FLST:Cognitive Foundations 2

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

Summary of cognitive issues

- Linguistic autonomy
  - Modularity vs localization in the brain (not the same)
  - Innate linguistic (domain specific) language “organ”
  - The relation between language and thought
- The genetic origins of language
  - The importance of small variation in FOXP2
- The emergence of the capacity for language
  - language specific versus general cognitive capacities
- Symbolic versus perceptually grounded meaning
  - Evidence for embodiment of mental representations

Human Language Processing

- We comprehend language word-by-word
  - *How do people construct incremental interpretations?*
- We must resolve local and global ambiguity
  - *How do people decide upon a particular interpretation?*
  - *What information sources are used? What is the time course?*
- Decisions are sometimes wrong!
  - *How do we find an alternative interpretation?*
- Answers can reveal important details about the underlying mechanisms

Language processing

- Psycholinguistics seeks cognitively plausible theories
  - about mental representations, and cognitive processes
- Computational psycholinguistics seeks to implement such theories
  - predictive models of human knowledge and behavior
- Model necessary to understand how people produce and comprehend language
  - *Competence*: How do utterances relate to meaning?
  - *Performance*: How do people establish this relationship during on-line language processing?

Competence & Performance?

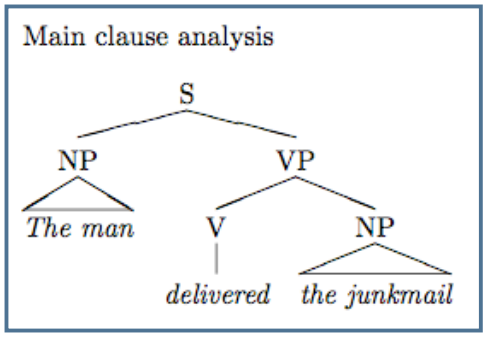
- Sometimes what we do
- Production: we say thin
  - Spoonerisms: "Mental le Mexican"
  - Agreement: "The key to the office doors are missing
- Comprehension: we can't understand things we know are ok
  - Centre embedding: "The mouse that the cat that the dog chased bit fled"
  - Garden paths: "The boat sailed down the river sank"

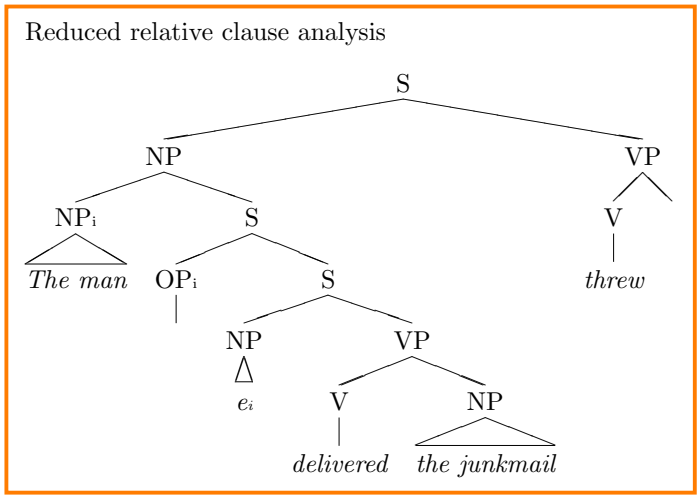
More Spoonerisms

- A lack of pies (A pack of lies)
- Wave the sails (Save the whales)
- Plaster man (Master plan)
- Bottle in front of me (Frontal Lobotomy)
- Rental Deceptionist (Dental Receptionist)
- Flock of bats (Block of flats)
- Chewing the doors (Doing the chores)

(Reduced) Relative Clauses

- One of the most studied syntactic ambiguities:
  - The man delivered the junkmail ... threw it away.

