Alma Mater Studiorum - Università di Bologna

Dipartimento di Elettronica, Informatica e Sistemistica

Dottorato di Ricerca in Bioingegneria
XXI Ciclo

Settore scientifico disciplinare di afferenza: ING-INF/06

MATHEMATICAL MODELS OF COGNITIVE PROCESSES

Dott. Ing. **Cristiano Cuppini**

Relatore:
Chiar.mo Prof. **Mauro Ursino**

Correlatore:
Prof. Ing. **Elisa Magosso**

Controrelatore:
Prof. Ing. **Fabio Babiloni**

Coordinatore del Corso di Dottorato
Prof. Ing. **Angelo Cappello**

Esame finale anno 2009

INDEX

ABSTRACT

INTRODUCTION

Part 1. SYNCHRONIZED OSCILLATORY ACTIVITY IN OBJECT RECOGNITION, SEMANTIC MEMORY AND LANGUAGE

1.1. Object segmentation and recovery via neural oscillators implementing the similarity and prior knowledge gestalt rules

1.2. Recognition of abstract objects via neural oscillators: interaction among topological organization, associative memory and gamma band synchronization

1.3. A neural network model of semantic memory linking feature-based object representation and word

Part 2. MULTISENSORY INTEGRATION

2.1. Multisensory integration in the superior colliculus: a neural network model

2.2. A neural network model describing sensory integration in the Superior Colliculus

CONCLUSIONS

PUBLICATIONS

ABSTRACT

The research activity carried out during the PhD course was focused on the development of mathematical models of some cognitive processes and their validation by means of data present in literature, with a double aim: i) to achieve a better interpretation and explanation of the great amount of data obtained on these processes from different methodologies (electrophysiological recordings on animals, neuropsychological, psychophysical and neuroimaging studies in humans), ii) to exploit model predictions and results to guide future research and experiments.

In particular, the research activity has been focused on two different projects: 1) the first one concerns the development of neural oscillators networks, in order to investigate the mechanisms of synchronization of the neural oscillatory activity during cognitive processes, such as object recognition, memory, language, attention; 2) the second one concerns the mathematical modelling of multisensory integration processes (e.g. visual-acoustic), which occur in several cortical and subcortical regions (in particular in a subcortical structure named Superior Colliculus (SC)), and which are fundamental for orienting motor and attentive responses to external world stimuli. This activity has been realized in collaboration with the Center for Studies and Researches in Cognitive Neuroscience of the University of Bologna (in Cesena) and the Department of Neurobiology and Anatomy of the Wake Forest University School of Medicine (NC, USA).

PART 1.

Objects representation in a number of cognitive functions, like perception and recognition, foresees distribute processes in different cortical areas. One of the main neurophysiological question concerns how the correlation between these disparate areas is realized, in order to succeed in grouping together the characteristics of the same object (binding problem) and in maintaining segregated the properties belonging to different objects simultaneously present (segmentation problem).

Different theories have been proposed to address these questions (Barlow, 1972). One of the most influential theory is the so called “assembly coding”, postulated by Singer (2003), according to which 1) an object is well described by a few fundamental properties, processing in different and distributed cortical areas; 2) the recognition of the object would be realized by means of the simultaneously activation of the cortical areas representing its different features; 3) groups of properties belonging to different objects would be kept separated in the time domain.

In **Chapter 1.1** and in **Chapter 1.2** we present two neural network models for object recognition, based on the “assembly coding” hypothesis. These models are networks of Wilson-Cowan oscillators which exploit: i) two high-level “Gestalt Rules” (the similarity and previous knowledge rules), to realize the functional link between elements of different cortical areas representing properties of the same object (binding problem); 2) the synchronization of the neural oscillatory activity in the γ-band (30-100Hz), to segregate in time the representations of different objects simultaneously present (segmentation problem).

These models are able to recognize and reconstruct multiple simultaneous external objects, even in difficult case (some wrong or lacking features, shared features, superimposed noise).

In **Chapter 1.3** the previous models are extended to realize a semantic memory, in which sensory-motor representations of objects are linked with words. To this aim, the network, previously developed, devoted to the representation of objects as a collection of sensory-motor features, is reciprocally linked with a second network devoted to the representation of words (lexical network)

Synapses linking the two networks are trained via a time-dependent Hebbian rule, during a training period in which individual objects are presented together with the corresponding words.

Simulation results demonstrate that, during the retrieval phase, the network can deal with the simultaneous presence of objects (from sensory-motor inputs) and words (from linguistic inputs), can correctly associate objects with words and segment objects even in the presence of incomplete information. Moreover, the network can realize some semantic links among words representing objects with some shared features.

These results support the idea that semantic memory can be described as an integrated process, whose content is retrieved by the co-activation of different multimodal regions. In perspective, extended versions of this model may be used to test conceptual theories, and to provide a quantitative assessment of existing data (for instance concerning patients with neural deficits).

PART 2.

The ability of the brain to integrate information from different sensory channels is fundamental to perception of the external world (Stein et al, 1993). It is well documented that a number of extraprimary areas have neurons capable of such a task; one of the best known of these is the superior colliculus (SC). This midbrain structure receives auditory, visual and somatosensory inputs from different subcortical and cortical areas, and is involved in the control of orientation to external events (Wallace et al, 1993).
SC neurons respond to each of these sensory inputs separately, but is also capable of integrating them (Stein et al, 1993) so that the response to the combined multisensory stimuli is greater than that to the individual component stimuli (enhancement). This enhancement is proportionately greater if the modality-specific paired stimuli are weaker (the principle of inverse effectiveness). Several studies have shown that the capability of SC neurons to engage in multisensory integration requires inputs from cortex; primarily the anterior ectosylvian sulcus (AES), but also the rostral lateral suprasylvian sulcus (rLS). If these cortical inputs are deactivated the response of SC neurons to cross-modal stimulation is no different from that evoked by the most effective of its individual component stimuli (Jiang et al 2001).

This phenomenon can be better understood through mathematical models. The use of mathematical models and neural networks can place the mass of data that has been accumulated about this phenomenon and its underlying circuitry into a coherent theoretical structure.

In **Chapter 2.1** a simple neural network model of this structure is presented; this model is able to reproduce a large number of SC behaviours like multisensory enhancement, multisensory and unisensory depression, inverse effectiveness. In **Chapter 2.2** this model was improved by incorporating more neurophysiological knowledge about the neural circuitry underlying SC multisensory integration, in order to suggest possible physiological mechanisms through which it is effected. This endeavour was realized in collaboration with Professor B.E. Stein and Doctor B. Rowland during the 6 months-period spent at the Department of Neurobiology and Anatomy of the Wake Forest University School of Medicine (NC, USA), within the Marco Polo Project.

The model includes four distinct unisensory areas that are devoted to a topological representation of external stimuli. Two of them represent subregions of the AES (i.e., FAES, an auditory area, and AEV, a visual area) and send descending inputs to the ipsilateral SC; the other two represent subcortical areas (one auditory and one visual) projecting ascending inputs to the same SC.

Different competitive mechanisms, realized by means of population of interneurons, are used in the model to reproduce the different behaviour of SC neurons in conditions of cortical activation and deactivation.

The model, with a single set of parameters, is able to mimic the behaviour of SC multisensory neurons in response to very different stimulus conditions (multisensory enhancement, inverse effectiveness, within- and cross-modal suppression of spatially disparate stimuli), with cortex functional and cortex deactivated, and with a particular type of membrane receptors (NMDA receptors) active or inhibited. All these results agree with the data reported in Jiang et al. (2001) and in Binns and Salt (1996).

The model suggests that non-linearities in neural responses and synaptic (excitatory and inhibitory) connections can explain the fundamental aspects of multisensory integration, and provides a biologically plausible hypothesis about the underlying circuitry.

INTRODUCTION

The aim of the research activity presented in this thesis is to study cognitive processes by means of mathematical models and computer simulations. The use of computational methods allows a mass of data obtained from different methodologies (electrophysiological recordings in animals, neuropsychological, psychophysical and neuroimaging studies in humans) to be reciprocally related and summarized into a unique theoretical structure; moreover, mathematical methods and computer simulations may generate new predictions about the behaviour and the function of cerebral structures hardly analyzed by means of classical approaches, and may drive future experiments.

During the PhD course, the activity has been focused on two different projects: 1) the first, described in the PART 1 of this thesis, concerns the development of neural networks of oscillators, in order to investigate the mechanism of synchronization of the oscillatory activity, especially in γ-band, in different brain areas and its role in cognitive processes, such as object recognition, memory, language, attention; 2) the second, reported in the PART 2 of this dissertation, concerns the mathematical modelling of multisensory integration processes (e.g. visual-acoustic) which occur in a subcortical structure named Superior Colliculus (SC), and which are fundamental in eliciting appropriate motor and attentive responses to external world stimuli. This activity has been realized in collaboration with the Center for Studies and Researches in Cognitive Neuroscience of the University of Bologna (in Cesena) and the Department of Neurobiology and Anatomy of the Wake Forest University School of Medicine (NC, USA).

Part 1. SYNCHRONIZED OSCILLATORY ACTIVITY IN OBJECT RECOGNITION, SEMANTIC MEMORY AND LANGUAGE

The brain is made of by a collection of neuronal assemblies, connected each others to form networks involved in different cognitive processes. In recent years there has been large agreement about the idea that cognitive processes do not rely on the activity of single areas but depend on large-scale circuits, networks of interconnected brain areas that become contemporary active during specific cognitive demands. Under this point of view, a coherent brain activity plays a fundamental role in human behaviour (for instance, several authors have shown that a strong connectivity between visual and premotor cortex is essential in driving a motor behavior response), and a similar important role is played during conscious visual perception, attentional processes and working memory tasks.

These networks can be of different size: small and bound to populations of neighbouring neurons (i.e. in visual-feature binding) or widespread including areas far from each other (i.e. in more complicate cognitive tasks like memorization, language and motor response).

Through the analysis of the EEG during several experiments, phase synchronization between cortical activities in regions involved in the same process was identified as a possible mechanism subserving large-scale cognitive integration. This synchronization may be distributed both within the same cortical area and among distant areas, and is not locked to external stimuli, i.e., it depends on internal connections among neurons (i.e., on an internal representation of objects). Further studies suggest also that synchronization increases with conscious perception compared with unconscious (subliminal) processes (Melloni L. et al., 2007).

Conversely, other brain activation measures such as EEG amplitude, turned out to be ineffective in discriminating between different kind of processes, or in signalling a good cognitive result (i.e. perceived vs. unperceived visual target task). For instance, Rodriguez E. et al. (1999) described a face recognition task in which EEG signals were recorded while subjects viewed ambiguous stimuli (faces or meaningless shapes). Results show that pattern of gamma activity was spatially homogeneous and similar between the perception and no-perception conditions over time. On the contrary, the pattern of synchrony was different between the two conditions. In particular

synchronization revealed the activation of a cognitive circuit between the left parieto-occipital and frontotemporal regions, during the visual perception and storage of a face. By contrast, the no-recognition event was associated with a no-synchrony activation in those regions.

In order to clarify the role of γ-band synchronization in high level cognitive processes (such as object recognition, memorization and semantic representation) some neural network models have been realized. Those presented in **chapter 1.1** and in **chapter 1.2** deal with the object representation problem. The model in **chapter 1.3** enlarge the viewpoint to lessical and semantic aspects.

Object representation (chapter 1.1 and chapter 1.2) is the first step performed by several higher-level cognitive processes like attention, memorization, working memory, and language.

An early hypothesis in literature was that the presence of objects would be signalled by specialized neurons, processing individual features via a feed-forward and hierarchically structured process, and would encode increasingly complex relationships (Barlow, 1972). According to this idea, the simultaneous presence of two objects in the same scene is signalled by activation of two distinct specialized neurons. This mechanism is generally rejected in the neurophysiological literature today, and in addition, it exhibits several drawbacks: first of all it would lead to a combinatorial explosion of possibilities, hence to an excessive number of individual neurons; furthermore, with this mechanism it is difficult to incorporate new knowledge and to deal with entirely novel objects (Singer, 1999).

The previous limitations may be overcome by the so called “assembly coding”: according to this hypothesis, the neural system utilizes a limited number of features to classify and recognize perceived objects, and the presence of an object is signalled by the simultaneous activity of many neurons, each encoding a single feature (von der Malsburg and Schneider, 1986; von der Malsburg and Buhmann, 1992; Singer, 1993; Singer and Gray, 1995; Eckhorn, 1999; Tallon-Baudry and Bertrand, 1999; Singer, 2003).

Several experiments have shown that cortical neurons are often engaged in synchronous activity in the γ-frequency band (40–60 Hz) (Gray and Singer, 1989; Singer, 1993). Synchronization of

cortical oscillatory activity in the gamma band has been observed in response to several classes of sensory stimuli (visual, somatosensory (Lebedev and Nelson, 1995), auditory (Brosch et al., 2002; Kaiser et al., 2002) and olfactory system (Freeman, 1978; Wehr and Laurent, 1996)) and also in high level cognitive tasks (such as recognition of music (Bhattacharya et al., 2001), recognition of word vs. non-words (Pulvermüller et al., 1996), during visual search tasks (Tallon-Baudry et al., 1997), during delayed-matching-to-sample-tasks (Tallon-Baudry et al., 1998)).

A central question in neurophysiology is how this distributed neuronal activity is functionally linked, to group the different features into a unitary and coherent object representation and how features of different objects, simultaneously present in the same scene or situation, are segregated to avoid interference and false conjunctions.

The aim of our models is to explore the mechanisms carrying out this synchronization between different neural groups, even belonging to distant cortical structures. In particular we investigate the possibility that this synchronization is realized by means of Gestalt rules of high level (like the similarity and the prior knowledge), implemented in the synaptic connections between elements in the same cortical area, but also between different distant cortical structures.

Lexical representation (chapter 1.3) A further extension of the previous models considers how object representation can be linked with words (lexical aspects) and how words can be related via common features in the corresponding objects (semantic aspects).

When discussing the organisation of memory, cognitive neuroscientists commonly distinguish the long-term declarative memory into two main classes: episodic and semantic (Tulving, 1983). The last term “semantic memory” is used to denote information which is context independent, is culturally shared and involves the comprehension of words and concepts.

Several theories of semantic memory have been proposed in past years, with a special focus on object representation. Actually, most of the available information used in the formulation of these theories is based on clinical trials on patients with lesions of the brain, who exhibit some deficits in recognizing objects from words or in evoking words from objects (Lambon Ralph et al.,

2007;Warrington and McCarthy, 1983;Warrington and Shallice, 1984) (see also (Gainotti, 2006) for a review). Additional information in most recent years is derived from functional neuroimaging studies and from neurophysiological measurements.

Although conceptual theories of semantic memory differ in many aspects, most of them agree in considering it as a distributed process, which engages many different cortical areas and exploits a multi-modal representation of objects (Warrington and McCarthy, 1983;Warrington and Shallice, 1984, Damasio, 1989, Caramazza et al., 1990, Lauro-Grotto et al.,1997, Caramazza and Shelton, 1998, Tyler et al., 2000, Snowden et al., 2004, Gainotti, 2000 and 2006).

Two fundamental points emerge from analysis of these theories (see also (Hart et al., 2007)): all assume that features are essential for the formation of concepts and that semantic memory encompasses a distributed representation of these features, over different modality domains.

Several problems, however, are implied in these conceptual models and must be solved for building a functioning semantic neural network. A first problem is how the different pieces of information, shared in a distributed and multimodal representation of features, can be linked together to form a coherent object description (this is the problem dealt with in chapt. 1.1 and 1.2). The second issue is how this object representation could be related with the use of words, and with the lexical aspects of our semantic. Of course, objects can be retrieved from words; similarly, the sensory presentation of objects can evoke the corresponding word. The third point is how different representations of objects and their relative words can be simultaneously maintained in memory, and correctly separated, preserving a distinction of their individual features (this is a semantic form of the classic integration vs. segregation problem of vision research, see (Singer and Gray, 1995)).

The models, that we have realized, give a valuable support for the clarification of these problems and allow the different conceptual theories to be implemented in quantitative terms, their mechanisms formulated in rigorous ways and the emergent behaviour analyzed via computer simulations.

CHAPTER 1.1. OBJECT SEGMENTATION AND RECOVERY VIA NEURAL OSCILLATORS IMPLEMENTING THE SIMILARITY AND PRIOR KNOWLEDGE GESTALT RULES

INTRODUCTION

Several experiments have pointed out that in various cognitive functions, object representation occurs in a highly parallel and distributed manner: the different features of an object are processed and coded in distinct and distant cortical areas (Tallon-Baudry C. and Bertrand O., 1999.

A recent influential hypothesis (Damasio, 1989; Singer and Gray, 1995; Singer, 1999; Varela et al., 2001; von der Malsburg and Schneider, 1986), named "Temporal Correlation Hypothesis", postulates that neuronal groups representing different aspects of the same object are bound together into a cell assembly through synchronization of their activity in the gamma range (30-100 Hz). According to this hypothesis, neurons that fire in phase would signal attributes of the same object, while neurons firing out of phase would signal attributes in different objects.

A role of gamma activity has been demonstrated in recognition of music (Bhattacharya et al., 2001), word vs. non-words (where it seems to reflect association between words and meanings) (Pulvermüller et al., 1996) as well as during visual search tasks (Tallon-Baudry et al., 1997) and delayed-matching-to-sample-tasks (Tallon-Baudry et al., 1998). Recent studies suggest that theta and gamma oscillations play an important role in formation of declarative memory and retrieval (Osipova et al., 2006; Salinas and Sejnowski, 2001) and that changes in synchrony might be important for processes, such as expectation and attention (Salinas and Sejnowski, 2001).

