Language Science & Technology: Cognitive Foundations

Matthew W. Crocker
Chair of Psycholinguistics

Department of Computational Linguistics
Saarland University

Computational *Psycho*linguistics

"To understand and model the processes that underlie the human capacity to understand language"

- How does the human language processor work?
- How is it realized in the brain?
- How can we model it computationally?
- Where does it come from?

How does language interact with other cognitive systems and the environment?

Cognitive Models of Language Processing

Not just about making computers understand language

- Using to computers to model *human* comprehension

Similarities to (Computational) Linguistics

- Competence hypothesis
  - Need to recover the meaning of language
  - Shared assumptions about representations
- Similar mechanisms: probabilistic, symbolic, learning …

Cognitive Models of Language Processing

Differences with Computational Linguistics:

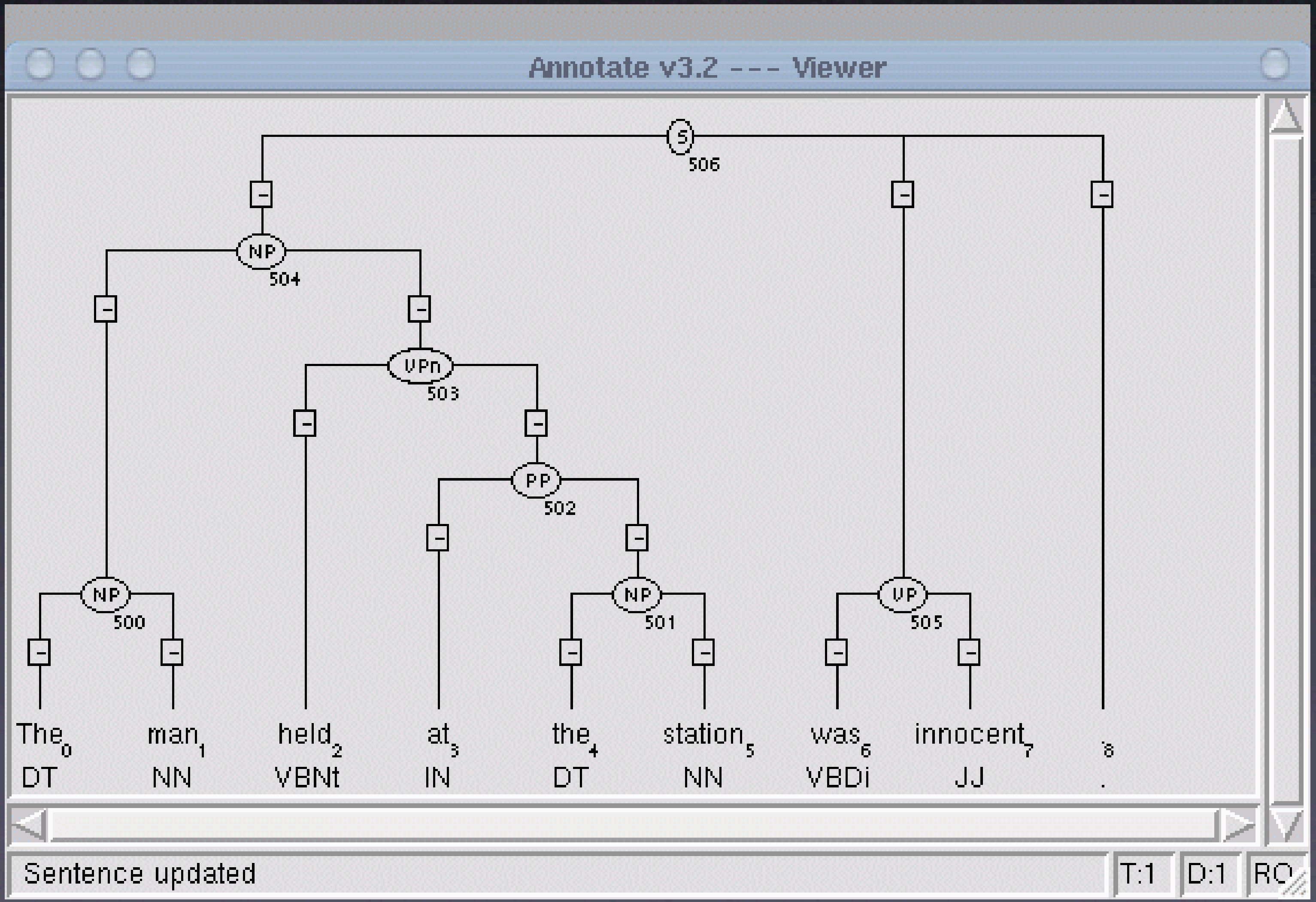
- People are highly adaptive, and context sensitive
- People are accurate and fast
- Incremental, word-by-word
- Some limitations that computers don't have: memory
- Psychological plausibility of computational mechanisms

In addition to understanding language, we want to model *on-line human behaviour*, or "performance"

The Problem

How do people recover the meaning of an utterance, with respect to a given situation, in real-time?

"The man held at the station was innocent"



Crocker & Brants, *Journal of Psycholinguistic Research*, 2000.

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

Decisions are sometimes wrong!

- What information is used to identify we made a mistake
- How do we search for an alternative

Methods for Investigating Human Behaviour

Whole sentence reading times:

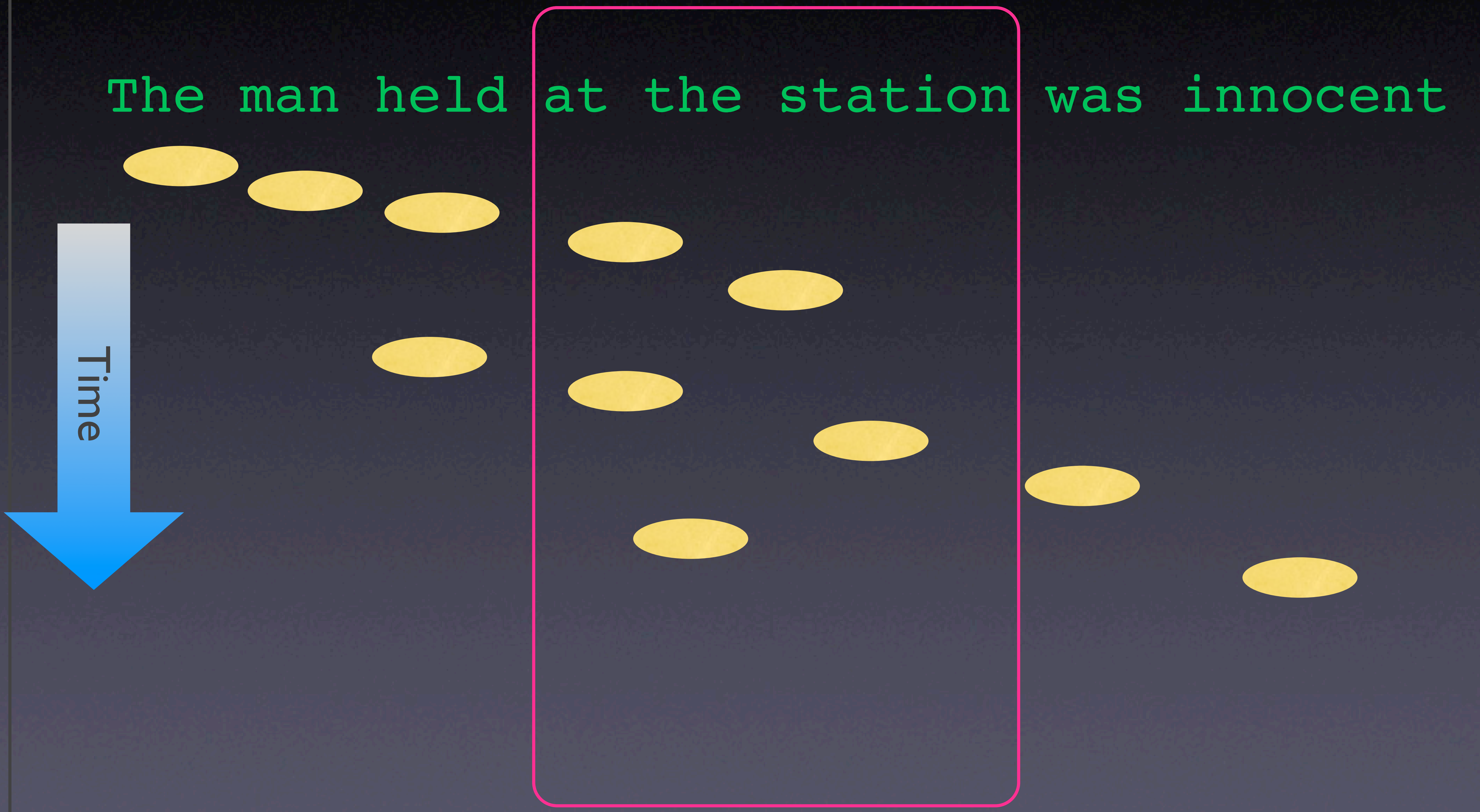
`The man held at the station was innocent`

Self-paced reading, central presentation:

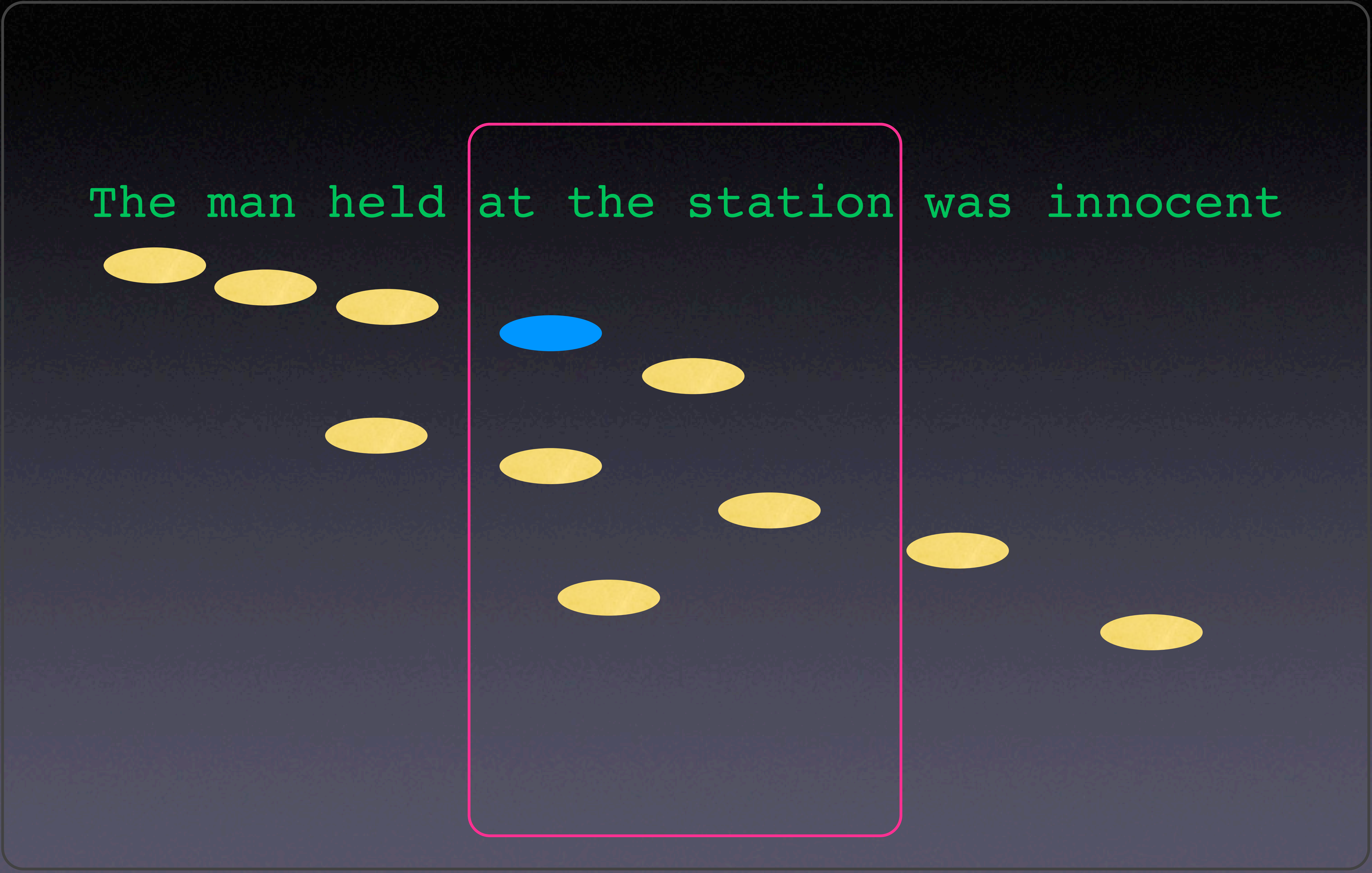
Self-paced reading, moving window:

`The man held at the station was innocent`

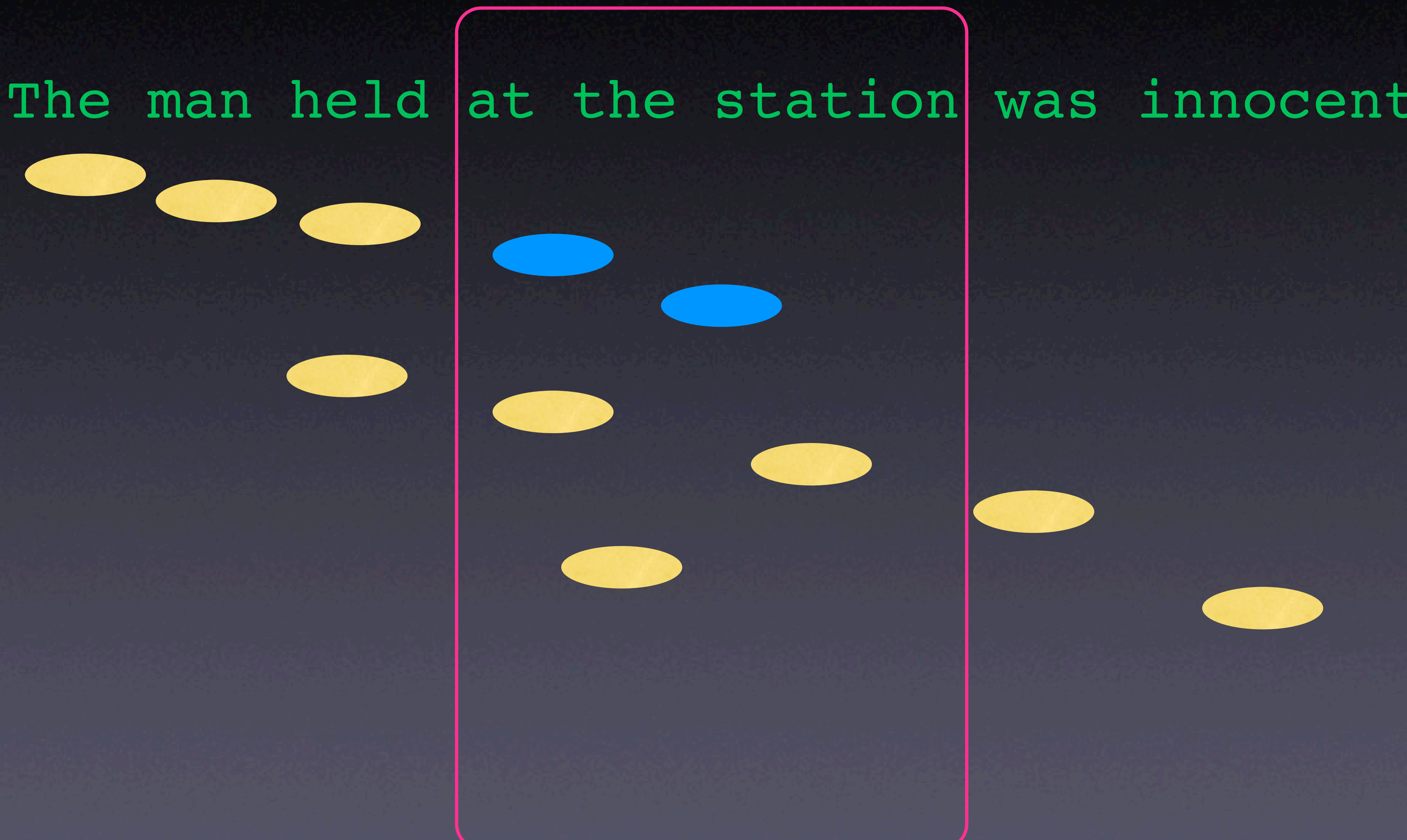
Eye-tracking: Difference Measures



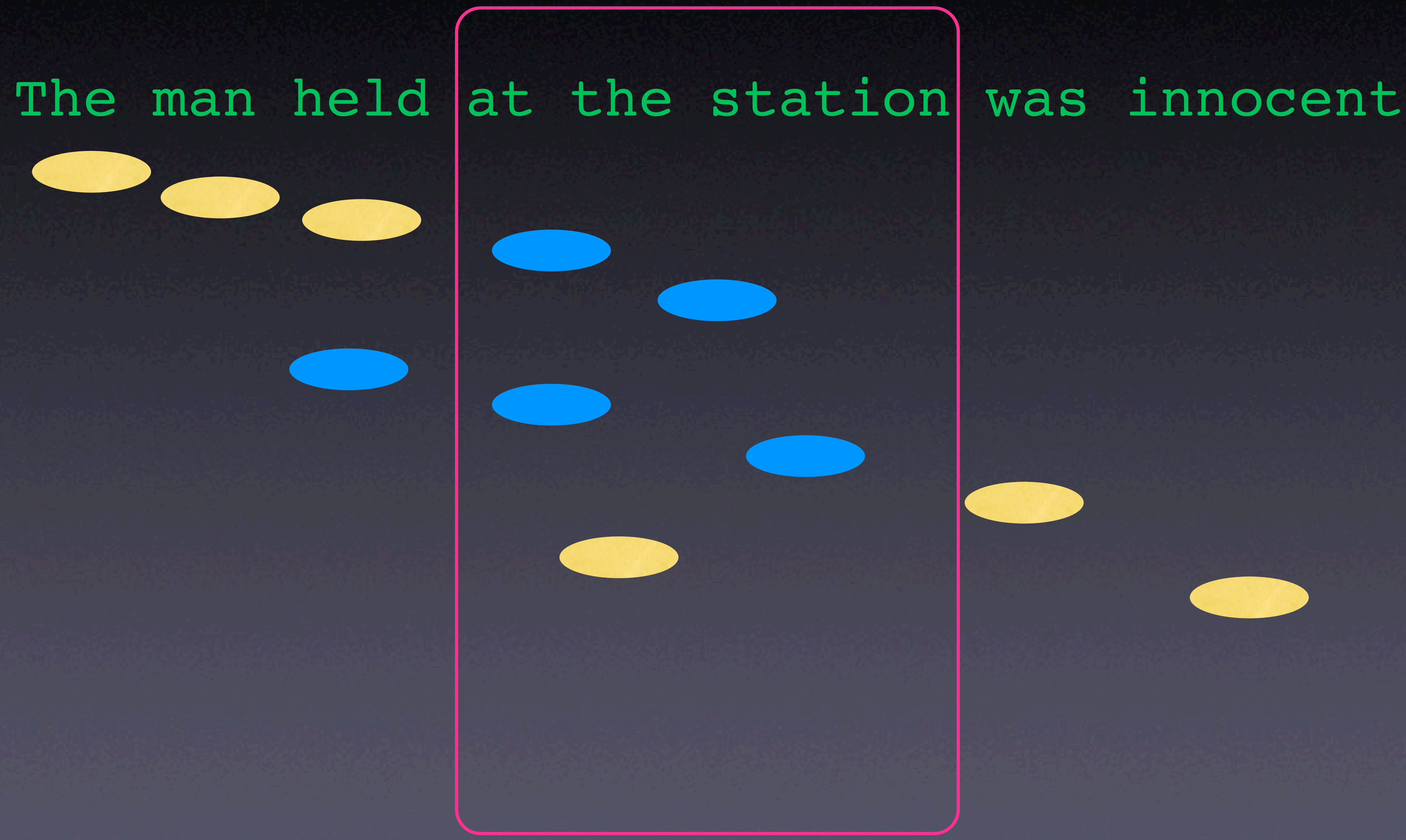
Eye-tracking: First Fixation



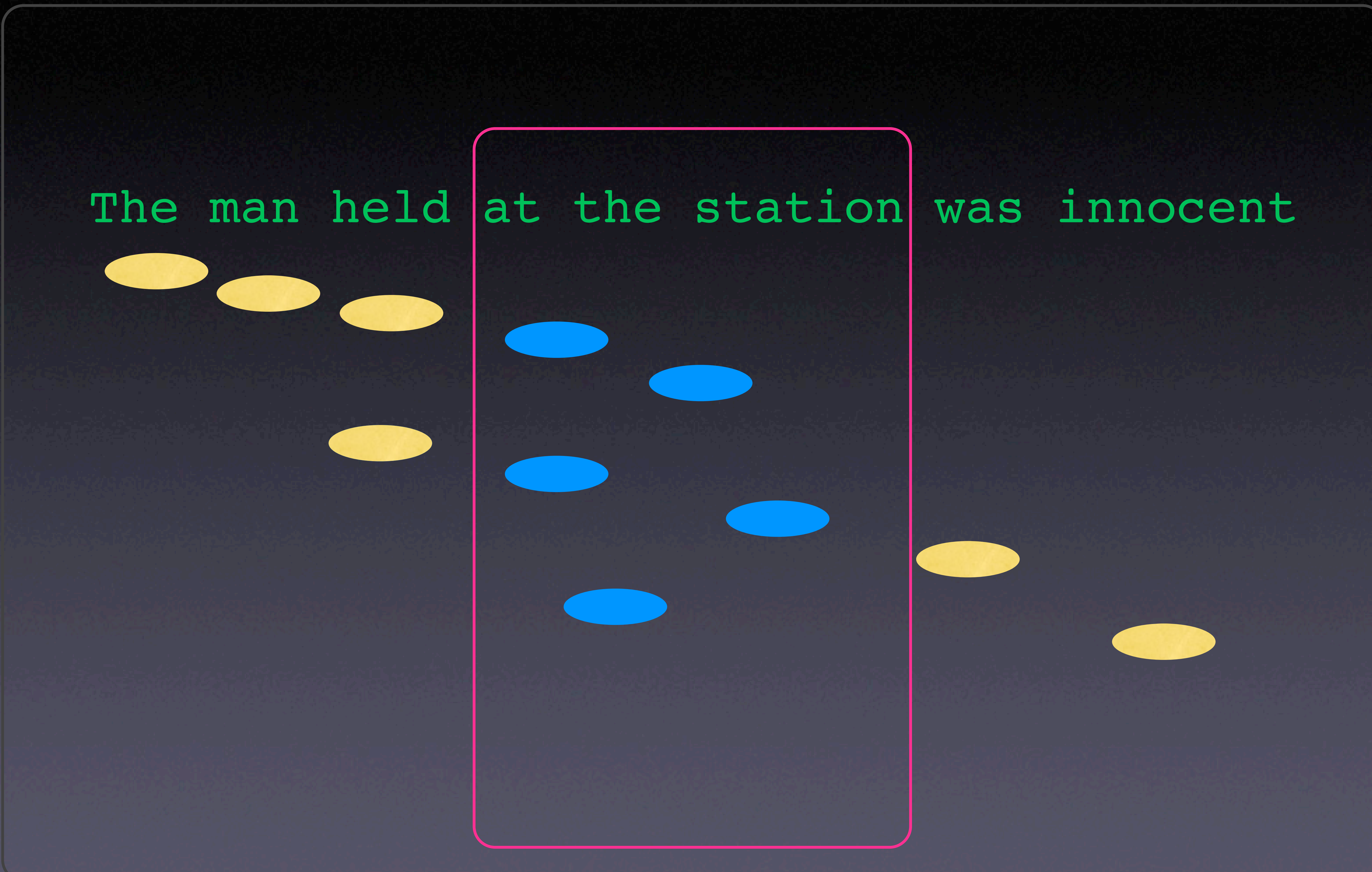
Eye-tracking: First Pass



Eye-tracking: Regression Path



Eye-tracking: Total time



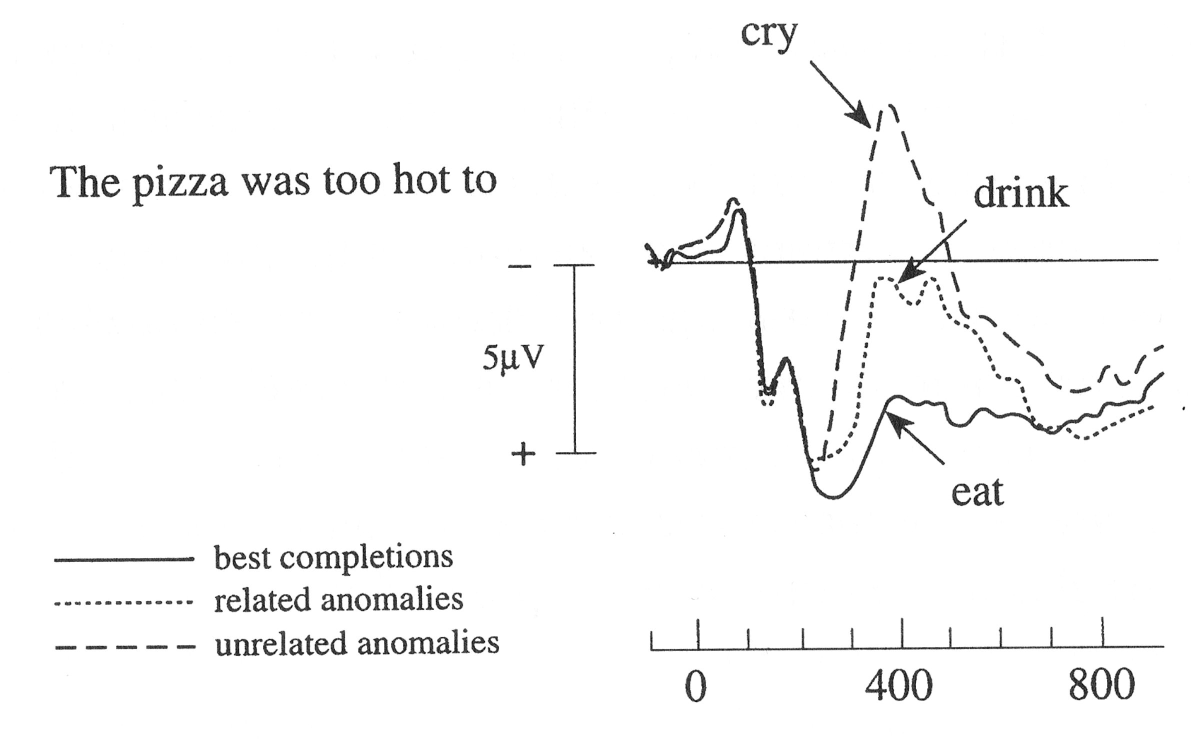
Neuroscientific Measures: ERPs

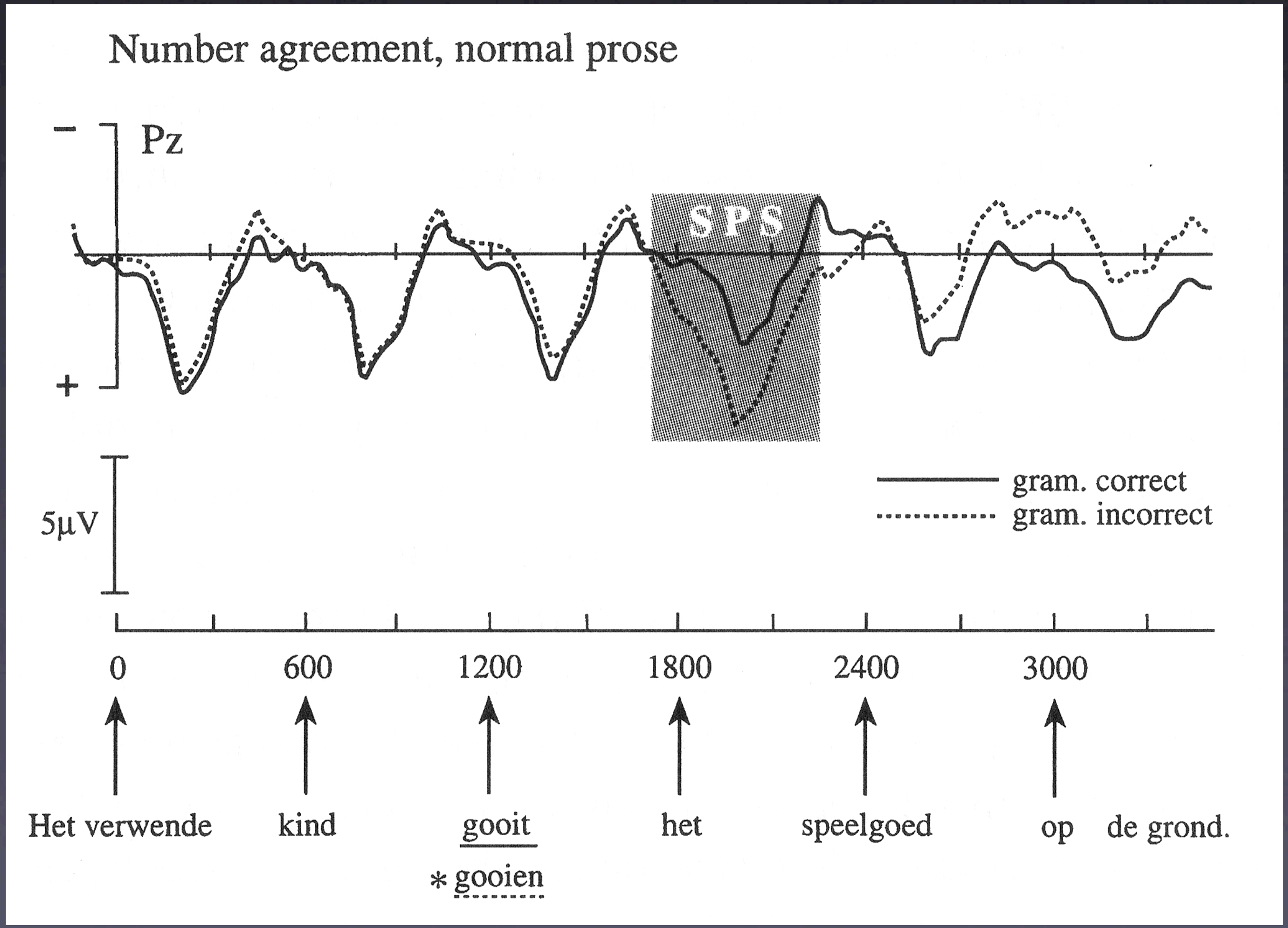
Syntactic and semantic processes are partially revealed by signature patterns in EEGs: Event-Related Potentials (ERPs)

Syntactic Anomaly: P600 or SPS

- `“The spoilt child throw(s) the toy on the ground”`

Semantic Anomaly: N400





Summary

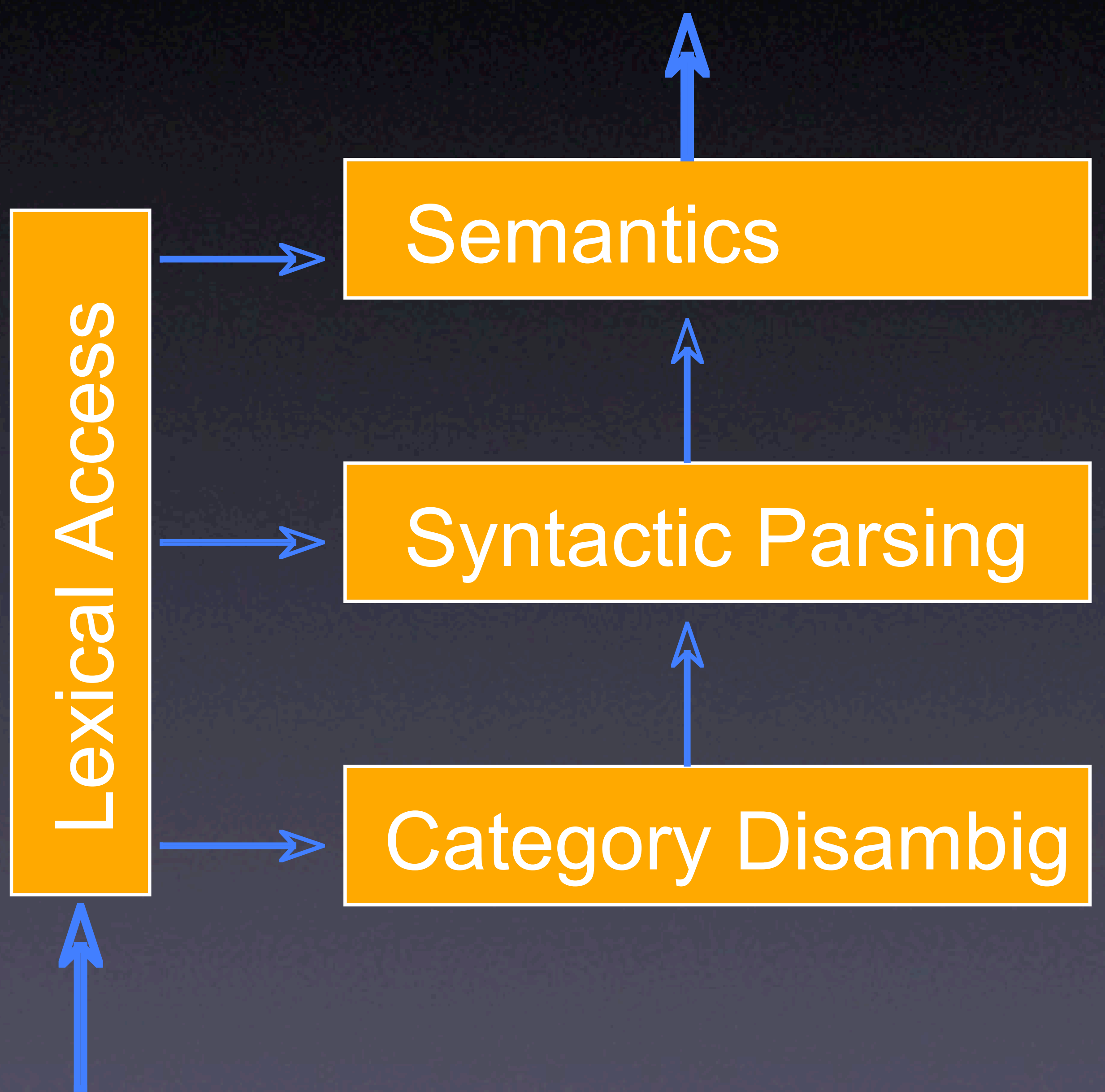
People construct an interpretation word-by-word

