Bio-Inspired Artificial Intelligence

Lecture 4: Bio-Inspired Language Processing

Knowledge Technology

http://www.informatik.uni-hamburg.de/WTM/

Humans despair to use the brain for communication

Language

- What is **frappucino**?
- Order in a restaurant a ”tall skinny mocha frappucino with a dash of vanilla” – you will most likely get a hot coffee drink
- ***Universal*** in human society and ***unique*** to humans
  - Thousands of languages and dialects throughout the world
  - Differ in word orders, word forms, phonemes, ...
  - Language represents information, experience, thoughts
- Hypothesis:
  “Universality of language is a consequence of the fact that the human brain has evolved special language-processing systems” [Bear et al. 2006]

Motivation

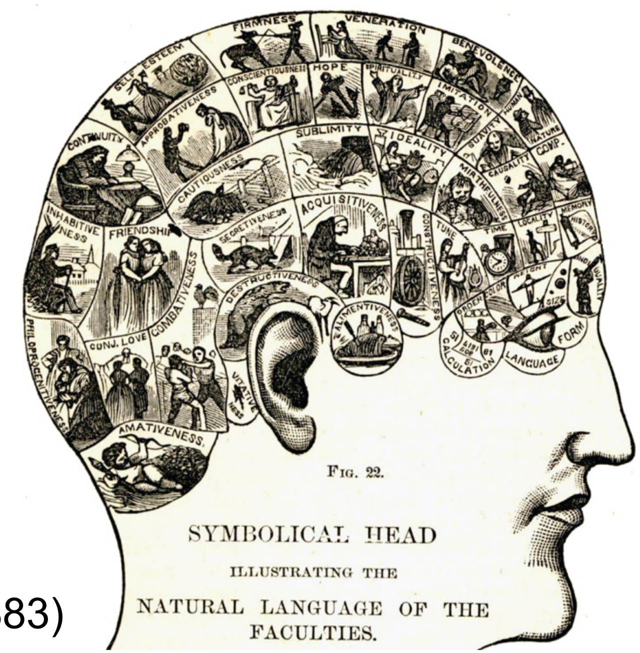
- Learn about psychological / ***neurobiological factors*** that enable us to acquire, use, and understand language
  - Language acquisition / production / comprehension
- Understand and ***model the process*** that is the basis for human language capacity
  - How does human language processing work?
  - How is linguistic knowledge represented in the brain?
- Gain from ***computational*** language ***models***
  - Improved understanding of the brain and human mind
  - Develop intelligent systems capable of using natural language

Example: Early word learning

Video from BBC documentary with Prof. Deb Roy
http://www.media.mit.edu/people/dkroy

Historical starts to explore language processing

- First systematic studies – relationship between brain and language
- Phrenology / craniology (Gall, ca. 1800)
  - Located language in anterior parts of the brain at the protrusion of the eye socket below the eye
  - Bumps of the skull taken to reflect areas of enlargement in the brain
- Empiricism of the 18th/19th century
  - Paul Broca: anatomical inspection of the brain
  - Study of aphasia

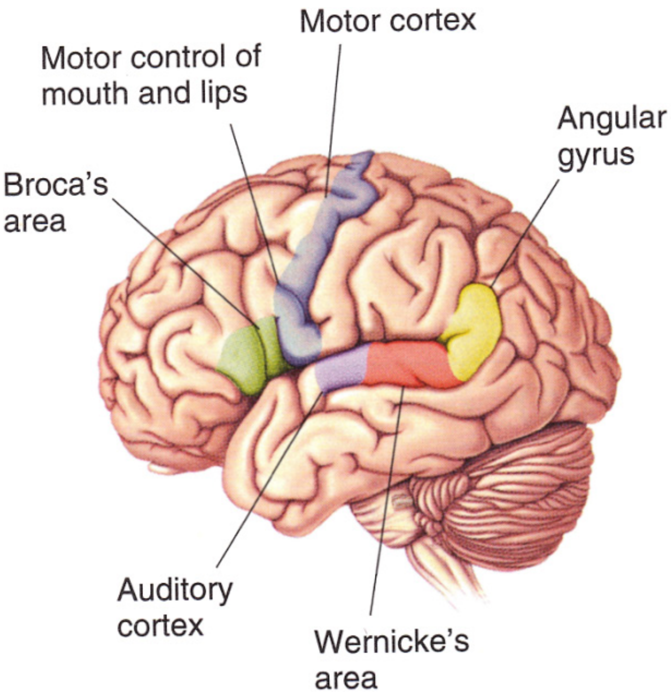


People's Cyclopedia of Universal Knowledge (1883)

Broca's area

- Paul Broca (physician, ca. 1860)
  - Localization of functions in the cerebral cortex

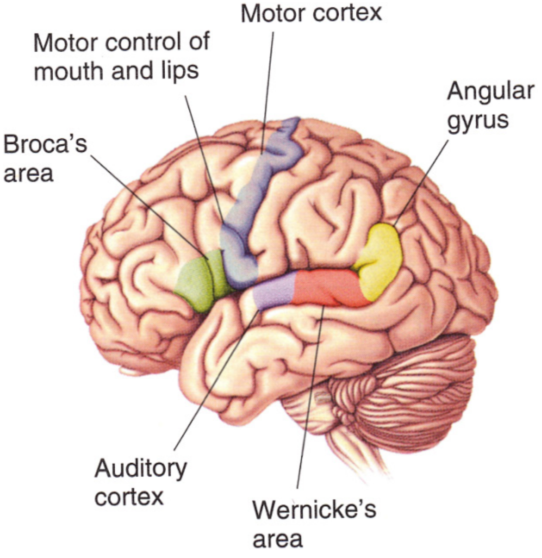
- Broca's aphasia
  - Sparse speech, nonfluent
  - Deficits of intonation and stress patterns
  - Lack of ***grammatical structure***



Based on [Bear et al. 2006]

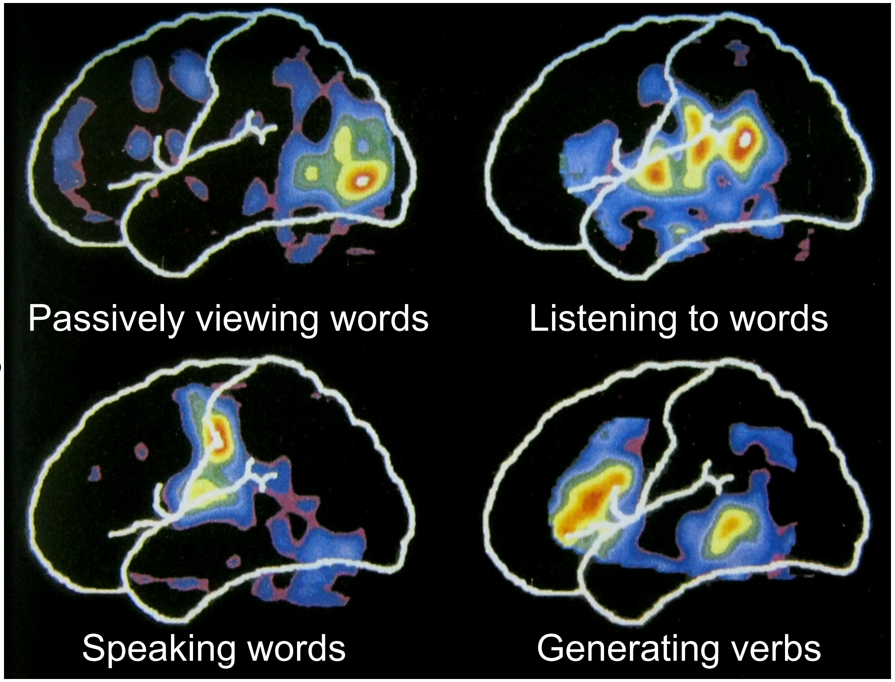
Wernicke's area

- Carl Wernicke (German physician)
  - Damage to an area of the left superior temporal lobe (part of auditory association cortex, and next to primary auditory cortex)
  - Result: Loss of ***language comprehension***
    - Wernicke's aphasia
  - Superior temporal area: where auditory words are stored
  - Proposed a connection between Wernicke and Broca



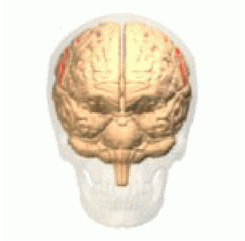
Non-invasive Brain Imaging

- Brain imaging (after 1950)
  - Positron emission tomography (***PET***) – ca. since 1950
  - functional Magnetic Resonance Imaging (***fMRI***) – ca since 1990
  - Allows to infer neural activity in different parts of the brain from regional blood flow
  - **Example**: PET visualises sensation and speech:

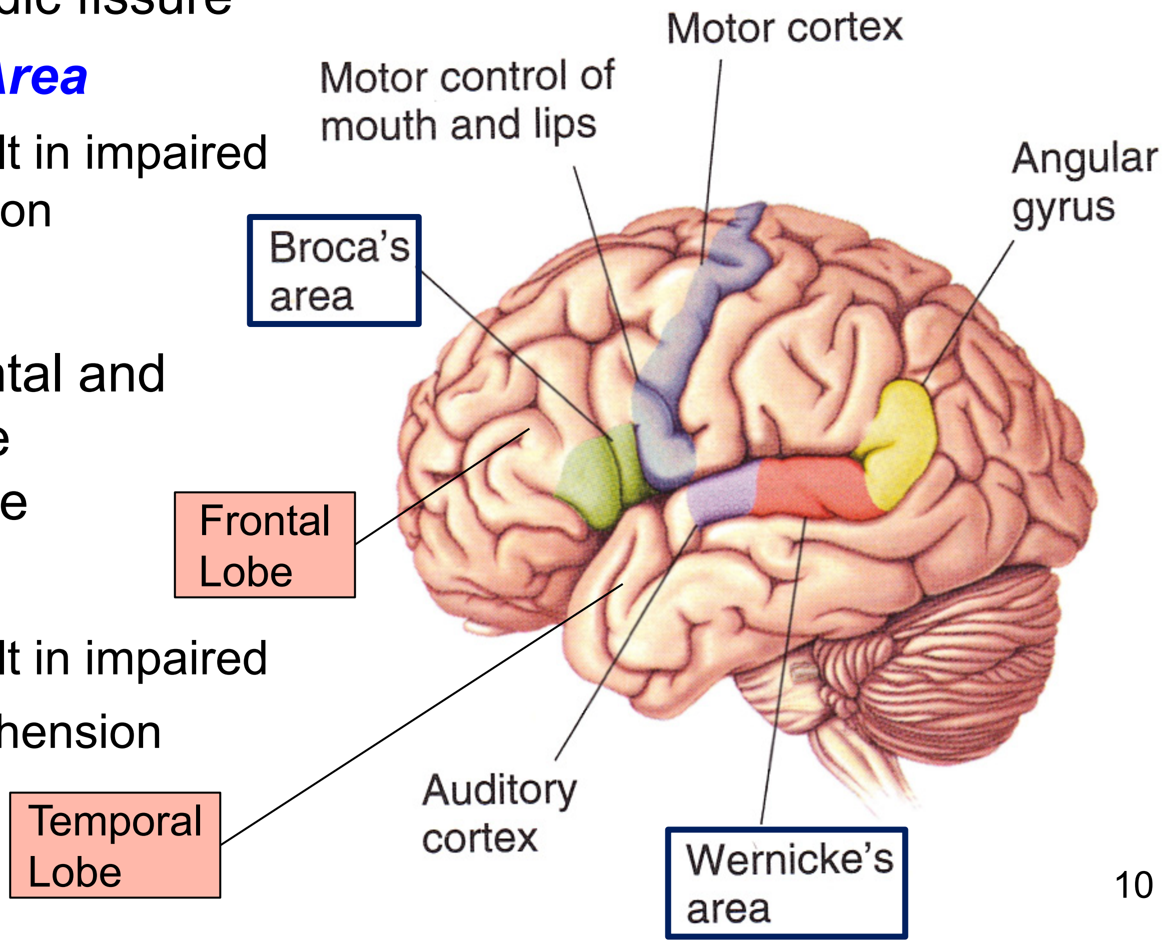


Based on [Posner & Raichle 1994]

Organization of language in the brain



- Frontal lobe
  - Separated from parietal lobe by central sulcus/rolandic fissure
  - Contains ***Broca's Area***
    - Damage can result in impaired language production
- Temporal lobe
  - Separated from frontal and parietal lobes by the lateral/sylvian fissure
  - ***Wernicke's area***
    - Damage can result in impaired language comprehension



Human language processing: Functions

- ***Comprehension***: Maps from “sound to meaning”
  - Speech/text to words
  - Words to structures
  - Structure to meanings
- ***Production***: Maps from “meaning to speech”
  - Meaning to grammatical encoding
  - Phonological encoding
  - Articulation
- Interactivist, parallel processing rather than sequential modules is the current state of the art

Language processing: Early Wernicke-Geschwind Model

Early visual processing:
Striate cortex

Visual word recognition:
Extrastriate gyrus

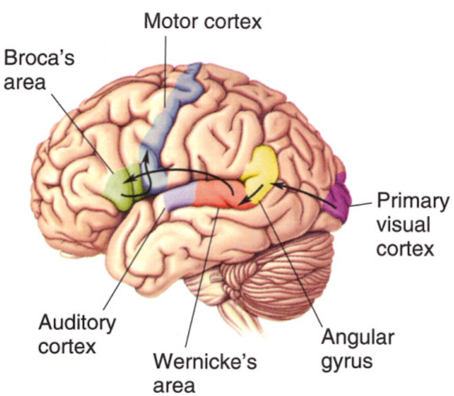
Early auditory processing:
Primary auditory cortex

Auditory word recognition:
Angular gyrus

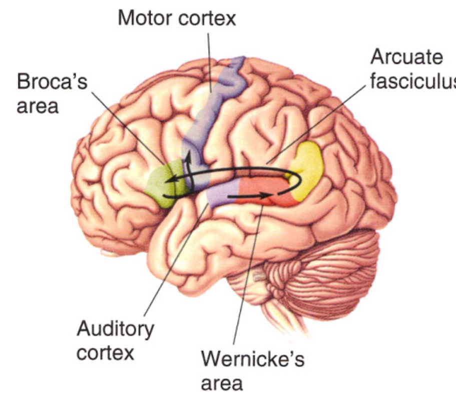
Semantic association:
Inferior frontal cortex

Premotor coding:
Supplementary motor area

Motor control of speech:
Primary motor cortex



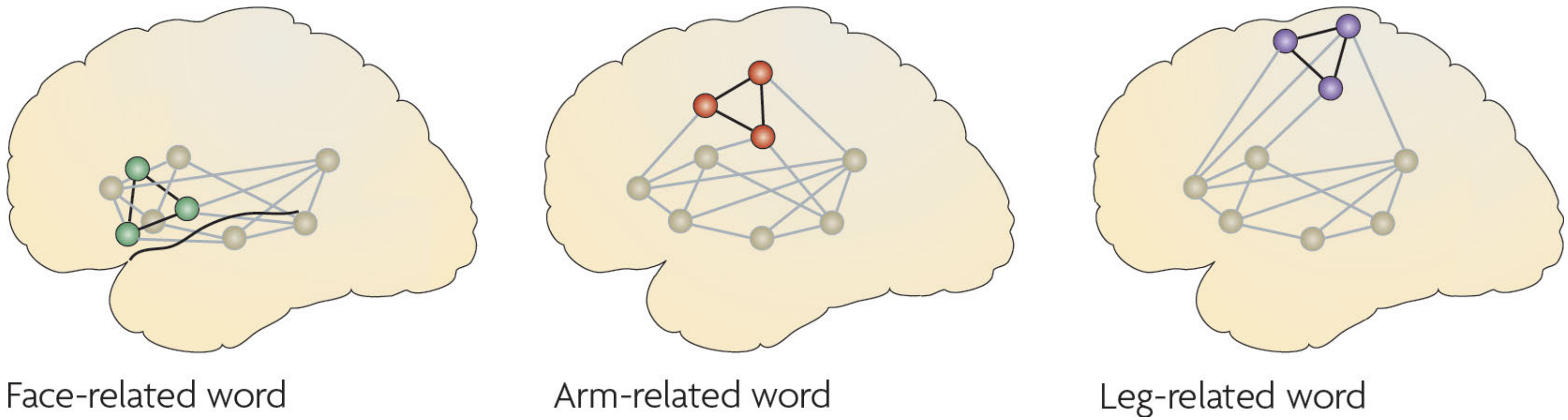
tell what you see



repeat what you hear

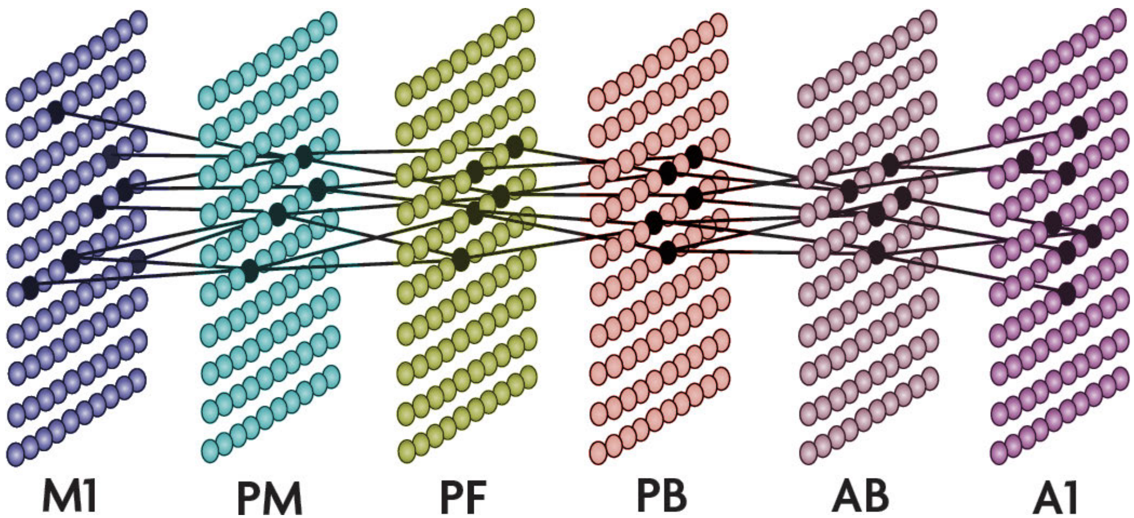
Language processing: Embodied Model

- Language is rather embodied and distributed over several areas of the brain: ***word-webs*** [Pulvermüller 2003]
  - Different areas are active during processing of words for different objects/subjects:
    - “Cat”: Activity in motor areas, because you have stroked a cat
    - “Shark”: Activity in vision areas, because you have seen one
  - Different areas are active during processing of words for different body parts (based on fMRI):

