# FLST:Cognitive Foundations II

Matthew W. Crocker

crocker@coli.uni-sb.de

---

# Summary of cognitive issues

- The relation between language and thought
  - language - culture mutually constraining
  - autonomy of language vs *mentalese*
- Linguistic autonomy
  - Modularity and localization in the brain (these aren't the same thing)
  - Innate linguistic (domain specific) language "organ"
- Distinction between animal "communication" and human language
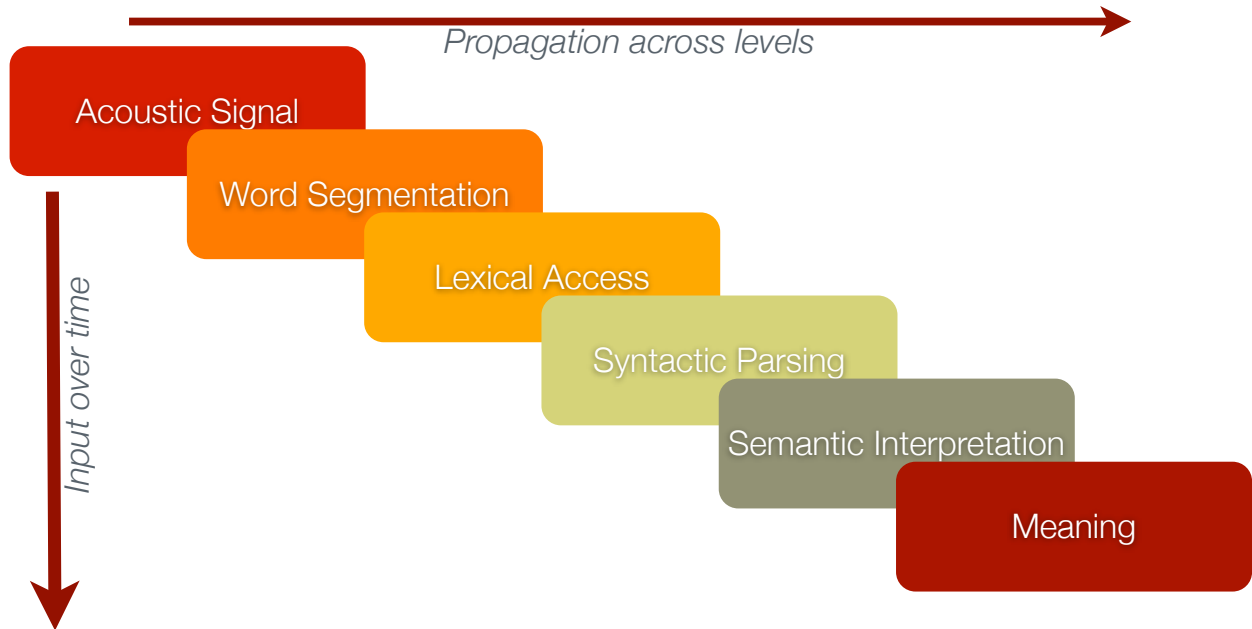- The evolution & emergence of the capacity for human language

# Human language processing

- ♣ People are highly accurate in understanding language

- ♣ People process language rapidly, in real-time

- ♣ People understand and produce language incrementally

- ♣ People even anticipate what's going to be said next

- ♣ People rapidly adjust to context, and are robust

- ♣ People achieve this despite limitations on processing resources

- ♣ People do make some interesting errors, and exhibit breakdown in certain situations ...

---

# Human language processing

- ♣ People are highly **accurate** in understanding language

- ♣ People process language **rapidly**, in **real-time**

- ♣ People understand and produce language **incrementally**

- ♣ People even **anticipate** what's going to be said next

- ♣ People rapidly adjust to **context**, and are **robust**

- ♣ People achieve this despite **limitations** on processing resources

- ♣ People do make some interesting **errors**, and exhibit **breakdown** in certain situations ...

# Sound to Meaning over Time

*Propagation across levels*

Acoustic Signal

Word Segmentation

Lexical Access

Syntactic Parsing

Semantic Interpretation

Meaning

*Input over time*

# Theories of Sentence Processing

♣ Language is complex & dynamic

   ♣ multiple levels of representation & knowledge

   ♣ each level has rich internal structure, unique constraints & representations

   ♣ processing unfolds over time: both across levels, and in response to signal

   ♣ levels interact in dynamically, and in complex ways

♣ We need computational models to understand ...

