Cognitive semantics and cognitive theories of representation:

Session 7: Meaning and brain

Martin Takáč
Centre for cognitive science
DAI FMFI Comenius University in Bratislava

Príprava štúdia matematiky a informatiky na FMFI UK v anglickom jazyku
ITMS: 26140230008

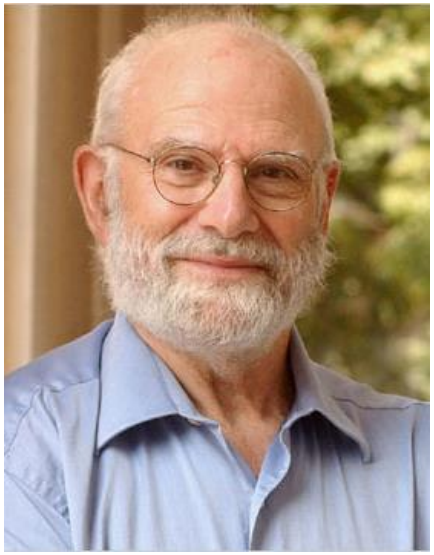
Main sources of knowledge

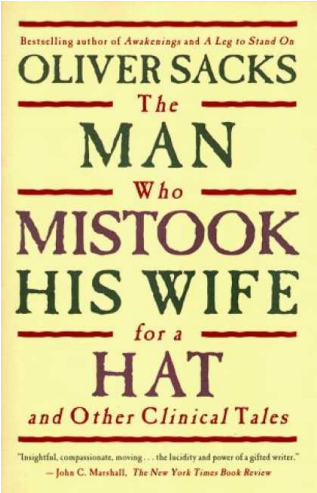
- Lesion studies (tracing back to A. Luria)

- Neuroimaging methods (EEG, ERP, CT, PET, fMRI)
- TMS

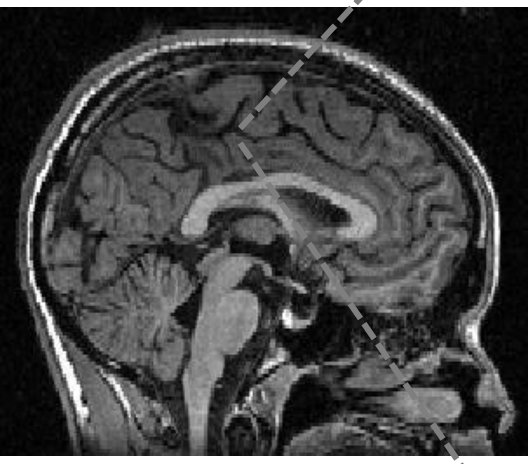
Popular science

- Oliver Sacks

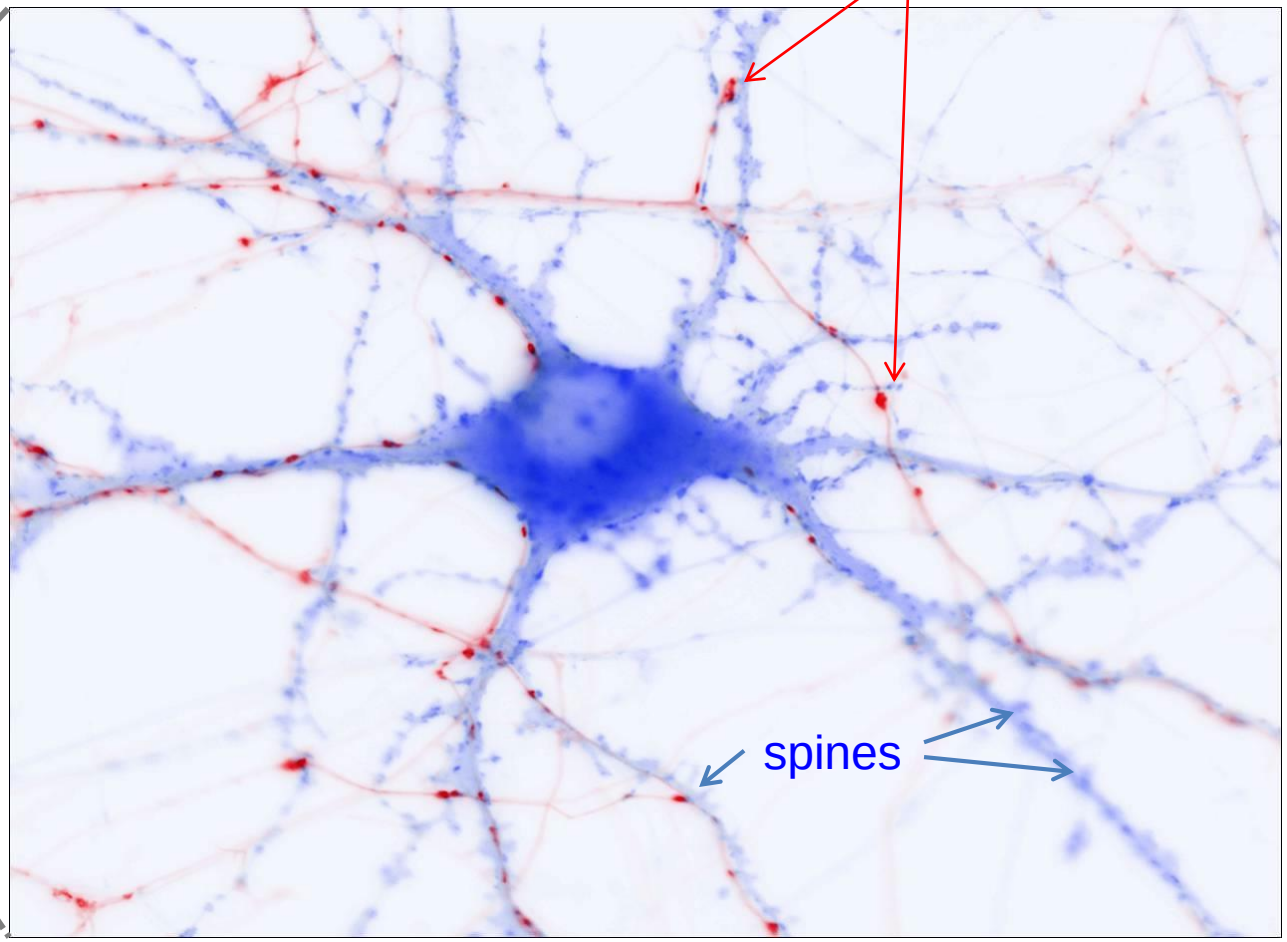




Brain is comprised of networks of neurons connected and communicating via synapses



