Dynamics and Modeling in Cognitive Science - II

Narayanan Srinivasan

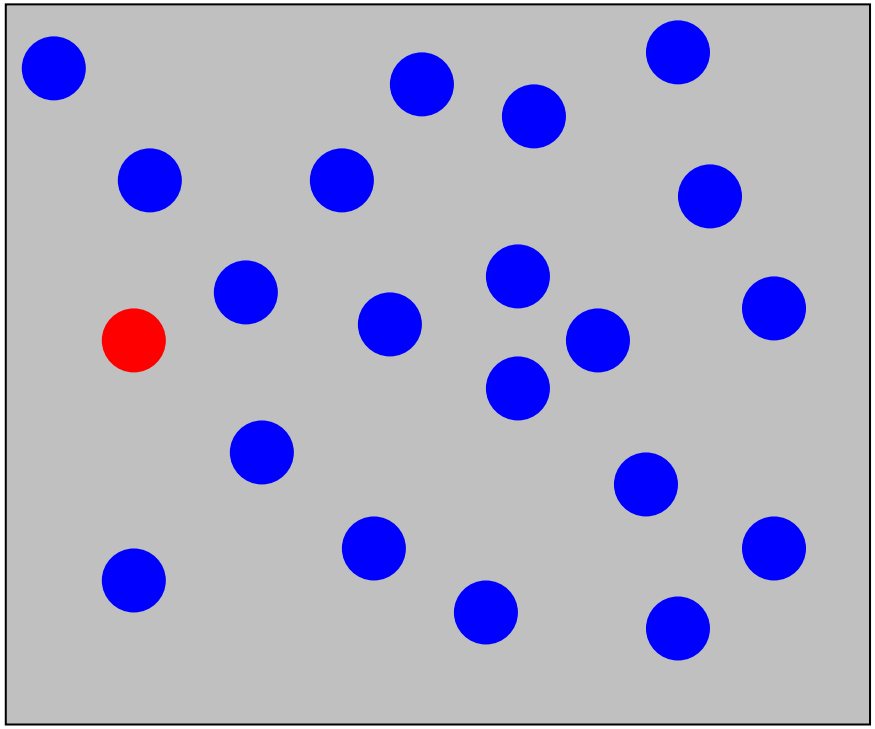
Centre of Behavioural and Cognitive Sciences

University of Allahabad, India

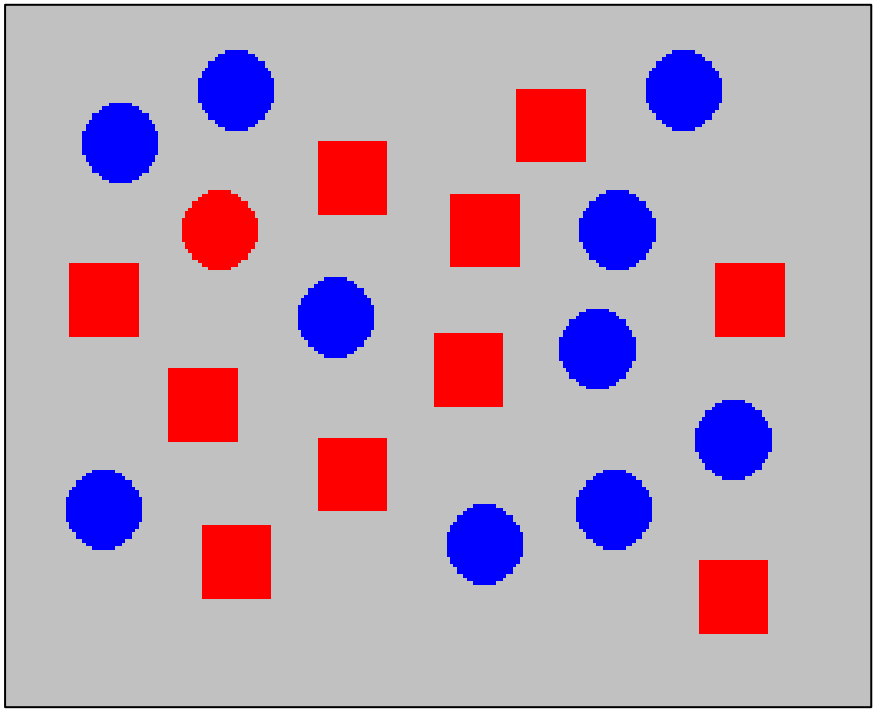
Outline

- Mathematical Modeling
- Symbolic Models
- Connectionist Models
  - Language
  - Perception
- Dynamics
  - Perception and Action
    - HKB Model
    - Bistability
    - Visual search
  - Consciousness

Visual Search



Poput

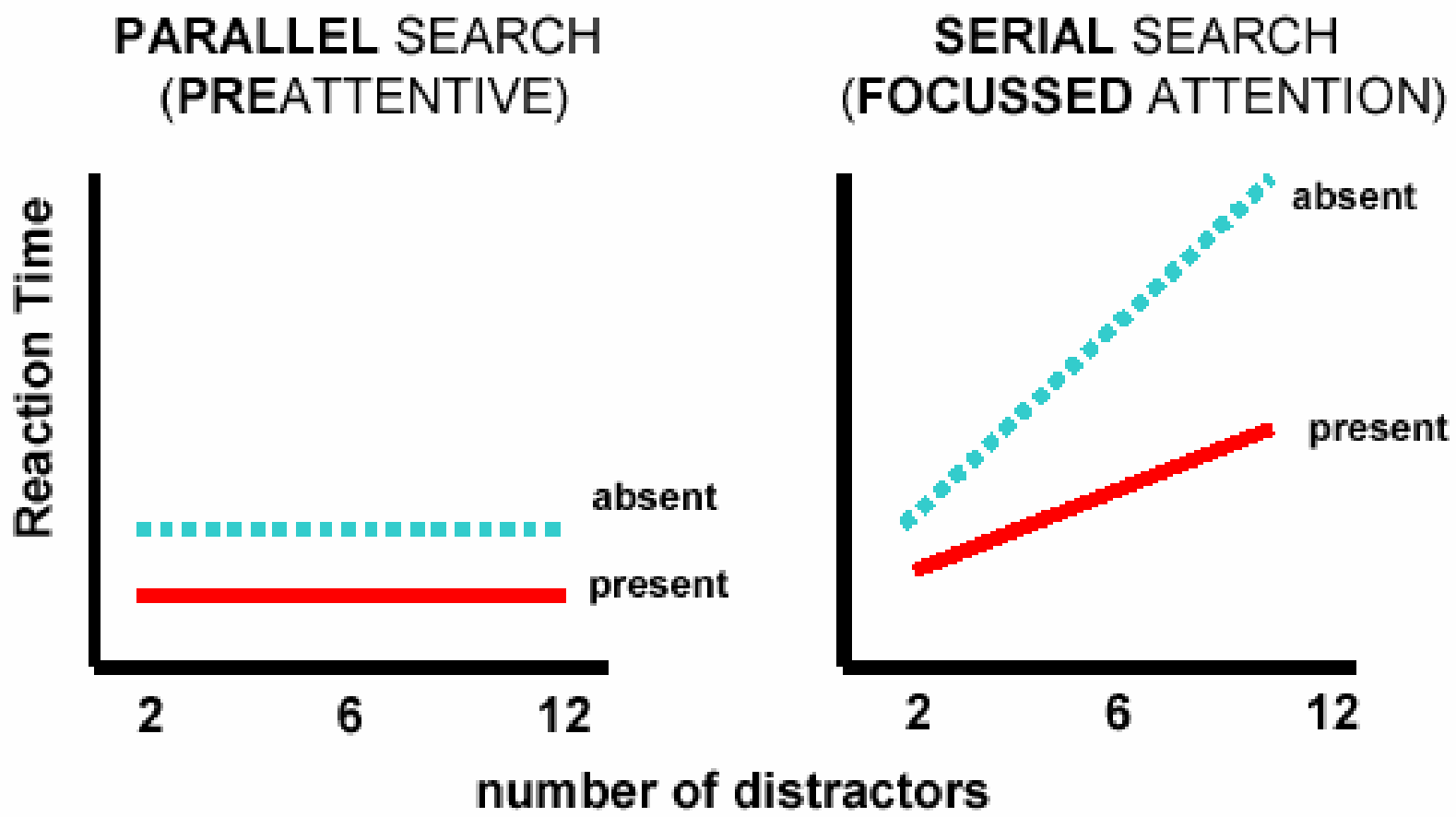


Conjunction – more difficult

Serial Search

- Assumes that items are examined one at a time.
- Search terminates when the target is found or all items have been examined.
- If the target is *present*, on average, how many items need to be examined?
  - roughly half
  - $(n+1)/2$
  - where n = size of the search set
- If *absent*, all items will be examined.

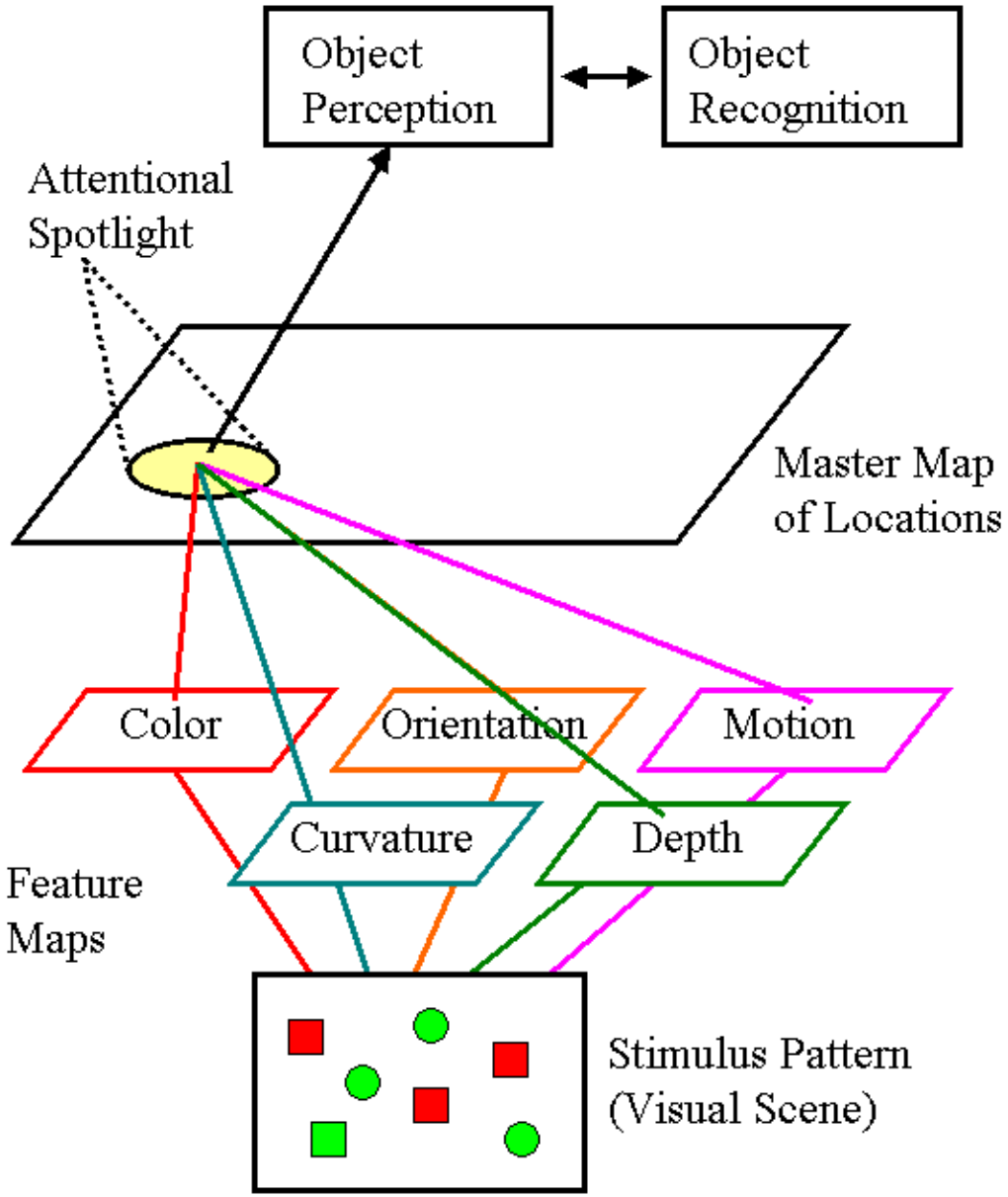
RT curves for parallel/serial search



Attention

- Information processing is supposed to happen in two stages
  - Preattentive stage
  - Attentive stage
- Preattentive processing can be defined as quick and basic feature analysis of the visual field, on which the attention can subsequently operate.

Feature Integration Theory (Treisman)



Guided Search (Wolfe)

- Computes saliency maps. Activation determined:
- Bottom up:
  - Attention is attracted to items that are highly dissimilar from their neighbors (local differences).
- Top-down attentional set.

Code Theory

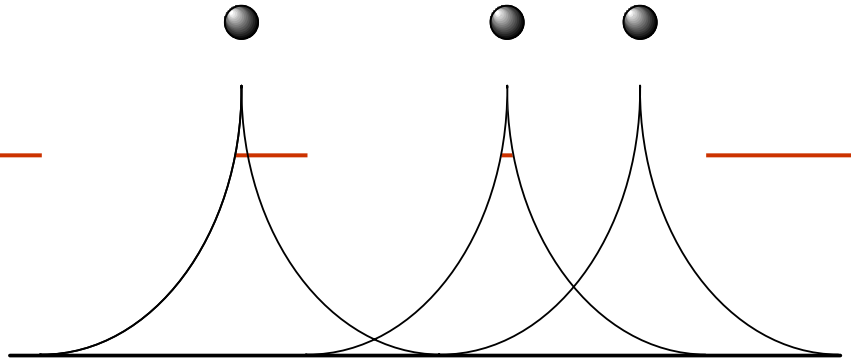
- The spatial element comes from CODE (COntour DEtection theory of grouping by proximity) (Offelen & Vos, 1982, 1983)

- The object-based input comes from Bundesen's TVA (Theory of Visual Attention) (Bundesen, 1990)

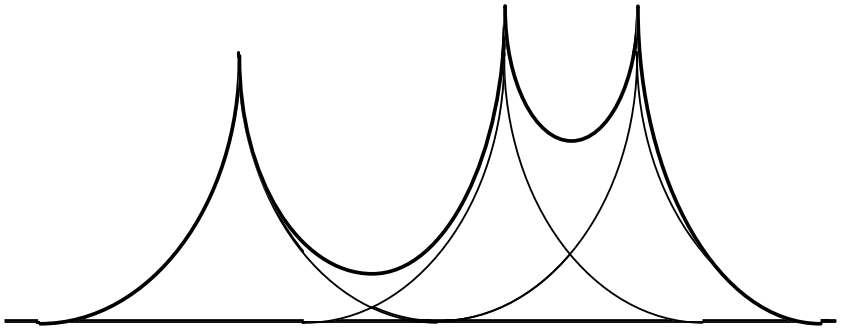
How CODE provides the input for CTVA?

- Objects are represented on an analogue map separated by a Euclidean metric.
- Features or items are not represented on the map as points, but as distributions across space.
- A threshold is applied to these distributions to turn the perceptual items into "quasi-discrete/quasi-analogue" (p606) representations of objects.

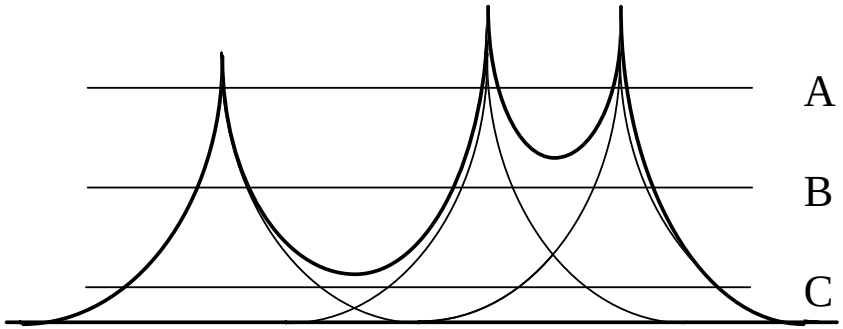
Three Dots:



...create impressions on a saliency map.