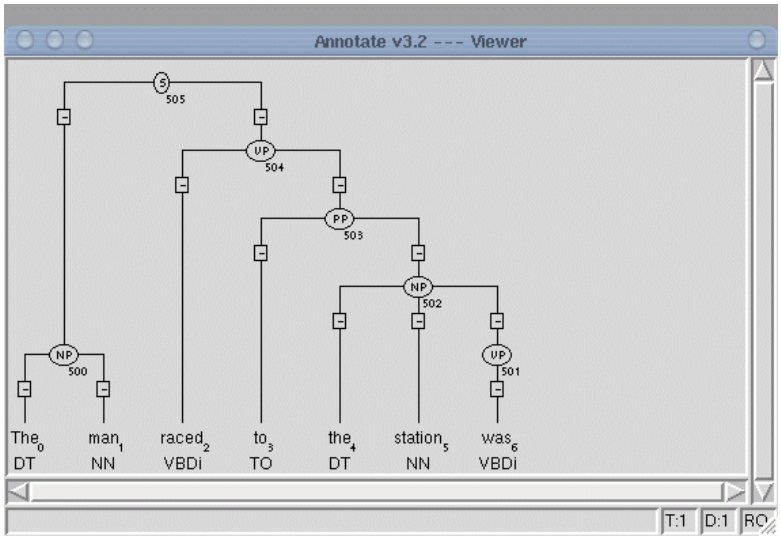




Reduced Relative Clause

- Parsers can make wrong decisions that lead them up the garden path

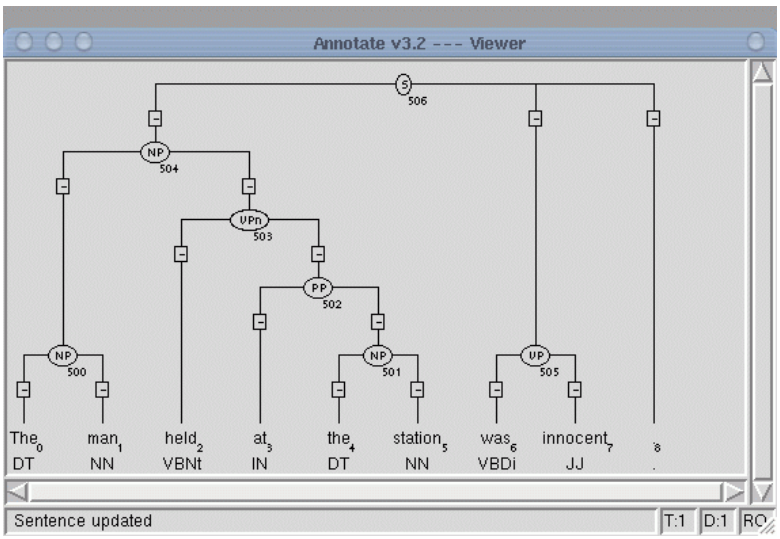
"The man raced to the station was innocent"



The Problem

- In some cases is may be possible to recover from the error earlier

"The man held at the station was innocent"



Experimental Methods

- We can use controlled experiments of reading times to investigate local ambiguity resolution
  - (a) The man held at the station was innocent (LA-trans)
  - (b) The man raced to the station was innocent (LA-intrans)
  - (c) The man that was held at the station was innocent (UA)

- Compare the reading times of ambiguous and unambiguous conditions.
  - Need a “linking hypothesis” from theory to measures
  - Can then manipulate other linguistic factors to determine their influence on on RTs in a controlled manner

Reading Methods

- Whole sentence reading times:

  The man held at the station was innocent

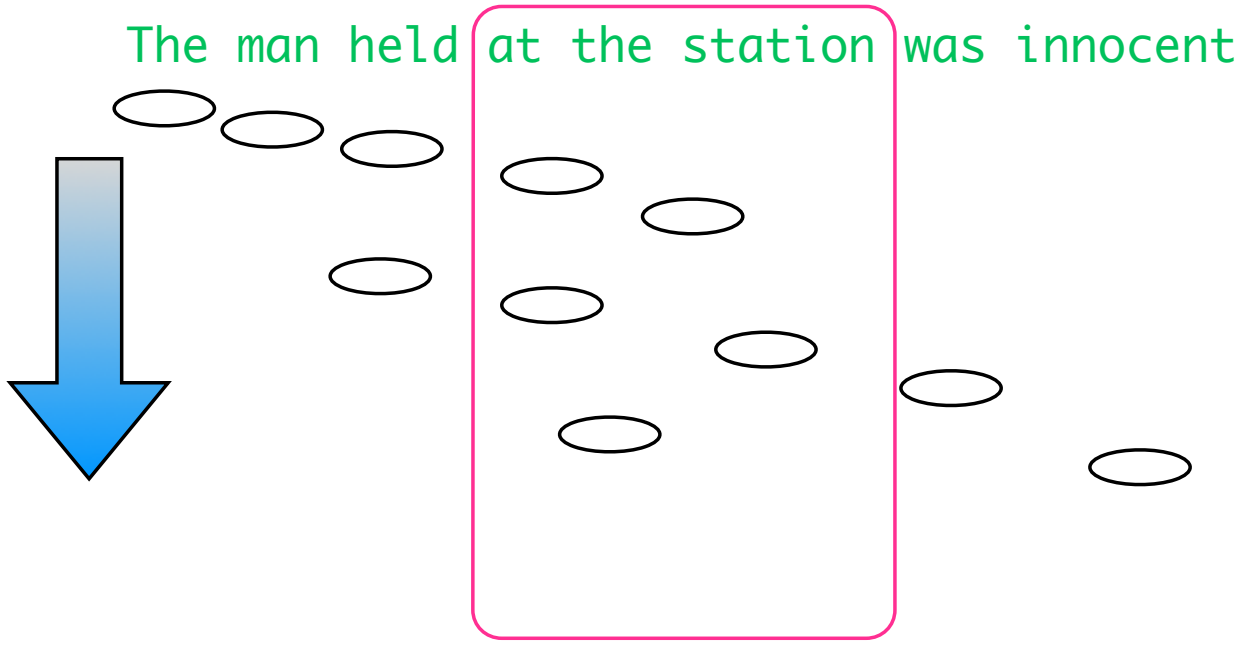
- Self-paced reading, central presentation:

  is the innocent

- Self-paced reading, moving window:

  The man held at the station was innocent

Eye-tracking: Difference Measures

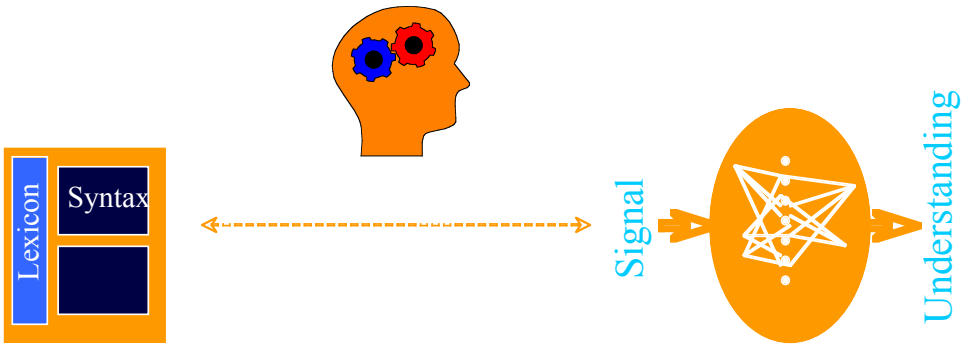


Summary of Methods

- People construct interpretations incrementally:
  - People must resolve ambiguity
  - Sometimes we must revise our interpretation of the sentence so far
- On-line measures can tell us about how/when this occurs
  - Reading times, ERPs, gaze in visual scene
- Experiments exploit these methods (and others) to investigate the underlying processing architectures and mechanisms

The Modularity Issue

- Is language distinct from other cognitive & perceptual processes?
  - e.g. vision, smell, reasoning ...
- Do distinct modules exist *within* the language processor?
  - e.g. word segmentation, lexical access, syntax ...
- What is a module anyway!?



