

Architectures and Mechanisms

Lecture 12: Review Introduction to Psycholinguistics

Matthew W. Crocker
Pia Knoeferle

*Department of Computational Linguistics
Saarland University*

Architectures and Mechanisms

Representational autonomy: e.g. phonological versus syntax representations

- Possibly interactive processes

Procedural autonomy: e.g. lexical access versus syntax

- Possibly shared representations

How are any such “distinct subsystems” for language processing organised? How do they interact?

- Does organisation affect possible mechanisms?
- Theoretical, computational and empirical arguments for and against ‘modularity’?

Models of Language Processing

Competence Assumptions

- What knowledge and representations do we assume?

Architectures

- The spectrum from modularity to interaction

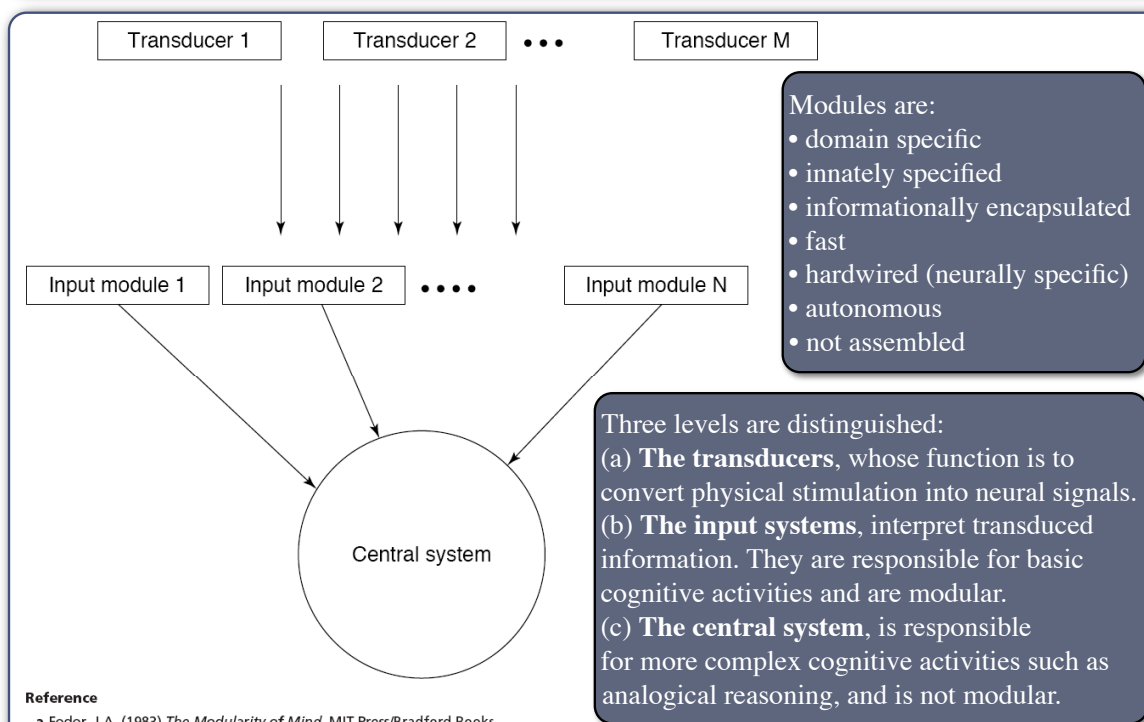
Mechanisms

- How are representations/interpretations recovered

Linking Hypothesis

- How do mechanisms map onto observed behaviour

Fodor's Modularity



Reference

a Fodor, J.A. (1983) *The Modularity of Mind*, MIT Press/Bradford Books

Linking Hypotheses

Different methods reveal different aspects of the underlying architectures and mechanisms