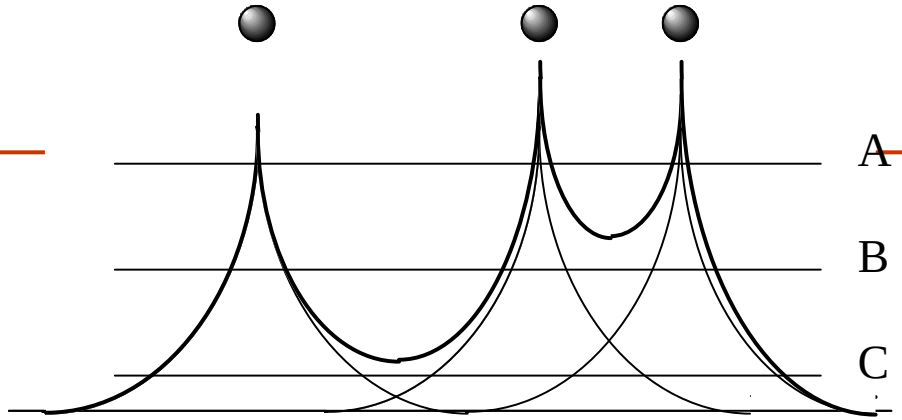


Then add the CODE surface.



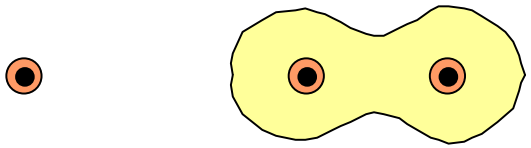
Finally add the threshold to give the feature catches

Three Dots:

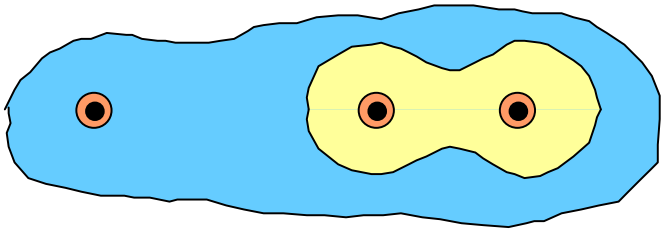


A

B



C

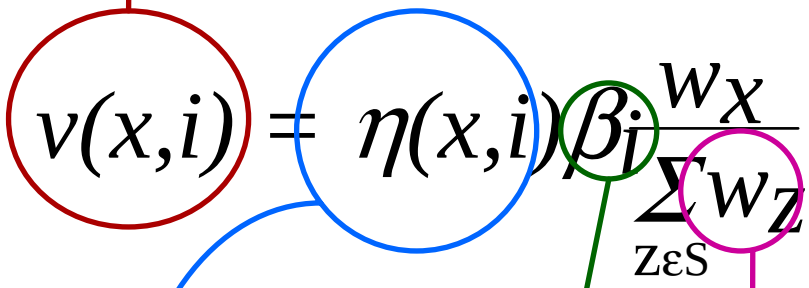


How TVA analyses CODE input:

- TVA assumes two levels of representation-
  - (a) a perceptual level that consists of features of display items;
  - (b) a conceptual level which consists of categorisations of both display items and features.
- They are linked by the following parameter- $\eta(x,i)$
- This parameter is used to select
  - (a) a catagorisation for that object and
  - (b) a within-group perceptual object.

(B) Selection of a ~~within-group perceptual~~ object:

The rate of categorisation is determined by...



(a) The strength of the perceptual evidence

(b) Perceptual Bias

(c) Attentional weight

Summary of CTVA

- Stimuli are represented as distributions in analogue space
- The CODE surface increases with overlapping distributions
- A threshold set by higher cognitive processes determines the size of the feature catches and whether stimuli will be grouped together
- The feature catches attract attention and provide the input for TVA - $h(x,i)$
- TVA adds Perceptual Bias and Attentional Weight
- A race for categorization within a feature catch determines which of two or more items within a feature catch is processed

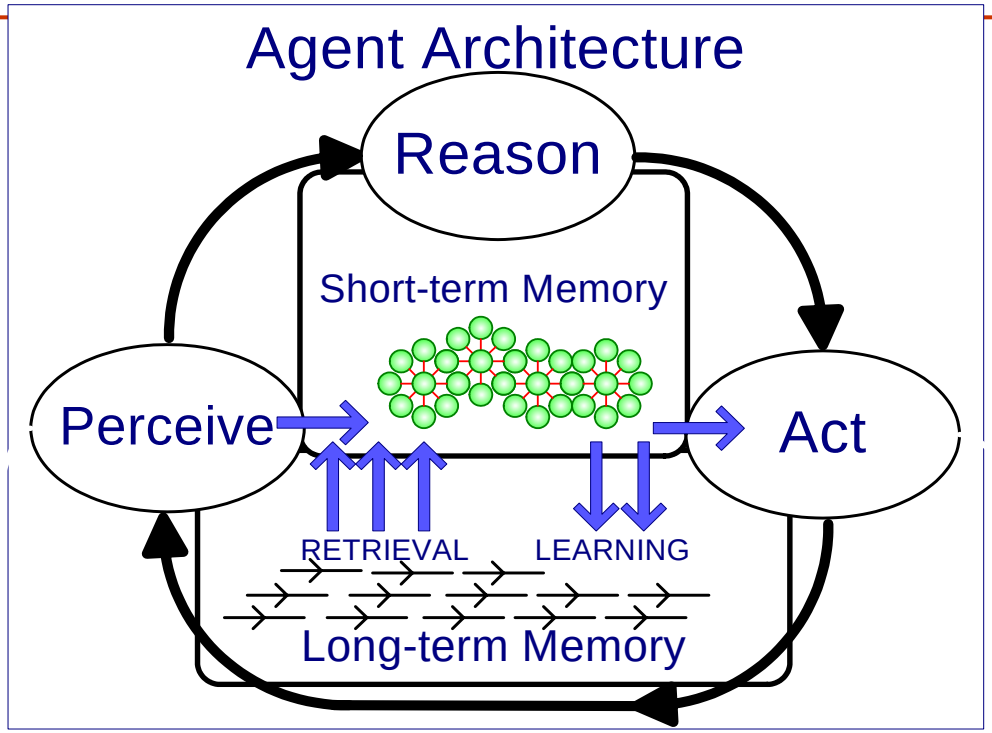
Cognitive Architectures

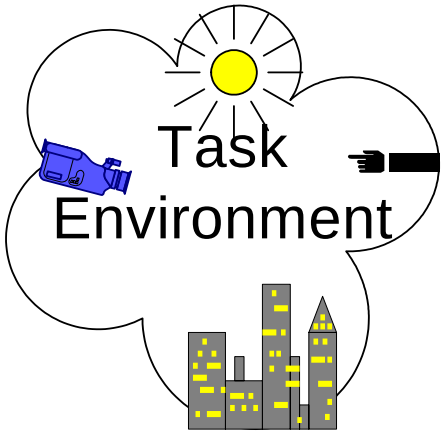
- A cognitive architecture specifies the infrastructure for an intelligent system that remains constant across different domains and knowledge bases.

Why Cog Arch?

A single system (mind) produces all aspects of behavior. It is one mind that minds them all. Even if the mind has parts, modules, components, or whatever, they all mesh together to produce behavior. Any bit of behavior has causal tendrils that extend back through large parts of the total cognitive system before grounding in the environmental situation of some earlier times. If a theory covers only one part or component, it flirts with trouble from the start. It goes without saying that there are dissociations, independencies, impenetrabilities, and modularities. These all help to break the web of each bit of behavior being shaped by an unlimited set of antecedents. So they are important to understand and help to make that theory simple enough to use. But they don't remove the necessity of a theory that provides the total picture and explains the role of the parts and why they exist (Newell, 1990).

What is an architecture?

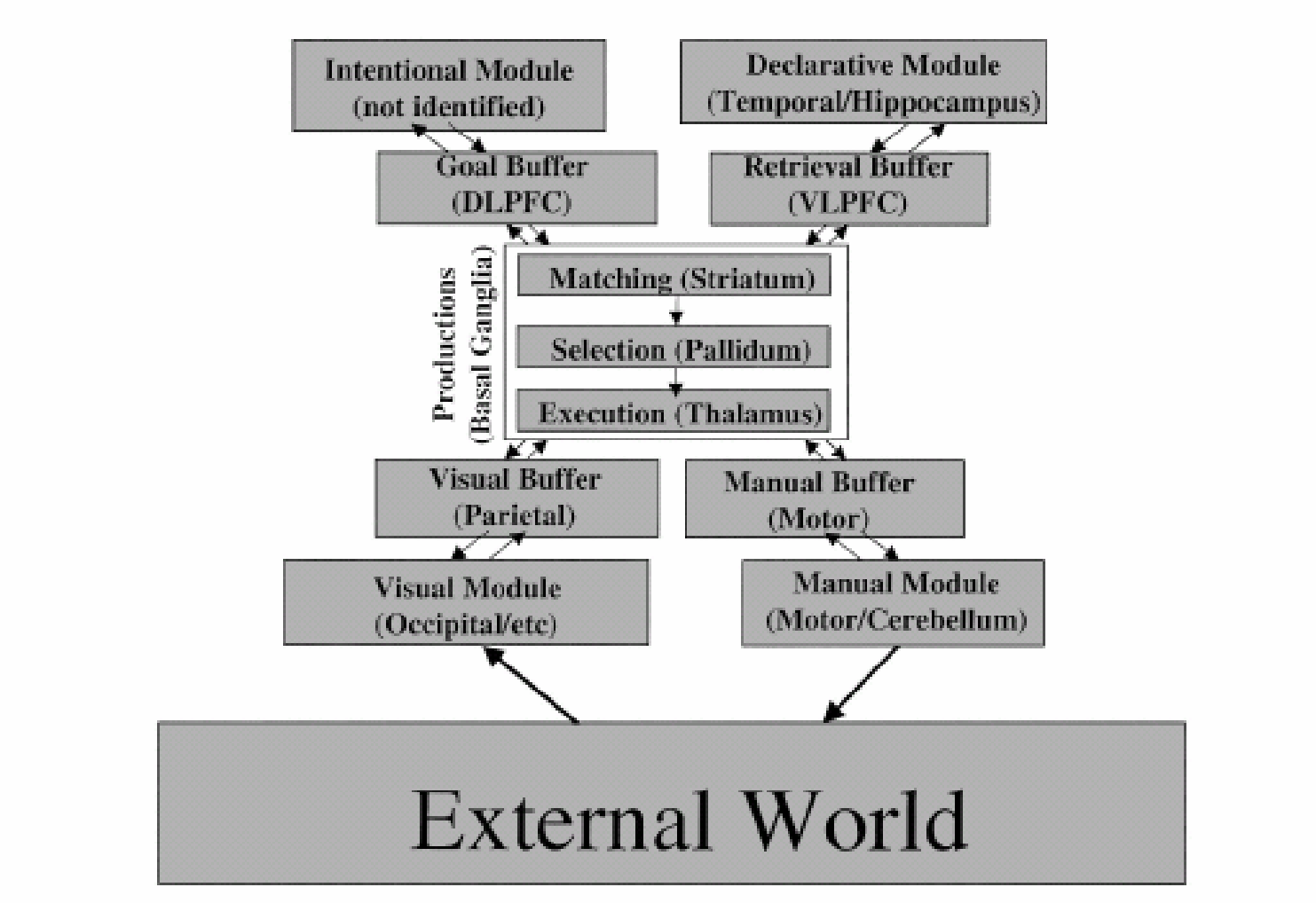




ACT-R

- Tight integration of symbolic and statistical
  - Symbolic level for structured cognition
  - Statistical level for learning and adaptivity
- Massive parallelism within each module
- Asynchronous interaction between modules
- Limited-capacity module interaction
- Central control of cortical areas through procedural module

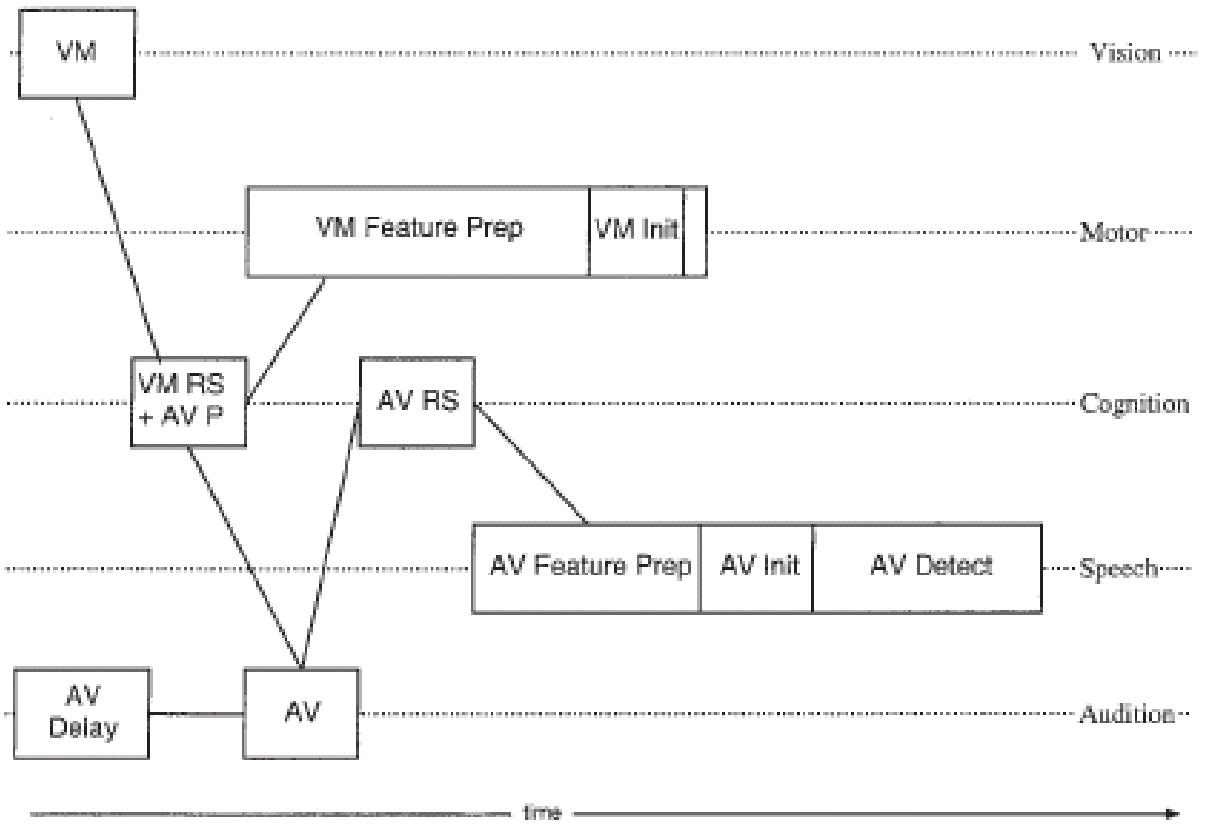
ACT-R 5.0


Intentional Module
(not identified)
Declarative Module
(Temporal/Hippocampus)
Goal Buffer
(DLPFC)
Retrieval Buffer
(VLPFC)
Productions
(Basal Ganglia)
Matching (Striatum)
Selection (Pallidum)
Execution (Thalamus)
Visual Buffer
(Parietal)
Manual Buffer
(Motor)
Visual Module
(Occipital/etc)
Manual Module
(Motor/Cerebellum)
External World