A recent hypothesis (Tallon-Baudry and Bertrand, 1999) assumes that the same mechanism can be extended to the more general idea of object representation, and that gamma activity participates in the activation, retrieval and rehearsal of an internal representation through top-down processes.

The role of neural synchronization in binding and segmentation, and its connection with memory, can be critically analyzed using mathematical models and computer simulation techniques. Indeed, many models of oscillating neural networks, with a different level of complexity and of physiological plausibility have been proposed in past-years, with encouraging results. In these models, the rules used for segmentation are generally inscribed into the synaptic connections linking oscillators. However, most of these studies are focused on low-levels Gestalt cues, such as proximity, smoothness and common fate to segment a visual scene at an early processing visual stage (Terman and Wang, 1995;Wang and Terman, 1997; Li, 1998; Kazanovich and Borisyuk, 2002; Ursino et al., 2003) whereas just a few attempts to use high-levels rules to classify more complex objects at a higher mental level, and to store them in memory, have been performed. Among the others, mention must be made of the pivotal work by von der Malsburg et al. (von der Malsburg and Schneider, 1986; Wang et al., 1990; von der Malsburg and Buhmann, 1992). In a first paper, the authors used a model of oscillating neurons, in which neural coupling reflects similarity of local quality (von der Malsburg and Schneider, 1986). In subsequent papers, the authors proposed models for sensory segmentation in which connections among oscillators encode prior knowledge (Wang et al., 1990) or in which oscillators are sensitive to the position of the cue and encode different features (von der Malsburg and Buhmann, 1992). Lourenc¸o et al. (2000) modified the model by Wang et al. (1990) introducing a law of synaptic change and discussed the problem of learning new memories. Fundamental works which analyze the problem of segmentation, feature extraction and memory inside a single model, are those by Wang and Liu (2002), and of Borisyuk and Kazanovich (2004). These works, however, are explicitly concerned with the problem of visual information, and so consider segmentation separate from recognition. In their models, segmentation is performed at an early processing layer (named the “segmentation

layer" (Wang and Liu, 2002) or "object selection layer" (Borisyuk and Kazanovich, 2004)), by using low-level spatial rules (such as proximity). The segmented objects are then sent to a feature extraction module, and then to a memory layer, which stores information and detects novelty. Moreover, in Wang and Liu (2002) the memory layer sends a feedback to the segmentation layer, to refine segmentation.

Although this subdivision into consecutive layers is certainly acceptable with reference to vision processing, it may be not adequate to represent object recognition involving multiple sensory modalities (such as audition, olfaction, taste, and their binding) which are less dependent on spatial rules.

Aim of the present work is to realize a single network, which implements segmentation of different objects and associative memory (i.e., recognition and recall of previously stored objects) within a single processing stage. The work further develops the same ideas as in the works by von der Malsburg et al. (von der Malsburg and Schneider, 1986; Wang et al., 1990; von der Malsburg and Buhmann, 1992), focusing attention especially on segmentation using high level cues with possible emphasis on higher cortical functions.

Another important aspect, not considered in previous models of autoassociative memory, is that perception of the external world, before being memorized, is ordered in the cortex according to topological maps (Anderson, 1995; Rolls and Treves, 1998). This topological organization of features is ubiquitous in the cortex, and implements a sort of similarity criterion: neurons which signal similar attributes tend to be reciprocally connected and activated together. A similar organization can emerge spontaneously in self-organized networks, such as in the well-known Kohonen's topological maps (Kohonen, 1982; Hertz et al., 1991; Anderson, 1995; Rolls and Treves, 1998). Although this aspect of organization of perception is certainly important for functioning of associative memories, and for segmentation too, we are not aware of any model of autoassociative memory which exploits these topological aspects. Hence, a second important characteristic of this work is the simultaneous use of two different kinds of information storage:

self-organizing maps, based on similarity and on a topological organization of features (as in Kohonen maps) and autoassociative memory, based on correlation and Hebbian learning (as in Hopfield nets). This aspect differentiates our approach from that used in classic autoassociative memories.

The essential concept of our model is that segmentation and recall of high-level objects may be realized by a single network starting from partial or incomplete information, by grouping together a limited set of fundamental features or attributes. These basic features are extracted at a former processing stage, and are arranged in a topologically ordered fashion at some areas of the cortex. Features are then linked together (binding) and separated (segmentation) by synchronization in the γ-range using the similarity and prior knowledge Gestalt rules, in order to arrive at high-level (semantic) object representation.

To illustrate the main ideas of our model, we propose a simple implementation, in which complexity is intentionally maintained at a minimum level. The objective is to show how this network may work, its virtues and robustness. More complex and physiologically founded networks may be naturally built in subsequent works.

MODEL DESCRIPTION

In this section we will first describe the general structure of the model, and its basic working principles, independently of the particular implementation adopted. Then, we will consider a particular very simple implementation, which is of course just a crude schematization of the physiological reality, but whose results are exemplary to show model behavior in a variety of circumstances.

General model structure

i) We assume that the model is composed of N oscillating neural oscillators, subdivided into H distinct cortical areas. Each neural group may be silent, if it does not receive enough excitation, or may oscillate in the γ-frequency band, if excited by a sufficient input.

ii) Each area is devoted to the representation of a specific attribute or feature of the object (for instance color, orientation, geometrical form in case of visual stimuli, tone in case of auditory stimuli, body position in case of somatosensory stimuli, etc…). Hence, one object is represented as the collection of H features (1 feature per each area). We assume that each attribute is not immediately present in the sensory input, but has been extracted from a previous processing stage in the neocortex.

iii) Neural groups within each area represent the value of that particular attribute according to a topological organization. This means that two proximal neural groups in the area signal the presence of two similar values, while distant groups signal the presence of different values. This topological organization is very frequent in the neocortex to represent sensory modalities (let us consider, for instance, the orientation map or the color map in the visual cortex, the tonotopic map in the auditory cortex, etc…).

iv) Neural groups within the same area are connected via lateral excitatory and inhibitory synapses. These lateral connections are organized according to a classical “Mexican hat” disposition. This means that a neuron excites (and is excited by) its proximal neurons in the area, whereas it inhibits (and is inhibited by) more distal neurons. As it is well known, excitatory neurons in the cortex may inhibit proximal neurons in the same area via inhibitory interneurons. Hence, all negative synapses within each area are realized via a bisynaptic connections, from excitatory units to inhibitory units, and then from the latter to other excitatory units.

v) Two neural groups belonging to different areas may be connected via symmetrical excitatory synapses. These reflect the existence of long range functional connections among different cortical areas. These synapses are normally equal to zero, but may assume a positive value when the two neural groups have been simultaneously active in the past during the learning phase. Hence, these synapses store a "prior knowledge" on whether different attributes occurred together in the past during the presentation of objects.

According to the points iii) and iv) before, lateral intra-area connections implement a similarity criterion, i.e., neural groups which signal a similar value for the attribute tend to be simultaneously active. According to point v, inter-area synapses implement a prior knowledge criterion, i.e., attributes which were collected together in the past tend to be grouped again in future experience.

Although in the present work we did not implement any learning process, i.e., model structure and synapses were assigned "a priori", we wish to stress that the hypotheses adopted might reflect real learning procedures in the cortex. In particular, the intra-area topological organization spontaneously emerges in the well-known Kohonen's self-organizing networks (Kohonen, 1982). The inter-area synapses may grow or decay on the basis of Hebbian learning, reflecting the correlation of activity between the pre-synaptic and post-synaptic neurons (Hertz et al., 1991). The latter hypothesis is supported by the recent observation that long-term potentiation and long-term depression strictly depend on the temporal correlation of neurons, with a precision of 10 ms or less (Markram et al., 1997)and can actually be driven by oscillations in the γ-frequency band.

Finally, we wish to stress that, in self-organizing networks, the input to a neuron is generally computed as the scalar product between a sensory vector and the vector of synapses entering the neuron (Hertz et al., 1991). In the present study, for the sake of simplicity, the input to each neuron is described as a scalar quantity, ranging between 0 and 1, which reflects the similarity of the input with the value signaled by the given neuron.

Quantitative model description

The model implemented in this work is composed of *H* different areas. Although each area should be represented as a bi-dimensional structure, in the present study we preferred to describe it as a mono-dimensional chain (see Fig. 1).

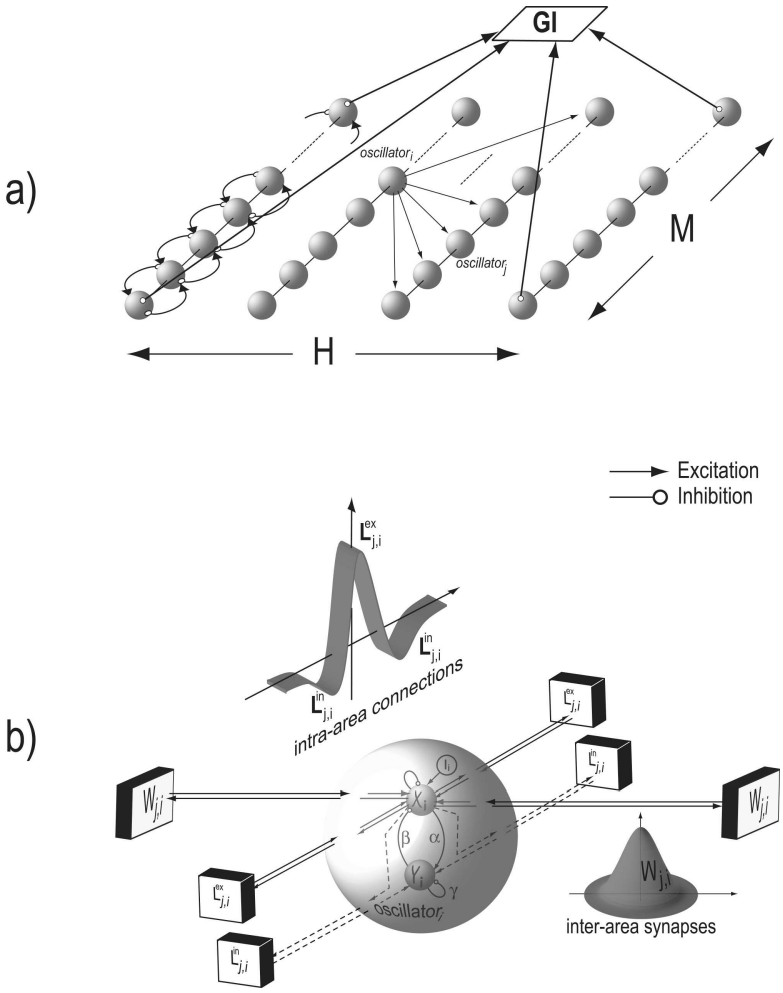


Figure 1 – Schematic diagrams describing the model structure. *Panel a*: the model is composed of *H* different cortical areas each represented as a mono-dimensional chain of *M* Wilson-Cowan oscillators. Each oscillator receives coupling terms both from oscillators in the same area (lateral intra-area connections), and from oscillators in different areas (inter-area synapses). Moreover each unit of the network receives an inhibitory signal from the global inhibitor (GI). *Panel b*: Detail of the coupling terms. L^{ex} represents lateral excitatory connections, while L^{in} lateral inhibitory connections. Both these terms come from other excitatory units in the same area. The lateral excitatory and inhibitory connections have been chosen to have a Mexican hat disposition for intra-are synapses. *W* represents inter-area synapses: they are normally set to zero and assume a positive value only when two oscillators in two distinct areas have been simultaneously activated during the learning phase (see text for an extensive description).

This choice has been adopted to reduce the mathematical complexity of the model. In this way, each neural group is described using a single index, and synapses among groups using two indexes. We think that the main properties of the model can be understood quite well with this simple structure. More complex physiological models, including a greater number of areas and bi-dimensional arrangements of neurons within each area, can be the subject of subsequent versions.

Each area in the model is composed of *M* oscillators. Hence, the total number of oscillators is $N = M \cdot H$. In the following, each area will be denoted with the symbol *h* or *k* (*h*, *k* = 1, 2, …, *H*) and each oscillator with the subscript *i* or *j* (*i*, *j* = 1, 2, …, *N*). Neurons which belong to the *h*-th area are

characterized by an index *j* ranging from (*h*-1)**M* + 1 to *h***M*. In the present study we adopted an exemplary network with 4 areas (*H* = 4) and 100 neural groups per area (*M* = 100).

As already described in our previous works (Ursino et al., 2003), each single oscillator consists of a feedback connection between an excitatory unit, x_i, and an inhibitory unit, y_i, while the output of the network is the activity of all excitatory units. The time derivatives are

$$\frac{d}{dt}x_i(t) = -x_i(t) + H\left(x_i(t) - \beta \cdot y_i(t) + E_i(t) + I_i - \varphi_x - z(t)\right) \tag{1}$$

$$\frac{d}{dt}y_i(t) = -\gamma \cdot y_i(t) + H\left(\alpha \cdot x_i(t) - \varphi_y\right) + J_i(t) \tag{2}$$

where H() represents a sigmoidal activation function defined as

$$H(\psi) = \frac{1}{1 + e^{-\frac{\psi}{T}}} \tag{3}$$

The other parameters in Eqs. (1) and (2) have the following meaning: α and β are positive parameters, defining the coupling from the excitatory to the inhibitory unit, and from the inhibitory to the excitatory unit of the same neural group, respectively. In particular, α significantly influences the amplitude of oscillations. Parameter γ affects the oscillation frequency. The self-excitation of x_i is set to 1, to establish a scale for the synaptic weights. Similarly, the time constant of x_i is set to 1, and represents a scale for time *t*. φ_x and φ_y are offset terms for the sigmoidal functions in the excitatory and inhibitory units. I_i represents an external stimulus for the oscillator in position *i*. E_i and J_i represent coupling terms from all other oscillators in the network. *z(t)* represents the activity of a global inhibitor. This is described with the following algebraic equation (see (Ursino et al., 2003) for more details):

$$z = \left[sign\left(\sum_i x_i - \vartheta \right) + 1 \right] \Big/ 2 \tag{4}$$

According to Eq. 4, the global inhibitor computes the overall excitatory activity in the network, and sends back an inhibitory signal when this activity overcomes a given threshold.

In the present model, the excitatory and inhibitory units within an oscillator receive the same excitation from other areas (to achieve rapid synchronization, see (Ursino et al., 2003)) but different inputs from other excitatory units in the same area, via lateral connections. The latter are arranged according to a Mexican hat. In the following we will denote with *W* the *NxN* matrix of inter-area excitatory connections, with L^{ex} the *NxN* matrix of lateral excitatory connections inside an area, and with L^{in} the *NxN* matrix of lateral inhibition inside an area. It is worth noting that we preferred the matrix form in the implementation by a computer, by giving the value 0 to all synapses which are not involved in the model.

Hence, we can write, in scalar form

$$E_i = \sum_{j=1}^{N} W_{ij} x_j + \sum_{j=1}^{N} L_{ij}^{ex} x_j \tag{5}$$

$$J_i = \sum_{j=1}^{N} W_{ij} x_j + \sum_{j=1}^{N} L_{ij}^{in} x_j \tag{6}$$

or, in matricial form

$$E = (W + L^{ex}) X \tag{5'}$$

$$J = (W + L^{in}) X \tag{6'}$$

where *X* denotes the *N*x1 vector of neuron outputs, and *E*, *J* are *N*x1 vectors of coupling terms.

The lateral excitatory and inhibitory synapses have been chosen to have a Mexican hat disposition for the intra-area connections. This has been realized as the difference of two Gaussian functions, with excitation stronger but narrower than inhibition. Hence, for a generic unit *i* belonging to the area *h*, we have

$$L_{ij}^{ex} = \begin{cases} L_0^{ex} e^{-(i-j)^2/(2\sigma_{ex}^2)} & \text{if } (h-1)M+1 \le j \le hM \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

$$L_{ij}^{in} = \begin{cases} L_0^{in} e^{-(i-j)^2/(2\sigma_{in}^2)} & \text{if } (h-1)M+1 \le j \le hM \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

The previous condition ensures that only units within the same area are connected via lateral syanpses, i.e., all lateral connections between neurons in different areas are set to zero. L_0^{ex}, L_0^{in}, σ_{ex} and σ_{in} are parameters, which establish the strength and extension of these synapses. To have a Mexican hat arrangement we must have: $L_0^{ex} > L_0^{in}$ and $\sigma_{ex} < \sigma_{in}$.

The synapses linking units in different areas (i.e., the quantities W_{ij} in Eqs. 5 and 6) are normally set to zero. We assume that these synapses have a positive value only if the pre-synaptic and post-synaptic units represent two attributes of a previously memorized object. Hence, matrix W embodies a prior knowledge stored within the network.

In our model each stored object is composed of H attributes (H = 4) one per each area. Each attribute is represented as a central or "exact" value, surrounded by an "activation bubble" , which spreads along a few adjacent neurons. As it will be shown in section "Results", the width of the activation bubble depends on parameters describing lateral intra-area synapses (Eq. 7 and 8). The following notation will be used throughout the manuscript, to represent the four "exact" attributes of one object:

$$obj = [a_1\ a_2\ a_3\ a_4] \tag{9}$$

where a_h means position of the neuron signaling the h-th attribute (i.e., the central neuron in the bubble), with $(h\text{-}1)M+1 < a_h < hM$.