Kinds of Models

- Constrained models
  - Emphasis on cognitive constraints such as memory & modularity
  - Evidence from difficulty structures, ambiguities, pathologies
  - Mechanisms are proposed which are optimized for these
- Unconstrained models
  - Evidence that comprehension is generally fast, accurate, effortless
  - Mechanisms emphasize optimal use of relevant information

Two Theories of Human Parsing

- What mechanisms is used to construct interpretations:
  - Frazier: Serial parsing, with reanalysis
  - McRae: Competitive activation of alternatives in parallel
- What information is used to determine preferred structure:
  - Frazier: General syntactic principles
  - McRae: Competitive integration of constraints

Models of Sentence Processing

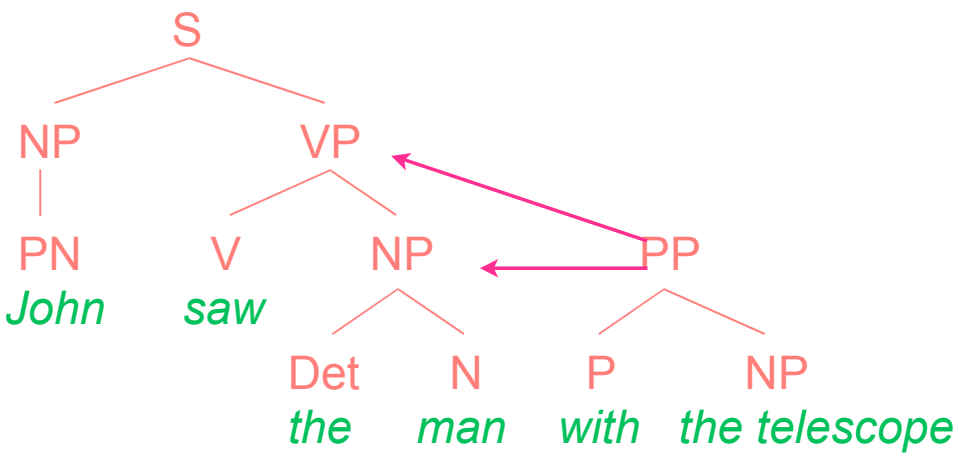
- Theories of parsing must specify …
  - what **mechanism** is used to construct interpretations?
  - which **information** sources are used by the mechanism?
  - which **representation** is preferred/constructed when ambiguity arises?
- **Linking Hypothesis**: Relate the theory/model to some observed measure
  - Preferred sentence structures should have faster reading times in the disambiguating region than dispreferred

Frazier: Garden Path Theory

- 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

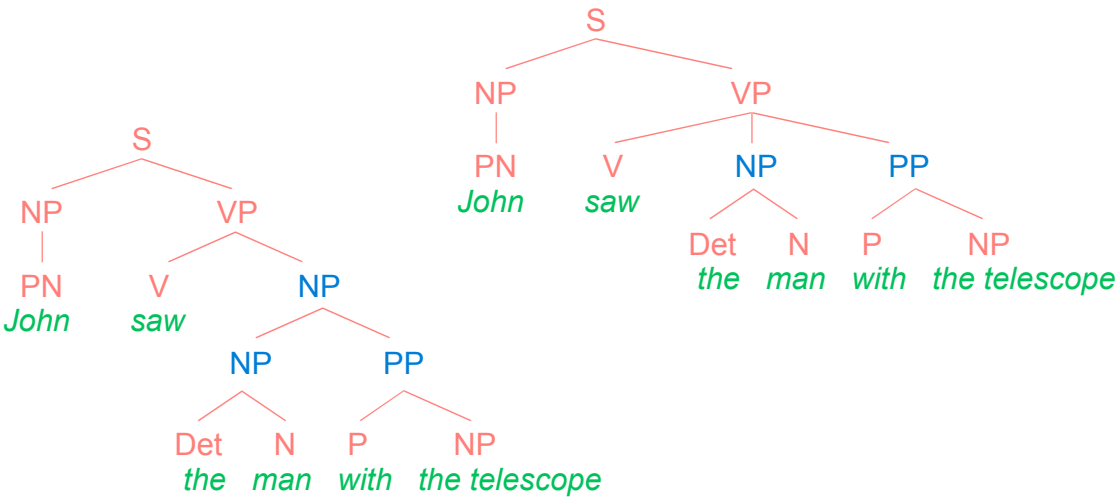
The Garden Path Theory (Frazier)