Reading times: relative processing difficulty

- correlated with processing complexity and reanalysis

Visual attention: reference and anticipation

- correlated with interpretation and inference

N-400: semantic anomaly

- correlated with semantic integration

P-600/SPS: syntactic anomaly

- correlated with disambiguation and reanalysis

Models of word recognition

Search models (e.g., Forster, 1976) - serial comparison

- Perceptual input has no direct access to lexical entries
 - a complete perceptual representation of the stimulus is constructed
- Forster's model is for both reading and listening

Direct Access models - parallel comparison

- Logogen Model (Morton, 1969), not domain specific
- Cohort Model (Marlsen-Wilson and Welsh, 1978), speech
- TRACE (McClelland and Elman, 1986), speech
- Shortlist/Merge (Norris, 1994), speech

Main similarities and differences

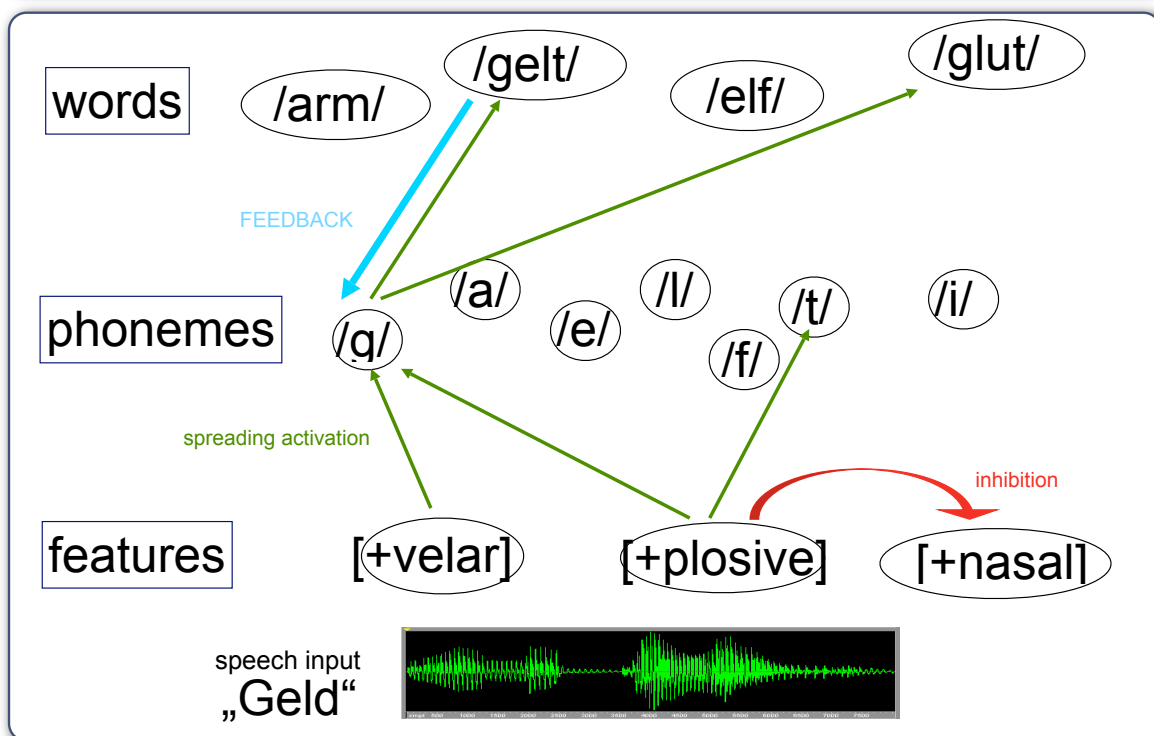
COHORT puts emphasis on word onsets

- limits the competitor set to
 - *beaker* and *beetle* but not *speaker*
- not very robust to noise and continuous speech

