Model: TRoPICALS: a computational embodied neuroscience model of compatibility effects

Authors		
First name	**Middle name**	**Last name**
Daniele		Caligiore
Anna	M.	Borghi
Domenico		Parisi
Gianluca		Baldassarre
Brief Description *	Evidence shows that the perception of objects automatically activates the representation of their affordances. This is a neural-network model that reproduces these results and also provides a general framework to account for several other types of compatibility effects. The model is based on four general principles: (a) visual perception and action are organized along a dorsal neural pathway (encoding affordances) and a ventral pathway; (b) within the ventral pathway, the prefrontal cortex biases action selection based on context and goals; (c) action selection results from neural dynamic competitions that cause variable reaction times; (d) words trigger â€œinternal simulationsâ€ of their referents. The model was designed within a methodological approach that aims at developing it cumulatively so as to furnish increasingly general and comprehensive accounts of compatibility effects. The approach imposed four types of constraints on the model: (a) neuroscientific constraints on their architecture and functioning; (b) reproduction of specific psychological experiments; (c) functioning within an embodied system; (d) reproduction of the learning processes that result in the target behaviors.	
Narrative *	Behavioral and brain imaging evidence indicates that visually perceived objects activate motor information. Seeing objects or pictures of objects elicits the actions that tend to be performed on or with objects. In this perspective, the notion of affordance (Gibson, 1979) has been given new relevance. An affordance can be described as a set of properties of an object which suggest possible actions and uses to an organism. An affordance is not an intrinsic property of an object, but rather a relational property: an object may provide different affordances depending on the features of the organism body and of the context.Recent evidence shows that the perception of objects automatically activates the representation of their affordances. For example, some experiments found compatibility effects between the size of objects (small/large) and the kind of grip (precision/power) required to categorize them as natural or as artifacts, and between common location of object parts (top or bottom) and the kind of movement (up and down) required to indicate whether or not these parts belong to a whole object. The model was designed within a methodological approach that aims at developing it cumulatively so as to furnish increasingly general and comprehensive accounts of compatibility effects. The approach imposed four types of constraints on the model: (a) neuroscientific constraints on their architecture and functioning; (b) reproduction of specific psychological experiments; (c) functioning within an	

	embodied system; (d) reproduction of the learning processes that result in the target behaviors. The claim on the generality of the model is supported by a critical comparison with other models that are related to the above four principles and by an analysis of how the model could be developed to account for other compatibility effects. The heuristic power of the model is also shown by presenting two testable predictions.
Tags	affordance, grasping, language, sound, visual perception

Architecture

Diagrams

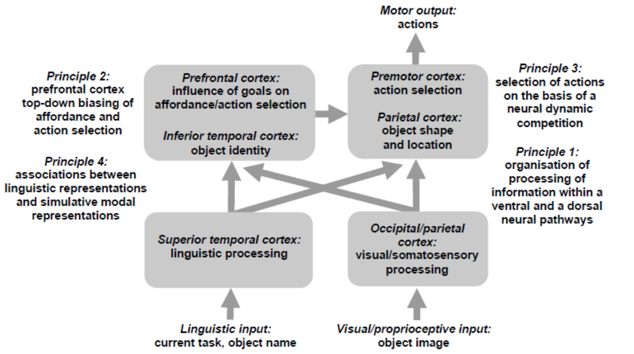


Figure 1: Abstract schema of the main functions implemented by TRoPICALS and the putative anatomical brain areas where they might be executed: (a) the occipital cortex and anterior parietal cortex which performs respectively visual pre-processing (e.g., edge detection) and somatosensory processing; (b) the inferior temporal cortex and the prefrontal cortex which respectively extracts information on identity of objects and biases action selection processes taking place in premotor cortex based on the current context and goal; (c) the parietal cortex and the premotor cortex which respectively extracts affordance information from objects and selects the desired final posture of the arm fingers and (sent to the muscle models, not shown in the figure); (d) the cortical areas involved in language processing, such as the superior temporal cortex involved in auditory processing and language comprehension. The graph also indicates the four main principles incorporated by the model

