We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

BOOK CITATION INDEX
CLARIVATE ANALYTICS
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Self-Organization of Object Categories in a Cortical Artificial Model

Alessio Plebe

Department of Cognitive Science, University of Messina
Italy

1. Introduction

Explaining the human capacity to construct mental categories is a central question in cognitive science, yet how reality is perceived and compartmentalized by our mind has puzzled philosophers since Aristotle. Today, categorization is believed to be a strategy humans have in common with many animals. In most species, individuals have the opportunity to directly experience only a tiny fraction of the entities, objects, features and events offered by their environments. By relying on categorization, organisms become able to determine which things in the world belong together or are alike in some way, therefore reacting in an appropriate manner, thanks to the previous experiences had with other instances of the same category.

The purpose of this chapter is to investigate computational models that can simulate mechanisms underlying the ability to develop categories. We argue that a neural mechanism for detecting and coding recurrent coincidences in stimuli could be the key component in our brain's ability to build the categories it uses to organize reality. The mathematical framework that has given the best interpretation so far on how signals with high levels of autocorrelation affect networks of neurons with local coincidence detection and coding abilities, is the one referred to as *self-organization*.

Being categorization a phenomena appearing at so many different cognitive levels, and that is pervasive in so many different modalities, we doubt that it could rely upon a single mechanism. Moreover, we believe that humans unlike any other animal reshape their categories and build new ones, by virtue of the information provided to them by language as soon as they have developed the most rudimentary understanding of it. Nevertheless, we do not see reasons for posing that the pre-linguistic development of categories in infants should happen as a result of mechanisms different from those found in other species. Even if we are convinced that different mechanisms might contribute, the basic mechanism briefly discussed is one we consider to play a key role.

1.1 The problem of mental categories

The research on categorization in humans can be traced back to traditional domains. In philosophy the objective of epistemology has included explaining the acquisition of knowledge, with empiricists attempting to formulate explanations in terms of mental mechanisms (Locke, 1690; Hume, 1739). It was within the field of psychology, however, that the search for this sort of mechanism became a central issue, employing the strategy of experimental analysis of the ontogenesis of categories in infants. While holding in high regard

the theoretical frameworks proposed within those two traditional disciplines in the past, we believe that the possibility of identifying the basic mechanisms of categorization is today more feasible thanks to recent advances in some specific domains. In particular, we are building on the following:

– Developmental psychology, that thanks to improved testing methodologies has produced an impressive amount of research on early categorization in the last 20 years.

– The neurobiology of learning, that has opened the door to the understanding of the mechanisms underlying the representational power of neural circuits.

– Computational models of emerging representations, that have contributed to understanding the modalities and constraints by which experiential inputs build representations in the brain.

Basing ourselves on evidence that has come from these three fields, we argue that a mechanism implementing coincidence detection of stimuli, and coding recurrent coincidences by self-organization, could be essential for the brain to build mental categories.

There are at least three crucial aspects of this mechanism. One is the special sensibility of organisms, in detecting that two or more elements in the stream of stimuli from the environment, happen with remarkable synchrony. A second is the ability of the organism to code in its nervous system traces of repeated coincidences of the detected elements. One effect of this internal representation, is that subsequent experiences of just one (or a few) of the coded elements, might elicit the activation of the internal representation of the other(s). Finally, in a population of neurons every reinforcement of the binding between two or more elements, should be compensated by reduction of other antagonistic elements. The combination of self-reinforcing interaction with competitive compensation may give rise, on the scale of cortical maps, to the emergence of self-organized representations corresponding to mental categories.

1.2 Computational models of categorization

The first attempt to simulate lexical semantics by means of self-organization was done by Ritter & Kohonen (1989). It was a very abstract model, in which names of animals and certain binary attributes (like "has feathers" or "likes to swim") were coded in vectors.

This approach was extended by Miikkulainen (1997), in a model that combined three SOM maps, two for orthography and phonology, and a semantic feature map. Associations between them were formed, based on co-occurrence of the lexical symbol and its meaning. The model was able to simulate dyslexic and category-specific aphasic impairments similar to those observed in human patients. More recently Li et al. (2004) proposed a model of early lexical categorization that combines two self-organizing maps as semantic map and phonological map, connected via associative links. Their model reproduced aspects of a growing lexicon during development. Mayor & Plunkett (2010) have also explored specific aspects of early lexical acquisition, using standard self-organizing maps. This model reproduces interesting aspects of language learning, like slips of the tongue and mispronunciation effects.

All these models, however, lack correspondence between their mathematics and how brain computations are performed, in particular, there is no account on how, on a large scale, processes relevant to categorization are distributed in the cortex.

The model here proposed is an attempt to simulate early phases of lexical categorization of visual objects, exploiting mathematical implementations of the self-organizing principle, in a way that takes into account the real structure of the cortex, with its pathways and hierarchies

of processes. In the next section, the core concept of the mechanism supporting categorization will be introduced, together with its available implementations and relevant theoretical issues. Section §3 will describe the model, with results shown in §4.

2. Artificial networks with self-organizing properties, and category formation

Computational models have been a valid complement to neurophysiology and neuroimaging in the enterprise of understanding how our brain works. Models, in abstracting biophysical phenomena at different levels, can help bridge neurophysiology with behavior and cognition. However, in the past the guiding principles of modeling cognitive functions were strongly based on the epistemological value of the algorithmic design, corroborated by the philosophical view of the mind as a computing machine, that prevailed in the 80's and '90s (Fodor, 1981; 1983). The serious weakness of this early form of computationalism was the risk of solving problems in our own head and then subsequently creating specific algorithms, that can be a far cry from how the brain solves the same problem.

Today, computational neuroscience has grown, reaching solid theoretical foundations Dayan & Abbott (2001), making it possible to constrain models, especially of low-level vision, in order to get relatively close to real brain mechanisms, avoiding the errors made by previously unconstrained computational models. However, the adherence to the known chemical and electrical processes of the nervous system is necessarily inverse to the degree of complexity of the problem at hand. In the case of categorization, even if limited to objects with visual salience, complexity is very high, and therefore, the replication of biology cannot be detailed. This means that the risk, mentioned above, of creating algorithms that are unrelated to actual brain functions, is always present. A good antidote for this, is to focus models on the repertoire of basic principles underlying cortical organization, and to study the ontogenesis of complex functions by implementing these principles in a simulated environment. This way of proceeding is characteristic of research done within a recent framework called neuroconstructivism (Quartz & Sejnowski, 1997; Westermann et al., 2007).

2.1 The mathematics of cortical development

In our view category formation is directly correlated with the development of cortical maps, in the meaning originally introduced by Mountcastle (1957), as a portion of the cortex unified functionally by the consistent responsiveness of its cells to contiguous sensorial stimulation. At this level, we consider the effect of being exposed to possible coincidences in a multitude of signals. In a sense, categories are ways of coding in neural aggregates, which signals co-occur or are co-incidental enough to be representative of the key traits of the objects or events that caused those signals. One of the most fruitful mathematical frameworks for simulating cortical development driven by the degree of autocorrelation of the incoming signals is the concept of self-organization. It is a vast topic, related to dynamic systems theory, with a long tradition in physics, physical chemistry, and astronomy (Haken, 1978). In general this framework studies the spontaneous formation of organized patterns in systems consisting of a large number of interacting elements, and initially in a relative undifferentiated state (Ball, 1999). In most organizing systems two fundamental facts are that global order can arise from local interactions, and that interactions work in two opposite directions: that of enhancing the degree of similarity (that in some systems can be interpreted intentionally as cooperation) between some elements, or by the differentiation between some other elements (intentionally read as competition). In the field of neural circuit development it was first introduced by von der Malsburg (1973), for modeling how the visual cortex can spontaneously develop its

mature organization, and was based exclusively on the local interaction of neurons, ruled by Hebb's principle. In the case of the development of orientation-sensitive patterns in the primary cortex, the model was based on the following system of differential equations:

$$\frac{\partial y_i}{\partial t} = -\alpha y_i + \vec{k}_i \cdot f(\vec{y}_i) + \vec{w}_i \cdot f(\vec{z}_i) + x_i, \tag{1}$$

$$\frac{\partial w_{i,j}}{\partial t} = \eta f(y_i) \left(f(z_j) - w_{i,j} \sum_{l \in \mathcal{Z}_i} f(z_l) \right), \tag{2}$$

where y and z are the activities of neurons in different layers of the same cortical area, the former in layer IV, where typically thalamocortical afferents project, the latter in other layers, where lateral intercortical interactions take place. $w_{i,j}$ is the connections strength between z_j and x_i, and the vector $\vec{w}_i$ is composed by all connections of intercortical neurons projecting to i. The vectors $\vec{y}_i$ and $\vec{z}_i$ are the activations of all neurons belonging to columns where intercortical connections with neurons i exist. Using the same convention for the corresponding scalars, $\vec{y}_i$ refers to cells in layer IV and $\vec{z}_i$ to cells in other layers. The constant kernel vector $\vec{k}$ modulates the activation of neighbors with the distance from i, typically with a Gaussian difference. The function f is the monotonic squeezing non-linearity, typically a sigmoid. In equation (2) the first term corresponds to Hebbian growth in response to coinciding activity on the presynaptic and postsynaptic sides of the connection between i and j. The second term, where $\mathcal{Z}_i$ is the set of cells out of layer IV connected to i, performs a competitive effect, ensuring that the sum of synaptic strengths converging on position i is kept constant at 1.

2.2 The SOM implementation

Kohonen (1982; 1995) has made the mechanism of self-organization popular, with his very simple but efficient SOM (*Self-Organizing Maps*) model.

The learning rule is on a *winner-takes-all* basis: if the input data are vectors $\vec{v} \in \mathbb{R}^N$, the SOM will be made of some M neurons, each associated with a vector $\vec{x} \in \mathbb{R}^N$ and a two dimensional (in vision applications) coordinate $\vec{r} \in \{< [0,1], [0,1] >\} \subset \mathbb{R}^2$. For an input v there will be a winner neuron w satisfying:

$$w = \arg \min_{i \in \{1,...,M\}} \{\|\vec{v} - \vec{x}_i\|\}. \tag{3}$$

The winner-takes-all mechanism is a mathematical substitution for the effect of lateral connections, which play a fundamental role in the emergence of complex response functions in the cortex. Once the winner is identified, neural vectors are updated during training using:

$$\Delta \vec{x}_i = \eta e^{-\frac{\|\vec{r}_w - \vec{r}_i\|^2}{2\sigma^2}} (\vec{v} - \vec{x}_i), \tag{4}$$

where η is the learning rate, σ the amplitude of the neighborhood affected by the updating. Often both η and σ are decreasing functions of the training epochs.

As said in §1.2 the SOM soon proved to be very effective in modeling categorization, at the lexical level as well, it is however, a significant departure from the behavior of biological cortical circuits. The winner-takes-all mechanism works as a mathematical substitution for the effect of lateral connections, but only assuming fixed connections and uniform neighborhoods. The input lacks a receptive field structure, therefore SOM cannot be used
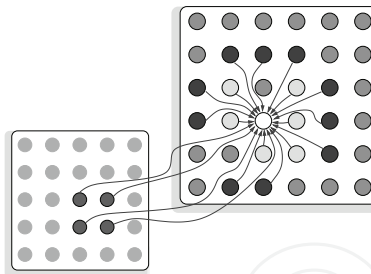
Fig. 1. The LISSOM architecture with lateral connections for one neuron (in white). Excitatory links are in pale gray and inhibitory links in dark gray.

as a building block in a hierarchy of cortical maps. However, there are concepts in SOM categorization that we believe to be relevant to brain structure as well. One is the mapping of meaning into two dimensions. The question concerning whether our disposition to better understand patterns depends on their being presented in a two dimensional manner due to the bidimensional structure of our cortex remains an open question (Chklovskii & Koulakov, 2004). Kohonen (1998) clearly defended the biological relevance of SOM-like maps in the cortex and proposed a modification of the SOM algorithm in order to relate the adaptation processes to physiological evidence in the brain, still preserving the winner-takes-all rule (Kohonen & Hari, 2000).