ACT-R

- Inspired by psychological models of memory, skills, and learning
- Optimization-oriented learning and memory
- Dual representations of knowledge
  - Procedural vs. Declarative knowledge units
  - Explicit vs. Semantic retrieval
- Integration of memory, action, and learning
- Highly parameterized
- Belief representation
  - Uniform representation of relational data and objects
  - Automatic semantic retrieval and deliberate management of assumptions
- Goal representations
  - Explicit goal buffer
  - Goals are “normal” memory objects
  - No goal stack
- Plan selection and representation
  - Explicit scripts may be encoded declaratively
  - Plans emerge from atomic conditional knowledge elements
  - Combined explicit selection with decision-theoretic memory model

ACT-R: An Example



Declarative Memory

$$A_i = B_i + \sum_j W_j S_{ji} \quad \text{(activation equation)}$$

where B_i is the base-level activation of the chunk i, the W_js reflect the attentional weighting of the elements that are part of the current goal, and the S_{ji}s are the strengths of association from the elements j to chunk i. Figure 5 displays the chunk encoding for 8 + 4 = 12

$$B_i = \ln\left(\sum_{j=1}^{n} t_j^{-d}\right), \quad \text{(base-level learning equation)}$$

where t_j is the time since the jth practice of an item. This equation

$$P_i = \frac{1}{1 + e^{-(A_i - \tau)/s}}, \quad \text{(probability of retrieval equation)}$$

where s controls the noise in the activation levels and is typically set at about .4. If a chunk is successfully retrieved, the latency of retrieval will reflect the activation of a chunk. The time to retrieve the chunk is given as

$$T_i = Fe^{-A_i}. \quad \text{(latency of retrieval equation)}$$

$$\text{recognition time} = I + Fe^{-A_i},$$

$$F \approx 0.35e^{r},$$

Procedural memory

ACT-R involve these utility calculations. The utility of a production i is defined as

$$U_i = P_i G - C_i \quad \text{(production utility equation)}$$

where P_i is an estimate of the probability that if production i is chosen the current goal will be achieved, G is the value of that current goal, and C_i is an estimate of the cost (typically measured in time) to achieve that goal. As we discuss, both P_i and C_i are learned from experience with that production rule.

$$P_i = \frac{e^{U_i/t}}{\sum_{j}^{n} e^{U_j/t}}, \quad \text{(production choice equation)}$$

where the summation is over all applicable productions and t controls the noise in the utilities. Thus, at any point in time there

$$P = \frac{Successes}{Successes + Failures}, \quad \text{(probability of success equation)}$$

Dynamic Approaches

- Connectionist Models
  - Associators
  - Back propagation
  - Recurrent networks
    - Simple recurrent networks
    - Hopfield networks/Boltzmann machines
- Non-connectionist
  - Dynamics
    - State-space, attractors, limit cycles, oscillations

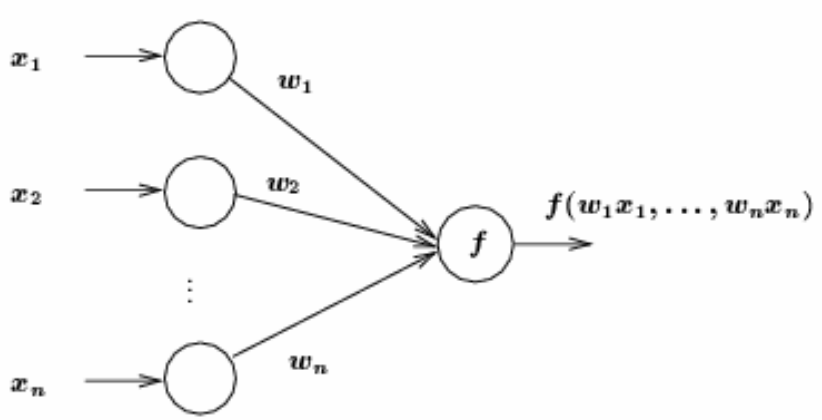
Introduction

- What is an (artificial) neural network
  - A set of nodes (units, neurons, processing elements)
    - Each node has input and output
    - Each node performs a simple computation by its node function
  - Weighted connections between nodes
    - Connectivity gives the structure/architecture of the net
    - What can be computed by a NN is primarily determined by the connections and their weights
  - A very much simplified version of networks of neurons in animal nerve systems

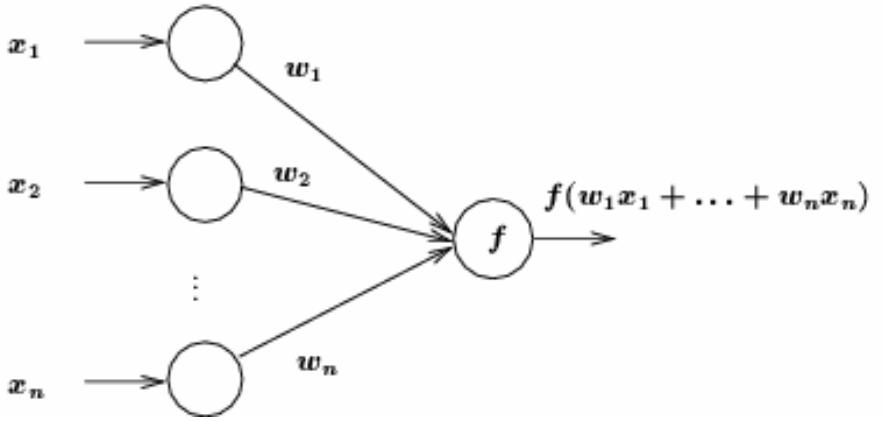
ANN Neuron Models

- Each node has one or more inputs from other nodes, and one output to other nodes
- Input/output values can be
  - Binary {0, 1}
  - Bipolar {-1, 1}
  - Continuous
- All inputs to one node come in at the same time and remain activated until the output is produced
- Weights associated with links
- $f(net)$ is the node function

  $net = \sum_{i=1}^{n} w_i x_i$ is most popular



General neuron model



Weighted input summation

Node Function

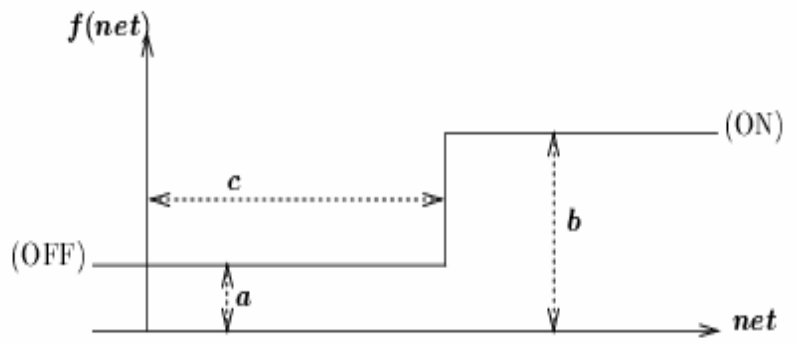
- Identity function : $f(net) = net.$
- Constant function : $f(net) = c.$
- Step (threshold) function

$$f(\text{net}) = \begin{cases} a & \text{if net} < c \\ b & \text{if net} > c \end{cases}$$

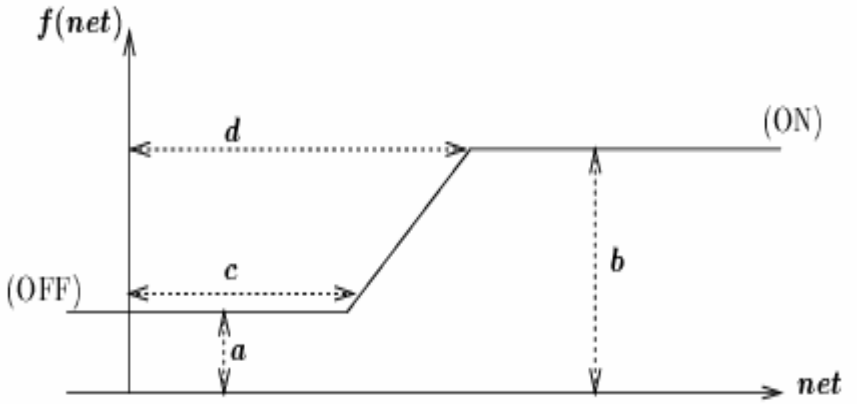
where c is called the threshold

- Ramp function

$$f(\text{net}) = \begin{cases} a & \text{if net} \leq c \\ b & \text{if net} \geq d \\ a + \frac{(\text{net} - c)(b-a)}{(d-c)} & \text{otherwise} \end{cases}$$



Step function



Ramp function

Node Function

- **Sigmoid function**
  - S-shaped
  - Continuous and everywhere differentiable
  - Rotationally symmetric about some point (*net* = *c*)
  - Asymptotically approach saturation points

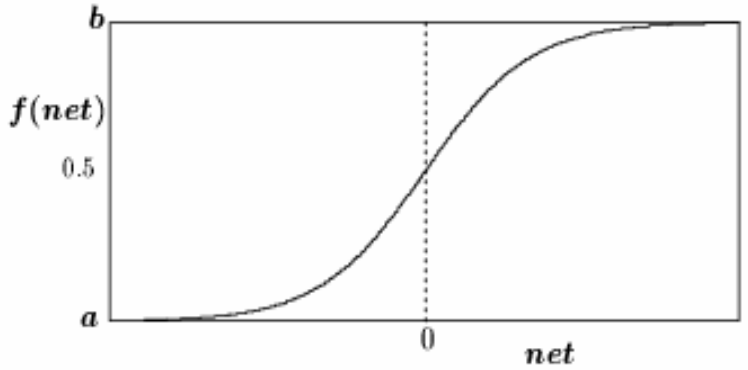
$$\lim_{\text{net}\to-\infty} f(\text{net}) = a$$

$$\lim_{\text{net}\to\infty} f(\text{net}) = b$$

  - Examples:

$$f(\text{net}) = z + \frac{1}{1+\exp(-x\cdot\text{net}+y)}$$

$$f(\text{net}) = \tanh(x\cdot\text{net}-y)+z,$$



Sigmoid function

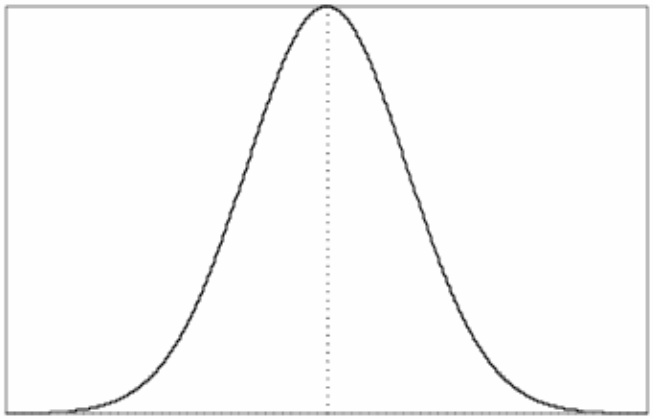
When $y = 0$ and $z = 0$:
$a = 0$, $b = 1$, $c = 0$.
When $y = 0$ and $z = -0.5$
$a = -0.5$, $b = 0.5$, $c = 0$.

Larger x gives steeper curve

Node Function

- **Gaussian function**
  - Bell-shaped (radial basis)
  - Continuous
  - *f*(*net*) asymptotically approaches 0 (or some constant) when |*net*| is large
  - Single maximum (when *net* = μ)
  - Example:

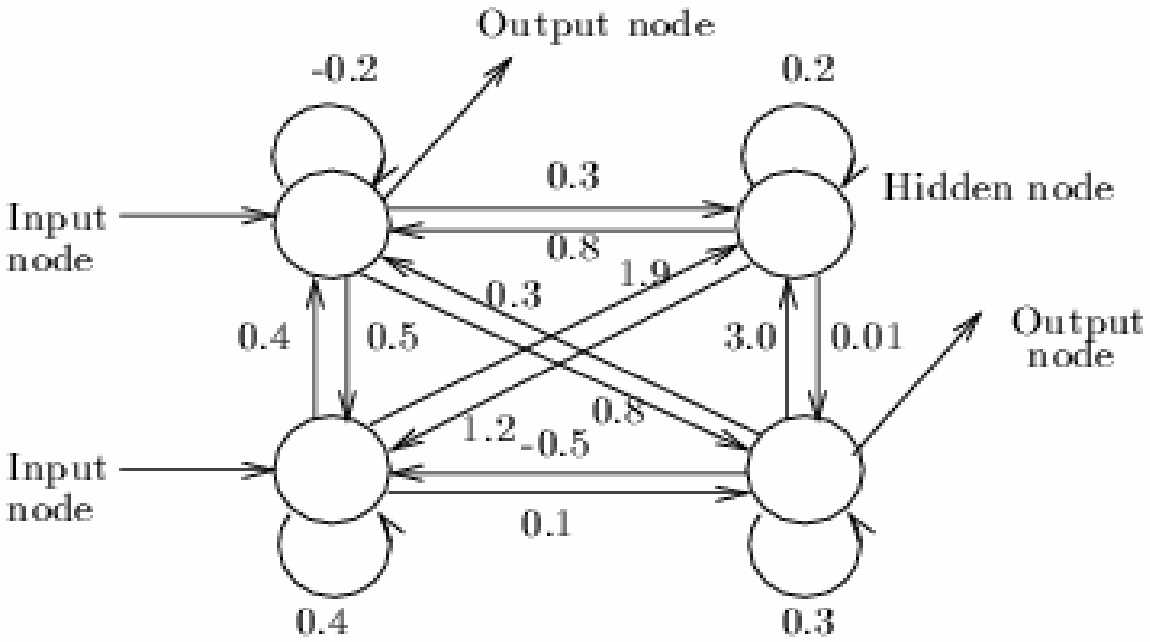
$$f(\text{net}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{\text{net}-\mu}{\sigma}\right)^2\right]$$



Gaussian function

Network Architecture

- **(Asymmetric) Fully Connected Networks**
  - Every node is connected to every other node
  - Connection may be excitatory (positive), inhibitory (negative), or irrelevant (≈ 0).
  - Most general
  - Symmetric fully connected nets: weights are symmetric ($w_{ij} = w_{ji}$)



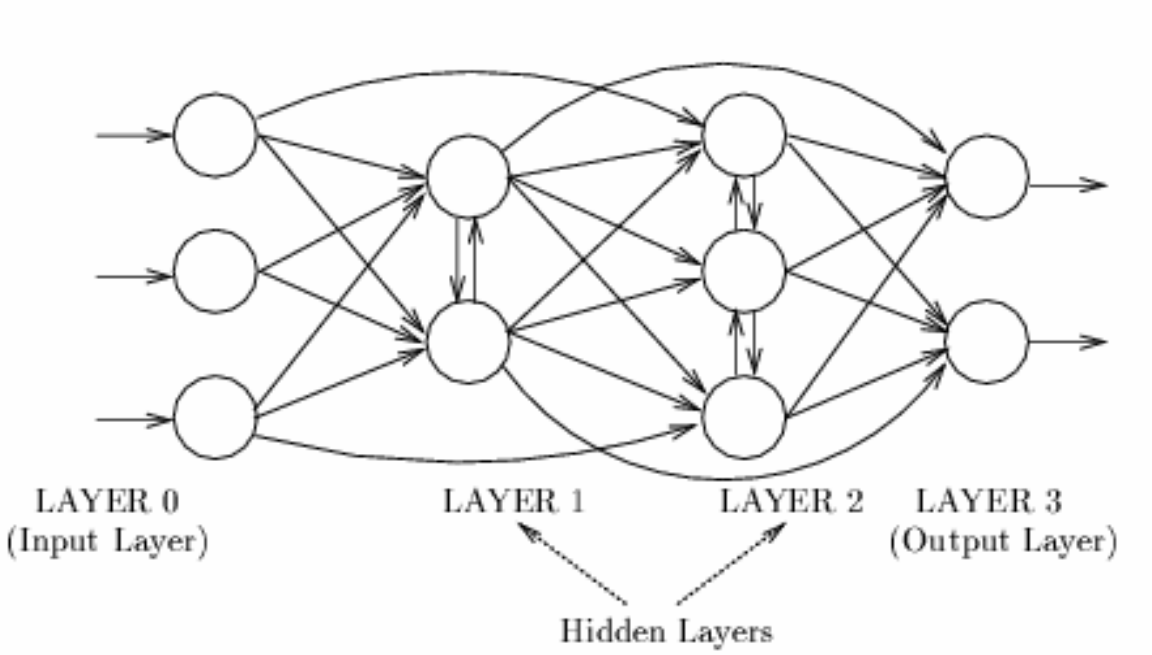
Input nodes: receive input from the environment