   ♣ the dynamics & interactions of processing; the role of processing limitations

   ♣ relate processing with empirical data; make predictions

# Sentence processing

- Sentence processing is the means by which the words of an utterance are combined to yield and interpretation

    - All people do it well

    - It is a difficult task: complexity and ambiguity

    - Not simple 'retrieval', like lexical access

- **Compositional**: interpretation must be constructed on-line, rapidly

    - Even for sentences with novel structures, or words used in novel positions

# Human Language Processing

- We understand language incrementally, word-by-word

    - *How do people construct 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

# Theories 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

---

# Theories of Linguistic Knowledge

- Theories of Syntax

    - **Representations**: Trees, feature structures, dependencies

    - **Structure building**: PS-rules, transformations, unification, composition, tree substitution

    - **Constraints on representations**: Case marking, theta-Criterion, c-command, binding principles, head-foot principle

- Competence Hypothesis

    - The mechanisms of language comprehension directly utilize the rules and representations of the linguistic theory

# Strong competence & modularity

🔶 Fodor's proposals emphasis language as a module, distinct from other perceptual cognitive abilities

🔶 Linguistic theories suggest that language itself may consist of sub-levels: phonology, morphology, syntax, semantics ...

🔶 Each with different <u>rules</u> and <u>representations</u>

🔶 Do these correspond to distinct processes?

🔶 Are these processes modules?

🔶 Which of Fodors characteristics might they have/not have?

# A Modular Architecture
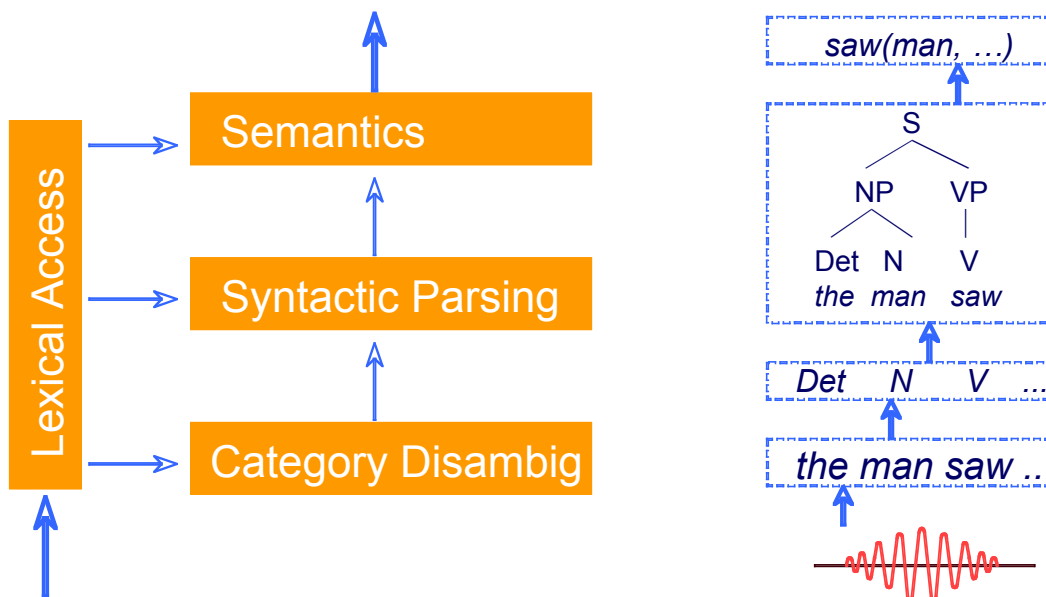
# Kind of Mechanisms

- Assume we believe that syntactic structure building is underlies sentence comprehension

- Questions:

  - What kinds of information are used:

    - lexical, grammatical, frequency, semantics, ...

  - What kinds of representations:

    - trees, dependencies, AVMs, distributed representations

  - What kind of mechanisms:

    - serial/parallel, symbolic/probabilistic/connectionist

---

# The Problem

- How do people incrementally recover the meaning of an utterance?



"The man held at the station was innocent"

Crocker & Brants, *Journal of Psycholinguistic Research*, 2000.
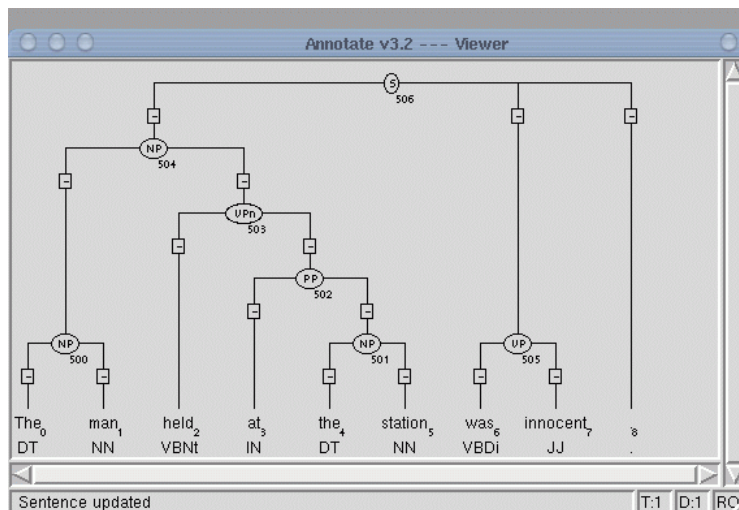
# Experimental Methods