Several new mathematical formulations have been provided for learning mechanisms in populations of neurons in cortical maps, that still achieve competitive self-organization, but without the SOM winner-takes-all mechanism. Wallis & Rolls (1997) designed a model with explicit lateral inhibitory effect on neighbor units. Learning in the network is achieved with a special modification of Hebb's rule, in which a decaying trace of previous cell activity is kept. The trace rule can be interpreted as a way of detecting regularities between appearances of objects in short time sequences, therefore coincidental features that regularly happen when the same object is under view. This architecture has been successful in recognizing categories of artificial objects and faces (Rolls, 1992; Wallis & Rolls, 1997). This architecture is closer to the reality of cortical maps than the SOM is, but still, due to the fixed modeling of inhibition, and the lack of excitatory connections, may miss the full set of functional and developmental phenomena of a real cortical map.

2.3 The LISSOM architecture

An attempt to implement flexible and modifiable lateral connections, while keeping the architecture relatively simple, is represented by the model developed by Sirosh & Miikkulainen (1997) called LISSOM (*Laterally Interconnected Synergetically Self-Organizing Map*). In this model, each neuron is not only connected with the afferent input vector, but receives excitatory and inhibitory inputs from several neighbor neurons on the same map, as shown in Fig. 1

The activation level x_i of a neuron i at a certain time step k will now be given by:

$$x_i^{(k)} = f\left(\frac{\gamma_A}{1 + \gamma_N \vec{I} \cdot \vec{v}_{r_A,i}} \vec{a}_{r_A,i} \cdot \vec{v}_{r_A,i} + \gamma_E \vec{e}_{r_E,i} \cdot \vec{x}_{r_E,i}^{(k-1)} - \gamma_H \vec{h}_{r_H,i} \cdot \vec{x}_{r_H,i}^{(k-1)}\right),$$ (5)

where the vectors $\vec{x}_{r_E,i}^{(k-1)}$ and $\vec{x}_{r_H,i}^{(k-1)}$ are the activations of all neurons in the map, where a lateral connection exists with neuron i of an excitatory or inhibitory type, respectively. Their fields are circular areas of radius, respectively, r_E, r_H. Vectors $\vec{e}_i$ and $\vec{h}_i$ are composed by

all connection strengths of the excitatory or inhibitory neurons projecting to i. The vectors $\vec{v}$ and $\vec{x}_i$, as before, are the input and the neural code. The scalars γ_A, γ_E, and γ_H, are constants modulating the contribution of afferents. The map is characterized by the matrices $\mathbf{A}, \mathbf{E}, \mathbf{H}$, whose columns are all vectors $\vec{a}$, $\vec{e}$, $\vec{h}$ for every neuron in the map. The scalar γ_N controls the setting of a push-pull effect in the afferent weights, allowing inhibitory effects without negative weight values. Mathematically, it represents dividing the response from the excitatory weights by the response from a uniform disc of inhibitory weights over the receptive field of neuron i. Vector $\vec{I}$ is just a vector of 1's of the same dimension of $\vec{x}_i$. The function f is any monotonic non-linear function limited between 0 and 1, in this model is implemented as a piece-wise linear sigmoid.

The final activation value of the neurons is assessed after a certain settling time K, typically about 10 time steps. It is easy to recognize the resemblance of (5) with the original formulation of von der Malsburg in (1). All connection strengths adapt according to the general Hebbian principle, but include a normalization mechanism that counterbalances the overall increase of connections of the pure Hebbian rule. The equations are the following:

$$\Delta \vec{a}_{r_A,i} = \frac{\vec{a}_{r_A,i} + \eta_A x_i \vec{v}_{r_A,i}}{\left\| \vec{a}_{r_A,i} + \eta_A x_i \vec{v}_{r_A,i} \right\|} - \vec{a}_{r_A,i}, \tag{6}$$

$$\Delta \vec{e}_{r_E,i} = \frac{\vec{e}_{r_E,i} + \eta_E x_i \vec{x}_{r_E,i}}{\left\| \vec{a}_{r_E,i} + \eta_E x_i \vec{x}_{r_E,i} \right\|} - \vec{e}_{r_E,i}, \tag{7}$$

$$\Delta \vec{h}_{r_H,i} = \frac{\vec{h}_{r_H,i} + \eta_A x_i \vec{x}_{r_H,i}}{\left\| \vec{h}_{r_H,i} + \eta_A x_i \vec{x}_{r_H,i} \right\|} - \vec{h}_{r_H,i}, \tag{8}$$

where $\eta_{\{A,E,H\}}$ are the learning rates for the afferent, excitatory, and inhibitory weights. The normalization in equations (6), (7), (8), is a computational account of the biological phenomena of homeostatic plasticity (Turrigiano & Nelson, 2004).

The LISSOM model has been extensively used in vision studies, for modeling several low-level visual phenomena (Miikkulainen et al., 1997), the simulation of aftereffects (Bednar & Miikkulainen, 2000), perceptual grouping (Choe & Miikkulainen, 2004), face detection (Bednar & Miikkulainen, 2003), color response development (Bednar et al., 2005), motion direction selectivity (Bednar & Miikkulainen, 2006), and 3D object recognition (Plebe & Domenella, 2006; 2007). A comprehensive review of LISSOM applications is in (Miikkulainen et al., 2005). The LISSOM framework has been used by Plebe (2006) in a model of the ventral visual cortex, to simulate learning invariance in vision. Invariance is a kind of perceptual categorization, with the final goal of recognizing shapes as such, despite large changes in their appearance on the sensory surface. In this simulation it is simply learned, by detecting coincidences of activation patterns in lower visual maps, when belonging to same shapes. The model has been extended to the recognition of real objects (Plebe & Domenella, 2007), categorized by their visual features. In a more complex model, that includes visual and auditory pathways, Plebe et al. (2007b) have simulated the early acquisition of object names. The model was exposed to the coincidence of objects in its field of view, and the utterance of the category the object belonged to. A similar model was able to reproduce to some extent, some of the phenomena characteristic of early word learning, such as fast mapping and shape bias (Plebe et al., 2007a).

One difficulty in using LISSOM instead of SOM for studying categorization is in losing the interpretation of the role of each neuron in the map, that in the SOM can be achieved

immediately by labeling (see §4.3.2). The point is that LISSOM resembles a biological cortical map more closely than SOM, sharing with it the problem of the interpretation of neural firing, especially in higher-level maps. Research in this direction has recently drawn on the idea that the power of representing categories in cortical circuits lies in the combinations of activities of many columnar units. This concept is usually named "distributed coding" Hinton et al. (1986), but also "population coding", "vector coding" and "state space representation", in the formulation by Churchland (1989). In the current interpretation of population coding, a higher level map may code for a kind of object or fact, and is the concurrent level of firing of a population of cells in that map that represents a specific category of the kind. There is large evidence of population coding in higher cortical areas (Rolls & Tovee, 1995; Abbott et al., 1996; Pasupathy & Connor, 2002). In §4.3.1 we will use the population coding concept in analyzing the organization of the higher-level LISSOM maps in the model.

2.4 Theoretical issues in explaining categorization by self-organizing networks

We will now discuss several critical issues in the enterprise of shedding light on human categorization by modeling the neural mechanism of coincidence detection and self-organization.

There is a classical argument against basing the organization of categories on feature associations, in that categories instead could form by distilling a few features from many, the ones typical of the category. Quoting from Gibson & Levin (1975,p.23), "[association] does not work for perceptual learning, because what is learned is not addition of something but rather extraction of something". It is arguable that associating, in the standard theory, would be equivalent to adding something, certainly detecting coincidence and organizing their representation accordingly has nothing to do with the notion of addition, while its net effect is typically of coding the discrimination of some feature.

A deeper challenge is against the whole assumption, here made, that early categorization is perceptually driven.

For Mandler (2004), an important feature of human cognition is conceptual analysis, that is said to be parallel to, and different from, ordinary perceptual learning, which is associative in essence. We do not deny that in building categories humans most probably rely on a variety of skills that cannot be ascribed to our putative mechanism, even in a pre-linguistic phase. We believe, however, that the coincidence mechanism plays a key role, for other reasons. The first reason is that coincidence not necessarily implies only perceptual elements. Regular conjunctions can be detected between signals elicited by stimuli and other signals recalled by activations of representations already coded, this is a necessary step for generalizing categories over a number of instances experienced in time. The second reason comes from developmental science evidence that casts doubt on the separation between perceptual and conceptual categorization (Eimas & Quinn, 1994; Rakison, 2003). As a result, Mandler's hypothesis that conceptual learning is different in nature from purely associative learning, not necessarily requires a different notion than that of coincidence detection and self-organized coding. One possibility is that conceptual analysis of experience is based on the same sort of coincidence detection and coding, but involving phenomena that emerge in widespread cortical circuits, with self-organization at a more global scale.

A more general criticism against the mechanism here proposed could be the objection that self-organization of perceptual regularities alone cannot act as the foundation for mental categories, because of an insufficiency of constraints. Keil et al. (1998) give the suggestive example that "virtually every washing machine ever encountered is white, yet we know that

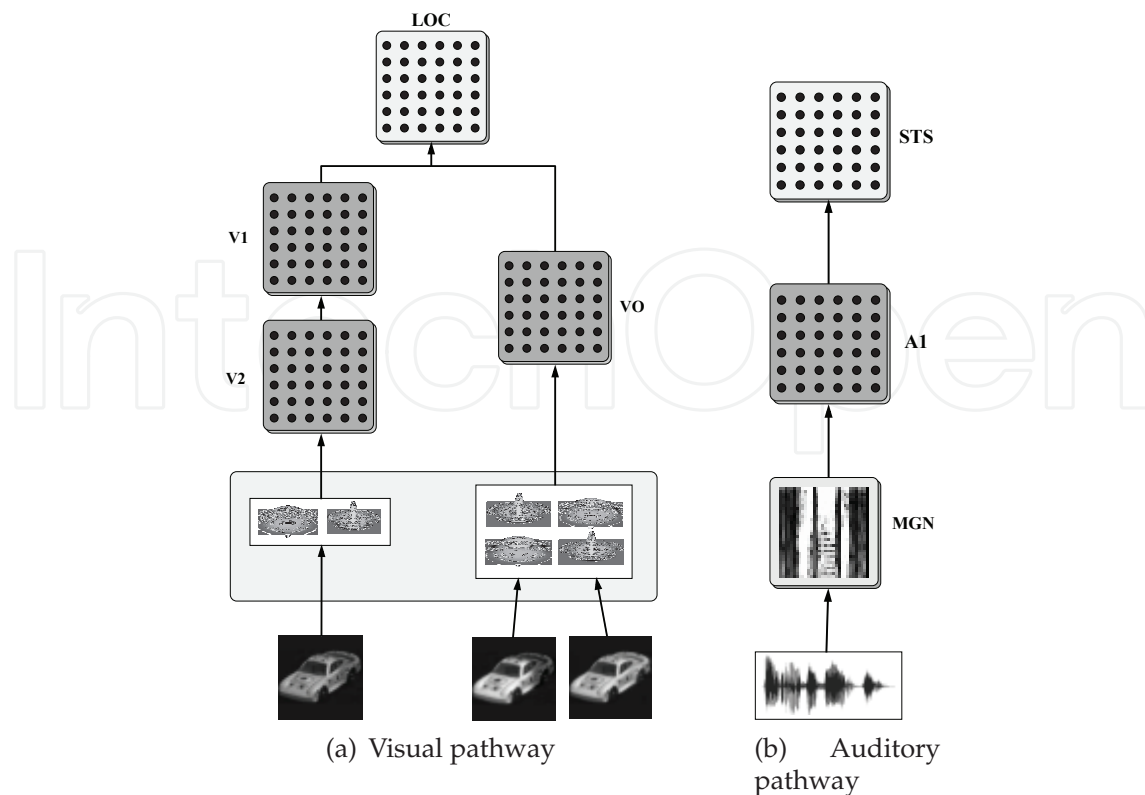(a) Visual pathway   (b)   Auditory pathway

Fig. 2. Scheme of the two pathways that make up the model.
All acronyms are listed in Tab. 1

whiteness is irrelevant in determining that an object is a washing machine.". This argument reveals to be poor when directed against the coincidence mechanism. In fact, it is entirely based on the assumption that perceptual features are equally neutral for the perceiver. Not only is this not the case, but coincidence detection and self-organizational coding is first and foremost the tool that allows the brain to rank the many features occurring with a category of objects or an event, according to how features are predictive of the category. Predictiveness is evaluated each time a next instance of the category is experienced, with consequent adjustment of the ranking of features. There is no doubt that early categories might include salient features that are not distinctive traits for the kind, in its normative semantics. We suspect that for many young children being white is indeed important for an object in order to be a washing machine, as important as being forbidden by their mothers. But this seems only to show that what is used to make a prediction depends on our global knowledge of the world, and that sometimes this knowledge could not be the best on the market.