Output nodes: send signals to the environment

Hidden nodes: no direct interaction to the environment

Network Architecture

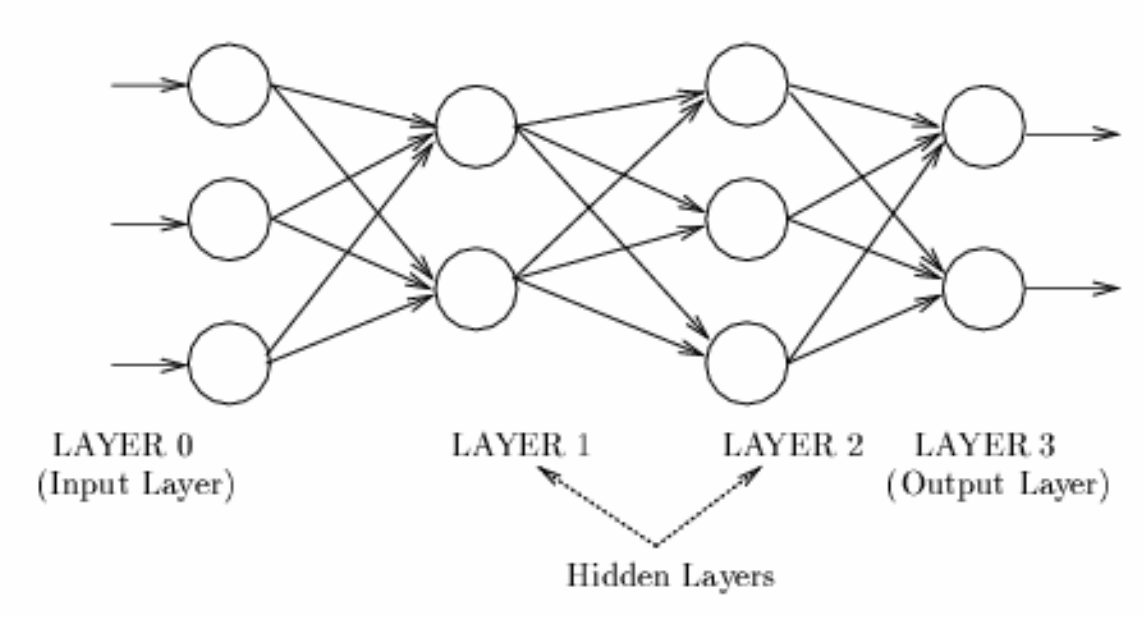
- **Layered Networks**
  - Nodes are partitioned into subsets, called layers.
  - No connections that lead from nodes in layer *j* to those in layer *k* if *j* > *k*.



- Inputs from the environment are applied to nodes in layer 0 (**input layer**).
- Nodes in input layer are place holders with no computation occurring (i.e., their node functions are identity function)

Network Architecture

- **Feedforward Networks**
  - A connection is allowed from a node in layer *i* only to nodes in layer *i* + 1.
  - Most widely used architecture.



Conceptually, nodes at higher levels successively abstract features from preceding layers

Network Architecture

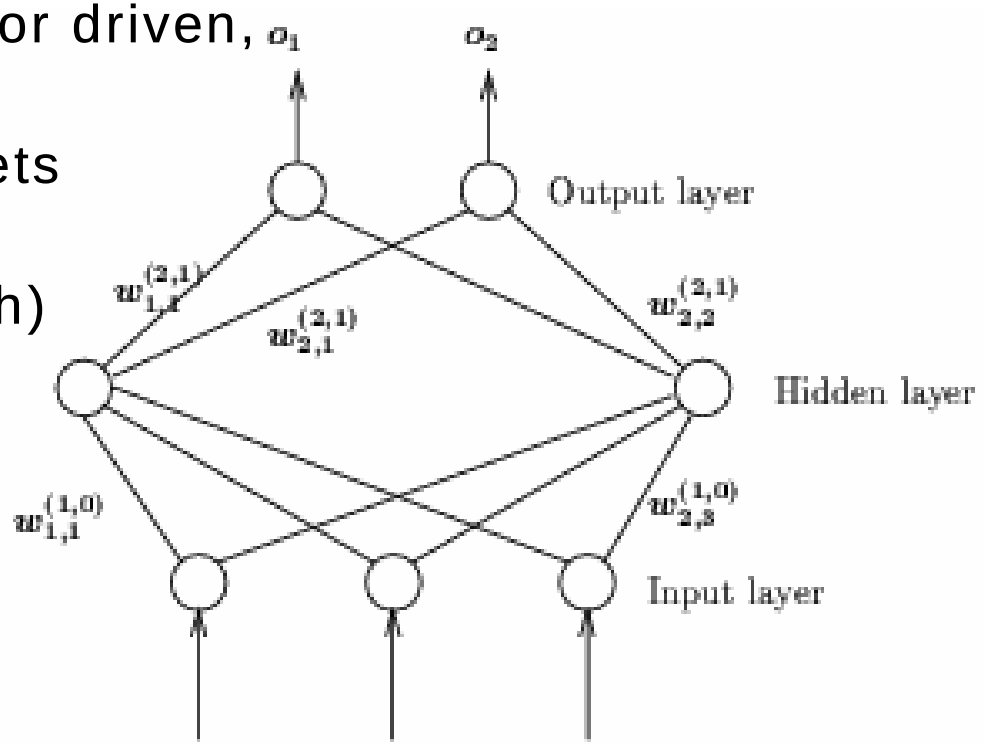
- Acyclic Networks
  - Connections do not form directed cycles.
  - Multi-layered feedforward nets are acyclic
- Recurrent Networks
  - Nets with directed cycles.
  - Much harder to analyze than acyclic nets.
- Modular nets
  - Consists of several modules, each of which is itself a neural net for a particular sub-problem
  - Sparse connections between modules

Backpropagation Learning

- **Architecture**:
  - **Feedforward** network of at least one layer of **non-linear** hidden nodes, e.g., # of layers $L \geq 2$ (not counting the input layer)
  - Node function is differentiable
    most common: **sigmoid function**

$$S(net) = \frac{1}{1 + e^{(-net)}}$$

- **Learning**: supervised, error driven, generalized delta rule
- Call this type of nets BP nets
- The weight update rule (gradient descent approach)
- Practical considerations
- Variations of BP nets
- Applications



Algorithm Backpropagation;

Start with randomly chosen weights;

while MSE is unsatisfactory

and computational bounds are not exceeded, do

for each input pattern x_p, $1 \leq p \leq P$,

Compute hidden node inputs ($net_{p,j}^{(1)}$);

Compute hidden node outputs ($x_{p,j}^{(1)}$);

Compute inputs to the output nodes ($net_{p,k}^{(2)}$);

Compute the network outputs ($o_{p,k}$);

Modify outer layer weights:

$$\Delta w_{k,j}^{(2,1)} = \eta(d_{p,k} - o_{p,k})\mathcal{S}'(net_{p,k}^{(2)})x_{p,j}^{(1)}$$

Modify weights between input & hidden nodes:

$$\Delta w_{j,i}^{(1,0)} = \eta \sum_k \left((d_{p,k} - o_{p,k})\mathcal{S}'(net_{p,k}^{(2)})w_{k,j}^{(2,1)} \right) \mathcal{S}'(net_{p,j}^{(1)})x_{p,i}$$

end-for

end-while.

Note: if S is a logistic function, then $S'(x) = S(x)(1 - S(x))$

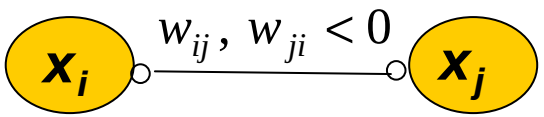
Unsupervised Learning

- Training samples contain only input patterns
  - No desired output is given (teacher-less)
- Learn to form classes/clusters of sample patterns according to similarities among them
  - Patterns in a cluster would have similar features
  - No prior knowledge as what features are important for classification, and how many classes are there.

Ways to realize competition in NN

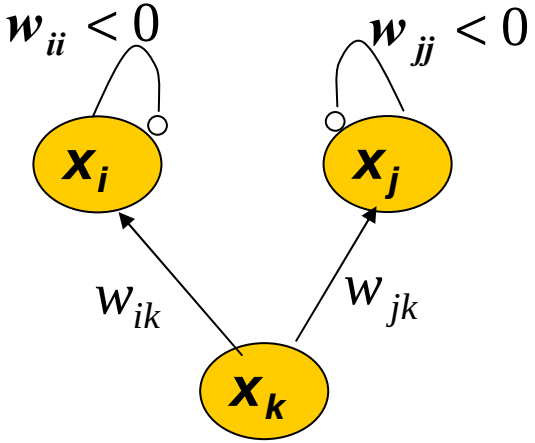
- **Lateral inhibition**

  output of each node feeds to others through inhibitory connections (with negative weights)



- **Resource competition**
  - output of *node k* is distributed to

    node i and *j* proportional to w_{ik} and w_{jk} , as well as x_i and x_j
  - self decay
  - biologically sound



Issues in Language

- Language acquisition
  - How is language acquired or learned?
- Language representation
  - How are the symbols of language represented in memory?
- Language processing
  - What factors influence the processing of language?

How do we learn language?

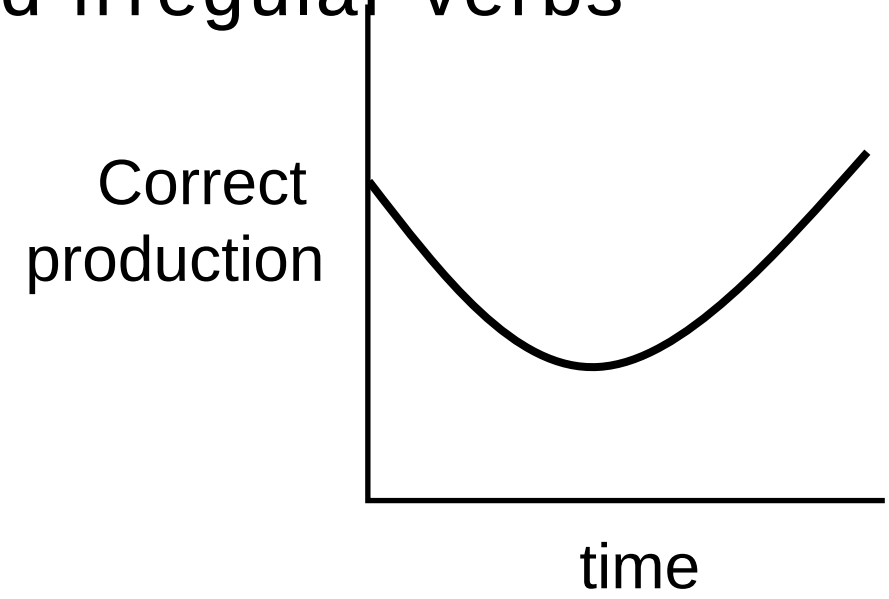
- Chomskyan view
  - Innate knowledge of possible rules of language
  - Children create hypotheses about how these rules apply to the language they are learning
  - We have mental representations of these rules
- Alternate view
  - No explicit representation of rules, although performance can be described in terms of rules

Symbolic vs Dynamic

- Rules
- Representations (symbolic/sub-symbolic/graded)
- Innate vs learned
- Competence vs Performance
- Time

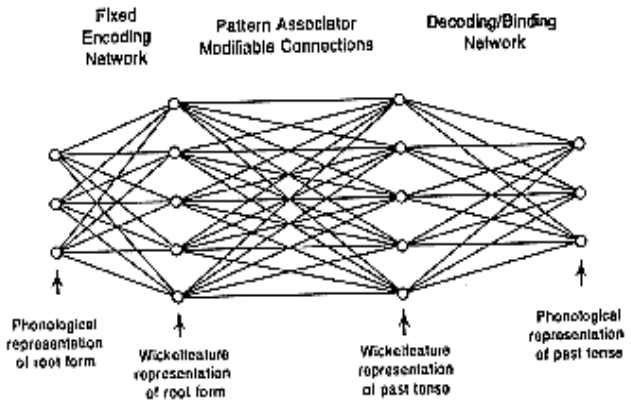
Past-tense acquisition (Brown, 1973)

- 1) specific forms learnt – both regular and irregular
- 2) overgeneralisation of irregular verbs
  - e.g. wented, goed, eated
- 3) correct pronunciation of both regular and irregular verbs



Rumelhart & McClelland's (1986) Model

- Architecture
  - Single layer pattern associator
    - Inputs: present tense (460 units)
    - Outputs: past tense (460 units)
  - Words represented as sets of Wickelfeatures
  - Extra networks at back & front of pattern associator to encode/decode Wickelfeatures from phonological representation



Rumelhart & McClelland's (1986) Model

- Word representation
  - Phonological form: /kAm/ = came
  - But, indistinguishable from /mAk/ or /Akm/
- Wickelphones
  - Context sensitive: #kA, kAm, Am#
  - Can be analysed along 4 dimensions
    - e.g. /A/ = long, low, vowel, front
- Wickelfeatures

Training & Results

- Training sets
  - 10 high frequency words (8 irregular)
  - 410 medium frequency words (76 irregular)
  - 86 low frequency words (14 irregular)
  - Trained on high frequency only; then medium frequency added; low frequency used later
- Results
  - U-shaped curve
  - Overgeneralisation
    - Come → comed, camed
    - Eat → eated

Implications

- Links between regular verb stems and past tense forms can be described using rules, but is governed by a mechanism which does not use explicit rules
- Knowledge of past-formation is distributed across the network
- Links between irregular verb stems and past tense forms are encoded in same set of weights
- In a rule-based account, there would need to be a rule for producing regular verbs and a list of exceptions (irregulars)

Pinker & Prince (1988)

- The U-shaped curve is a result of the way in which the input was presented, not anything to do with the properties of the network
  - The middle of the curve coincides with the addition of the medium frequency verbs
  - Network is flooded by regular verbs – forces network to generalise
- In real language input, there is no such discontinuity

Pinker & Prince (1988)

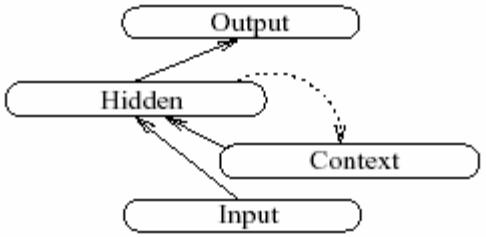
- R&M model does a poor job of generalizing to some novel verbs
  - mail → membled
  - tour → toureder
  - Model doesn't conceive of stem+suffix
    - Cannot encode the formula for creating a past-tense ending
- Task decomposition
  - Past tense treated as autonomous
- Wickelphones & Wickelfeatures

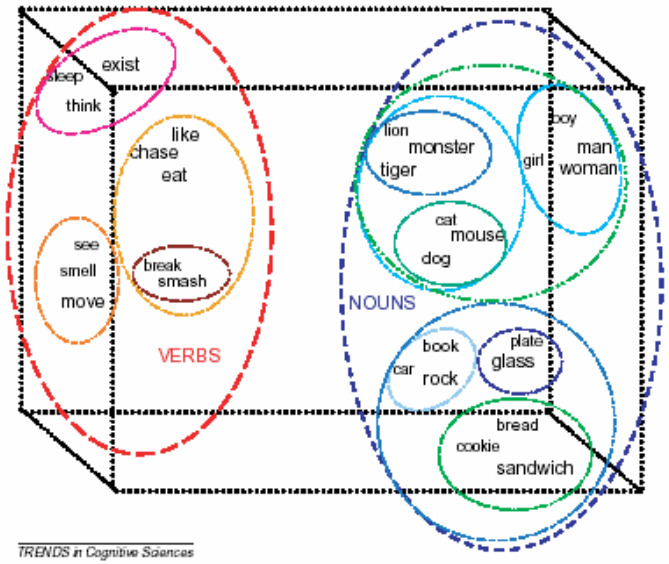
Plunkett & Marchman's (1991) Response

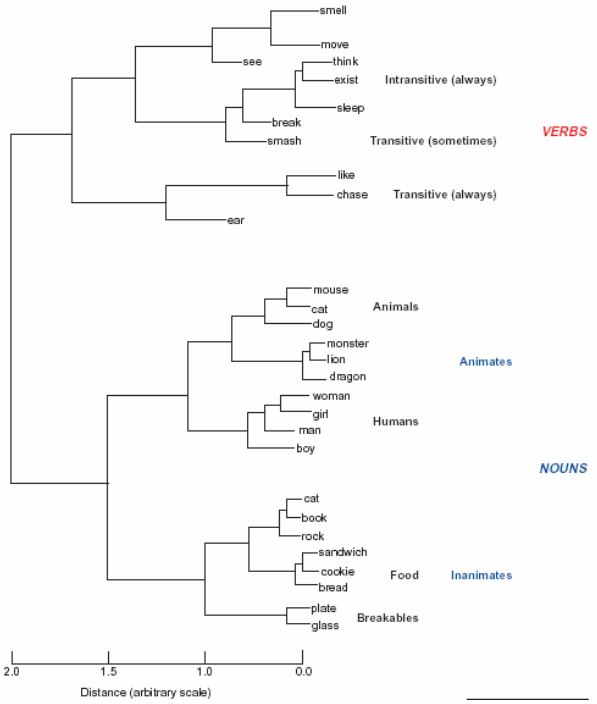
- Model of past-tense acquisition using back-propagation network
  - 3 layer network, 20 units per layer
  - No discontinuity in input
  - Didn't use wickelfeatures
- Parametric studies
  - 74% of tokens irregular – regular not learned
  - 74% of tokens regular – irregular not learned
  - 50/50 (about the same as parental input)– network performed well
  - No global U-shaped curve
  - Micro U-shaped curves – corresponds better to child data as global U-shaped curve is a myth.