In order to build the matrix of inter-area synapses, W, we assume that neurons in two different areas may be connected only if they belong to two bubbles simultaneously active in the same object. Moreover, the strength of this connection decreases with the distance from the "exact attribute". In other words, attributes at the center of the bubble are more strongly connected than neurons at the periphery of the bubble. This hypothesis may correspond to Hebbian learning.

Let us consider two neurons, i and j, which belong to two distinct areas h and k, respectively. The matrix W is initially set to zero. After presentation of the object, synapses connecting neurons which belong to two simultaneously active bubbles are given the value

$$W_{ij} = W_0 \exp\left(-\frac{(i-a_h)^2 + (j-a_k)^2}{2B^2}\right) \qquad \text{if} \quad |i-a_h| \le B \quad and \quad |j-a_k| \le B \tag{10}$$

Otherwise, the synapse is left at the previous value. Parameter B in Eq. (10) represents the width of the bubble, during the learning phase.

An example of the lateral synapses in the model (L_{ij}^{ex} and L_{ij}^{in}) and of the excitatory inter-area synapses (W_{ij}) is shown in Fig. 2, by using the same parameters as In Table I.

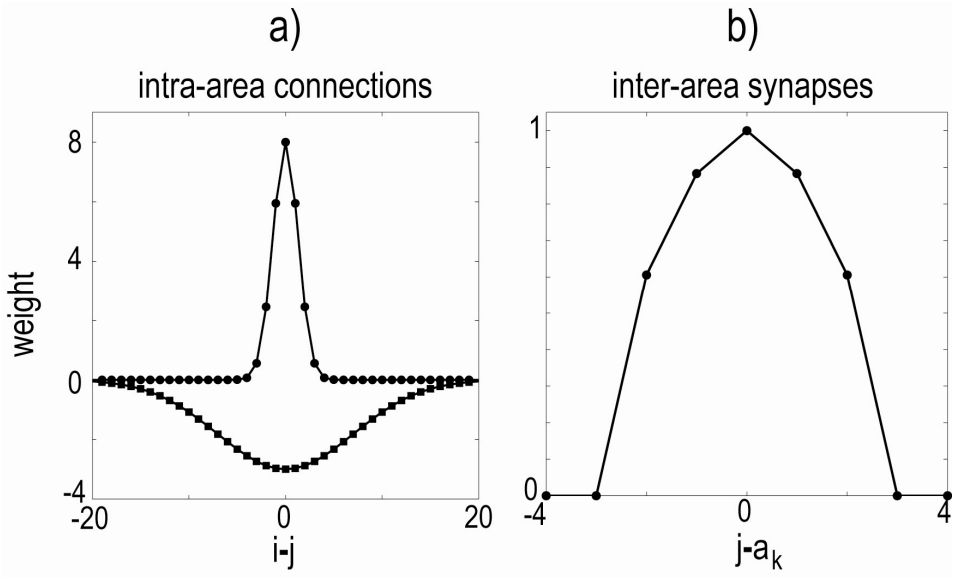


Figure 2 – Synaptic weights of the intra-area (*panel a*) and inter-are connections (*panel b*) when parameters L_0^{ex}, L_0^{in} and *B* are set at their basal values (see Table 1). *Panel a*: Weights of the excitatory (●) and inhibitory (■) lateral connections between neuron *j* and neuron *i* belonging to the same area *h* (*(h-1)M+1≤i,j≤hM*) as a function of their relative position. *Panel b*: Weights of the inter-area synapses between neuron a_h in area *h* (signalling the h-th attribute of an object) and neuron *j* in area *k*, as a function of the position of *j* with respect to neuron a_k (signalling the k-th attribute of the same object). The strength of the connection is maximal (= W_0) between neurons signalling exact properties, and it decreases as the distance from the exact value increases. Outside of the activation bubble (of length *2·B*), the connection is null.

Table 1 - Values for parameters

Wilson-Cowan oscillators			
$\alpha = 0.3$	$\beta = 2.5$	$\gamma = 0.6$	$T = 0.025$
$\varphi_x = 0.7$	$\varphi_y = 0.15$	$\theta = 0.3$	
Lateral intra-area connections			
$L_0^{ex} = 8$	$\sigma_{ex} = 1.3$	$L_0^{in} = 3$	$\sigma_{in} = 7$
Inter-area synapses			
$W_0 = 1$	$B = 2$		

Parameter assignment

All parameters characterizing a single oscillator (Eqs. 1 and 2) have been given the same value used in the previous work (Ursino et al., 2003), with the exception of parameter γ. The value of this parameter determines the maximum number of objects which can be simultaneously perceived. We used $\gamma = 0.6$, which allows perception of three simultaneous objects. Parameters which summarize lateral connections (L_0^{ex}, L_0^{in}, σ_{ex} and σ_{in}) have been initially assigned to have a moderate "activation bubble" within each area, which spreads just along two neural groups at both sides of the central value. A change in these parameters has the effect of varying the dimension of the activation bubble (see section "Results"). Parameter W_0, which represents the strength of prior knowledge, has been assigned so that a previously memorized object can be restored starting from 3 original attributes. Increasing or decreasing this parameter modifies the number of attributes necessary to completely recognize one object (see section "Results"). Parameter B in Eq. 10 has ben given the value $B = 2$, which agrees with the dimension of the bubble during the learning phase.

In section "Results" we will consider initially the following three objects, characterized by different attributes, which are quite distant within each area:

Obj1=[5, 112, 208, 317]; Obj2=[54, 141, 251, 361]; Obj3=[94, 181, 292, 390];

These objects are stored in the inter-area synaptic matrix, W, according to Eq. 10. During the simulations phase, these objects are recalled by stimulating the attributes listed above (via the input I_i in Eq. 1), or proximal attributes.

The case of an object which shares two identical attributes with another one (i.e., the case of strong correlation among objects) will be treated in the last section.

In all subsequent simulations the state variables, x_i, y_i (i = 1, 2, …, N), and the global inhibitor, z, are given a random initial value, ranging between 0 and 1, obtained from a uniform random distribution.

A list of parameters in basal conditions is given in Table 1.

RESULTS

Role of the global inhibitor

The first simulation has been performed in order to elucidate the role of the global inhibitor (GI). For the sake of simplicity, in the first simulation we assumed no lateral connections within the same area, i.e., parameters L_0^{ex} and L_0^{in} are set to zero in Eqs. 7 and 8. This choice has been adopted to emphasize the role of a single oscillator within each area, by avoiding contextual influences between oscillators in the same area. In other words, just the Gestalt property of "prior knowledge" is assumed in this simulation. Accordingly, we set $B = 0$, in Eq. 10. This simulation, however, has been performed using parameter $W_0 = 5$, to have the same total excitation for each oscillating neural group that would occur in the presence of an activation bubble.

Fig. 3 shows the time pattern of all oscillators in the network, assuming that neurons belonging to the first object receive the input $I_i = 0.8$ (with $i \in$ Obj1), while neurons belonging to the second object receive a greater input $I_j = 1.0$ (with $j \in$ Obj2). The input to all other neurons is set to zero, i.e., these neurons do not receive any external stimulus and do not oscillate (including those of the third object).

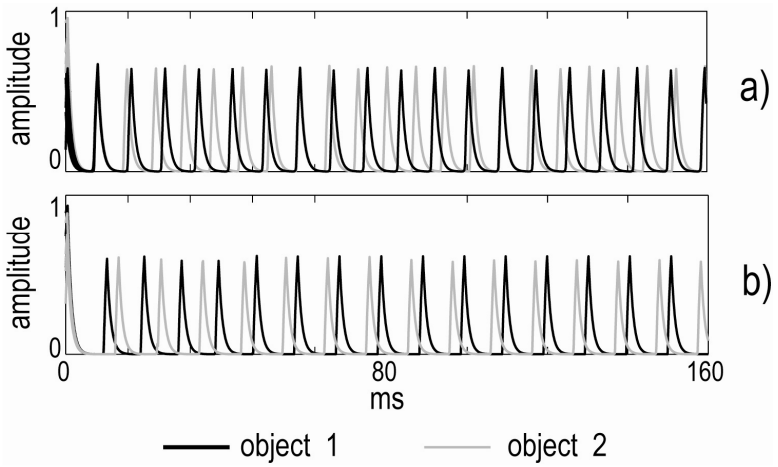


Figure 3 – Time pattern of the network output without (*panel a*) and with (*panel b*) the global inhibitor. In these simulations, two objects (*Obj1* and *Obj2*) are present in the visual scene. *Obj1* receives input *I* = 0.8, while *Obj2* receives input *I* = 1.0. Parameters L_0^{ex}, L_0^{in} and *B* are set at 0. The synchronisation between oscillators within the same object is rapidly achieved in both situations. However, when the global inhibitor is absent (*panel a*), the two objects oscillate with different periods (the object receiving a lower input oscillates more slowly) and in some instants they emerge simultaneously (that is the segmentation problem is not solved). In the presence of the global inhibitor (*panel b*), the objects oscillate with the same period and different phases (that is they are clearly distinguishable).

The upper panel in Fig. 3 shows the time pattern of all oscillators in the network in the absence of the global inhibitor (i.e., we excluded Eq. 4 and set z = 0 in Eq. 1). As it is clear from this figure, neurons reach a rapid synchronization within each object, i.e., the network can easily solve the binding problem thanks to the prior knowledge stored in the synaptic matrix. However, the two objects oscillate with a different period: neurons with a higher input show a higher oscillation frequency and vice versa. As a consequence, the objects cannot be easily separated, and we can observe instants in which two objects emerge simultaneously.

The lower panel of Fig. 3 shows activity in the network vs. time in the presence of the global inhibitor. In this case, objects oscillate with the same frequency, independently of their input, and their phase is clearly distinguishable. Hence, the GI acts as a "metronome", which establishes a single frequency of oscillation for all objects independently of their actual input.

The role of prior knowledge: completion of incomplete knowledge

The subsequent simulations will be performed by including all lateral connections within an area (i.e., parameters L_0^{ex}, L_0^{in} and B have the same value as in Table 1). The presence of lateral connections produces an "activation bubble" within each area, i.e., not only the stimulated neurons oscillate, but also neurons in the same area signaling similar properties. The width of the excitation bubble, hence the degree of specificity depends on a balance between lateral excitation and lateral inhibition, as will be more deeply investigated below.

In this condition, synapses W in the model, which incorporate prior knowledge, not only ensure a rapid synchronization between properties of the same object, but also allow restoration of lacking information. In order to underline this aspect, we performed some simulations by assuming that only some properties of the objects are provided as input to the network, while other properties are lacking, i.e., the network must deal with "incomplete information". Furthermore, some attributes

may be a little changed compared with the exact value, i.e., the network must deal with "corrupted information".

The first simulations (Fig. 4) have been performed assuming the absence of one property in each object.

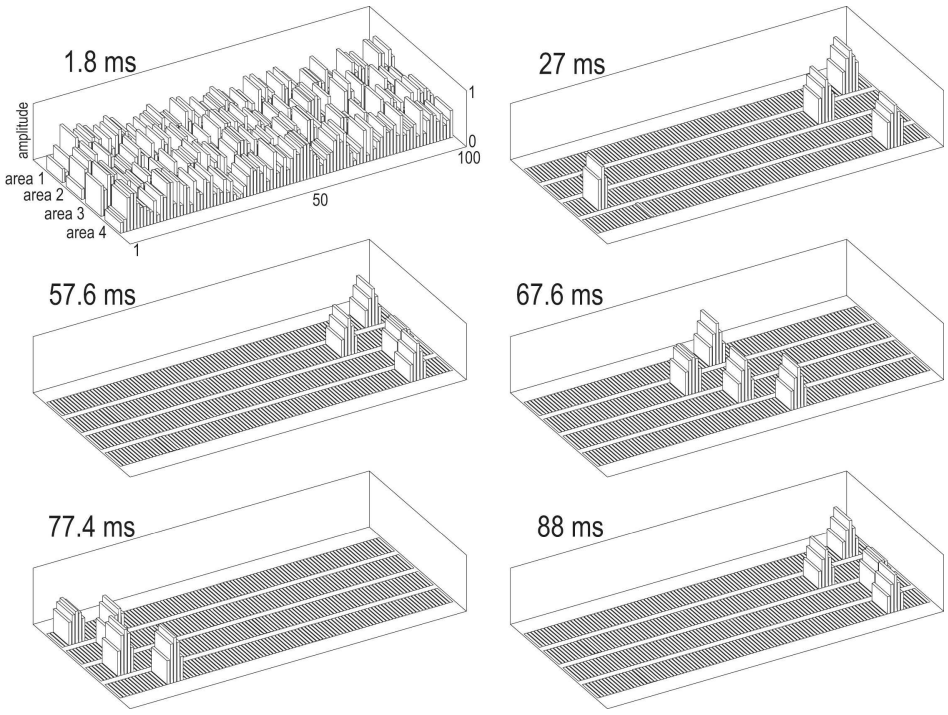


Figure 4 – Network activity at different snapshots during the numerical simulation. Each pixel represents an oscillator. The emerging height is proportional to the corresponding oscillator's activity, i.e., to the value of the excitatory variable, x_i. In the simulation, three objects are present in the sensory input, each lacking of one attribute (that is input to the neuron signalling one attribute is set to 0, while the other three attributes within each object are stimulated with input 0.8). In particular, *Obj1* lacks property a_1, *Obj2* lacks property a_2 and *Obj3* lacks property a_3. After an initial transient, the three objects are perfectly reconstructed by the network, recovering the fourth lacking property. Separation among the three objects is achieved via synchronisation of neurons responding to the same object, but desynchronisation of neurons coding for different objects. The activity of the network in the first snapshot (t = 1.8 ms) derives from the random initial state value (ranging between 0 and 1) assigned to all state variables x_i and y_i and to the global inhibitor z.

In particular, object 1 lacks the property a_1, while object 2 lacks the property a_2 and object 3 lacks the property a_3. The other three properties are stimulated with an input I_i = 0.8. The figure shows network activity in all neural groups at different snapshots during the simulation. The network recovers the lacking property in each object; in other words, the object can be completely reconstructed, re-creating the property which is not given as input.

Fig. 5 shows the same simulation, assuming the absence of two properties in objects 2 and 3 (i.e., only 2 properties over 4 are given as to these objects). With the value W_0 = 1 for the synapses, the information is insufficient to recuperate the entire object.

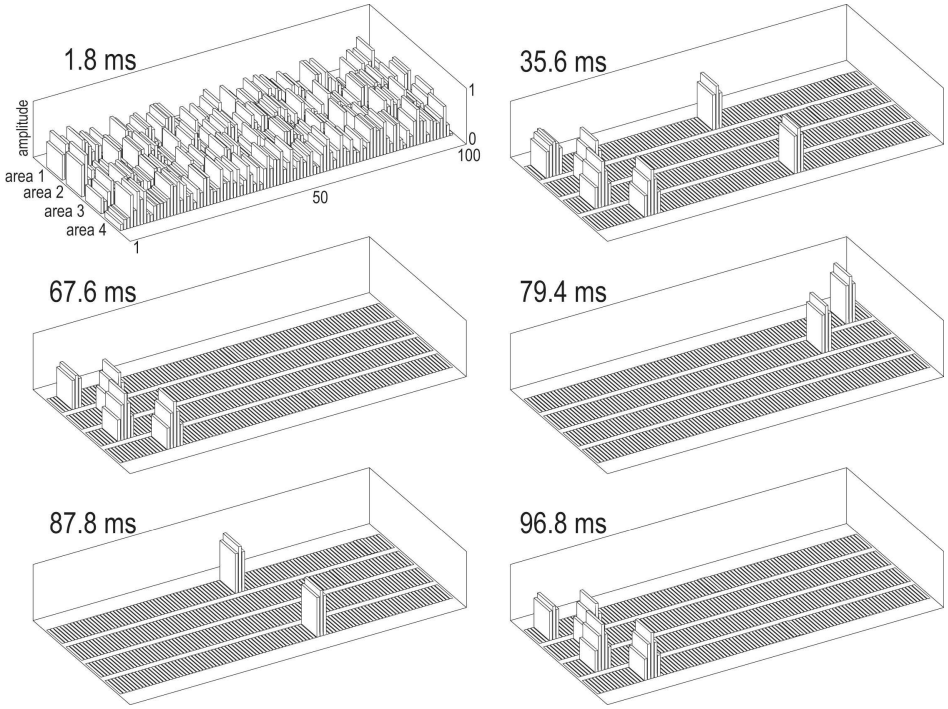


Figure 5 – Network activity at different snapshots during the simulation. Simulation is similar to that of Fig.4; in this case, we assumed the absence of one property in *Obj1* (*a1*), and the absence of two properties in *Obj2* (*a2* and *a3*) and in *Obj3* (*a3* and *a4*). By using the basal value for *W0*, the network is not able to recover the two lacking properties, hence only the two assigned attributes emerge in *Obj2* and in *Obj3*.

Just two properties may be sufficient to recover an entire object from prior knowledge, if we assume a stronger value for the synapses, i.e., $W_0 = 1.5$. The simulation results are not shown for the sake of brevity, since they are almost indistinguishable from those presented in Fig. 4.