3. Modeling categorization in the cortex

This section will introduce the model, the rationale for its design, describe each component, and discuss in detail the higher component of the model, where a SOM or a LISSOM module will be used and compared.

3.1 Overall structure of the model

The overall model is a combination of artificial cortical sheets that reproduce essentially the part of the brain involved in the categorization of objects with perceptual salience: the visual and the auditory pathways. Each of the two processing streams is fed by realistic stimuli, and therefore also comprises, in an approximate simulation, the components that transduce the external signals, and the subcortical components. An outline of the modules that make up the visual and auditory pathways is shown in Fig. 2, the names of the component are in Tab. 1, together with their parameters as in equation (5). The two higher cortical maps in the visual and the auditory paths, LOC and STS, will carry the best representation coded by the models on object visual features and word sound features. The way of combining the two representations is varied in the model, with the alternative of including a higher LISSOM-type artificial cortical map, or a standard SOM module. While in the visual and auditory pathways the criteria of adhering realistically to the structure of the human cortex is possible, in higher categorization correlation between functions and cortical areas becomes more fuzzy and uncertain. Categorization processes are by definition abstract representations, where what is coded at a neural level has no obvious correspondence with features in the sensorial inputs. Furthermore, in the mapping between words and objects, the native language acts in a complex way, with its own established category structure, that initially mismatches with the emerging mental categories of the developing mind, making it hard to trace back to the level of neural processes. It is therefore, justifiable to adopt at this level a very abstract neural architecture such as the SOM.

3.2 The visual pathway

The visual path is built upon and extends a previous model of visual object recognition (Plebe & Domenella, 2007). The visual system encompasses the ventral stream only, in the broad partition established by Ungerleider & Mishkin (1982), and commonly accepted in current vision science. They proposed the well known dichotomy of "what" and "where" pathways, in that the ventral pathway is specialized for object recognition whereas the dorsal pathway is specialized for spatial perception. While the "where" has been quite controversial and has been subject to other proposed interpretations (Goodale & Milner, 1992), there has always been a large consensus on the purpose of the "what" pathway. For the purposes of this model, therefore, limiting ourselves to the ventral path only, is a valid approximation. One of the simplifications used in the model is the processing of sensorial input coming from one side only. Stereopsis contributes to the perception of 3-dimensional shapes, and is therefore,

acronym	full name	size	r_A	r_E	r_H	γ_A	γ_E	γ_H	γ_N
LGN	Lateral Geniculated Nucleus	112	2.6	-	-	-	-	-	-
MGN	Medial Geniculated Nucleus	36	2.2	-	-	-	-	-	-
V1	Primary Visual Cortex	96	8.5	1.5	7	1.5	1.0	1.0	0.0
A1	Auditory Primary Cortex	24	6.5	1.5	5.0	1.7	0.9	0.9	0.0
V2	Secondary Visual Cortex	30	7.5	8.5	3.5	50.0	3.2	2.5	0.7
VO	Ventral Occipital	30	24.5	4	8	1.8	1.0	1.0	0.0
LOC	Lateral Occipital Complex	16	6.5	1.5	3.5	0.5	1.1	1.6	0.0
STS	Superior Temporal Sulcus	16	2.5	1.5	5.5	1.8	1.0	1.2	0.0
PFC	Pre-Frontal Cortex	24	2.5	1.5	5.5	1.5	3.2	4.1	0.1
ACM	Abstract Conceptual Map	18	-	-	-	-	-	-	-

Table 1. Legend of acronyms of maps used in the model, and their main parameters.

a piece that is missing in the model. It is, however, much less useful than 2-dimensional contours when there is a single, unmoving object in the foveal view of a scene. This is the case of the samples used in our experiments.

This pathway includes the thalamic module, that is not the object of this study and is therefore hardwired, according to what is known about its functions. The governing equations for the subcortical components are the following:

$$x^{(\odot)} = f\left(\left(\vec{l}_{r_A} + \vec{m}_{r_A}\right) \cdot \left(\vec{g}_{r_A}^{(\sigma_N)} - \vec{g}_{r_A}^{(\sigma_W)}\right)\right) \tag{9}$$

$$x^{(\circledcirc)} = f\left(\left(\vec{l}_{r_A} + \vec{m}_{r_A}\right) \cdot \left(\vec{g}_{r_A}^{(\sigma_W)} - \vec{g}_{r_A}^{(\sigma_N)}\right)\right) \tag{10}$$

$$x^{(R^+G^-\odot)} = f\left(\vec{l}_{r_A} \cdot \vec{g}_{r_A}^{(\sigma_N)} - \vec{m}_{r_A}\vec{g}_{r_A}^{(\sigma_W)}\right) \tag{11}$$

$$x^{(R^+G^-\circledcirc)} = f\left(\vec{l}_{r_A} \cdot \vec{g}_{r_A}^{(\sigma_W)} - \vec{m}_{r_A}\vec{g}_{r_A}^{(\sigma_N)}\right) \tag{12}$$

$$x^{(G^+R^-\odot)} = f\left(\vec{m}_{r_A} \cdot \vec{g}_{r_A}^{(\sigma_N)} - \vec{l}_{r_A}\vec{g}_{r_A}^{(\sigma_W)}\right) \tag{13}$$

$$x^{(G^+R^-\circledcirc)} = f\left(\vec{m}_{r_A} \cdot \vec{g}_{r_A}^{(\sigma_W)} - \vec{l}_{r_A}\vec{g}_{r_A}^{(\sigma_N)}\right) \tag{14}$$

where x is the activation of a neuron, without indexes for clarity. There are two distinct subcortical pathways, one achromatic, processed by equations (9), (10), and another sensitive to colors, limited here to medium and long wapath velengths. The equations are: (11), (12), (13), (14). The symbol $\odot$ refers to on-center type receptive fields, and symbol $\circledcirc$ to off-center receptive fields. The profile of all visual receptive fields is given by differences of two Gaussians $\vec{g}^{(\sigma_N)}$ and $\vec{g}^{(\sigma_W)}$, with $\sigma_N < \sigma_W$. This is an approximation of the combined contribution of gangliar cells and LGN (Dowling, 1987).

The lower cortical maps included in the model are V1, V2 and VO, the primary, secondary, and ventral-occipital visual areas. Their equations are the following:

$$x^{(V1)} = f\left(\gamma_A^{(V1)}\left(\vec{a}_{r_A}^{(V1\leftarrow\odot)} \cdot \vec{x}_{r_A}^{(\odot)} + \vec{a}_{r_A}^{(V1\leftarrow\circledcirc)} \cdot \vec{x}_{r_A}^{(\circledcirc)}\right) + \right.$$
$$\left. \gamma_E^{(V1)}\vec{e}_{r_E}^{(V1)} \cdot \vec{x}_{r_E}^{(V1)} - \gamma_H^{(V1)}\vec{h}_{r_H}^{(V1)} \cdot \vec{x}_{r_H}^{(V1)}\right) \tag{15}$$

$$x^{(V2)} = f\left(\gamma_A^{(V2)}\vec{a}_{r_A}^{(V2\leftarrow V1)} \cdot \vec{x}_{r_A}^{(V1)} + \gamma_E^{(V2)}\vec{e}_{r_E}^{(V2)} \cdot \vec{x}_{r_E}^{(V2)} - \gamma_H^{(V2)}\vec{h}_{r_H}^{(V2)} \cdot \vec{x}_{r_H}^{(V2)}\right) \tag{16}$$

$$x^{(VO)} = f\left(\gamma_A^{(VO)}\left(\vec{a}_{r_A}^{(VO\leftarrow R^+G^-\odot)} \cdot \vec{x}_{r_A}^{(R^+G^-\odot)} + \vec{a}_{r_A}^{(VO\leftarrow R^+G^-\circledcirc)} \cdot \vec{x}_{r_A}^{(R^+G^-\circledcirc)} + \right.\right.$$
$$\left. \vec{a}_{r_A}^{(VO\leftarrow G^+R^-\odot)} \cdot \vec{x}_{r_A}^{(G^+R^-\odot)} + \vec{a}_{r_A}^{(VO\leftarrow G^+R^-\circledcirc)} \cdot \vec{x}_{r_A}^{(G^+R^-\circledcirc)}\right) +$$
$$\left. \gamma_E^{(VO)}\vec{e}_{r_E}^{(VO)} \cdot \vec{x}_{r_E}^{(VO)} - \gamma_H^{(VO)}\vec{h}_{r_H}^{(VO)} \cdot \vec{x}_{r_H}^{(VO)}\right) \tag{17}$$

In all equations the activation x has to be taken as the activation of a generic i-th neuron of that level, and all receptive fields have to be intended as referring to that neuron, the index i has been omitted for clarity. For the same reason, the indication of the radius r of the circular receptive fields has not been superscripted with the indication of the layer, for example in equation (16) all r_A should be read $r_A^{(V2)}$, as the radii are different in each map.

The use of still pictures justifies the lack of a V3 area, whose main role is the processing of motion information (Gegenfurtner et al., 1997). In exploiting the modularity of the model, enacted by the correspondence with cortical maps, a simplification has been introduced by way of the separation of the processing of shape and color. Shape is elaborated through V1 and V2, by equations (15) and (16), and the processing of color is entrusted to VO, with equation (17).

There is evidence, in fact, that suggests that in the visual system no segregation of functions such as shape or color processing takes place, and that almost all visual cortical maps cooperate in analyzing form, color, motion and stereo information (Van Essen & DeYoe, 1994; Schiller, 1996). On the other hand, it is clear that visual areas are not equally involved in all aspects of object recognition. It is possible to identify specialization in one main function in certain maps, although the same maps might also marginally extend to other operations, and conversely, other maps may overlap in the elective function of that map as well. This is the case in what has been called the color center area by Zeki (1983a;b), who named it "V4". The precise localization of this area in humans, as well as its name, have been the object of a yet unsettled controversy. Sereno et al. (1995) redefined the borders of a color area in humans in the fusiform gyrus, and called it "V4v" ("v" for ventral), Hadjikhani et al. (1998) argued for an anterior area , which they called "V8". Here we are using the more general name of VO (Ventral Occipital), given by Wandell et al. (2005); Brewer et al. (2005) to the cluster of maps in the fusiform gyrus involved in color processing. V1 is the well-known primary visual cortex. It has been the most studied part of the brain (Hubel & Wiesel, 1962; 1959). One of its main functions is the organization of the map into domains of orientation tuned neurons (Blasdel, 1992; Vanduffel et al., 2002), which are fundamental for early shape analysis. For the reasons discussed above, our model discards the contributions of V1 to color processing (Landisman & Ts'o, 2002) and ocularity (Miller et al., 1989). The main projection from V1 is to its immediately anterior area, V2, and this connection is maintained in the model. V2 is almost as large as V1, but much less understood. A generally shared idea is that V2 is responsible for shape analysis at a level of complexity and scale larger than that of V1 (Kobatake & Tanaka, 1994). This is compatible with findings of V2 cells responding to end-lines and corners (Heider et al., 2000), gratings spirals and other complex shapes (Hegdé & Van Essen, 2000), Glass patterns Smith et al. (2007), and, with shared stronger evidence, angles (Ito & Komatsu, 2004; Anzai et al., 2007).