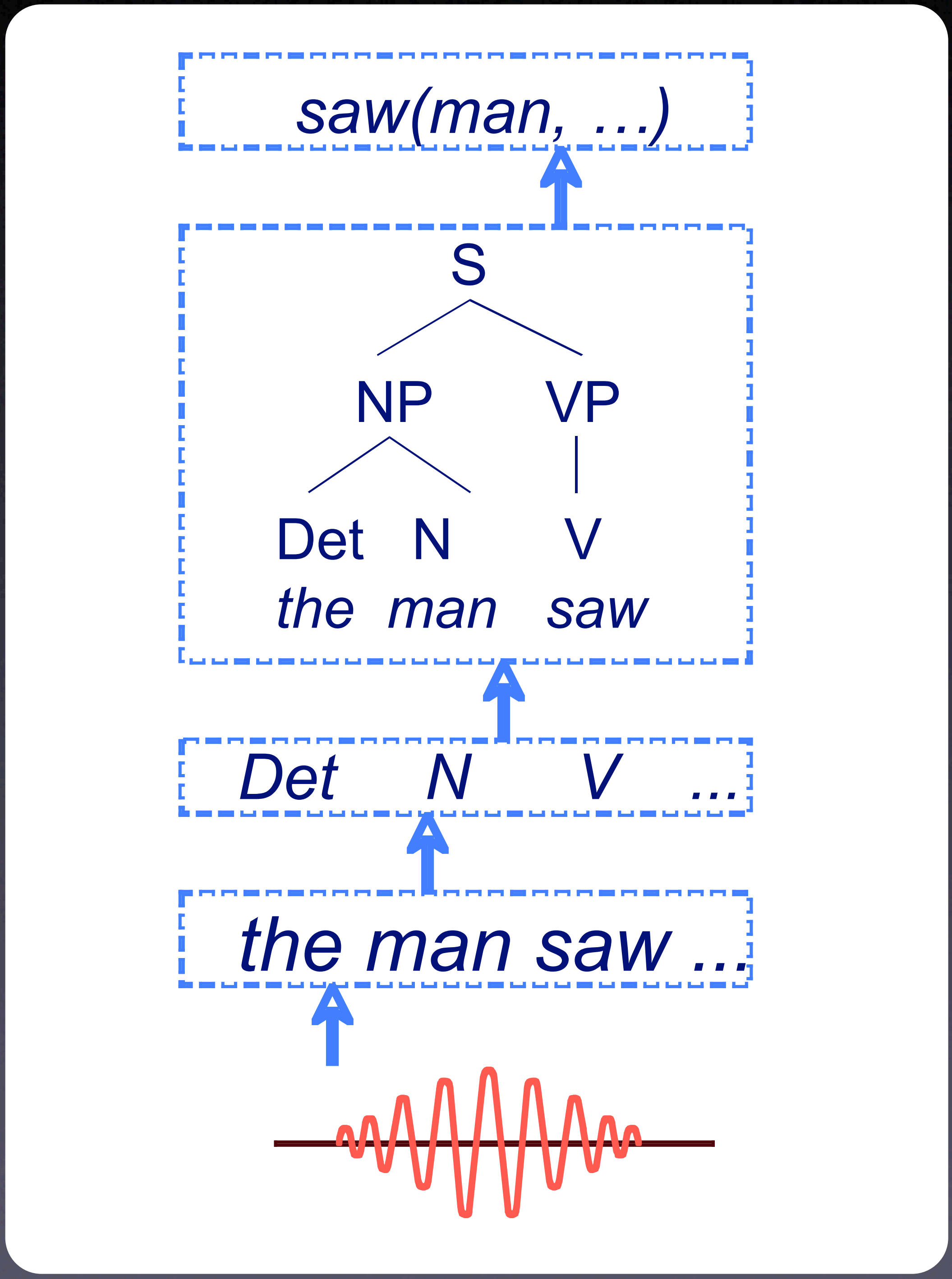
People must resolve ambiguity, and sometime reanalyse

Reading times and ERPs can tell us when this occurs

We can design experiments which exploit this to investigate the underlying processing architectures and mechanisms

A Modular Architecture





Theories of Sentence Processing

What mechanisms is used to construct interpretations:

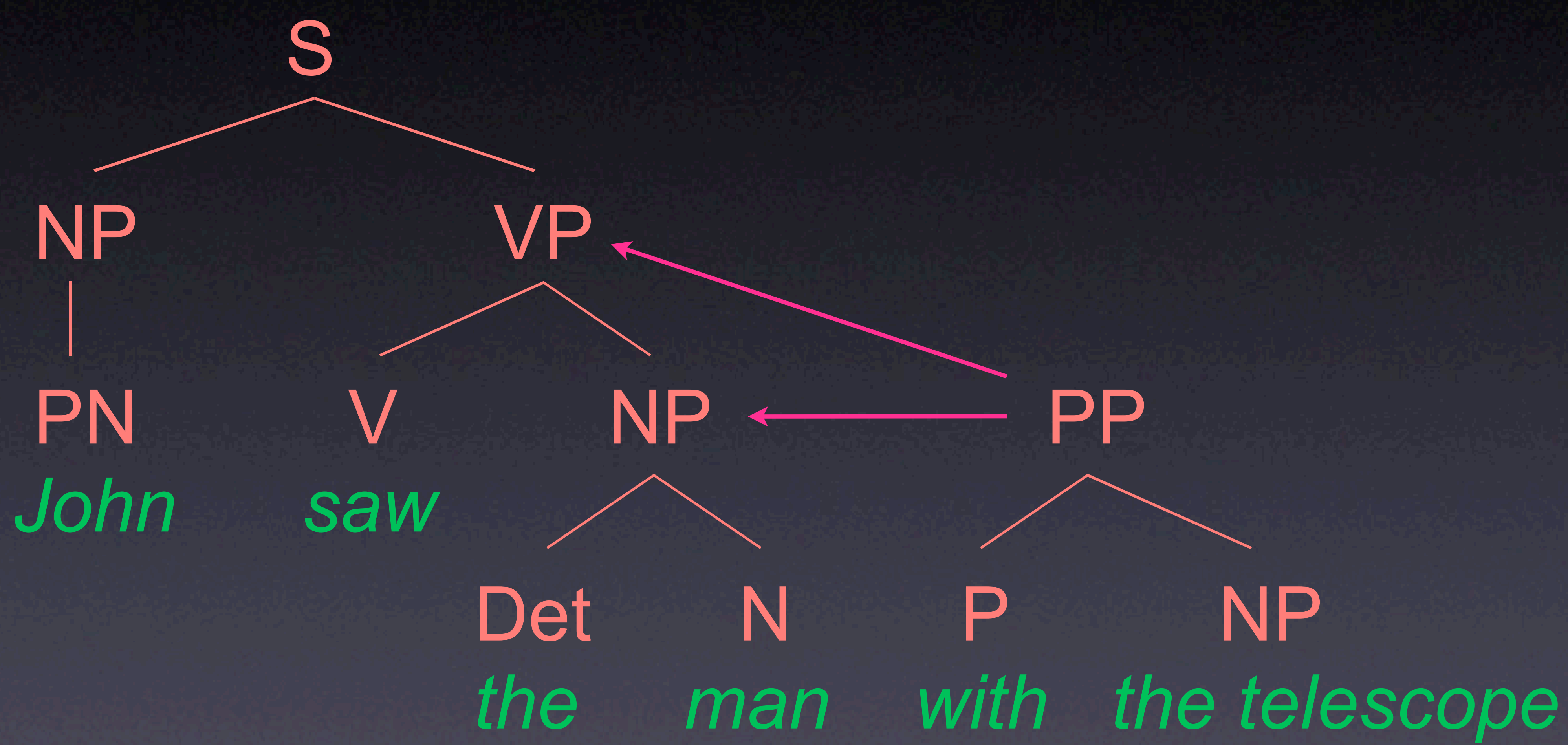
- Frazier: Serial parsing, with reanalysis
- Jurafsky: Parallel parsing, with reranking

What information is used to determine preferred structure:

- Frazier: General syntactic principles
- Jurafsky: Relative probabilities of alternative structures

The Garden Path Theory (Frazier)

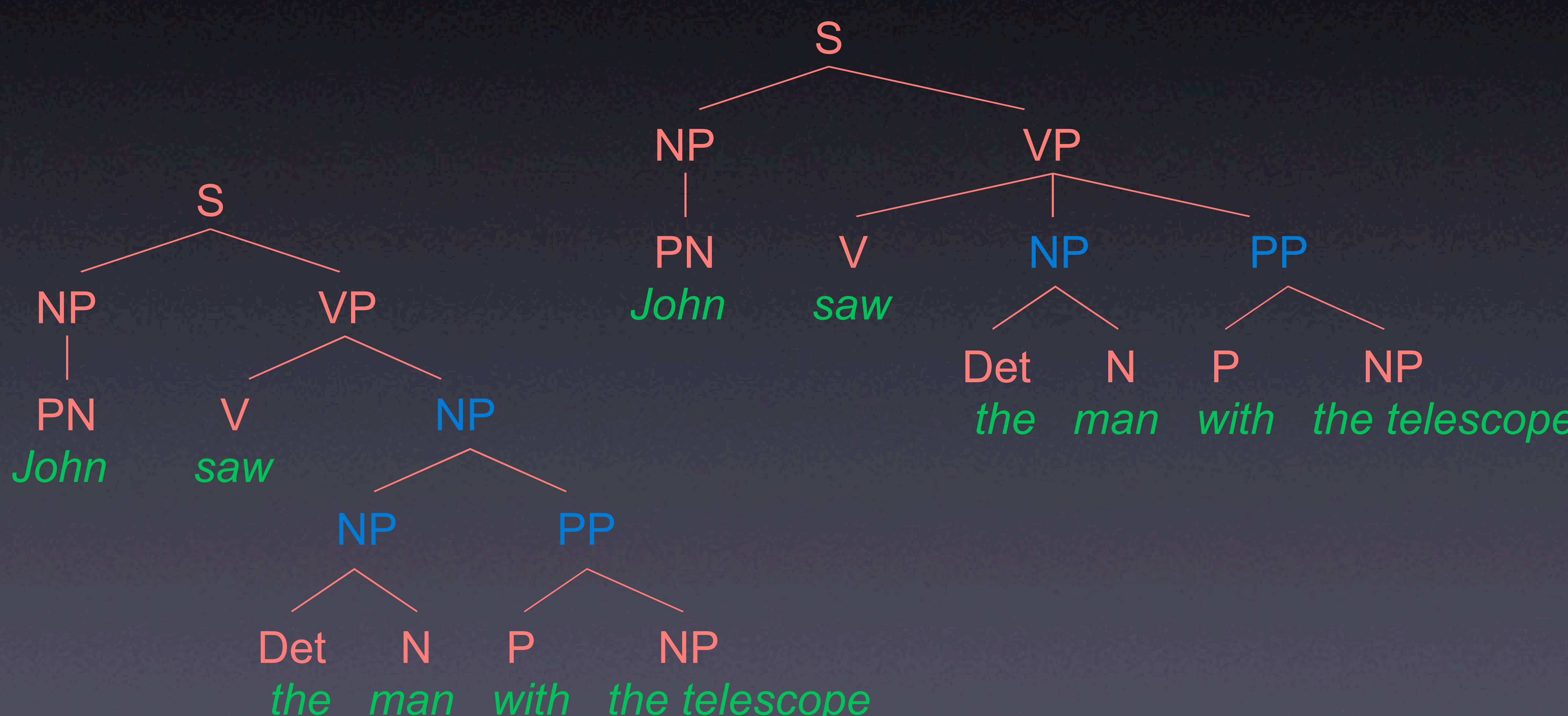
Prepositional Phase Attachment:



Which attachment do people initially prefer?

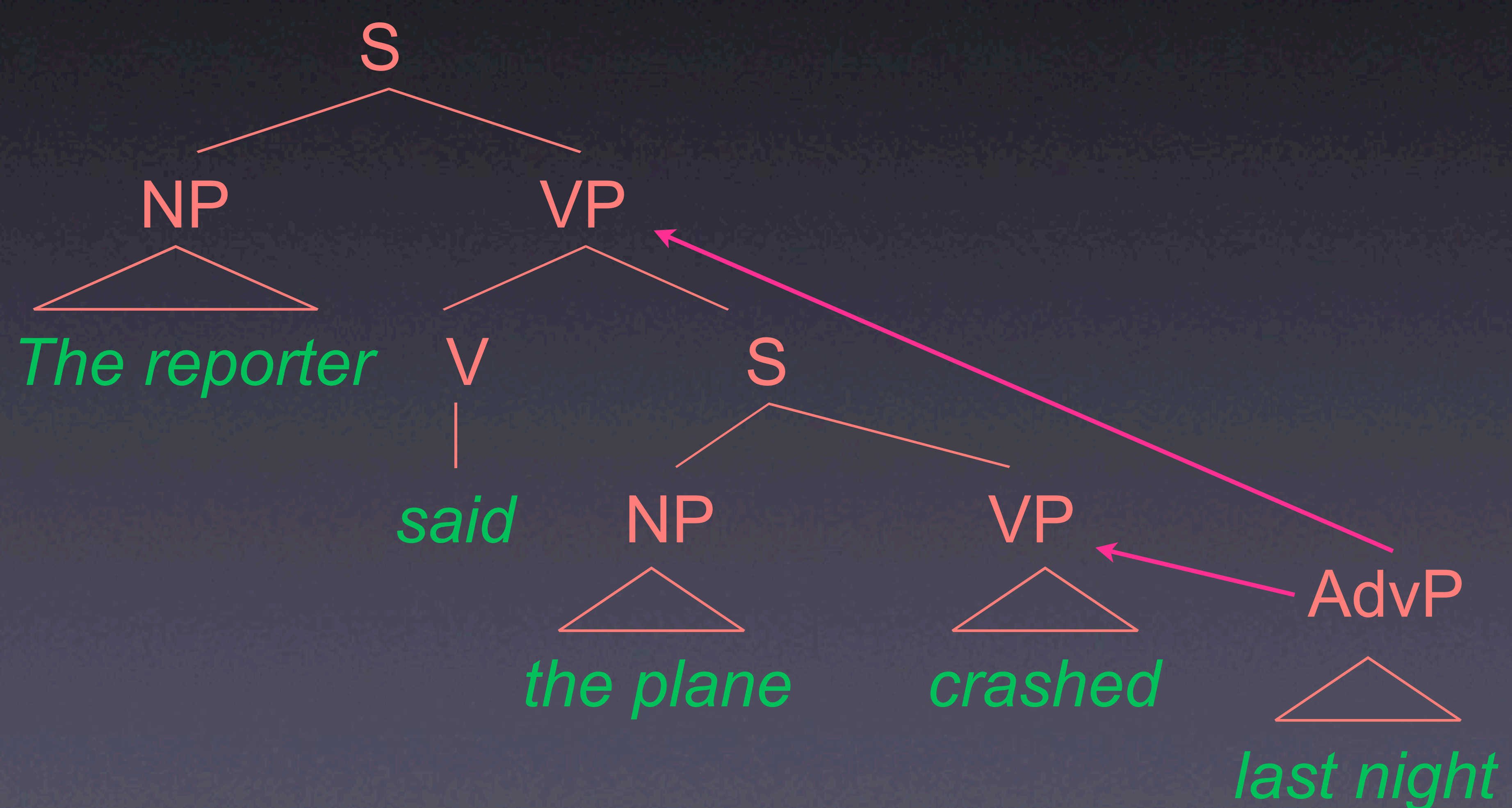
First Strategy: Minimal Attachment

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



Second Strategy: Late Closure

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



Summary of Frazier

Parsing preferences are guided by general principles:

- Serial structure building
- Reanalyse based on syntactic conflict
- Reanalyse based on low plausibility

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

Probabilistic Theories of Processing

Task of comprehension: recover the correct interpretation

Goal: Determine the most likely analysis for a given input:

$$\arg\max_i P(s_i) \text{ for all } s_i \in S$$

P hides a multitude of sins:

- P corresponds to the degree of belief in a particular interpretation
- Influenced by recent utterances, experience, non-linguistic context

Implementation

Interpretation of probabilities

- Likelihood of structure occurring, P can be determined by frequencies in corpora or human completions

Estimation of probabilities

- Infinite structural possibilities = sparse data
- Associate probabilities with grammar (finite): e.g. PCFGs

What mechanisms are required:

- Incremental structure building and estimation of probabilities
- Comparison of probabilities entails parallelism

Probabilistic Grammars

Context-free rules annotated with probabilities

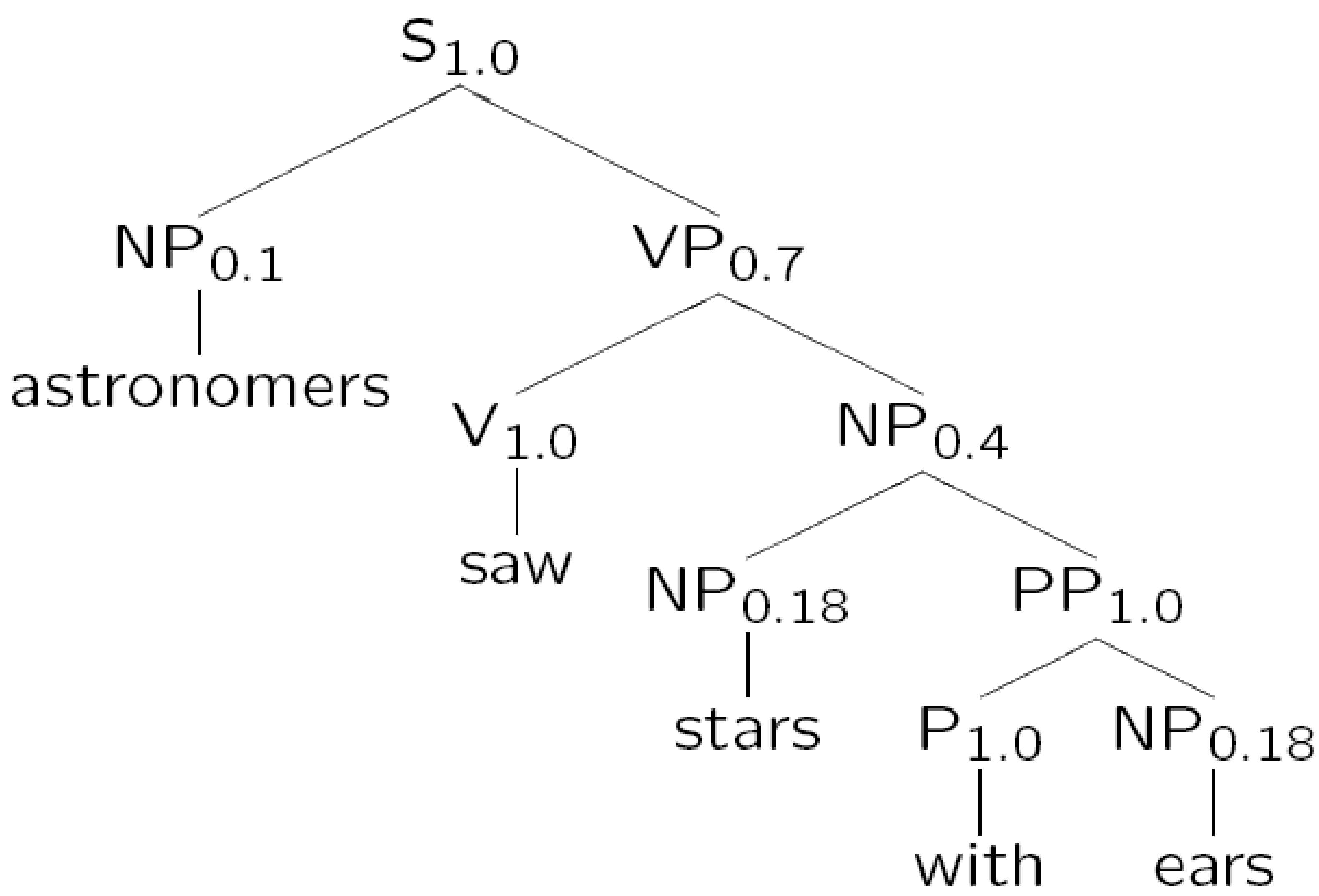
- Probabilities of all rules with the same LHS sum to one;
- Probability of a parse is the product of the probabilities of all rules applied in the parse.