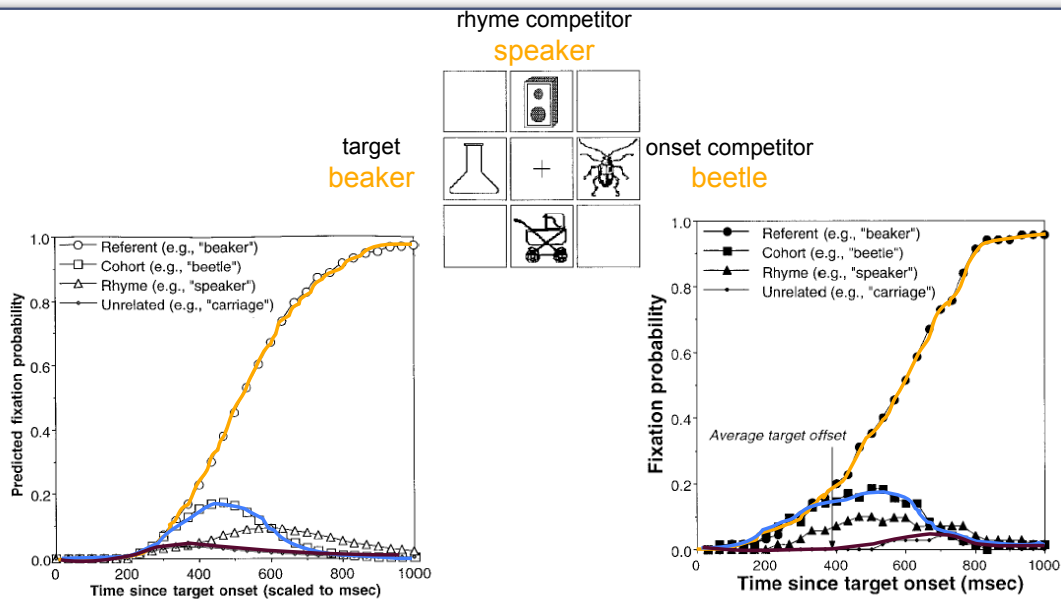
TRACE and ShortList allow all perceptual features to activate competitors ... e.g. predicts rhyme competition

- *beaker* and *beetle* and *speaker*
- TRACE allows top-down feedback
- ShortList does not

Sketch of TRACE



TRACE predictions, Allopenna et al. (1998)



Mechanisms for Language Processing

Syntactic processing requires a solution to the problem of:

- How structures are incrementally constructed
- How local and global ambiguity

Incremental Parsing

- Top-down
- Bottom-up
- Mixed strategy

Ambiguity and parsing

- Serial (deterministic/non-deterministic)
- Parallel (bounded/unbounded)

Comments on Left-Corner

Mixed data-driven and hypothesis driven approaches

- Eager corresponds to composition of partial structures

Oracle increases the top-down component, reduce ambiguity

Arc Standard: less ambiguity

- attach when constituents are complete: safer
- delayed attachment means more is kept on the stack

Arc Eager: less memory

- early composition reduces stack growth
- eager attachments are less bottom-up

Summary of Behaviour

Node	Arcs	Left	Centre	Right
Top-down	Either	$O(n)$	$O(n)$	$O(1)$
Shift-reduce	Either	$O(1)$	$O(n)$	$O(n)$
Left-corner	Standard	$O(1)$	$O(n)$	$O(n)$
Left-corner	Eager	$O(1)$	$O(n)$	$O(1)$
People		$O(1)$	$O(n)$	$O(1)$

Theories of Sentence Processing

Structure-based theories

- Disambiguation based on structural heuristics

Grammar-based theories

- Preferred structure based on grammatical principles

Experience-based theories

- Structural preferences are based on prior experience

Interactive accounts

- Disambiguation draws on diverse knowledge sources

Resources-based accounts

- Preferred structure involves the least resources

Theories of Sentence Processing

Relate the theory/model to some observed measure

- Typically impossible to predict measures completely

Theories of parsing typically determine ...

- 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:

- Preferred sentence structures should have faster reading times in the disambiguating region than dispreferred

Garden-Path Theory: Frazier (1978)

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 structure

Linking Hypothesis:

- Parse complexity and reanalysis cause increased RTs

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

Garden-Path Theory: Jurafsky (1996)

What **architecture** is assumed?