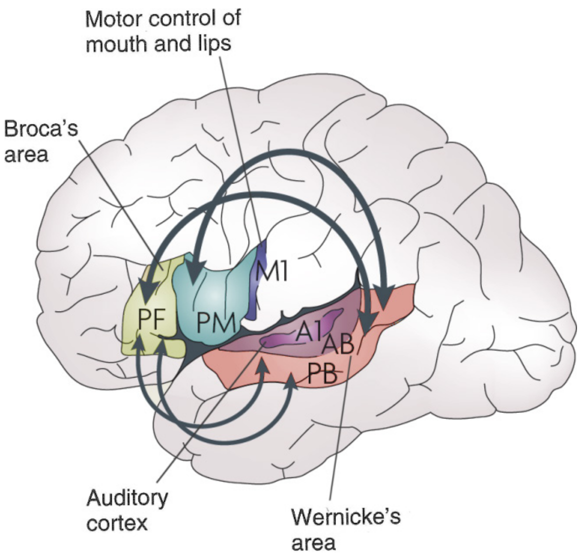


Associations between auditory word recognition and motor control of speech

- Associator: Multi-layer Neural Network
  - Mimics neuroanatomical, ***connectivity***, and neurophysiological ***properties of language*** related areas



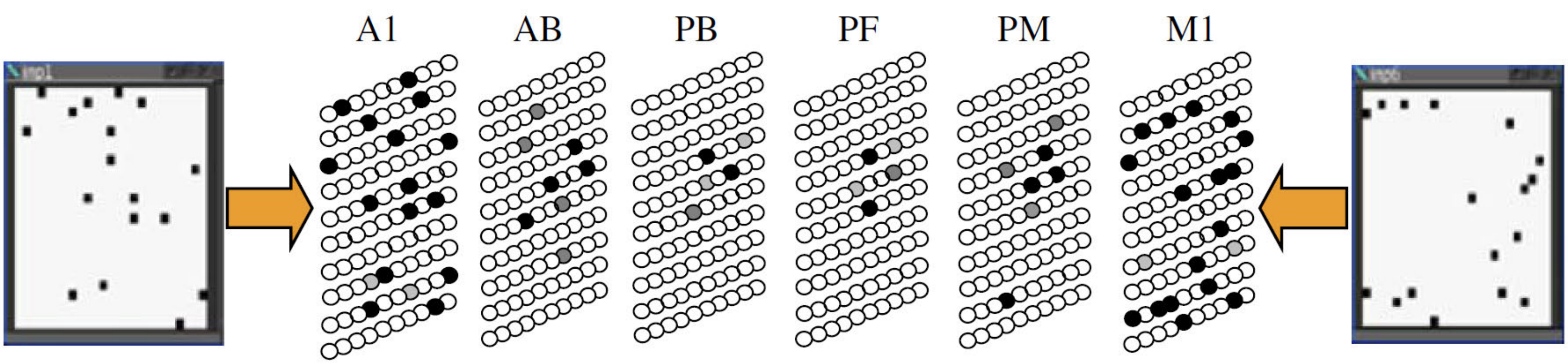
[Garagnani et al. 2009]



Primary motor cortex
Premotor
Prefrontal (Broca)
Parabelt (Wernicke)
Auditory belt
Primary auditory cortex

Establishing associations

- Neural Net for world learning via action-perception correlation
  - A1 represents a morpheme/auditory form
  - M1 represents an articulatory form
- Sparseness:
  - 25x25 cells each layer (= 625 cells/layer)
  - 17 cells active for representing an entity
- Tested with two different learning rules



Learning Rules

- Hebbian (or Co-variance) Learning Rule:

  $$\Delta w_{i,j} = \alpha\left(x_i - \langle x_i \rangle\right)\left(x_j - \langle x_j \rangle\right)$$

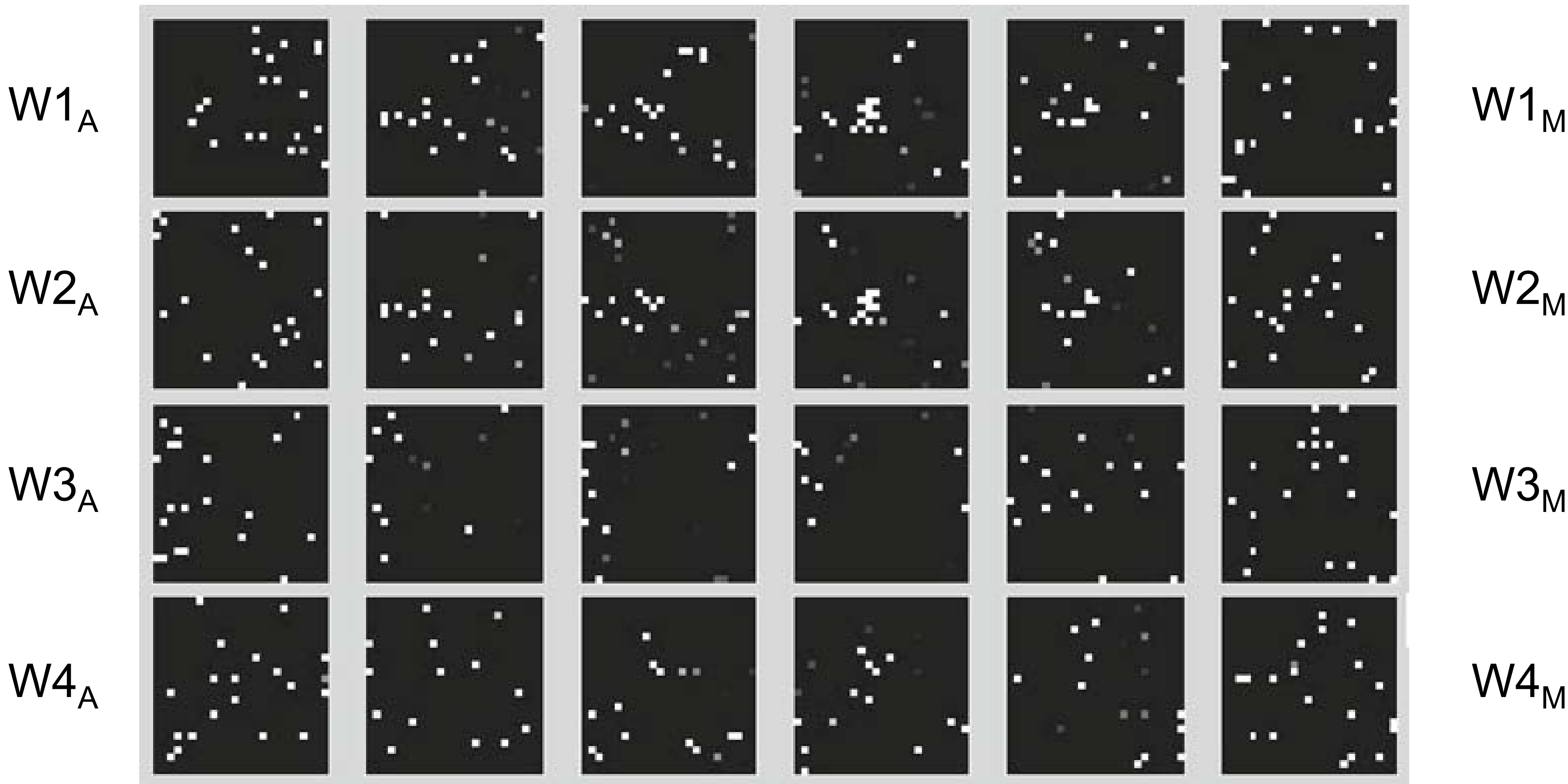
$\Delta w_{i,j}$: weight adaptation
α : learning rate
x_i : output of cell i
$\langle x_i \rangle$: time averaged output of cell i

- ***What wires together, fires together***

Training and results

- Hebbian rule:

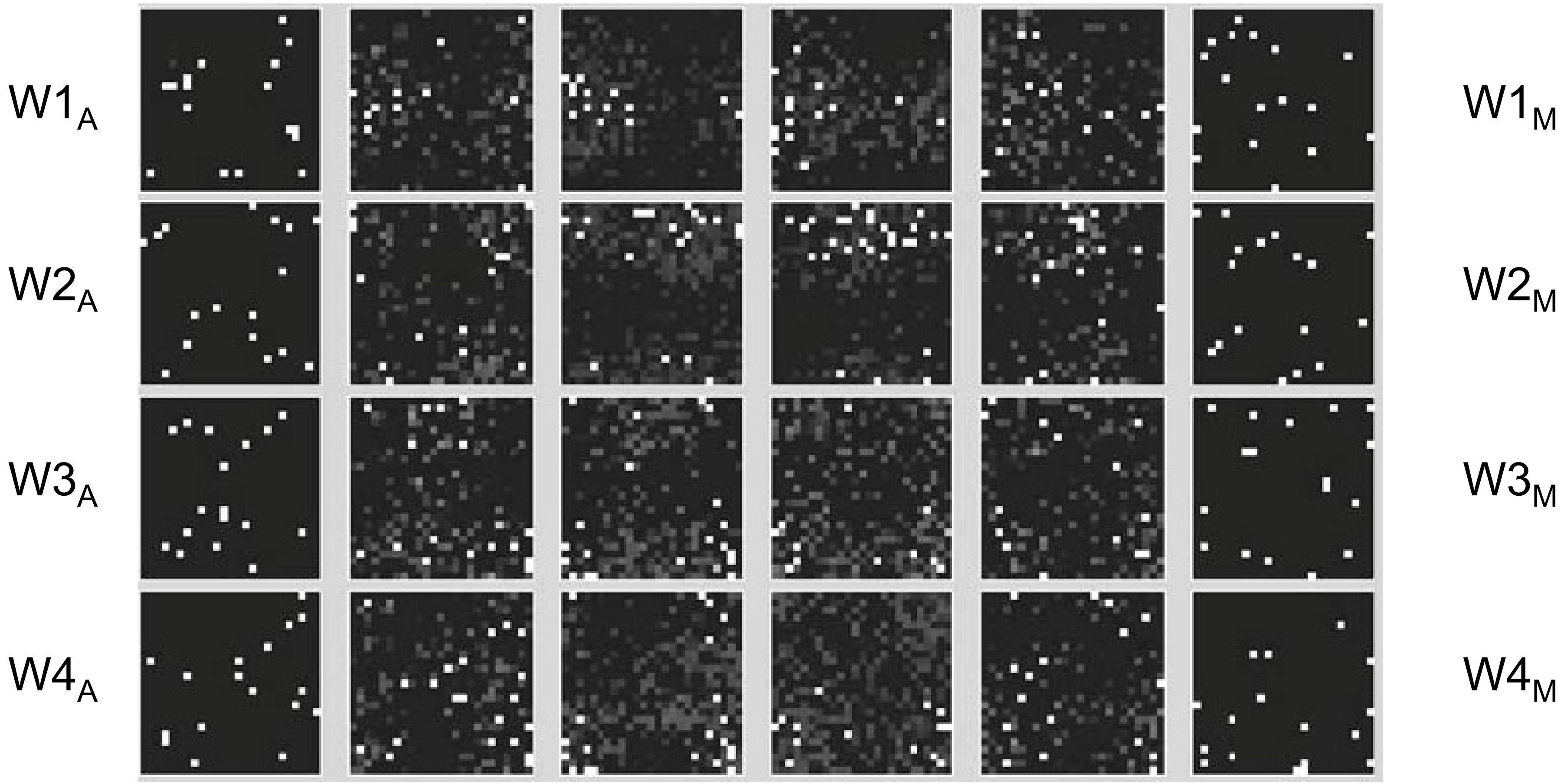
t=3000



Training and results (cont.)

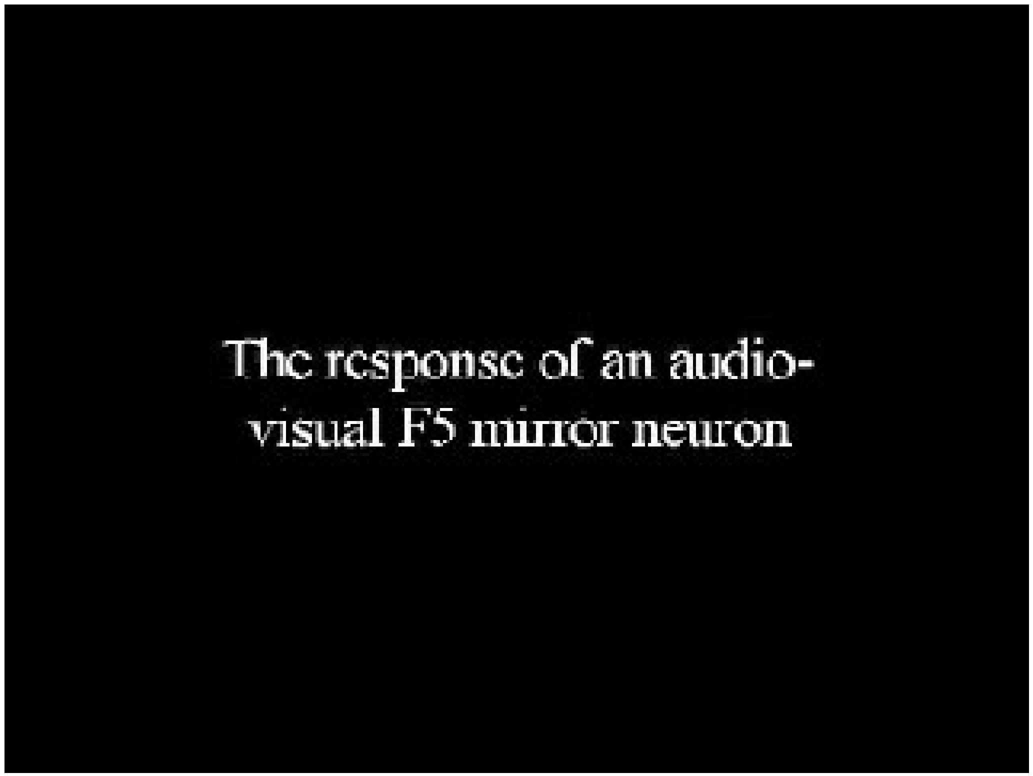
- ABS rule:

t=5000



Break for discussion on language and animals

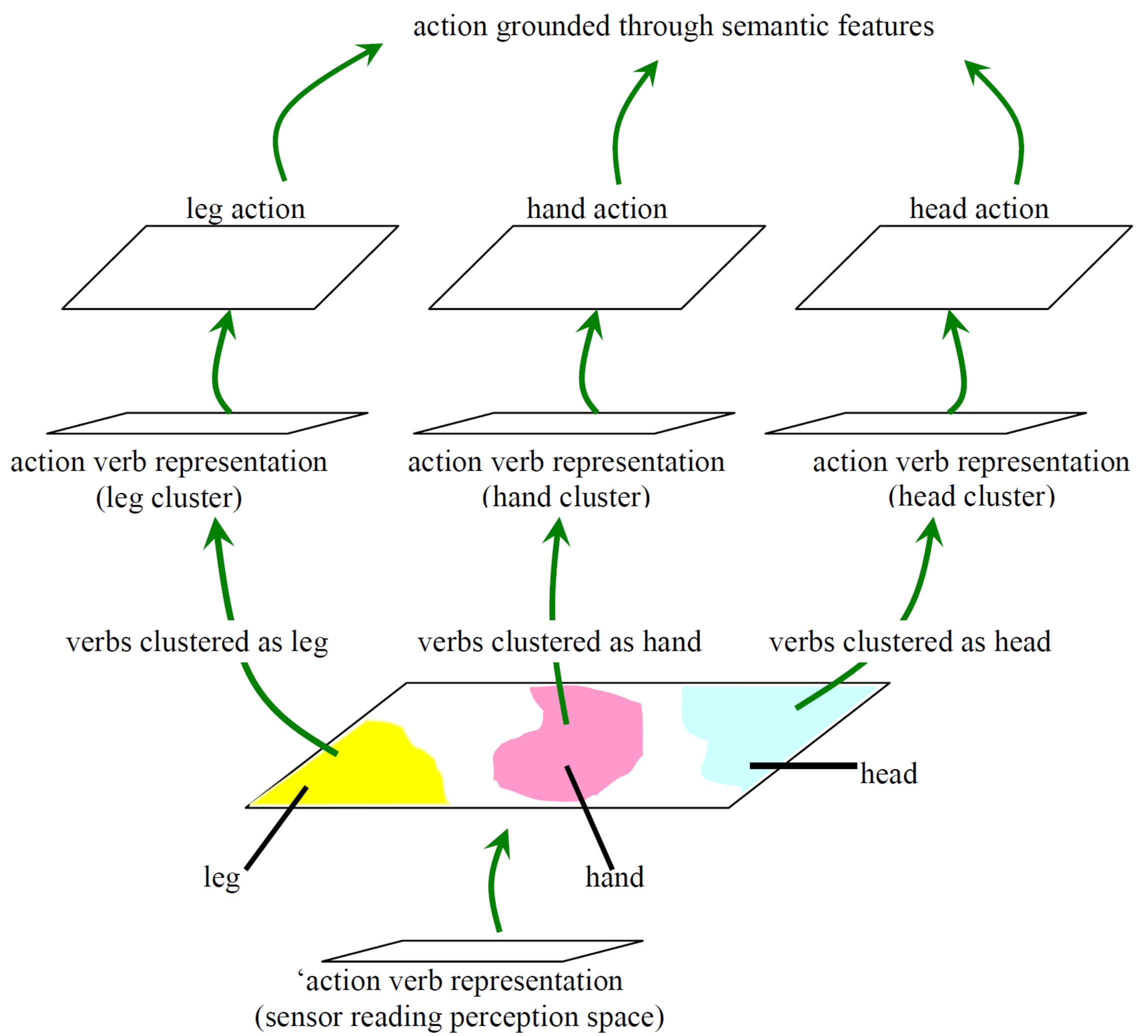
Audio visual mirror neurons firing



Rizzolatti, Gallese and team, Parma

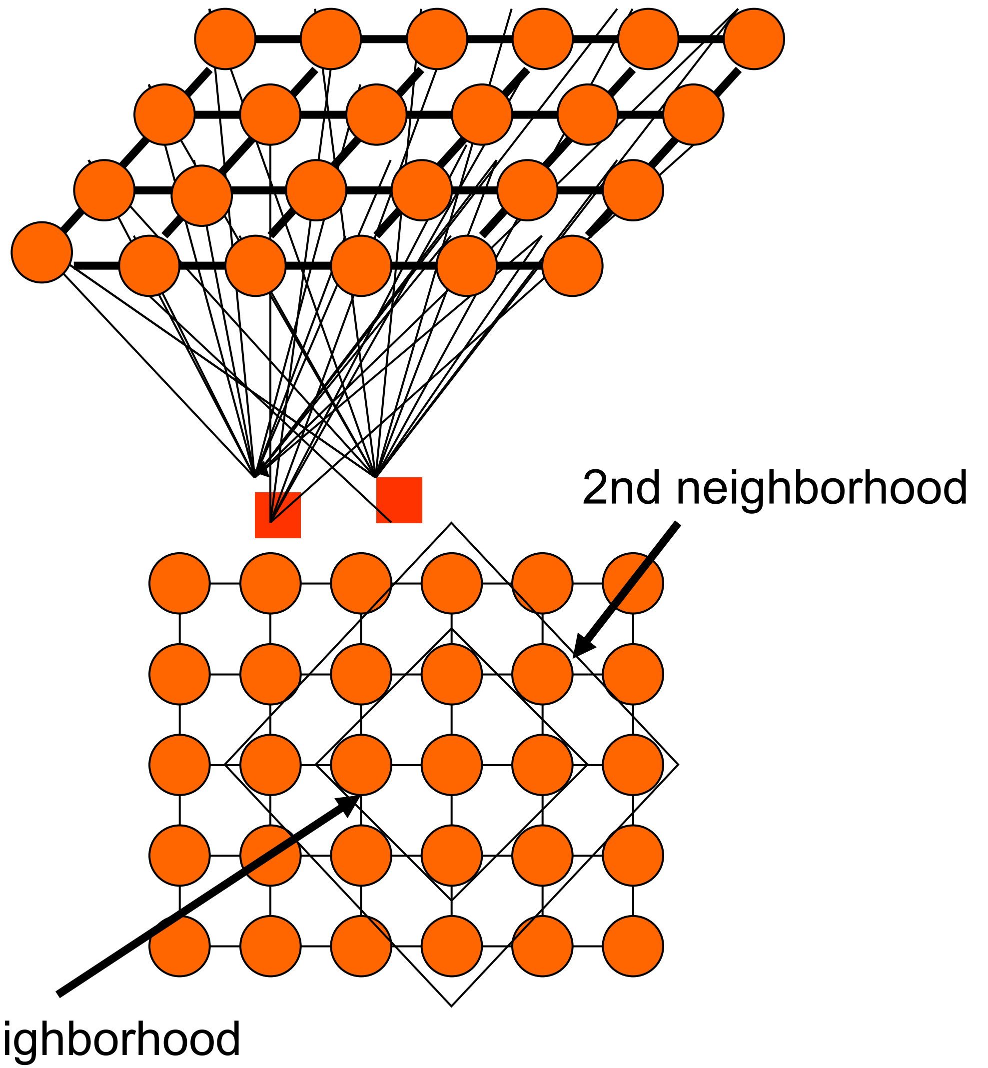
Semantic association: Grounding language in action (Elshaw 2005)

- Learning of semantic ***features***
  - Modular system
  - Distributed processing
  - Reflect cortex principles
- Modelled with SOMs
  - Coarse clustering at lower level
  - Fine clustering at higher level



Self-Organising Maps: Activation spreading

- The activation of the neuron is spread in its direct ***neighborhood***
  - neighbors become sensitive to the same input patterns
- The size of the neighborhood is initially large but reduced over time during training as the network neurons become more specialized

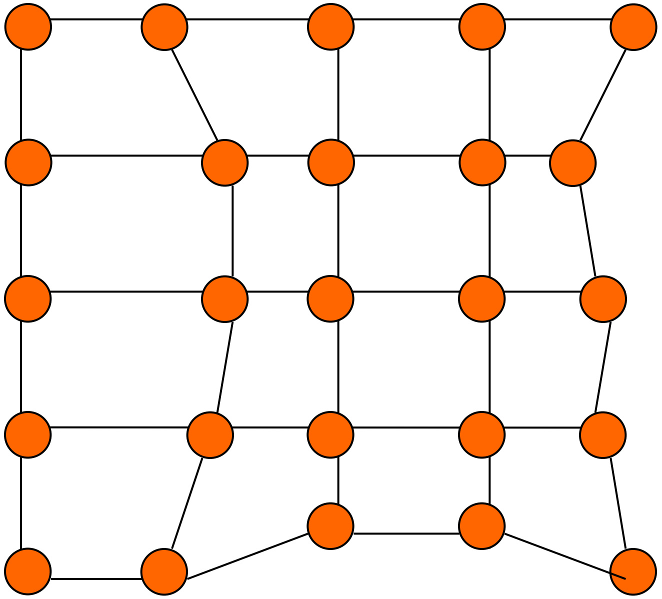


Self-Organizing feature Maps (SOMs)

- SOMs, also called topological ordered maps, or Kohonen Self-Organizing Feature Map
- Map points from high-dimensional source space into a 2 to 3-d target space; distance and proximity relationship (i.e., topology) preserved as much as possible
- Clustering is performed by having several units competing for the current object
  - Unit whose weight vector is closest to current object wins
  - Adjust winner and its neighbours
- SOMs are believed to resemble self-organization in the brain at a high level
- Useful for visualizing high-dimensional data in 2- or 3-D space

Self-Organizing Maps: Adaptation

- During training, “winner” neuron and its neighborhood adapts to make their weight vector more similar to the input pattern
- Neurons are moved closer to the input pattern
- The magnitude of the adaptation is controlled via a learning parameter which decays over time



Self-Organizing Maps: Learning

- The weight vectors are randomly initialised
- Input vectors are presented to the network
  - Neurons activated proportional to the Euclidean distance between the input and the weight vector

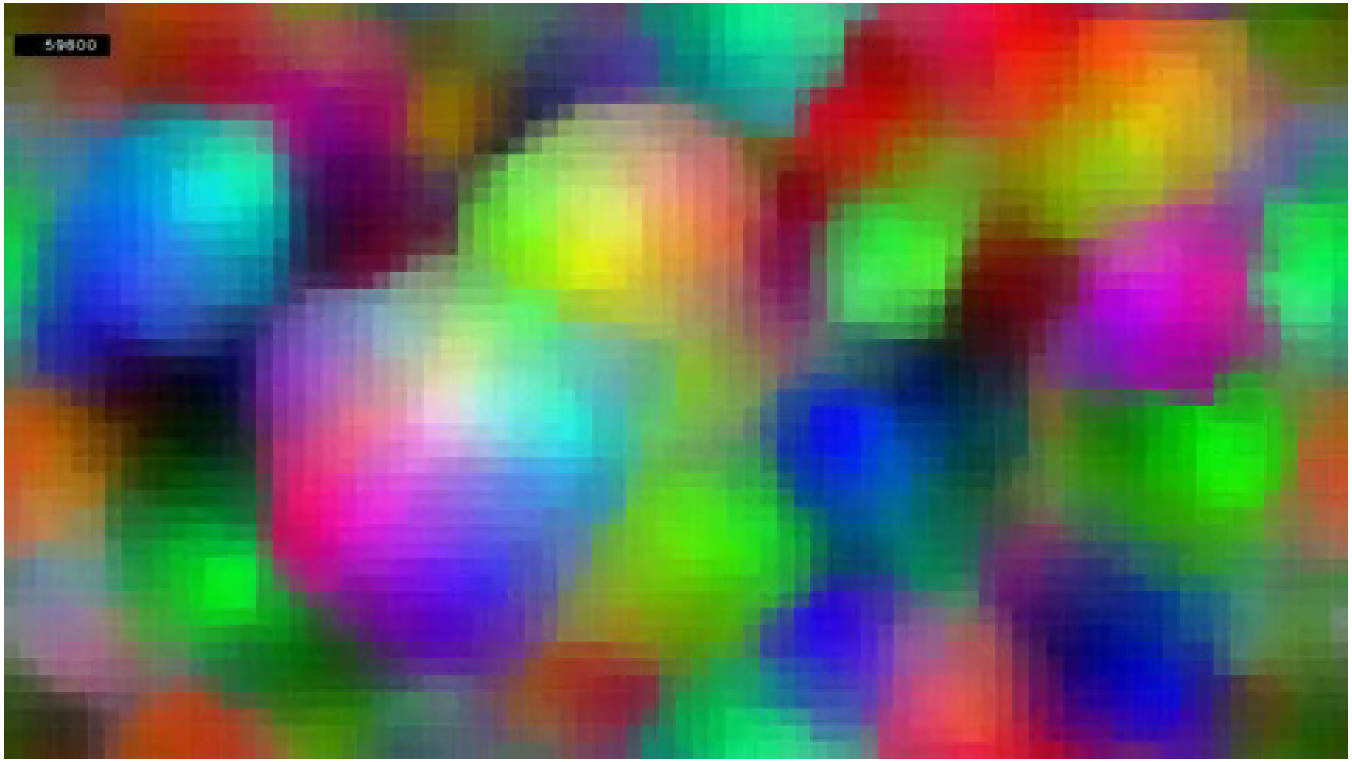
$$\min_j \left\| x - w_j^T \right\|$$

  - The winning node and neighbours have weight vector moved closer to the input

$$w_j^T \leftarrow w_j^T + \eta(t)\left(x - w_j^T\right)$$

- Over time, the network *self-organises* so that the input topology is preserved

Self-Organizing Maps: Learning



Grounding language in action: Scenario

- Intelligent Systems is supposed to perform actions related to Head, Hand and Leg body parts based on
  - Features ***embodied*** in motor and sensors of the system
  - Symbolic ***language description*** stating the actions

- Desired behaviour:
  - Recognition of the action and the action verbs based on the performed action
  - Self-organising of semantic association

