Computational Psycholinguistics

Lecture 5: Reanalysis

Matthew W Crocker

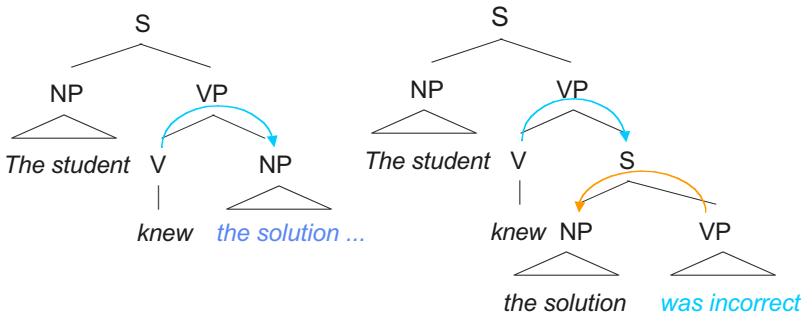
Computerlinguistik
Universität des Saarlandes

Pritchett (1992)

- Theta-Attachment:
  - The theta-criterion attempts to be maximally satisfied et every point during processing, given the maximal theta-grid
- Theta-Criterion:
  - Each argument must receive exactly one theta-role, and each theta role must be assigned to exactly one argument
- Theta Reanalysis Constraint:
  - Reanalysis of a constituent out of its theta-domain results in a garden-path effect
- Generalised Theta Attachment:
  - Every principle of the Syntax attempts to be maximally satisfied at every point during processing.

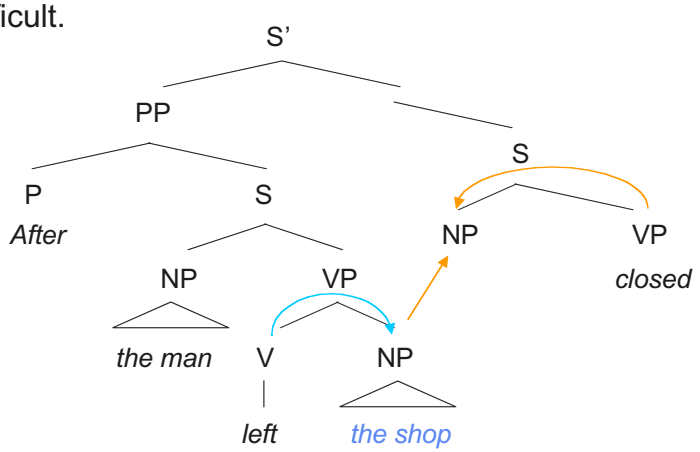
Theta-Reanalysis: Easy

- Reanalysis to a position within the original theta-domain is easy.



Theta-Reanalysis: Difficult

- Reanalysis to a position outside the original theta-domain is difficult.



Ambiguities revisited: [preferred/dis-preferred]

- NP/VP Attachment Ambiguity:
  - "The cop [saw [the burglar] [with the binoculars]]"
  - "The cop saw [the burglar [with the gun]]"
- NP/S Complement Attachment Ambiguity:
  - "The athlete [realised [his goal]] last week"
  - "The athlete realised [[his shoes] were across the room]"
- Clause-boundary Ambiguity:
  - "Since Jay always [jogs [a mile]] the race doesn't seem very long"
  - "Since Jay always jogs [[a mile] doesn't seem very long]"
- Red. Relative-Main Clause Ambiguity:
  - "[The woman [delivered the junkmail on Thursdays]]"
  - "[[The woman [delivered the junkmail]] threw it away]"
- Relative/Complement Clause Ambiguity:
  - "The doctor [told [the woman [that he was in love with]] [to leave]]"
  - "The doctor [told [the woman] [that he was in love with her]]"