Abstract schema of the main functions implemented by TRoPICALS and the putative anatomical brain areas where they might be executed: (a) the occipital cortex and anterior parietal cortex which performs respectively visual pre-processing (e.g., edge detection) and somatosensory processing; (b) the inferior temporal cortex and the prefrontal cortex which respectively extracts information on identity of objects and biases action selection processes taking place in premotor cortex based on the current context and goal; (c) the parietal cortex and the premotor cortex which respectively extracts affordance information from objects and selects the desired final posture of the arm fingers and (sent to the muscle models, not shown in the figure); (d) the cortical areas involved in language processing, such as the superior temporal cortex involved in auditory processing and language comprehension. The graph also indicates the four main principles incorporated by the model

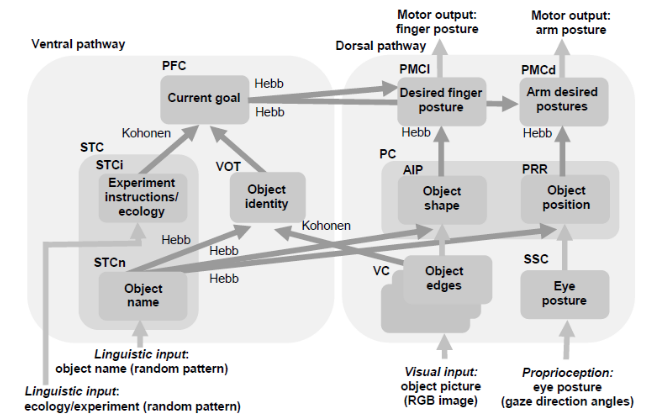


Figure 3: Architecture of TRoPICALS. The boxes indicate the components of the model. The label inside each box indicates the type of information encoded by the component, whereas the acronym at its top-left corner indicates the brain anatomical area putatively corresponding to it (the acronyms are explained in the label of Table 1). Light and dark grey arrows respectively indicate connections which were hardwired and connections which were updated by learning processes based on a Hebb covariance learning rule or a Kohonen learning rule. The input of the model is formed by three RGB visual neural maps (VC) and a somatosensory map (SSC). Downstream the VC and the SSC, the model divides into two main neural pathways: the dorsal pathway, which implements suitable sensorimotor transformations needed to perform action on the basis of perception, and the ventral pathway, which allows flexible control of behaviour thanks to the biasing effects exerted by the PFC on action selection. In turn, the dorsal pathway is formed by a pathway controlling grasping and a pathway controlling reaching.

Architecture of TRoPICALS. The boxes indicate the components of the model. The label inside each box indicates the type of information encoded by the component, whereas the acronym at its top-left corner indicates the brain anatomical area putatively corresponding to it (the acronyms are explained in the label of Table 1). Light and dark grey arrows respectively indicate connections which were hardwired and connections which were updated by learning processes based on a Hebb covariance learning rule or a Kohonen learning rule. The input of the model is formed by three RGB visual neural maps (VC) and a somatosensory map (SSC). Downstream the VC and the SSC, the model divides into two main neural pathways: the dorsal pathway, which implements suitable sensorimotor transformations needed to perform action

		on the basis of perception, and the ventral pathway, which allows flexible control of behaviour thanks to the biasing effects exerted by the PFC on action selection. In turn, the dorsal pathway is formed by a pathway controlling grasping and a pathway controlling reaching.
Inputs		
Name	**Data Type**	**Description**
Experimental Instructions	sound	Linguistic input of experiment instructions/ecology
Eye Posture	gaze angles	Used to determine view of object and to map it to the visual cortex.
Object Name	sound	Linguistic input of object names. In experiment with words: Similar in color shape and size to: apple, potato, grape, raisin, cup, mallet, eraser, needle. In experiment with object parts: Doll, Car, Head, Feet, Roof, Wheels, Plum, Mandarin, Head, Feet, Root, Wheels
Visual input	rbg image	Image of object to grasp. Similar in color shape and size to: apple, potato, grape, raisin, cup, mallet, eraser, needle.
Outputs		
Name	**Data Type**	**Description**
Arm posture	angle	Motor output to change the angle of the arm.
Figure posture	angle	Motor output to change the angle of a finger.
States		
Name	**Data Type**	**Description**
Current goal	array	The neurons of the prefrontal cortex (PFC) are assumed to encode information about the current goal of action depending on both the task (STCi) and the identity of the object (VOT), similarly to what is done in Deco and Rolls (2003
Desired arm posture	angle	Angles are mapped onto the two dimensions of the map and are â€œread outâ€ from the map as a weighted average of the position of neurons in it (suitably remapped onto the joint space) with the weights of the average corresponding to the activation of the neurons
Desired finger posture	angle	Angles are encoded by the two dimensions of the map. These are â€œread outâ€ from the map as a weighted average of the position of neurons within it, suitably remapped onto the angles space, with the weights of the average corresponding to the