Time and Recurrent Neural Nets

- SRNs can learn language based on statistical information available in the stimuli.





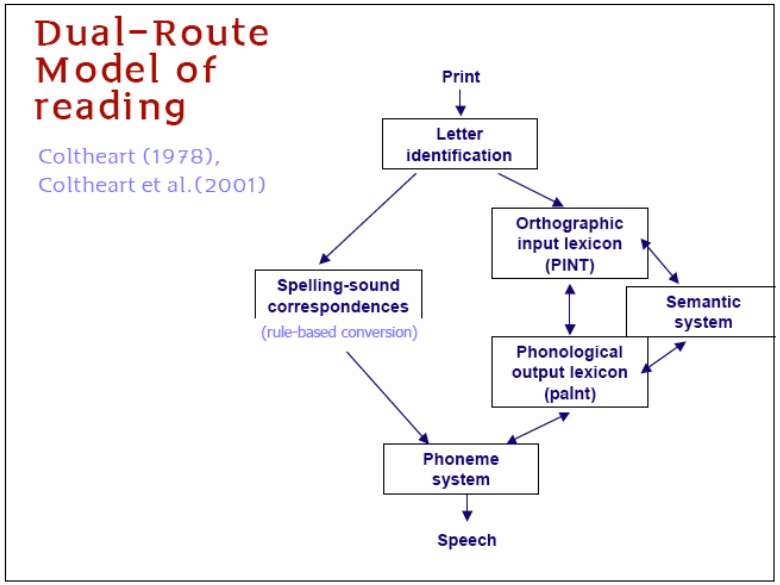


Time and Recurrent Neural Nets

- SRN and mental lexicon
- SRN with complex sentences which contained number agreement between nouns and verbs, different types of verbs (transitive/intransitive), and nouns modified by relative clauses (Elman, 1991).
- Rules as attractors
- Recurrent neural networks can “learn” distinctions such as *subject* and *object*, and generalisations of words at positions not experienced by the network (Elman, 2004).

A model of Reading

- Max Coltheart's multi-route model of reading aloud
- Symbolic (rule based), not a connectionist, model (McClelland, Seidenberg, Plaut, Kello, etc.)
- Passes psychological reality test because it predicts specific error types in reading aloud regular words (e.g., "few"), exception words (e.g., "sew"), and pronounceable non-words (e.g., "tew")
- Passes neural reality test because it predicts specific effects of brain damage on reading
  - Phonological dyslexia
  - Surface dyslexia
  - Deep dyslexia
  - etc



Can connectionism account for human cognition?

- Learning driven by examples
- Knowledge of rules is emergent
  - Multitude of sub-symbolic representations
  - Complex interaction produces behaviour which is rule-like
- Knowledge of rules remains implicit
  - Cannot analyse own activity
  - Cannot form symbolic representations of rules
- To model human development adequately, connectionist systems must be able to:
  - Treat own representations as objects for further manipulation
  - Do so independently of continual training input
  - Retain copies of original networks
  - Form new structured representations

Clark & Karmiloff-Smith (1993)

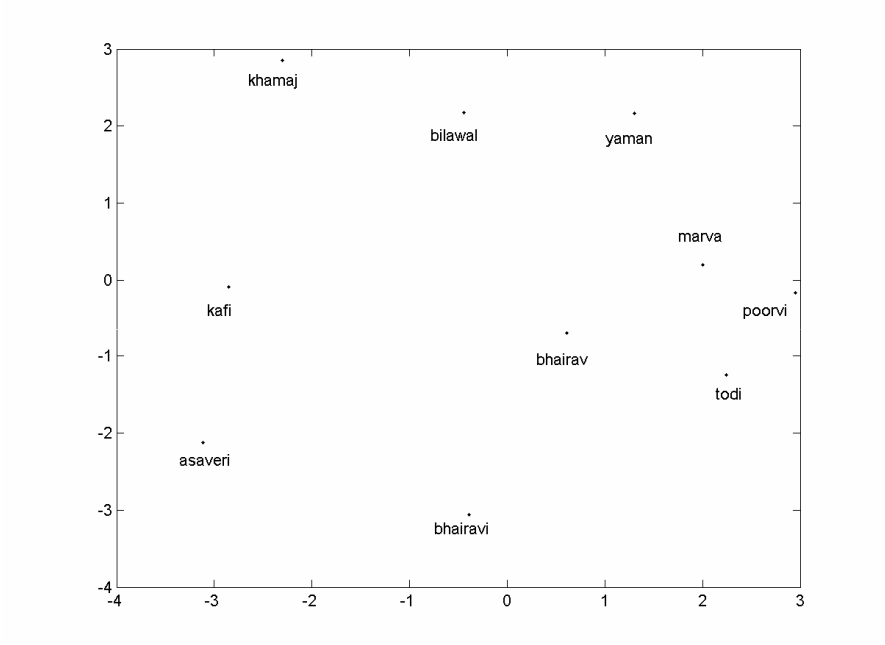
Promise of Connectionist Models

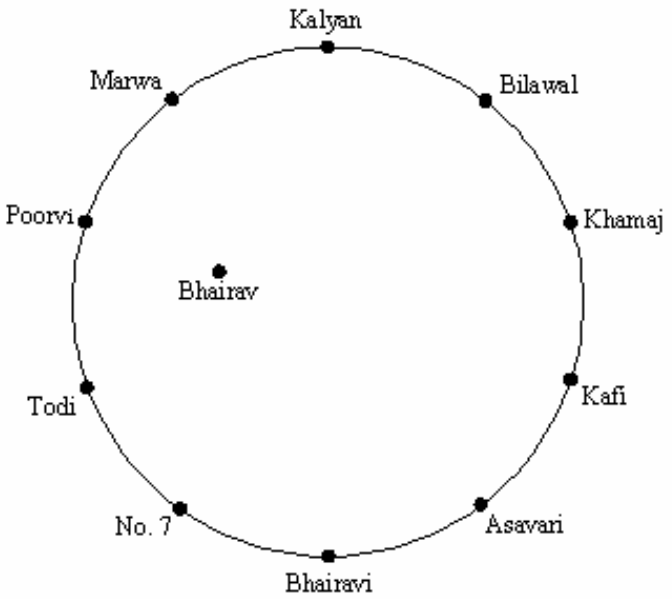
- From the early association models to recurrent models incorporating time, models have been proposed for specific language related phenomena.
- Some models for phenomena like past tense acquisition show reasonable similarities to human behaviour showing the promise of these models.
- These approaches also show a possible way that rules could develop in a system without explicit knowledge of those rules.
- Neural nets show that simple cognitive tasks can be performed without employing features that could correspond to beliefs, desires and plans.
- These models also hold promise in integrating findings from neuroscience and other areas into models for language processing.
- Connectionist models may also provide a seam less interface for combining models for other aspects of cognition like perception, attention and memory.

Dynamical Approaches

Markov Models

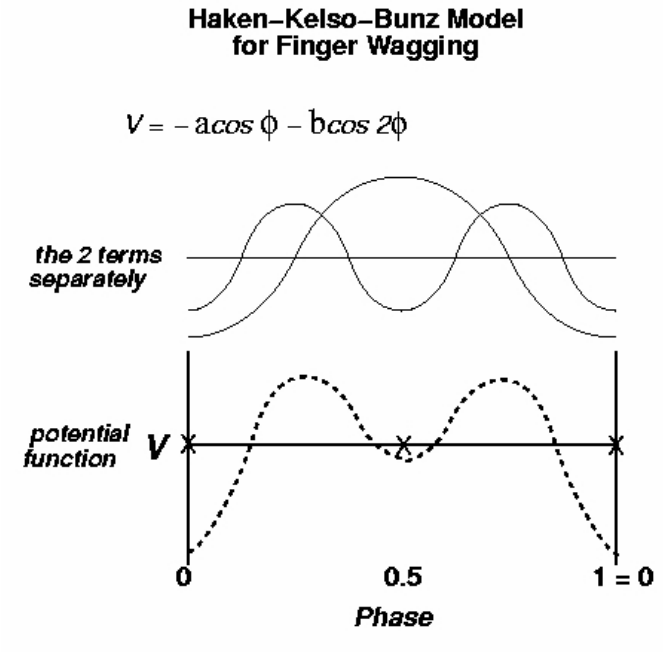
- Hindustani Music





Dynamic Modelling: An example

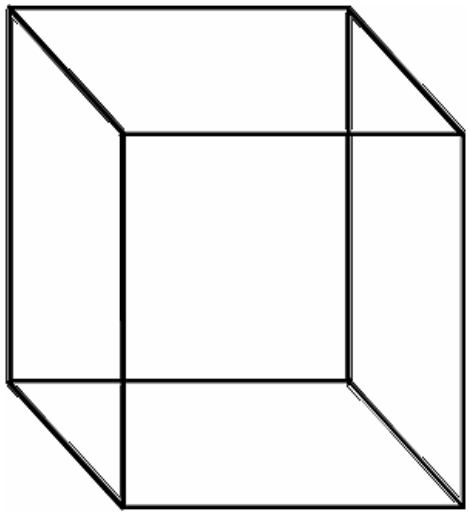
- Two stabla patterns in finger wagging experiments, relative phase 0 and ½.
- The proposed nonlinear model has two attractors at relative phases 0 and ½.
- Nature of stable points dependent on rate
- Model predictions
  - At higher rates, the finger wagging will settle to the 0 relative phase pattern.
  - only two stable patterns.
  - Critical slowing down
  - Critical fluctuations
  - Similar effects in speech coupling

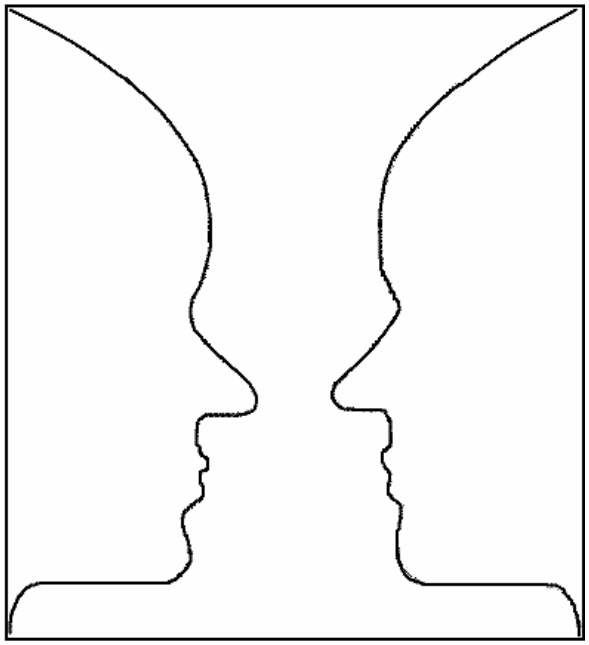


Arguments from Phonology

- Look for periodicity in behaviour and patterns.
- Speech typically shows cyclic repetition of similar events and a task that has been used to study speech timing is the “speech cycling” task.
- Subjects were asked to say a phrase (Give the dog a bone) and were asked to time it so that they begin the phrase with a beep from a metronome. The rate of beeps in the metronome was increased and the phase angle of the onset of a particular word (bone) with the first word (Give) of the phrase and the first word of the next phrase was measured. It was found that there was tendency for the phase onset of the syllable of interest to be around 1/2.
- Other experiments manipulating speech timing and results were found that were similar to those on motor behaviour (Kelso, 1995). These results show the influence of timing in speech perception and production arguing for the necessity of dynamic approaches to speech

Ambiguous Figures - Reversals





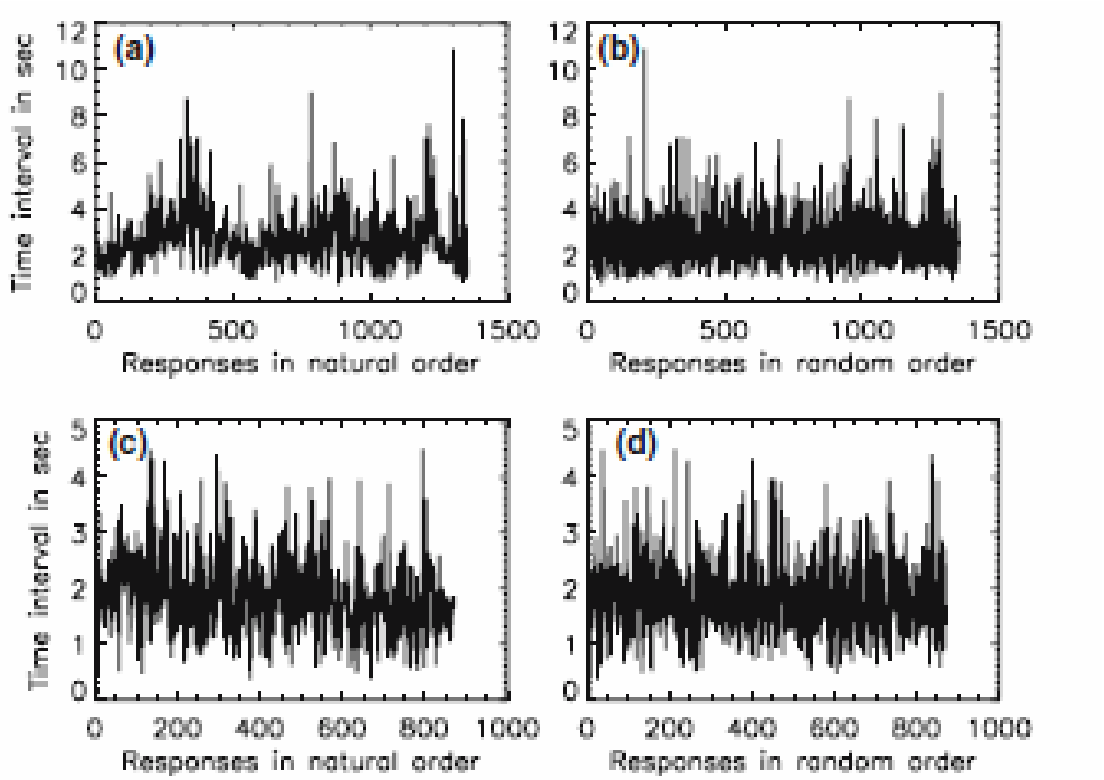
Multistability

- Multistability is a phenomenon in visual perception that occurs when your percept varies from one state to another though the stimulus in front of you is non-variant
- Multistablity in the percept formation – underlying mechanism Perceptual organization
- Philosophically – to do with consciousness.

