Human Sentence Processing

Lexicalized Parsing

- peel the apple on the towel
  - ambiguous
- put the apple on the towel
  - put loves on (is the other reading even possible?)
    - put the apple on the towel in the box
    - VP[head=put] ? V[head=put] NP PP
    - VP[head=put] ? V[head=put] NP PP[head=on]
- study the apple on the towel
  - study dislikes on (how can the PCFG express this?)
    - VP[head=study] ? VP[head=study] PP[head=on]
- study it on the towel
  - it dislikes on even more – PP can't attach to pronoun

Lexicalized Parsing

- the plan that Natasha would swallow
  - ambiguous between content of plan and relative clause
- the plan that Natasha would snooze
  - snooze dislikes a direct object (plan)
- the plan that Natasha would make
  - make likes a direct object (plan)
- the pill that Natasha would swallow
  - pill can't express a content-clause the way plan does
  - pill is a probable direct object for swallow

- How to express these distinctions in a CFG or PCFG?

Human Performance: Self-Paced Reading Experiments

Pretend you're pressing a key to get each word:

- The shop sold to the bank was quite old.
- We included this sentence just to distract you.
- The necklace sold to the bank was quite old.

- The lawyer examined by the judge was silly.
- The evidence examined by the judge was silly.

Human Performance: Self-Paced Reading Experiments

- What happens to a self-paced reader when she goes down the garden path?
  - Eventually she has to backtrack.
  - That's when she pauses: "point of disambiguation."
  - The lawyer examined by the judge …
- Why isn't the second sentence a garden path:
  - The shop sold to the bank was quite old.
  - The necklace sold to the bank was quite old.
- People are sensitive to frequency!

Big Human-Parsing Debate of the 1990's: How Soon Does Semantics Come Into Play?

- How fast is semantics? Use it constantly, or only as a last resort?
- Hypothesis 1: Rely mainly on syntactic heuristics
  - Get a parse this way, *then* interpret it semantically
  - Backtrack & fix if we can't finish the parse or it makes little semantic sense
  - **Sample heuristic:** When you build a PP, attach it to the most recently built thing you can
  - **Sample heuristic:** When an NP starts a sentence, it's the subject
  - If true, people should backtrack on "the necklace sold to the bank was …"
- Hypothesis 2: Rely mainly on syntactic probabilities using head words
  - Smarter version of hypothesis 1: "necklace" isn't a common subject for "sell"
  - Explains why no backtracking on "the necklace sold to the bank was …"
- Hypothesis 3: Consider full semantics of a constituent as soon as it's built
  - Full interpretation as soon as we build a constituent
  - So semantic analysis and backtracking are *never* delayed
  - Garden paths result from genuinely ambiguous prefix, not slow semantics
- Hypothesis 4: Consider full semantics of a constituent even before it's built
  - Start interpreting a constituent before hearing it all
  - Semantics before syntax! (opposite of hypothesis 1)

Eye Tracking

- Self-paced reading data too flaky to answer these questions.
- Brain imaging is too slow and coarse.

- Track people's eye movements as they read.
  - They don't backtrack on "the necklace sold to the bank was ..."
  - Cleaner data; eliminates hypothesis 1.
- But reading is an artificial task – people didn't evolve to be good readers.

Eye Tracking

- The eyes are the window to the soul.
- You are constantly scanning your environment.
- Why? Evolution!

- highly accurate at jumping to objects
- 3-4 of these "saccades" per second on average
- fast motion (90 degrees of arc in 100 millisec)
- low latency – short wires connect eyes to brain

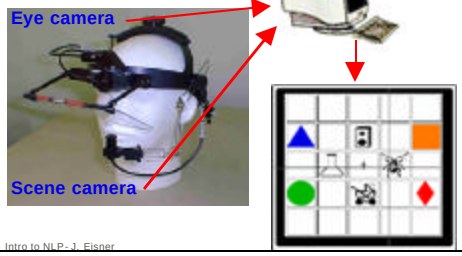
Head-Mounted Eye Tracker

- Like looking into someone's thoughts
- As they happen, in a real environment!

slide courtesy of M. Tanenhaus (modified)

Videotape

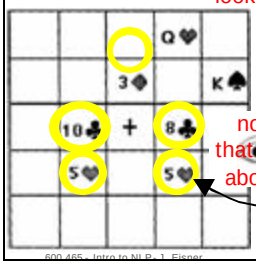
- From Mike Tanenhaus's lab – University of Rochester