10^{11} neurons



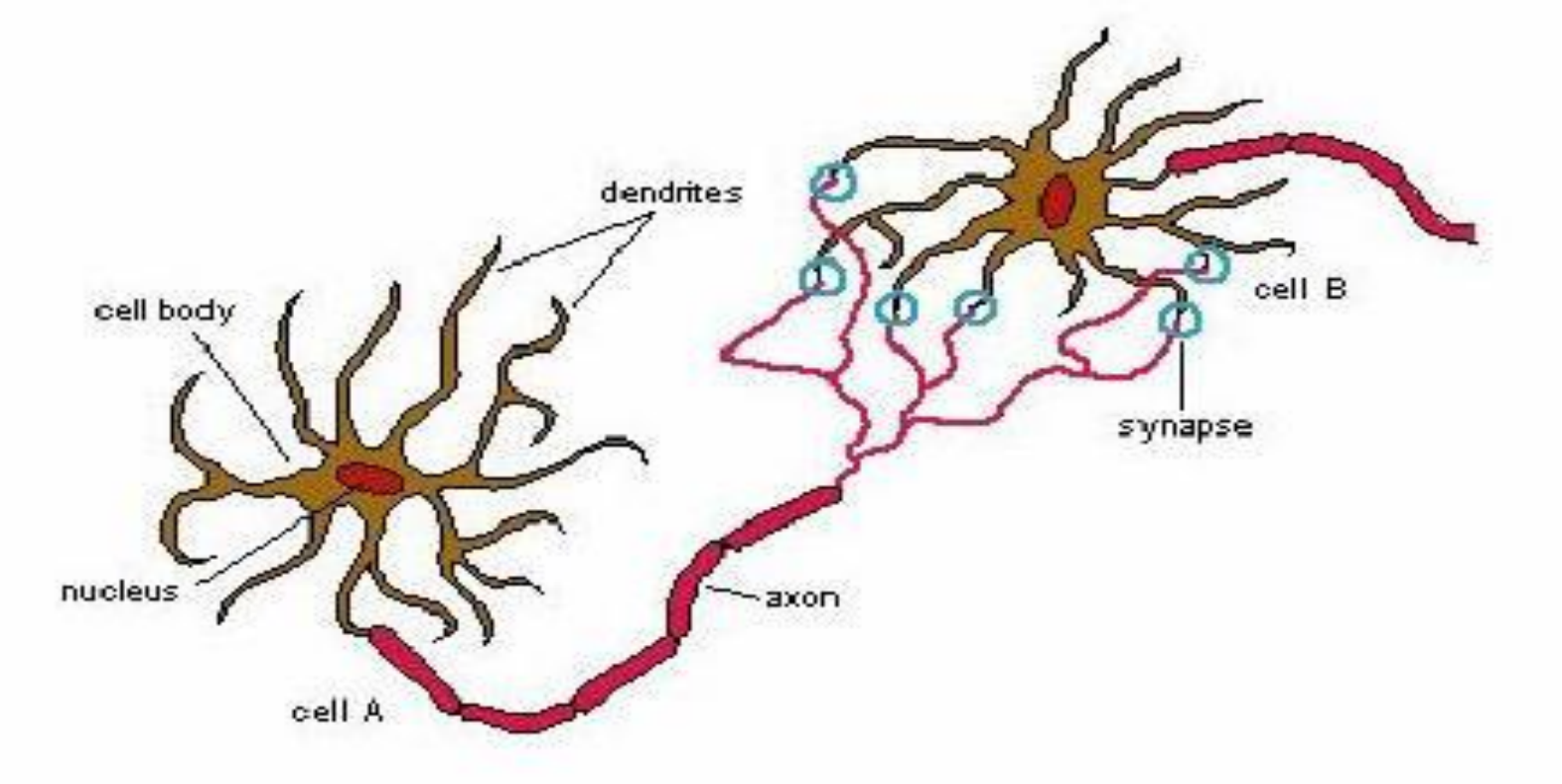
10^{4} synapses in and out

Learning and brain plasticity

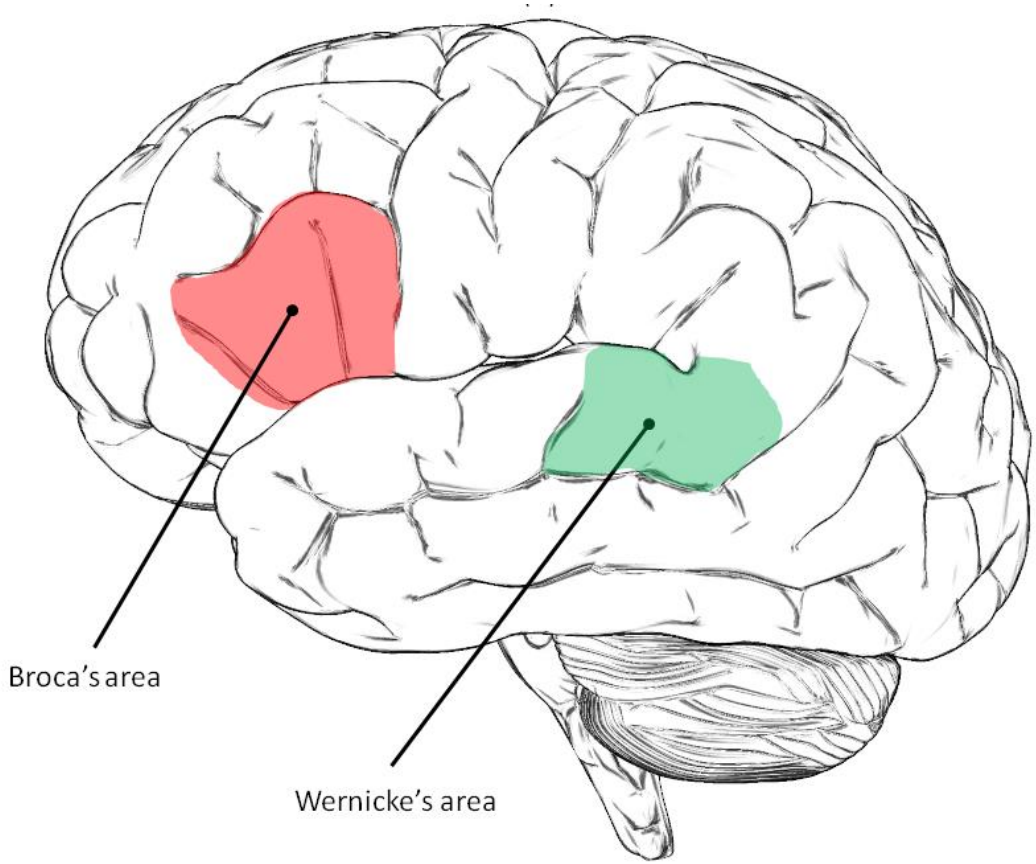
- Traditionally: learning is an acquisition of memories.
- Memory is an organism's ability to store, retain, and subsequently recall information.
- Neural correlate of learning is **brain plasticity**.
- Brain plasticity (neuroplasticity) is a lifelong ability of the brain to reorganize connectivity of neural circuits based on new experience.
- Brain plasticity is based on **synaptic plasticity**.

Hebb's rule of synaptic plasticity (1949):

- *When an axon of cell A is near enough to excite a cell B and* ***repeatedly or persistently*** *takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased."*
- ***"Fire together – wire together"***



Traditionally language related areas



Neuropsychologist theories of cortical functioning

- Localizationist
- Holistic
- Hebbian
  1. Coactivated neurons become associated.
  2. Associations can occur between adjacent or distant neurons; that is, the entire cortex is an associative memory.
  3. If neurons become associated, they will develop into a functional unit, a **cell assembly**.

(Pulvermüller, BBS 1999)

Associations

- LTP
- But also LTD

Table 1. *Associative synaptic learning according to a Hebbian coincidence rule*

		neuron L	
		active	inactive
neuron M	active	+w*	--
	inactive	--	--

*+w indicates an increase in connection strength between neurons L and M; hyphens indicate no change in connection strength.

Table 2. *Associative synaptic learning according to a correlation rule*

		neuron L	
		active	inactive
neuron M	active	$+w_1$*	$-w_2$
	inactive	$-w_3$	--

*$+w_1$, $-w_2$, and $-w_3$ indicate positive or negative changes in connection strength.

Cell assemblies

- If neurons in an associative network exhibit correlated activity, they will be a stronger influence on each other. This implies that these neurons will be more likely to act together as a group. = **CELL ASSEMBLY**
- The strong within-assembly connections are likely to have two important functional consequences:
  1. If a sufficiently large number of the assembly neurons are stimulated by external input (either through sensory fibers or through cortico-cortical fibers), activity will spread to additional assembly members and, finally, the entire assembly will be active. = **IGNITION**
  2. After an assembly has ignited, activity will not stop immediately (because of fatigue or regulation processes), but the strong connections within the assembly will allow activity for some time. = **REVERBERATION**, spatiotemporal pattern of high frequency (>20 Hz).

- On the cognitive level, **ignition** may correspond to **perception of a meaningful stimulus** and to **activation of its representation**.
- Sustained activity of the assembly and **reverberation** of activity therein may represent an elementary process underlying **short-term or active memory**

Updated Hebbian paradigm

1. Simultaneous pre- and postsynaptic activity of cortical neurons leads to synaptic strengthening. However, pre- or postsynaptic activity alone leads to synaptic weakening.
2. Associations can occur between adjacent neurons and between cortical neurons located far apart, provided there is a synapse connecting them. The cortex is an associative memory although it is not fully connected.
3. If synaptic strengthening occurs among many neurons, they will develop into an assembly that can ignite and exhibit well-timed reverberatory activity.

Potential problems

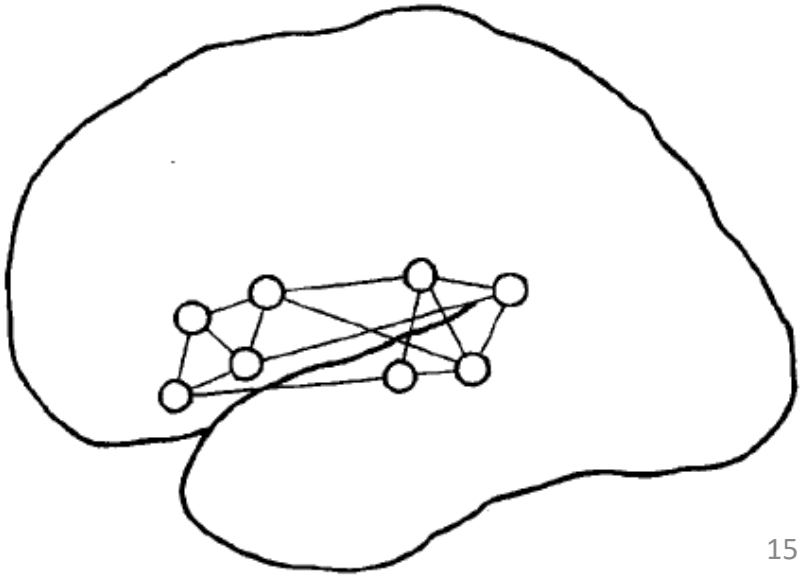
- Activity spreading to the entire cortex
  - it is necessary to have a control device regulating the cortical equilibrium of activity – a “threshold control mechanism”; its neuroanatomical substrate has been proposed to be located in the basal ganglia or in the hippocampus.
- Clamping together all assemblies
  - LTD in correlation-based learning prevents this
- How to represent complex objects (e.g. house with doors and windows)
  - Hierarchy (subsets) of assemblies and global activation threshold

Cortical distribution of cell assemblies

- The cortical localization of a representation is a function of where in the cortex simultaneous activity occurred when the representation was acquired or learned.
- Attention to stimuli necessary

Assemblies representing word forms

- neuronal activity can be assumed to be present almost simultaneously in defined primary and higher-order **motor** and sensory (**auditory** and **somatosensory**) **cortices**. All in **perisylvian cortex**



Lateralization

- Content vs. function words
- Abstract content words somewhere in between

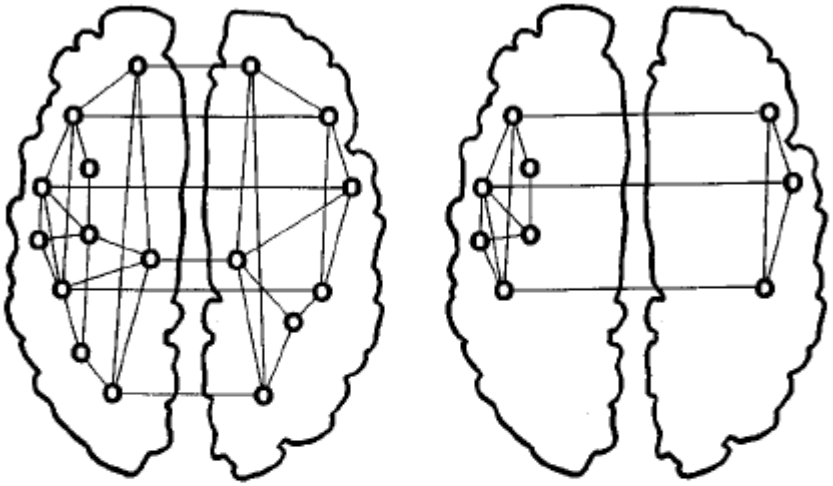


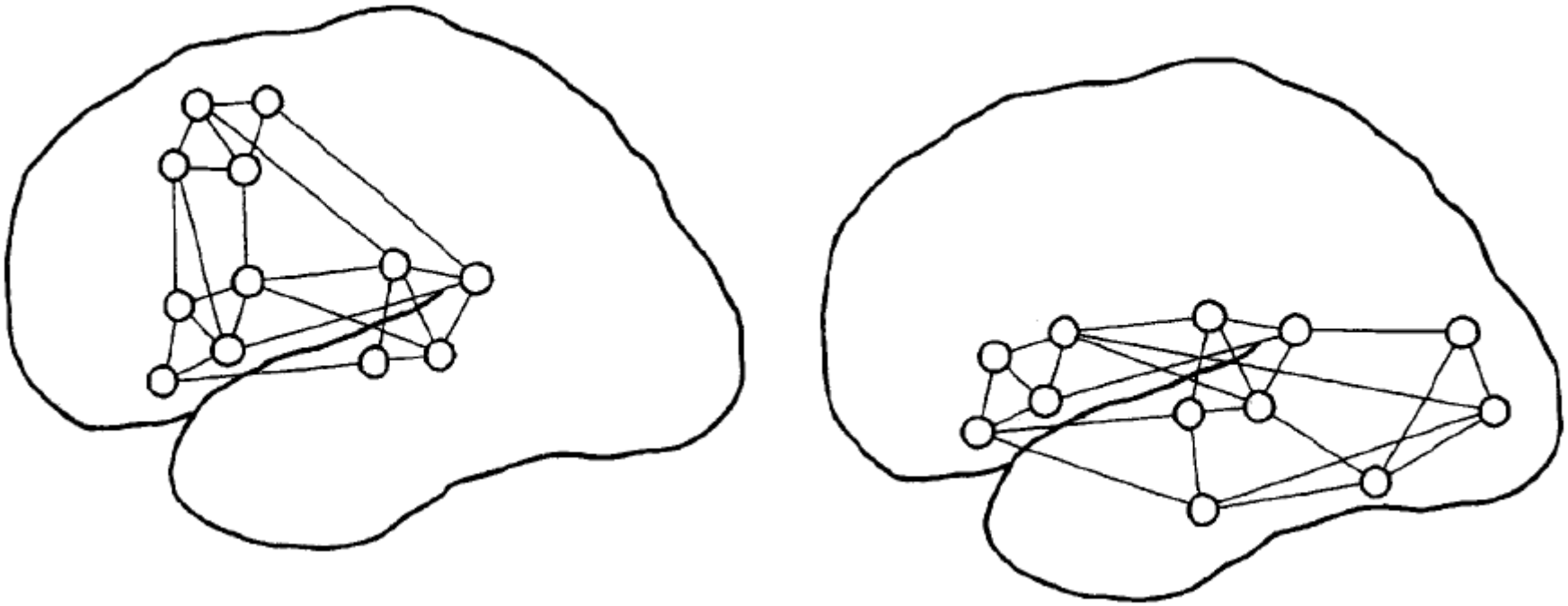
Figure 4. Cell assemblies relevant for cognitive processing may be distributed over both hemispheres and may be lateralized to different degrees. Whereas for cell assemblies representing phonological word forms and grammatical function words a high degree of laterality appears likely (right), an assembly representing a concrete content word may exhibit a reduced degree of laterality (left). (Adopted from Pulvermüller & Mohr 1996.)