♣ We can use controlled experiments of reading times to investigate local ambiguity resolution

♣ (a) The man held at the station was innocent (LA)

♣ (b) The man who was held at the station was innocent (UA)

♣ Compare the reading times of (b) where there is no ambiguity, with (a) to see if and when the ambiguity causes reading difficulty.

   ♣ 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:

                 isthldehelabont

♣ Self-paced reading, moving window:

   The man held at the station was innocent

# Eye-tracking: Difference Measures

The man held at the station was innocent
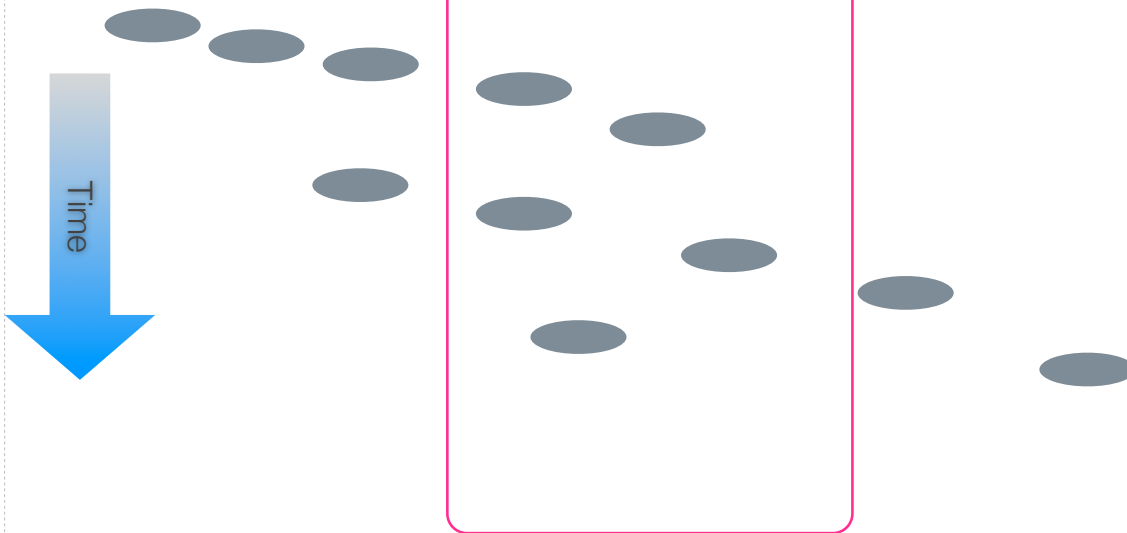
Time

# Eye-tracking: First Fixation

The man held at the station was innocent
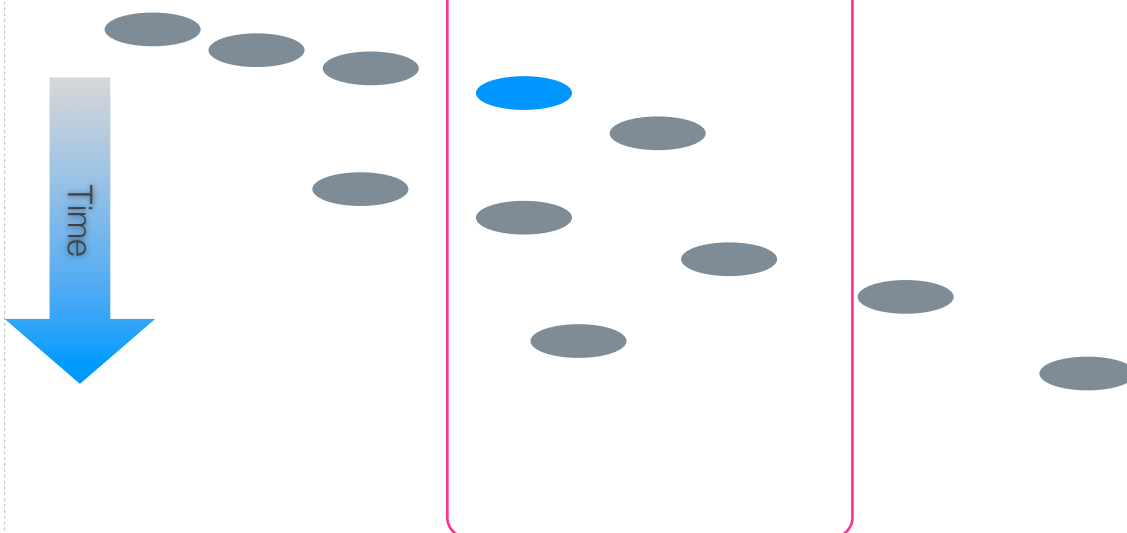
Time