In conclusion, the previous simulations illustrate the possibility to reconstruct an entire object from prior knowledge starting from partial information, still satisfying the binding and segmentation problem. Reconstruction from partial information depends on information stored in the synaptic matrix W: The higher the values of the trained synapses (i.e., parameter W_0) the smaller the number of properties necessary to recover an object.

Further simulations were performed assuming that some attributes are corrupted from the "exact" value. These simulations are summarized in Table 2. This table shows the percentage of success is 10 different trials (with random initial values for the network) and the settling time, i.e., the time required for achieving a synchronization.

Table 2 - Percentage of success in 10 trials for 3 simulated conditions, by using basal parameter values (Table 1).

	condition A	**condition B**	**condition C**
success event	recognition of three objects	recognition of two objects	recognition of one object
trial 1	31 ms	40ms	31ms
trial 2	86 ms	30ms	37ms
trial 3	39 ms	27ms	113ms
trial 4	26 ms	53ms	79ms
trial 5	no	36ms	27ms
trial 6	36 ms	51ms	45ms
trial 7	26 ms	28ms	55ms
trial 8	27 ms	27ms	37ms
trial 9	27 ms	37ms	no
trial 10	27 ms	56ms	18ms
number of successes	9/10	10/10	9/10

In the three simulated conditions (A, B and C), all the three objects are present in the sensory input. Each object receives two exact properties and a third corrupted property which may differ by one position (configuration (*i*)) or by two positions (configuration (*ii*)) from the exact value (the fourth property is lacking). In condition A, configuration (*i*) holds for all the three objects. In condition B, configuration (*i*) holds for two objects, whereas configuration (*ii*) holds for one object. In condition C, configuration (*i*) holds for one object, and configuration (*ii*) holds for two objects. The desired behaviour of the network (success event) is the recognition of the objects belonging to configuration (*i*). Values in ms represent the settling time, that is the time necessary to achieve synchronisation in the success event. *No* indicates that the success event was not occurred.

In the first simulation (Table 2, first column) we assumed that the network receives two correct properties for each object with the parameter W_0 set to 1. We remind that, according to Fig. 5, two properties with this value of synapses are insufficient to recover the entire object. However, we now assume that object 1 also receives a property which by just 1 position from one of the lacking properties. Moreover, we assume that also objects 2 and 3 receive a "corrupted" property, which differ by just 1 position from the exact one. Thanks to lateral connections, the existence of this "similar " property is sufficient to evoke the overall object, including the fourth lacking property. Table 2 shows that the percentage of success is 90%. The time required for achieving a synchronization is short (average 30-40 ms).

However, if one property is shifted by 2 from the original one (Table 2, second and third columns), the object cannot be reconstructed, using the values for the activation bubble as in Tab. 1. The remaining two objects, however, are correctly reconstructed. Hence, the network works well to reconstruct correct objects, avoiding reconstruction of objects with excessive corruption.

An interesting characteristic of the model is that the degree of similarity, required to evoke an object from prior knowledge, can be enlarged, acting on the extension of lateral synapses (for instance, on parameters σ_{ex}) even if parameter B in Eq. 10 is still maintained at the original value ($B = 2$). This means that the "activation bubble" was quite small during the storage phase, but it can be modulated during the recovering phase.

When we use low values for the standard deviations σ_{ex} and σ_{in}, an object can be reconstructed from incomplete knowledge only if the input properties are very similar to the original ones. In this case, neuron activity generates a very narrow activation bubble (high specificity, but low sensitivity). By contrast, the use of a larger value for the excitatory standard deviation, $\sigma_{ex,}$ allows object reconstruction even in the presence of a larger dissimilarity between the input properties (low selectivity, high sensitivity) while the activation bubbles spread along some contiguous features.

This aspect is investigated in Tab. 3, while an example is shown in Fig. 6. Here we show the activation bubble obtained with σ_{ex} =1.7, which allows reconstruction of an object although one property is different by 2 positions from the exact one.

From examination of Table 3, it is interesting to observe that the percentage of success decreases dramatically if several objects have simultaneously wrong properties (for instance, two objects have 1 property which differs by 2 positions from the exact one: 80% of success; three objects with 1 property which differs by 2 positions from the exact one: 40% of success), whereas the network can easily deal with a single object which exhibits three properties corrupted by 1 (100% of success).

Table 3 Percentage of success and settling time in 10 trials for five simulated conditions by using a higher value for σ_{ex}, which has been increased from 1.3 to 1.7. All the other parameters are set at their basal value.

	condition A	**condition B**	**condition C**	**condition D**	**condition E**
success event	recognition of three objects	recognition of three objects	recognition of three objects	recognition of three objects	recognition of three objects
trial 1	29 ms	34 ms	33 ms	no	34 ms
trial 2	no	22 ms	29 ms	no	28 ms
trial 3	31 ms	42 ms	no	no	29 ms
trial 4	30 ms	43 ms	38 ms	47 ms	40 ms
trial 5	33 ms	49 ms	34 ms	38 ms	18 ms
trial 6	27 ms	66 ms	63 ms	no	18 ms
trial 7	29 ms	29 ms	42 ms	34 ms	25 ms
trial 8	41 ms	46 ms	34 ms	no	29 ms
trial 9	30 ms	45 ms	51 ms	no	18 ms
trial 10	29 ms	29 ms	no	38 ms	47 ms
number of successes	9/10	10/10	8/10	4/10	10/10

In each simulated condition, three objects are present in the visual scene. The conditions A, B and C are the same as in Table 2. In condition D, the three objects received two exact properties and a third property which differs by two positions from the exact value (that is configuration *(ii)* holds for the three objects). Finally, in the last simulated condition (E), two objects are perfect, whereas the third receives only three corrupted properties, each differing by one position from the exact value. The success event is recognition of all the three objects.

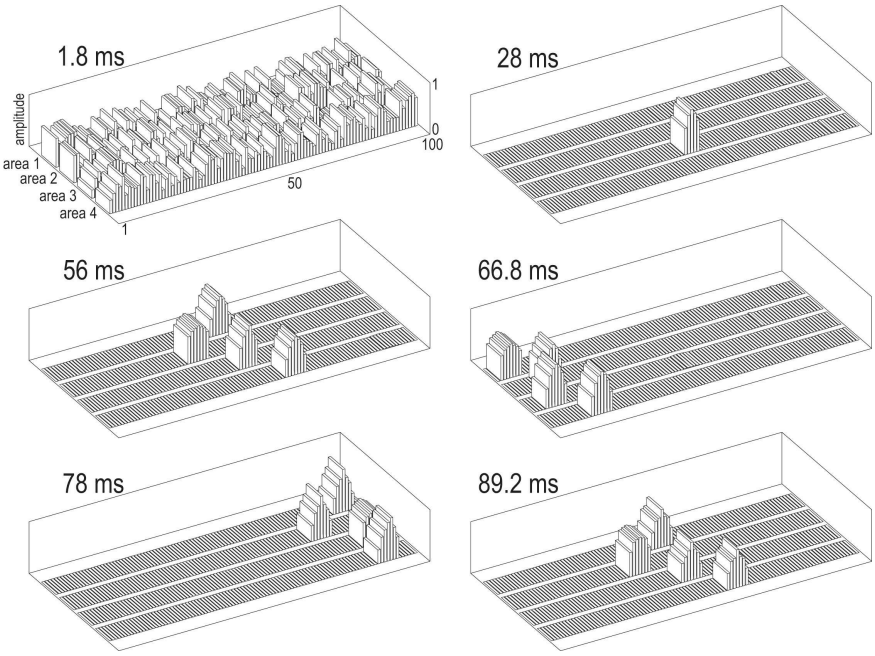


Figure 6 – Effect of increasing the extension of lateral synapses, by augmenting parameter σ_{ex} from 1.3 to 1.7. In this simulation, we assumed that each object receives three properties: two are exact, while the third is corrupted with respect to the exact value. In particular, *Obj1* lacks property *a1* and receives property *a2*+1 (shifted by one position from the exact value). *Obj2* lacks property *a2* and receives property *a3*+1. *Obj3* lacks property *a3* and receives property *a4*+2. As the time snapshots show, all the three objects are recognized, even the object with a property differing by two from the exact value (*Obj3*). Indeed, activation bubble has been enlarged, spreading along several neurons coding similar attributes.

The best compromise between specificity and sensitivity depends on the particular purposes of the network. In the last session we will discuss the possibility that the extension of lateral synapses may be controlled by a feedback from higher centers, reflecting a mechanism of attention, which adapts the network to the specific problem and objective.

Correlation among objects

In all previous simulations we used three distinct objects, which had no common attributes. In other words, input patterns were completely uncorrelated. It is now interesting to analyze the behavior of the network assuming that the objects used in the learning process (i.e., the objects which constitute the "prior knowledge" stored in the synaptic matrix *W*) have some common attributes. Since previous simulations show that 3 attributes may be sufficient to restore an entire object, whereas 2 attributes are insufficient, we assumed that, in the memorization phase, objects can also have two common attributes (this means a 50% correlation). Our main hypothesis is that these objects can still be recognized by the network, provided they are not simultaneously present in the visual scene.

In order to test this hypothesis, we "trained" the network with the objects Obj1 and Obj2 described above, whereas we modified the third memorized object as follows: Obj3 = [54, 141, 292, 390]. In this way, the third object shares the first two attributes with the second. The information about these 3 objects is stored in the synaptic matrix *W*, using the parameter value W_0 = 1 and B = 2. The model behavior has then been tested still using σ_{ex} =1.7 and assuming that the network receives three exact properties of objects 1 and 2, while object 3 is not present in the sensory input. In particular, object 2 receives the two properties in common with object 3, plus an additional property. Results (percentage of success 10/10) show that the network recognizes not only object 1, but also object 2, and correctly restores its fourth property, despite the fact that two input attributes were shared between objects 2 and 3.

However, if objects 2 and 3 were simultaneously present in the input scene, they could not be separated: the network recognizes a single object formed by 6 attributes. A multi-layer network, exploiting also low-level Gestalt rules (see Discussion) may be necessary to solve this problem.

Detection of a different number of objects: role of the parameter γ

Some simulations have been performed to analyze the role of parameter γ in Eq. 2. To this end, we varied γ in the range 0.3 – 0.8 (which corresponds to an oscillation frequency approximately between 30 and 80 Hz, although the true oscillation frequency also depends on the number of objects, see Tab. 4). Simulations have been repeated using two, three, four or even five objects simultaneously present in the same scene. Results are summarized in Tab. 4.

In the absence of any corruption (i.e., when objects are presented with all four exact properties) the network succeeds in segmenting two objects for all values of γ. By contrast, segmentation of three objects requires a value of γ less than approximately 0.8, segmentation of four objects require a value of gamma less than 0.5, while the network never succeeds in managing 5 objects simultaneously. This result suggests that a value of $\gamma < 0.5$ is preferable, since it allows detection of 2, 3 or 4 objects indifferently.

The situation, however, becomes more complex if the objects have corrupted properties (i.e., if at least one property is shifted by one from the correct position). In this case, we observed that the use of a low oscillation frequency produces worst results in case of only two objects. In fact, by using $\gamma = 0.3 – 0.5$ (oscillation frequency approximately 40 Hz), the corrupted property sometimes emerges separately from the rest of the object, as a new independent oscillator. In other words, depending on its initial conditions, a corrupted property might exploit the dead time between oscillators to come into view independently of the other properties. This reconstruction error disappears if the oscillation frequency is increased by raising γ up to 0.8. Similarly, when using three objects with one corrupted property, the best performance is obtained with $\gamma = 0.6$. A higher

value of γ does not allow the synchronisation of the three objects separately, while a lower value of γ sometimes consents the appearance of isolated properties. Finally, if four objects are presented, γ must be decreased down to 0.3 to allow their separate synchronisation, but the management of corrupted objects becomes difficult. We can conclude that: i) the network cannot recognize more than 4 objects simultaneously, a result which agrees with psychophysical studies (Anderson, 1995); ii) in order to have a good reconstruction even in the presence of corrupted information, the oscillation frequency must be increased if the number of objects is reduced; iii) the capacity to manage corrupted objects worsens the greater the number of objects simultaneously present.

Table 4 - Performance of the network as a function of parameter γ and of the number of objects simultaneously present.

	γ = 0.3	γ = 0.5	γ = 0.6	γ = 0.7	γ = 0.8	γ = 0.9
2 objects	YES f = 45 Hz	YES f = 70 Hz	YES f = 80 Hz	YES f = 90 Hz	YES f = 100 Hz	YES f = 105 Hz
3 objects	YES f = 45.4 Hz	YES f = 70 Hz	YES f = 70 Hz	NO f = 70 Hz	NO f = 95 Hz	NO f = 105 Hz
4 objects	YES f = 45 Hz	YES f = 60 Hz	NO f = 70 Hz	NO f = 70 Hz	NO f = 75 Hz	NO f = 90 Hz
5 objects	NO f = 45 Hz	NO	NO	NO	NO	NO

	γ = 0.3	γ = 0.5	γ = 0.6	γ = 0.7	γ = 0.8	γ = 0.9
2 objects	NO f = 60 Hz	NO f = 80 Hz	YES f = 85 Hz	YES f = 90 Hz	YES f = 90 Hz	YES f = 95 Hz
3 objects	NO f = 50 Hz	YES f = 65 Hz	YES f = 65 Hz	YES f = 70 Hz	YES f = 70 Hz	NO f = 85 Hz
4 objects	NO f = 50 Hz	YES/NO * f = 55 Hz	NO f = 60 Hz	NO f = 60 Hz	NO f = 80 Hz	NO f = 85 Hz
5 objects	NO	NO	NO	NO	NO	NO

	γ = 0.3	γ = 0.5	γ = 0.6	γ = 0.7	γ = 0.8	γ = 0.9
2 objects	NO f = 55 Hz	NO f = 85 Hz	NO f = 85 Hz	NO f = 85 Hz	YES f = 85 Hz	YES f = 90 Hz
3 objects	NO f = 55 Hz	NO f = 60 Hz	YES f = 65 Hz	YES f = 70 Hz	NO f = 70 Hz	NO f = 90 Hz
4 objects	NO f = 45 Hz	NO f = 60 Hz	NO f = 60 Hz	NO f = 65 Hz	NO f = 75 Hz	NO f = 85 Hz
5 objects	NO	NO	NO	NO	NO	NO

In the upper table, all objects received all four exact properties. In the middle table, all objects received three exact properties, while one property was shifted by 1. In the bottom table, all objects received two exact properties, while one property was shifted by one and one property was lacking. The oscillation frequency is also reported, when possible (If oscillation frequency is not reported, oscillations were irregular). It is evident the dependence of the optimal value of γ on the number of objects, and the difficulty the segment more than 4 objects.

DISCUSSION

Objects are defined as a collection of different features, which must be grouped together to achieve a correct object reconstruction, but must be taken apart from features of different objects to avoid confusion. Moreover, these features are processed in distinct areas of the brain, and are generally reproduced via a topologically ordered organization. The problem still remains open on how the brain can integrate this sparse and highly distributed information to achieve a coherent and cohesive perception of the external world.

Several authors in past-years have linked fast oscillatory activity to learning and memory, especially in the perception of previously recognized objects. An increase in gamma-band activity has been observed in subjects which identify a fragmented picture after having previously seen the complete one (Gruber et al., 2002).The perception of a meaningful and usual stimulus is able to elicit long-distance gamma-band synchronization, whereas meaningless images fail to induce this synchronism (Rodriguez et al., 1999). Enhanced activity in the gamma and beta bands distinguishes visual memory tasks from non-memory conditions (Tallon-Baudry et al., 1998; Tallon-Baudry et al., 1999).

Aim of this work is to propose a simple model for high-level object representation, which exploits two fundamental Gestalt rules: prior knowledge and similarity, together with synchronization among oscillatory neural populations. Prior knowledge is incorporated into the model in the synapses linking properties in one area to properties in another area. The similarity principle ensues from the presence of lateral (excitatory and inhibitory) synapses within the same area, which are arranged according to the classical “Mexican hat”.

Although in the present work we did not model synaptic plasticity, i.e., the arrangement of synapses has been established “a priori” to reflect the previous principles, this particular arrangement may originate from well-known learning rules: a Hebbian rule, which reflects the coexistence of activation between pre-synaptic and post-synaptic neurons within a narrow time scale (less than 10 ms, reflecting the duration of the active phase of an oscillator, see Fig. 3), may

be responsible for inter-area connectivity and storage of previous-knowledge; a topological representation of features in each area can originate as in the well-known Kohonen self-organizing topological maps (Kohonen, 1982; Hertz et al., 1991). Of course, it is possible that plastic behavior of synapses is associated with some additional attentive or emotional mechanisms, so that important situations are stored in the network and grouped together, by neglecting unimportant or invalid conditions.

The consequence of this specific disposition of synapses is that excitation of a neural group causes the occurrence of an excitation bubble, i.e., activation of one feature is always associated with the activation of similar features in the same area. Assuming the existence of Hebbian reinforcement of synapses among different areas, based on temporal correlation, similarity interferes with prior knowledge: not only the exact features of a perceived object are linked together via inter-area synapses, but also similar features which lie inside the activation bubble, and so are simultaneously co-active.