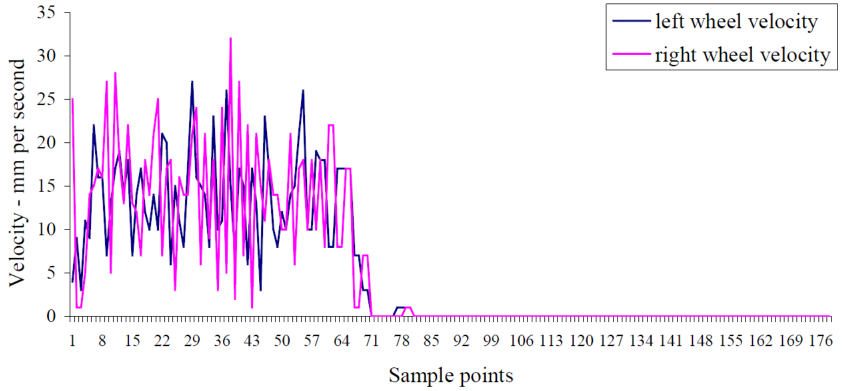
Scenario: MIRA robot for grounding language in robot actions

- Cognitive robot
  - PeopleBot platform, 120° pan-tilt camera
  - 2-DOF gripper, two side wheels, one back wheel
- Language
  - Action words are related to the capabilities of gripper, wheels and camera joints

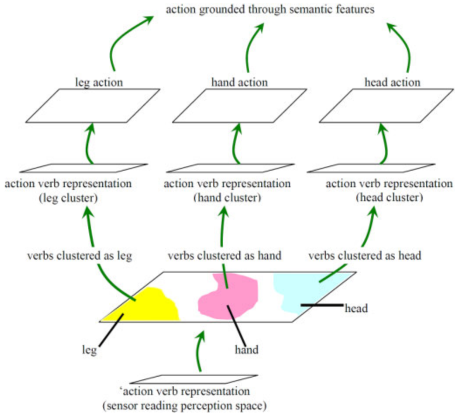
Leg	Head	Hand
turn left	head up	pick
turn right	head down	put
forward	head right	lift
backward	head left	drop
		touch

Training the hierarchical SOMs

- Preprocessing:
  - Transform sensor readings and language descriptions in numerical values
  - Collection of sensor readings for every action

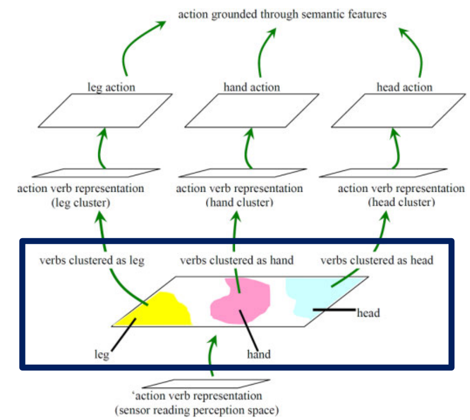


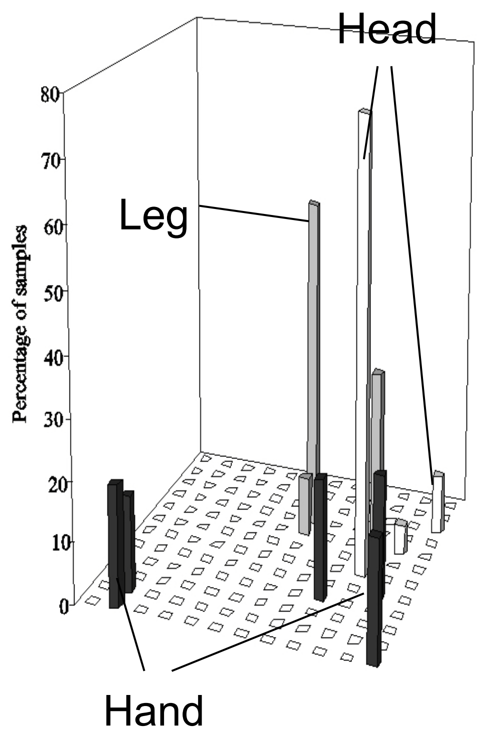
- Training the SOMs:
  - Learning Rule: Competitive learning [Nowlan 1990]
  - SOMs with size up to 13x13 neurons
  - Fixed learning rate at 0.2
  - Trained in batch mode
  - Lower level: 500 epochs of training
    Higher level: 200 epochs of training

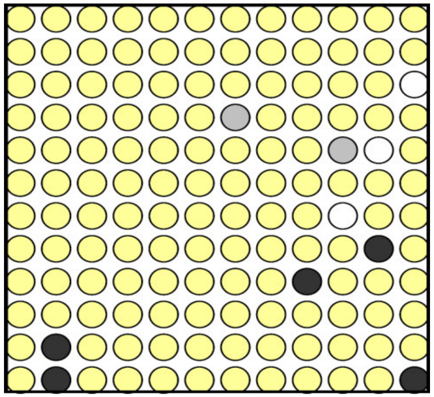


Results after learning

- Coarse clustering of verb representation
  - Clear clustering into three body parts
  - Test samples are located in the appropriate body-part regions
- Finding: corresponds to observations in neuroscience
  - ***Regional clustering*** emerges based on grounded sensor/motor features.

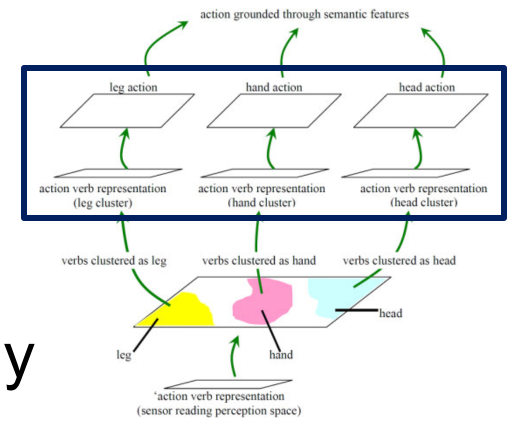




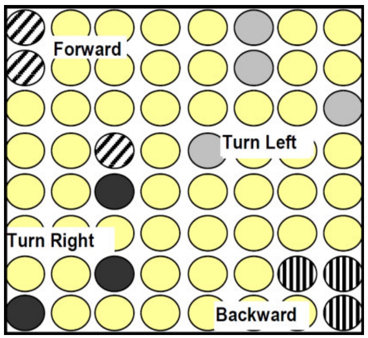


Results after learning (cont.)

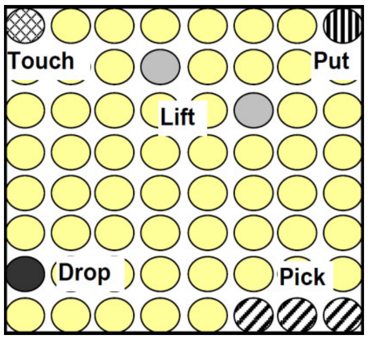
- Fine clustering of specific action words
  - Regional modularity
  - Multiple SOMs are performing subtasks



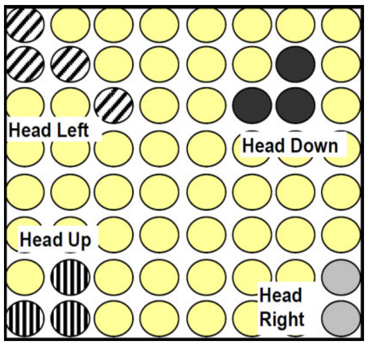
- Finding: clustered action verbs consistently
  - Model have shown a more fine granularity than the neuroscientific model of Pulvermüller 2003



Leg



Hand

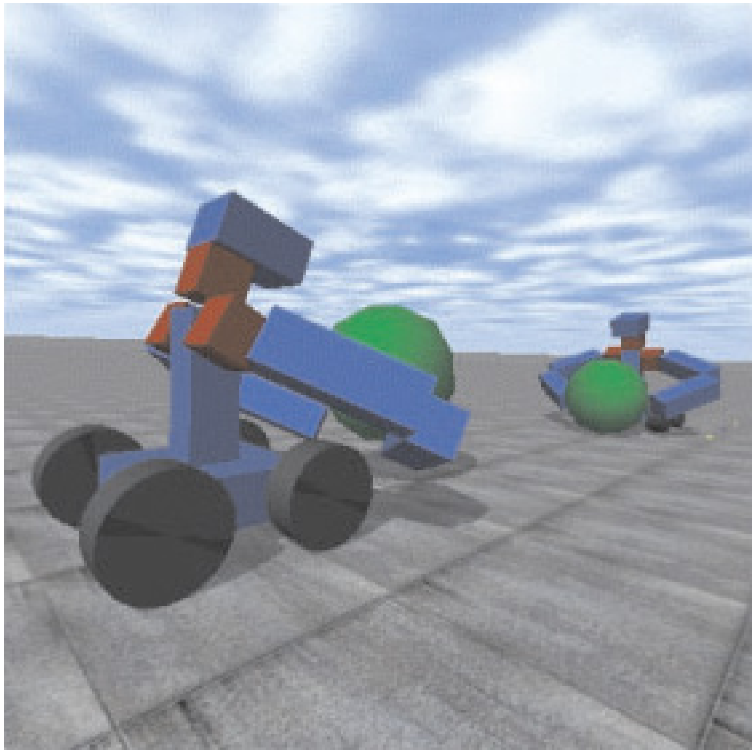


Head

Results after learning (cont.)