Example (Manning and Schütze 1999)

S → NP VP	1.0	NP → NP PP	0.4
PP → P NP	1.0	NP → astronomers	0.1
VP → VP NP	0.7	NP → ears	0.18
VP → VP NP	0.3	NP → saw	0.04
P → with	1.0	NP → stars	0.18
V → saw	1.0	NP → telescopes	0.1

Parse Ranking

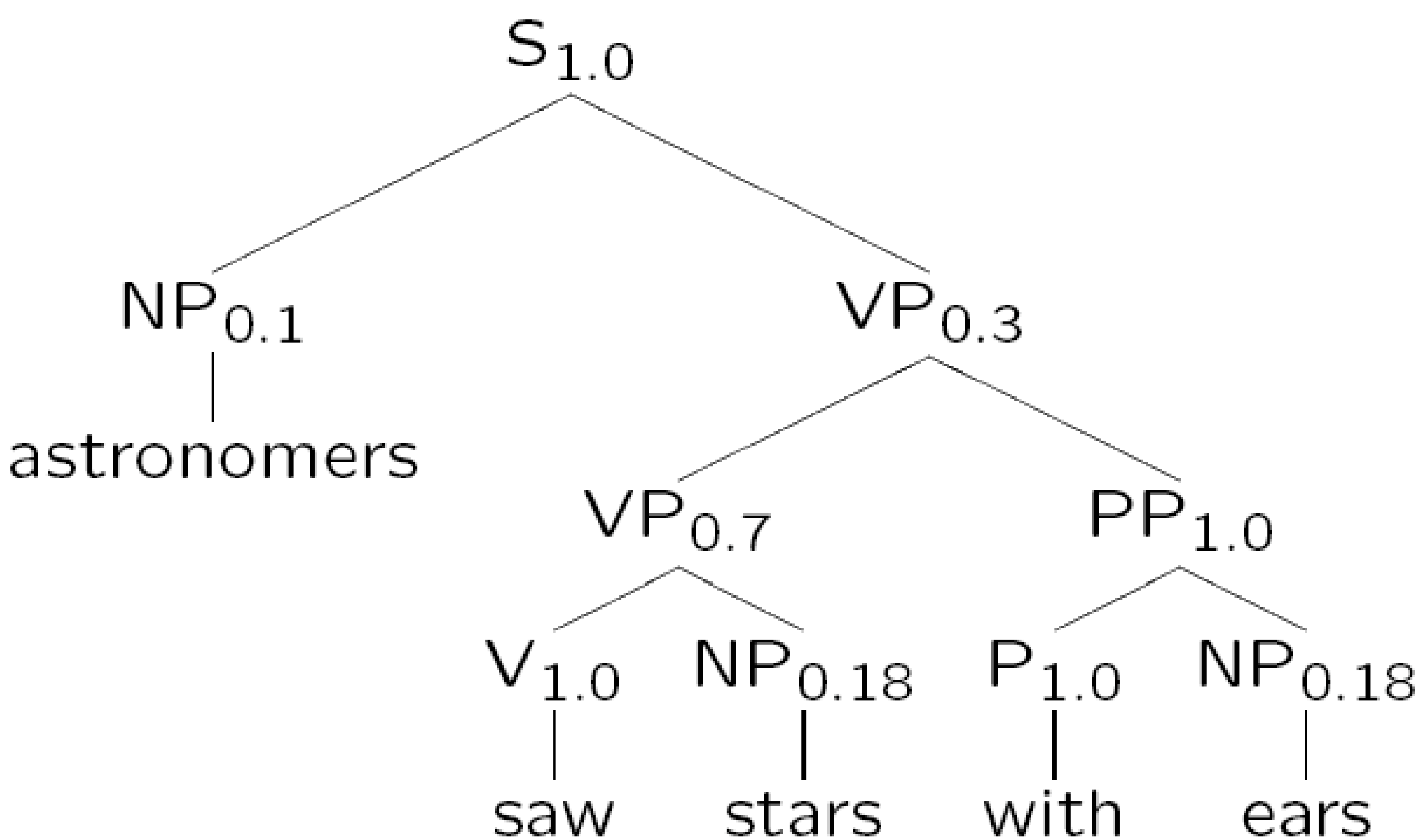
t_1:



$$P(t_1) = 1.0 \times 0.1 \times 0.7 \times 1.0 \times 0.4 \times 0.18 \times 1.0 \times 1.0 \times 0.18 = 0.0009072$$

Parse Ranking

t_2:



$$P(t_1) = 1.0 \times 0.1 \times 0.3 \times 0.7 \times 1.0 \times 0.18 \times 1.0 \times 1.0 \times 0.18 = 0.0006804$$

Jurafsky (1996)

Jurafsky's (1996) approach:

- psycholinguistic model of lexical and syntactic access and disambiguation;
- exploits concepts from computational linguistics: probabilistic CFGs, Bayesian modeling frame probabilities;

Overview of issues:

- data to be modeled: frame preferences, garden paths;
- architecture: serial, parallel, limited parallel;
- probabilistic CFGs, frame probabilities;
- examples for frame preferences, garden paths;
- comparison with other models; problems and issues.

Frame Preferences

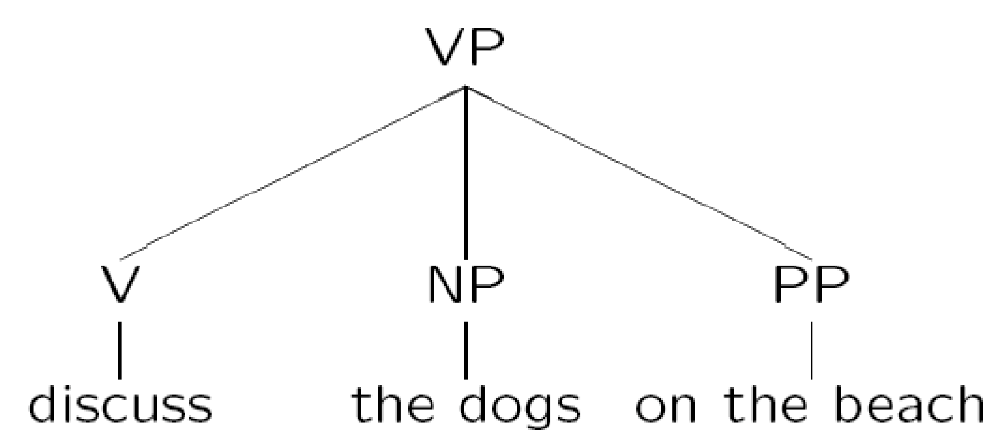
The women discussed the dogs on the beach.

- t1. The women discussed them (the dogs) while on the beach. (10%)
- t2. The women discussed the dogs which were on the beach. (90%)

$p(\text{discuss}, \langle \text{NP PP} \rangle) = 0.24$

VP → V NP XP 0.15

t_1:

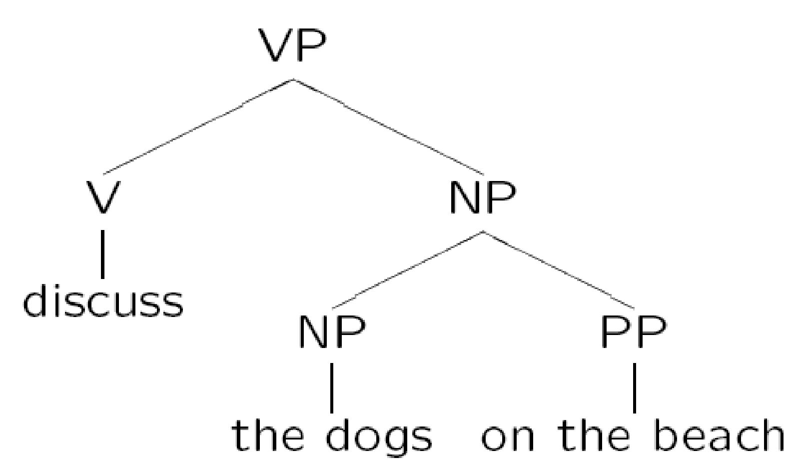


$p(t_1) = 0.15 \times 0.24 = 0.036$ (dispreferred)

$p(\text{discuss}, \langle \text{NP} \rangle) = 0.76$

VP → V NP 0.39
NP → NP XP 0.14

t_2:



$p(t_2) = 0.76 \times 0.39 \times 0.14 = 0.041$ (preferred)

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 goal]] last week”

“The athlete realised [[his shoes] were across the room]”

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]”

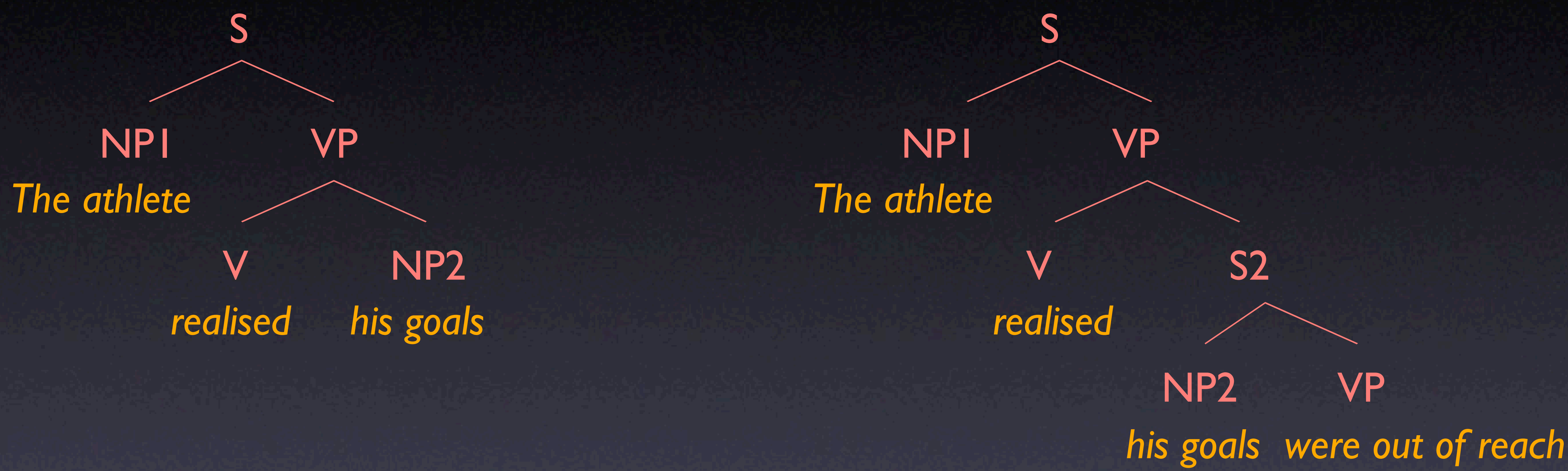
Relative/Complement Clause Ambiguity:

“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]]”

A Problem for Likelihood?

NP/S Complement Ambiguity: *The athlete realised his goals ...*



Evidence for object attachment: (Pickering, Traxler & Crocker 2000)

- Despite S-comp bias of verb, NP is attached as D-object
- Ideal likelihood model and Jurafsky predict the opposite
- realised is initially tagged at S-comp, but the simpler DO analysis is then given higher probability, when NP is found

Traditional Approaches

Experimental research:

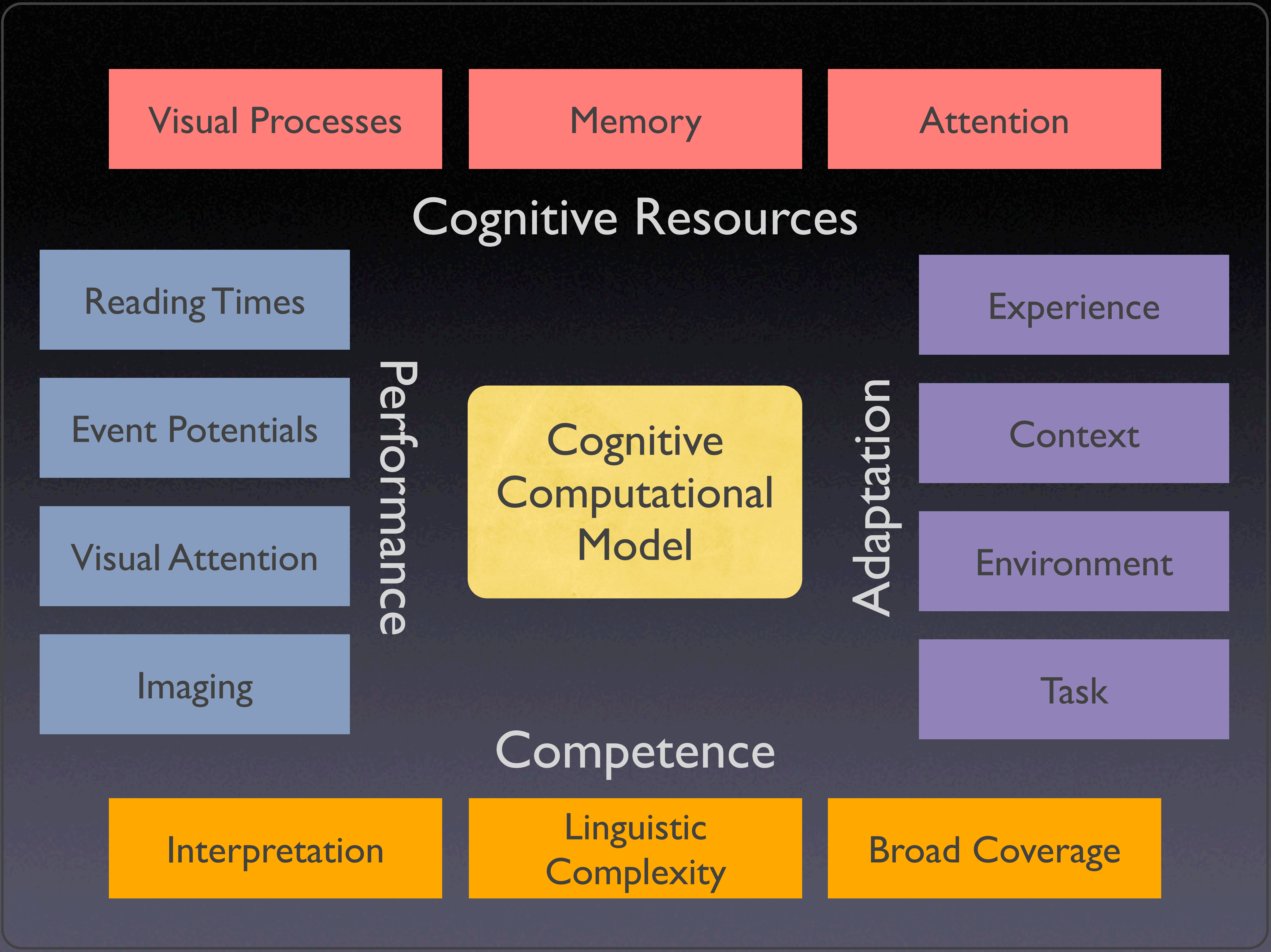
- Reading: self-paced and eye-tracking paradigms
- Measure: reading times = processing complexity

Computational models:

- Emphasis on linguistic processing (lex, syn, sem)
- Theories strive to explain processing complexity

✴ Emphasis on the weaknesses of human comprehension

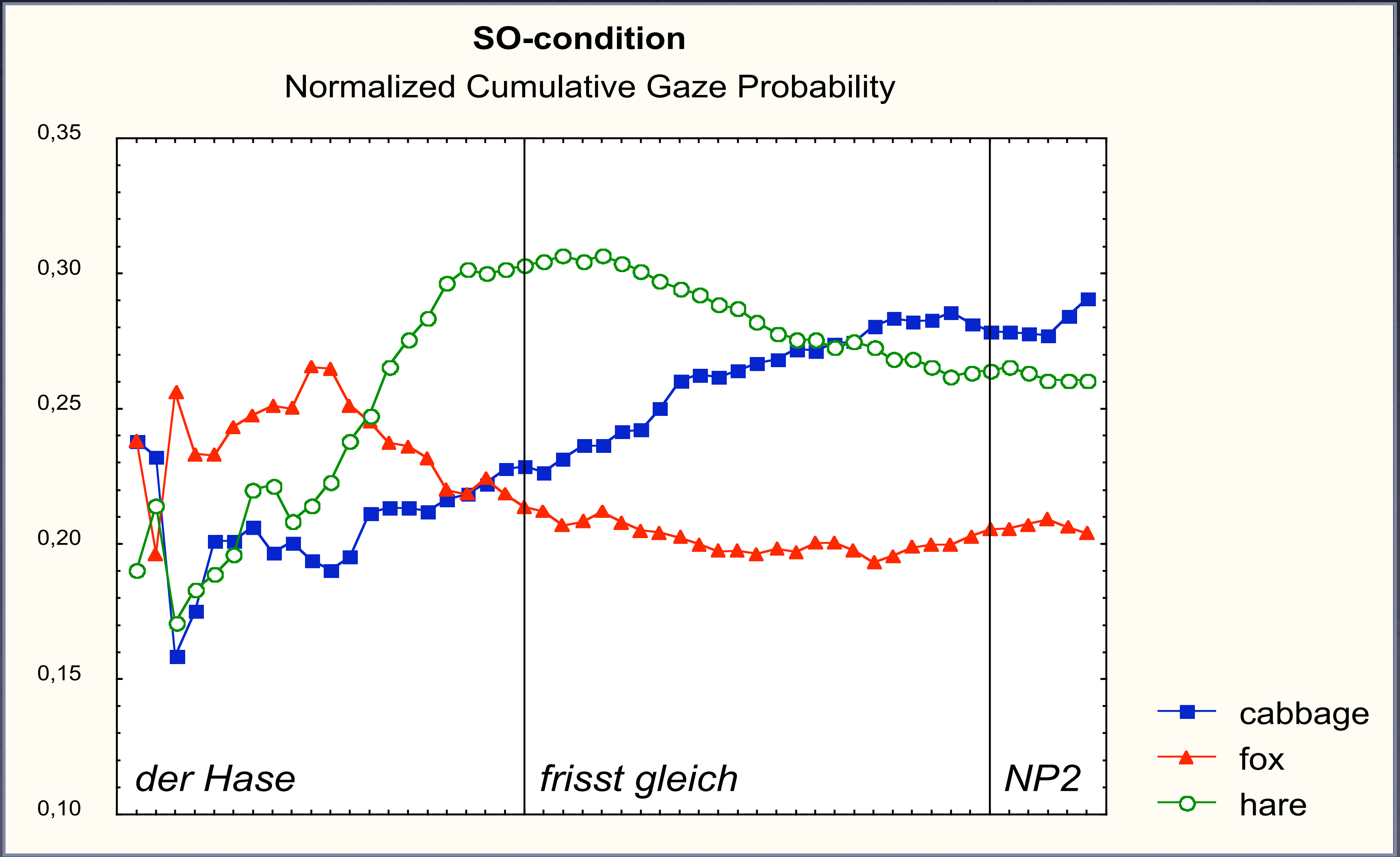
✴ Failure to situate the human language processor