Thanks to this implementation, the model exhibits a good robustness in the separation between binding and segmentation. Robustness depends on the particular dynamics included in the model, especially on the input that each oscillatory unit receives from the other units and from the global inhibitor. As shown in a previous paper using a representation of oscillator dynamics in the state plane (Ursino et al., 2003), two oscillators which are connected by strong excitatory and inhibitory synapses tend to synchronise quite rapidly, thus ensuring a good solution of the binding problem. In the present model, this synchronisation is warranted both by inter-area synapses (to bind different attributes of the same object) and by lateral synapses (to bind similar attributes inside a single activation bubble). By contrast, segmentation ensues from the activity of the global inhibitor. Without the global inhibitor, only a poor segmentation may be achieved. The global inhibitor stops activity of all oscillators which are not enough synchronized with those maximally active in the present instant, thus separating attributes which do not meet the previous-knowledge and similarity

criteria. Hence, robustness crucially depends on the connection among oscillators, which implement a peculiar dynamics in the network.

In the following, we will first analyze the main elements which differentiate the present model from previous similar ones. Then, the performance of the model, as emerging from the simulations, is discussed. Finally, lines for future extensions and improvements are pointed out.

Differences between the present model and previous ones – Important recent models couple associative memory with oscillation based segmentation, and include prior-knowledge and novelty detection (Wang and Liu, 2002; Borisyuk and Kazanovich, 2004). However, there are profound differences between these models and the present one.

First, in the models mentioned above segmentation and recognition are performed at two different processing stages. Generally, a first layer of neurons segments a visual image on the basis of proximity and spatial connection laws. Subsequently, the information segmented among different objects is sent to a feature extraction layer, and then to an associative memory layer, which recognizes objects and implement prior knowledge rules. By contrast, in the present model, contrarily to the models mentioned above, segmentation and object recognition occur at the same processing stage. We just assumed that a previous layer (not included here, but similar to a classic Kohonen's self-organizing map) extracts the main features of the input, and orders these feature in a topological way. Hence, the present model does not directly utilize spatial information (i.e., it does not aspire to simulate segmentation of visual images). Rather, it aspires to simulate the perception and binding of different sensory modalities (such as hearing, smell, taste, or a combination of them) which lead to the formation of complex high-level concepts, when spatial information plays a minor role (let us consider, for instance, the case of a person perceiving a burning flame at night: features like the smell of smoke, the heat, the brightness, and the crackling of burning wood should be bind together into the complex perception of a flame).

A further important aspect, which differentiates our model from those by (Wang and Liu, 2002; Borisyuk and Kazanovich, 2004), is that, in our work, prior knowledge is implemented assuming a

classic Hebbian reinforcement of synapses. By contrast, prior knowledge in other networks is implemented using more complex rules, which do not have a clear neurophysiological counterpart. For instance, Borisyuk and Kazanovich, (2004) explicitly say that their memory "works without modification of connection strength, which makes the memory different from traditional connectionist learning models".

A further novel aspect of our model is that we exploit a topological organization of features, which is a well-known characteristic of sensory representation in the cortex. Thank to this aspect, our network is able to recall previously memorized objects not only in the presence of partial information, but also when input information exhibits a moderate shift in the input space. The last item significantly differentiates the present model from other associative memory models, based on Hebbian learning rules, such as the well-known Hopfield model (Hopfield, 1982; Hertz et al., 1991) and the associative memory model with oscillatory dynamics and prior knowledge by (Wang et al., 1990) . Indeed, in the Hopfield net, a small shift in the input pattern would result in a wrong recall. By contrast, our network can sustain a small shift in the input properties (together with the total absence of some properties) in virtue of lateral connections, which implement the similarity principle, and in virtue of the use of "activation bubbles". In other words, our model exploits cooperation between topological maps and Hebbian learning, to implement a stronger auto-associative memory which is partly insensitive to an input shift. Oscillatory dynamics is further used to recall and separate multiple objects simultaneously present. We are not aware of other networks which implement Hebbian learning, topological organization (with the related concept of activation bubble) and oscillatory dynamics inside a single consistent structure.

Analysis of model performance - Simulation results, obtained by using a simple network with a minimum of internal complexity, and using an abstract representation of objects (as a collection of 4 features) demonstrate that the proposed mechanism may actually work, producing a high percentage of success (more than 90%). Moreover, simulation results provide some interesting indications on the ease or difficulty to recognize multiple objects, which, if confirmed on subsequent more

physiological models, may represent the subject for future validation studies via psychophysical tests.

A first important aspect of our model is the possibility to recognize objects starting from an initial incomplete representation. This mechanism in part resembles that exploited in auto-associative memories (Hertz et al., 1991). This aspect is controlled by the strength of synapses among different areas, i.e., by parameter W_0 in Eq. (10): the higher the value of W_0, the smaller the number of properties necessary to achieve reconstruction. The disadvantage in the use of higher value of synapses, however, is that objects with a large level of correlation cannot be discriminated. With the value $W_0 = 1$ in our model, we can reconstruct objects even if they exhibit 50% correlation, i.e., if they share half of their properties.

An important point in our model is the trade-off between sensitivity and specificity. Sensitivity is the proportion of true positives that are correctly identified by the test as meeting a certain condition, while specificity is the proportion of true negatives that are correctly identified by the test as not meeting the same condition. Of course, a trade off between sensitivity and specificity is common in pattern recognition systems, and it depends on the choice of the cut-off between positives and negatives. As you increase your sensitivity (true positives) and can identify more cases with a certain condition, you also sacrifice accuracy on identifying those without the condition (specificity). In the present model this trade off basically depends on the implementation of the similarity and previous knowledge rules, i.e., on which difference between the "exact" properties of an object and the "actual" ones can be tolerated to recognize the object itself. We can increase the number of positives by increasing the strength of inter-area synapses (i.e., increasing parameter W_0 in Eq. 10) and/or by using a wider extension for the lateral intra-area synapses (i.e., increasing parameter σ_{ex} in Eq. 8). In the first case, a single object can be recognized by a limited number of properties; in the second case, a single object can be recognized even if some properties are modified compared with the original ones. Of course, the best choice depends on the particular requirements of the system: if objects with similar properties must be separated and recognized as different, one needs

to improve specificity, thus reducing synaptic strength and extension, but obtaining a less robust system.

In our model, if the activation bubble is small and $W_0 = 1$ (as in the simulations summarized in Table 2) the network exhibits a good compromise between sensitivity and specificity. It is able to reconstruct objects if one feature is lacking and another feature is corrupted by one position, while an object is not reconstructed if a property is corrupted by two positions. A greater sensitivity is achieved by extending the lateral excitatory synapses (see Table 3, where just parameter σ_{ex} has been raised from 1.3 to 1.7). In this case, network can recognize an object also if one property is shifted by 2 positions in the topological scale, and one property is completely lacking. It is important to observe that, in moving from Table 2 to Table 3, we did not change synapses reflecting prior knowledge, but only acted on the competitive mechanism within one area. This change might reflect an influence from higher hierarchical centers, for instance an attention mechanism. Reciprocal influences among neurons in the same area, in fact, are not necessarily caused by lateral synapses (as assumed in our mathematical model), but may also reflect top-down strategies from higher centers: for instance, a bi-synaptic connection from one population in one area to a higher level cortical region, and then back from the high-level region to another population in the same original area (Angelucci et al., 2002). According to this idea, the width of the activation bubble in our model (as in Fig. 6) may be controlled by different levels of attention: with a large activation bubble an object is recognized even in the presence of large dissimilarities from the original one. A narrow activation bubble may signify greater awareness to reject objects: hence, only objects quite close to the original ones are detected and reconstructed.

Further important parameters of the model, which may be adjusted to improve the flexibility and comprehensiveness of the proposed mechanism, are the threshold of the global inhibitor (i.e, parameter ϑ in Eq. 4) and the frequency of oscillations (especially affected by parameter γ in Eq. 2). In a previous study (Ursino and La Cara, 2004b) we have shown that the threshold of the global inhibitor must be closely related with the dimension of the object to be detected (that is, with the

number of oscillators simultaneously active). In the present work we used objects represented by just four properties (i.e., four activation bubbles); this justifies the choice of a low threshold $\vartheta = 0.3$. A higher threshold may be required in case of more complex objects, characterized by many features.

In the present simulations, the value of $\gamma = 0.6$ has been chosen to recognize three simultaneous objects. Simulations performed with a different number of objects suggest that the value of γ (hence the oscillation frequency) must be decreased to detect a greater number of objects. Hence, these parameters too may be the target of sophisticate top-down strategies, to optimize network performance depending on the particular context. The dependence of the optimal oscillation frequency on the number of objects is especially important in case of corrupted features. If oscillation frequency is too high, the objects have not enough time to appear separately in time division. However, if the oscillation frequency if too small some corrupted properties may benefit of the dead time between one object and another, and appear as isolated properties instead of synchronizing with the other properties of the same object. We suggest that oscillation frequency should be controlled by higher centers (maybe by an attentive or concentration mechanism) on the basis of the number of objects simultaneously scrutinized.

The previous results can also be figured out in the frequency domain (i.e., in the EEG spectrum). In particular, the model assumes that recognition of familiar objects produces synchronization of neural activities in different regions of the brain, with the appearance of a clear increase of power spectral density in the EEG in the γ-band (20-80 Hz). This aspect is supported by recent results, showing that the early phase-locked gamma-activity might reflect the activation of the neural representation of the familiar target stimulus (Stefanics et al., 2004). In particular, EEG at the scalp surface shows an augmentation of induced gamma-activity after the presentation of meaningful (familiar) as opposed to meaningless (unfamiliar) stimuli, which is accompanied by a dense pattern of significant phase-locking values between distant recording sites (Gruber et al.,

2005). Model results also suggest that the frequency of γ-band oscillations should depend on the number of objects simultaneously detected, and on the level of corruption of the detected objects (i.e., on the complexity of the task). Recent papers suggest that attention enhances gamma-band response (Senkowski et al., 2005). We are not aware of experimental results showing a change in frequency with the complexity of the task. The latter aspect may be the subject of future verification.

Model limitations and lines for future work - Finally, we wish to discuss some lines to improve and enrich the present model. First, the present study utilizes highly schematized stimuli, which are much simpler than those used by the brain to face real scenes. The use of more complex inputs stimuli will represent the major direction for future research.

An important limitation of our model, strictly related with the previous point, is that the network is able to discriminate highly correlated objects (see Results, section "Correlation among objects") only if the objects are not presented together. The simultaneous presentation of two objects with several common features leads to the reconstruction of a unique object with all features shared together. We think that this problem (which may be typical, for instance, of analysis of visual images) may be overcome with the use of spatial information. In fact, two identical features, simultaneously present, may be discriminate according to their spatial position.

The last consideration introduces to an important limitation of our study, i.e., the absence of a spatial organization for the input features. Of course, spatial information may be implemented in future works by using a segmentation layer, similar to that already used in our previous studies (Hopfield, 1982; Hertz et al., 1991; Ursino et al., 2003; Ursino and La Cara, 2004a; Ursino and La Cara, 2004b), based on proximity rules, followed by a feature extraction layer, and finally by the same associative memory layer implemented here. We wish to stress that, in the latter case, the performance of the associative memory layer would be facilitated compared with the present study, since the input properties would be presented already segmented (i.e., in temporal division). The

condition simulated here, in which the input features have not been previously segmented, is certainly more difficult to be managed by the memory net.

In the present simulations we always assumed that all parameters (i.e., the strength and extension of inter-area and intra-area synapses) are the same for all memorized objects. This means that all objects, and all features within an object, have the same importance. Of course, it is possible that these parameters may be different from one object to another (or even from one feature to another in the same object) reflecting its importance or emotional impact. For instance, the use of an higher value for W_0 in some objects (or only in some features) implies that the object can be immediately recognized with presentation of a limited number of essential features. Low values of W_0 signify that many features are necessary to recognize the object. This differentiation, together with the use of a different width for the activation bubble, may produce a flexible and highly specialized system.

In conclusion, we stress that our model does not aspire to reflect present neurophysiological or neuroanatomical knowledge in detail, but rather to propose a computational mechanism, which exploits and extends similar ideas developed in earlier similar models (see (von der Malsburg and Schneider, 1986; Wang et al., 1990; von der Malsburg and Buhmann, 1992)). However, the model general structure can have some support from the present knowledge of memory and learning (Kaiser and Lutzenberger, 2003). Future lines may be directed both toward an improvement of the computational aspects (i.e., the capacity to recognize objects in different conditions with a flexible and reliable performance) or toward a more precise connection with neurophysiology and neuroanatomy.

REFERENCES

Anderson JA. (1995) An introduction to neural networks. MIT Press, Cambridge, MA.

Angelucci A, Levitt JB, Walton JSE, Hupe´ JM, Bullier J, Lund JS. (2002) Circuits for Local and Global Signal Integration in Primary Visual Cortex. J.Neurosci. 22: 8633-8646.

Barlow HB. (1972) Single units and cognition: A neuron doctrine for perceptual psychology. Perception 1: 371-394.

Bhattacharya J, Petsche H, Pereda E. (2001) Long-range synchrony in the gamma band: role in music perception. J.Neurosci. 21: 6329-6337.

Borisyuk R and Kazanovich Y. (2004) Oscillatory model of attention-guided object selection and novelty detection. Neural Networks 17: 899-915.

Eckhorn R. (1999) Neural Mechanisms of Scene Segmentation: Recordings from the Visual Cortex Suggest Basic Circuits for Linking Field Models. IEEE transactions on neural networks 10: 464-479.

Engel AK, Fries P, Singer W. (2001) Dynamic predictions: oscillations and synchrony in top-down processing. Nat.Rev.Neurosci. 2: 704-716.

Gray CM and Singer W. (1989) Stimulus-specific neuronal oscillations in orientations columns of cat visual cortex. Proc.Natl.Acad.Sci.USA 86: 1698-1702.

Gruber T, Müller MM, Keil A. (2002) Modulation of induced gamma band responses in a perceptual learning task in the human EEG. J.Cogn.Neurosci. 14: 732-744.

Gruber T, Trujillo-Barreto NJ, Giabbiconi CM, Valdes-Sosa PA, Muller MM. (2005) Brain electrical tomography (BET) analysis of induced gamma band responses during a simple object recognition task. Neuroimage 19, Epub ahead of print.

Hertz J, Krogh A, Palmer RG. (1991) Introduction to the theory of neural computation. Addison-Wesley, Redwood, CA.

Hopfield JJ. (1982) Neural networks and physical systems with emergent collective computational abilities. Proc.Natl.Acad.Sci.USA 79: 2554-2558.

Kaiser J and Lutzenberger W. (2003) Induced gamma-band activity and human brain function. The Neuroscientist 9: 475-484.

Kazanovich Y and Borisyuk R. (2002) Object selection by an oscillatory neural network. BioSystems 67: 103-111.

Koffka K. (1935) Principles of Gestalt Psychology. Harcourt, New York.

Kohonen T. (1982) Self-organized formation of topologically correct feature maps. Biol.Cybern. 43: 59-69.

Krause CM, Korpilahti P, Porn B, Jantti J, Lang HA. (1998) Automatic auditory word perception as measured by 40 Hz EEG responses. Electroencephalogr.Clin.Neurophysiol. 107: 84-87.

Li Z. (1998) A neural model of contour integration in the primary visual cortex. Neural Computation 10: 903-940.

Lourenço C, Babloyantz A, Hougardy M. (2000) Pattern segmentation in a binary/analog world: unsupervised learning versus memory storing. Neural Networks 13: 71-89.

Lutzenberger W, Pulvermüller F, Birbaumer N. (1994) Words and pseudowords elicit distinct patterns of 30 Hz EEG responses in humans. Neurosci.Lett. 176: 115-118.

Markram H, Lübke J, Frotscher M, Sakmann B. (1997) Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSSs. Science 275: 213-215.

Palmer SE. (1999) Vision Science. MIT Press, Cambridge, MA.

Pulvermüller F, Preissl H, Lutzenberger W, Birbaumer N. (1996) Brain rhythms of language: nouns versus verbs. Eur.J.Neurosci. 8: 937-941.

Rodriguez E, George N, Lachaux JP, Martinerie J, Renault B, Varela FJ. (1999) Perception's shadow: long-distance synchronization of human brain activity. Nature 397: 430-433.

Rolls ET and Treves A. (1998) Neural Networks and Brain Function. Oxford University Press, Oxford.

Senkowski D, Talsma D, Herrmann CS, Woldorff MG. (2005) Multisensory processing and oscillatory gamma responses: effects of spatial selective attention. Exp Brain Res. 166: 411-426.

Singer W. (1993) Synchronization of cortical activity and its putative role in information processing and learning. Annu.Rev.Physiol. 55: 349-374.

Singer W. (1999) Neuronal synchrony: a versatile code for the definition of relations? Neuron 24: 49-65.

Singer W. (2003) Synchronization, binding and expectancy. In: Arbib, M. A., ed., The handbook of brain theory and neural networks. The MIT Press, Cambridge, MA, pp. 1136-1143.

Singer W and Gray CM. (1995) Visual Feature integration and the temporal correlation hypothesis. Ann.Rev.Neurosci. 18: 555-586.

Stefanics G, Jakab A, Bernath L, Kellenyi L, Hernadi I. (2004) EEG early evoked gamma-band synchronization reflects object recognition in visual oddball tasks. Brain Topogr. 16: 261-264.

Tallon-Baudry C, Bertrand O, Peronnet F, Pernier J. (1998) Induced gamma-band activity during the delay of a visual short-term memory task in humans. J.Neurosci. 18: 4244-4254.