Emotion words

- Visual stimuli (faces)
- Patterns of activity of facial muscles and autonomous nervous system
- Additional links to subcortical neurons in limbic system

Action vs. vision words

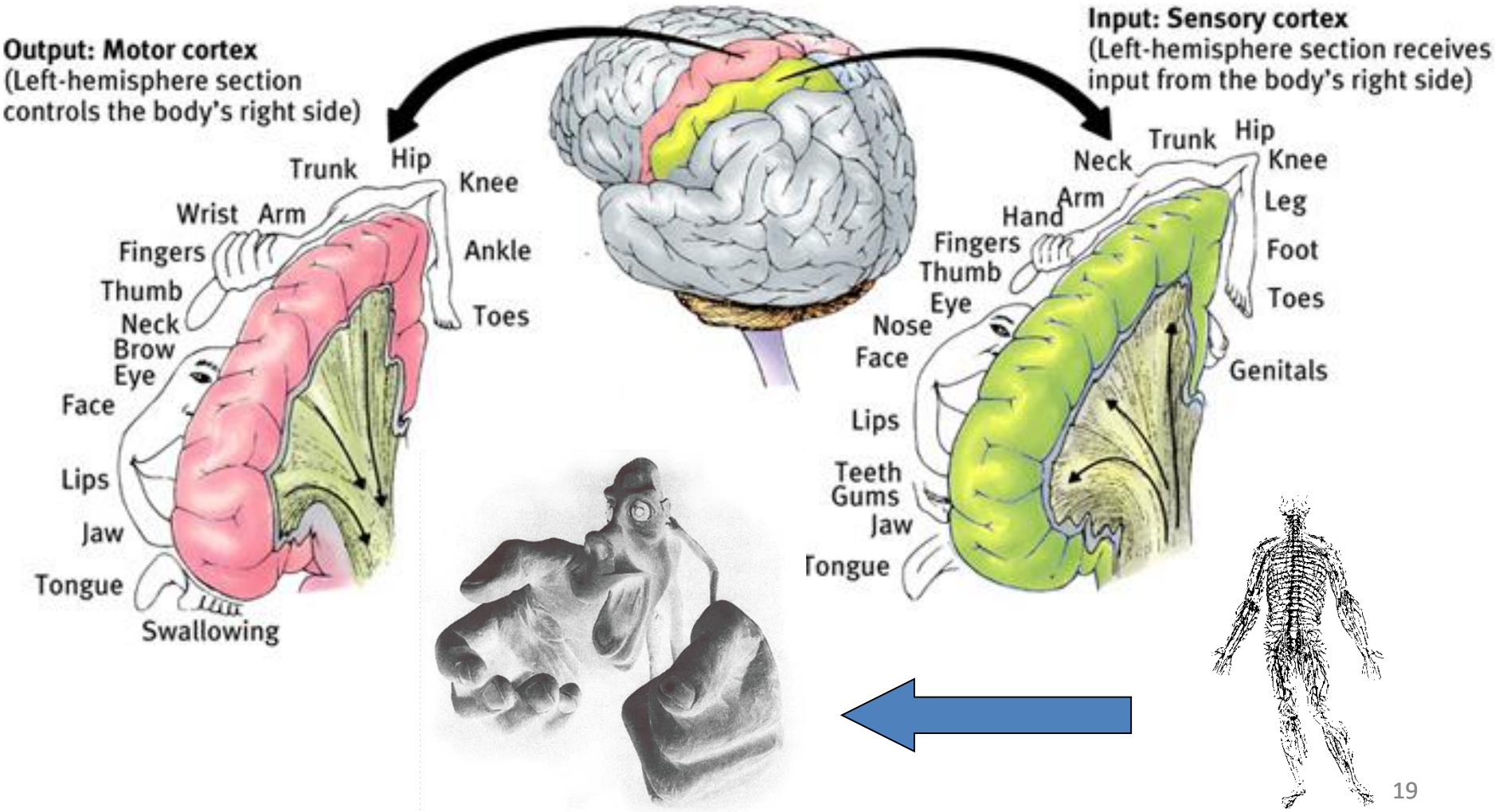
- Action (left), vision (right) cell assemblies



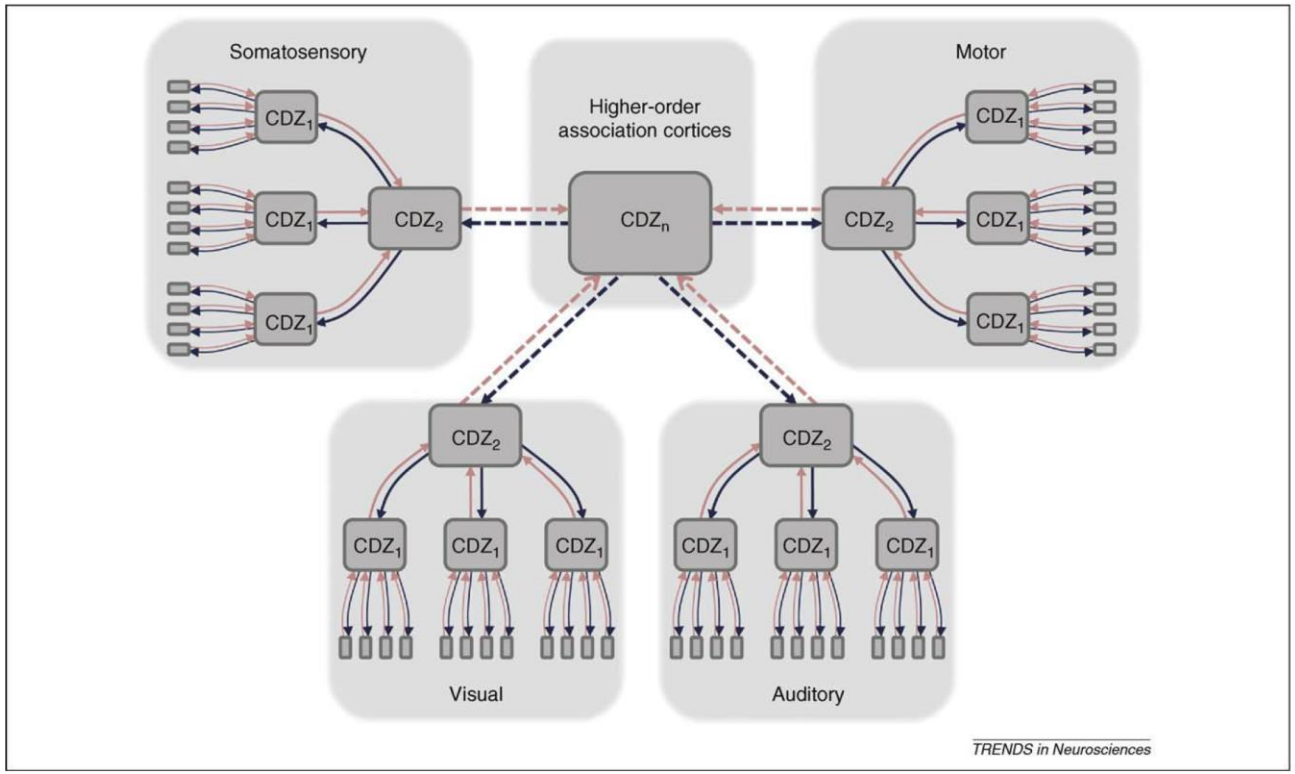
Localist / sparse representations

Somatosensory and motor systems in animals

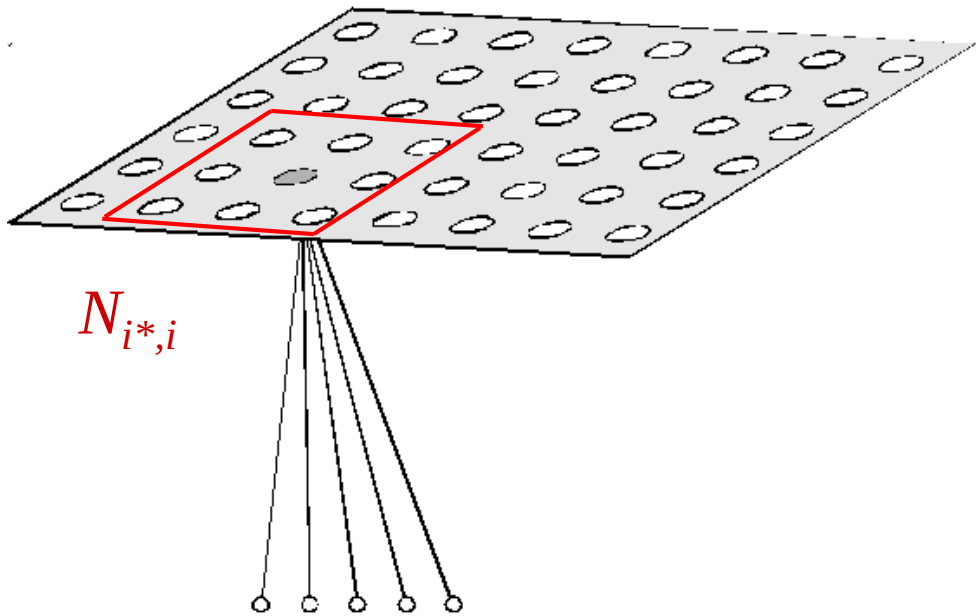
- topological mapping from body to cortex



Convergence zones (Damasio)



Self-organizing maps – Unsupervised learning



The weights of the winner and its neighbours in $N_{i*,i}$ move closer to the input:

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + \alpha(t) \cdot N_{i*,i}(t) \cdot [\mathbf{x}(t) - \mathbf{w}_i(t)]$$

So is there any “grand-mother” cell in the brain?