Nonlinear dynamics of bistability

- Reversal rate histogram for Necker cube - RFR depicts a gamma distribution (De Marco et al. 1977)
- Switching related gamma band synchrony between parietal and frontal areas with alpha band activity in occipital (Nakatani, 2006)
- Switching times behave as a 1/f noise and possess very long range correlation (Gao et al. 2006)

Switching times



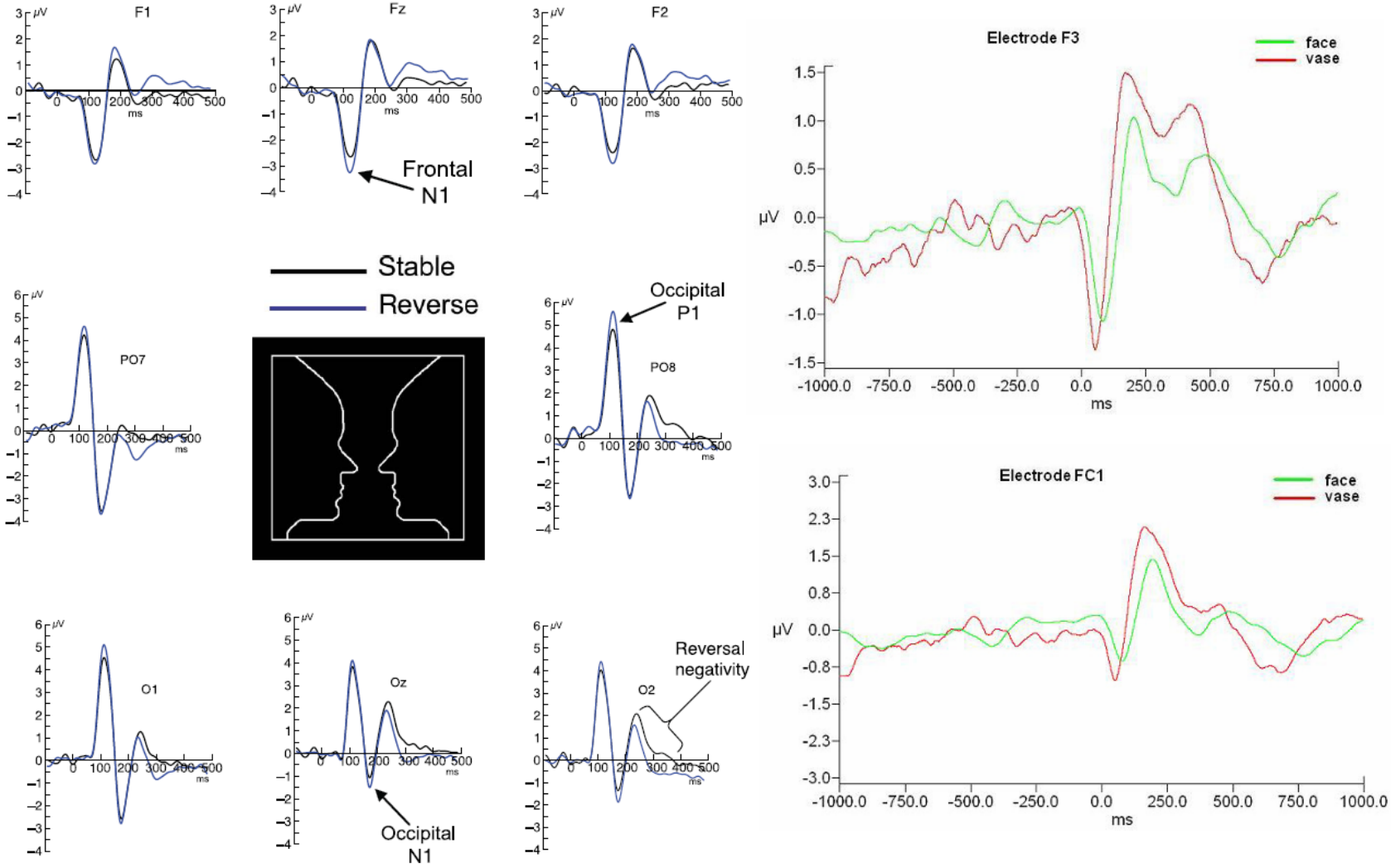
Hurst Parameter Results

Table 1 The Hurst parameter for natural time series from ten Necker cube and ten binocular rivalry subjects

Necker	Rivalry
0.63	0.59*
0.64	0.62
0.64	0.63
0.69	0.66
0.72	0.66
0.72	0.67
0.75	0.73
0.76	0.77
0.84	0.77
0.84	0.78

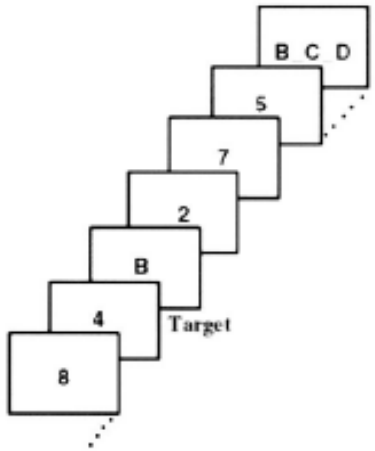
For 19 of these 20, the estimated likelihood was less than 0.001 (zero occurrences in 1,000 or more shuffles) that H in their shuffled time series was at least as large as H in their naturally ordered time series. For the 20th subject (indicated by *), whose $H_{natural}$ was the smallest found, the estimated likelihood was 0.003 (30 such occurrences out of 9,000 shuffles)

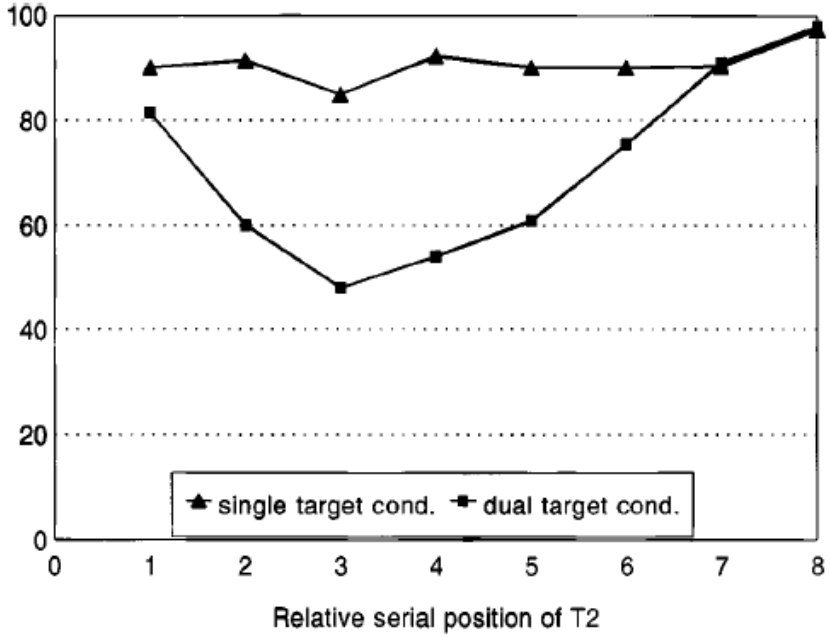
Multistability – ERP Results



Attentional Blink

- Targets are presented one at a time very briefly.
- Typically the presentation of the target as well as the blank interval is of a duration around 100 ms.
- Participants have to detect two targets (T1 and T2) and the rest of the stimuli are distractors.
- Target T1 appears first followed by target T2 and the temporal gap (lag) between T1 and T2 are varied.
- The basic finding is that accurate identification of target T2 is poor for lag 2.
- The performance improves with higher lag and reaches asymptote around lag 6 or 7.





AB Results

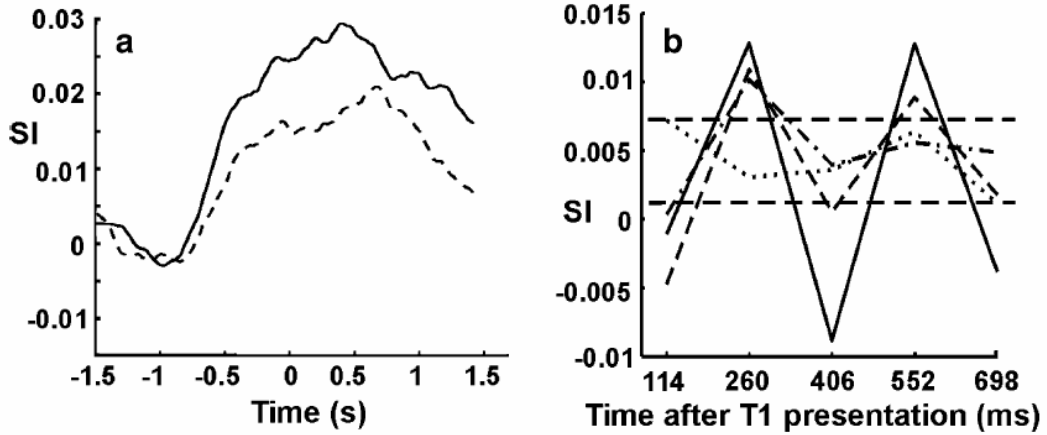


Fig. 4 Synchronization (SI) in the target-related network. **a** The "noAB" condition (*solid line*) shows a stronger SI during stimulus presentation compared to the "AB" condition (*dashed line*). 0 ms corresponds to the onset of the first target. The beginning of the letter stream ranges from −880 to −580 ms. The SI time courses were smoothed with a Savitzky-Golay filter (polynomial order: 3, frame length: 600). **b** SI for the components of five successive stimuli. Zero on the *x*-axis corresponds to a 60 ms long window centered at 260 ms after presentation of the first target. The other windows were shifted by multiples of the SOA (146 ms). Position 2 corresponds to the target component of the second target (for AB and noAB condition). For each window the mean SI is shown. Conditions are pattern-coded (noAB: *solid*; AB: *dashed*; target: *dash-dot*; distractor: *dotted*). The dashed horizontal lines mark the extent of SI in trials containing only distractors

AB Results

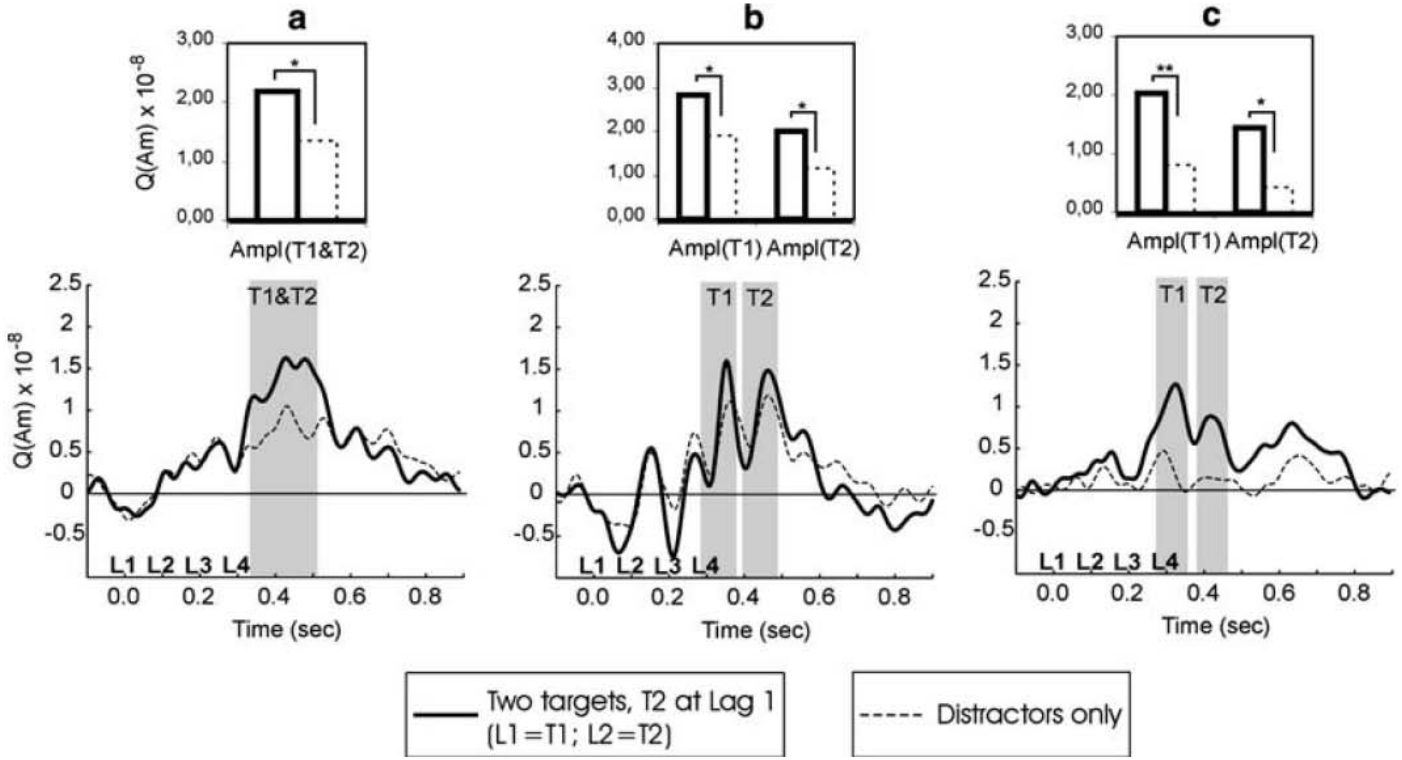
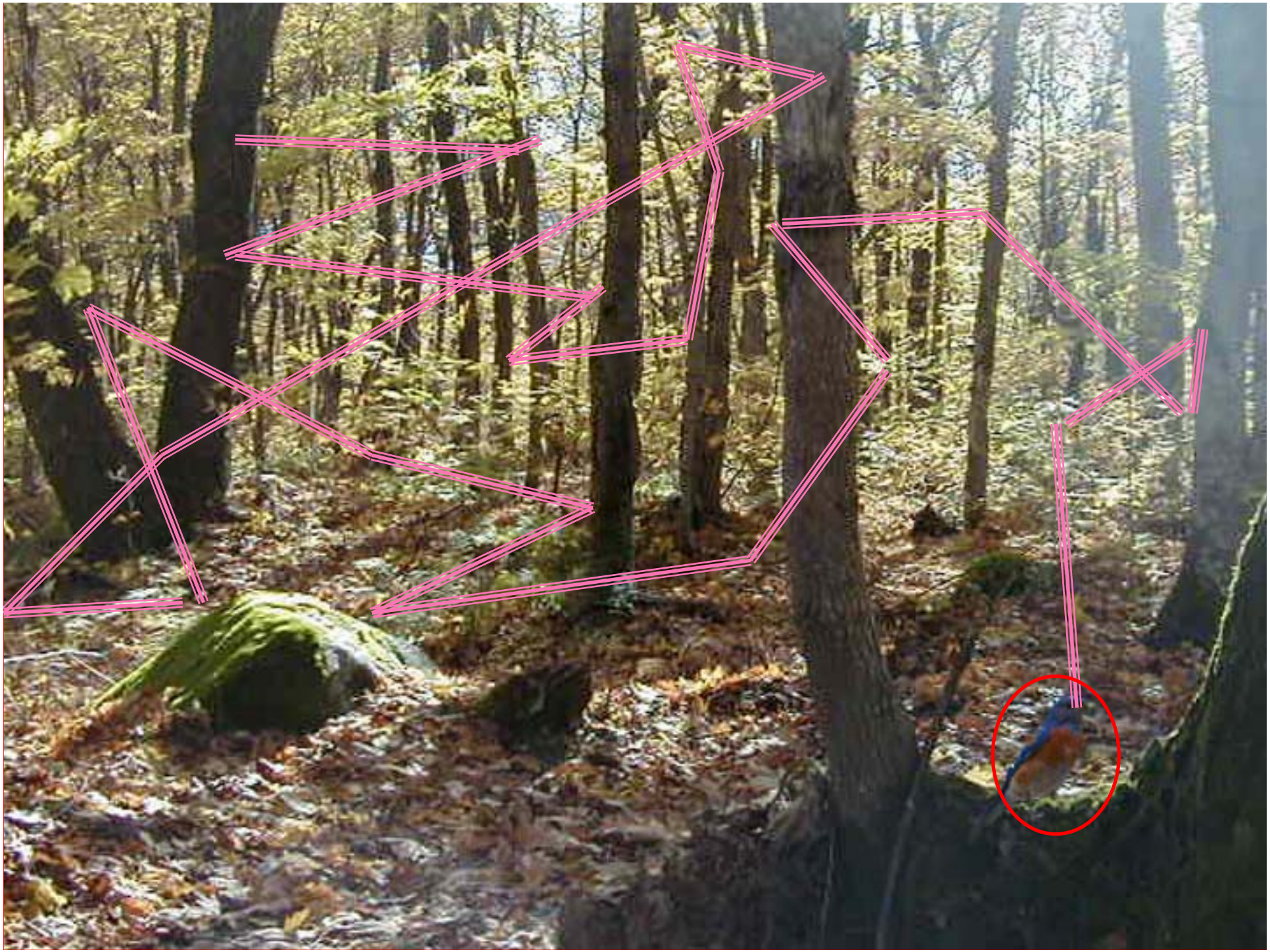