(as shown on video)

The Visual World Paradigm

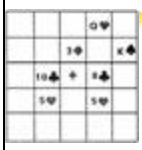
look at the five of hearts
look at the other five of hearts



now put the five of hearts
that is below the eight of clubs
above the three of diamonds

total time: 4.15 seconds

The Visual World Paradigm

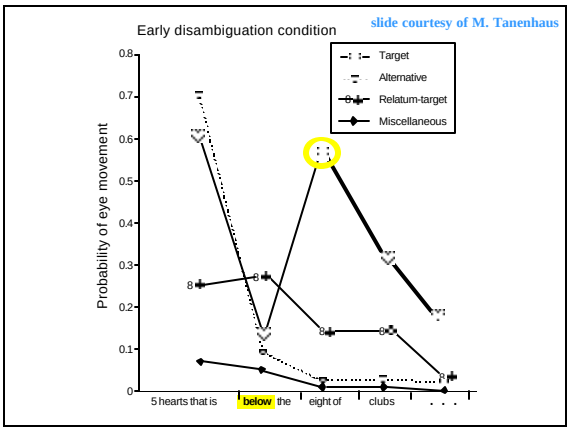


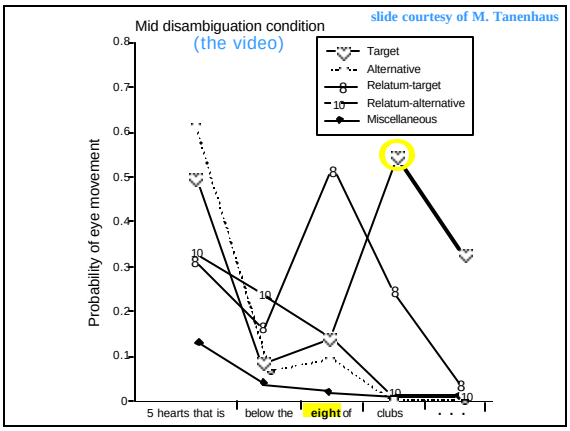
now put the five of hearts
that is below the eight of clubs
above the three of diamonds

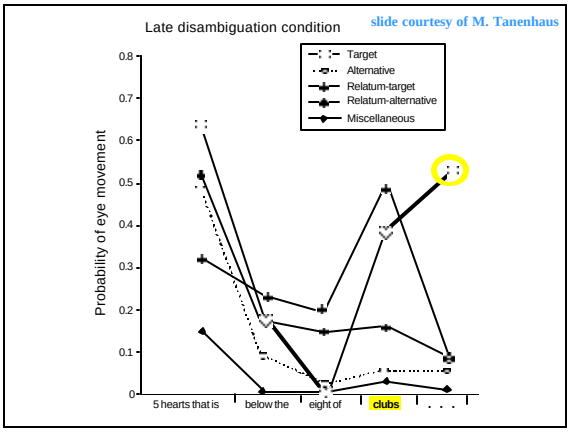
Subject looks at 5? shortly after point of disambiguation (underlined) – only one 5? below an 8

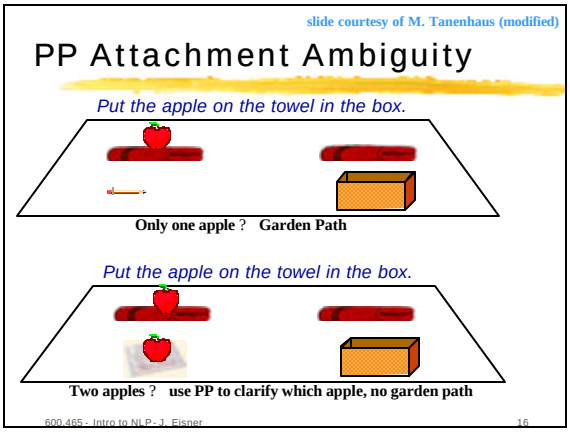
Where would point of disambiguation be if only one of the 5? was below something?