Shape and color processing paths converge in a higher visual map, with the following equation:

$$
\begin{aligned}
x^{(\mathrm{LOC})} = f\Bigg(& \gamma_{\mathrm{A}}^{(\mathrm{LOC})} \left(\vec{a}_{r_{\mathrm{A}}}^{(\mathrm{LOC} \leftarrow V2)} \cdot \vec{x}_{r_{\mathrm{A}}}^{(V2)} + \vec{a}_{r_{\mathrm{A}}}^{(\mathrm{LOC} \leftarrow VO)} \cdot \vec{x}_{r_{\mathrm{A}}}^{(VO)} \right) + \\
& \gamma_{\mathrm{E}}^{(\mathrm{LOC})} \vec{e}_{r_{\mathrm{E}}}^{(\mathrm{LOC})} \cdot \vec{x}_{r_{\mathrm{E}}}^{(\mathrm{LOC})} - \gamma_{\mathrm{H}}^{(\mathrm{LOC})} \vec{h}_{r_{\mathrm{H}}}^{(\mathrm{LOC})} \cdot \vec{x}_{r_{\mathrm{H}}}^{(\mathrm{LOC})} \Bigg)
\end{aligned}
\tag{18}
$$

The LOC model map corresponds to the area in the human cortex thought to be crucial for the task of recognition in vision. Malach et al. (1995) first identified this region, an area located anterior to Brodmann's area 19, near the lateral occipital sulcus, extending into the posterior and mid fusiform gyrus and occipital-temporal sulcus, with an overall surface size similar to V1. They called it "lateral occipital complex", where the term "complex" denotes the uncertainty regarding whether this region is a single visual map or a cluster of several maps. Being an object of research that is still quite recent, a clear picture of LOC characteristics and its role in vision is still lacking, however, the impetus given by its discovery

has lead to an outstanding body of knowledge. Perhaps the most important idea, one that has obtained a certain amount of consensus, is that this area is involved in visual behavior in which recognition is the main task (Grill-Spector et al., 1998; 1999; Vuilleumier et al., 2002; Kourtzi et al., 2003; Weigelt et al., 2007).

3.3 Auditory pathway

A division in dorsal and ventral streams, similar to that in the visual channel, has been found in the auditory system as well, with two streams originating in the caudal and rostral parts of the auditory cortex, targeting spatial and non-spatial domains of the frontal lobe (Romanski & Goldman-Rakic, 1999). Here again, curiously, the "where" interpretation is debated (Belin & Zatorre, 2000), but not the "what" processing path. This is the only contribution taken into account in the auditory system of this model.

As in the ventral pathway, a simplifications used in the model is the processing of sensorial input coming from one side only. This is not too much of a limitation for the auditory system, since biaural hearing supports the "where" processing but not the "what" detection. An outline of the modules that make up this pathway is shown in Fig. 2(b). The auditory path includes as thalamic component the medial geniculate nucleus, with the following equation:

$$x_{\tau,\omega}^{(\square)} = \left| \sum_{t=t_0}^{t_M} v(t)w(t-\tau)e^{-j\omega t} \right|^2 \tag{19}$$

where the symbol $\square$ refers to the spectrotemporal representation of the auditory signal, the horizontal dimension τ is time, and the vertical dimension ω is frequency. Function $w(\cdot)$ in 19 is a short term temporal window, that performs a spectrogram-like response, similar to that given by the combination of cochlear and MGN processes (Brown, 2003).

The cortical modules in this pathway include the auditory primary cortex, and the superior temporal sulcus, the equations are the following:

$$x^{(A1)} = f\left(\gamma_A^{(A1)} \vec{a}_{r_A}^{(A1\leftarrow\square)} \cdot \vec{x}_{r_A}^{(\square)} + \gamma_E^{(A1)} \vec{e}_{r_E}^{(A1)} \cdot \vec{x}_{r_E}^{(A1)} - \gamma_H^{(A1)} \vec{h}_{r_H}^{(A1)} \cdot \vec{x}_{r_H}^{(A1)} \right) \tag{20}$$

$$x^{(STS)} = f\left(\gamma_A^{(STS)} \vec{a}_{r_A}^{(STS\leftarrow A1)} \cdot \vec{x}_{r_A}^{(A1)} + \gamma_E^{(STS)} \vec{e}_{r_E}^{(STS)} \cdot \vec{x}_{r_E}^{(STS)} - \gamma_H^{(STS)} \vec{h}_{r_H}^{(STS)} \cdot \vec{x}_{r_H}^{(STS)} \right) \tag{21}$$

Very little is known about the computational organization of the auditory primary cortex, compared to the early visual maps just discussed (Linden & Schreiner, 2006). Our model discards biaural interaction, and preserves the main connectivity from single cochlear signals in the Medial Geniculate nucleus to A1. A large body of evidence points to an organization of A1 with a fundamental dependency on sound frequencies along one cortical dimension, and a distribution of neural responses to temporal properties (Atzori et al., 2001; Miller, Escab, Read & Schreiner, 2002; Winer et al., 2005), model A1 has frequency and time as map dimensions.

Equation (21) states the projection from the primary auditory cortex to STS. This is the model's correlate of a region in the cortical ventral auditory stream, on which there is accumulating evidence and a convergence of opinion on its role in representing and processing phonological information (Belin et al., 2000; 2004; Liebenthal et al., 2005; Hickok & Poeppel, 2007).

3.4 Higher categorial maps

The higher part of the model is implemented in two versions, as shown in Fig. 3. In version (a) the two higher cortical maps in the visual and the auditory paths, LOC and STS, are the

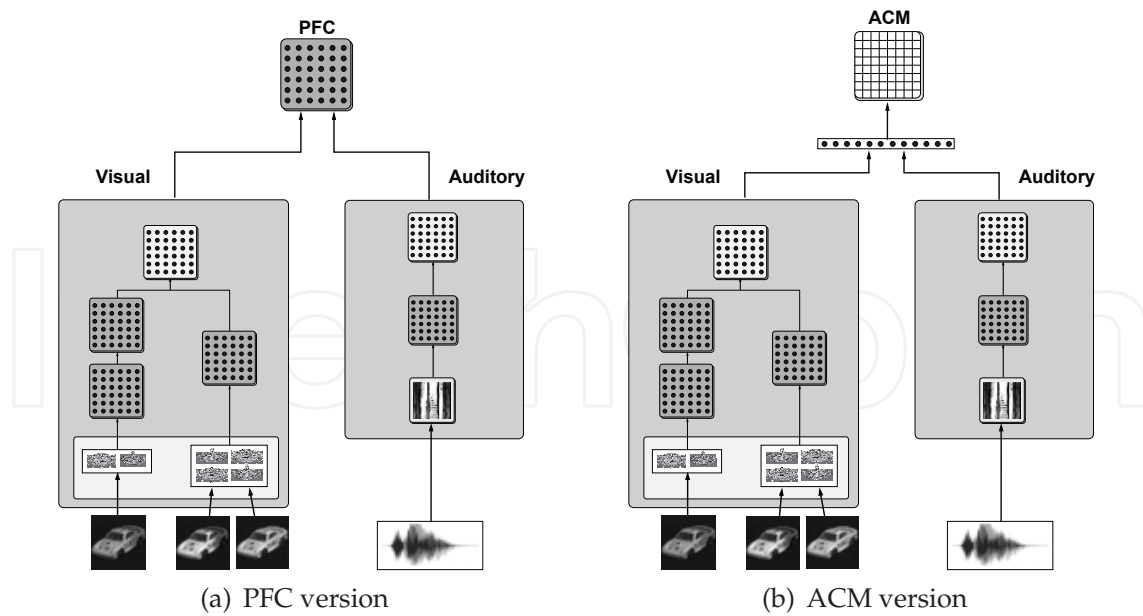(a) PFC version   (b) ACM version

Fig. 3. Scheme of the complete model, in two versions. In version (a) LOC and STS converge to PFC (PreFrontal Cortex), of type LISSOM. Version (b) aggregates LOC and STS contents in a single vector, presented to ACM (Abstract Conceptual Map), of type SOM.

afferents to a cortical map, still modeled with LISSOM, called PFC. In version (b) all neural activations in LOC and STS are aggregated in a single vector, that is fed to an abstract SOM type component, called ACM. Therefore there are the following two different equations for the higher cortical maps in the model:

$$x^{(PFC)} = f\left(\gamma_A^{(PFC)}\left(\vec{a}_{r_A}^{(PFC\leftarrow LOC)} \cdot \vec{x}_{r_A}^{(LOC)} + \vec{a}_{r_A}^{(PFC\leftarrow STS)} \cdot \vec{x}_{r_A}^{(STS)}\right) + \right.$$
$$\left. \gamma_E^{(PFC)}\vec{e}_{r_E}^{(PFC)} \cdot \vec{x}_{r_E}^{(PFC)} - \gamma_H^{(PFC)}\vec{h}_{r_H}^{(PFC)} \cdot \vec{x}_{r_H}^{(PFC)}\right) \tag{22}$$

$$w^{(ACM)} = \arg\min_{i\in\{1,...,M\}}\left\{\left\|\begin{bmatrix}\vec{x}_{r_\infty}^{(LOC)}\\\vec{x}_{r_\infty}^{(STS)}\end{bmatrix} - \vec{x}_i^{(ACM)}\right\|\right\} \tag{23}$$

There are actually several areas where visual and auditory signals converge, and more than one area activated in categorization tasks, justifying the solution of an abstract architecture like the SOM as the final map, with equation (22). Its relevance in principle for the brain categorization mechanism has been discussed in §2.2.

The other assumption, made with equation (22), is to expect a specific area to be at least partially involved in aspects of abstraction and integration of object information typical of categorization, even if not the only area where this process takes place. This is a role which can be ascribed to biological PFC, Freedman et al. (2003) have suggested that PFC could also be a crucial component of high-level visual categorization. While ITC (Inferior Temporal Cortex) is more involved in the analysis of viewed shapes, lateral PFC could form a more abstract kind of categorization (Miller, Freedman & Wallis, 2002; Fuster, 2002; Huey et al., 2006). Needless to say, biological PFC deals with larger sets of object information than those used in the model, and these would include for example, the functional information pertaining to what objects can be used for.

The choice of ACM allows a more abstract, and powerful way of categorizing information processed by the visual and auditory pathways, at the price of losing the topological information pertaining to the LOC and STS maps, and population coding organization, that can be studied in the model PFC.

4. Results of the model

In this section we will describe how the model has been trained in this experiment, we will report on the functions developed in all the lower cortical maps of the model, and subsequently discuss the various aspects of categorization that emerged in the higher maps. For this purpose additional methods of analysis are required, that will be described. In the upper map, representative of the joint visual and linguistic aspects of object categorization, PFC and ACM results will be compared.

4.1 Simulation of inner and environmental experiences

The model has been exposed to a variety of stimuli, in different stages of its development, that to some extent parallel periods of human development from the pre-natal stage to that of early language acquisition. Initially only V1, VO and A1 maps are allowed to modify their synaptic weights, by equations (6), (7), and (8). The stimuli presented to V1 and VO are synthetic random blobs that mimic waves of spontaneous retinal activity, that are known to play a fundamental role in the ontogenesis of the visual system (Mastronarde, 1983; Katz & Shatz, 1996; Thompson, 1997; Gödecke & Bonhoeffer, 1996; Chapman et al., 1996). Blobs presented to V1 are elongated along random directions, to stimulate orientation selectivity. Blobs to VO are circular, with constant hues, and random size, position, and intensity. The A1 map is exposed to short trains of waves sweeping linearly around a central frequency. Time durations, central frequencies and sweeping intervals are changed randomly. The next period of development involves V2 and STS maps as well. The visual stimuli comprises pairs of elongated blobs with a coinciding end point, to enhance the experience of patterns that are slightly more complex

bird	boat	bottle
can	car	cheese
chewinggum	chicken	cigarettes
cup	dog	drink
dummy	frog	fruit
hanger	jam	jug
kitten	medicine	mug
pepper	piece	pig
plug	pot	pottery
ring	sauce	soap
spoon	sweets	tape
telephone	toast	tomato
tool	truck	

Fig. 4. The samples used as stimuli for the models, on the bottom all 100 objects of the COIL-100 image library, in their frontal view, on the top the 38 names of categories in which the objects has been classified.