# Eye-tracking: First Pass

The man held at the station was innocent
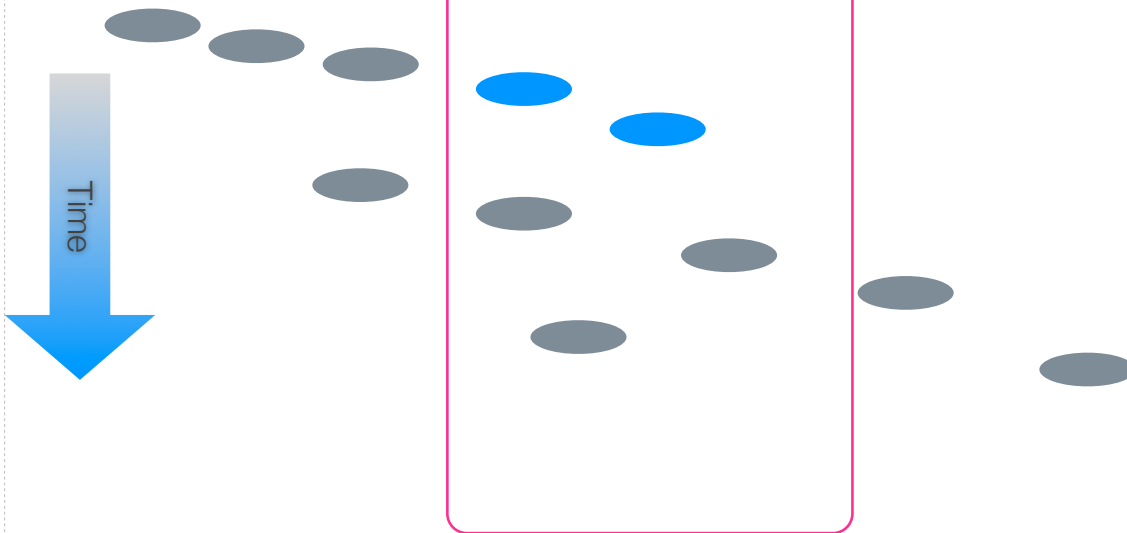
# Eye-tracking: Total time

The man held at the station was innocent
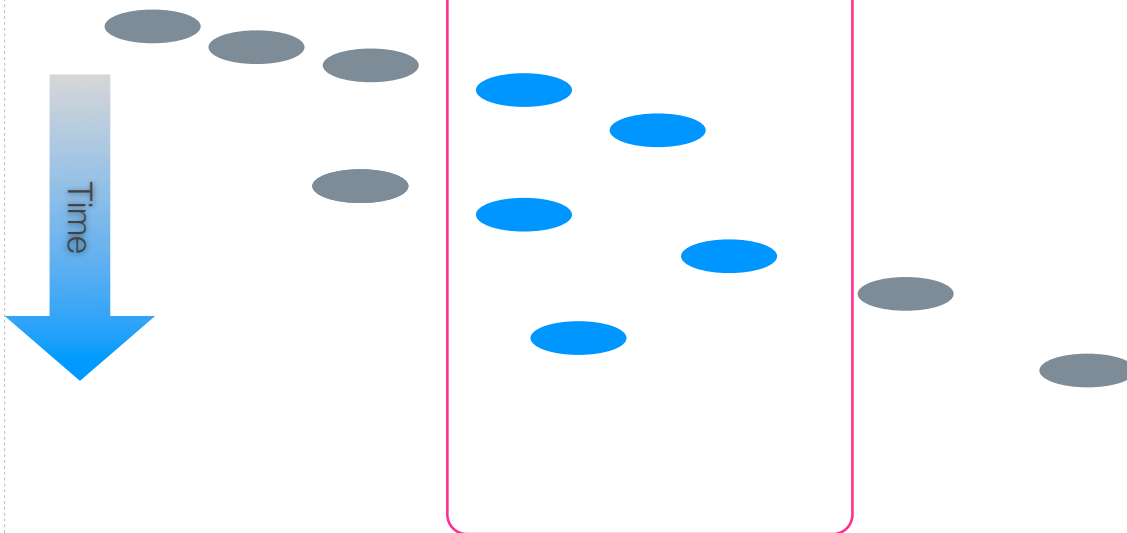
# Eye-tracking: Regression Path

