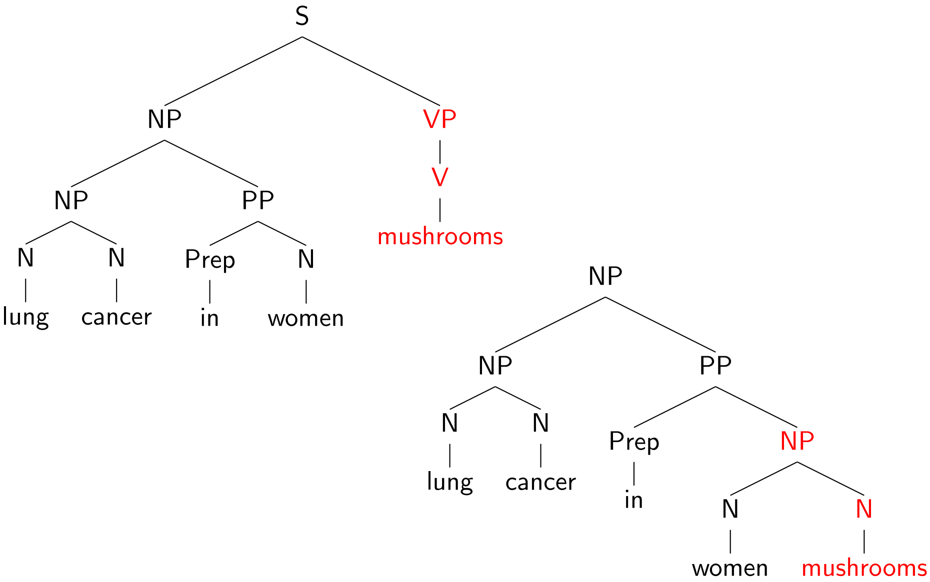
Global Ambiguity



Global Ambiguity



Global Ambiguity



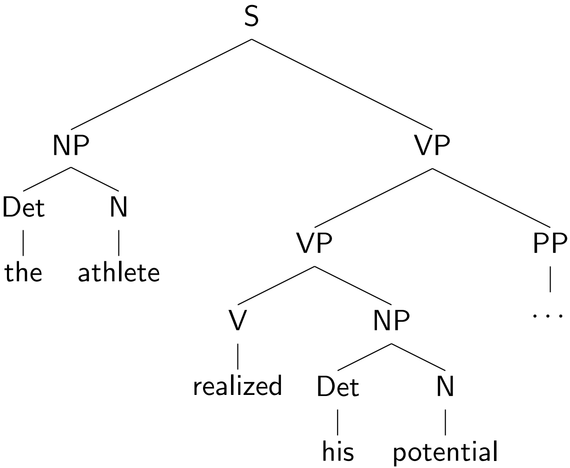
Local Ambiguity

When only an initial substring is structurally ambiguous, the sentence is said to have local ambiguity; once the remainder of the string is known, only one tree remains possible.

Example

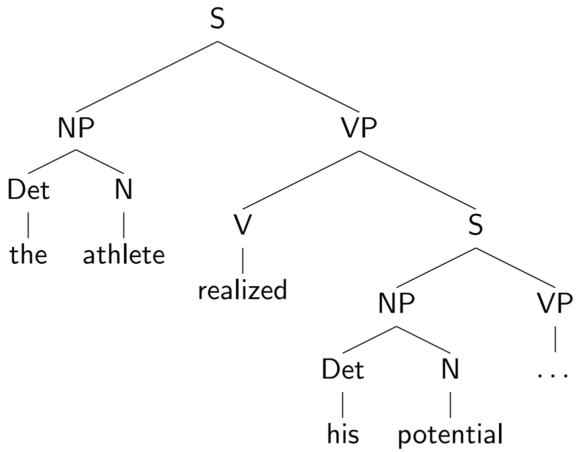
The athlete realized his potential ...
a. ... at the competition.
b. ... could make him a world-class sprinter

Local Ambiguity: Structure 1 (VP $\rightarrow$ V NP)



The athlete realized his potential at the competition.

Local Ambiguity: Structure 1 (VP → V S)



The athlete realized his potential could make him a world-class sprinter.

Garden Paths

This is an example of a garden path:

- both structures are compatible with the input up until *potential*; only the next word disambiguates;
- however, the processor commits to a single (wrong) structure early on, and trips up when later input is inconsistent with that structure;
- presumably, the processor now has to compute a new structure that is consistent with the input;
- garden path sentences result in longer reading times, reverse eye-movements, lower comprehension accuracies, etc.;
- some garden paths are so strong that the parser fails to recover from them.