V1   VO   V2
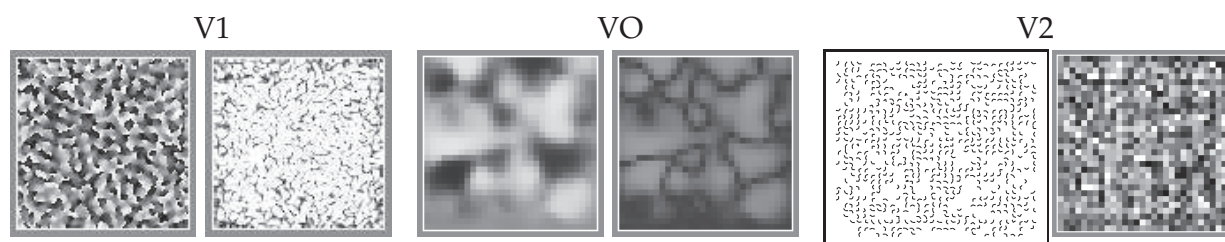

Fig. 5. Organization of neural responses in the lower maps of the model. The two maps on the far left show orientation selectivity in V1. Of these, the map at the far left shows the preferred orientation for each neuron, coded in gray level from black→horizontal to white→vertical. The map on the right contains the degree of neuronal selectivity in a gray scale from black→equally responsive to all orientations, to white→highly selective to a single orientation. The two middle maps represent the hue-responding domains in VO. The map on the far left encodes the preferred hue, from black→red to white→green, the rightmost map is again the degree of selectivity. The pair of maps on the far right show angle selectivity in V2. The map on the far left is overlaid with patterns inducing peak responses in the neurons, the map on the far right is the degree of selectivity.

than lines, such as corners. The auditory stimuli are synthesized waves of the 7200 most common English words (from `http://www.bckelk.uklinux.net/menu.html`), with length of 3 to 10 characters. All words are converted from text to waves using *Festival* software (Black & Taylor, 1997), with cepstral order 64 and a unified time window of 2.3 seconds. In the development stage that corresponds to that just after eye opening, natural images are used. In order to include the primary and most realistic difficulty in recognition, which is the identification of an object under different views, the COIL-100 collection has been used (Nayar & Murase, 1995), where for each of the 100 objects, 72 different views are available. In most experiments, unless otherwise stated, only 8 views per object have been used during the learning phase of the model, and all 72 views are used in the testing phases.

The last stage of the experiment simulates events in which an object is viewed and a word corresponding to its basic category is heard contemporaneously. The 100 objects have been grouped manually into 38 categories. Certain categories, such as `cup` or `medicine` have 5 exemplars in the object collection, while others, such as `telephone`, have only one exemplar. Each category word is converted from text to waves using the `en1` "Roger" male voice, and the `us1` female American speaker in the *Festival* software. Both male and female utterances are duplicated at standard and slower speeds, using the 1.3 value of the `Duration_Stretch` parameter in *Festival*. The 100 objects and the names of the categories can be seen in Fig. 4.

4.2 Emergence of organization in the lower maps

At the end of development, different organizations are found in the lower maps that enable the performance of processes that are essential to recognition, and that are similar to those found in corresponding brain areas. The model's V1 map organized orientation selectivity, with responsiveness of neurons to oriented segments arranged over repeated patterns of gradually changing orientations, broken by few discontinuities, as shown in Fig. 5.

This ordering resembles the one known to be found in biological primary cortex (Blasdel, 1992; Vanduffel et al., 2002). In the VO map of the model, most neurons respond to specific hues, regardless of intensity. This is one of the basic features of color processing. Color constancy is crucial in object recognition and is known to develop somewhere between two and four months of age (Dannemiller, 1989). The organization in the model's VO is shown in Fig. 5.

The kind of mapping found in A1 is typically tonotopic, and it encodes the dimensions of frequency and time sequences in a sound pattern. This is known to be the main ordering of neurons in biological A1 (Verkindt et al., 1995). The main organization found in the V2 map is responsiveness to angles, especially in the 60 and 150 degrees range, this is shown in Fig. 5. This kind of selectivity is one of the major phenomena recently discovered in biological V2 (Ito & Komatsu, 2004; Anzai et al., 2007). We refer to other published works for further details on the functions that emerged in V1 and VO (Plebe & Domenella, 2007) and V2 (Plebe, 2007).

4.3 Categorization and coding analysis

In the scope of this experiment the most interesting results are found in the three higher maps. It is here where we expect types of categorization to take place that concern both visual and word forms, and mostly conceptual categorization at the boundary between the two. In order to understand how a categorial representation about world objects can be carried by a sheet of artificial neurons, it is necessary to apply mathematical tools which will be introduced below.

4.3.1 Population coding

A common method of analysis has been carried out for the upper maps of LISSOM type to reveal possible distributions of neural activation as population coding of categories, as discussed in §2.3. Let us introduce the following function:

$$x_i(s) : S \in \mathcal{S} \to \mathbb{R}; \quad s \in S \in \mathcal{S}, \tag{24}$$

that gives the activation x of a generic neuron i in one of the higher maps, in response to the presentation of the stimulus s to the system. This stimulus is an instance of a class S, belonging to the set of all classes of stimuli $\mathcal{S}$ available in the experiment. For a class $S \in \mathcal{S}$ we can define the two sets:

$$X_{S,i} = \left\{ x_i(s_j) \mid s_j \in S \right\}; \quad \overline{X}_{S,i} = \left\{ x_i(s_j) \mid s_j \in S' \neq S \in \mathcal{S} \right\}. \tag{25}$$

We can then associate to the class S a set of neurons in the map, by ranking it with the following function:

$$r(S,i) = \frac{\mu_{X_{S,i}} - \mu_{\overline{X}_{S,i}}}{\sqrt{\frac{\sigma_{X_{S,i}}}{|X_{S,i}|} - \frac{\sigma_{\overline{X}_{S,i}}}{|\overline{X}_{S,i}|}}}, \tag{26}$$

where μ is the average and σ the standard deviation of the values in the two sets, and $|\cdot|$ is the cardinality of a set. Now the following relation can be established as the population code of a class S:

$$p(S) : \mathcal{S} \to \left\{ \langle i_1, i_2, \cdots, i_M \rangle \mid r(S,i_1) > r(S,i_2) > \cdots > r(S,i_M) \right\}, \tag{27}$$

where M is a given constant, typically one order of magnitude smaller than the number of neurons in the map, and α is a constant that is close, but smaller, than one. The population code $p(S)$ computed with (27) can be used to classify a stimulus s in an expected category:

$$c(s) = \arg\max_{S \in \mathcal{S}} \left\{ \sum_{j=1 \cdots M} \alpha^j x_{p(S)_j}(s) \right\}, \tag{28}$$

where $p(S)_j$ denotes the j-th element in the ordered set $p(S)$. It is possible to evaluate how the population code of a map is effective in discriminating a category S by measuring the fraction

of hits in classifying stimuli belonging to that category:

$$a(S) = \frac{|\{s \mid c(s) \in S\}|}{|S|}, \tag{29}$$

where $|\cdot|$ is the cardinality of a set.

4.3.2 SOM coding

We will provide here some of the mathematical details used in our interpretation of the results in the ACM map, that we consider most relevant to lexical categorization. The ACM is a SOM map, and at the end of its development can be analyzed by labeling its neurons with elements from the set of objects and names.

Since the only SOM map is ACM, and its purpose is to organize categories, the labeling function could be directly specified:

$$l^{(\mathrm{ACM})}(i) = \arg\max_{C \in \mathcal{C}} \left\{ \left| \left\{ c = \langle o, u \rangle \in C = \left(\bigcup_{O \in \mathcal{O}_C} O \right) \times U_C \mid i = w^{(\mathrm{ACM})}(c) \right\} \right| \right\}, \tag{30}$$

where $w_i^{(\mathrm{ACM})}(c)$ is computed by (23), when the utterance u is presented to the auditory path, and the object o to the visual path. The 38 object categories introduce a partition in the set of object's views $\mathcal{O}$, such that all sets of views in the partition, $O \in \mathcal{O}_C$ are of objects of that category C, while the set $U_C \subset \mathcal{U}$ collects all utterances naming the category C.

From the labeling functions the possibility of judging the correctness of the identification or the categorization of an object presented to the model immediately follows, by simply checking the label of the winner neuron in the ACM map. It is also possible to estimate the accuracy of the categorization simply by weighing the number of cases where the category or the object has been classified as the prevailing one in each neuron of the SOM, as done by the following equation:

$$a^{(\mathrm{ACM})}(C) = \frac{\left| \left\{ c \mid l^{(\mathrm{ACM})}(c) = w^{(\mathrm{ACM})}(c) \right\} \right|}{|C|}. \tag{31}$$

4.4 Categorization in the top visual and auditory maps

It is possible to begin evaluating the type of representation established in LOC, by using as stimuli view o of an object O, therefore specializing the equation (24) as:

$$x_i^{(\mathrm{LOC})}(o) : O \in \mathcal{O} \to \mathbb{R}; \quad o \in O \in \mathcal{O}, \tag{32}$$

where $x_i^{(\mathrm{LOC})}(o)$ is computed by (18), when the image o is presented to the visual path. In this case a category S of (24) is just an individual object, whose instances can be ascribed to the same entity, or mistaken for a different entity, if a specific point of view does not resemble the typical view of that object. Samples of population coding of individual objects are shown in Fig. 6. As can be seen from the comparison of the population coding map with LOC responses to different views of the same object, most coding neurons exhibit invariant responses. Invariance, the ability to recognize known objects despite large changes in their appearance on the sensory surface, is the most important property to fulfill the requirement for an object-recognition cortical area. It has been identified in human LOC by several studies (Grill-Spector et al., 1999; Vuilleumier et al., 2002; Kourtzi et al., 2003; Weigelt et al., 2007).
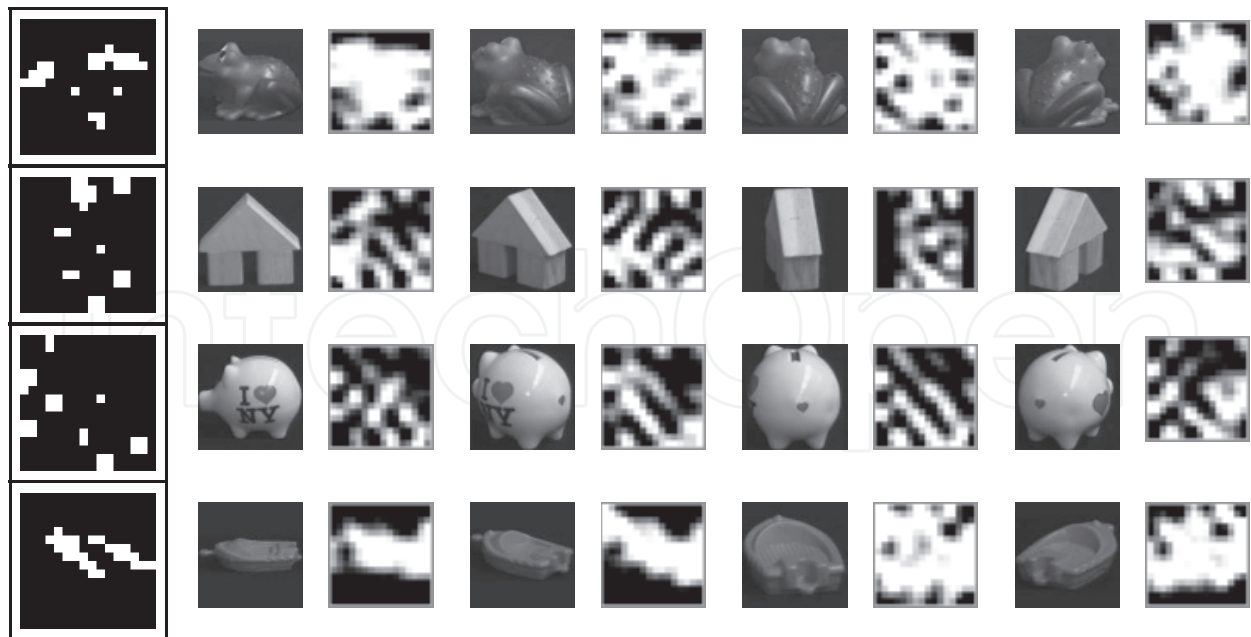
Fig. 6. Examples of population coding of objects in the LOC map. In each row the map on the far left displays the coding neurons. The following images in the row, are samples of views of the same object, and the corresponding response patterns in LOC.