The man held at the station was innocent

Time

---
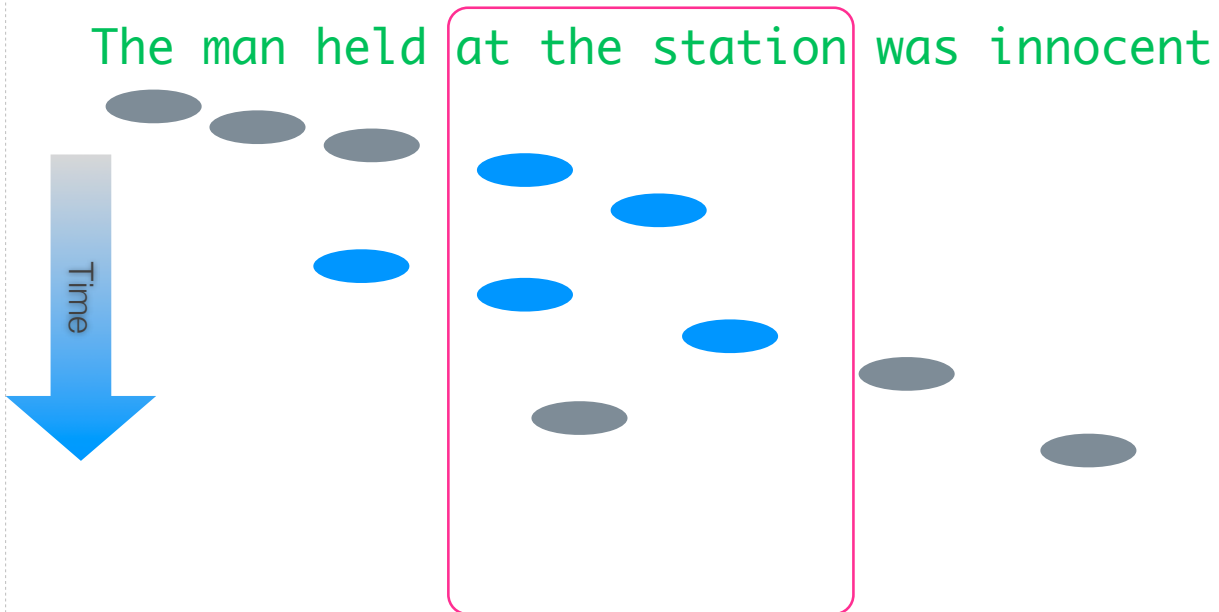
# Experiments (continued)

- Think about what "confounds" might limit your interpretation of the results (e.g. length, meanings ... )

- Create a set of similar sentence pairs that minimize confounds

  - add "filler" sentences

- Choose the right experimental method based on the behavior you're expecting

- Difference in reading times in the disambiguating region?

  - Yes: support for your theory!

  - No: "null result", no support for your theory, but also doesn't prove the

# Two Theories of Human Parsing

♣ What mechanisms is used to construct interpretations:

♣ Frazier: Serial parsing, with reanalysis

♣ McRae: Competitive activation of alternatives

♣ What information is used to determine preferred structure:

♣ Frazier: General syntactic principles

♣ McRae: Competitive integration of constraints

# The 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)
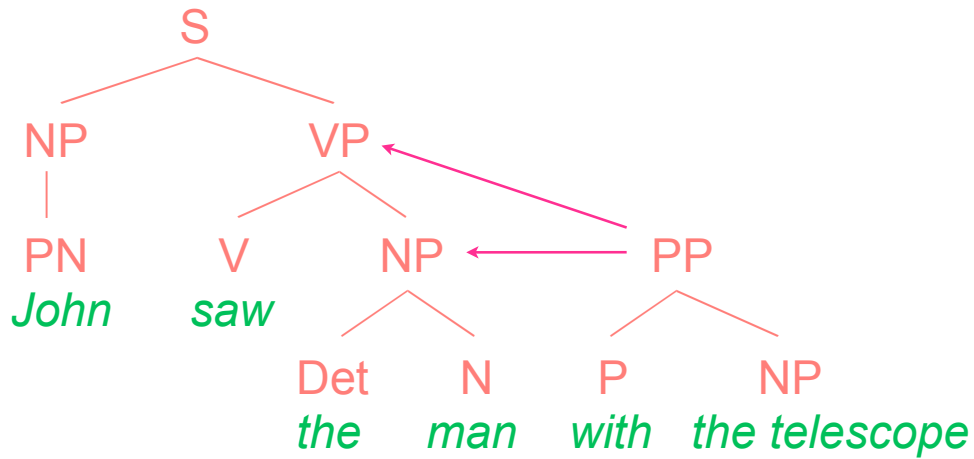


Which attachment do people initially prefer?