(Quiroga et al, Nature, 2005)

nature Vol 435|23 June 2005|doi:10.1038/nature03687

LETTERS

Invariant visual representation by single neurons in the human brain

R. Quian Quiroga[1,2]†, L. Reddy[1], G. Kreiman[3], C. Koch[1] & I. Fried[2,4]

- Single-cell recordings from 8 patients with pharmacologically intractable epilepsy
- Responses of neurons (993 units - 343 single units and 650 multi-units) from the hippocampus, amygdala, entorhinal cortex and parahippocampal gyrus to images shown on a laptop computer in 21 recording sessions.
- Stimuli: were different pictures of individuals, animals, objects and landmark buildings presented for 1 s in pseudorandom order, six times each.

nature Vol 435|23 June 2005|doi:10.1038/nature03687

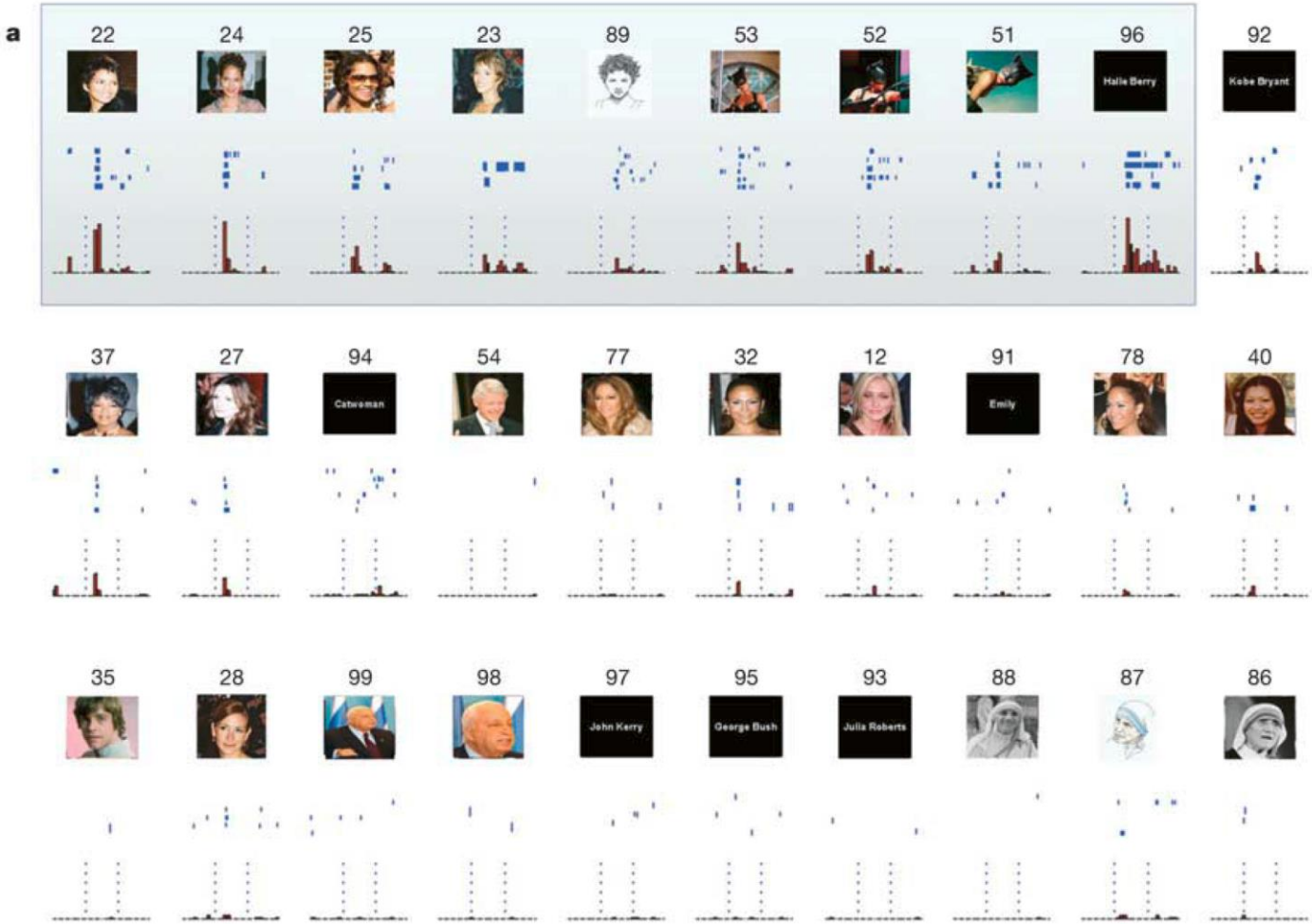
LETTERS

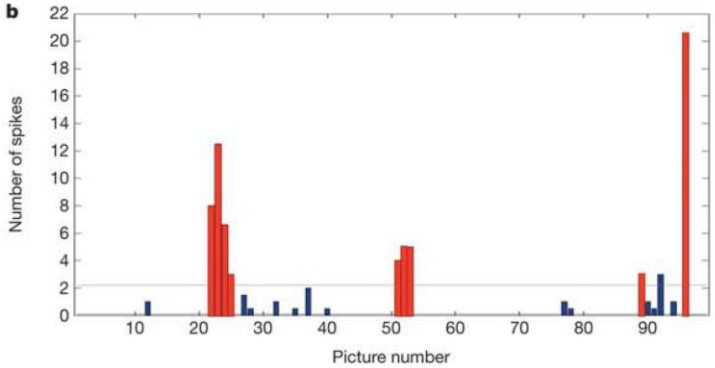
Invariant visual representation by single neurons in the human brain

R. Quian Quiroga[1,2]†, L. Reddy[1], G. Kreiman[3], C. Koch[1] & I. Fried[2,4]

- Neurons responding to Bill Clinton, The Simpsons, Michael Jordan
- Highly selective responses (to 2.8% of images on avg)
- Suggest a sparse encoding rather than fully distributed
- Resembles hippocampal place cells in rodents

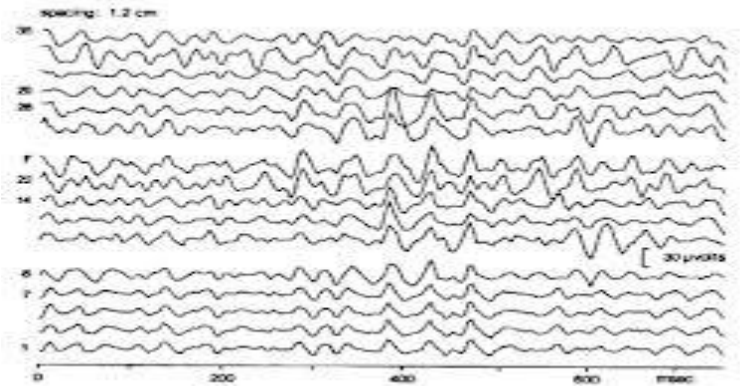

a
32
29
5
4
31
30
28
7
6
67
43
44
45
33
36
37
38
25
22
21
40 Hz
1 s
72
8
1
2
57
18
26
85
65
87
b
Number of spikes
Picture number
0
1
2
3
4
5
6
7
8
9
10
10
20
30
40
50
60
70
80


a
22
24
25
23
89
53
52
51
96
92
Halle Berry
Kobe Bryant
37
27
94
54
77
32
12
91
78
40
Catwoman
Emily
35
28
99
98
97
95
93
88
87
86
John Kerry
George Bush
Julia Roberts


b
Number of spikes
Picture number
0
2
4
6
8
10
12
14
16
18
20
22
10
20
30
40
50
60
70
80
90


a
22
82
24
23
81
80
79
2
8
14
Sydney Opera
Sydney Opera
29
72
73
66
34
95
85
75
37
7
Bahai
19
92
76
74
32
26
84
57
78
1
Grapes
Eiffel Tower
b
Number of spikes
Picture number
0
2
4
6
8
10
12
10
20
30
40
50
60
70
80
90

Do we have similar brains to monkeys?



(Kriegeskorte et al, Neuron, 2008)

Please cite this article in press as: Kriegeskorte et al., Matching Categorical Object Representations in Inferior Temporal Cortex of Man and Monkey, Neuron (2008), doi:10.1016/j.neuron.2008.10.043