Although the term invariance is commonly used in vision, a better term would be "degree of tolerance" with respect to classes of changes in appearance of the same object, since in no cortical map are responses absolutely invariant to transformations of the same object (Cox et al., 2005). There is also psychophysical evidence of inabilities to recognize objects viewed from less common perspectives, even in adult humans (Farah, 1990). In the model's LOC the supposed coding by population of individual objects is efficient in discriminating most of the views for many objects, although not in an absolute way. By means of equation (29), adapted to stimuli conditions described in (32), the discrimination power of LOC has been quantitatively assessed, and is shown in Tab. 2. In order to test the ability of the representation to cope with invariance, two versions of the model have been trained, exposing them to just a few views for each object only. During the test phase, all 72 views per object have been used. As a comparative figure, discrimination by chance of this task would be of 0.01.

In the STS map it is possible to evaluate the population coding of words, as heard in four different voices, therefore using stimuli that specialize equation (24) as:

$$x_i^{(STS)}(u) : U \in \mathcal{U} \to \mathbb{R}; \quad u \in U \in \mathcal{U}, \tag{33}$$

where $x_i^{(STS)}(u)$ is computed by (18), when the utterance u is presented to the auditory path.

# of views	discriminative power	
	average over all objects	standard deviation
4	0.611	0.334
8	0.638	0.315

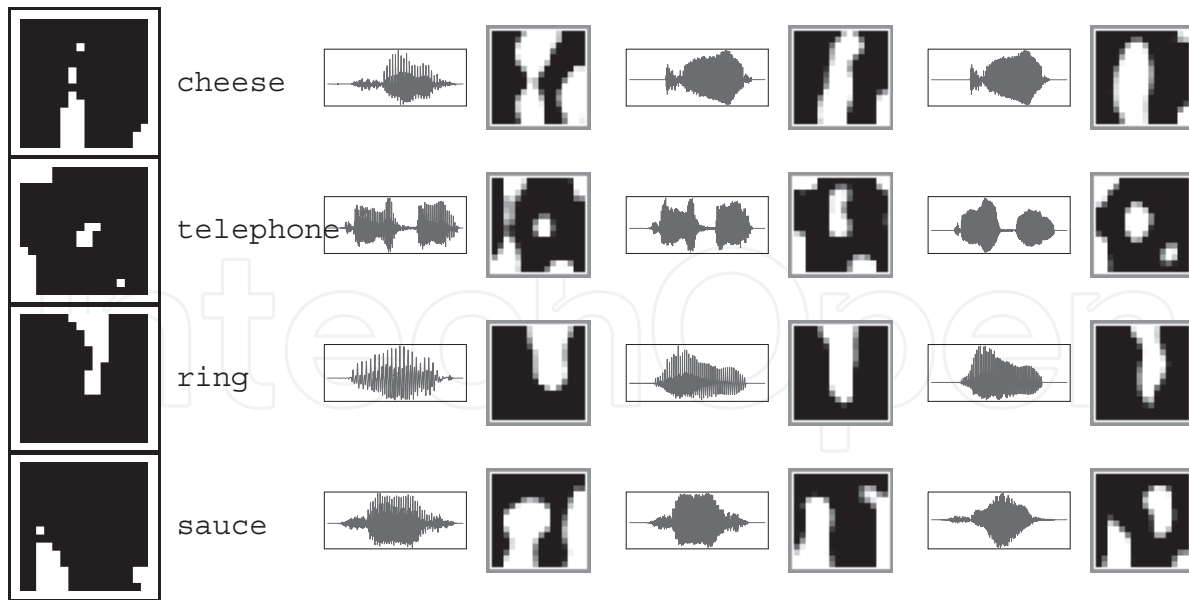Table 2. The model LOC's discriminative power of individual objects by population coding.

Fig. 7. Examples of population coding of word sounds in the STS map. In each row the leftmost map displays the coding neurons. The following images in the row are samples of sound of the same name, and corresponding response patterns in STS.

There are 38 classes U of sounds in $\mathcal{U}$, corresponding to the names of 38 object categories. Samples of population coding of names are shown in Fig. 7. In Tab. 3 the accuracy of STS

voices	discriminative power	
	average over all names	standard deviation
female	0.882	0.242
male	0.895	0.234
both	0.658	0.300

Table 3. The model STS's discriminative power of words heard by different voices.

in discriminating between the different names is shown. The development of STS has been experimented using three different sets of stimuli: female voices only, male voices only, or using the full set of voices. Listening to voices of a single gender facilitates the identification of names, as can be observed in children. As a comparative figure, discrimination by chance in this task would be 0.026.

4.5 Lexical categorization in PFC or ACM
Lexical categorization in the model takes place where the visual and the linguistic information meet, implemented with the two alternatives of PFC or ACM maps. In the case of PFC population coding is the computed specializing equation (24) as:

$$x_i^{(\text{PFC})}(c) : C \in \mathcal{C} \to \mathbb{R}; \quad c = \langle o, u \rangle \in C = \left(\bigcup_{O \in \mathcal{O}_C} O \right) \times U_C, \tag{34}$$

where $x_i^{(\text{PFC})}(c)$ is computed by (22), when the utterance u is presented to the auditory path, and the object o to the visual path. The 38 object categories introduce a partition in the set

of object's views $\mathcal{O}$, such that all sets of views in the partition, $O \in \mathcal{O}_C$ are of objects of that category C, while the set $U_C \in \mathcal{U}$ collects all utterances naming the category C.

Samples of population coding and actual activations are shown in Fig. 9. Shown are how both the population coding, and the map activations, do not preserve any explicit trace of visual features, and the activations of neurons seem to code the categories in a rather abstract way, as well as the peculiarities of the single samples. For each couple of waves and images, the activations in the PFC map exhibit an important overlap with the population coding of the corresponding category, with contingent additional or lacking activations, denoting the specific view and the specific voice.

The development of both PFC and ACM followed that of all the lower hierarchies in the model, and has been replicated, as in the case of STS, for single gender voices, and for the full set. The objects have been presented using 8 views, of the 72 available, for each. Tab. 4 displays the accuracy achieved at the end of the development, as for STS, the discrimination is quite more accurate when words are spoken by a person of a single gender. Nevertheless, both models achieve a remarkable ability of categorizing objects taking into account their names, even when thrown into a two gender speaking environments. The ACM version exhibits a better accuracy in every condition with respect to PFC, with absolute accuracy on most categories, confirming how the SOM architecture is suitable for simulating categorization.

5. Conclusion

The model here described demonstrates that the perceptual experiences of sound and lights are enough for a first emergence of categories of visual objects and their names, where sounds are uttered by people naming objects, and the lights are the views of those objects. All processes implemented in the model rely on fundamental mechanisms that enable detection of recurrent coincidences in stimuli features, and the self-organization of crucial representations of entities in the external world and of lexical categories.

A strength in the model is the attempt to adhere to the chain of computational processes taking place in the brain, especially in the cerebral cortex, and to rely on fundamental plasticity mechanisms that allow functions to emerge from sensorial experience. The model benefits
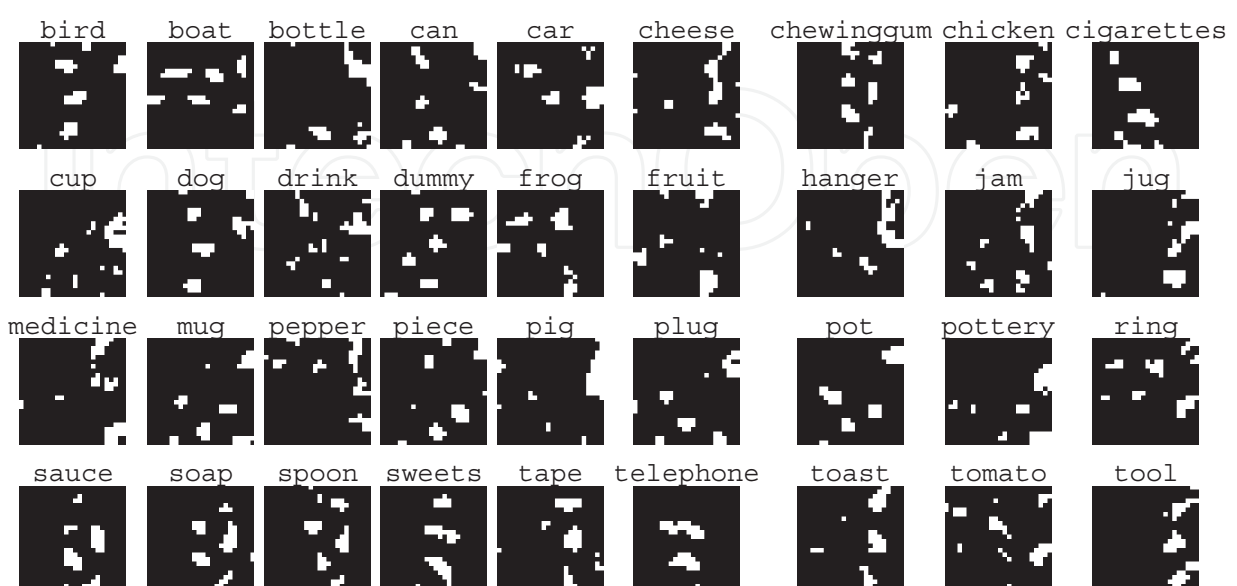
Fig. 8. Population coding in the PFC map for most categories.

category	PFC			ACM		
	all	male	female	all	male	female
bird	0.986	0.903	1.000	1.000	1.000	1.000
boat	0.304	0.198	0.844	0.972	0.965	1.000
bottle	0.450	1.000	0.564	1.000	1.000	1.000
can	0.764	1.000	0.833	0.750	1.000	1.000
car	0.316	0.423	1.000	0.996	1.000	1.000
cheese	1.000	1.000	0.963	1.000	1.000	1.000
chewinggum	0.500	1.000	0.676	1.000	1.000	1.000
chicken	0.750	1.000	0.972	1.000	1.000	1.000
cigarettes	1.000	1.000	1.000	1.000	1.000	1.000
cup	0.293	0.942	1.000	0.994	1.000	1.000
dog	0.986	1.000	1.000	1.000	1.000	1.000
drink	0.484	1.000	0.462	1.000	1.000	1.000
dummy	0.819	1.000	0.986	1.000	1.000	1.000
frog	0.743	0.958	1.000	1.000	1.000	1.000
fruit	0.650	0.560	1.000	1.000	1.000	1.000
hanger	0.576	1.000	0.764	0.875	1.000	1.000
jam	0.616	0.514	0.764	0.938	1.000	1.000
jug	0.389	1.000	0.486	0.500	1.000	0.500
kitten	0.424	0.722	1.000	1.000	1.000	1.000
medicine	0.969	1.000	1.000	1.000	1.000	1.000
mug	0.823	1.000	0.941	1.000	1.000	1.000
pepper	0.823	1.000	1.000	1.000	1.000	1.000
piece	0.935	0.994	0.997	1.000	1.000	1.000
pig	0.979	1.000	1.000	1.000	1.000	1.000
plug	0.521	0.889	0.611	1.000	1.000	1.000
pot	0.694	0.833	1.000	1.000	1.000	1.000
pottery	0.790	0.733	1.000	1.000	1.000	1.000
ring	0.868	1.000	1.000	1.000	1.000	1.000
sauce	0.984	1.000	1.000	1.000	1.000	1.000
soap	0.979	0.950	0.976	1.000	1.000	1.000
spoon	0.750	1.000	0.944	0.750	1.000	1.000
sweets	0.580	0.771	1.000	0.847	1.000	1.000
tape	0.750	0.361	0.806	1.000	1.000	1.000
telephone	0.750	1.000	1.000	0.500	1.000	1.000
toast	1.000	1.000	0.951	1.000	1.000	1.000
tomato	0.743	1.000	1.000	1.000	1.000	1.000
tool	1.000	1.000	0.986	1.000	1.000	1.000
truck	0.483	0.858	1.000	0.977	1.000	1.000
average	0.690	0.853	0.902	0.950	0.999	0.987