# First Strategy: Minimal Attachment
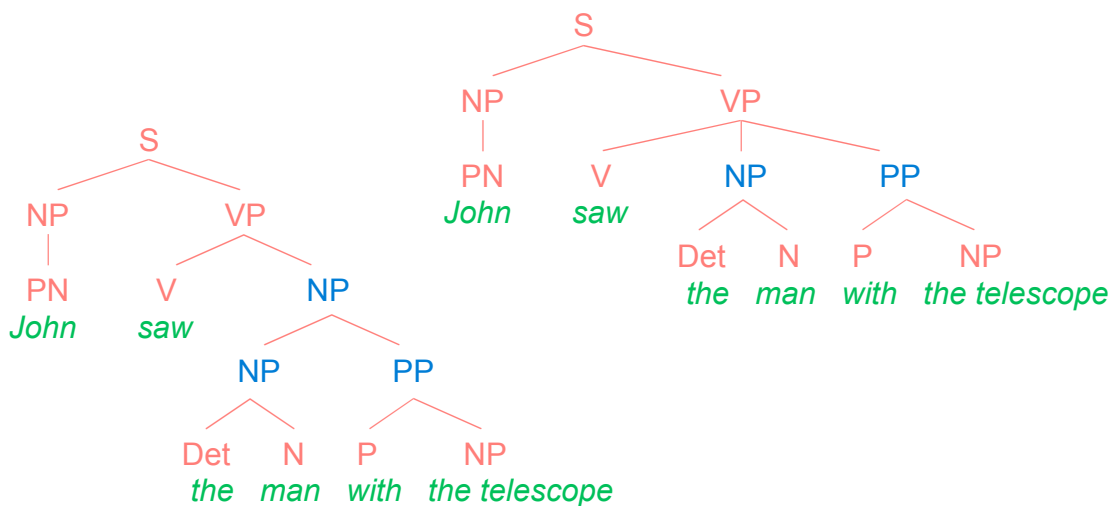
# Second Strategy: Late Closure



---
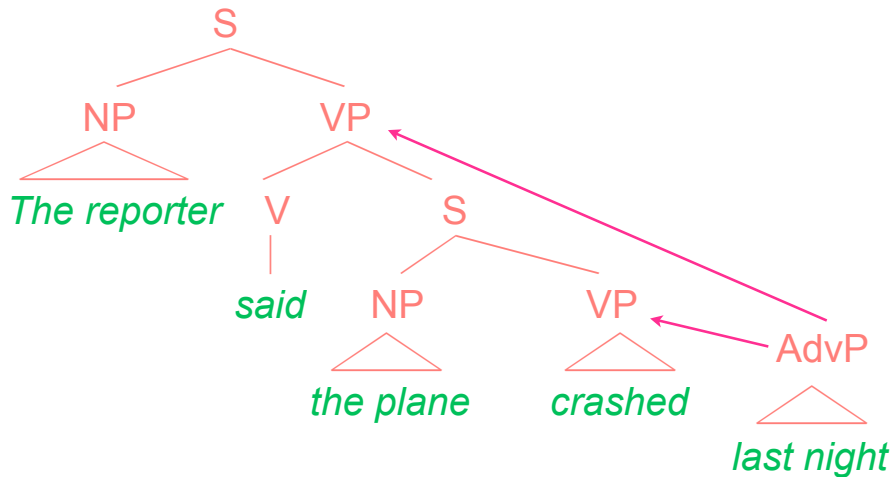
# Support for Linguistic Modularity

- Modular lexical access versus syntax: Forster

  - all possible word meanings temporarily available

  - no immediate influence of syntactic context

- Modular syntax versus semantics: Frazier

  - initial attachment ambiguities resolved by purely structural preferences

  - no immediate effect of semantics or context

- Dissociation in language impairment at different levels

  - lexical, syntactic, semantic; production versus comprehension

# Against linguistic modularity

- Empirical evidence from on-line methods

  - later evidence for "immediate" (very early) interaction effects of animacy, frequency, plausibility, discourse context …

    - *The woman/patient sent the flowers was pleased*

- 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

---

# Multiple constraints

"The man/lecture held/fought/given at the station …

         *… a copy of the NY times that he had bought at the airport"* [Main Clause]

         *… was rather boring"*                         [Relative Clause]

**Prosody**: intonation can assist disambiguation, does it in this case?

**Lexical** preference: *held* = {Past, **PastPart**}, *fought* = {**Past**, PastPart}, *given* = {PastPart}

**Subcat**: *held* = { [ _ NP] [ _ NP PP]}**,** *fought* = { [ _ ] [ _ NP]}
         *given* = { [ _ NP PP] [ _ NP NP]}

**Semantics**: Referential context, plausibility

- **Reference**: is there more than one man in the context? Yes: prefer relative clause. Why?