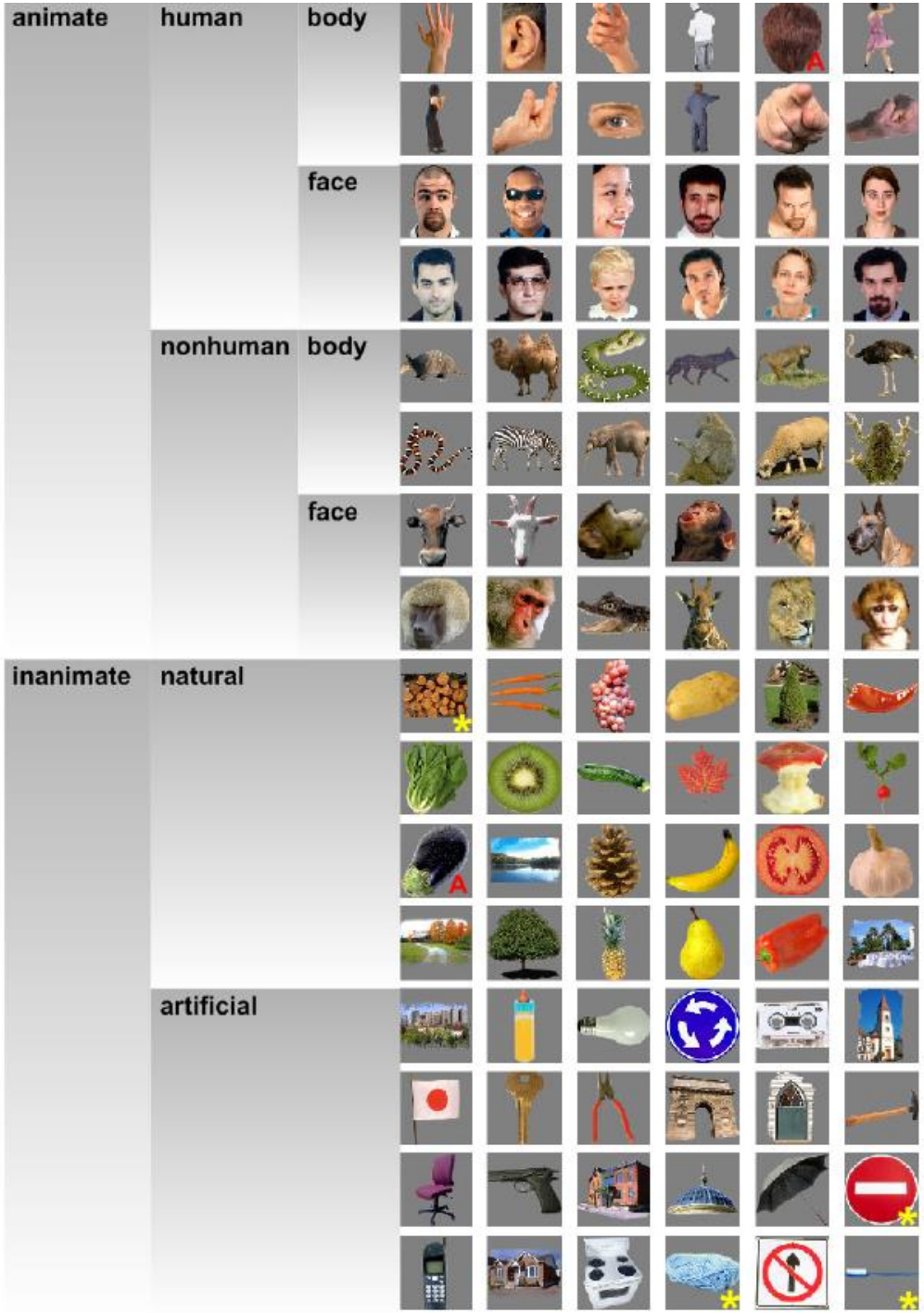
Neuron
Article

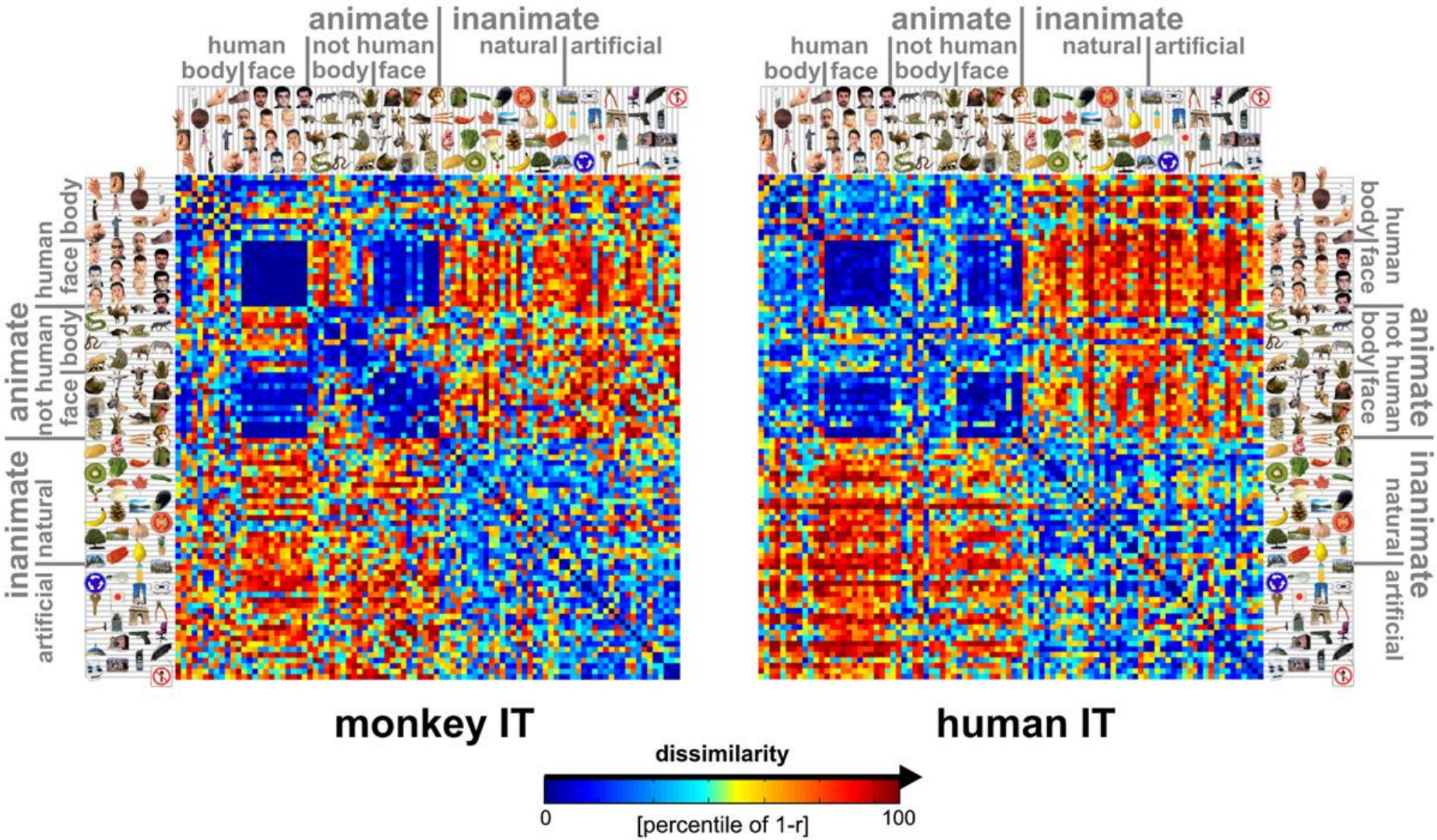
Matching Categorical Object Representations in Inferior Temporal Cortex of Man and Monkey

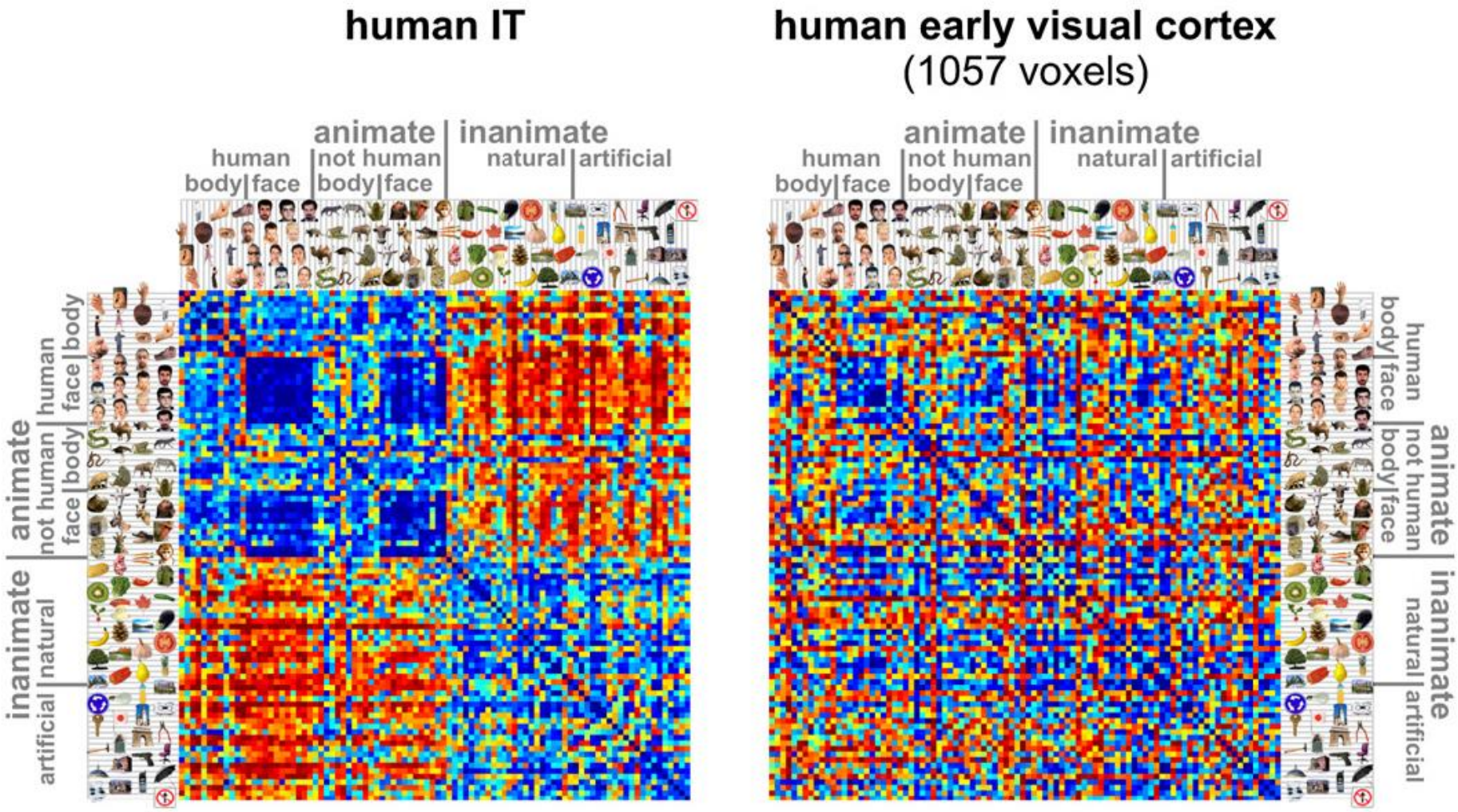
Nikolaus Kriegeskorte,[1,*] Marieke Mur,[1,2] Douglas A. Ruff,[1] Roozbeh Kiani,[3] Jerzy Bodurka,[1,4] Hossein Esteky,[5,6] Keiji Tanaka,[7] and Peter A. Bandettini[1,4]

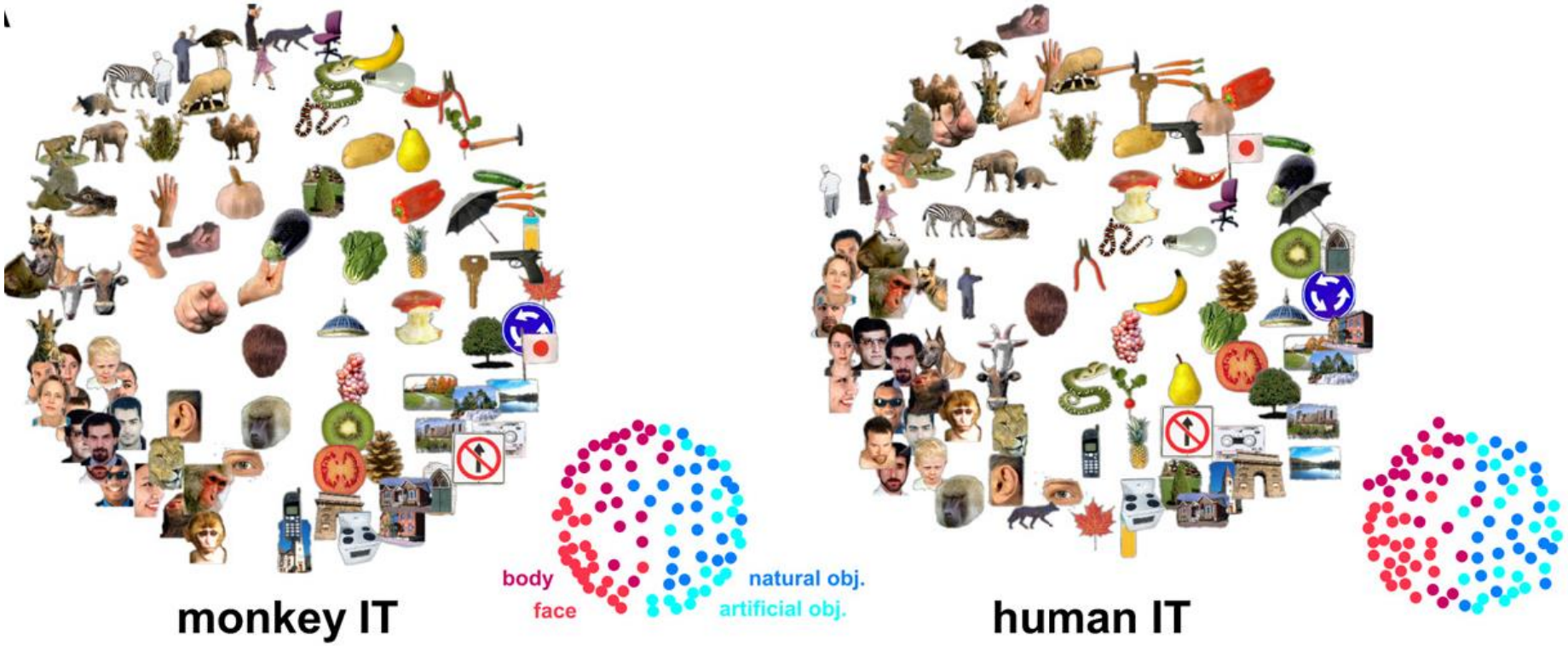
- 92 images presented to subjects in rapid succession
- 2 monkeys, single-cell recordings from 674 neurons in anterior IT (left in Monkey1, right in Monkey2)
- 4 humans, BOLD fMRI – IT voxels taken
- RDM (representational dissimilarity matrix) computed using correlation distance (1-PearsonR) across voxel populations