Table 4. Categorization accuracy of the two versions of the top map in the model, trained with different sets of voices.
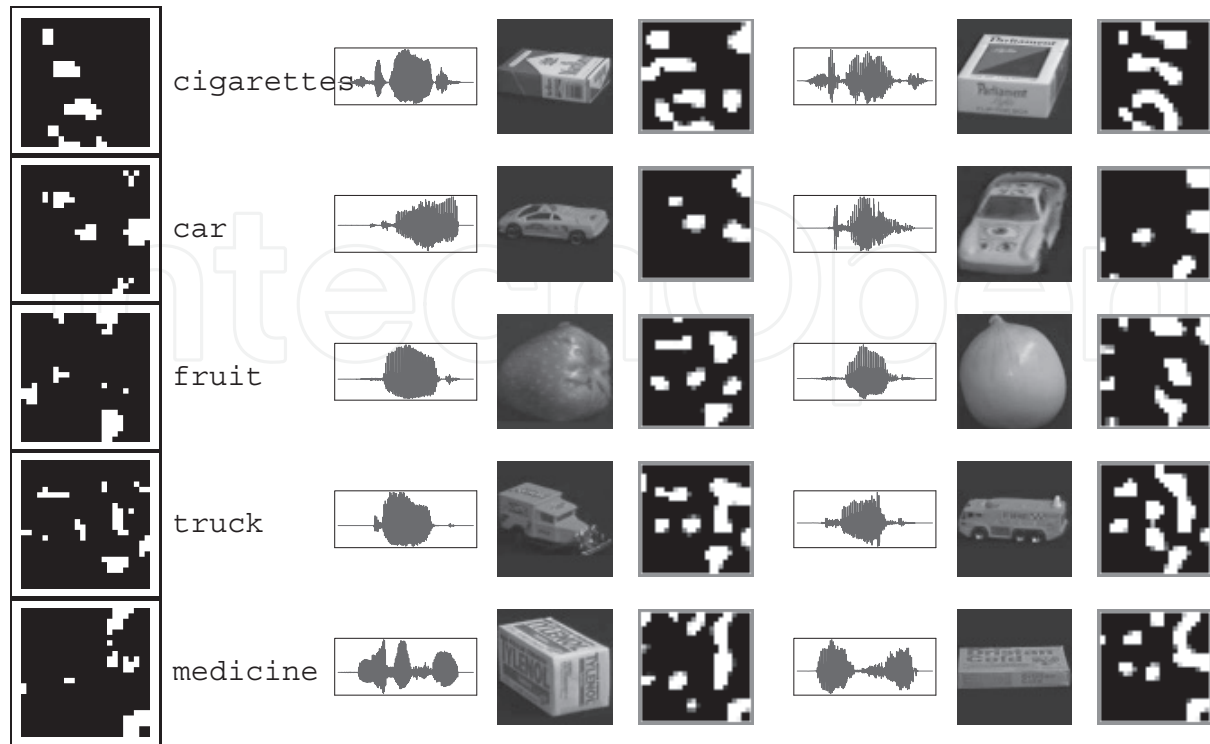
Fig. 9. Examples of population coding in the PFC map. In each row the leftmost map displays the coding neurons. Following images in the row are samples of a sound of the same name, a view of an object of the relevant class, and the corresponding response patterns in PFC. The first sample always has a male voice sound, and the second sample a female voice.

from this effort to different degrees, due to the different levels of knowledge we have of the involved cortical processes. Much more is known about visual recognition, and this explains why this path is much more detailed in the model. In vision, there is also a better knowledge on how the process is organized in the lower part of the path, but we have less information regarding the higher one. In the higher level of the model, where lexical categorization takes place, two different versions have been implemented and compared, one with a classical SOM map, and another that keeps the same LISSOM architecture closer to the cortical structure, even if in this case there is no strict correspondence between model map and a single brain area. We conceive the model here described as an important step in the effort of developing a brain-informed simulation of phenomena in the early development of fundamental human cognitive abilities, that rely on the same coincidence detection and self-organization mechanism. Its novelty is not just in demonstrating the emergence of concepts from the association of visual information and sounds, that has been already simulated in previous work, but rather in its showing how this could happen in neural structures that are similar to the cortical architecture of the human brain.

cigar	cigar	tape	plug	pot	bottl	bottl	bottl		bottl	bottl	bottl		drink		sauce	sweet	sauce
cigar		cup	plug				telep		bottl	bottl			drink	drink		sweet	sweet
potte	cup	cup	car	pot	potte	potte	pot		cup			bird			dummy		sauce
potte				mug	mug	potte	pot	cup	cup	bird	bird		drink	drink			sauce
truck	tape	cup	cup	car	mug	mug	mug	cup	cup	tool		can		truck	dog	dog	dummy
	pig	cup		mug	mug		pig		can	piece	piece		mug	mug	truck	truck	dummy
plug		kitte	mug	mug		peppe		piece	piece		piece		car	mug	car		jam
truck	hange	hange	hange	hange	hange	peppe	peppe			piece	piece		potte		car	jam	jam
tool	hange	kitte			chees			piece	piece			potte	potte		peppe	boat	
tool	truck		jam		chees	chees	jam			fruit	fruit		potte		peppe		tomat
tool	hange	jug		chees			jam	sweet		fruit		fruit	potte	cup	ring	boat	tomat
tool		tape	tool	chees	chees	chick	jam	sweet		fruit	fruit	boat	boat	truck	ring	boat	fruit
soap	soap		tool	chick		chees	jam		toast	toast	boat	boat	boat		truck		truck
sauce			soap		chewi	chick		toast	toast	toast			cup	car	car		soap
sauce	spoon	spoon	soap		chewi	jam			toast	car	car		car		car	car	frog
soap	soap	truck		chewi	chewi	kitte	kitte	medic	medic	medic		car		car	car	boat	frog
	chewi				chewi		kitte		medic	medic		car	car		car	truck	truck
soap	soap	tape	chewi	fruit		bottl		medic		medic	medic	car	car	boat	boat	boat	ring

Fig. 10. Organization of categories in the ACM SOM map, in the case of the model trained with all voices. Category nouns are abbreviated for better readability, see Fig. 4 for full nouns.

6. References

Abbott, L. F., Rolls, E. & Tovee, M. J. (1996). Representational capacity of face coding in monkeys, *Cerebral Cortex* 6: 498–505.

Anzai, A., Peng, X. & Essen, D. C. V. (2007). Neurons in monkey visual area V2 encode combinations of orientations, *Nature Neuroscience* 10: 1313–1321.

Atzori, M., Lei, S., Evans, D. I. P., Kanold, P. O., Phillips-Tansey, E., McIntyre, O. & McBain, C. J. (2001). Differential synaptic processing separates stationary from transient inputs to the auditory cortex, *Neural Networks* 4: 1230–1237.

Ball, P. (1999). *The Self-Made Tapestry: Pattern Formation in Nature*, Oxford University Press, Oxford (UK).

Bednar, J. A., De Paula, J. B. & Miikkulainen, R. (2005). Self-organization of color opponent receptive fields and laterally connected orientation maps, *Neurocomputing*

65-66: 69–76.

Bednar, J. A. & Miikkulainen, R. (2000). Tilt aftereffects in a self-organizing model of the primary visual cortex, *Neural Computation* 12: 1721–1740.

Bednar, J. A. & Miikkulainen, R. (2003). Learning innate face preferences, *Neural Computation* 15: 1525–1557.

Bednar, J. A. & Miikkulainen, R. (2006). Joint maps for orientation, eye, and direction preference in a self-organizing model of v1, *Neurocomputing* 69: 1272–1276.

Belin, P., Fecteau, S. & Bédard, C. (2004). Thinking the voice: neural correlates of voice perception, *Trends in Cognitive Sciences* 8: 129–135.

Belin, P. & Zatorre, R. J. (2000). 'what', 'where' and 'how' in auditory cortex, *Nature Neuroscience* 3: 965–966.

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P. & Pike, B. (2000). Voice-selective areas in human auditory cortex, *Cognitive Brain Research* 403: 309–312.

Black, A. W. & Taylor, P. A. (1997). The festival speech synthesis system: System documentation, *Technical Report HCRC/TR-83*, Human Communciation Research Centre, University of Edinburgh, Edinburgh, UK.

Blasdel, G. G. (1992). Orientation selectivity, preference, and continuity in monkey striate cortex, *Journal of Neuroscience* 12: 3139–3161.

Brewer, A. A., Liu, J., Wade, A. R. & Wandell, B. A. (2005). Visual field maps and stimulus selectivity in human ventral occipital cortex, *Nature Neuroscience* 8: 1102–1109.

Brown, M. C. (2003). Audition, *in* L. R. Squire, F. Bloom, S. McConnell, J. Roberts, N. Spitzer & M. Zigmond (eds), *Fundamental Neuroscience*, Academic Press, New York, pp. 699–726.

Chapman, B., Stryker, M. P. & Bonhoeffer, T. (1996). Development of orientation preference maps in ferret primary visual cortex, *Journal of Neuroscience* 16: 6443–6453.

Chklovskii, D. B. & Koulakov, A. A. (2004). Maps in the brain: what can we learn from them?, *Annual Review of Neuroscience* 27: 369–392.

Choe, Y. & Miikkulainen, R. (2004). Contour integration and segmentation with self-organized lateral connections, *Biological Cybernetics* 90: 75–88.

Churchland, P. M. (1989). *A Neurocomputational perspective: The Nature of Mind and the Structure of Science*, MIT Press, Cambridge (MA).

Cox, D. D., Meier, P., Oertelt, N. & DiCarlo, J. J. (2005). 'breaking' position-invariant object recognition, *Nature Neuroscience* 8: 1145–1147.

Dannemiller, J. L. (1989). A test of color constancy in 9- and 20-weeks-old human infants following simulated illuminant changes, *Developmental Psychology* 25: 171–184.

Dayan, P. & Abbott, L. F. (2001). *Theoretical Neuroscience*, MIT Press, Cambridge (MA).

Dowling, J. E. (1987). *The Retina: An Approachable Part of the Brain*, Cambridge University Press, Cambridge (UK).

Eimas, P. D. & Quinn, P. C. (1994). Studies on the formation of perceptually based basic-level categories in young infants, *Child Development* 3: 903–917.

Farah, M. J. (1990). *Visual Agnosia: Disorders of Object Recognition and What They Tell Us About Normal Vision*, MIT Press, Cambridge (MA).

Fodor, J. (1981). *Representations: Philosofical Essay on the Foundation of Cognitive Science*, MIT Press, Cambridge (MA).

Fodor, J. (1983). *Modularity of Mind: and Essay on Faculty Psychology*, MIT Press, Cambridge (MA).

Freedman, D. J., Riesenhuber, M., Poggio, T. & Miller, E. K. (2003). Visual categorization

and the primate prefrontal cortex: Neurophysiology and behavior, *Journal of Neurophysiology* 88: 929–941.

Fuster, J. M. (2002). Frontal lobe and cognitive development, *Journal of Neurocytology* 31: 373–385.

Gegenfurtner, K. R., Kiper, D. C. & Levitt, J. B. (1997). Functional properties of neurons in macaque area V3, *Journal of Neurophysiology* 77: 1906–1923.

Gibson, E. J. & Levin, H. (1975). *The psychology of reading*, MIT Press, Cambridge (MA).

Gödecke, I. & Bonhoeffer, T. (1996). Development of identical orientation maps for two eyes without common visual experience, *Nature* 379: 251–254.

Goodale, M. A. & Milner, A. D. (1992). Separate visual pathways for perception and action, *Trends in Neuroscience* 15: 20–25.

Grill-Spector, K., Kushnir, T., Edelman, S., Avidan-Carmel, G., Itzchak, Y. & Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex, *Neuron* 24: 187–203.

Grill-Spector, K., Kushnir, T., Edelman, S., Itzchak, Y. & Malach, R. (1998). Cue-invariant activation in object-related areas in the human occipital lobe, *Neuron* 21: 191–202.

Hadjikhani, N., Liu, A. K., Dale, A. M., Cavanagh, P. & Tootell, R. B. H. (1998). Retinotopy and color sensitivity in human visual cortical area v8, *Nature Neuroscience* 1: 235–241.

Haken, H. (1978). *Synergetics – An Introduction, Nonequilibrium Phase Transitions and Self-organization in Physics, Chemistry and Biology*, Springer-Verlag, Berlin.

Hegdé, J. & Van Essen, D. C. (2000). Selectivity for complex shapes in primate visual area V2, *Journal of Neuroscience* 20: 4117–4130.

Heider, B., Meskenaite, V. & Peterhans, E. (2000). Anatomy and physiology of a neural mechanism defining depth order and contrast polarity at illusory contours, *The European Journal of Neuroscience* 12: 4117–4130.