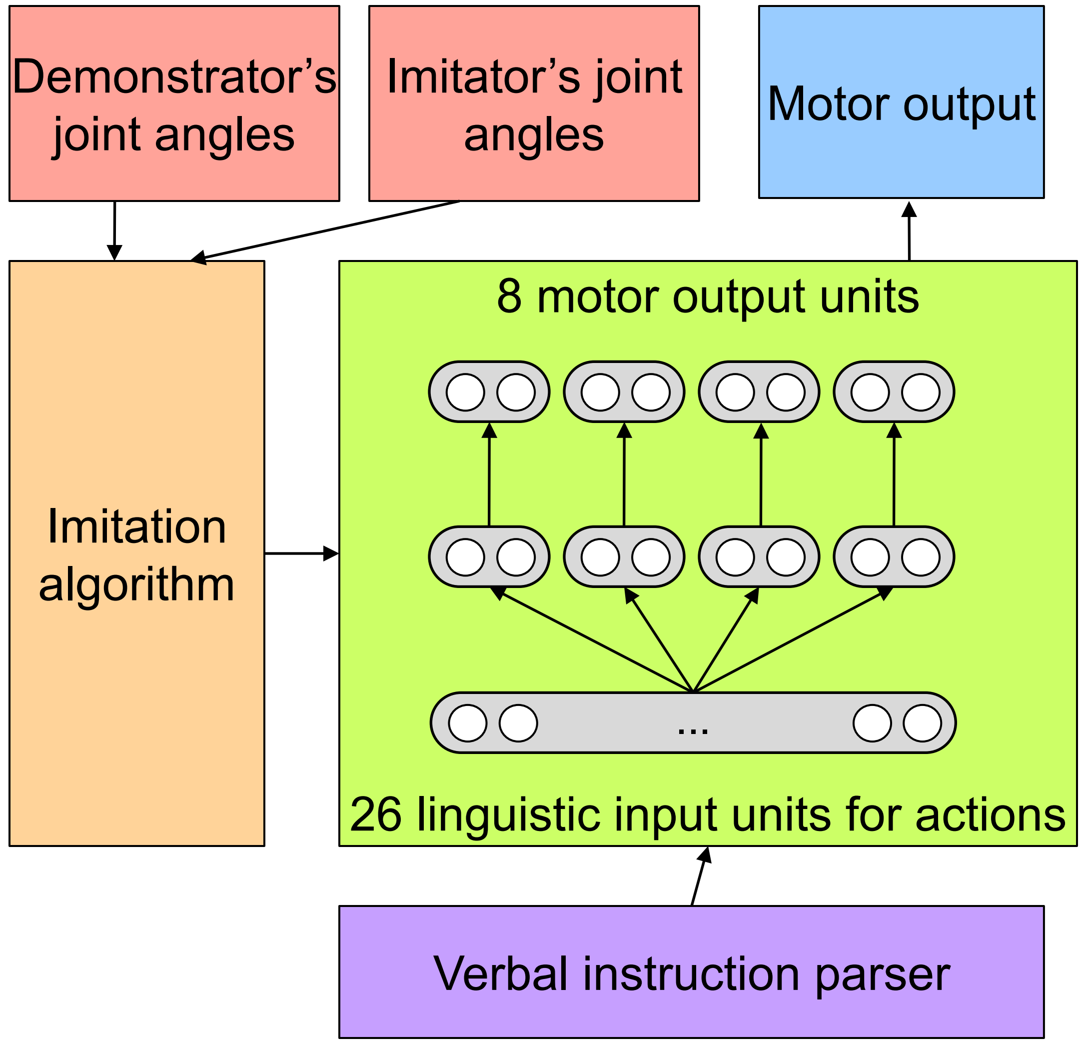
Grounding language in action: Towards grounding transfer (Cangelosi 2006)

- Embodied model for grounding of language in action
  - Epigenetic autonomous robot learns basic actions and names.
  - Reflect semantic ***associations*** at ***sensori-motor*** level
  - Linguistic abilities strictly depend on motor skills and behaviours
- Modelled with stick-figure robots based on neural network controller
  - Demonstrator and Imitator



Neural network controller

- Fully connected Feedforward NN
- Four pairs of motor neurons:
  - Left upper arm and forearm;
  - Right upper arm and forearm;
  - Shoulder and upper arm
  - Wheels



Language processing and imitation algorithm

- Symbolic parser filters linguistic input
  - Localist encoding of one word per linguistic input unit
- Estimate necessary force to apply to each motor joint

$$f(t+1) = f(t) + g\big(x(t), y(t)\big)$$

$$g\big(x(t), y(t)\big) = \alpha\left(\frac{2}{1+\exp\big(-2\beta\big(x(t)-y(t)\big)\big)}\right) - 1$$

  - Difference of motor output and estimate is reduced through error back propagation

$x(t)$: joint angles demonstrator

$y(t)$: joint angles imitator

$f(t)$: motor forces imitator

α : scale

β : gain

If difference between demonstrator and imitator is large so is the force applied

Learning: Basic grounding

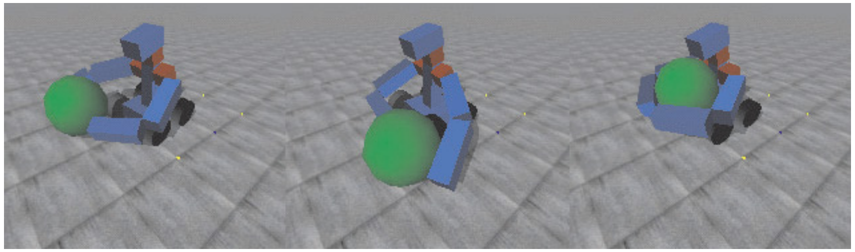
- Imitator learns to execute actions
  - Observing the demonstrator
  - Mimicking movement
  - Learns actions and names simultaneously
- Grounding of ***words*** in perception and production
- Training: 50 training epochs
- Transfer:
  - Link hidden units' activations of words
  - Next step: input the respective motor commands to the network

Actions/Words
Open_Left_Arm
Close_Left_Arm
Open_Right_Arm
Close_Right_Arm
Lift_Left_Arm
Lift_Right_Arm
Move_Forward
Move_Backward

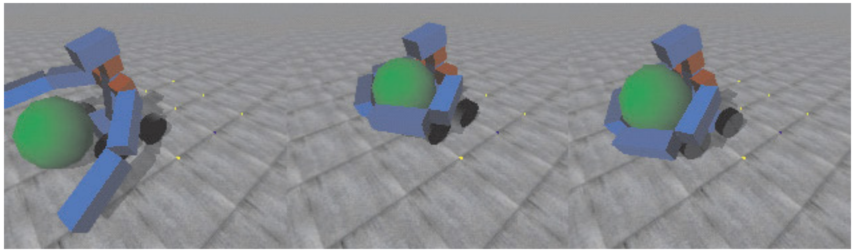
Learning: Higher order grounding

- Imitator learns behaviours
  - Combined actions
  - Linguistic descriptions in natural language
- Grounding of ***names*** and ***concepts*** of new actions
- Training: Up tp150 epochs due to increased difficulty

Behaviours
Grab
Push_Left
Push_Right
Open_Arms
Arms_up
Carry
Pull
Cheer

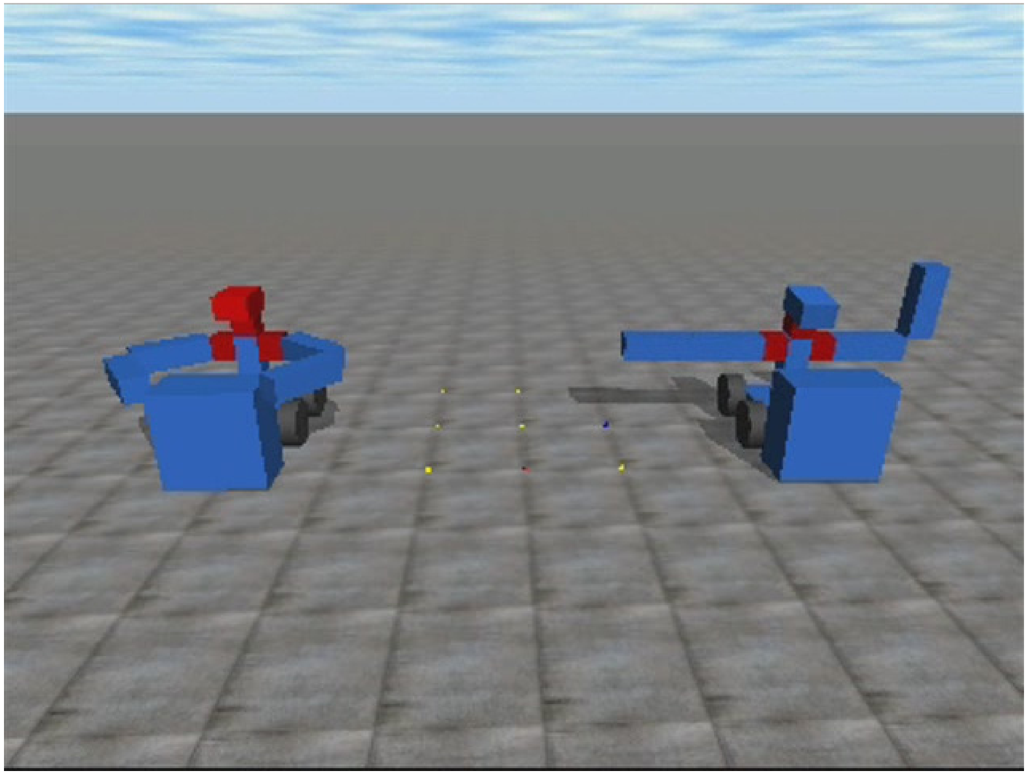


Example 1: `Grab Is Close_Left_Arm And Close_Right_Arm`



Example 2: `Carry Is Move_Forward And Grab`

Resulting behaviour



Left:
Demonstrator

Right:
Imitator

Centre for Interactive Intelligent Systems, University of Plymouth
http://www.tech.plym.ac.uk/Research/computer_science_and_informatics/