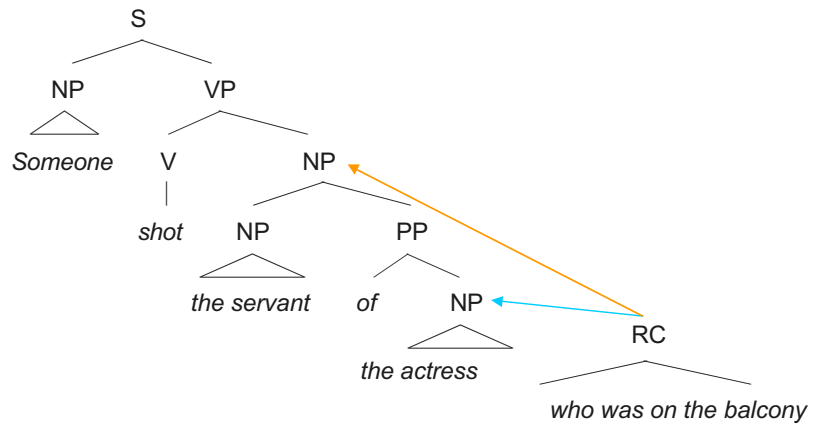
Grammar-Based

- Theta-Attachment: relies on theta-grids
  - Head driven
  - Fine for English
  - Not incremental for head-final languages
  - Same problem for head-driven parsers, e.g. Marcus (1981), Abney (1989)
- Crocker: A-Attachment
  - Prefer attachment into 'potential' A-positions, over A'-positions
  - A-position: a position which is visible for theta-role assignment

Experienced-based Models

- Resolve ambiguities according to linguistic experience:
  - Lexical Guidance Hypothesis: (Ford et al)
    - Resolve subcategorisation ambiguities in favour of the most likely frame for the ambiguous verb
  - Linguistic Tuning Hypothesis: (Mitchell et al)
    - Resolve structural ambiguities according to the structure which has previously prevailed
- Relative clause attachment
  - Someone shot the servant of the actress who was on the balcony

Relative Clause Attachment



Alguien disparo contra el criado de la actriz que estaba en al balcon

Cross-linguistic RC Preferences

	Off-Line	On-Line
Spanish	high	low
French	high	low
Italian	high	low
Dutch	high	
German	high	low(early), high(late)
English	low	??, probably low
Arabic	low	
Norwegian	low	
Swedish	low	
Romanian	low	

- Immediate low attachment, possibly revised quickly (even on-line) … seems the best account

The Grain Problem

- Experience-based models rely on frequency of prior linguistic exposure to determine preferences.
- There are many ways to realise experience-based models
  - Possibilities: What kinds of things do we count?
    - Actual sentence/structure occurences? Data too sparse?
    - Head driven: I.e. verb subcategorization frequencies
      - Do we distinguish tenses? Senses?
    - Word level, part-of-speech
    - Tuning is structural: NP P NP RC *vs* NP P NP RC
      - High Low
- Interesting issues:
  - Does all experience have equal weight (old vs. new)?
  - Are more frequent “words” or “strings” (idioms) dealt with using finer grain statistics that less frequent?

Summary

- Shared features:
  - Parsing is a distinct process
  - Decisions are resolved using only syntactic information:
    - structural principles
    - grammar constraints
    - frequency
  - "Dependency formation" seems to be important:
    - Associate argument to heads ASAP
    - Associate fillers with gaps/heads ASAP
  - Argument/Adjunct assymetries
- Are these models successful?
- Do they support the notion of a modular architecture?

Reanalysis Revisited

- Incrementality leads to local ambiguity
- Local ambiguity leads to reanalysis:
  - how does the HSPM recover?
- Two possibilities
  - Serial: Deterministic, Backtracking or 'repair'
  - Parallel: Re-rank or change activation
- Are there different 'sorts' of reanalysis:
  - "John knows the truth **hurts**" *easy*
  - "While John was walking the dog **barked**" *hard*
  - "The boat sailed down the river **sank**" *v. hard*

Deterministic Parsing

- Alternative: make sure there is only one possible parsing action at any point in the derivation.
  - Avoid/delay rule selection until it is fully determined
  - How: fully bottom up (e.g. S/R) plus lookahead
  - Marcus, Berwick and Weinberg, and Abney
- Advantage: very fast, clear predictions
- Disadvantages:
  - not fully incremental
  - unsuccessful for head-final languages
  - wrong predictions!