Hickok, G. & Poeppel, D. (2007). The cortical organization of speech processing, *Nature Reviews Neuroscience* 8: 393–402.

Hinton, G. E., McClelland, J. L. & Rumelhart, D. E. (1986). Distributed representations, *in* D. E. Rumelhart & J. L. McClelland (eds), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 1, MIT Press, Cambridge (MA), pp. 77–109.

Hubel, D. H. & Wiesel, T. N. (1959). *Brain and Visual Perception: The Story of a 25-Year Collaboration*, Oxford University Press, Oxford (UK).

Hubel, D. & Wiesel, T. (1962). Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex, *Journal of Physiology* 160: 106–154.

Huey, E. D., Krueger, F. & Grafman, J. (2006). Representations in the human prefrontal cortex, *Current Directions in Psychological Science* 15: 167–171.

Hume, D. (1739). *A Treatise of Human Nature*, John Noon, London. Vol 1,2.

Ito, M. & Komatsu, H. (2004). Representation of angles embedded within contour stimuli in area V2 of macaque monkeys, *Journal of Neuroscience* 24: 3313–3324.

Katz, L. & Shatz, C. (1996). Synaptic activity and the construction of cortical circuits, *Science* 274: 1133–1138.

Keil, F. C., Smith, W. C., Simons, D. J. & Levin, D. T. (1998). Two dogmas of conceptual empiricism: implications for hybrid models of the structure of knowledge, *Cognition* 65: 103–135.

Kobatake, E. & Tanaka, K. (1994). Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex, *Journal of Neurophysiology* 71: 856–867.

Kohonen, T. (1982). Self-organizing formation of topologically correct feature maps, *Biological Cybernetics* 43: 59–69.

Kohonen, T. (1995). *Self-Organizing Maps*, Springer-Verlag, Berlin.

Kohonen, T. (1998). The self-organizing map, a possible model of brain maps, *in* K. H. Pribram (ed.), *Brain and Values: Is A Biological Science of Values Possible?*, Psychology Press, East Sussex (UK), pp. 207–236.

Kohonen, T. & Hari, R. (2000). Where the abstract feature maps of the brain might come from, *Trends in Neurosciences* 22: 135–139.

Kourtzi, Z., Erb, M., Grodd, W. & Bülthoff, H. H. (2003). Representation of the perceived 3-d object shape in the human lateral occipital complex, *Cerebral Cortex* 13: 911–920.

Landisman, C. E. & Ts'o, D. Y. (2002). Color processing in macaque striate cortex: Relationships to ocular dominance, cytochrome oxidase, and orientation, *Journal of Neurophysiology* 87: 3126–3137.

Li, P., Farkas, I. & MacWhinney, B. (2004). Early lexical development in a self-organizing neural network, *Neural Networks* 17: 1345–1362.

Liebenthal, E., Binder, J. R., Spitzer, S. M., Possing, E. T. & Medler, D. A. (2005). Neural substrates of phonemic perception, *Cerebral Cortex* 15: 1621–1631.

Linden, J. F. & Schreiner, C. E. (2006). Columnar transformations in auditory cortex? a comparison to visual and somatosensory cortices, *Cerebral Cortex* 13: 83–89.

Locke, J. (1690). *An essay concerning human understanding*, Meridian Books, Cleveland.

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., Ledden, P. J., Brady, T. J., Rosen, B. R. & Tootell, R. B. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex, *Proceedings of the Natural Academy of Science USA* 92: 8135–8139.

Mandler, J. M. (2004). *The Foundations of Mind*, Oxford University Press, Oxford (UK).

Mastronarde, D. N. (1983). Correlated firing of retinal ganglion cells: I. spontaneously active inputs in X- and Y-cells, *Journal of Neuroscience* 14: 409–441.

Mayor, J. & Plunkett, K. (2010). A neurocomputational account of taxonomic responding and fast mapping in early word learning, *Psychological Review* 117: 1–31.

Miikkulainen, R. (1997). Dyslexic and category-specific aphasic impairments in a self-organizing feature map model of the lexicon, *Brain and Language* 59: 334–366.

Miikkulainen, R., Bednar, J. A., Choe, Y. & Sirosh, J. (1997). Self-organization, plasticity, and low-level visual phenomena in a laterally connected map model of the primary visual cortex, *in* R. L. Goldstone, P. G. Schyns & D. L. Medin (eds), *Psychology of Learning and Motivation*, Vol. 36, Academic Press, New York, pp. 257–308.

Miikkulainen, R., Bednar, J., Choe, Y. & Sirosh, J. (2005). *Computational maps in the visual cortex*, Springer-Science, New York.

Miller, E. K., Freedman, D. J. & Wallis, J. D. (2002). The prefrontal cortex: Categories, concepts and cognition, *Philosophical Transactions: Biological Sciences* 357: 1123–1136.

Miller, K. D., Keller, J. B. & Stryker, M. P. (1989). Ocular dominance column development: Analysis and simulation, *Science* 245: 605–615.

Miller, L. M., Escab, M. A., Read, H. L. & Schreiner, C. E. (2002). Spectrotemporal receptive fields in the lemniscal auditory thalamus and cortex, *Journal of Neurophysiology* 87: 516–527.

Mountcastle, V. (1957). Modality and topographic properties of single neurons in cats somatic sensory cortex, *Journal of Neurophysiology* 20: 408–434.

Nayar, S. & Murase, H. (1995). Visual learning and recognition of 3-d object by appearence,

International Journal of Computer Vision 14: 5–24.

Pasupathy, A. & Connor, C. E. (2002). Population coding of shape in area v4, *Nature Neuroscience* 5: 1332–1338.

Plebe, A. (2006). Learning visual invariance, *in* M. Verleysen (ed.), *ESANN 2006 – 14th European Symposium on Artificial Neural Networks*, d-side Publications, Evere (BE), pp. 71–76.

Plebe, A. (2007). A model of angle selectivity development in visual area V2, *Neurocomputing* 70: 2060–2066.

Plebe, A., De la Cruz, V. & Mazzone, M. (2007a). Artificial learners of objects and names, *in* Y. Demiris, B. Scassellati & D. Mareschal (eds), *Proceedings of the 6th International Conference on Development and Learning*, IEEE, pp. 300–305.

Plebe, A., De la Cruz, V. & Mazzone, M. (2007b). Simulating the acquisition of object names, *in* P. Buttery, A. Villavicencio & A. Korhonen (eds), *Proceedings of the Workshop on Cognitive Aspects of Computational Language Acquisition*, Association for Computational Linguistics, Stroudsburg (PA), pp. 57–64.

Plebe, A. & Domenella, R. G. (2006). Early development of visual recognition, *BioSystems* 86: 63–74.

Plebe, A. & Domenella, R. G. (2007). Object recognition by artificial cortical maps, *Neural Networks* 20: 763–780.

Quartz, S. R. & Sejnowski, T. J. (1997). The neural basis of cognitive development: a constructivist manifesto, *Behavioral and Brain Science* 20: 537–596.

Rakison, D. H. (ed.) (2003). *Early Category And Concept Development*, Oxford University Press, Oxford (UK).

Ritter, H. & Kohonen, T. (1989). Self-organizing semantic maps, *Biological Cybernetics* 61: 241–254.

Rolls, E. (1992). Neurophysiological mechanisms underlying face processing within and beyond the temporal cortical visual areas, *Philosophical transactions of the Royal Society B* 335: 11–21.

Rolls, E. & Tovee, M. J. (1995). Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex, *Journal of Neurophysiology* 73: 713–726.

Romanski, L. M. & Goldman-Rakic, P. S. (1999). Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex, *Nature Neuroscience* 2: 1131–1136.

Schiller, P. H. (1996). On the specificity of neurons and visual areas, *Behavioural Brain Research* 76: 21–35.

Sereno, M., Dale, A., Reppas, J., Kwong, K., Belliveau, J., Brady, T., Rosen, B. & Tootell, R. (1995). Borders of multiple visual areas in human revealed by functional magnetic resonance imaging, *Science* 268: 889–893.

Sirosh, J. & Miikkulainen, R. (1997). Topographic receptive fields and patterned lateral interaction in a self-organizing model of the primary visual cortex, *Neural Computation* 9: 577–594.

Smith, M. A., Kohn, A. & Movshon, J. A. (2007). Glass pattern responses in macaque V2 neurons, *Journal of Neurophysiology* 97: 4284–4295.

Thompson, I. (1997). Cortical development: A role for spontaneous activity?, *Current Biology* 7: 324–326.

Turrigiano, G. G. & Nelson, S. B. (2004). Homeostatic plasticity in the developing nervous system, *Nature Reviews Neuroscience* 391: 892–896.

Ungerleider, L. & Mishkin, M. (1982). Two cortical visual systems, *in* D. J. Ingle, M. A. Goodale

& R. J. W. Mansfield (eds), *Analysis of visual behavior*, MIT Press, Cambridge (MA), pp. 549–586.

Van Essen, D. C. & DeYoe, E. A. (1994). Concurrent processing in the primate visual cortex, *in* M. S. Gazzaniga (ed.), *The Cognitive Neurosciences*, MIT Press, Cambridge (MA).

Vanduffel, W., Tootell, R. B., Schoups, A. A. & Orban, G. A. (2002). The organization of orientation selectivity throughout the macaque visual cortex, *Cerebral Cortex* 12: 647–662.

Verkindt, C., Bertrand, O., Echallier, F. & Pernier, J. (1995). Tonotopic organization of the human auditory cortex: N100 topography and multiple dipole model analysis, *Electroencephalography and Clinical Neurophisiology* 96: 143–156.

von der Malsburg, C. (1973). Self-organization of orientation sensitive cells in the striate cortex, *Kibernetic* 14: 85–100.

Vuilleumier, P., Henson, R. N., Driver, J. & Dolan, R. J. (2002). Multiple levels of visual object constancy revealed by event-related fmri of repetition priming, *Nature Neuroscience* 5: 491–499.

Wallis, G. & Rolls, E. (1997). Invariant face and object recognition in the visual system, *Progress in Neurobiology* 51: 167–194.

Wandell, B. A., Brewer, A. A. & Dougher, R. F. (2005). Visual field map clusters in human cortex, *Philosophical transactions of the Royal Society of London* 360: 693–707.

Weigelt, S., Kourtzi, Z., Kohler, A., Singer, W. & Muckli, L. (2007). The cortical representation of objects rotating in depth, *Journal of Neuroscience* 27: 3864–3874.

Westermann, G., Mareschal, D., Johnson, M. H., Sirois, S., Spratling, M. S. & Thomas, M. S. C. (2007). Neuroconstructivism, *Developmental Science* 10: 75–83.

Winer, J. A., Miller, L. M., Lee, C. C. & Schreiner, C. E. (2005). Auditory thalamocortical transformation: structure and function, *Neuron* 28: 255–263.

Zeki, S. (1983a). Colour coding in the cerebral cortex: The reaction of cells in monkey visual cortex to wavelenghts and colours, *Neuroscience* 9: 741–765.

Zeki, S. (1983b). Colour coding in the cerebral cortex: the responses of wavelength-selective and colour-coded cells in monkey visual cortex to changes in wavelength composition, *Neuroscience* 9: 767–781.

Self Organizing Maps - Applications and Novel Algorithm Design

Edited by Dr Josphat Igadwa Mwasiagi

Kohonen Self Organizing Maps (SOM) has found application in practical all fields, especially those which tend to handle high dimensional data. SOM can be used for the clustering of genes in the medical field, the study of multi-media and web based contents and in the transportation industry, just to name a few. Apart from the aforementioned areas this book also covers the study of complex data found in meteorological and remotely sensed images acquired using satellite sensing. Data management and envelopment analysis has also been covered. The application of SOM in mechanical and manufacturing engineering forms another important area of this book. The final section of this book, addresses the design and application of novel variants of SOM algorithms.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Alessio Plebe (2011). Self-Organization of Object Categories in a Cortical Artificial Model, Self Organizing Maps - Applications and Novel Algorithm Design, Dr Josphat Igadwa Mwasiagi (Ed.), ISBN: 978-953-307-546-4, InTech, Available from: http://www.intechopen.com/books/self-organizing-maps-applications-and-novel-algorithm-design/self-organization-of-object-categories-in-a-cortical-artificial-model

INTECH
open science | open minds