- Modular lexico-syntactic processor with lexical (category and subcategory), no semantic knowledge

What **mechanisms** is used to construct interpretations?

- Incremental, bounded parallel parsing, with reranking

What **information** is used to determine preferred structure?

- Lexical and structural probabilities

Linking Hypothesis:

- Parse reranking causes increased RTs, if correct parse has been eliminated, predict a garden-path

Constraint-based Models

What **architecture** is assumed?

- Non-modular: all levels of representation 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

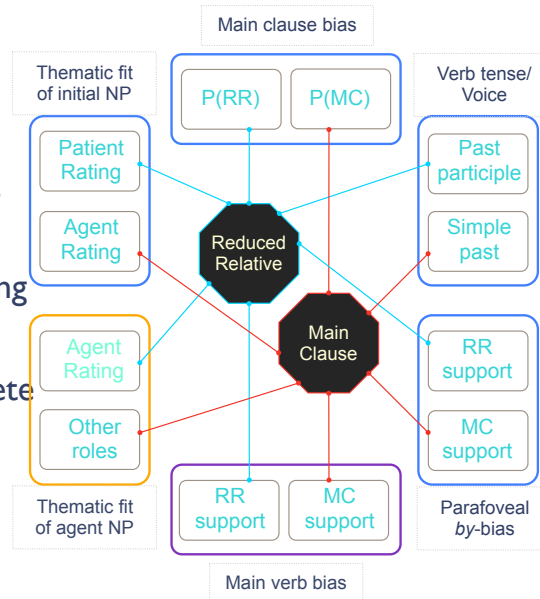
Linking Hypothesis:

- Comprehension is easy when constraints support a common interpretation, difficult when they compete