- 92 stimuli

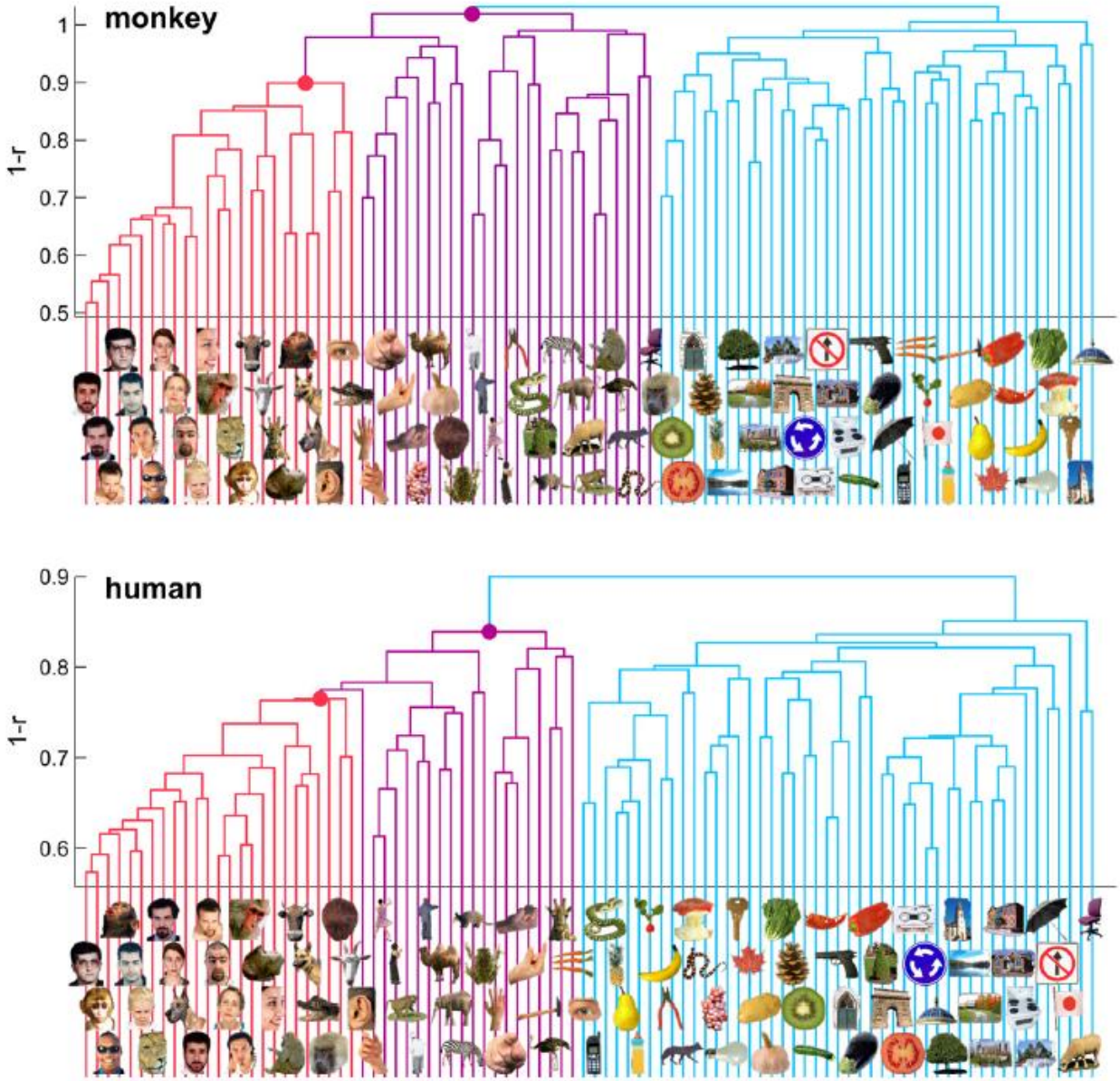

animate
human
body
face
nonhuman
body
face
inanimate
natural
artificial


animate
inanimate
human body
face
not human body
face
natural
artificial
monkey IT
human IT
dissimilarity
0
[percentile of 1-r]
100


human IT
animate
inanimate
human body
face
not human body
face
natural
artificial
human face
body
animate not human face
body
inanimate natural
artificial
human early visual cortex
(1057 voxels)
animate
inanimate
human body
face
not human body
face
natural
artificial
human body
face
animate not human body
face
inanimate natural
artificial



- MDS



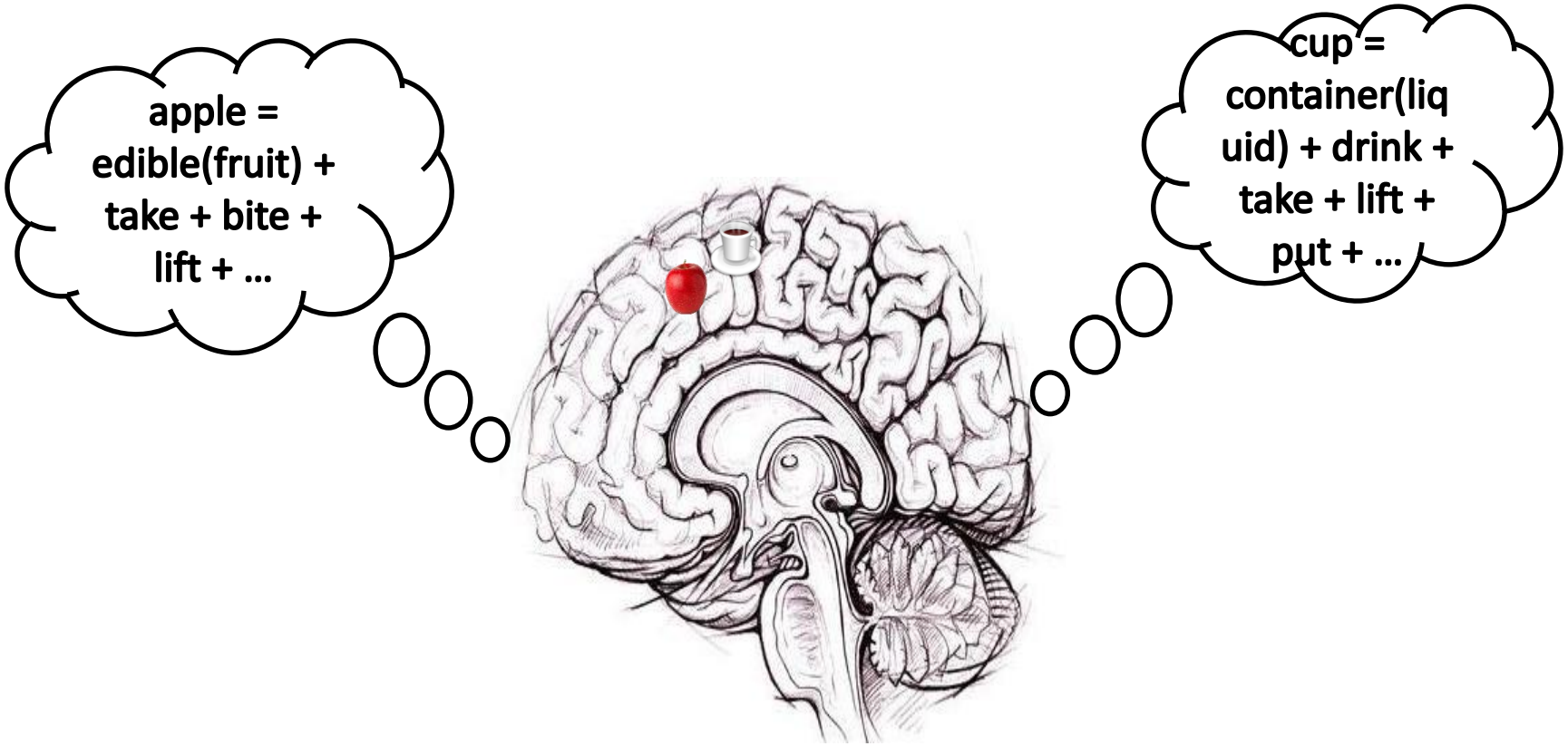
Hierarchical clustering

Can we use fMRI for mind reading?

- Mitchell et al, Science, 2008.
- Pereira et al, Frontiers in Human Neuroscience, 2011.
- Nishimoto et al, Cur. Biology, 2011.

Representing nouns & predicting brain activity associated with their meaning

(Mitchell et al., 2008)



Conceptual knowledge in the brain

- Distinct spatial patterns of fMRI activity are associated with viewing pictures of semantic categories (e.g. tools, animals)

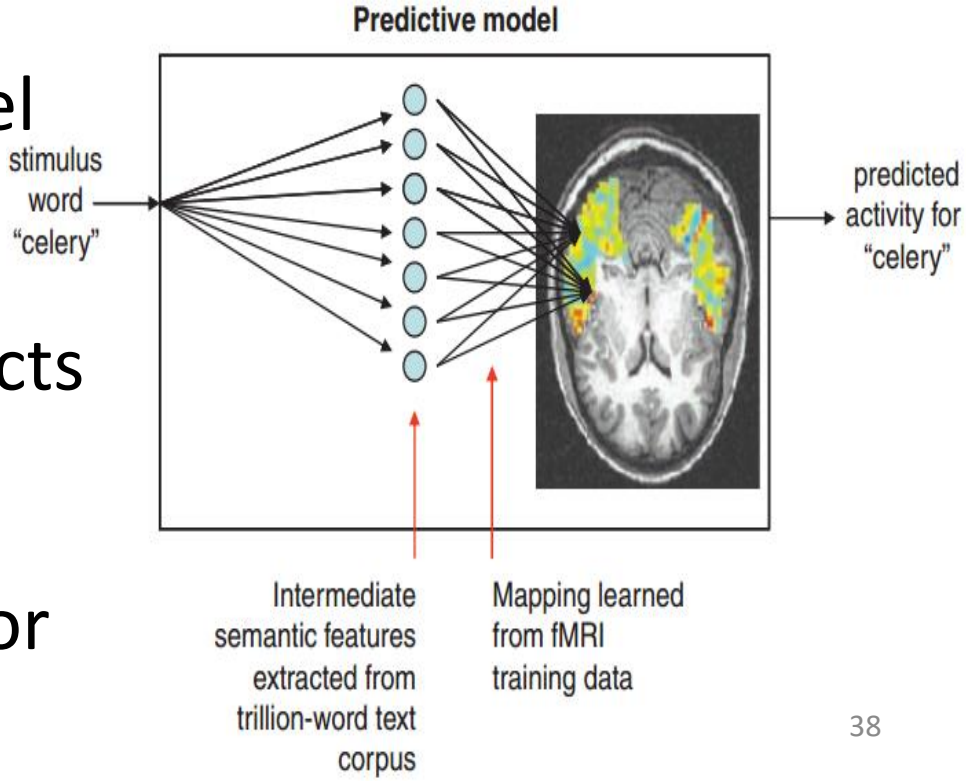
- Groupings of verbs and nouns, categorized by semantic roles
  (e.g. VerbNet & WordNet)