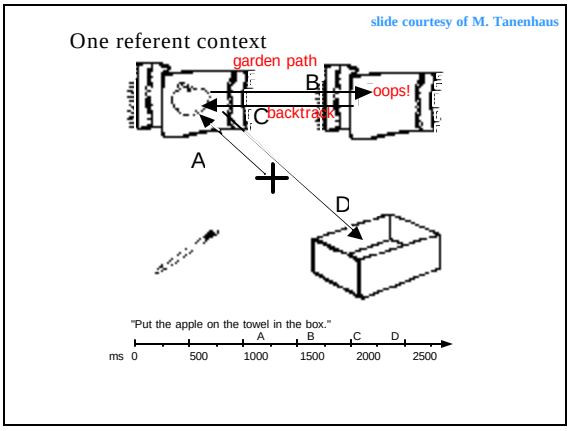
What if both 5? were below an 8? (8? , 8?)

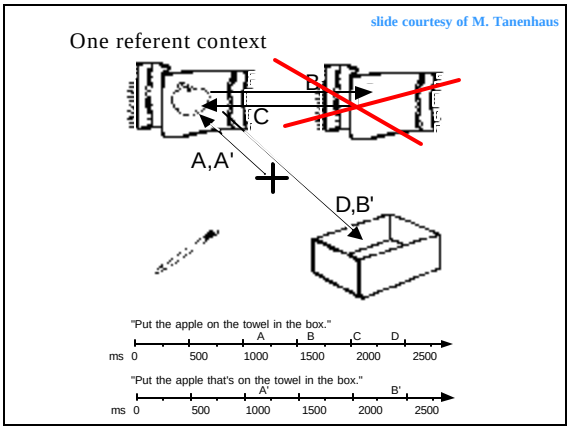

slide courtesy of M. Tanenhaus
Early disambiguation condition
Target
Alternative
Relatum-target
Miscellaneous
Probability of eye movement
5 hearts that is
below
the
eight of
clubs


slide courtesy of M. Tanenhaus
Mid disambiguation condition
(the video)
Target
Alternative
Relatum-target
Relatum-alternative
Miscellaneous
Probability of eye movement
5 hearts that is
below the
eight
of
clubs


slide courtesy of M. Tanenhaus
Late disambiguation condition
Target
Alternative
Relatum-target
Relatum-alternative
Miscellaneous
Probability of eye movement
5 hearts that is
below the
eight of
clubs

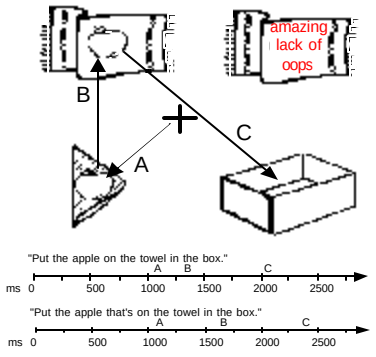

slide courtesy of M. Tanenhaus (modified)
PP Attachment Ambiguity
Put the apple on the towel in the box.
Only one apple ? Garden Path
Put the apple on the towel in the box.
Two apples ? use PP to clarify which apple, no garden path
600.465 - Intro to NLP - J. Eisner
16


slide courtesy of M. Tanenhaus
One referent context
garden path
B
oops!
C
backtrack
A
D
"Put the apple on the towel in the box."
ms
A
B
C
D


slide courtesy of M. Tanenhaus
One referent context
B'
C
A,A'
D,B'
"Put the apple on the towel in the box."
ms
A
B
C
D
"Put the apple that's on the towel in the box."
A'
B'

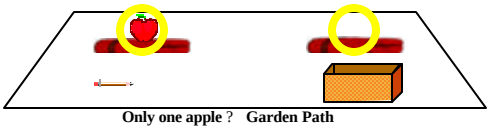
slide courtesy of M. Tanenhaus

Two-referent context



(same picture as before – as shown on video)

One-Referent Context



Only one apple ? Garden Path

put the apple on the towel in the box

put the candy bar on the saucer in the bowl
put the salt shaker on the envelope in the bowl

Processing at Syllable Level

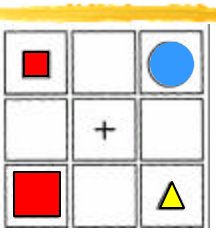
Don't wait for constituent to finish
Don't even wait for word to finish!

pick up the candy

They're already looking based on "can" –
we know this because if there's also a
candle, they're 50% likely to look at it first!

Processing Stress Information

what's the point
of disambiguation?



Touch the large red square.

Touch the LARGE red square.