Which attachment do people initially prefer?

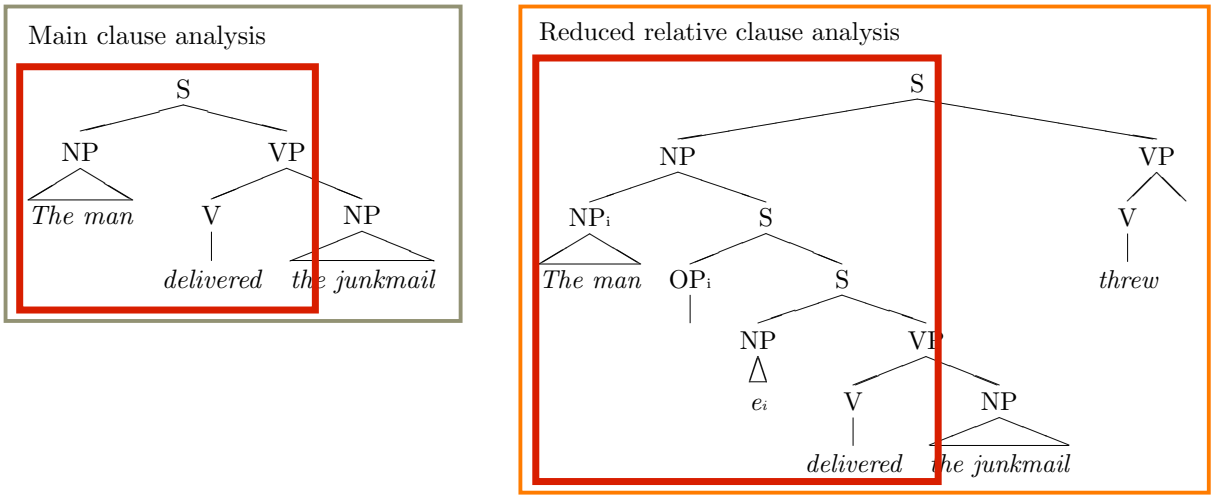
First Strategy: Minimal Attachment

Minimal Attachment: *Adopt the analysis which requires postulating the fewest nodes*



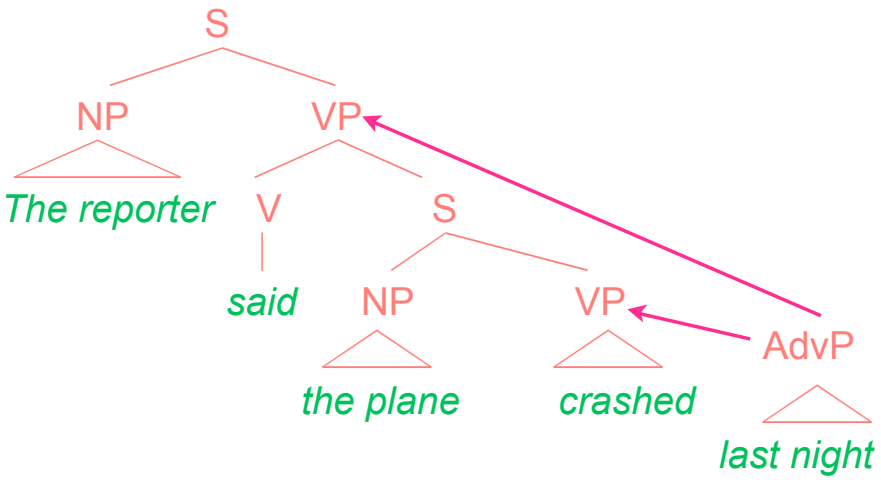
(Reduced) Relative Clauses

- One of the most studied syntactic ambiguities:
  - The man delivered the junkmail ... threw it away.