- word meanings can be captured by the distribution of words and phrases it commonly co-occurs with

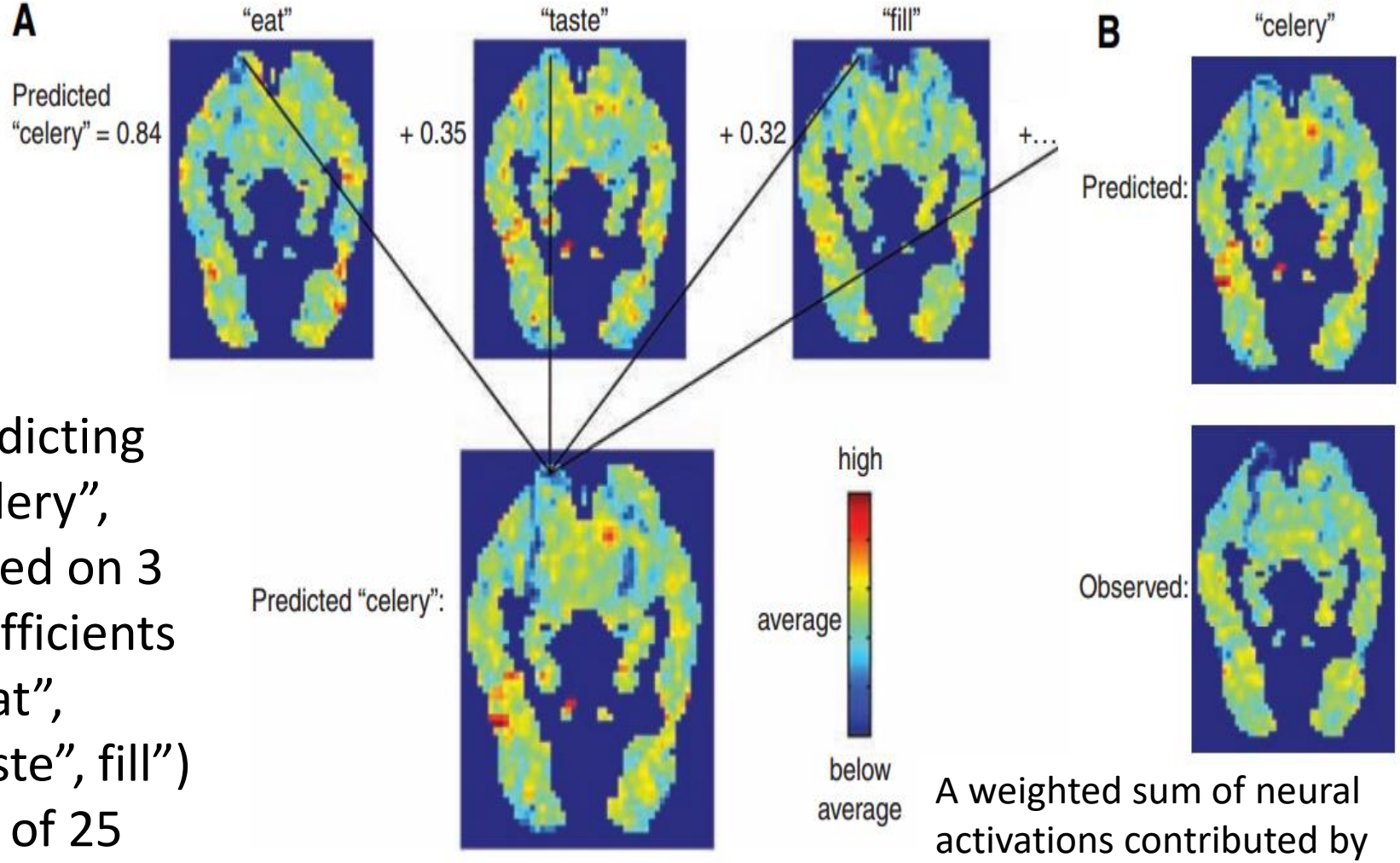
Predicting fMRI activation patterns

- Representation of concrete nouns is related to the distributional properties of those words in broadly based groups of the language

➔ A computational model trained to predict brain activity (fMRI) for the meaning of concrete objects (e.g. “celery”), even for objects that no fMRI activation patterns exist for



Example Prediction



Predicting "celery", based on 3 coefficients ("eat", "taste", fill") out of 25 semantic features

A weighted sum of neural activations contributed by each of these features is used to compute the fMRI activation at every voxel

Assumptions/Basic functionality

- meaning of the word is encoded as a vector of semantic features, based on the occurrences of that word within a (very large) corpus

- A weighted sum of neural activations contributed by each of these features is used to compute the fMRI activation at every voxel

- Model instantiated via a set of semantic features, “grounded” in sensory-motor actions (e.g. “hear”, “lift”, “fill”, “open”, etc.)

Model Performance

- Mean accuracy (training sets vary) for
  - words within the training set: **0.77**
  - words outside the training set: **0.70**
    - The model needs to extrapolate from words that are distant from those it was trained on
  - discriminating similar words (e.g. “celery” vs. “corn”): **0.62**

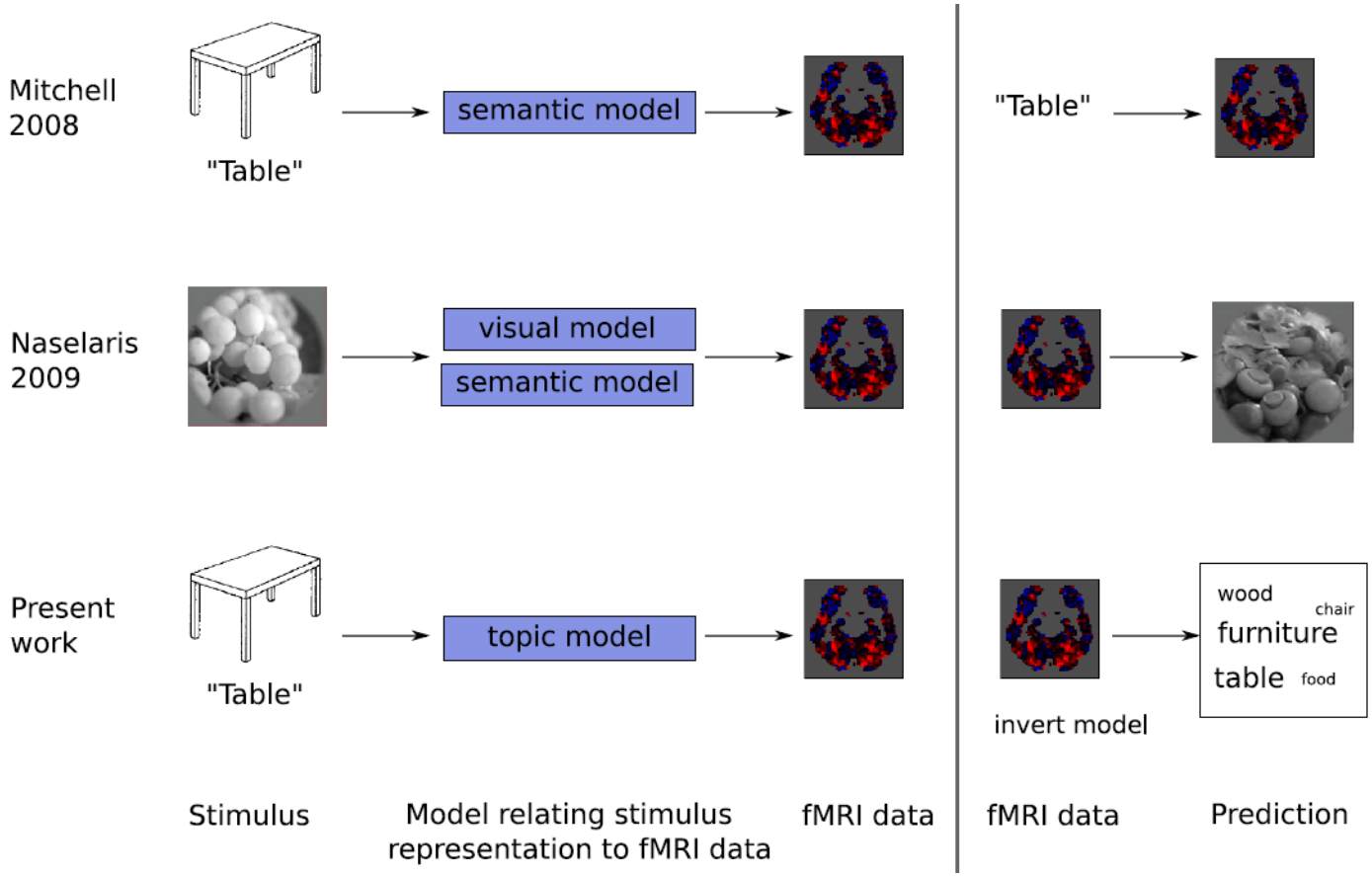
frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 23 August 2011
doi: 10.3389/fnhum.2011.00072

Generating text from functional brain images

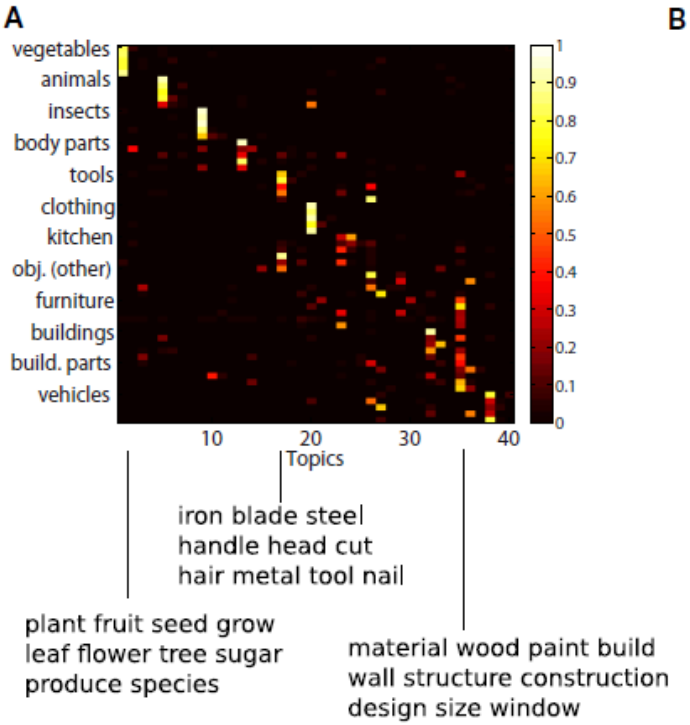
Francisco Pereira[1,2]*, Greg Detre[1,2] and Matthew Botvinick[1,2]