- **Plausibility**: of *man* versus *lecture* as Agent/Patient of the verb

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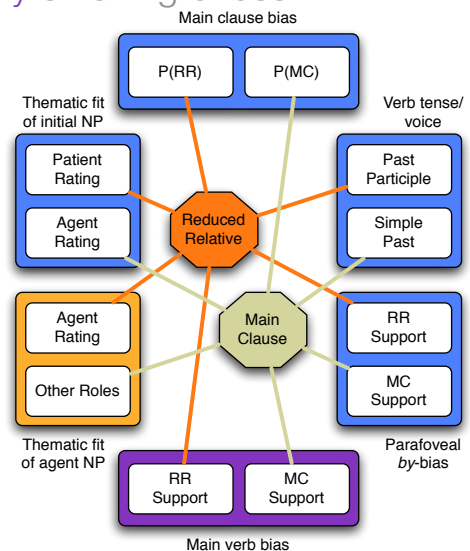
*"The crook/cop **arrested by** the detective was guilty of taking bribes"*

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= .8 MC= .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).

---

- $S_{c,a}$ is the raw activation of the node for the $c^{th}$ constraint, supporting the $a^{th}$ interpretation,
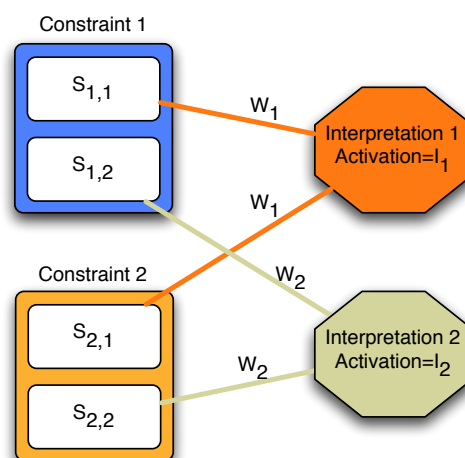
- $w_c$ is the weight of the $c^{th}$ constraint

- $I_a$ is the activation of the $a^{th}$ interpretation

- 3-step normalized recurrence mechanism:

  - Normalize: $S_{c,a}(norm) = \dfrac{S_{c,a}}{\sum\limits_{a} S_{c,a}}$

  - Integrate: $I_a = \sum\limits_{c}\left[w_c \cdot S_{c,a}(norm)\right]$

  - 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

# 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!

♣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]]"

---

# Other experimental methods

♣ Reading-time experiments:

   ♣ Natural: reading is an important comprehension modality

   ♣ Intuitive: reading times reveal processing complexity

♣ Neuroscientific methods:

   ♣ associate certain processes with regions of the brain

   ♣ certain kinds of EEG components indicate different kinds of cognitive processing

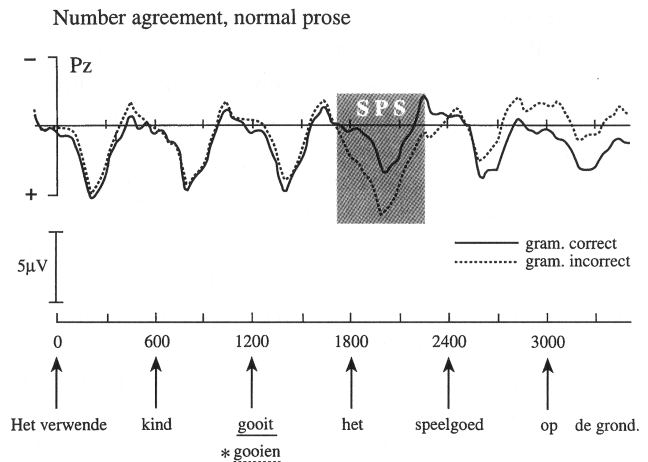♣ Visual attention: reveals interpretation more directly

   ♣ These methods can be used with spoken language

# Neuroscientific Measures: ERPs

♣ Syntactic and semantic processes are partially revealed by signature patterns in EEGs: Event-Related Potentials (ERPs)