Visual Processes
Memory
Attention
Cognitive Resources
Reading Times
Event Potentials
Visual Attention
Imaging
Performance
Cognitive Computational Model
Adaptation
Experience
Context
Environment
Task
Competence
Interpretation
Linguistic Complexity
Broad Coverage

Spoken comprehension in visual scenes

Monitor gaze in the scene as people hear a spoken utterance

- Listeners fixate objects which are mentioned (180ms)
- Anticipatory eye-movements reflect interpretation



"The rabbit will eat the fox"

German Word Order

Subject Verb Object (SVO) and Object Verb Subject (OVS)

Case-marking reflects grammatical function:

- "*Der Hase*" (rabbit): Nominative/Subject
- "*Den Hasen*" (rabbit): Accusative/Object

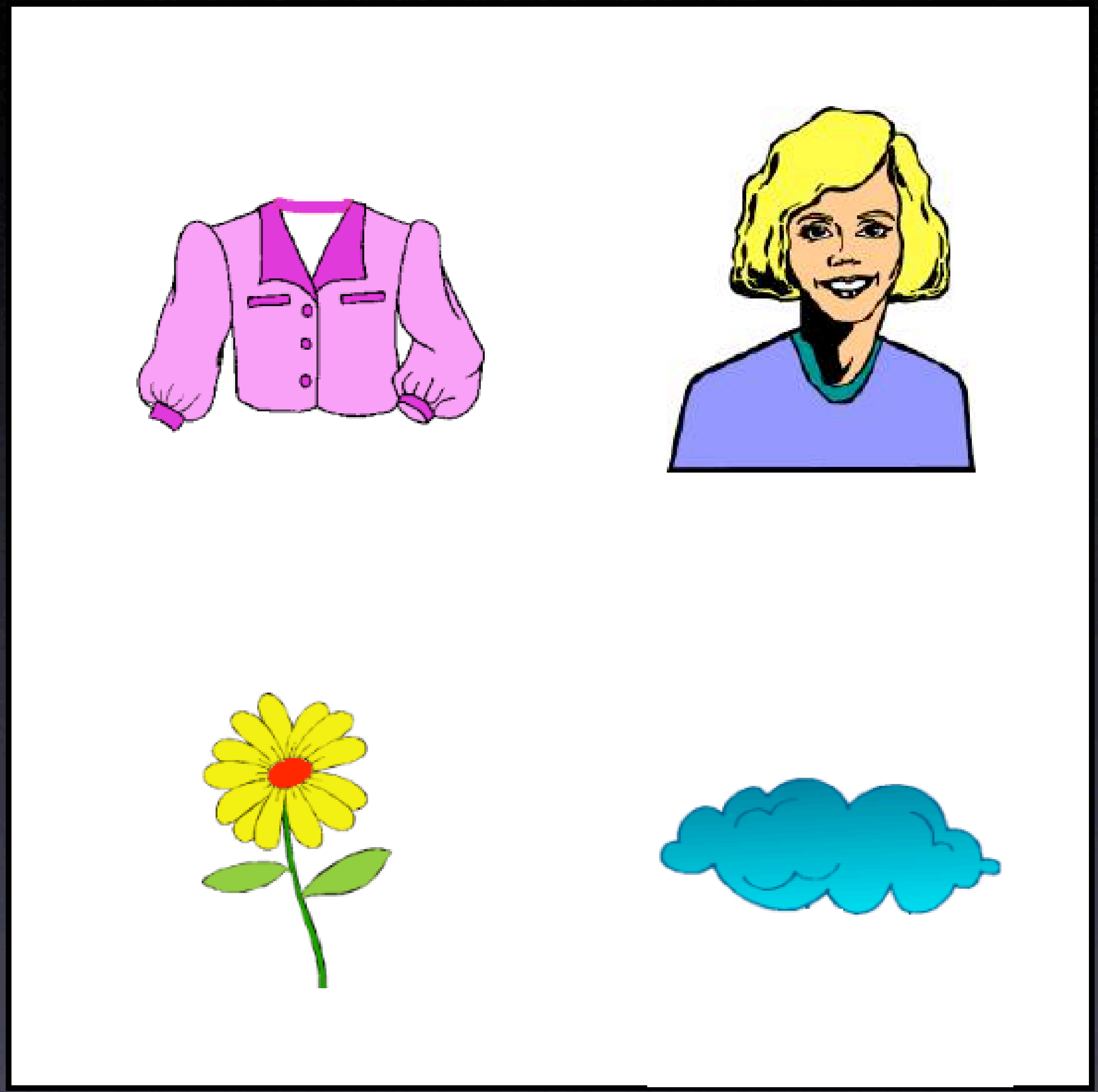
But some ambiguity in case-marking:

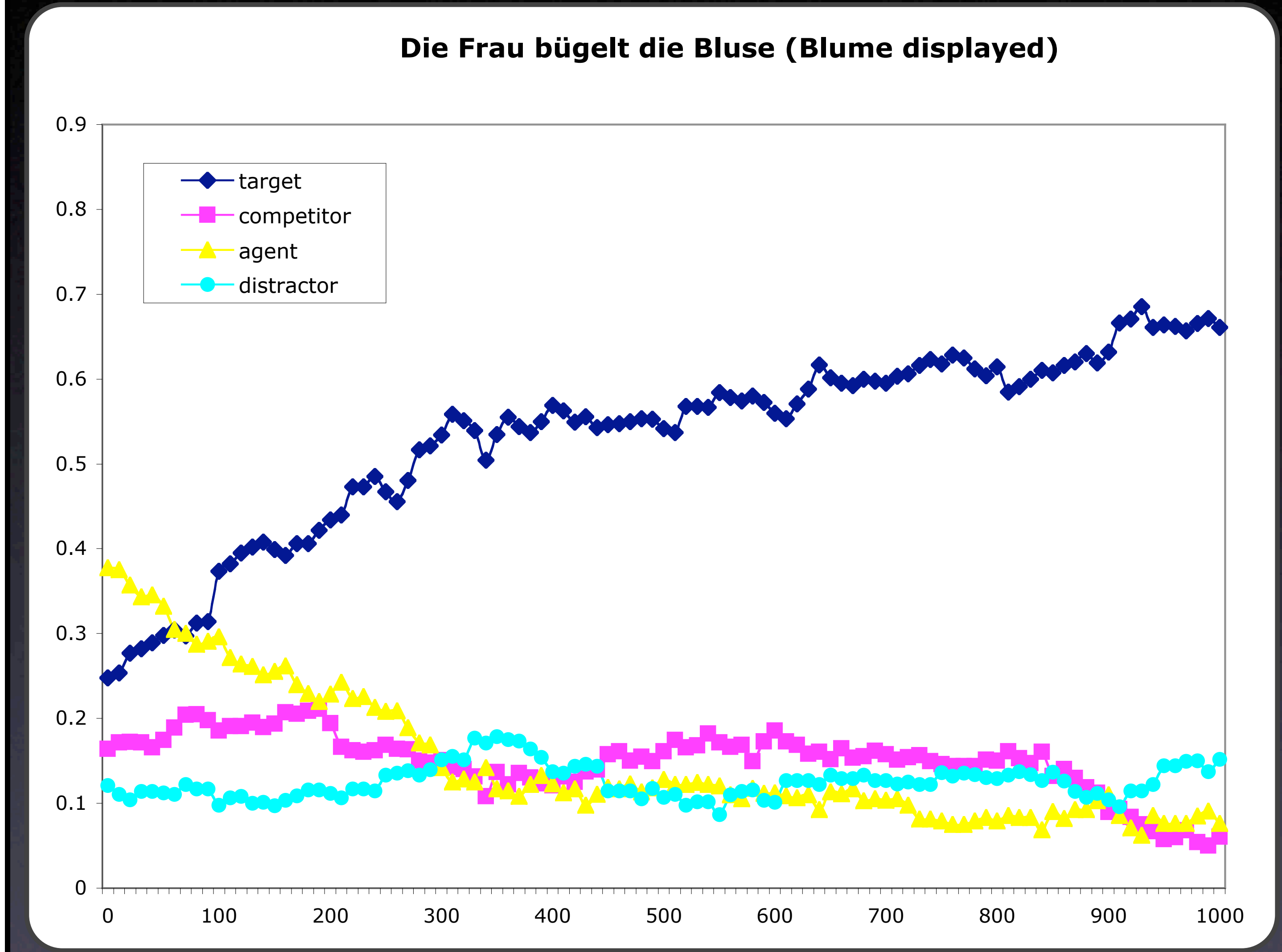
- "*Die Prinzessin*" (princess): Nominative or Accusative

Preferred SVO word order, OVS is marked:

- SVO: "*Die Prinzessin sah den Hasen*" easy
- OVS: "*Die Prinzessin sah der Hase*" difficult

Anticipation versus Lexical Access



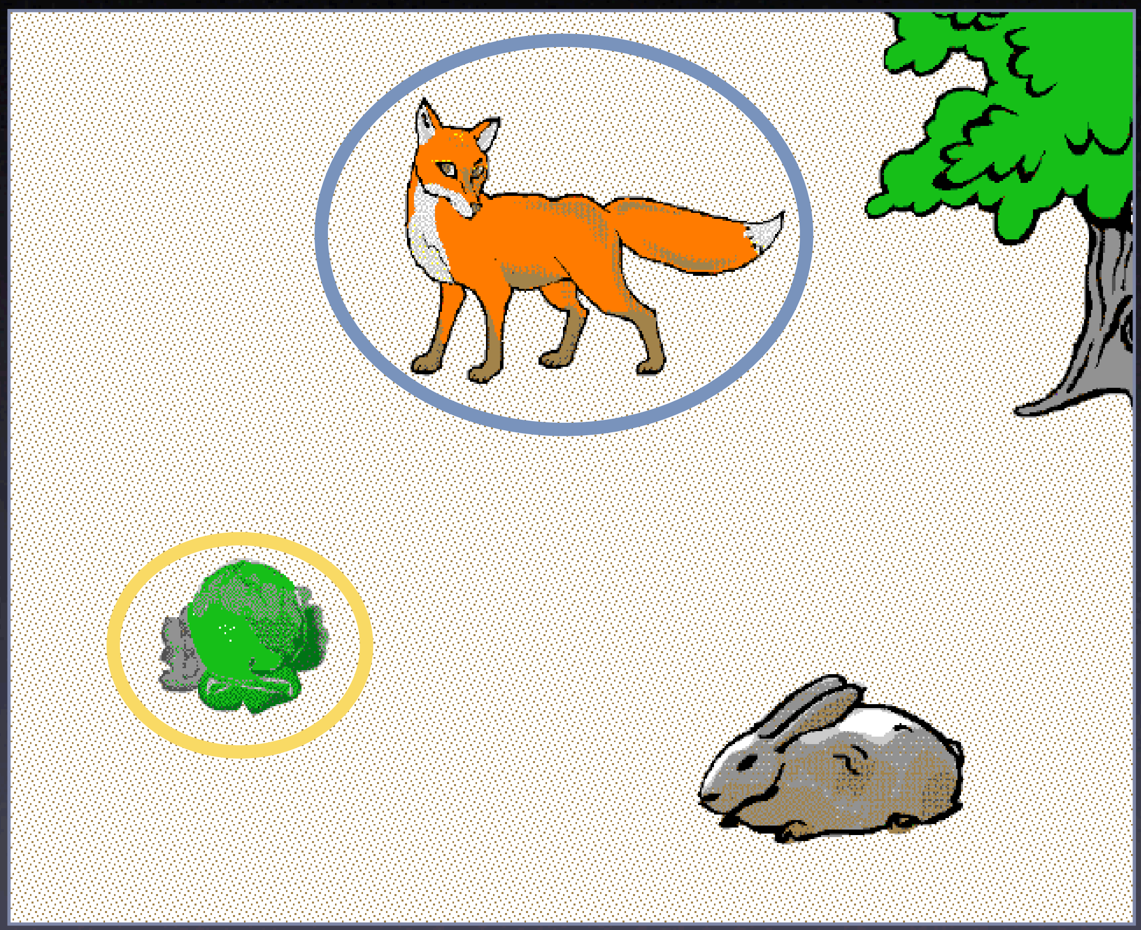


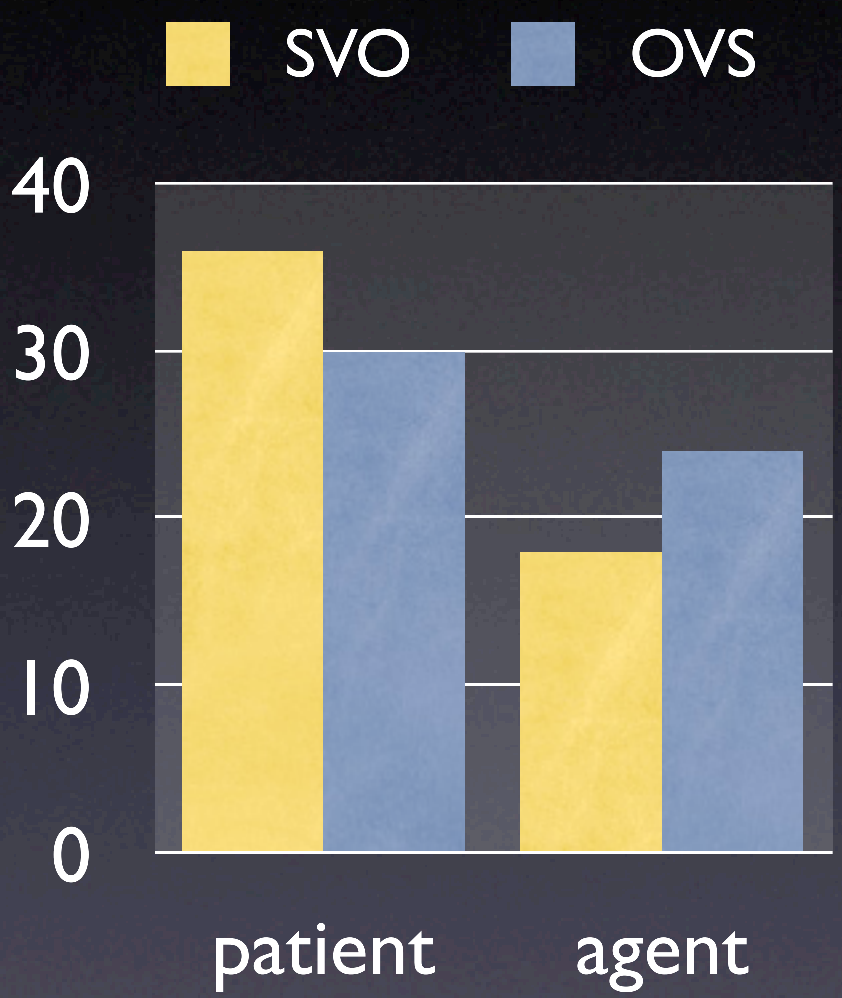
Non-restrictive: *"Die Frau **findet** die Bluse"*
The woman finds the blouse