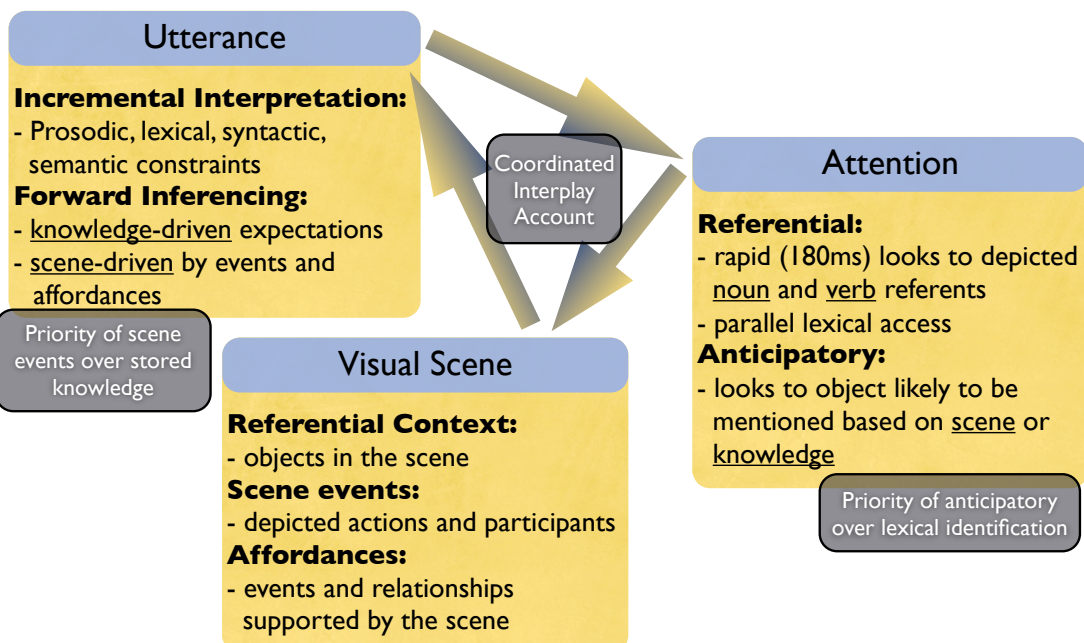
The Computational Model

The crook arrested by the detective was guilty of taking bribes

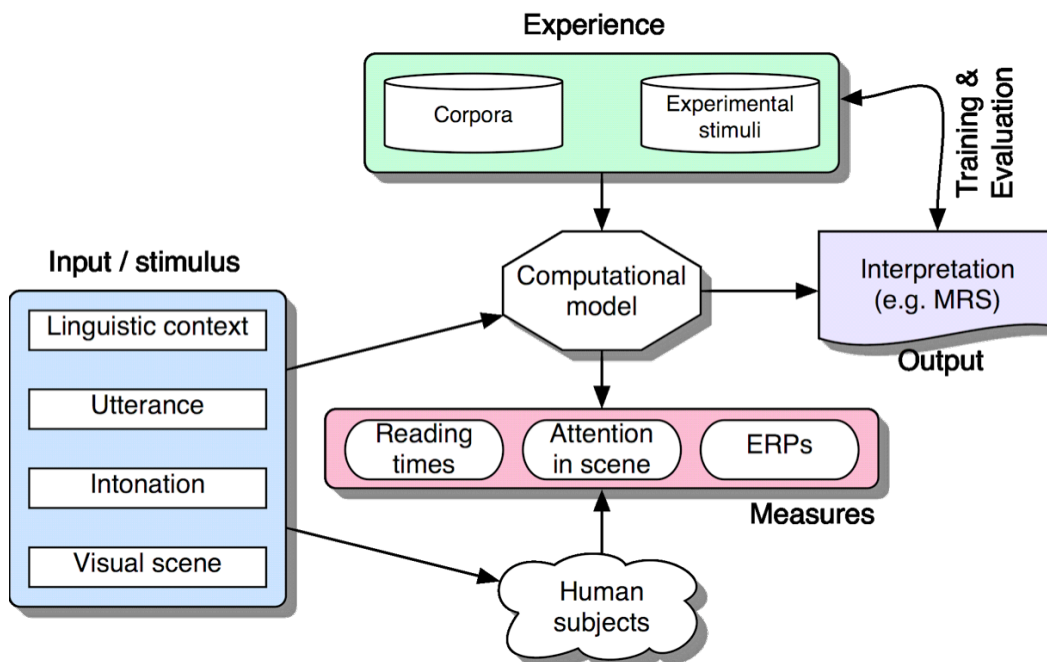
1. Combines constraints as they become available in the input
2. Input determines the probabilistic activation of each constraint
3. Constraints are weighted according to their strength
4. Alternative interpretations compete to a criterion
5. Cycles of competition mapped to reading times



Situated Language Processing



Situated models



A Connectionist Model of Scene & Sentence

Trained to model materials from 5 visual world studies

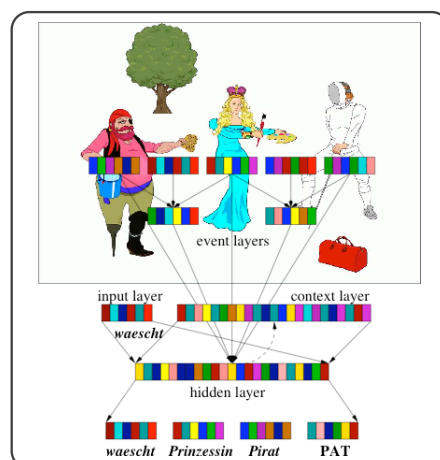
- SRN + Scene

Successfully models the use of:

- experience
- immediate scene
- sentence alone
- priority of the scene

Exhibits anticipatory behaviour

What about modeling attention in the scene?



The Exam

26th July at 9:00, Musiksaal

- **90 minutes**
- **Notes: one sheet of A4 paper**
 - **self-written!**

Format

- Part I (40%): 8 short answer questions (~5 mins each)
- Part II (60%): 3 long questions (20% --15 mins each)
 - Choose 3 from the 4 questions provided