		activation of the neurons
Eye Posture	array	The proprioception of the eye is encoded in terms of its gaze direction angles.
Object Edges	array	3 arrays indicating the object's edges based on the colors red, green, and blue. Each array is a 2d representation of the view with individual values in the range of [0,1].
Object Identity	array	The ventral occipito-temporal cortex (VOT) encodes the â€œidentity of objectsâ€□ on the basis of their colour and shape.
Object Names	text	The superior temporal cortex (STC) encodes both linguistic information about names of the seen objects and linguistic information about the instructions of the experiment. To differentiate these two kinds of information two neural maps were used which were assumed to represent two distinct sets of neuron clusters of STC: STCn, corresponding to neuron clusters encoding the (phonological aspects of the) names of objects, and STCi, corresponding to neuron clusters encoding the (phonological aspects of the) experiment instructions or, alternatively, the ecological conditions. Note that the distinction between STCn and STCi is a functional one, not an anatomical one, and is due to the lack of empirical data on where and how the nouns of objects and the experiment instructions are encoded within STC.
Object Position	array	A copy of the Somatosensory cortex gaze angles.
Object Shape	array	2D array with each node containing the average activation of the topologically correspondent RGB edge-encoding neurons of the visual cortex.
Superior temporal cortex for instructions (STCi)	array	The neurons of STCi are activated with either one of two random patterns (each pattern was formed by 20 randomly chosen neurons set equal to one and the rest set equal to zero).

Submodules

Name	**Description**
A neurodynamical cortical model of visual attention and invariant object recognition	A model of visual attention in complex natural scenes relies on adjusting the size of the receptive field of inferior temporal cortex (IT) neurons. An attractor network models IT where translation invariant representations of objects are stored. Eccentricity from the fovea, attentional bias, and complexity of the scene change the effective size of the receptive field and magnitude of the response of IT neurons.
Inferior Parietal Cortex	Serves as a repository for knowledge of skills, including information about stored sequences

Posterior Parietal Cortex	Guides the execution of skills that involve action sequences
Temporoparietal Cortex	Stores word sounds
Anterior Prefrontal Cortex	Involved in selection or retrieval of declarative memories, including lexical knowledge
Action recognition	The Action Recognition schema receives as input a distributed representation of the object and hand state match as well as the currently recognized sound. It outputs a vector of mirror neuron firing rates indicating the level of confidence that the grasps encoded by each neuron are observed or indicated by the auditory input.
Hand-Object spatial relation analysis schema	The Hand-Object spatial relation analysis schema receives object-related signals from the Object features schema, as well as input from the Object Location, Hand shape recognition and Hand motion detection schemas. Its output is a set of vectors relating the current hand preshape to a selected affordance of the object. The schema computes such parameters as the distance of the object to the hand, and the disparity between the opposition axes of the object and the hand.