Garden Paths

More examples of garden paths

1. I convinced her children are noisy.
2. Until the police arrest the drug dealers control the street.
3. The old man the boat.
4. Fat people eat accumulates .
5. The cotton clothing is usually made of grows in Mississippi.
6. The prime number few.

Examples from `http://www.fun-with-words.com/ambiguous_garden_path.html`

Garden Paths

More examples of garden paths

1. I convinced her that children are noisy.
2. Until the police arrest the drug dealers control the street.
3. The old man the boat.
4. Fat people eat accumulates .
5. The cotton clothing is usually made of grows in Mississippi.
6. The prime number few.

Examples from `http://www.fun-with-words.com/ambiguous_garden_path.html`

Garden Paths

More examples of garden paths

1. I convinced her children are noisy.
2. Until the police make the arrest , the drug dealers control the street.
3. The old man the boat.
4. Fat people eat accumulates .
5. The cotton clothing is usually made of grows in Mississippi.
6. The prime number few.

Examples from `http://www.fun-with-words.com/ambiguous_garden_path.html`

Garden Paths

More examples of garden paths

1. I convinced her children are noisy.
2. Until the police arrest the drug dealers control the street.
3. The old people man the boat.
4. Fat people eat accumulates .
5. The cotton clothing is usually made of grows in Mississippi.
6. The prime number few.

Examples from `http://www.fun-with-words.com/ambiguous_garden_path.html`

Garden Paths

More examples of garden paths

1. I convinced her children are noisy.
2. Until the police arrest the drug dealers control the street.
3. The old man the boat.
4. The Fat that people eat accumulates in their bodies.
5. The cotton clothing is usually made of grows in Mississippi.
6. The prime number few.

Examples from `http://www.fun-with-words.com/ambiguous_garden_path.html`

Garden Paths

More examples of garden paths

1. I convinced her children are noisy.
2. Until the police arrest the drug dealers control the street.
3. The old man the boat.
4. Fat people eat accumulates .
5. The cotton that clothing is usually made of grows in Mississippi.
6. The prime number few.

Examples from `http://www.fun-with-words.com/ambiguous_garden_path.html`

Garden Paths

More examples of garden paths

1. I convinced her children are noisy.
2. Until the police arrest the drug dealers control the street.
3. The old man the boat.
4. Fat people eat accumulates .
5. The cotton clothing is usually made of grows in Mississippi.
6. The prime people number few.

Examples from http://www.fun-with-words.com/ambiguous_garden_path.html

Two Theories of Human Parsing

1. What mechanism is used to construct interpretations?
2. What information is used to determine preferred structure?

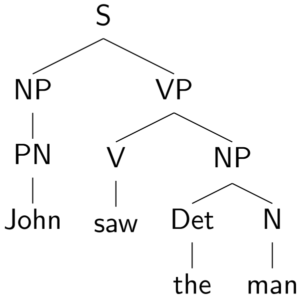
Garden Path Theory

Parsing is autonomous and takes place in two stages; during stage 1 only syntactic information is taken into account; stage 2 takes additional information sources into account if single analysis turns out to be incorrect.

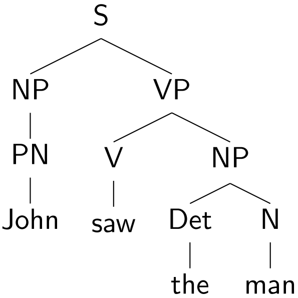
Constraint-based Theory

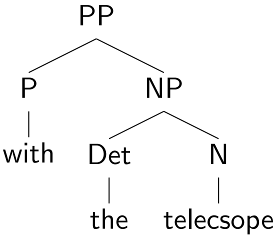
Parsing is interactive and takes place in one stage. Processor uses multiple sources of information at once, structure most supported by constraints is active, plausible alternatives also remain active.

The Garden Path Theory (Frazier, 1987)

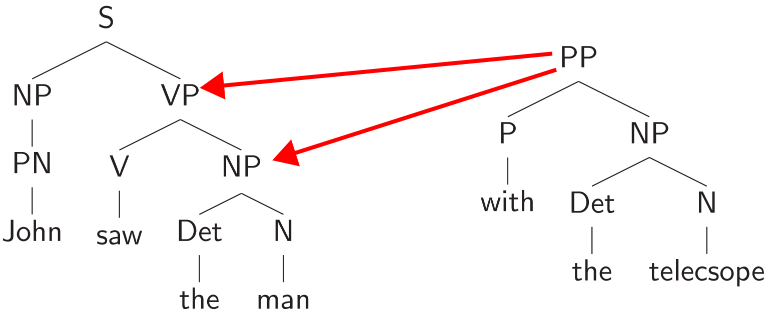


The Garden Path Theory (Frazier, 1987)