Second Strategy: Late Closure

Late Closure: *Attach material into the most recently constructed phrase marker*



Garden-Path Theory: Frazier

- What architecture is assumed?
  - Modular syntactic processor, with restricted lexical (category) and semantic knowledge
- What mechanisms is used to construct interpretations?
  - Incremental, serial parsing, with reanalysis
- What information is used to determine preferred structure?
  - General syntactic principles based on the current phrase stucture
- Linking Hypothesis:
  - Parse complexity and reanalysis cause increased RTs

Against linguistic modularity

- Empirical evidence from on-line methods
  - later evidence for “immediate” (very early) interaction effects of animacy, frequency, plausibility, discourse context
    - *The cop/crook arrested for taking bribes was found guilty*
- Appropriate computational frameworks:
  - symbolic constraint-satisfaction systems
  - connectionist systems & competitive activation models
- Homogenous/Integrative Linguistic Theory: HPSG
  - multiple levels of representation within a unified formalism

The Competitive-Integration Model (McRae et al, 1998)

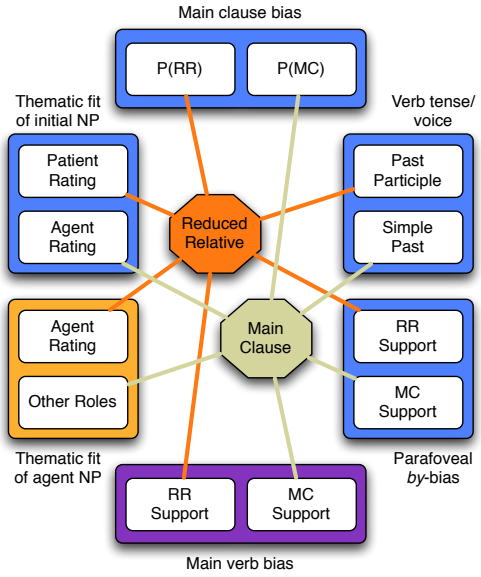
- Claim: Diverse constraints (linguistic and conceptual) are brought to bear simultaneously in ambiguity resolution.
- The Model: *Assumes the all analyses are constructed*
  - Constraints provide “probabilistic” support for analyses
    - Constraint are weighted and normalized
    - Lexical & structural bias, parafoveal cues, thematic fit ...
- Goal: Simulate reading times
  - RTs are claimed to correlate with the number of cycles required to settle on one of the alternatives

“No model-independent signature data pattern can provide definitive evidence concerning when information is used”

The Computational Model

The crook arrested by the detective was guilty of taking bribes

1. Combines constraints as they become available in the input
2. Input determines the probabilistic activation of each constraint
3. Constraints are weighted according to their strength
4. Alternative interpretations compete to a criterion
5. Cycles of competition mapped to reading times



Constraint Parameters

"The crook/cop arrested by the detective was guilty of taking bribes"

Verb tense/voice constraint: verb bias towards past or past participle
Relative log frequency is estimated from corpora: RR=.67 MC=.33

Main clause bias: general bias for structure for "NP verb+ed …"
Corpus: P(RR|NP + verb-ed) = .08, P(MC|NP + verb-ed) = .92