Restrictive: *"Die Frau **bügelt** die Bluse"*
The woman irons the blouse

Anticipation in Visual Worlds

Anticipatory eye-movements in visual scenes





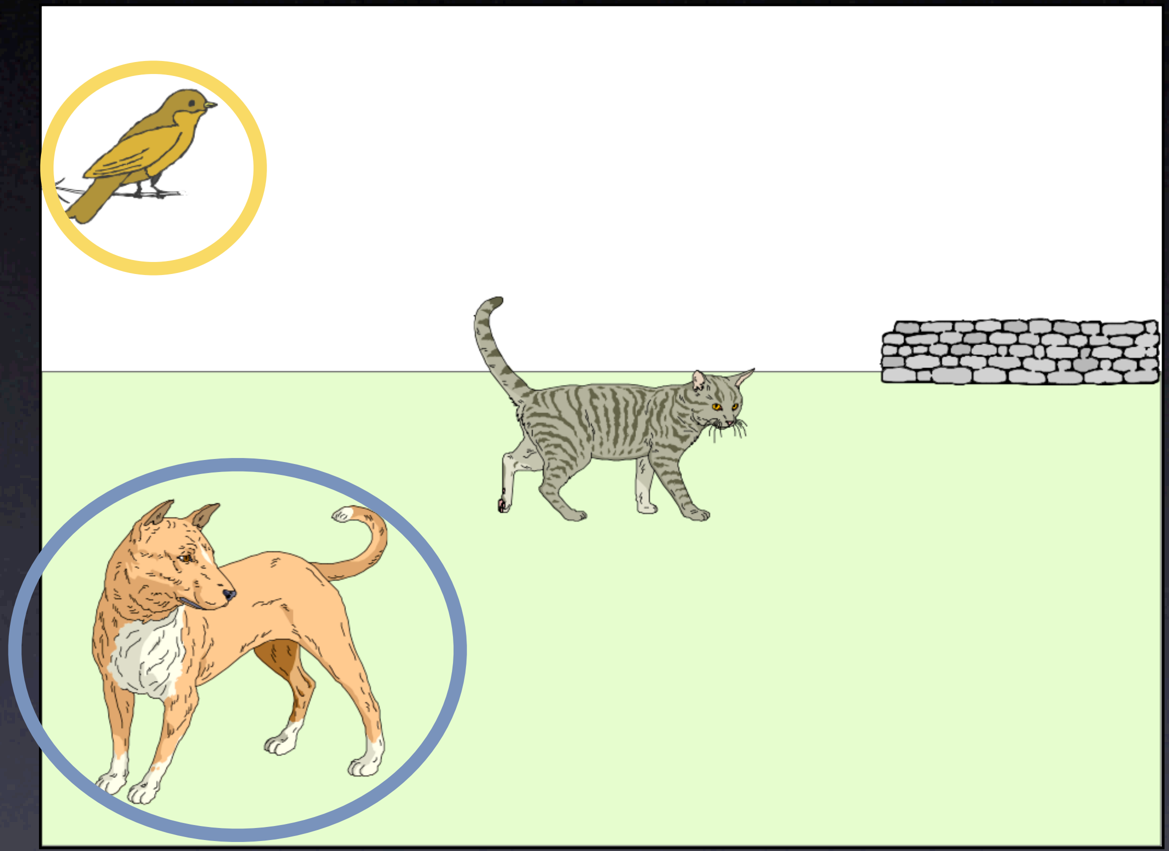
SVO: *Der Hase frisst <u>gleich</u>* ***den Kohl***

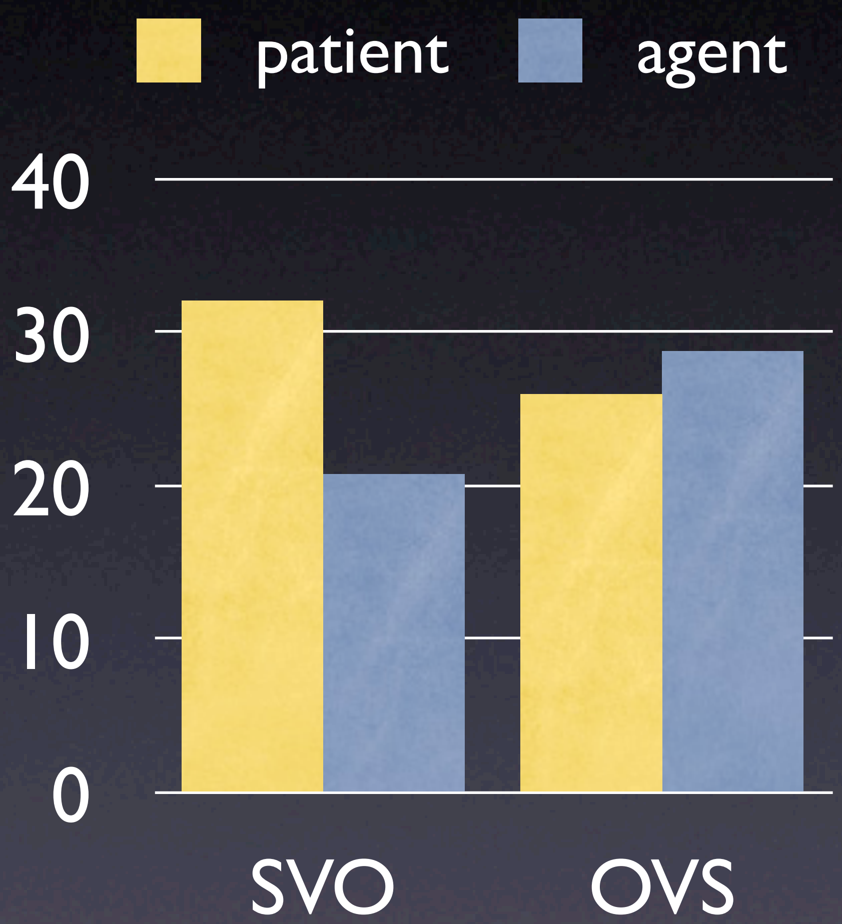
OVS: *Den Hasen frisst <u>gleich</u>* ***der Fuchs***

Kamide, Scheepers & Altmann, *JPR*, 2003

Intonation in spoken sentence comprehension

Can prosodic cues resolve local ambiguity?





SVO: *Die Katze jagt <u>womöglich</u> **den Vogel***

Low pitch accent on *cat*, high focal pitch accent on *chases*.

OVS: *Die Katze jagt <u>womöglich</u> **der Hund***

Focal high pitch accent on *cat*.

Weber, Grice & Crocker, *Cognition, (accepted)*

Utterance mediated visual attention

Speech Contingent Eye Movements

- Rapid saccades to mentioned objects (180ms)

Immediate Integration of Diverse Information Sources

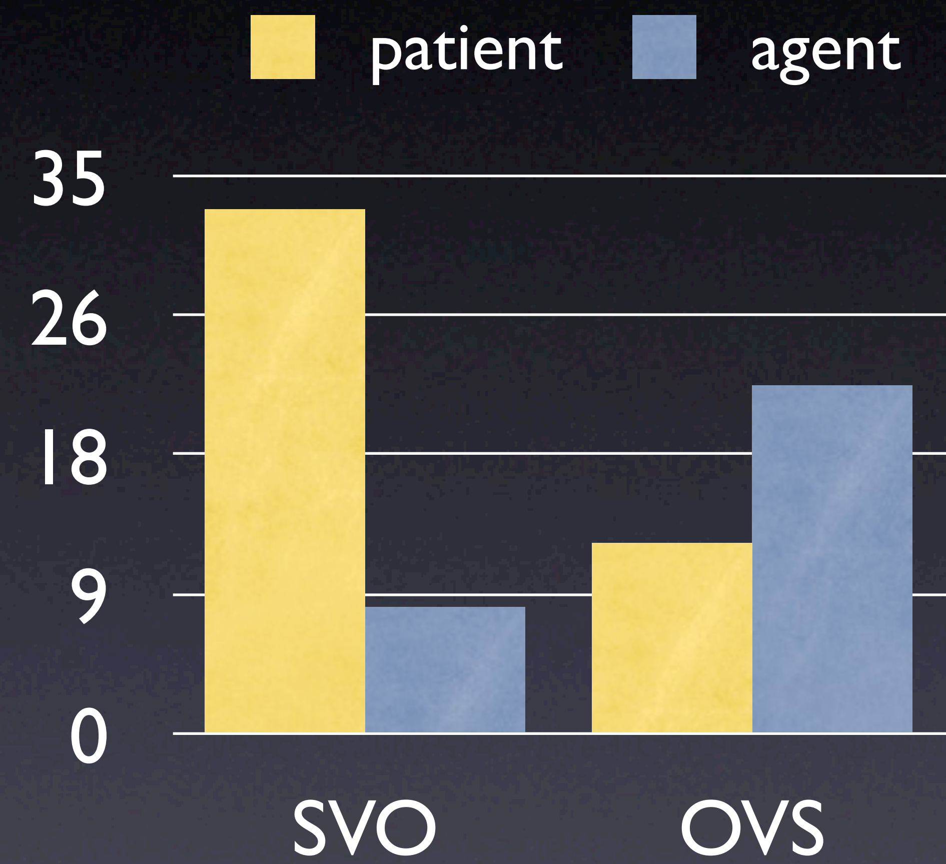
- Linguistic
- Prosodic
- World knowledge

Anticipatory Eye Movements

- Looks to expected entities reflects incremental understanding mechanisms

The Influence of the Scene

Do scene events influence interpretation?



SVO: *Die Prinzessin wäscht gleich* ***den Pirat***

OVS: *Die Prinzessin malt gleich* ***der Fechter***

Knoeferle, Crocker, Scheepers & Pickering, *Cognition*, 2005.

Scene influence on comprehension

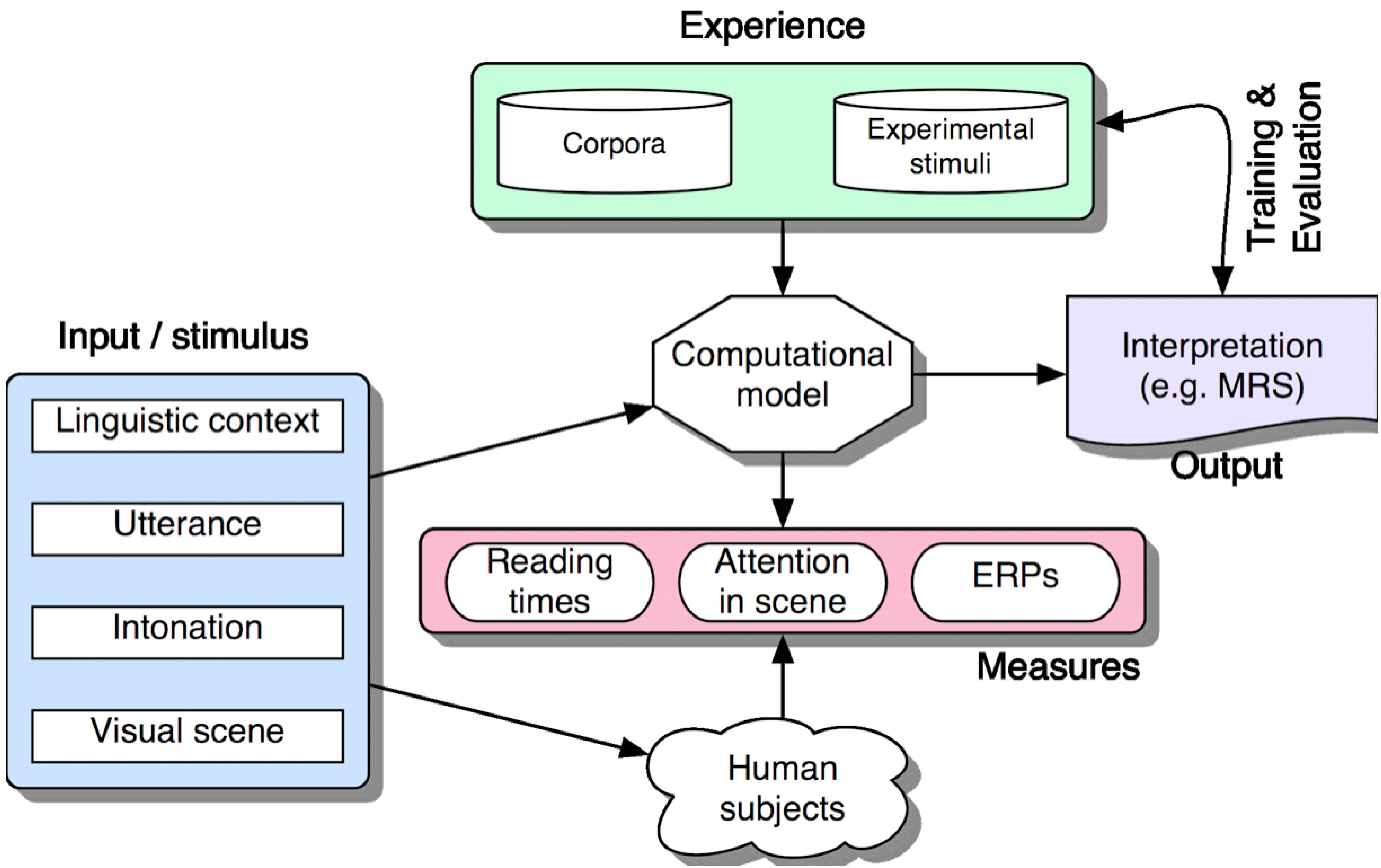
Verbs trigger looks to depicted events, just as nouns trigger looks to objects in scene

Event information influences comprehension

- Anticipation of role-fillers
- Triggers syntactic reanalysis from SVO to OVS

Evidence for early (scene induced) P600 confirms our interpretation of eye movements.

Adaptive models of language



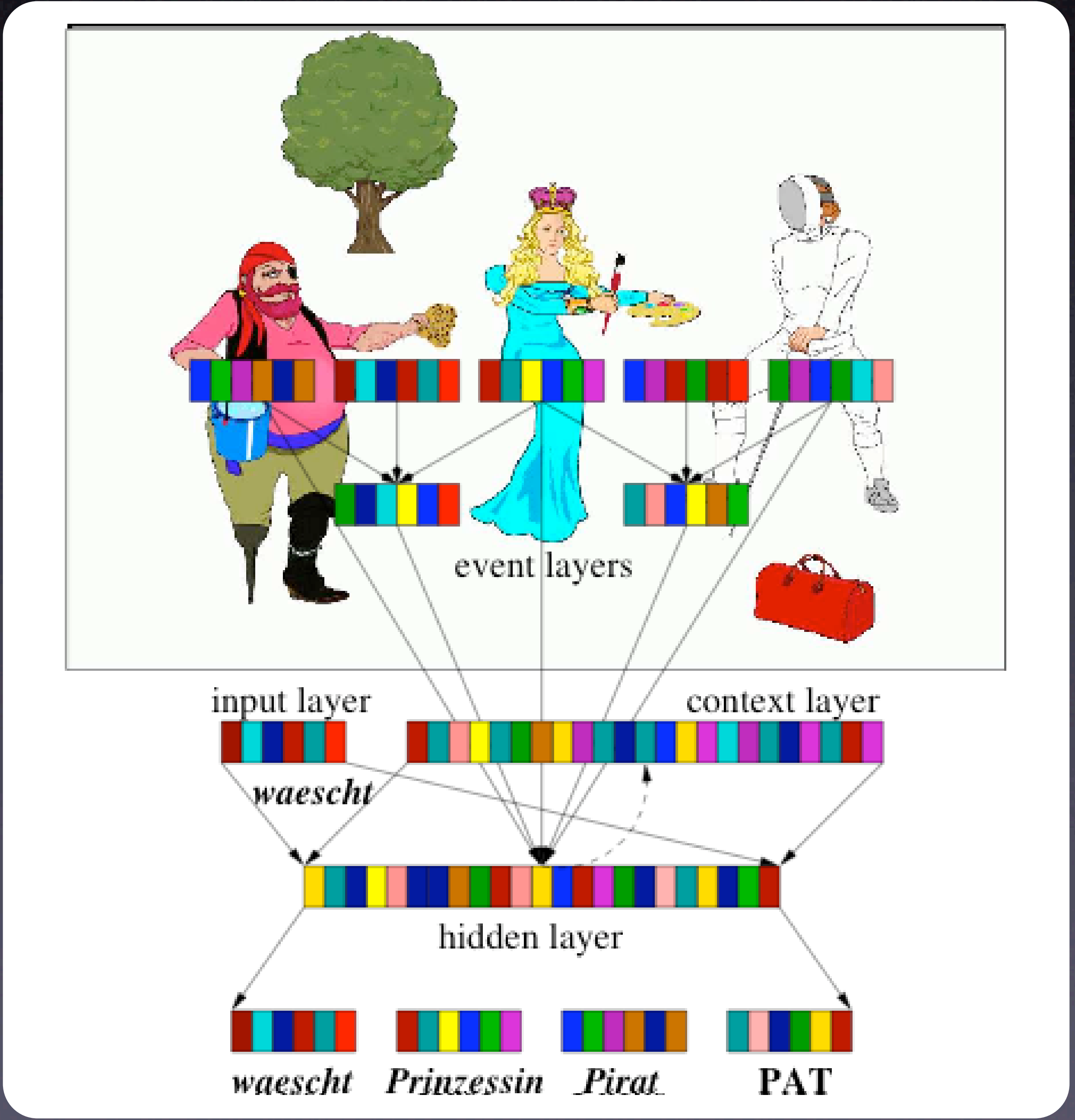
Network Architecture

Simple Recurrent Network trained with BPTT

- Enhanced with encoding of scene

Entities: (characters)

Events: (agent-action-patient)



Training

Training data

- Generated from experimental materials as templates
- All combinations of referents

Test data

- Actual experimental sentences are held out
- All test sentences have matching scene

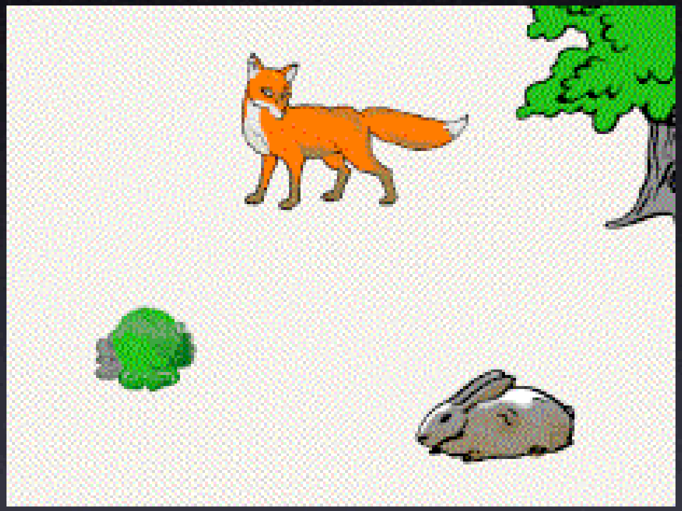
Training Regime

- Trained on final interpretation
- Scene provided as context 50% of time
  - reflect experience, and adaption to availability of scene

Simulation 1

One network to model four experiments simultaneously

Exp 1 & 2: Linguistic and Stored knowledge



- No event information available
- 32 verbs
- 48 nouns
- 96 extra nouns (for stereotypicality w/o overfitting)