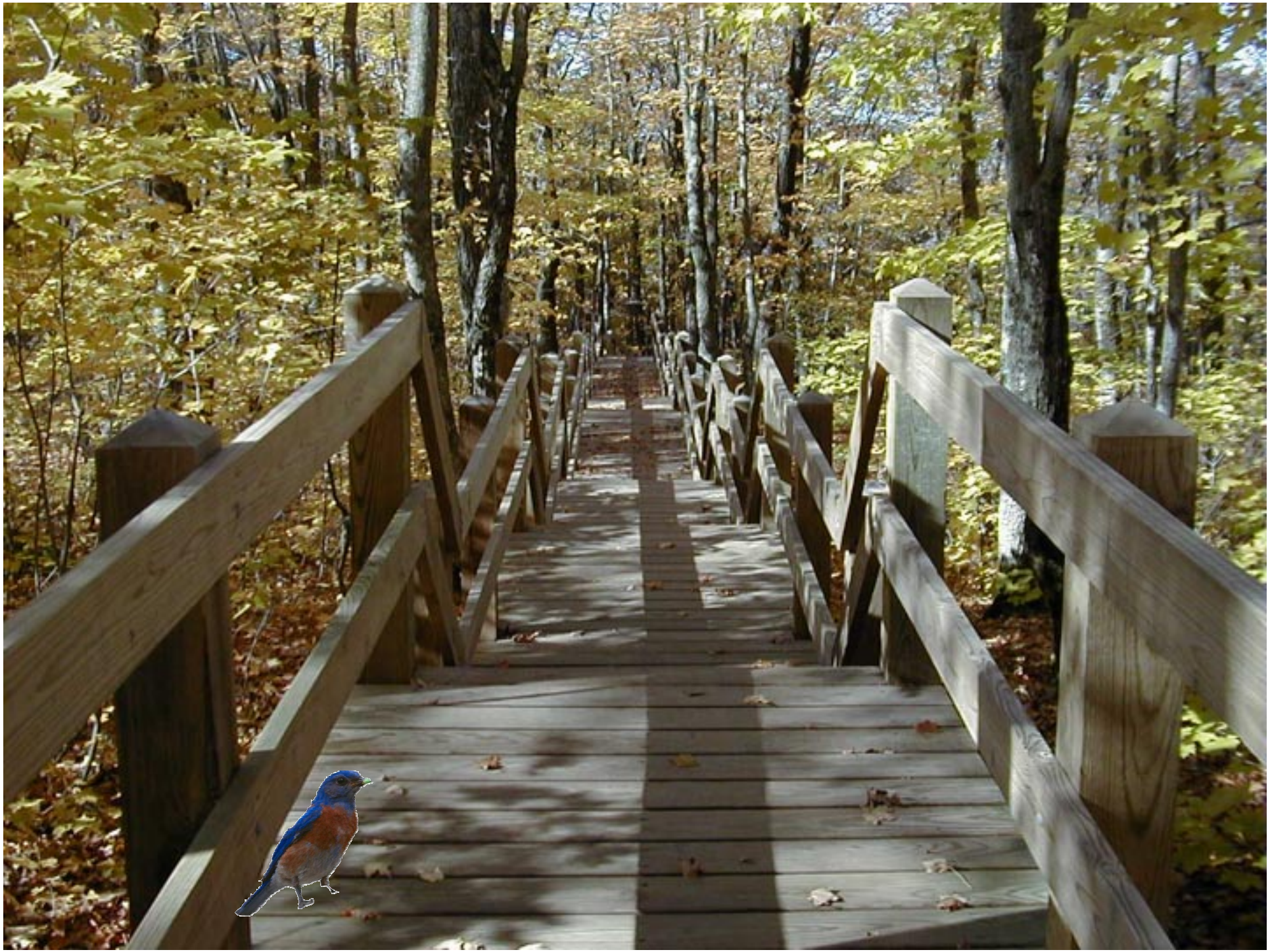
Fig. 5 Source waveforms from left temporo-parietal (**a**), right-temporo-parietal (**b**) and frontal (**c**) areas for Lag1 trials (T2 immediately following T1; T1–T2 SOA = 100 ms, ISI = 50 ms), adapted from Kessler et al. (2005b). *Letter onsets* are indicated by "L1" to "L4" on the *x*-axis. At the top of each panel a graph compares amplitude means in the Lag 1 and in the distractor condition. *Asterisks* denote the 5% (*single*) and the 1% (*double*) significance levels. In left temporo-parietal areas (**a**) only one M300 component for T1 and T2 is observed that might reflect a single or two overlapping target-related processes. In right temporo-parietal areas (**b**) two target-related M300 components (*grey bars*) are observed on top of regular biphasic responses that mirror the occipital pattern (cf. Fig. 2). In PFC (**c**) two distinct target-related M300 components are the dominant waveform patterns

Visual Search (Deborah Acks)

Find

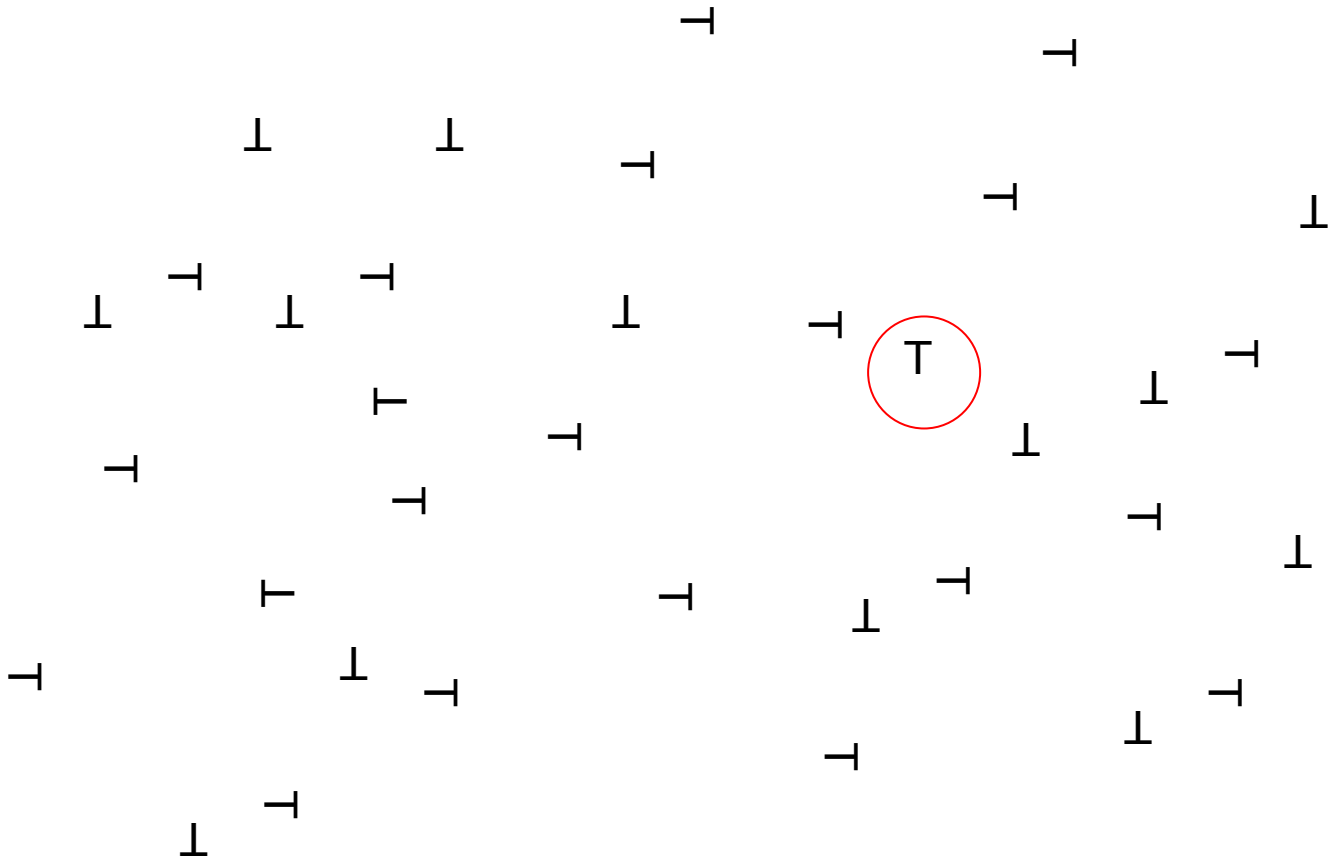
:





Visual Search Task

Find the upright “T”



Method

- Each trial contained 81 Ts.
- 400 trials lasting 2.5 hours.
- Eight 20-minute sessions separated by 5-minute rest
- Generation V dual purkinje-image (DPI) tracker

Map trajectory of eyes:

- Duration & x,y coordinates for each fixation.

  - Differences between fixations
    - $x_n - x_{n+1}$ & $y_n - y_{n+1}$
  - Distance = $(x^2 + y^2)^{1/2}$
  - Direction = Arctan (y/x).

Results

Conventional search stats…

What's the central tendency?

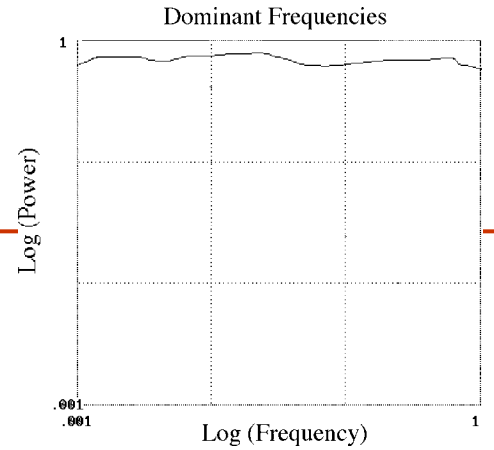
- 24 fixations per trial (on average)
- 7.6 seconds (SD =6.9 sec) per trial
- Mean fixation duration = 212 ms (SD = 89 ms)

Focusing on the dynamic…

- 10,215 fixations across complete search experiment.

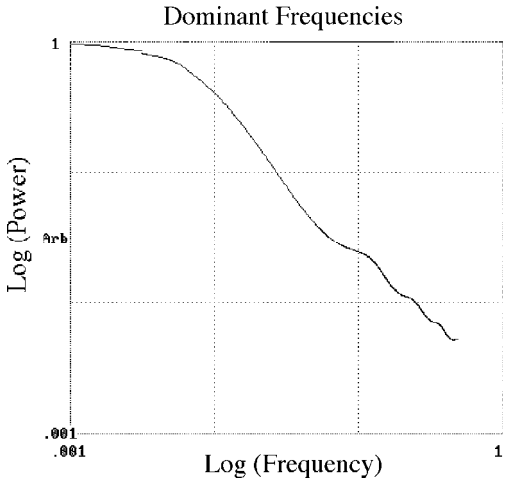
"Color' of noise

1/f ^{0} noise -- flat spectrum= no correlation across data points
Short & Long range = 0



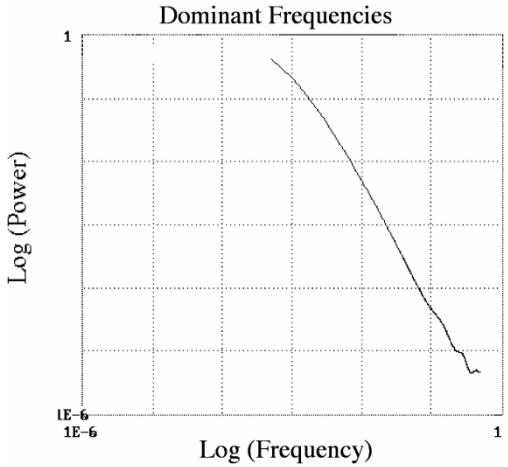
White Noise

1/f noise --shallow slope = subtle long range correlation



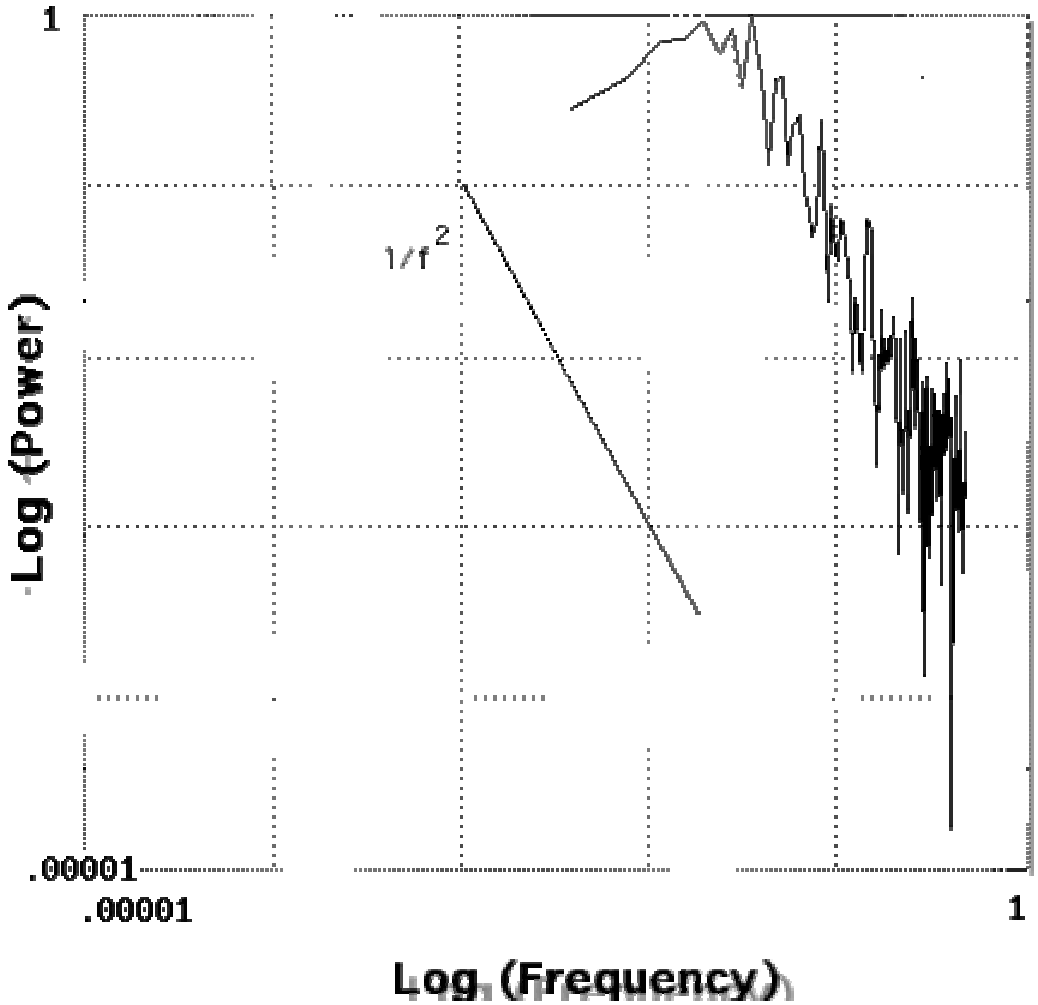
Pink Noise

1/f ^{2} noise-- steep slope = Predictable long-range, 'undulating' correlation
Short range = 0 (successive events uncorrelated)



