

Syntactic Parsing

based on: van Gompel, R., Pickering, M. Syntactic parsing.
The Oxford handbook of psycholinguistics.

May 6, 2014

Introduction

- ▶ To explain why a sentence such as (1) has two possible interpretations, we need to construct its syntactic structure.

(1) The hunter killed the poacher with the rifle.

- ▶ Processes involved in constructing syntactic structures during language comprehension are referred to as **parsing** or **syntactic processing**.

Incrementality

- ▶ Parsing is **incremental** = language comprehenders incorporate each word into the preceding syntactic structure as they encounter it.
- ▶ Evidence for incrementality: language comprehenders experience difficulty with temporarily ambiguous sentences well before the end of the sentence (the garden-path effect).
 - (2) The defendant examined by the lawyer turned out to be unreliable.

- ▶ People use different sources of information during sentence processing.

- ▶ People use different sources of information during sentence processing.
- ▶ Do they immediately use all relevant sources, or are some sources of information delayed relative to others?

Sentence processing theories

... can roughly be divided into

- ▶ **interactive accounts** (information can be used immediately)
- ▶ **modular accounts** (some information can be used immediately, some not)

Modular models

- ▶ Assumption: the mind consists of modules that perform specific processes.
- ▶ These processes use only information represented within this module.

Modular models

- ▶ Assumption: the mind consists of modules that perform specific processes.
- ▶ These processes use only information represented within this module.

- ▶ Most sentence processing studies use globally and temporarily ambiguous sentences
 - ▶ syntactic 'hints' allow multiple interpretations

Modular models

Garden-path model

1. The sentence processor initially employs only information about the syntactic structure
 - ▶ he adopts a single analysis in (temporarily) ambiguous sentences

Modular models

Garden-path model

1. The sentence processor initially employs only information about the syntactic structure
 - ▶ he adopts a single analysis in (temporarily) ambiguous sentences
2. Other, non-structural sources of information are employed in later stages of processing

Modular models

Garden-path model

1. The sentence processor initially employs only information about the syntactic structure
 - ▶ he adopts a single analysis in (temporarily) ambiguous sentences
2. Other, non-structural sources of information are employed in later stages of processing
3. The processor has to reanalyse when the initial analysis becomes inconsistent with information that becomes available later
 - ▶ processing difficulty follows

Modular models

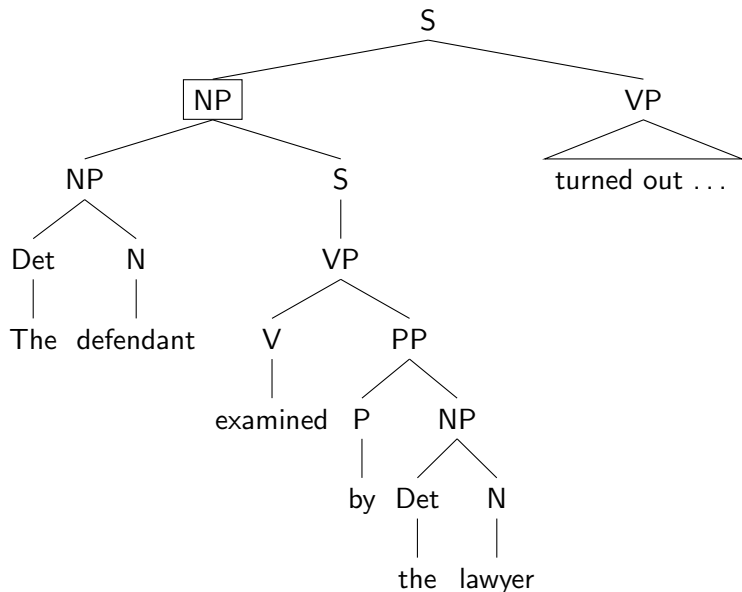
Garden-path model: Minimal attachment principle

- ▶ The processor incorporates an ambiguous phrase into the preceding syntactic tree structure using the fewest number of nodes.
 - ▶ This explains the garden-path effect in the first example:

(1) The defendant examined by the lawyer turned out to be unreliable.

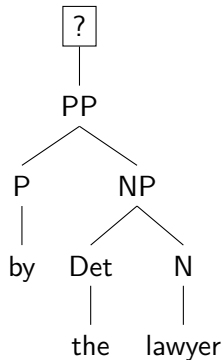
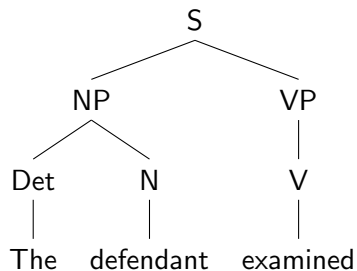
Modular models

Minimal attachment principle: Tree structure of the reduced relative analysis



Modular models

Minimal attachment principle: Tree structure of the main clause analysis



Modular models

Garden-path model: Late closure principle

If two analyses of an ambiguous structure have an equal number of tree structure nodes, the **late closure principle** applies.