Exp 3 & 4: Depicted actions

- Depicted events encoded
- 48 verbs
- 72 nouns

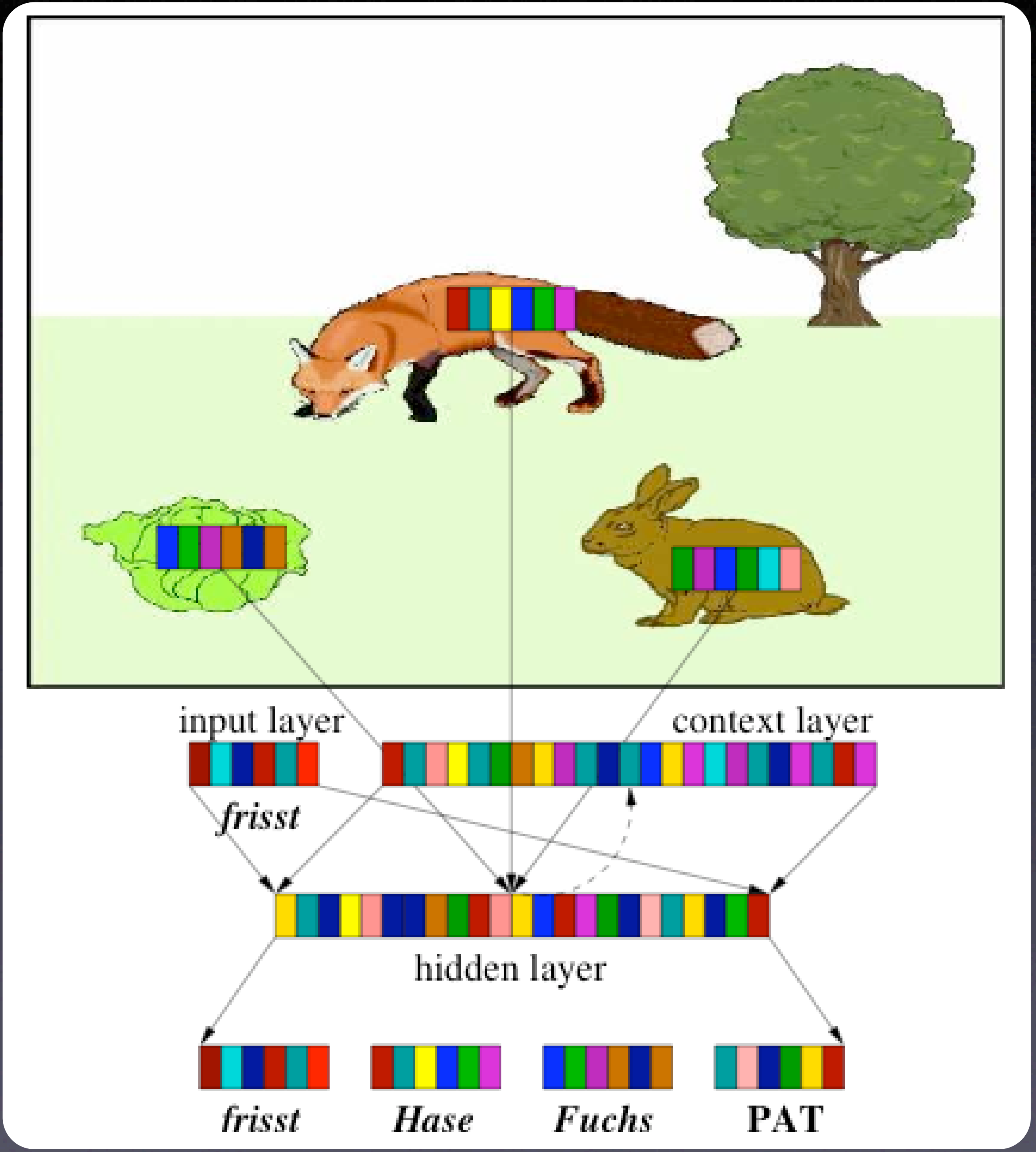
Den Hasen *frisst* gleich der Fuchs

⇒ Fox-like Agent

anticipated based on experience and stereotypicality

⇒ Fox Agent

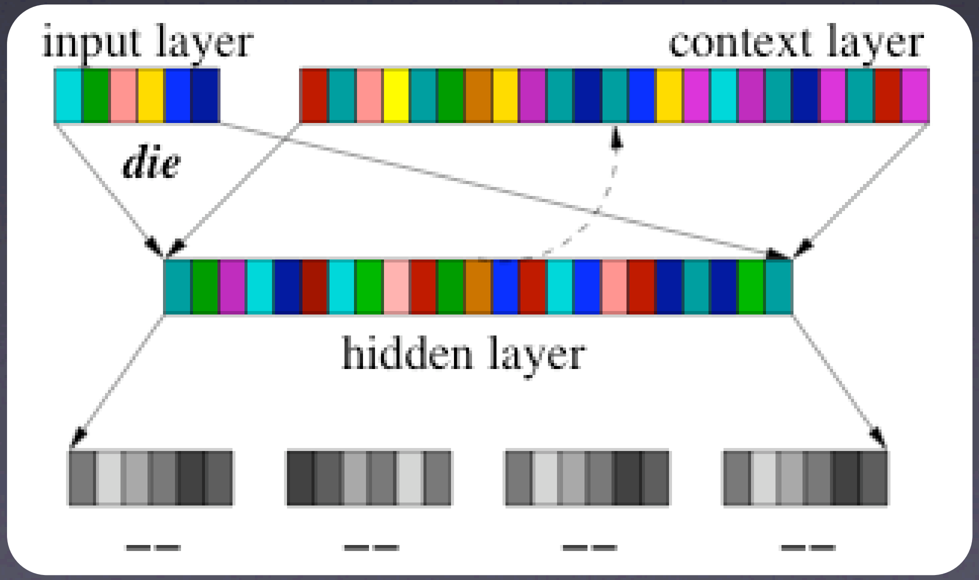
anticipated based on experience and scene

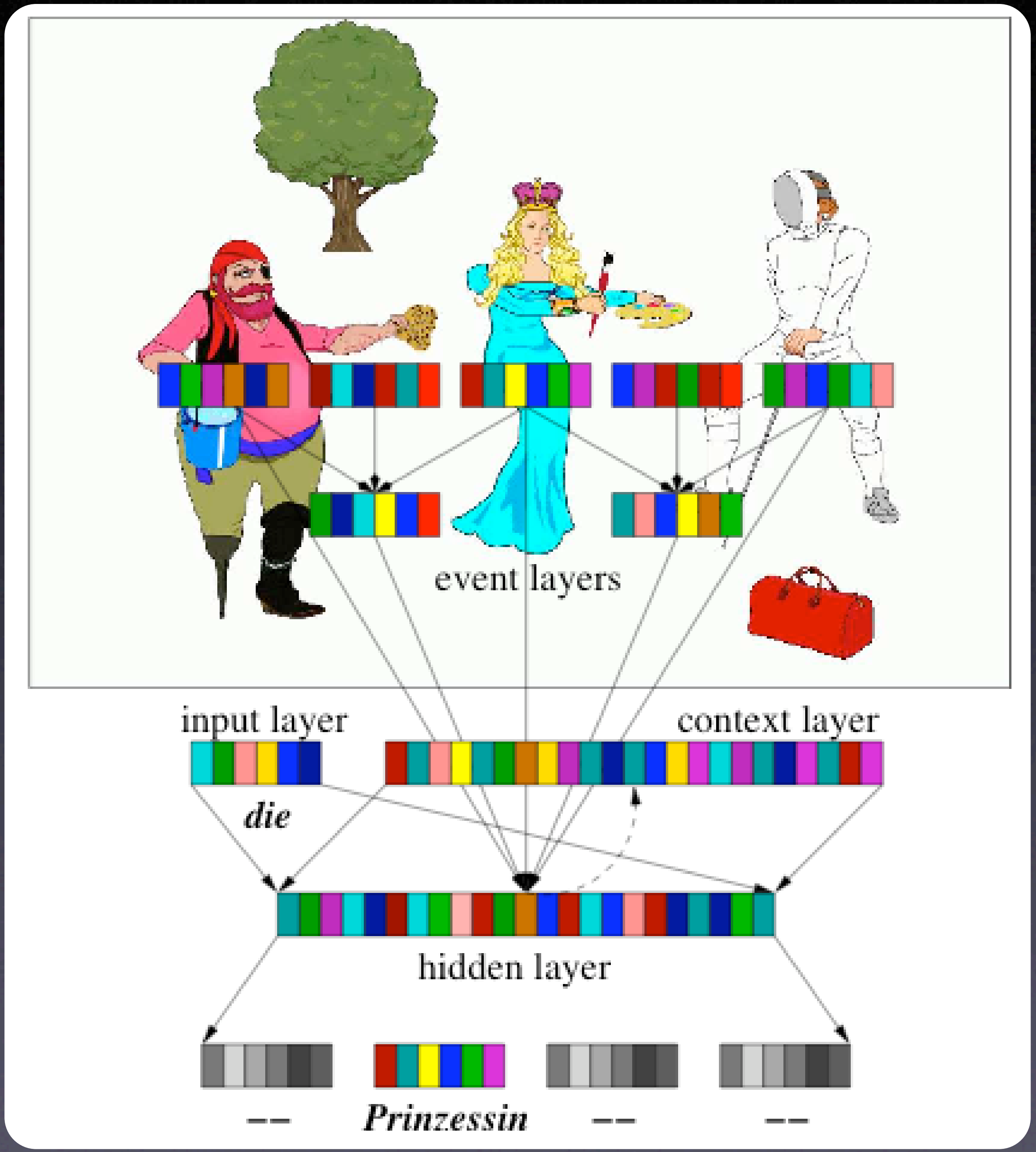


die Prinzessin waescht der Pirat

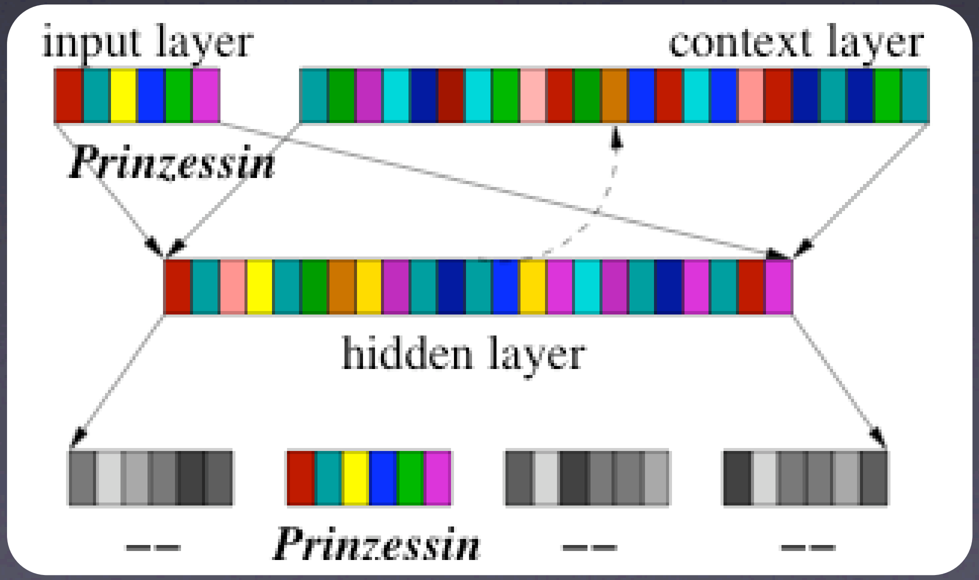
⇒ Princess

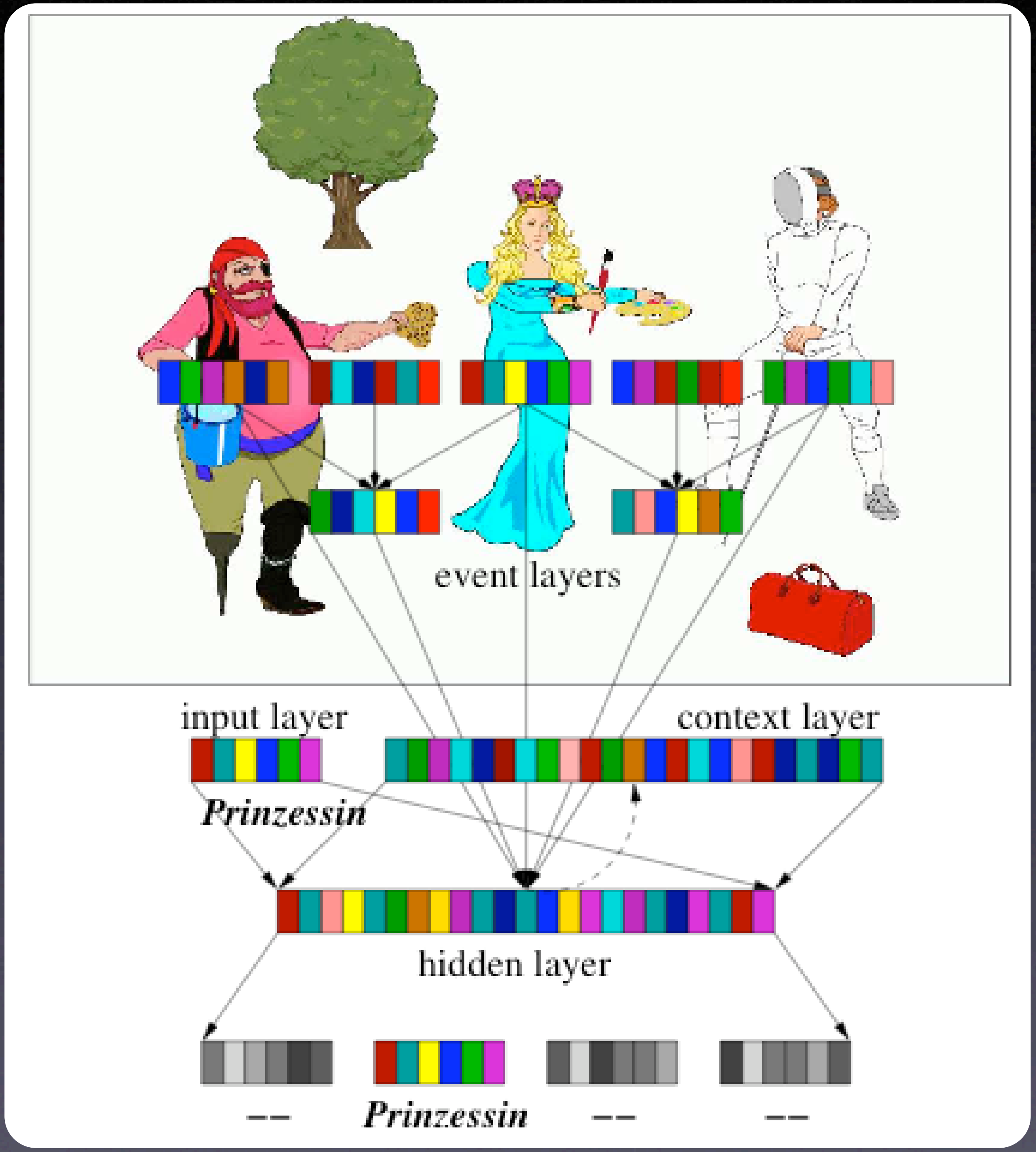
anticipated because it is the only feminine object depicted