Tallon-Baudry C, Kreiter A, Bertrand O. (1999) Sustained and transient oscillatory responses in the gamma and beta bands in a visual short-term memory task in humans. Vis.Neurosci. 16: 449-459.

Terman D and Wang D. (1995) Global competition and local cooperation in a network of oscillators. Physica D 81: 148-176.

Tovee MJ and Rolls ET. (1992) Oscillatory activity is not evident in the primate temporal visual cortex with static stimuli. NeuroReport 3: 369-372.

Ursino M and La Cara GE. (2004a) A model of contextual interactions and contour detection in primary visual cortex. Neural Networks 17: 719-735.

Ursino M and La Cara GE. (2004b) Modeling segmentation of a visual scene via neural oscillators: fragmentation, discovery of details and attention. Network: Comput Neural Syst. 15: 69-89.

Ursino M, La Cara GE, Sarti A. (2003) Binding and segmentation of multiple objects through neural oscillators inhibited by contour information. Biol.Cybern. 89: 56-70.

von der Malsburg C and Buhmann J. (1992) Sensory segmentation with coupled neural oscillators. Biol.Cybern. 67: 233-246.

von der Malsburg C and Schneider W. (1986) A neural cocktail-party processor. Biol.Cybern. 54: 29-40.

Wang D, Buhmann J, and von der Malsburg C. (1990) Pattern segmentation in associative memory. Neural Computation 2: 94-106.

Wang D and Liu X. (2002) Scene analysis by integrating primitive segmentation and associative memory. IEEE Transactions Man and Cybernetics - Part B: Cybernetics 32: 254-268.

Wang D and Terman D. (1997) Image Segmentation Based on Oscillatory Correlation. Neural Computation 9: 805-836.

CHAPTER 1.2. RECOGNITION OF ABSTRACT OBJECTS VIA NEURAL OSCILLATORS: INTERACTION AMONG TOPOLOGICAL ORGANIZATION, ASSOCIATIVE MEMORY AND GAMMA BAND SYNCHRONIZATION

INTRODUCTION

Object representation in various cognitive functions occurs in an activation of distinct and distant cortical areas (Tallon-Baudry and Bertrand, 1999). How can be group together different features, processed in different cortical areas, of the same object and how the activities of neural elements, coding properties of different objects simultaneously presented in the same scene, can be maintained segregated to avoid a false recognition, is a fundamental issue in neurophysiology.

A recent influential hypothesis (Damasio, 1989; Singer and Gray, 1995; Singer, 1999; Varela et al., 2001; von der Malsburg and Schneider, 1986), named “Temporal Correlation Hypothesis”, postulates that neuronal groups representing different aspects of the same object are bound together into a cell assembly through synchronization of their activity in the gamma range (30-100 Hz). According to this hypothesis, neurons that fire in phase would signal attributes of the same object, while neurons firing out of phase would signal attributes in different objects.

Cortical neurons are often engaged in synchronous activity in the γ-frequency band (Gray and Singer, 1989; Singer, 1993). Experimental studies show that oscillations in the γ-range become more evident when subjects perform cognitive tasks, which involve feature binding and/or short term memory (Tallon-Baudry et al., 1998; Engel et al., 2001).

In this line of thinking, Tallon-Baudry and Bertrand (Tallon-Baudry and Bertrand, 1999) formulated a stimulating hypothesis, named the “representational hypothesis”, which assumes that

the same mechanism can be extended to the more general idea of an internal object representation, through top-down processes. Furthermore, this mechanism may also apply across different sensory modalities.

The representational hypothesis, however, involves several problems. In many high-level cognitive tasks multiple objects must be simultaneously recognized from external stimuli: to this end, input stimuli have to be compared with an own internal representation, to recover lacking information from previous experience, and maintain this information in memory avoiding confusion. Furthermore, recognition of objects must be independent of spatial attributes (such as the position, distance and prospective), must spread across different sensory modalities and the objects should be recognized even if they exhibit some moderate changes compared with a previous prototypical representation.

Hence, the problem of multiple object recognition is similar to the classic binding and segmentation problem of sensory perception, with two main differences: objects are considered as collection of features (which allows spatial invariance); the main rules for recognition are high-level Gestalt rules such as prior knowledge and similarity, whereas low-level Gestalt rules are used for early sensory processing (as proximity, collinearity and common fate).

The similarity law is implemented in the cortex through topological maps of features (e.g., the color map and the orientation map in the visual system, the tonotopic map in the auditory system, the somatotopic ‘homunculus’ in the somatosensory system), in which proximal neurons signal similar values of the feature and tend to be reciprocally connected and co-activated. As in self-organizing Kohonen maps, topological organization in the cortex may arise naturally as the result of the long-range inhibition and short-range excitation which characterizes brain connectivity (Rolls and Treves, 1998). Since this kind of connectivity is present at most processing stages in the cortex, topological maps may be involved in high-level functions; for instance, similarity law can apply not only to elementary properties of the sensory stimulus (e.g., edge-orientation, color, tone, etc.), but

also to more complex features or abstract concepts (such as shapes, faces, animate and inanimate objects) (Kohonen and Hari, 1999; Rolls and Deco, 2002).

Implementation of prior knowledge inside the structure of the temporal correlation hypothesis requires connections susceptible of use-dependent modifications with a temporal precision in the millisecond range (that is the temporal resolution of gamma band synchronization). These conditions have been recently supported by experimental data. The tendency of neurons to synchronize their response increases if they are repeatedly engaged in synchronous oscillatory firing in the gamma band, while the synchronizing tendency decreases if neurons are repeatedly engaged in desynchronized oscillatory firing (Singer, 1999). Recently, it was discovered that the temporal order of the presynaptic and the postsynaptic spikes is essential to have synaptic potentiation or depression (Markram et al., 1997). In particular, in order to have potentiation, synaptic inputs must be activated in a short critical window 10-20 ms before post-synaptic spiking (Abbott and Nelson, 2000; Zhang et al., 1998). Such rules have been named “spike timing dependent synaptic plasticity”, and are naturally engaged in oscillation networks (Paulsen and Sejnowski, 2000).

Numerous neural network models of oscillators have been proposed since the mid-eighties in the attempt to elucidate the role of synchronization in sensory information processing and in object recognition and retrieval (von der Malsburg and Schneider, 1986; Eckhorn et al., 1990; Sompolinsky et al., 1990; Wang et al., 1990; Grossberg and Somers, 1991; Hummel and Biederman, 1992; von der Malsburg and Buhmann, 1992; Horn and Opher, 1996; Grossberg and Grunewald, 1997; Wang and Terman, 1997; Hendin et al., 1998; Hoshino, 1998; Campbell et al., 1999; Kuntimad and Ranganath, 1999; Li, 1999; Ranganath and Kuntimad, 1999; Cesmeli and Wang, 2000; Lourenço et al., 2000; Hummel, 2001; Knoblauch and Palm, 2001; Levy et al., 2001; Chen and Wang, 2002; Kazanovich and Borisyuk, 2002; Wang and Liu, 2002; Ursino et al., 2003; Borisyuk and Kazanovich, 2004; Kuzmina et al., 2004; Nakano and Saito, 2004; Ursino and La Cara, 2004; Yazdanbakhsh and Grossberg, 2004; Zhang and Minai, 2004; Zhang et al., 2007; Wu and Chen, 2008; Rao et al., 2008). These models differ as to several aspects: the type of oscillators

used in the network (spiking neurons in pulse-coupled neural networks, networks of relaxation or Wilson-Cowan oscillators); the level of complexity; the physiological reliability. Most of these papers are concerned with segmentation in sensory perception. Von der Malsburg and Schneider in their pivotal paper (von der Malsburg and Schneider, 1986), examined segmentation in auditory modality by a network of oscillators in which neural coupling reflect similarity of local quality. Many models faced the problem of segmentation of a visual image by using primitive grouping Gestalt rules such as proximity, good continuation, pixel/orientation similarity, coherent motion, without involving memory and recognition (Campbell et al., 1999; Kazanovich and Borisyuk, 2002; Kuntimad and Ranganath, 1999; Kuzmina et al., 2004; Li, 1999; Nakano and Saito, 2004; Wang and Terman, 1997; Yazdanbakhsh and Grossberg, 2004; Zhang and Minai, 2004). Some of these models involve multiple layers and areas, either to process separately different cues of the visual input (such as motion and brightness) (Cesmeli and Wang, 2000; Zhang and Minai, 2004) or to reflect the anatomical laminar structure of the visual cortex (Yazdanbakhsh and Grossberg, 2004). Hummel and Biederman (Hummel and Biederman, 1992; Hummel, 2001) developed an oscillatory neural network for shape recognition: a viewpoint-invariant structural description of an object is made possible through temporary synchronization of independent oscillator units representing the parts of the object and the relations among them. More advanced models (Borisyuk and Kazanovich, 2004; Knoblauch and Palm, 2001; Wang and Liu, 2002) address the scene analysis problem via multilayer systems, encompassing initial primitive segmentation and associative memory: objects are first segmented at an early processing layer by using low-level spatial rules, then the segmented objects are sent to a memory layer which performs recognition and learns new memories. Some models based on oscillators synchronization have been proposed with application to the problem of segmentation, recognition and memorization of odors in the olfactory system (Hendin et al., 1998; Hoshino et al., 1998; Lourenço, 2000; Wang et al., 1990). Wang et al. (1990) proposed a model for sensory segmentation in which connections among oscillators encode prior knowledge, but dynamical learning was lacking. The work by Hoshino et al. (1998) introduces a

law of synaptic modifications and considers the problem of recognizing previously learned odors, but not segmentation (i.e. the network deals with one input at a time). Odors segmentation and learning of new patterns are addressed in a single model by Hendin et al. (1998) and by Lourenceo et al. (2000). Finally, some papers by Horn et al. (Horn and Opher, 1996; Levy et al., 2001) investigated the properties of binding, segmentation and learning in oscillatory neural networks both by simulations and analytical calculations. Among the several results, they found that the processes of binding and segmentation were facilitated if the inputs to the system, representing simultaneously activated memories, possessed noisy components.

Inspired by the previous encouraging results of neural network modeling and by recent studies supporting the role of gamma band synchronization in higher cortical functions, we have realized, as described in the Chapter 1.1, a network of Wilson-Cowan oscillators which aspires to simulate segmentation at high cognitive levels, rather than at low sensory levels (Ursino et al., 2006). The network realizes separation of simultaneous objects and their recognition at a single processing layer, by grouping together a set of fundamental features on the basis of two high-level Gestalt rules: similarity and previous knowledge. We assumed that these basic features are extracted at a former processing stage and arranged topologically in some areas of the cortex. Accordingly, the network consists of L (L=4) cortical areas, each devoted to the representation of a specific feature of an object. In order to reduce the mathematical complexity, each area was described as a monodimensional chain of oscillators, arranged in a topological order. The similarity law was realized via intra-area connections disposed as a Mexican hat; the prior knowledge was realized via inter-area connections which link attributes of previously memorized objects. It is worth noticing that both intra and inter-area synapses were assigned “a priori”, that is they were not subjected to any learning process. The network showed good ability both in recovering an object starting from partial and corrupted input information, and in separating multiple objects.

In this paper, the previous model is significantly improved by introducing some important new aspects. The improvements can be summarized as follows:

i) The oscillators are placed in a bi-dimensional lattice. This structure more closely resembles that found in the cerebral cortex. In particular, a bi-dimensional map is more suitable to represent the columnar organization of the cortex, where features may vary both within a column, and from one column to another (Tanaka, 2003). Furthermore, a two-dimensional map encodes a more rich and flexible description of similarity, in which a feature has several neighbors. In particular, each neural oscillator has four nearest neighbors with a distance, *d*, conventionally assumed equal to 1, and four neighbors with a distance $d = \sqrt{2}$. By contrast, in the previous model (see Chapter 1.1) we used a monodimensional chain in which each oscillator has just two nearest neighbors.

ii) The inter-area connections, which reflect previous knowledge, are not assigned a priori as in the former paper, but are subjected to synaptic plasticity. The adopted learning rule is a modified Hebbian rule according to spike timing dependent synaptic plasticity: in order to have potentiation, the presynaptic activity must occur in a short-temporal window (10 ms) before the post-synaptic activity (Abbott and Nelson, 2000). As a consequence, objects can be memorized in the network with different synaptic weights, to mimic the presence in memory of more or less familiar objects, or of objects with different attentive and emotional impact. Furthermore, synapses may be asymmetrical.

iii) In order to recognize or reject objects, we developed a simple “decision network” which works downstream of the network of oscillators. This further layer verifies a certain number of requirements for the oscillator network activity and provides an output signal which automatically indicates the correct or missed detection of each object.

In the following, the general structure of the network will be first presented. Subsequently, we will focus on the learning procedure adopted to train the modifiable synapses. Finally, the

downstream decision network will be described and justified. Simulation results with incomplete, modified and correlated objects are presented and discussed to point out the network capabilities and limitations.

METHOD

The bidimensional network of oscillators

We assume that the model is composed of N neural oscillators, subdivided into L distinct cortical areas (see Figure 1). Each area in the model is composed of $M_1 x M_2$ oscillators. Hence, the total number of oscillators is $N = L \cdot M_1 \cdot M_2$. Each oscillator may be silent, if it does not receive enough excitation, or may oscillate in the γ-frequency band, if excited by a sufficient input.

Each area is devoted to the representation of a specific attribute or feature of the object, according to a topological organization. This means that two proximal neural groups in the area signal the presence of two similar values, while distant groups signal the presence of different values. One object is represented as the collection of L features (one feature per each area). We assume that each attribute is not immediately present in the sensory input, but it has been extracted from a previous processing stage in the neocortex.

Neural groups within the same area are connected via lateral excitatory and inhibitory synapses. These lateral connections are organized according to a classical “Mexican hat” disposition: a neuron excites (and is excited by) its proximal neurons in the area, whereas it inhibits (and is inhibited by) more distal neurons, through the presence of inhibitory interneurons. Such a lateral disposition of synapses is not subject to learning. Furthermore, neural groups belonging to different areas can be connected via excitatory synapses. These synapses are initially equal to zero, but may assume a positive value during a learning phase, to memorize a “prior knowledge” on attributes occurring together in the past during the presentation of objects.

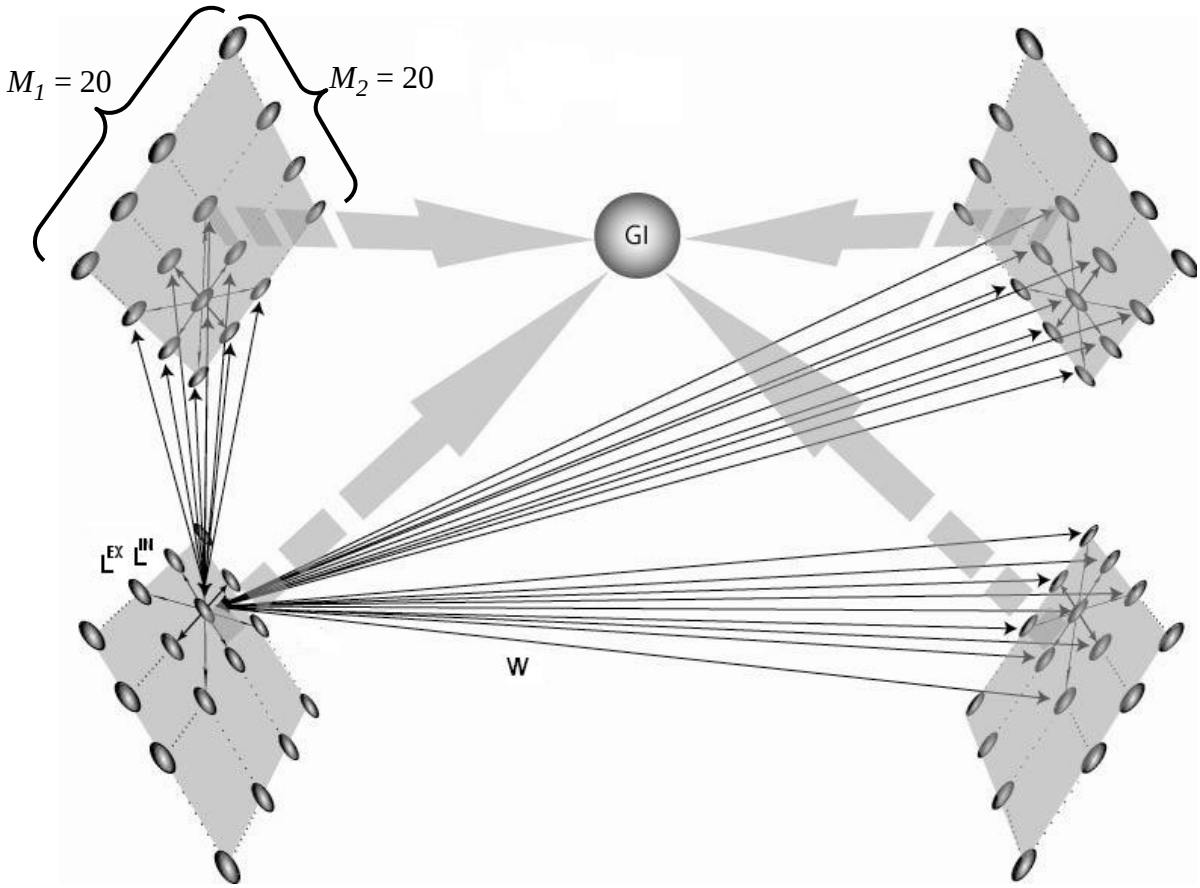


Figure 1 – Schematic diagram describing the general structure of the network. Each grey circle represents an oscillator. Oscillators are organized into 4 distinct areas (shadow squares) of 20x20 elements. Each oscillator is connected with other oscillators in the same area via lateral excitatory and inhibitory intra-area synapses (arrows L^{EX} and L^{IN} within the area) and with other oscillators in different areas via excitatory inter-area synapses (arrows *W* among areas). *GI* represents the *G*lobal *I*nhibitor (see text for details).