♣ Syntactic Anomaly: P600 or SPS
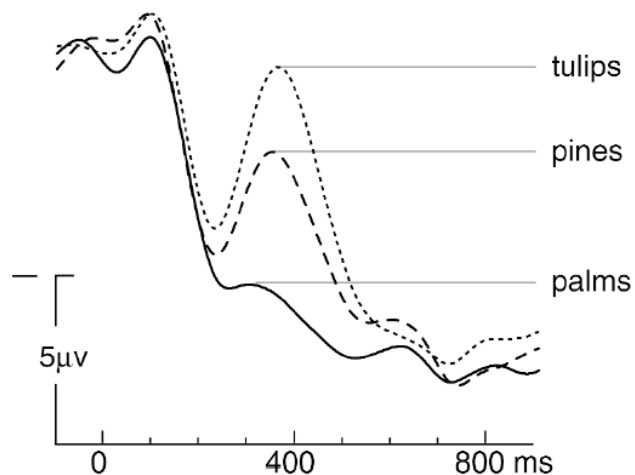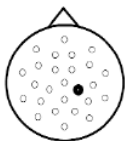


Number agreement, normal prose
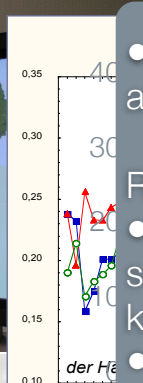
"The spoilt child throw(s) the toy on the ground"

# Semantic Anomaly: N400

'They wanted to make the hotel look more like a tropical resort. So along the driveway they planted rows of ...'



trends in Cognitive Sciences

# Anticipation in Visual Worlds
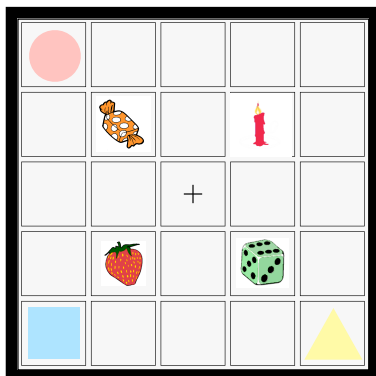


- On-line mediation of visual attention by spoken language

Rapid use of:
- morpho-syntax, verb semantics and world knowledge
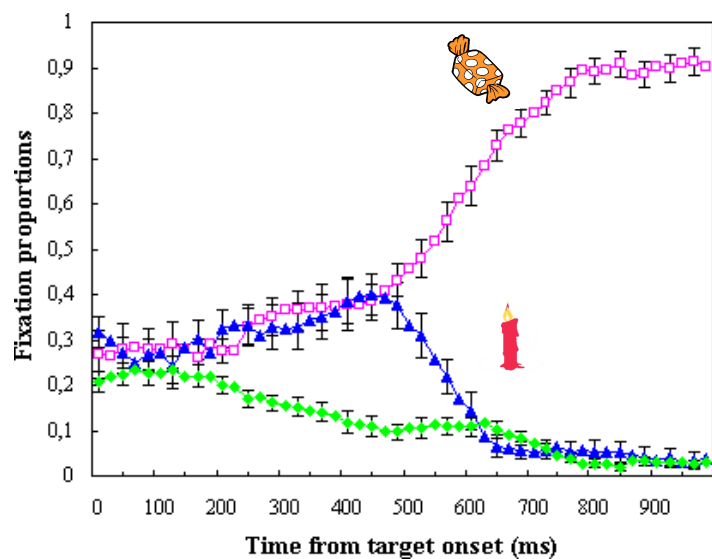- trigger anticipation of role-

| | | | | |
|---|---|---|---|---|
| SVO | Der Hase *The hare (nom)* | frisst *eats* | gleich *soon* | den Kohl *the cabbage* |
| OVS | Den Hasen *The hare (acc)* | frisst *eats* | gleich *soon* | der Fuchs *the fox (nom)* |

Kamide *et al*, *JPR*, 2003.

---

# Lexical access over time



"Pick up the *candle*"

# 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

- We can design experiments which exploit these methods (and others) to investigate the underlying processing architectures and mechanisms

# For the exam ...

- Be familiar with the lecture & tutorial material !

- Supplement it with the following reading:

    - Gerry T. M. Altmann. Ambiguity in Sentence Processing. *Trends in Cognitive Sciences*, Vol. 2, Num. 4, 1988.

Visual Processes

Memory

Attention

Cognitive Resources

Reading Times

Event Potentials

Visual Attention

Imaging

Performance

Human Language Processor

Context

Experience

Discourse/Dialogue

Environment

Task

Competence

Interpretation

Linguistic Complexity

Broad Coverage