Monotonic Parsing

- Inspired by determinism and reanalysis:
  - many local ambiguities seem to cause little difficulty
  - contra predictions of naive backtracking
  - Gorrell, Weinberg, Sturt & Crocker
- Provides a richer set of 'tree-building' operations which means destructive backtracking is not always required
- Predicts 'reanalysis' outwith these operations to be difficult

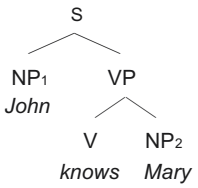
Talking about Talking about Trees

- Traditional theories locate the reanalysis process in the parser: “re-parsing”
- Can local ambiguity be handled using *underspecified* representations?
  - *Representations which allows some ambiguity to remain, and be later removed without reparsing.*
- Description-Theory: (Marcus, Hindle & Fleck, 1983)
  - Uses tree descriptions: dominance, rather than immediate dominance
  - Allows subsequent insertion of nodes & branches

[Gorrell (1995), Weinberg (1993), Sturt & Crocker (1996)]

Monotonic Parsing (Gorrell; Sturt & Crocker)

- Trees are described as a set of nodes, and a set of precedence and dominance relations:
  - *John knows Mary*



  - {dom(S,NP_1), dom(S,VP), dom(S,V), dom(S,NP_2), prec(NP_1,VP), dom(VP,V), dom(VP,NP_2), prec(V,NP_2) …}

Properties of Trees

- Single root condition:

$$\exists x \forall y \cdot dom(x, y)$$

- Exclusivity condition:

$$\forall x, y \cdot prec(x, y) \vee prec(y, x) \leftrightarrow \neg dom(x, y) \wedge dom(y, x)$$

- Inheritance

$$\forall w, x, y, z \cdot prec(x, y) \wedge dom(x, w) \wedge dom(y, z) \rightarrow prec(w, z)$$

- *dom* and *prec* are transitive relations
- *dom* is reflexive, *prec* is irreflexive

Constraints on the Model

- Strict incrementality:
  - words are connected to the tree description as they are encountered
- Coherence:
  - tree properties must always be satisfied, the tree must be grammatically licensed
- Full specification of nodes:
  - no features on nodes (e.g. bar-level) can be left unspecified
- Informational monotonicity:
  - the tree description at state$_n$ is a subset of the description at state$_{n+1}$
- Obligatory assertion of precedence:
  - precedence must be specified for sisters

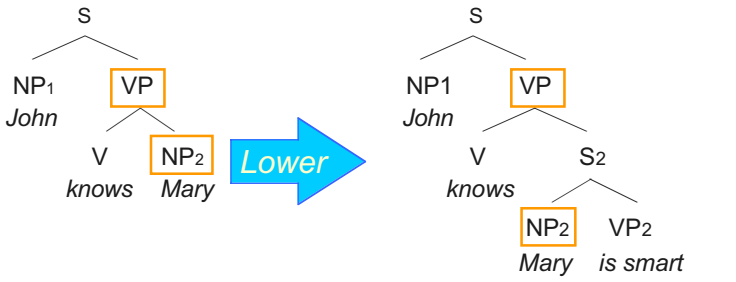
Monotonic Parsing and Reanalysis

- Easy: *Monotonic* reanalysis
  - *John knows Mary is smart*
- Hard: *Non-monotonic* reanalysis
  - *While John walked the dog barked*
- How does the parser work:
  - The monotonic parser is realised via specialised “macro”-like operations, which meet the constraints.
  - This simplifies the parsers search process
- Non-monotonic operations are not permitted
- General Predictions:
  - Easy reanalysis is simply parsing
  - Hard reanalysis is not possible