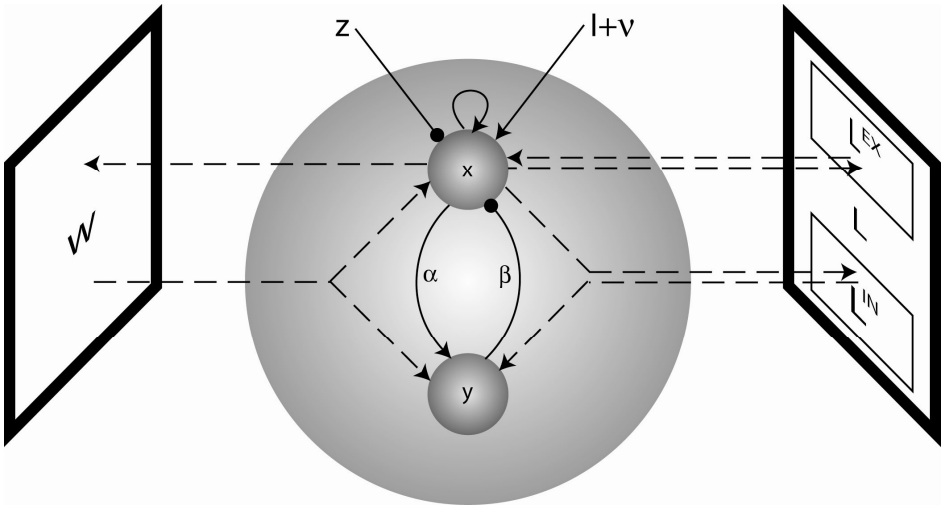


Figure 2 – Structure of the single oscillator and its synaptic connections. Symbol –➤ represents excitatory connections, while symbol —● represents inhibitory connections. The dashed lines indicate synapses to and from other oscillators within the network. *x*: activity of the excitatory population; *y*: activity of the inhibitory population; α and β: strength of the connections between the two populations; *I*: external stimulation; *ν*: random noise superimposed over external input; *z*: activity of the global inhibitor; L^{EX} and L^{IN}: excitatory and inhibitory lateral synapses among oscillators inside the same area. Note that the lateral excitatory synapses connect the excitatory units of two oscillators, while the lateral inhibitory links occur from the excitatory to the inhibitory populations. *W*: long-range synapses among oscillators belonging to different areas. These inter-area synapses link the excitatory unit of one oscillator to the excitatory and inhibitory units of another oscillator in a different area.

In the following, each area will be denoted with the symbol l (l = 1, 2, …, L) and each oscillator with the subscripts ij or hk (i, h = 1, 2, …, M_1; j, k = 1,2,…, M_2). In the present study we adopted an exemplary network with 4 areas (L = 4) and 400 neural groups per area ($M_1 = M_2$ = 20).

As already described in our previous works (Ursino et al., 2003; Ursino and La Cara, 2004), each single oscillator consists of a feedback connection between an excitatory unit, x_{ij}, and an inhibitory unit, y_{ij} (see Figure 2), while the output of the network is the activity of all excitatory units.

The time derivatives are

$$\frac{d}{dt}x_{ij}(t) = -x_{ij}(t) + H\left(x_{ij}(t) - \beta \cdot y_{ij}(t) + E_{ij}(t) + I_{ij} + v_{ij}(t) - \varphi_x - z(t)\right) \quad (1)$$

$$\frac{d}{dt}y_{ij}(t) = -\gamma \cdot y_{ij}(t) + H\left(\alpha \cdot x_{ij}(t) - \varphi_y\right) + J_{ij}(t) \quad (2)$$

where H() represents a sigmoidal activation function defined as

$$H(\psi) = \frac{1}{1 + e^{-\frac{\psi}{T}}} \quad (3)$$

The other parameters in Eqs. (1) and (2) have the following meaning: α and β are positive parameters, defining the coupling from the excitatory to the inhibitory unit, and from the inhibitory to the excitatory unit of the same neural group, respectively. In particular, α significantly influences the amplitude of oscillations. Parameter γ is the reciprocal of a time constant and affects the oscillation frequency. The self-excitation of x_{ij} is set to 1, to establish a scale for the synaptic weights. Similarly, the time constant of x_{ij} is set to 1, and represents a scale for time t. φ_x and φ_y are offset terms for the sigmoidal functions in the excitatory and inhibitory units. I_{ij} represents an external stimulus for the oscillator in position ij, while v_{ij} represents random noise. E_{ij} and J_{ij} represent coupling terms from all other oscillators in the network. $z(t)$ represents the activity of a global inhibitor. This is described with the following algebraic equation (see (Ursino et al., 2003; Ursino and La Cara, 2004) for more details):

$$z = \left[sign\left(\sum_i \sum_j x_{ij} - \theta_z \right) + 1 \right] \Big/ 2 \qquad (4)$$

According to Eq. 4, the global inhibitor computes the overall excitatory activity in the network, and sends back an inhibitory signal (z = 1) when this activity overcomes a given threshold (say θ_z). Its role is to ensure separation among the objects simultaneously present. In particular, the inhibitory signal prevents a subsequent object to pop up as long as a previous object is still active.

Each neuron receives fixed (i.e., not modifiable) excitatory and inhibitory synapses from other neurons in the same area. In the following these synapses will be denoted with the symbols $L^{EX}_{ij,hk}$ and $L^{IN}_{ij,hk}$, respectively, where *ij* denotes the position of the postsynaptic (target) neuron, and *hk* the position of the presynaptic neuron, both in the same area. We assume that the excitatory lateral synapses goes from the presynaptic excitatory unit at position *hk* to the postsynaptic *excitatory* unit at position *ij*. By contrast, the inhibitory synapses go from the presynaptic excitatory unit at position *hk* to the postsynaptic *inhibitory* unit at position *ij*. Hence, according to physiological knowledge (Rolls and Treves, 1998), inhibitory links are realized by means of inhibitory interneurons.

Moreover, a neuron group at position *ij* can also receive a long range excitatory synapse from a neural group located in a different area. These synapses, named $W_{ij,hk}$, link a presynaptic excitatory unit at position *hk*, to both the excitatory and inhibitory units at position *ij* in another area. According to the learning rule used in this work, these synapses are symmetrical. Moreover, we assume that long-range synapses to excitatory and inhibitory units are identical (i.e., we used the same learning rule for both). This corresponds to a parsimony principle. A more complex choice, using different learning rule for long-range excitatory and inhibitory synapses, may be attempted in future works. We did not introduce any delay in the present connections. The effect of delay may also be analyzed in future studies.

As described above, in the present model we do not use direct connections from presynaptic inhibitory units in one neural group to post-synaptic excitatory units in another neural group, i.e.,

we assume that connections originating from interneurons are always confined within their neural group. Actually, in the Wilson-Cowan equations a single oscillator represents the average activity in a population of proximal neurons and, according to physiology, interneurons can make synapses only to their neighbors. Hence, the only physiological way to simulate a long-distance inhibition is to send a connection from an excitatory unit at position *hk* to the inhibitory interneurons at different position *ij*.

The existence of long-range synapses from excitatory units to distal inhibitory interneurons (either in the same area or among different areas in our model) is important to favor a fast synchronization between the corresponding neural groups (see our original paper (Ursino et al., 2003) for a mathematical analysis in the state plane). This kind of connectivity is supported by recent experimental and theoretical studies at least in sensory cortices (Angelucci and Bressolf, 2006). In particular, recent models suggest that this "far inhibition" is not always suppressive but may facilitate the response, depending on the amount of excitatory input to local inhibitors (Schwabe et al., 2006).

According to the previous statements, the coupling terms, E_{ij} and J_{ij} in Eqs. (1) and (2) are computed as follows

$$E_{ij} = \sum_h \sum_k W_{ij,hk} \cdot x_{hk} + \sum_h \sum_k L^{EX}_{ij,hk} \cdot x_{hk} \tag{5}$$

$$J_{ij} = \sum_h \sum_k W_{ij,hk} \cdot x_{hk} + \sum_h \sum_k L^{IN}_{ij,hk} \cdot x_{hk} \tag{6}$$

All terms $W_{ij,hk}$, linking neurons *ij* and *hk* in the *same* area, are set to zero. Similarly, we set to zero all terms $L^{EX}_{ij,hk}$ and $L^{IN}_{ij,hk}$ in which neurons *ij* and *hk* belong to *different* areas.

The Mexican hat disposition for the intra-area connections (see Figure 3) has been realized as the difference of two Gaussian functions, with excitation stronger but narrower than inhibition.

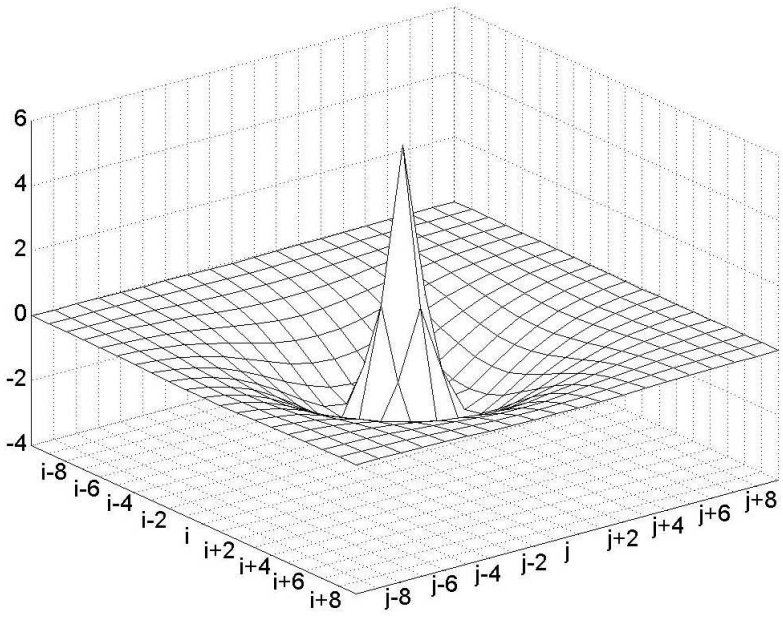


Figure 3 – Intra-area connections: weights of the synapses linking oscillators in position *ij* with the surrounding oscillators in the same area.

Hence,

$$L_{ij,hk}^{EX} = \begin{cases} L_0^{EX} e^{-\left[(i-h)^2+(j-k)^2\right]/(2\sigma_{ex}^2)} & \text{if } ij \text{ and } hk \text{ are in the same area} \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

$$L_{ij,hk}^{IN} = \begin{cases} L_0^{IN} e^{-\left[(i-h)^2+(j-k)^2\right]/(2\sigma_{in}^2)} & \text{if } ij \text{ and } hk \text{ are in the same area} \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

where L_0^{EX} and L_0^{IN} are constant parameters, which establish the strength of lateral (excitatory and inhibitory) synapses, and σ_{ex} and σ_{in} determine the extension of these synapses.

Training of inter-area synapses

Synapses linking neural groups in different areas are trained in order to store and recover objects. In the following, an object will be represented with the notation:

$$obj = [i_1, j_1 \quad i_2, j_2 \quad i_3, j_3 \quad i_4, j_4]$$

where i_l, j_l represent the position of the neuron signaling the *l*-th attribute (l = 1,2,..,L, with L = 4 in our examples).

We distinguish a learning phase, in which inter-area synapses are modified, from a recall phase in which connection strength does not change. We assume that these inter-area synapses are initially set to zero, and that they are increased on the basis of the correlation between the activity of the presynaptic and postsynaptic neurons (time-dependent Hebbian learning).

During the learning phase each object is presented separately from the others. This means that L neural groups (one per each area) receive an input, I_{ij}, different from zero. The input must be high enough to induce oscillation in the same group, but the four inputs can also be different, causing neurons to oscillate with a different frequency. The input to all other groups is set to zero. However, as a consequence of lateral intra-area connections, neural groups close to the stimulated one are also excited and start to oscillate, thus forming an activation bubble. The width of the activation bubble depends on parameters describing lateral intra-area synapses (Eqs. 7 and 8).

A fundamental problem of the learning phase is that the L "activation bubbles" are initially out of phase, due to the absence of any inter-area connection among them (see Figure 4). As a consequence, a simple Hebbian rule based on the instantaneous activity of the presynaptic and post synaptic groups cannot work.

Recent experimental data, however, suggest that synaptic potentiation occurs if the presynaptic inputs precede post-synaptic activity by 10 ms or less (Markram et al., 1997; Abbott and Nelson, 2000). Hence, in our learning phase we assumed that the Hebbian rule depends on the present value of post-synaptic activity, $x_{ij}(t)$, and on the moving average of the presynaptic activity (say $m_{hk}(t)$) computed during the previous 10 ms. We define a moving average signal, reflecting the average activity during the previous 10 ms, as follows

$$m_{hk}(t) = \frac{\sum_{m=0}^{N_s-1} x_{hk}(t - mT_c)}{N_s} \qquad (9)$$

where T_c is the sampling time (in milliseconds), and N_S is the number of samples contained within 10 ms (i.e., $Ns = 10/T_c$). The synapses linking two neurons (say *ij* and *hk*) are then modified as follows during the learning phase

$$\Delta W_{ij,hk}(t+T_c) = W_{ij,hk}(t) + \beta_{ij,hk} \cdot x_{ij}(t) \cdot m_{hk}(t) \tag{10}$$

where $\beta_{ij,hk}$ represents a learning factor.

In order to assign a value for the learning factor, $\beta_{ij,hk}$, in our model we assumed that inter-area synapses cannot overcome a maximum saturation value. This is realized assuming that the learning factor is progressively reduced to zero when the synapse approaches its maximum saturation. Furthermore, neurons belonging to the same area cannot be linked by a long-range synapse. We have

$$\beta_{ij,hk} = \begin{cases} \beta_0 \left(W_{\max} - W_{ij,hk}\right) & \text{if } ij \text{ and } hk \text{ belong to different areas} \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

where W_{max} is the maximum value allowed for any synapse, and $\beta_0 W_{\max}$ is the maximum learning factor (i.e., the learning factor when the synapse is zero).

Eqs. (10) and (11) require a few comments. First, the synapses are modified according to the Hebb rule only if the presynaptic and postsynaptic neurons belong to different areas. Second, according to Eq. (10), the array of inter-area synapses can be asymmetrical. Third, Eq. (11) implies that each synapse approximately increases according to a sigmoidal relationship, with upper saturation W_{max}. The slope of this sigmoidal relationship (hence the increasing rate) is determined by parameter β_0.

The strength of the synapse $W_{ij,hk}$ at the end of the presentation of one object, depends on two factors: parameter β_0 and the duration of the period along which the object is presented to the network. The longer is this period, the higher is the value of synapses, and the strength of memorization.

Parameter β_0 is assumed to be the same for all synapses at a given instant. However, this parameter can be modified from one object to the next during the learning phase. In this way, the model can account for objects with a different relevance (for instance, for the effect of attention, emotion, expectation and for all other factors which may affect storage).

In the present work, we did not use a decay term in the Hebbian rule. This may improve the biological plausibility of the model and could be introduced in future versions.