die *Prinzessin* *waescht der Pirat*



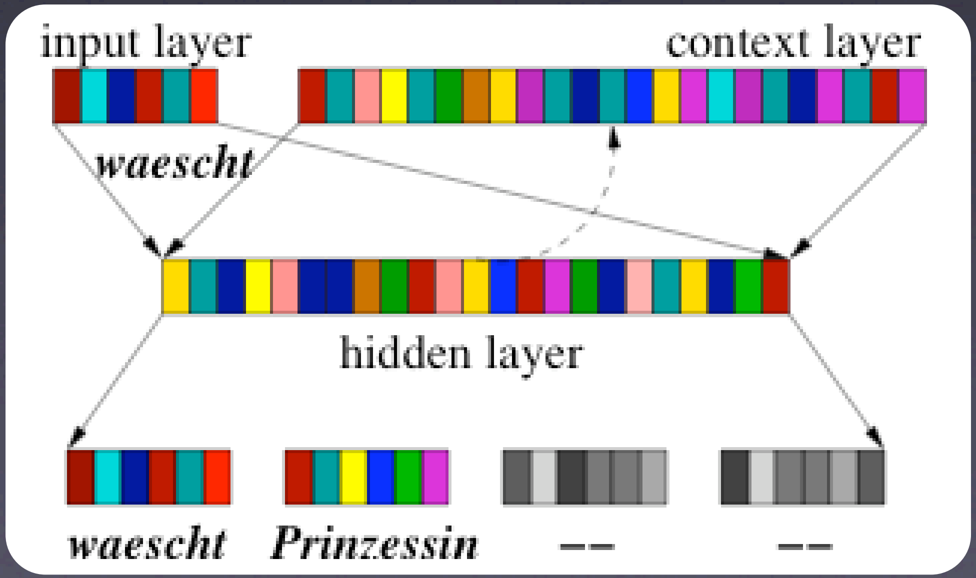


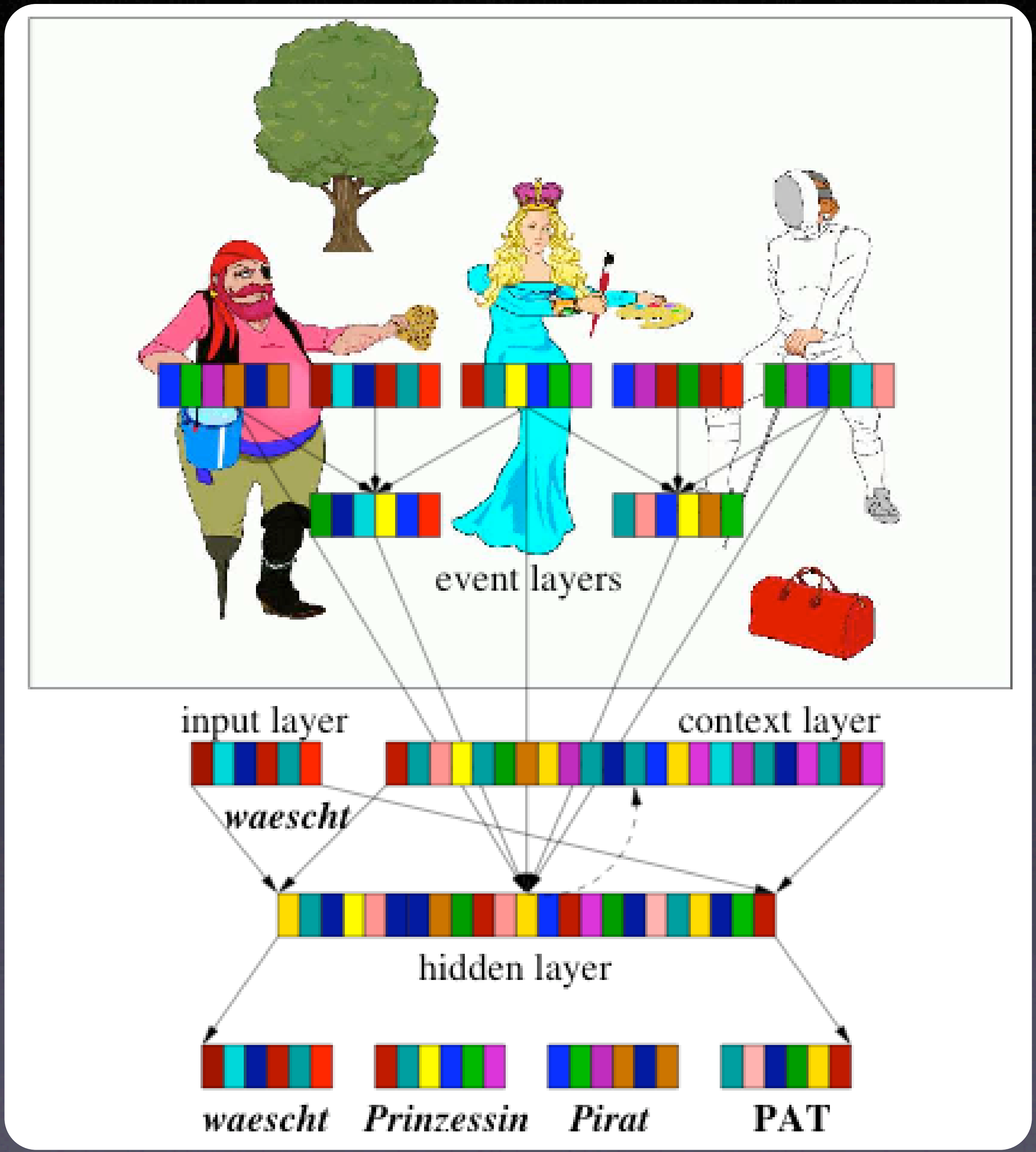
die Prinzessin waescht der Pirat

⇒ Processing of

washes

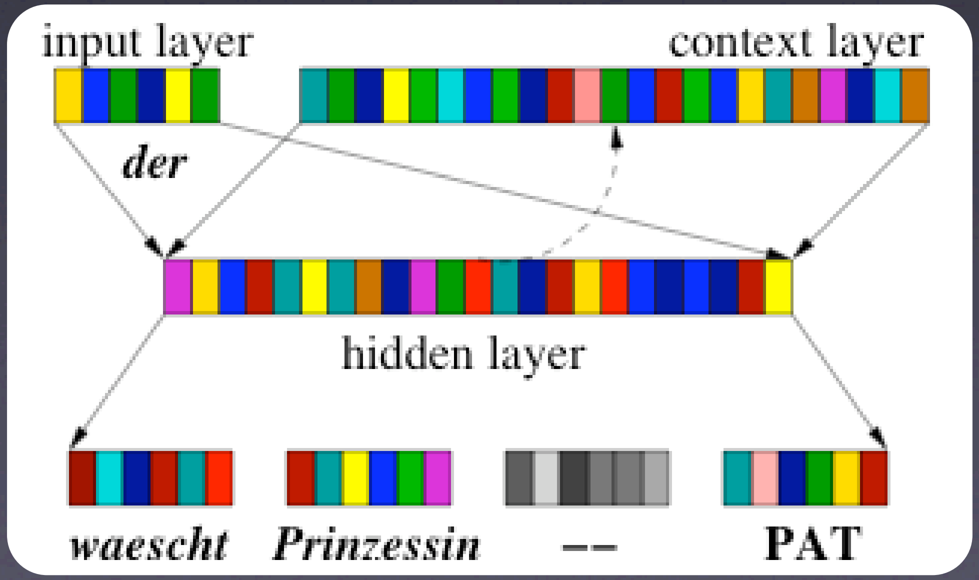
enables recovery of all event information from depicted event

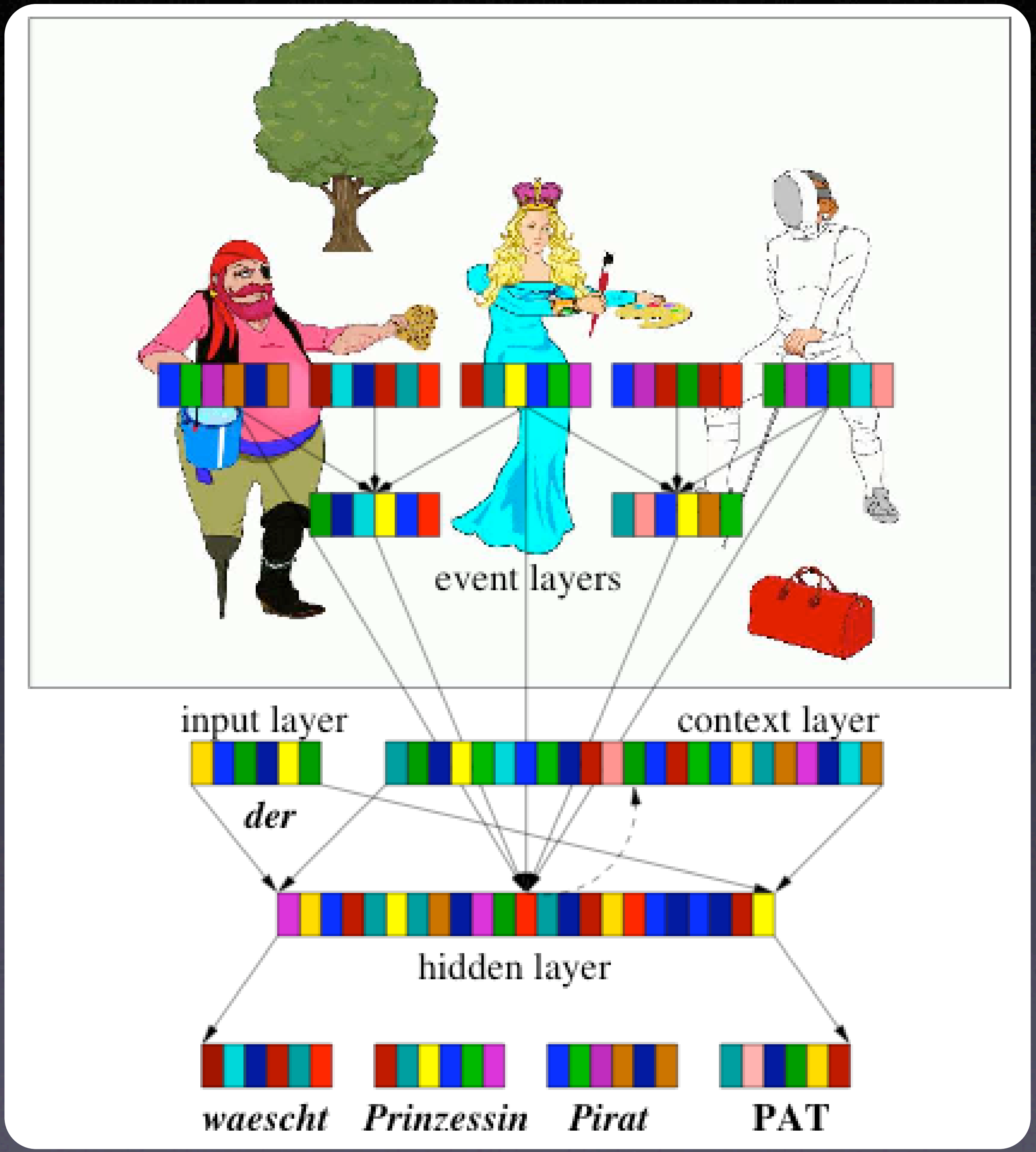




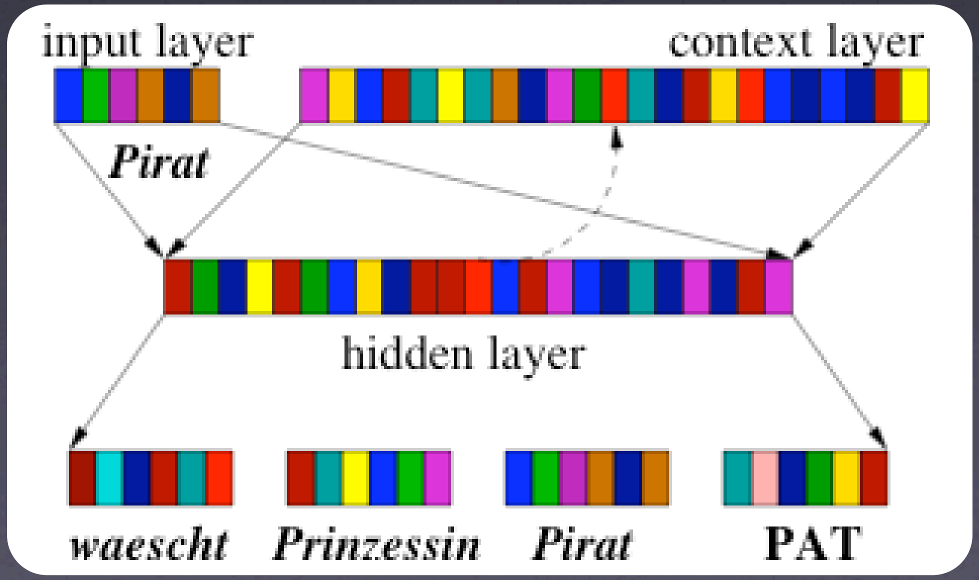
die Prinzessin waescht der Pirat

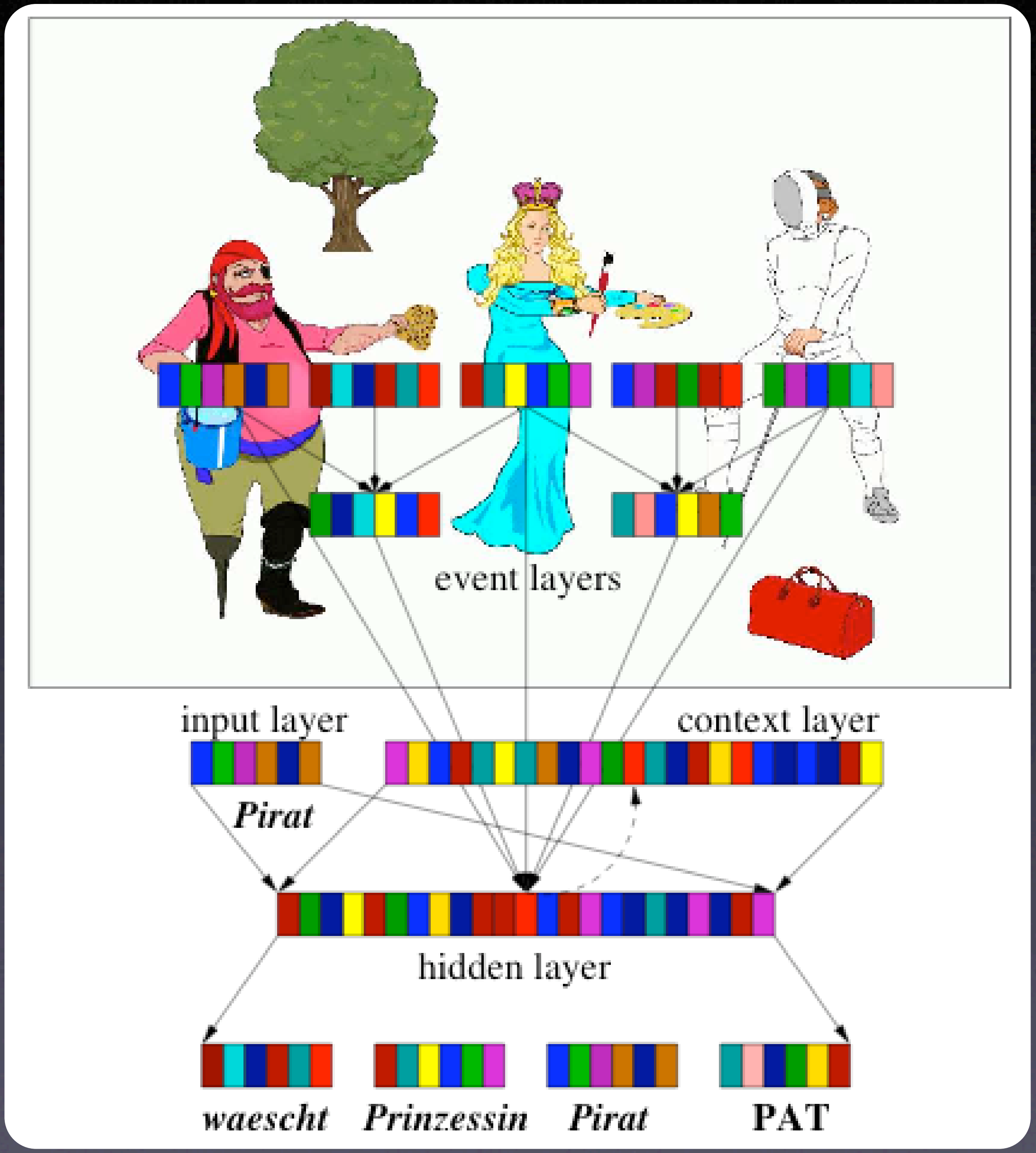
⇓ Processing of nominative article *der* establishes role of *Princess* as Patient





die Prinzessin waescht der Pirat





A Connectionist Model of Scene & Sentence

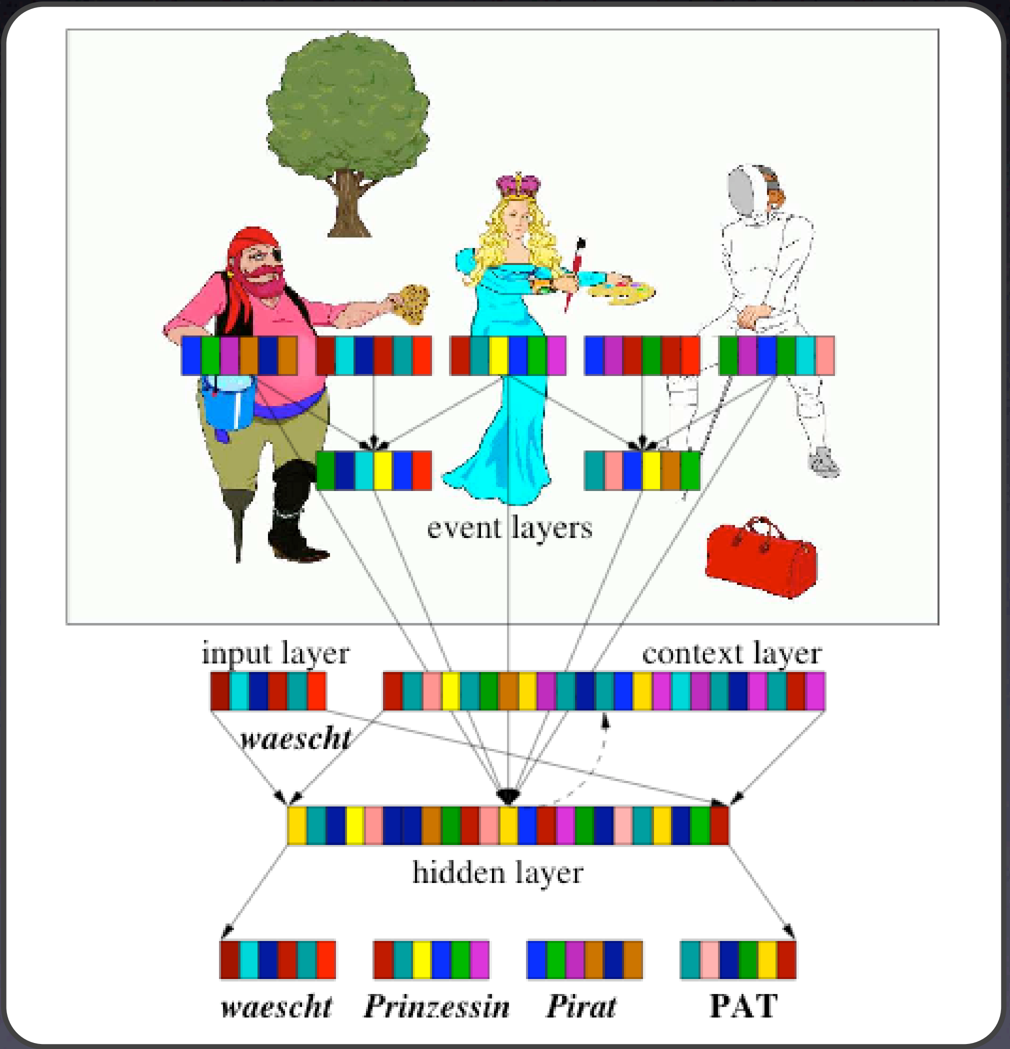
Trained to model materials from 5 visual world studies

- SRN + Scene

Successfully models the use of:

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

Exhibits anticipatory behaviour



Mayberry & Crocker, *CUNY*, 2005.
Mayberry, Crocker & Knoeferle, 2005a,b,c.

Conclusions

Linguistic, Prosodic and World Knowledge incrementally integrated: forward inferences guide visual attention

Event information in the visual scene can immediately influence role assignment and structural disambiguation

Connectionist model can capture integration, adaptation, acquisition, and predict the preferred reliance on scene

Language Science & Technology

Computational Linguistics

Deep, general language processing

- formal representations
- parsing and generation
- semantics and discourse

Language Technologies

Task specific, robust, efficient

- shallow representations
- broad coverage
- probabilistic

Shared Methods:
deep competence
use of context
algorithms for processing
complexity theories

Cognitive Models

Deep, robust, efficient, situated

- model human behaviour
- learn from experience
- adaptive to information & task

Shared Techniques:
probabilistic mechanisms
efficiency
task oriented
experience-based