Submodule: Inferior Parietal Cortex

Brief Description *	Serves as a repository for knowledge of skills, including information about stored sequences
Tags	

Submodule: Posterior Parietal Cortex

Brief Description *	Guides the execution of skills that involve action sequences
Tags	

Submodule: Temporoparietal Cortex

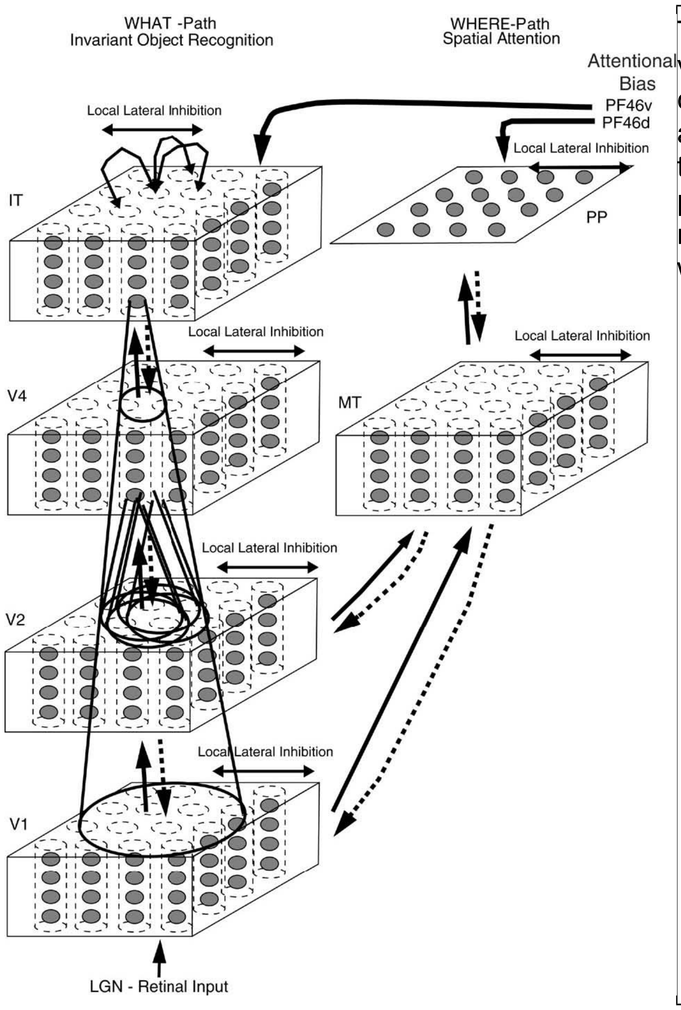
Brief Description *	Stores word sounds
Tags	

Submodule: F5

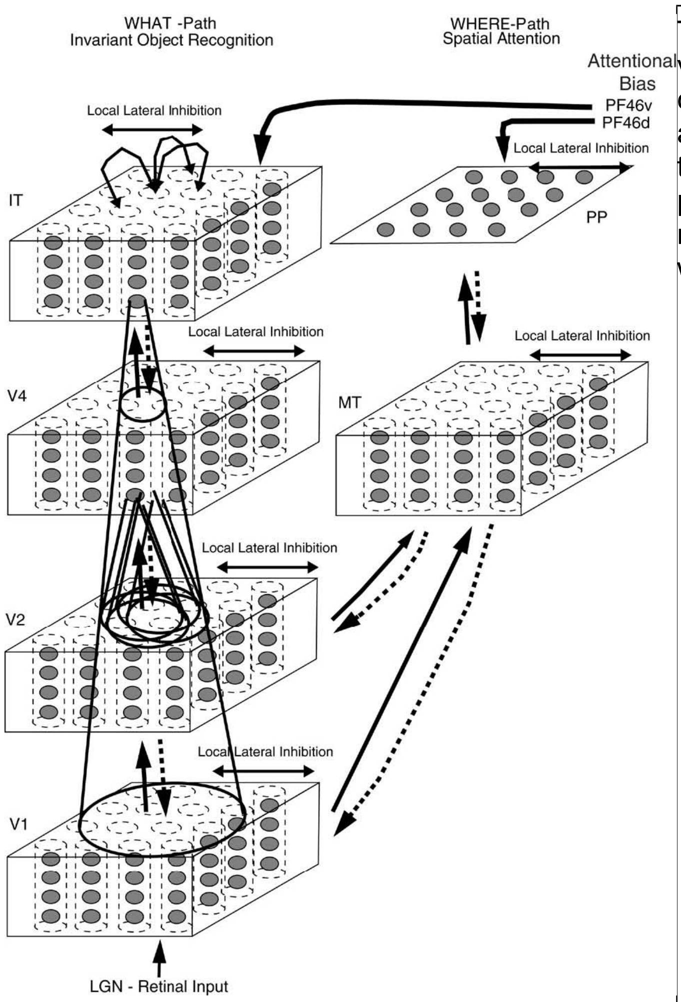
Brief Description *	The F5 module contains mirror neurons and responds during observation and performance of grasping
Tags	