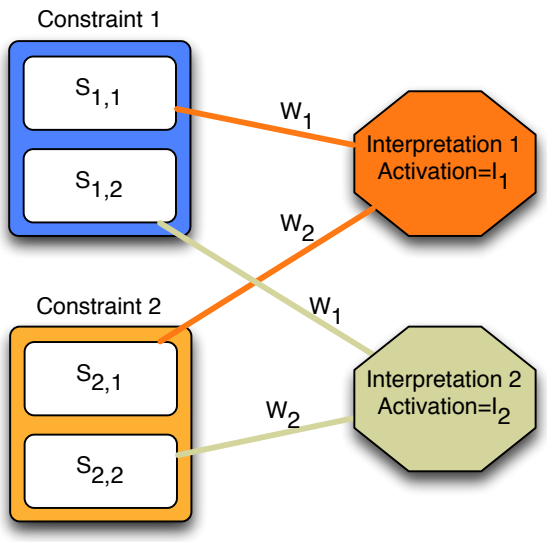
by-Constraint: extent to which 'by' supports the passive construction
Estimated for the 40 verbs from WSJ/Brown: RR= .8MC= .2

Thematic fit: the plausibility of crook/cop as an agent or patient
Estimated using a rating study

by-Agent thematic fit: good Agent is further support for the RR vs. MC
Same method as (4).

The recurrence mechanism

- $S_{c,a}$ is the raw activation of the node for the c^{th} constraint, supporting the a^{th} interpretation,
- w_c is the weight of the c^{th} constraint
- I_a is the activation of the a^{th} interpretation
- 3-step normalized recurrence mechanism:
  - Normalize: $S_{c,a}(norm) = \frac{S_{c,a}}{\sum_a S_{c,a}}$
  - Integrate: $I_a = \sum_c \left[w_c \cdot S_{c,a}(norm) \right]$
  - Feedback: $S_{c,a} = S_{c,a}(norm) + I_a \cdot w_c \cdot S_{c,a}(norm)$



$$\sum_i w_i = 1$$

Constraint-based Models

- What architecture is assumed?
  - Non-modular: all levels are constructed and interact simultaneously
- What mechanisms is used to construct interpretations?
  - Parallel: ranking based on constraint activations
- What information is used to determine preferred structure?
  - All relevant information and constraints use immediately
- Linking Hypothesis:
  - Comprehension is easy when constraints support a common interpretation, difficult when they compete. Determined by cycles required at each word.

A Good Model ...

- Should make independently motivated assumptions
  - architectures, representations, memory etc.
  - identify arbitrary decisions needed for implementation
- Should have clearly specified parameters
  - clearly defined objective methods for setting parameters
- A clear linking hypothesis to empirical measures
- Should "predict" unseen data, generate testable hypotheses

Falsifying model predictions

- A central goal of building models is to generate testable predictions
- How do we respond when data disconfirms predictions?
  - At what level were the predictions made?
  - Is the problem with the model, or with the linking hypothesis?
  - Can the model/link be easily revised to account for the data (and still account for previous data)
- Beware of skeet shooters!

Summary

- People are extremely good at understanding language
  - fast, accurate, robust and adaptive to context
- There are some “pathologies”, where processing is imperfect
  - centre-embedding, ambiguity resolution, garden paths
- These findings are used to shape the development of models
  - serial, parallel, competitive activation -- modular, interactive
  - rule-based, constraint-based or probabilistic
- Models make predictions, so we run more experiments!

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]]”

Tutorial Tomorrow

- Look at the syntactic alternatives locally and globally ambiguous sentences
  - Identify at which word ambiguity arises / at which word disambiguation occurs?
  - How do you think people resolve the local ambiguity (what's your preference)?
  - At the point of ambiguity, which structure does Frazier's theory predict will be constructed?
  - Tell what additional kinds of information would influence processing in an interactive model like McRae's.
- How to design an experiment to test whether Frazier or McRae is right?

For the exam ...

- Be familiar with the lecture & tutorial material !
- A good supplement to the lecture is:
  - Gerry Altmann. Ambiguity in Sentence Processing. *Trends in Cognitive Sciences*, 2(4), 1998.

    – available from course web page –