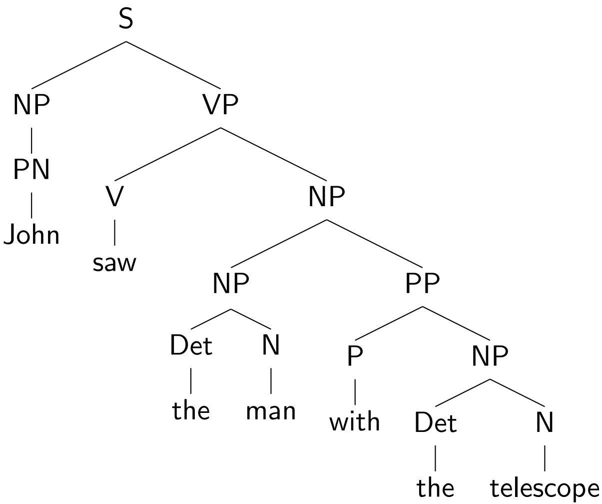


The Garden Path Theory (Frazier, 1987)



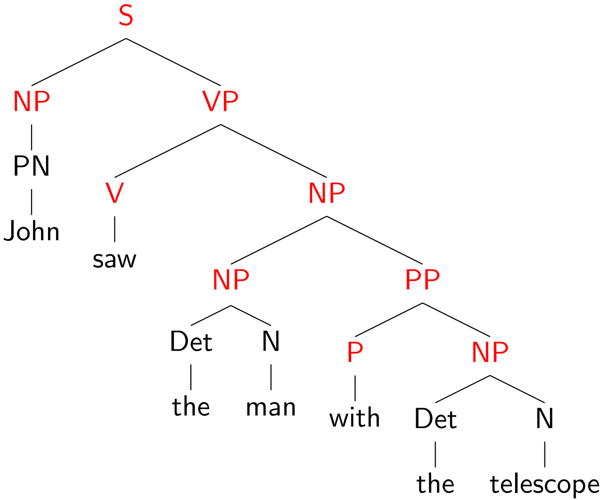
Which attachment do people initially prefer?

First Strategy: Minimal Attachment



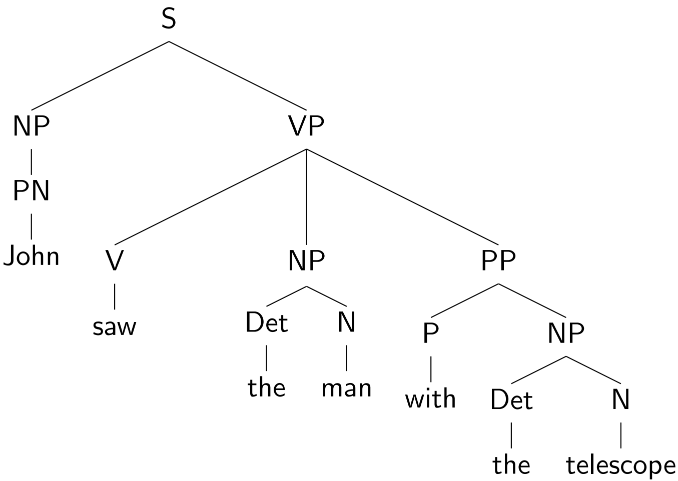
Adopt structure containing fewest number of nodes

First Strategy: Minimal Attachment



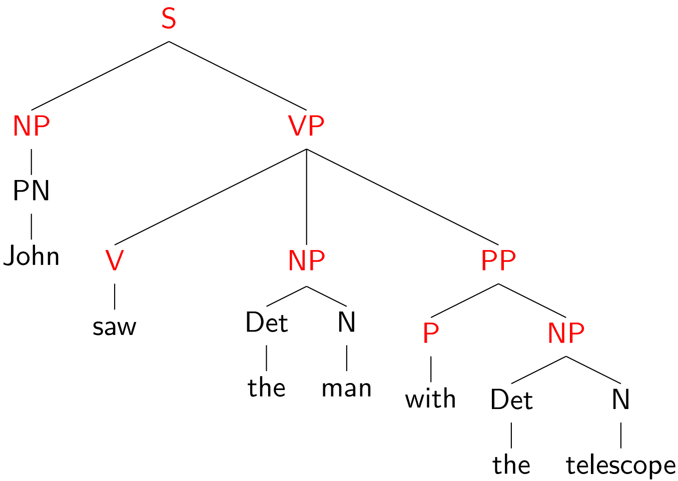
Adopt structure containing fewest number of nodes