- ▶ Prediction: people attach an ambiguous phrase to the currently processed phrase.
- ▶ In the following sentence the principle predicts that the relative clause 'that was tasty' prefers to attach low to the most recent clause 'the sauce' rather than high to 'the steak'.

(3) The steak with the sauce that was tasty didn't win a prize.

Modular models

Alternative accounts

- ... claim that the processor ignores certain sources of information.
- ▶ Most assume that the processor prefers analyses that involve a thematic relationship = prefers arguments over adjuncts.
 - ▶ (An argument is required by a predicate, an adjunct provides extra information.) See example:
- (4) a. The company lawyers considered employee demands *for a month* but they did not act.
- b. The company lawyers considered employee demands *for a raise* but they did not act.

Modular models

Reanalysis

What happens once the processor discovers that the initial analysis is inconsistent with later information and has to reanalyse?

- ▶ Inconsistency with later syntactic information = syntactic information makes the initial analysis ungrammatical.
- ▶ Misanalysis detection and subsequent reanalysis result in processing difficulty.

There is little evidence for reanalysis when the initial choice remains syntactically possible and semantically plausible, but other sources of information favour the alternative analysis.

Modular models

Reanalysis

Some results about reanalysis:

- ▶ The length of the temporarily ambiguous phrase affects reanalysis cost
- ▶ Attachment to a recent phrase is preferred (re-attachment to a distant phrase is costly)

Modular models

Reanalysis

People do not always successfully abandon their initial analysis after encountering a syntactic disambiguation:

(5) While Anna dressed the baby that was small and cute spit up on the bed.

- ▶ Who spit up on the bed?
- ▶ Did Anna dress the baby?
- ▶ Readers adopted the subject analysis for 'the baby' (as the subject of 'spit up'), while at the same time they retained the incorrect analysis on which this phrase was the object of the verb 'dressed'.

Interactive models

Constraint-based theories/models

- ▶ Assumption: the processor immediately draws upon all possible sources of information
 - ▶ semantics, discourse context and information about the frequency of syntactic structures
- ▶ All syntactic alternatives are activated in parallel
- ▶ **Processing difficulty** appears in consequence of two analyses having a similar activation at the point of disambiguation
 - ▶ No true reanalysis: both analyses activated → disambiguation does not result in the construction of a initially not considered analysis.

Interactive models

Semantic effects

Semantic information provides strongly constraining information for syntactic analysis.

Is this information used immediately to guide sentence processing?

- (6) a. The defendant examined by the lawyer turned out to be unreliable.
- b. The evidence examined by the lawyer turned out to be unreliable.

Does similar difficulty as in (a) occur in (b), where semantic information rules out the main analysis?

- ▶ Experiments and studies show that semantic information fails to override syntactic preferences.

Interactive models

Frequency effects

Different languages have different relative clause attachment preferences - this might be explained by structure frequency:

- (7)
- a. The journalist interviewed the daughter of the colonel who had the accident.
 - b. El periodista entrevistó a la hija del coronel que tuvo el accidente.
-
- ▶ English: low attachment preference (the colonel had the accident)
 - ▶ Spanish: high attachment preference (the daughter had the accident) - also Dutch and French

Cross-linguistic differences present a problem for the garden-path theory: late closure predicts a universal preference for low attachment.

Interactive models

Discourse effects

Many parsing preferences occur because sentences are presented in isolation.

- (8)
- a. The burglar blew open the safe with the dynamite and made off with the loot.
 - b. The burglar blew open the safe with the new lock and made off with the loot.
- ▶ Absence of context: people initially prefer to attach the PP 'with the dynamite/new lock' to the VP containing 'blew open' rather than to the NP containing 'the safe'
 - ▶ Specific discourse context (one or two safes had been mentioned): preferences change

Adopting ungrammatical syntactic structures

People may misanalyse sentences that are not locally ambiguous:

- ▶ People often misanalyse passive sentences as active:

(9) The dog was bitten by the man.

The strategy is to interpret the first NP as the agent and the second as the patient, despite the fact that this is ungrammatical. This strategy is particularly strong if plausibility information supports this analysis.

- ▶ People may also sometimes misanalyse active sentences as passive

Conclusions and future directions

- ▶ An important aim in sentence processing research: investigate whether the parser is modular or interactive.
 - ▶ It appears that non-syntactic information has an effect on sentence processing (especially discourse and frequency information)
 - ▶ The use of semantic plausibility appears to be less rapid

Conclusions and future directions

- ▶ Researchers have become more interested in natural conversation: sentences produced during natural conversation are generally very different from those in well-constructed texts.
- ▶ Participants considered sentences such as the following more often grammatical when a disfluency preceded 'waiter' than when it preceded 'busboy'.

(10) Sandra bumped into the busboy and the waiter told her to be more careful.