Monotonic Parsing Operation

- Tree-lowering
  - John knows Mary …

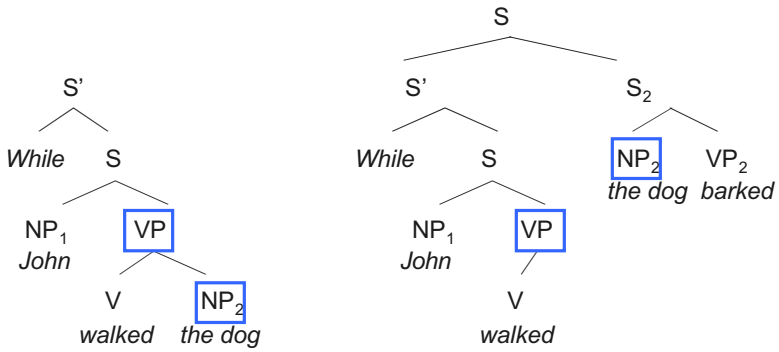
John knows Mary is smart



  - {dom(S,NP1), dom(S,VP), dom(S,V), dom(S,NP2), prec(NP1,VP), dom(VP,V), dom(VP,NP2), prec(V,NP2), dom(VP,S2, dom(S2,NP2), prec(NP2,VP2) …}

Non-Monotonic Parsing

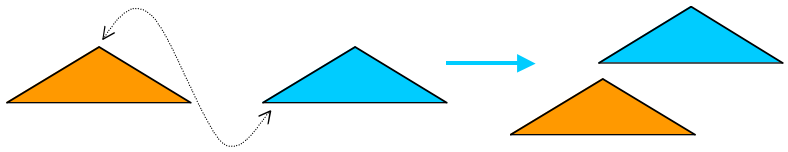
- Predicting difficult reanalysis
  - While John walked the dog … barked.



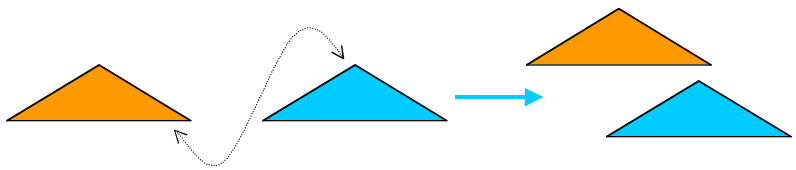
  - … *dom*(VP,NP_2) … ***but*** … *dom*(VP,NP_2) & *prec*(VP,NP_2) ...

Parsing Operations: Attachment

- Left attachment:

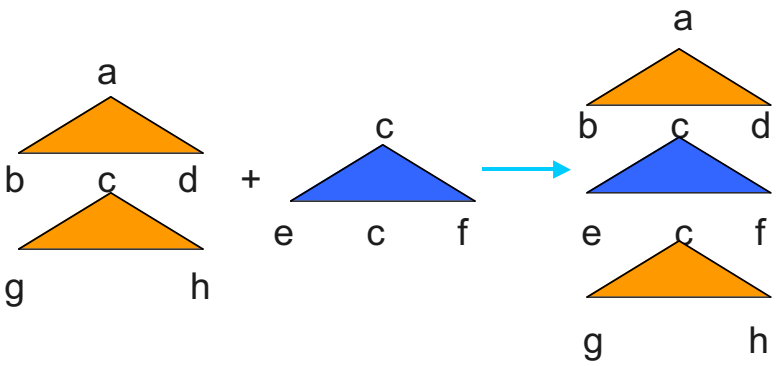


- Right attachment:

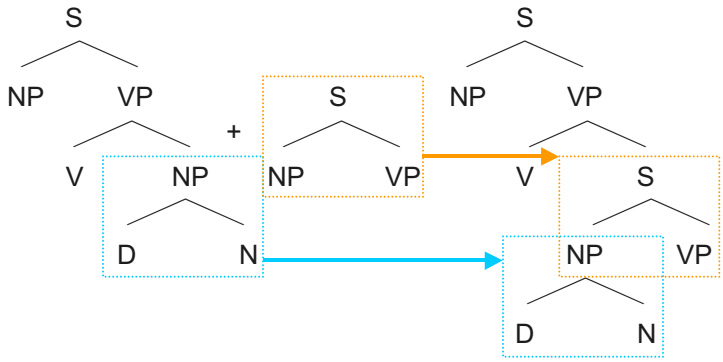


TAG Adjunction

- The operations of the monotonic parser resemble those of Tree Adjoining Grammar.



Tree Lowering

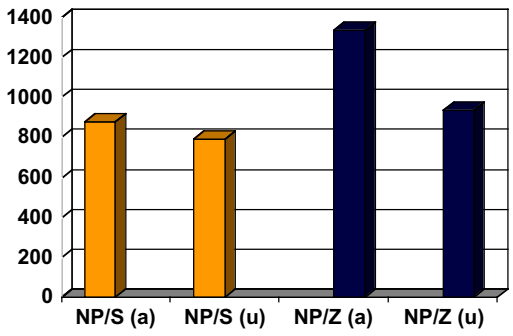


Psycholinguistic Evidence

- Are there really two types of reanalysis?
  - NP/S Ambiguous:
    The woman saw the famous doctor had been drinking
  - NP/Z Ambiguous:
    Before the woman visited the famous doctor had been drinking
  - NP/S Unambiguous:
    The woman saw that the famous doctor had been drinking
  - NP/Z Unambiguous:
    Before the woman visited, the famous doctor had been drinking
- All verbs are biased (BNC) towards NP complement
  - To make sure the object attachment is initially adopted, forcing reanalysis
- Plausibility of the the direct object analysis is similar (pre-test).

Results

- Reading times:

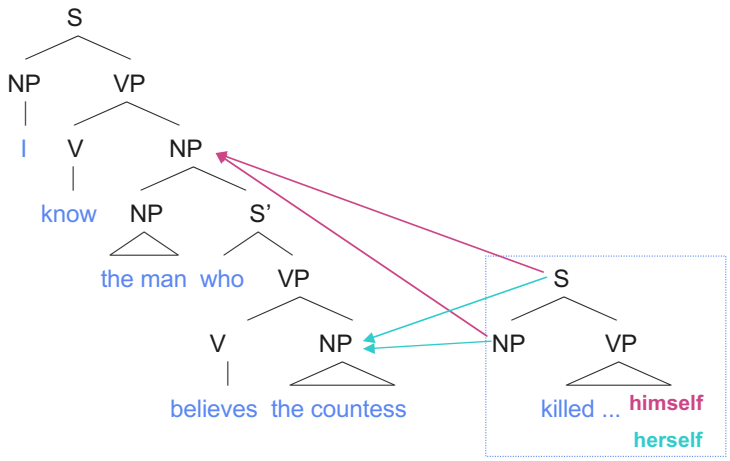


  - Main effects of construction type, ambiguity, and a significant interaction
  - GP effect: NP/Z (400ms) vs. NP/S (87ms)

Search in Parsing/Reanalysis

- (Some) reanalysis is simply monotonic attachment: e.g. tree-lowering.
- What if there are multiple such attachments
- Consider a double NP/S ambiguity:
  - I know the man who believes the countess killed *herself*
  - I know the man who believes the countess killed *himself*
- Which is easier?
- How does the parser search for an attachment?

The Trees



Search

- English appears to use a bottom-up search strategy for attachment
  - late closure, recency

  *Someone shot **the servant** of **the actress** who was on the balcony*
- Possibly other influences:
  - verb bias? predicate proximity (Gibson)?
- Japanese seems to be top-down
  - Head final language
  - Sturt & Crocker(1996)