First Strategy: Minimal Attachment



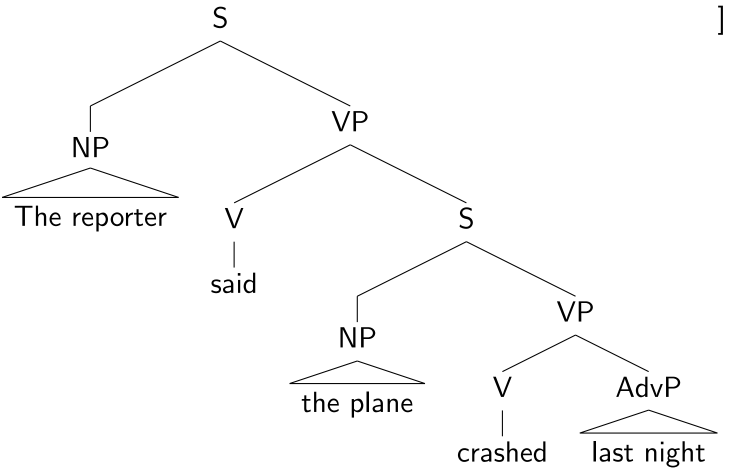
Adopt structure containing fewest number of nodes

First Strategy: Minimal Attachment



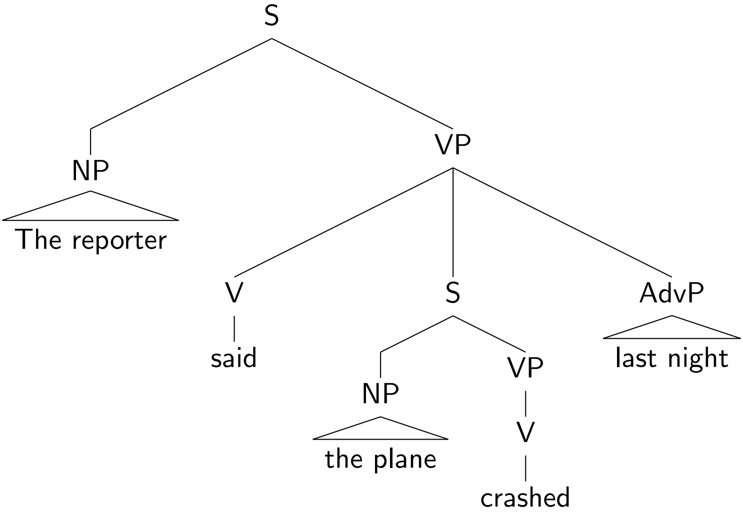
Adopt structure containing fewest number of nodes

Second Strategy: Late Closure



Add incoming material to clause/phrase currently processed.

Second Strategy: Late Closure



Add incoming material to clause/phrase currently processed.

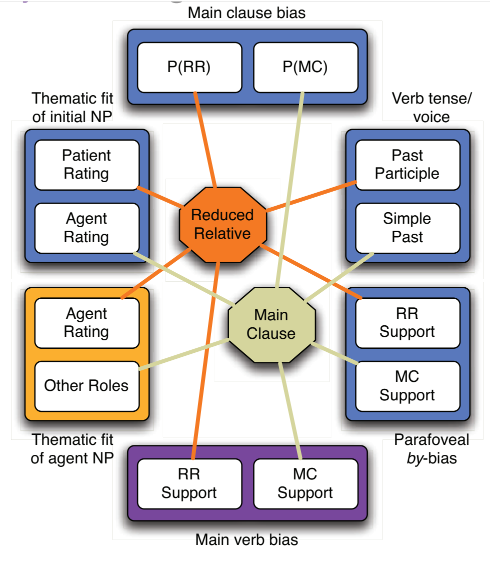
The Competition Integration Model (McRae et al., 1998)

Diverse constraints (linguistic and conceptual) are brought to bear simultaneously in ambiguity resolution.

- Model assumes all possible analyses are constructed
- Constraints provide probabilistic support for analyses
- There are multiple processing cycles given each input
- On each cycle, evidence in support of the syntactic alternatives is computed.
- Competition ends when the activation of one alternative reaches a threshold.
- Processing time is assumed to be a linear function of the duration of competition.

The Competition Integration Model (McRae et al., 1998)

The crook arrested by the detective was guilty of taking bribes.



Summary

- Sentence processing (parsing) is the task of assigning a structure to a string of words;
- Human sentence processing is incremental (word by word);
- It can encounter global vs local ambiguity.
- Garden paths derive from local ambiguities that are hard to resolve; they lead to longer processing times
- Theories of sentence processing: serial vs parallel, autonomous vs. interactive.