Computational Psycholinguistics Lecture 4: **Reanalysis**

Matthew W. Crocker
crocker@coli.uni-sb.de

Summary of Frazier

- Parsing preferences are guided by general principles:
  - Serial structure building
  - Reanalyze based on syntactic conflict
  - Reanalyze based on low plausibility (“thematic fit”)
- Psychological assumptions:
  - Modularity: only syntactic (not lexical, not semantic) information used for initial structure building
  - Resources: emphasizes importance of memory limitations
  - Processing strategies are universal, innate

Grammar-Based Strategies

- Not concerned with representation or 'form', but defined in terms of syntactic 'content' – i.e. not based on "trees" but "principles of grammar"
- Strategies are modular, but 'knowledge-based'
- Motivation: strategies are derived from the purpose of the task, not e.g. computational efficiency
- Closer competence-performance relationship
- Defined w.r.t. to deeper syntactic notions: less sensitive to minor structural details (cf. Minimal Attachment)
- Pritchett (1988), Abney(1989), Crocker(1991;1996), Gibson (1992)

Pritchett (1992)

- Incrementally establish primary syntactic dependencies
- Theta-Criterion: (GB theory, also in LFG + HPSG)
  - Each argument must receive exactly one theta-role, and each theta role must be assigned to exactly one argument
- Consider:

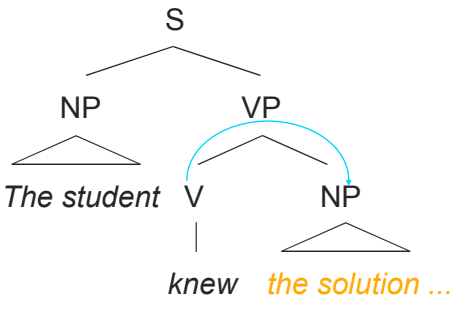
The boy put the candy on the table in his mouth

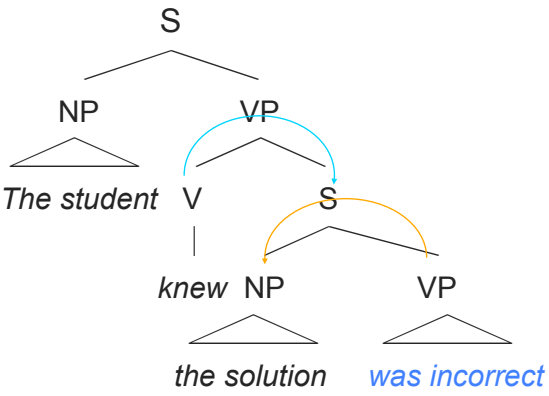
Pritchett (1992)

- **Theta-Attachment:**
  - Maximally satisfy the theta-criterion at every point during processing,given the maximal theta-grid of the verb
- **Theta Reanalysis Constraint:**
  - Reanalysis of a constituent out of its theta-domain results in a conscious garden-path effect

Theta-Reanalysis: Easy

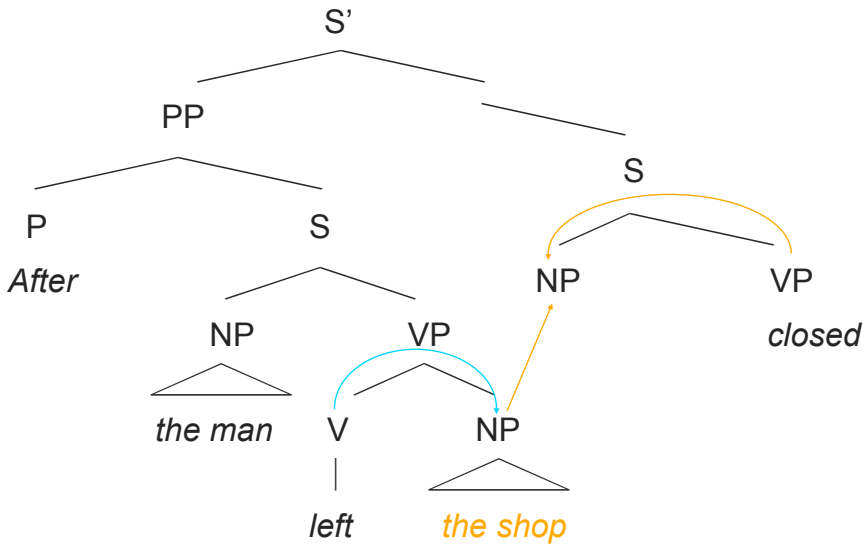
Reanalysis to a position **within** the original theta-domain is easy.