Submodule: A neurodynamical cortical model of visual attention and invariant object recognition

Brief Description *	A model of visual attention in complex natural scenes relies on adjusting the size of the receptive field of inferior temporal cortex (IT) neurons. An attractor network models IT where translation invariant representations of objects are stored. Eccentricity from the fovea, attentional bias, and complexity of the scene change the effective size of the receptive field and magnitude of the response of IT neurons.
Tags	
Diagrams	



The system models the two known main visual pathways of the mammalian visual cortex. Main modules include areas V1, V2, and V4 of the visual cortex; the inferior temporal cortex (IT) and the posterior parietal cortex (PP). V1, V2, V4, and IT are mainly concerned with object recognition while PP handles spatial attention.



The system models the two known main visual pathways of the mammalian visual cortex. Main modules include areas V1, V2, and V4 of the visual cortex; the inferior temporal cortex (IT) and the posterior parietal cortex (PP). V1, V2, V4, and IT are mainly concerned with object recognition while PP handles spatial attention.

Inputs

Name	**Data Type**	**Description**
Attentional Bias	Scalar	Bias for target object.
LGN Retinal Input	Array	Simulated retinal input.

Outputs

Name	**Data Type**	**Description**
IT Firing Rate	Array	Firing rate for each neuron in inferior temporal cortex (IT).

Submodule: PF

Brief Description *	The PF module contains mirror neurons and associates visual representations of grasping in STS with motor representations in F5
Tags	

Submodule: Anterior Prefrontal Cortex

Brief Description *	Involved in selection or retrieval of declarative memories, including lexical knowledge
Tags	

Submodule: STS

Brief Description *	The STS module encodes a visual representation of actions and can distinguish between actions of the other and self-generated

	actions
Tags	

Submodule: Hand-Object spatial relation analysis schema

Brief Description *	The Hand-Object spatial relation analysis schema receives object-related signals from the Object features schema, as well as input from the Object Location, Hand shape recognition and Hand motion detection schemas. Its output is a set of vectors relating the current hand preshape to a selected affordance of the object. The schema computes such parameters as the distance of the object to the hand, and the disparity between the opposition axes of the object and the hand.	
Tags		
Inputs		
Name	**Data Type**	**Description**
Hand shape	Vector	Shape of the hand in terms of joint angles and finger positions
Object affordance location	3d vector	Location of the center of the opposition axis for the chosen affordance of the observed object, in hand-centered coordinates
Object geometry	Vector	Coarse coding of geometrical features of the observed object
Wrist position	3d vector	Position of the wrist in three dimensional space
Outputs		
Name	**Data Type**	**Description**
d	Scalar	Distance between the wrist and the chosen affordance of the observed object
o1	Scalar	Angle between the object axis and the(index finger tip-thumb tip)vector
o2	Scalar	Angle between the object axis and the (index finger tip-thumb tip) vector

Submodule: Action recognition

Brief Description *	The Action Recognition schema receives as input a distributed representation of the object and hand state match as well as the currently recognized sound. It outputs a vector of mirror neuron firing rates indicating the level of confidence that the grasps encoded by each neuron are observed or indicated by the auditory input.	
Tags		
Inputs		
Name	**Data Type**	**Description**
F5 canonical	Vector	Firing rate of cells determining which type of grasp to perform
Object-hand state match	Vector	Distributed representation of the object and hand state match

Sound recognition	Vector	Vector of firing rates uniquely identifying different sounds
Outputs		
Name	**Data Type**	**Description**
F5 mirror	Vector	Firing rate of cells indicating which grasp was recognized