Current: Reinforcement learning for mapping instructions to actions (Branavan 2009, 2012)

Input

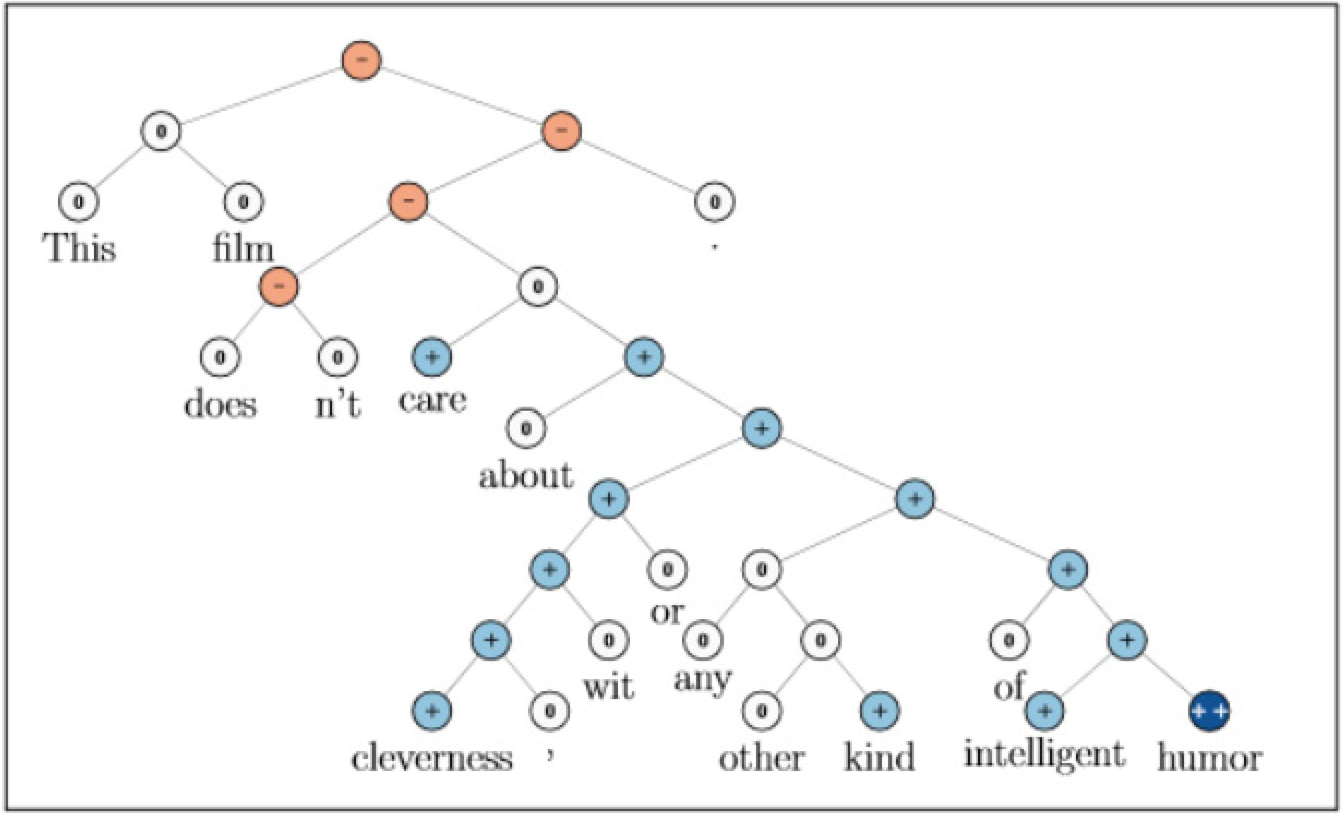
Instructions:
step-by-step
descriptions of actions

1. Click **Start**, point to **Search**, and then click **For Files or Folders.**
2. In the **Search for** box, type "msdownld.tmp"
3. In the **Look in** list, click **My Computer**, and then click **Search Now**.
4. ...

Target environment:
where actions
need to be executed

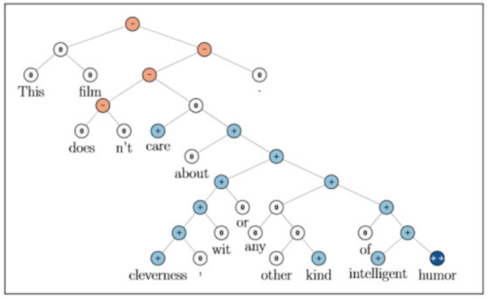
Research on deep neural language learning

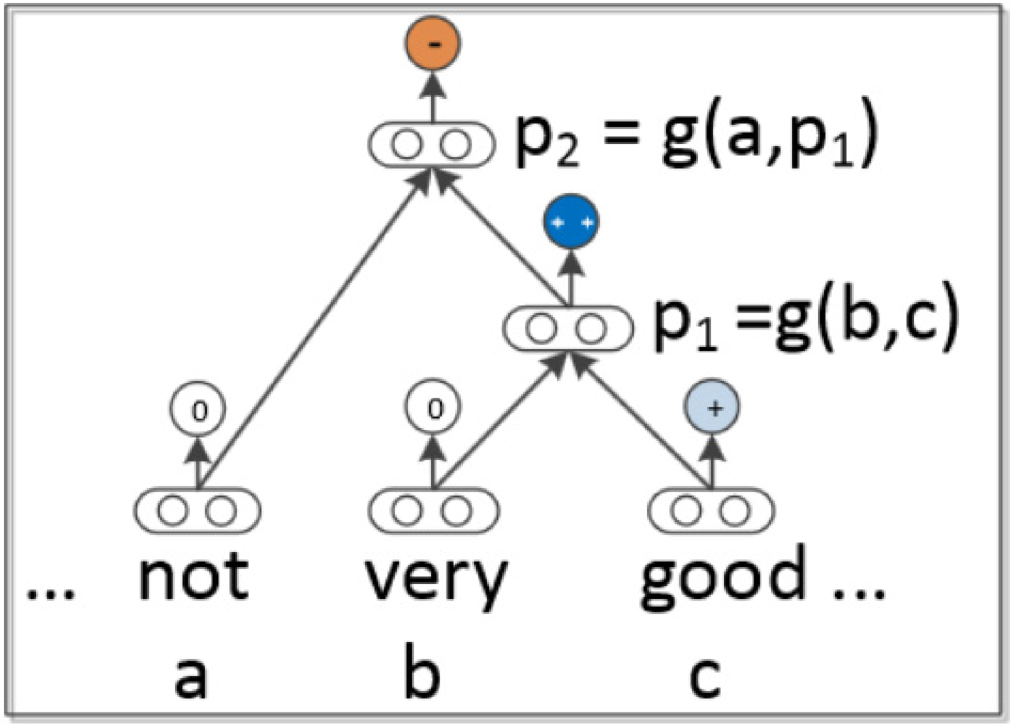
- Recursive deep models for semantic compositionality over a sentiment tree-bank (Socher et al 2013)



Research on deep learning and NLP

- Approach of Recursive Neural Network models for sentiment (Socher et al 2013)
- Compute parent vectors in a bottom up fashion using a compositionality function g and use node vectors as features for a classifier at that node



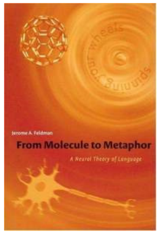


Summary

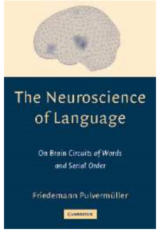
- Knowledge about phoneme classification and processing can be used for speech recognition
- Computational models can support and extend neuro-scientific hypotheses:
  - Sensor to actuator mapping
  - Semantic association
- Understanding language is a complex issue
  - Research into strong methods for understanding language processing in the human brain is still ongoing
  - Bio-inspired computational language processing can put our understanding about language to the next level

Further optional references

- Cangelosi, A. & Riga, T. An Embodied Model for Sensorimotor Grounding and Grounding Transfer: Experiments with Epigenetic Robots. *Cognitive Science*, 2006.



- Feldman, J.A. From Molecule to Metaphor: A Neural Theory of Language. MIT Press, 2006.
- Garagnani, M., Wennekers, T., & Pulvermüller, F. Recruitment and Consolidation of Cell Assemblies for Words by Way of Hebbian Learning and Competition in a Multi-Layer Neural Network. *Cognitive Computation*, Vol. 1(2), pp. 160–176, Springer, 2009.



- Pulvermüller, F. *The Neuroscience of Language: On Brain Circuits of Words and Serial Order*. Cambridge Univ. Pr., 2003.
- Wermter S., Weber C., Elshaw M., Gallese, V. Pulvermüller F. **Grounding Neural Robot Language in Action**. In Wermter S., Palm G., Elshaw M. *Biomimetic Neural Learning for Intelligent Robots*, pp. 162-181, 2005.

Outlook: Research at WTM