Theta-Reanalysis: Difficult

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



Pritchett: Another example

- *"Without her contributions the orphanage closed"*
  - 'Without': a Prep with a single thematic role
  - 'her': an determiner of an unseen NP head, or a Full NP (Pronoun) [Theta-attach]
  - 'contributions': head of a new NP, with no role, or combine with 'her' for a Full NP [Theta-attach]
- *"Without her contributions failed to come in"*
  - 'contributions' becomes subject of 'failed', violating [Theta-reanalysis Constraint]

Well-known local ambiguities

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 goals]] last week"

"The athlete realised [[his goals] were unattainable]"

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]"

Reduced 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 her]]"

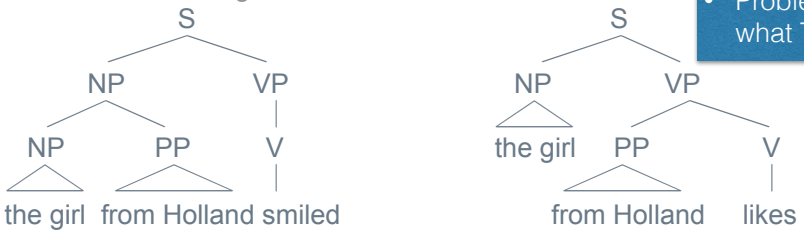
"The doctor [told [the woman [that he was in love with]] [to leave]]"

Grammar-Based (cont'd)

- Theta-Attachment: reliance on theta-grids means it's head driven
  - O.k. for English, but not incremental for head-final languages
  - Same problem for Abney (1989), and other head-driven models
- Argument-Attachment: Attach constituent into potentially role-receiving positions (Crocker, 1992)

- That study used phrase-by-phrase self-pace reading.
- Eye-tracking studies suggest the modifier attachment is actually preferred.
- Problematic for A-Attachment, unclear what Theta-Attachment would predict.

"... dat het meisje van Holland glimlachte/houdt"
… that the girl from Holland smiled/likes



Pritchett's Theory (1992)

- What **architecture** is assumed?
  - Modular lexico-syntactic processor with syntactic and thematic role features
- What **mechanisms** is used to construct interpretations?
  - Incremental, serial parsing, with reanalysis
- What **information** is used to determine preferred structure?
  - Grammar principles and thematic role information
- **Linking Hypothesis**:
  - TRC violation causes garden-path, reanalysis without TRC is relatively easy

Quick Recap: Syntactic Accounts

- **Frazier**: early parsing decisions driven by purely syntactic heuristics
  - **MA** and **LC** were argued to be by-products of a race mechanism
- **Pritchett**:
  - **Theta-attachment** determines preferred parse
  - **Theta-reanalysis** constraint determines when reanalysis of difficult
- Eager dependency-formation plays a strong role in driving parsing decisions:
  - What about **long-distance dependencies**?

Long Distance Dependencies

- Wh-Fillers:
  - *Who$_{i}$ did Fred tell Mary e$_{i}$ left the country?* **dispreferred**
  - *Who$_{i}$ did Fred tell e$_{i}$ Mary left the country?* **preferred**
- Subject-Relative preference:
  - *I met the man$_{i}$ that John likes e$_{i}$.* **dispreferred**
  - *I met the man$_{i}$ that e$_{i}$ likes John.* **preferred**
- **Active Filler Strategy:** ("*Gap as a first resort*")
  - When a filler has been identified, rank the possibility of a assigning it to a gap above all other options.

Further observations ...

- Filled-Gap effect:
  - *My brother wanted to know who$_{i}$ Ruth will bring (*e$_{i}$) <u>us</u> home to e$_{i}$ at Christmas*
  - *My brother wanted to know if Ruth will bring <u>us</u> home to Mom at Christmas*
- Found an increased reading time at **us**, interpreted this as surprise
- Intuitively easy:
  - *Who$_{i}$ (e$_{i}$) did you want (e$_{i}$) Mother to bake (e$_{i}$) a cake for e$_{i}$?*
  - ... despite 3 possible earlier gap locations

Gaps versus Dependencies

- Consider:
  - *[In which tin]$_i$ did you put the cake e_i ?* **Gap account**
  - *[In which tin]$_i$ did you put$_i$ the cake ?* **Dependency account**
- If keeping the filler in memory causes difficulty, we can compare:

[In which tin]$_i$ did you put$_i$ the cake that your little sister baked for you e_i ?