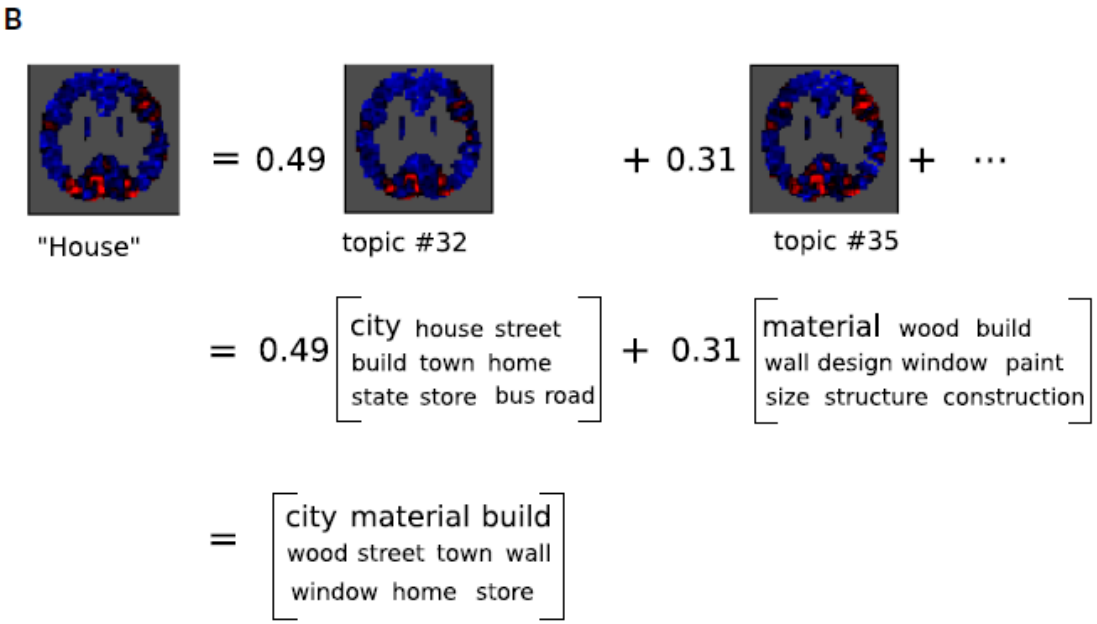
[1] *Department of Psychology, Princeton University, Princeton, NJ, USA*
[2] *Princeton Neuroscience Institute, Princeton University, Princeton NJ, USA*



Methodology

- Corpus of text (Wikipedia articles) used to create a *topic model*, a latent factor representation of each article (the *concept* the article is about).
- Learn a mapping from each latent factor in the model from Step 1 to a corresponding brain image that captures how the factor gives rise to sub-patterns of distributed brain activity, using a training set of brain images.
- For each brain image in a new, test set, the mapping from Step 2 can be used to infer a weighting over latent factors. Given this, *invert* the generative model from Step 1 in order to map from the latent factor representation to *text*, in the shape of a probability distribution over words.

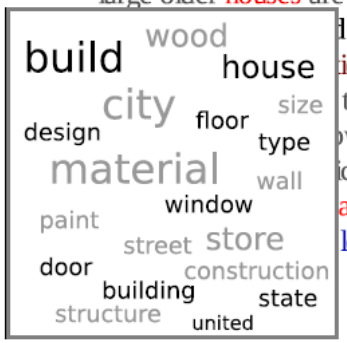

A
vegetables
animals
insects
body parts
tools
clothing
kitchen
obj. (other)
furniture
buildings
build. parts
vehicles
10
20
30
40
Topics
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
iron blade steel
handle head cut
hair metal tool nail
plant fruit seed grow
leaf flower tree sugar
produce species
material wood paint build
wall structure construction
design size window


B
"House"
= 0.49
topic #32
+ 0.31
topic #35
+ ⋯
= 0.49 [city house street build town home state store bus road]
+ 0.31 [material wood build wall design window paint size structure construction]
= [city material build wood street town wall window home store]

Keywords derived from brain activity of a person reading the article

Apartment

An apartment is a self-contained housing unit that occupies only part of a building. Apartments may be owned (by an "owner occupier") or rented (by "tenants"). In the US, some apartment-dwellers own their own apartments, either as co-ops, in which the residents own shares of a corporation that owns the building or development; or in condominiums, whose residents own their apartments and share ownership of the public spaces. Most apartments are in buildings designed for the purpose, but large older houses are sometimes divided into



l "apartment" connotes a
tion in a building. In some
the United States, the word
owned by the building
ically used for a
artment landlords, each
loss of income from

Hammer

A hammer is a tool meant to deliver an impact to an object. The most common uses are for driving nails, fitting parts, and breaking up objects. Hammers are often designed for a specific purpose, and vary widely in their shape and structure. Usual features are a handle and a head, with most of the weight in the head. The basic design is hand-operated, but there are also many mechanically operated models for heavier uses. The hammer is a basic tool of many professions, and can also be used as a weapon. By analogy, the name '"hammer"' has also been used for devices that are designed to deliver blows, e.g. in the caplock

s. History. The use of simple
,400,000 BCE when various
sed to strike wood, bone, or
them apart and shape them.
cks with strips of leather or
ng used as hammers by about
e middle of the Paleolithic Stone

Current Biology *21*, 1641–1646, October 11, 2011 DOI 10.1016/j.cub.2011.08.031

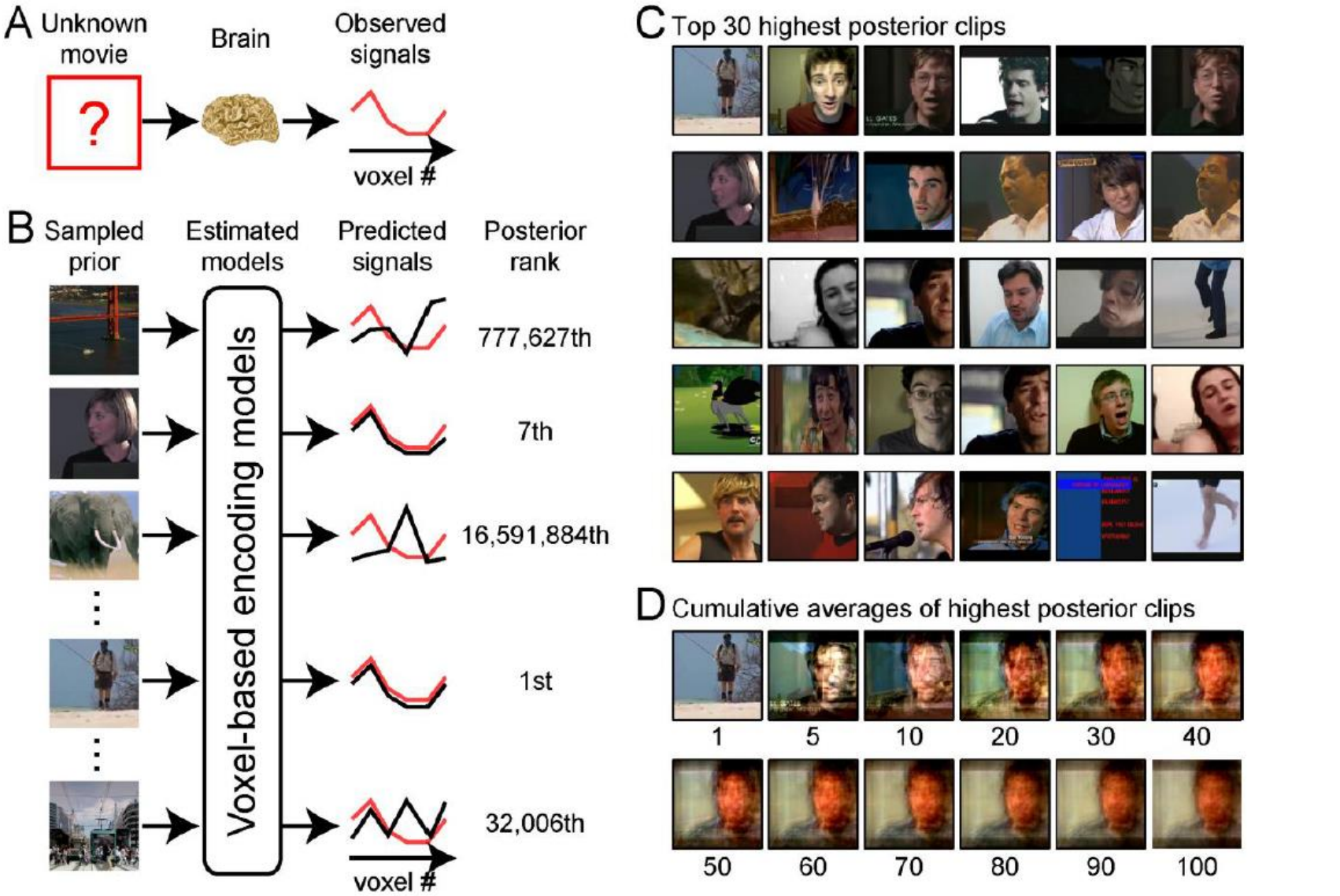
Report

Reconstructing Visual Experiences from Brain Activity Evoked by Natural Movies

Shinji Nishimoto,[1] An T. Vu,[2] Thomas Naselaris,[1] Yuval Benjamini,[3] Bin Yu,[3] and Jack L. Gallant[1,2,4,*]

mental processes. It has therefore been assumed that fMRI data would not be useful for modeling brain activity evoked

- Record brain activity while the subject watches several hours of movie trailers.
- Build dictionaries (regression model) to translate between the shapes, edges and motion in the movies and measured brain activity. A separate dictionary is constructed for each of several thousand points in the brain at which brain activity was measured.
- Record brain activity to a new set of movie trailers that will be used to test the quality of the dictionaries and reconstructions.
- Build a random library of ~5000 hours of video downloaded at random from YouTube (that have no overlap with the movies subjects saw in the magnet). Put each of these clips through the dictionaries to generate predictions of brain activity. Select the 100 clips whose predicted activity is most similar to the observed brain activity. Average those clips together. This is the reconstruction.


A
Unknown movie
Brain
Observed signals
?
voxel #
B
Sampled prior
Estimated models
Predicted signals
Posterior rank
Voxel-based encoding models
777,627th
7th
16,591,884th
1st
32,006th
voxel #
C
Top 30 highest posterior clips
D
Cumulative averages of highest posterior clips
1
5
10
20
30
40
50
60
70
80
90
100

- video