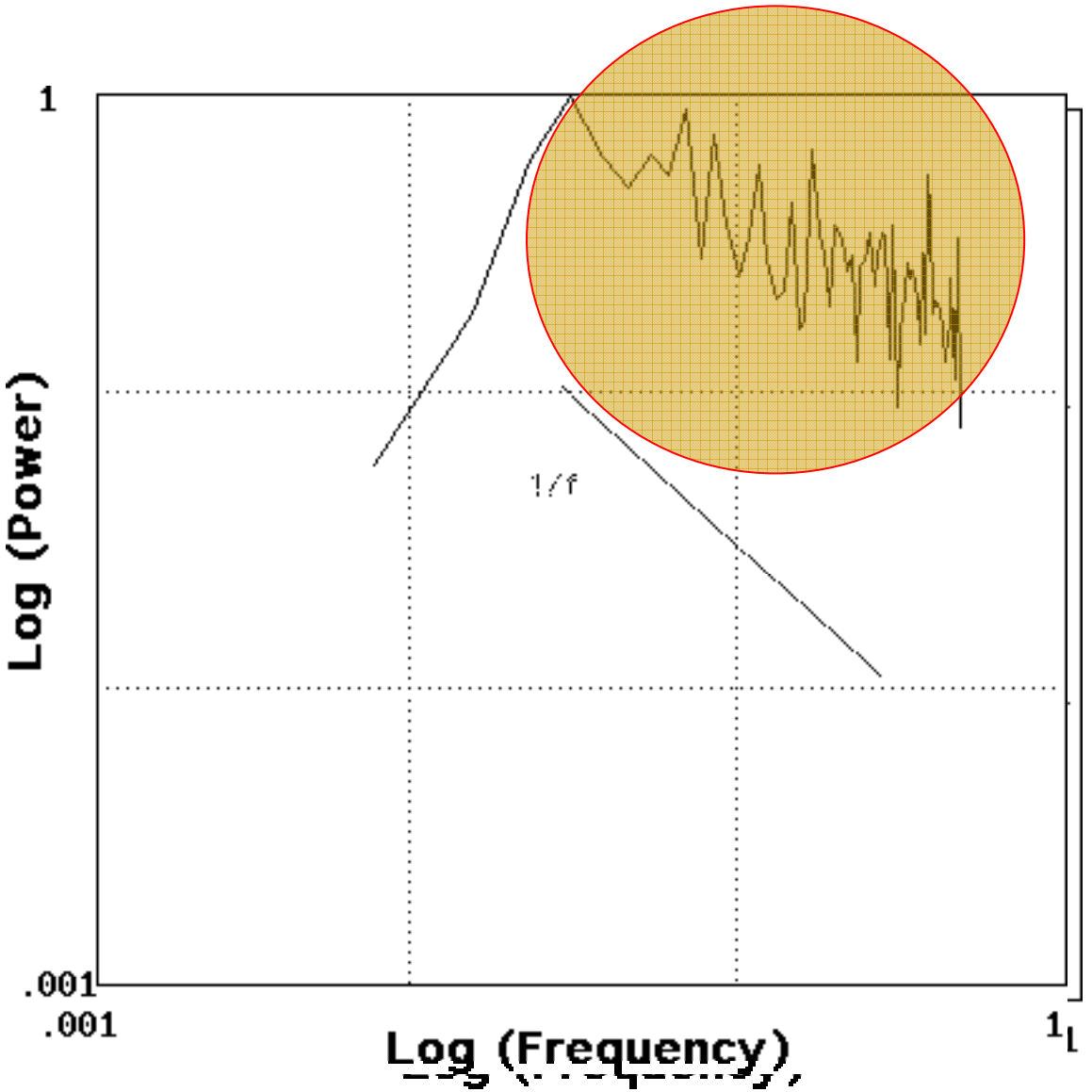
Brown Noise

Power Spectra of raw fixations



$\alpha = -2$

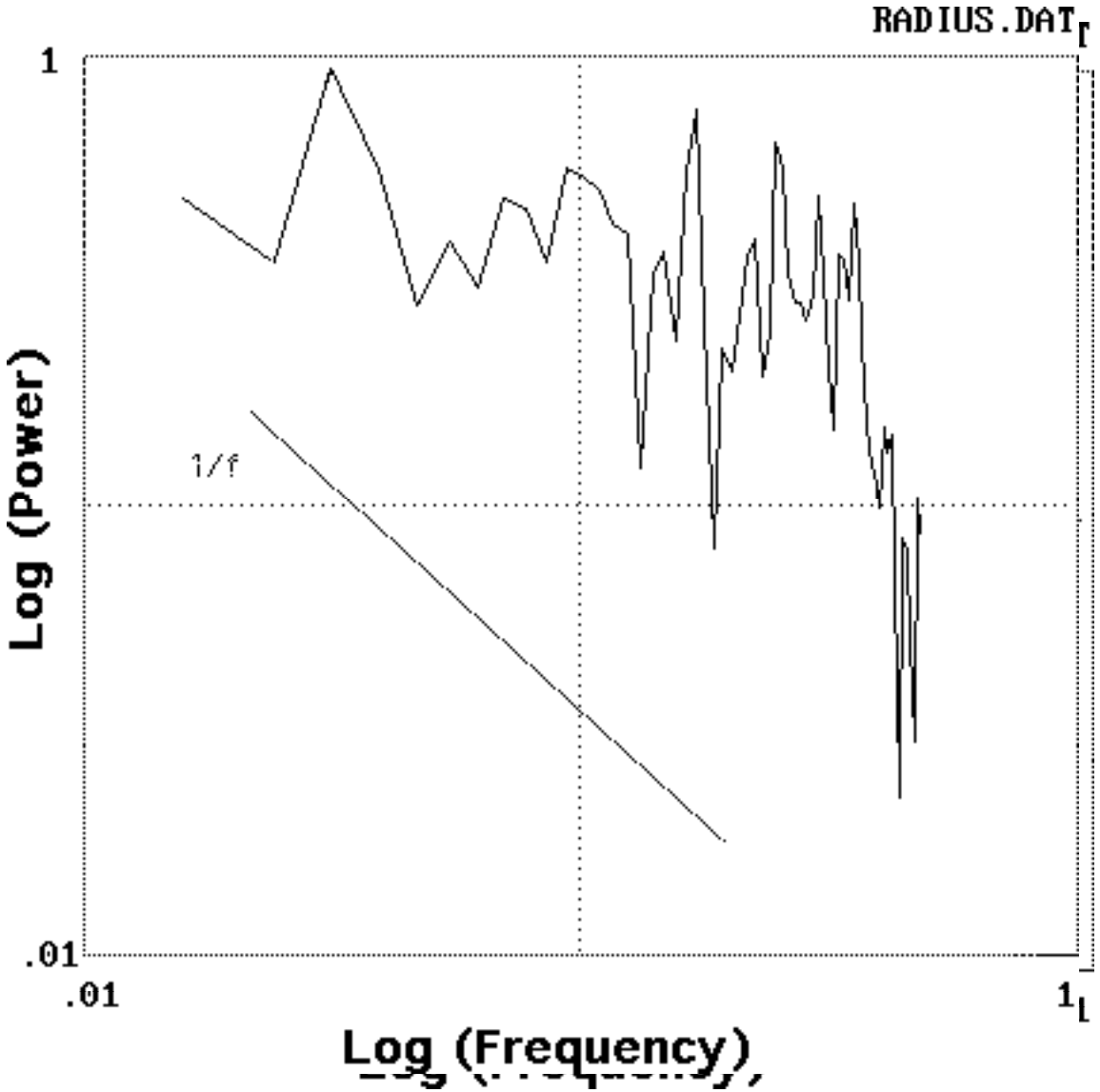
Power Spectra of first differences across fixations



$\alpha = -.6$

Distance across eye fixations

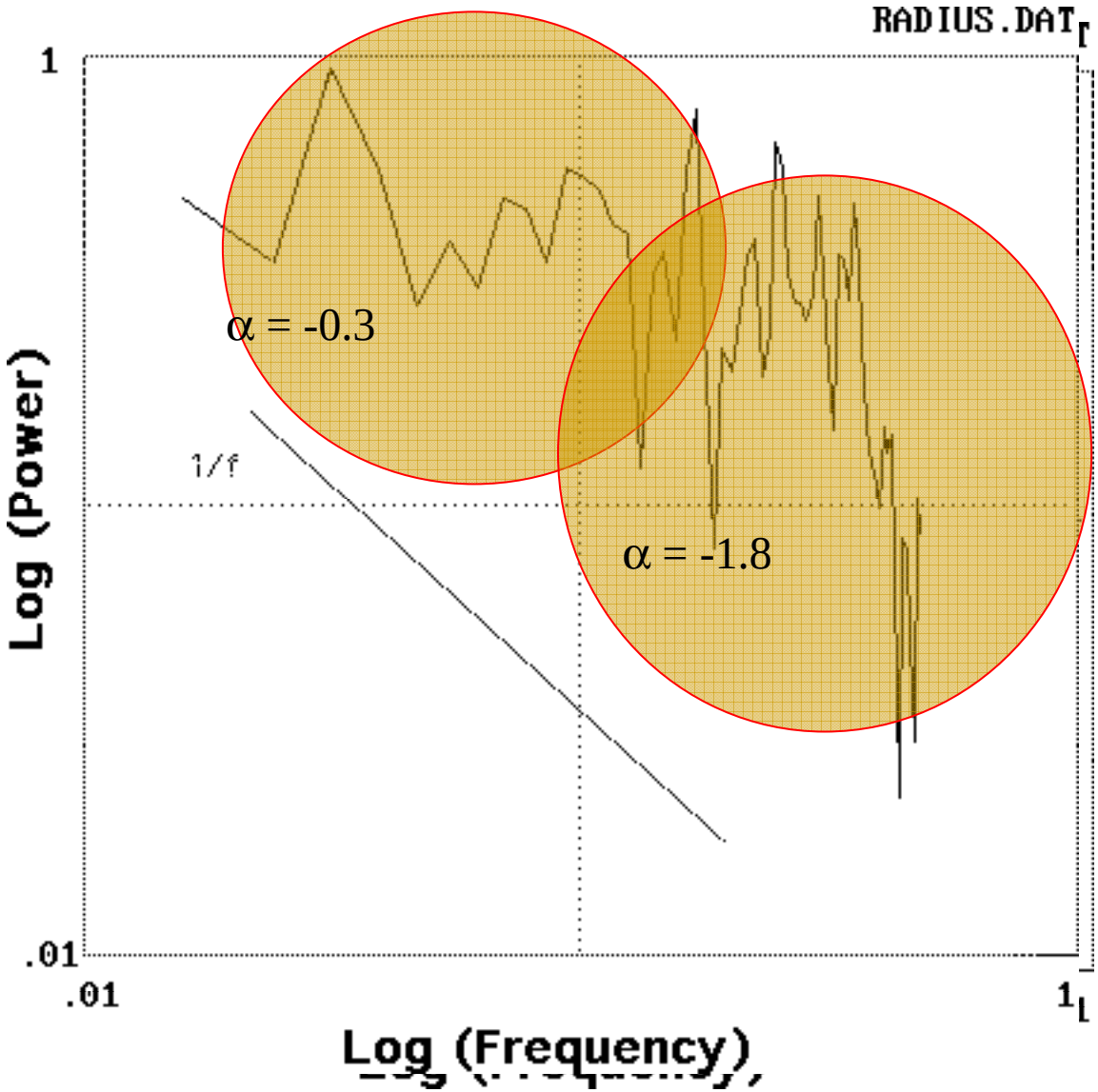
$(x^2 + y^2)^{1/2}$



α = -.47

Distance across eye fixations

$$(x^2 + y^2)^{1/2}$$



α = -.47

Summary of results:

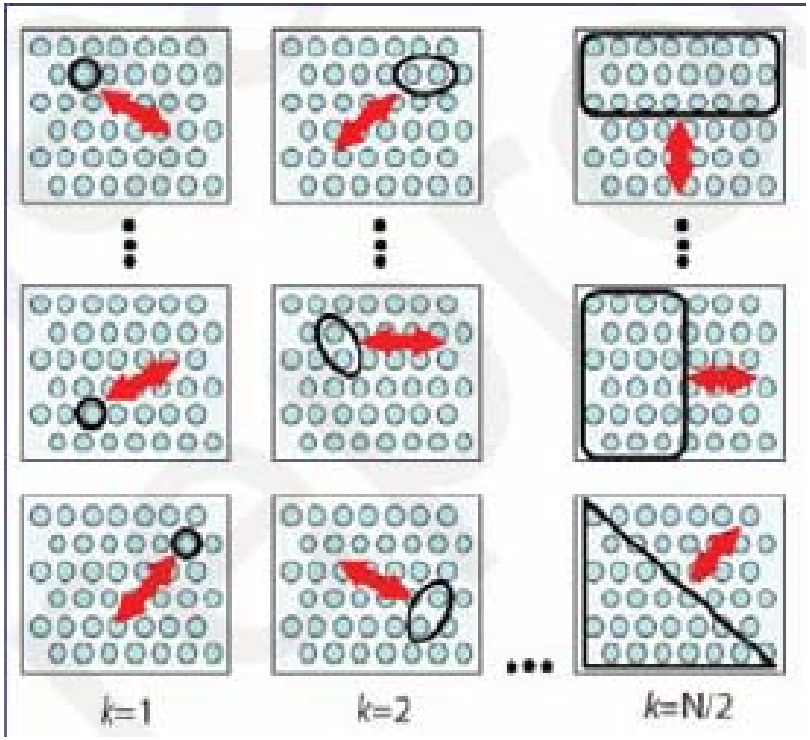
- Sequence of...
  - Absolute eye positions --> 1/f brown noise
  - local random walk
- Differences & distance-across-fixations --> ~1/f pink noise
  - Subtle long-term memory.

From perceptual to brain dynamics

- Are the spontaneous changes experienced in perceptual patterns the outcome of similar events in brain activity?
- what kind of dynamics governs these changes in the brain?
- Brain dynamics as a necessary requirement for the dynamics of our mental states

Consciousness and Complexity

- Tononi, Edelman, Anil Seth
- Simultaneous Differentiation and Integration
- “Small parts of a system are independent, large parts are comparatively integrated”



$$C_N(\mathrm{X}) = \sum_k \langle MI(\mathrm{X}_j^k; \mathrm{X} - \mathrm{X}_j^k) \rangle,$$

Coherence Intervals

- Subsystems of the brain will remain in quasi-stable phase synchrony for as long as it takes to pass information between them. Periods with this function are called coherence intervals (van Leeuwen & Baaker, 1995).
- Coherence intervals as “filter” for information

Thanks

Acknowledgments

- This presentation has been prepared with material from many other presentations from the Internet. I thank all those for making their work available for the transfer of knowledge.