Easy

[Which tin]$_i$ did you put the cake that your little sister baked for you in$_i$ e_i ?

Hard

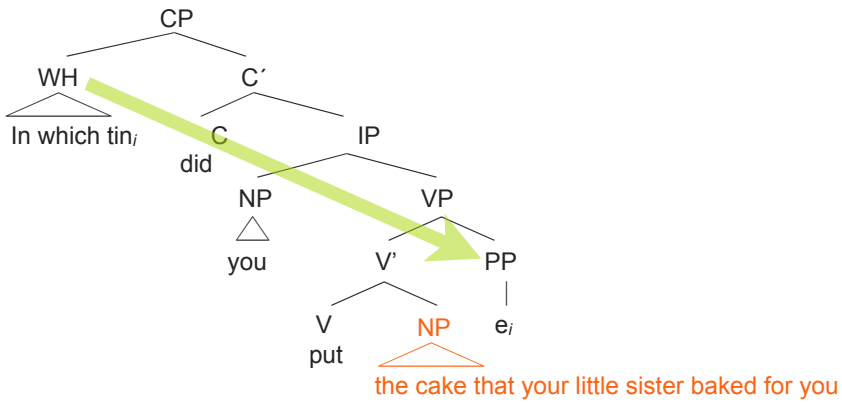
- **Intuitive support for the dependency account, and against gaps.**

Other evidence

- Implausibility detected immediately at the verb, as shown by increased reading times.
  - *That's the [pistol/**garage**]$_i$ with which the heartless killer shot$_i$ the hapless man e_i yesteday afternoon.*
- Garnsey et al (1998) found an N400 at the verb, for the implausible condition
  - *The businessman knew which [customer/**article**]$_i$ the secretary called e_i at home*
- Pickering and Barry (1996) argued that a dependency-based account was preferable to a trace-based account.

Parsing in 2 dimensions

- Gaps don't exist in the input, so we needn't wait until they are found
- We can associate a filler & gap as soon as the structure licenses it:



Consider: *Den Hund$_i$ sah$_j$ Maria e$_j$ e$_i$.*

Psycholinguistic Evidence

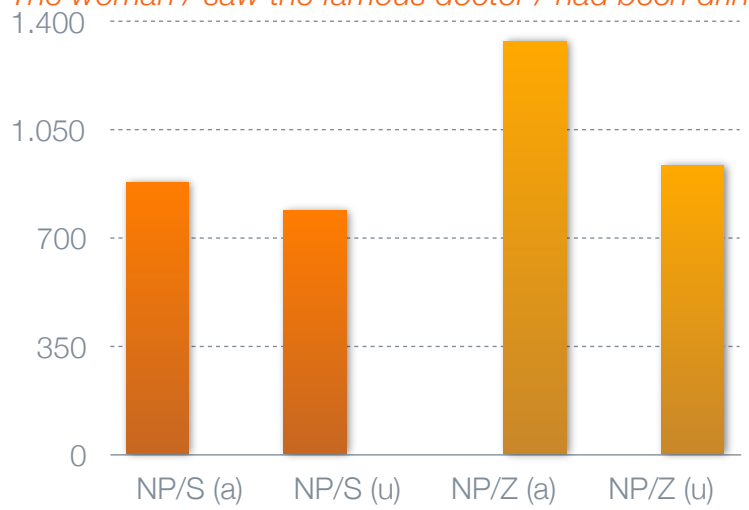
- Are there really two types of reanalysis?
  - **NP/S (A):** *"The woman saw the famous doctor had been drinking"*
  - **NP/Z (A):** *"Before the woman visited the famous doctor had been drinking"*
  - **NP/S (U):** *"The woman saw that the famous doctor had been drinking"*
  - **NP/Z (U):** *"Before 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).

Sturt, Pickering & Crocker, *JML*, 1999

Results

- Reading times: Region 3

"The woman / saw the famous doctor / had been drinking / all day"



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