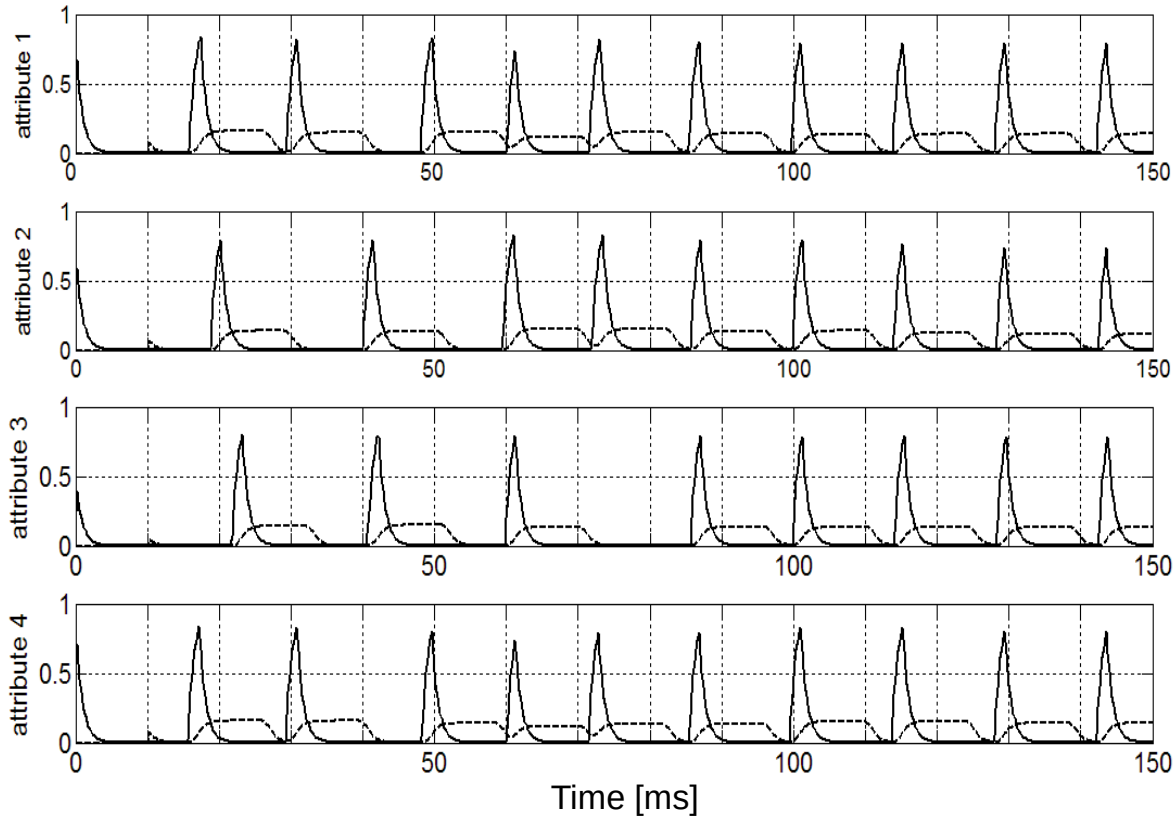


Figure 4 – Learning phase: instantaneous activity (solid black line) and moving averaged activity (dashed grey line) of four oscillators representing the exact attributes of an object, during the training of the inter-area synapses. Oscillators receive external input 0.8 (attribute 1 and 4), 0.7 (attribute 2) and 0.6 (attribute 3). and are affected by a random noise term. Initially, the first and the fourth attributes oscillate with a higher frequency than the other two; then, rapidly the four oscillators synchronize due to the formation of inter-area synapses. These inter-area synapses are created according to a Hebbian rule, thanks to the partial temporal superimposition of the moving average presynaptic signals with the instantaneous activity of the post-synaptic neuron.

An example of the training phase is shown in Figure 4, where the temporal activity of the four central neurons (i.e., the neurons which signal the exact attributes) is presented. In this figure we display also the "moving average signal" (i.e., the quantity $m_{hk}(t)$ in Eq. 9) for the four neurons. The neurons received the input values 0.8, 0.7, 0.6 and 0.8, respectively; hence, the first and the last

initially oscillate with higher frequency than the other two. However, as it is clear from Figure 4 , the activities of the four neurons rapidly synchronize, due to the formation of inter-area synapses among them.

An example of the synapses linking the neuron 5,5 in object 1 (Obj1 = [5,5 5,35 35,35 35,5]) with all the other neurons in the network after the learning phase is present in Figure 5 (in particular, this figure shows the values of the array $W_{55,hk}$ with h = 1, 2, … , 40; k = 1, 2, … , 40; i.e., , it represents the inter-area synapses which *target* into the neuron 5,5). As it is clear from this figure, after the learning phase the neuron receives synapses from the other neurons in the same object (i.e., from neurons 5,35 35,35 35,5) and, although with smaller strength, from the other neurons in activation bubbles.

After the learning phase, the network can be used to detect and reconstruct objects, even in the presence of lacking or modified information.

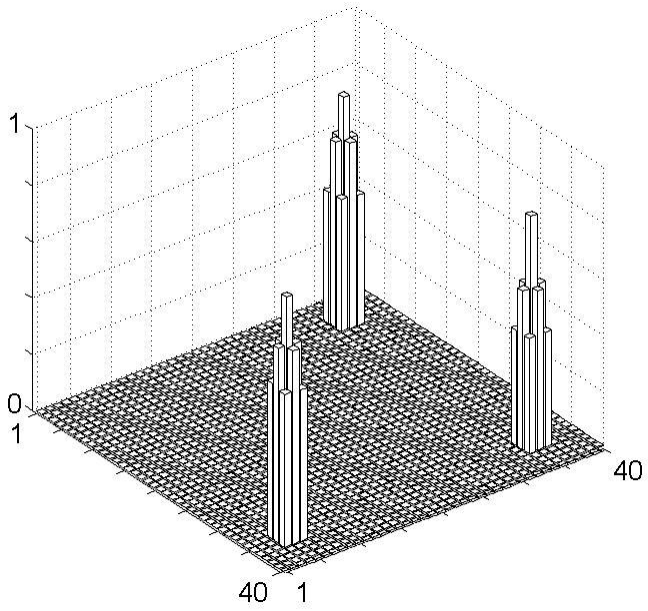


Figure 5 – Learning of object 1 (Obj1) and inter-area synapses. The four exact attributes of Obj1 are codified by the oscillators in positions [5,5 5,35 35,35 35,5]. The figure shows the values of the synapses linking oscillator 5,5 with oscillators in the other areas at the end of the learning process of Obj1. In particular, the figure displays the array $W_{55,hk}$ (representing the inter-area synapses directed to oscillators 5,5) by means of a three-dimensional graph: the x,y plane represents the coordinate *hk* within the network, and the height of the pixel in position *hk* represents the value of the synapse linking oscillator *hk* with oscillator 5,5. Note that the oscillator 5,5 receives the strongest synapses from the oscillators in the other areas signalling the exact attributes (that is oscillators 5,35 35,35 35,5) and weaker synapses from the other oscillators within the activation bubbles.

The decision network

In our network, object recognition is considered well-done if the L neural oscillators which signal the attributes of the object oscillate in synchrony, and with a different phase from attributes of different objects (this means that the network produces four synchronous activation bubbles,

located in the correct positions). In previous works, the success or failure of the network in recognizing objects was simply decided by the user, looking at the temporal pattern of network activity. A new aspect of the present model is that we designed a "decision network", located downstream the network of Wilson-Cowan oscillators, with the task of producing a "true" output only during the correct detection of an object.

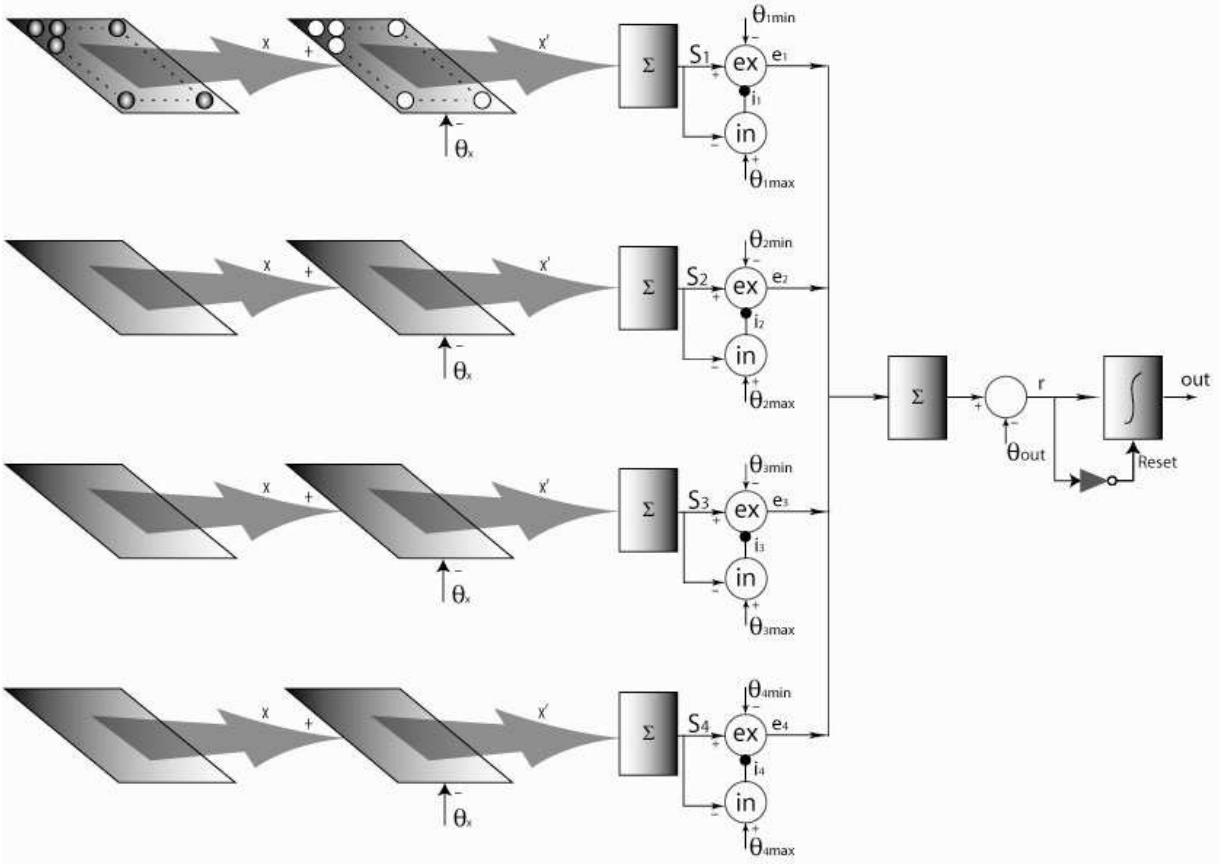


Figure 6 – Schematic diagram of the decision network. The white circles represent binary neurons, which produce output 0 or 1 depending on whether the input is below or above the given threshold (indicated by θ). The activity of all oscillators within each cortical area is given as input to a layer of binary neurons and compared with a threshold θ_x to detect those oscillators that are active at any instant. The overall number of active oscillators within each area is compared with two thresholds (θ_{min} and θ_{max}, whit $\theta_{min} < \theta_{max}$), by means of two binary neurons, labelled as excitatory (*ex*) and inhibitory (*in*) neuron, respectively. The inhibitory neuron inhibits the excitatory neuron; therefore, the output e_l of the excitatory neuron assumes value 1 only if the number of active oscillators in the corresponding area lies between θ_{min} and θ_{max}. The downstream portion of the network compares the sum of all the excitatory neurons with a threshold θ_{out} (= *L*-0.5), to detect whether *L* attributes are simultaneously present within the network; then, the obtained signal (*r*) is integrated in order to verify that the *L* attributes are maintained for a certain time interval. The integrator is reset as soon as the signal *r* is switched off.

In order to detect an object, this decision network must verify the following requirements: i) there is an "activation bubble" in any area. To this end, the network verifies that the number of active oscillators in any area, at a given instant, overcomes a first threshold, assuming that a single

bubble is composed at least of a minimum number of coactive units; ii) any area must produces just a single activation bubble at a given instant. In order to check this requirement, the network verifies that the number of active oscillators within an area, at a given instant, does not overcome a second threshold, assuming that two simultaneous activation bubbles (or a single bubble with excessive width) would produce too much activity in the area. Of course, the second threshold is higher than the first; iii) the conditions i and ii must be verified along a certain time interval, to ensure the continuity of object perception.

The previous requirements are detected by the network illustrated in Figure 6, which corresponds to the following equations:

a) a first layer of binary neurons compares the activity of all oscillators, $x_{ij}(t)$, with a threshold, to detect only those oscillators which are sufficiently active at a given instant. By denoting with x'_{ij}(t) the activity of a neuron in this network, located at position ij, we have:

$$x'_{ij}(t) = \lfloor sign(x_{ij}(t) - \theta_x) + 1 \rfloor / 2 \tag{12}$$

According to Eq. (12), neurons of this layer are binary in type, producing output 0 or 1 depending on whether the activity in the corresponding oscillator is below or above θ_x.

b) L downstream *inhibitory* binary neurons (one per area) control whether the number of active oscillators in that area at a given instant is above a maximum tolerated threshold (say θ_{max}). By denoting with i_l the output of these neurons (l = 1, 2, …, L) we have

$$S_l(t) = \sum_i \sum_j x'_{ij}(t) \qquad \text{with } ij \text{ belonging to the area } l \tag{13}$$

$$i_l(t) = \lfloor sign(S_l(t) - \theta_{\max}) + 1 \rfloor / 2 \tag{14}$$

According to Eqs. (13) and (14), the output i_l of the neuron is set to 1 when the activity in that area is too high (this is the case of two simultaneous activation bubbles in that area, or of an activation bubble too large). Otherwise, its activity is zero. The activity of this neuron is used to inhibit object detection (see point c, below).

c) Other *L excitatory* binary neurons (one per area) verify whether the number of active oscillators in that area at a given instant is not too low (i.e., it is above a given minimum threshold, say θ_{min}). However, this neuron is inhibited by the activity i_l of the inhibitory neuron (see Eq. 14). By denoting with e_l (l = 1, 2, …, L) the activity of these neurons, we have

$$e_l(t) = \frac{[sign(S_l(t) - \theta_{\min}) + 1]}{2} - i_l(t) \tag{15}$$

where the quantity $S_l(t)$ has the same meaning as in Eq. (13).

In conclusion, the binary quantity $e_l(t)$ signals that there is one and only one property in that area

d) Finally, a downstream decision neuron scrutinizes whether the network exhibits L attributes simultaneously, to detect an acceptable object. Moreover, we require that these attributes are maintained for a certain time interval. In order to achieve these requirements, this decision neuron computes the sum of all excitatory neurons ($e_l(t)$), compares this sum with a threshold $\theta_{out} = L$ -0.5, (to detect that L attributes are present) and integrates the signal thus obtained. We can write:

$$r(t) = \left[sign\left(\sum_{l=1}^{L} e_l(t) - \theta_{out} \right) + 1 \right] \Big/ 2 \tag{16}$$

$$out(t) = \int_{t_{reset}}^{t} r(t)dt \tag{17}$$

The symbol t_{reset} in Eq. (17) designates the previous instant in which the integrator has been reset to zero. We decided that the reset occurs as soon as the signal $r(t)$ is switched off (i.e., the signal $r(t)$ is used both as the input to the integrator, and as a reset signal to start integration again). According to Eq. (17), the output signal starts to increase as soon as L attributes are simultaneously detected. The longer the time during which these attributes keep on together, the higher the value reached by $out(t)$. A high value of $out(t)$ (for instance above 6, see section Results) signify that the object is perceived with good reliability.

The parameter values are reported in Table 1. A justification for the parameters concerning the Wilson-Cowan oscillators can be found in previous works (Ursino et al., 2003). Lateral intra-area connections were given to ensure network stability, and to have an activation bubble which spreads for a few units around the central oscillator. The value of β_0 in the Hebb rule was given to have learning periods of the order of 1 second. W_{max} was assigned to ensure a strong synchronization among two interconnected bubbles in case of maximal synapses. Finally, parameters of the decision network were given on an empirical basis, by considering the dimension of the bubbles, in order to maximize the performance of the detector.

Table 1 – Values for parameters

Wilson Cowan oscillators	
α	0.3
β	2.5
γ	0.5
T	0.025
φ_x	0.7
φ_y	0.15
θ_z	1.9
Lateral intra-area connections	
L_0^{EX}	9
σ_{ex}	0.8
L_0^{IN}	3
σ_{in}	3.5
Hebbian rule for inter area synapses	
β_0	0.033
W_{max}	1
Decision network	
θ_x	0.2
θ_{mi}	3
θ_{max}	11
θ_{out}	3.5
Gaussian noise	
σ	0.02
Object attributes	
Obj1	[5,5 5,35 35,35 35,5]
Obj2	[15,15 15,25 25,25 25,15]
Obj3	[15,5 15,35 25,35 25,5]
Obj4	[5,15 5,25 35,25 35,15

Simulations

We have performed different kinds of simulations to test the limits and the capabilities of the model. In the first simulations, we tested the capability of the model to reconstruct and recognize multiple objects starting from partial information, both in case of strong and weak memorization of objects. To this end, we stimulated the network with several objects simultaneously, but assigning a different number of properties (from 1 to 4) to each object. Furthermore, we repeated the simulations with different values of the parameter γ, which establishes the frequency of oscillations.

In another set of simulations we introduced a Gaussian white noise term both in the learning phase and during the recall phase. In these simulations network started from a random initial condition, and we repeated every simulations tenfold to perform a statistical analysis.

In a subsequent set of tests, we tested model capacity to recognize objects also in presence of some changes in the input properties (i.e., we assumed that objects were modified compared with those used in the learning phase). Further simulations analyze the capability of the network to reconstruct and recognize a different number of objects (from 1 to 4), correlated objects (i.e., objects sharing some properties) and objects with different distances among their properties in the topological space.

RESULTS

Simulations with incomplete objects

The network was trained with three objects (named Obj1, Obj2 and Obj3), represented by the following properties: Obj1 = [5,5 5,35 35,35 35,5], Obj2 = [15,15 15,25 25,25 25,15], Obj3 = [15,5 15,35 25,35 25,5]. Two different training sessions were performed, starting with null inter-area synapses. In the first, the three objects were separately presented for 1600 ms (Obj1), 1400 ms (Obj2) and 1100 ms (Obj3), while we used the same value of parameter β_0 for all objects. With these values, synapses linking attributes of the first and the second object are stronger than the synapses in the last one: however all objects have strong memorization and can be reconstructed starting from two attributes only. In a second training session, duration of the learning phase was 1300 ms for the first object, 1100 ms for the second and 800 ms for the third. As a consequence, the third object is stored more weakly than the others, and requires at least three attributes to be correctly reconstructed. Objects 1 and 2 require just two attributes.

Table 2 – Results of 17 different simulations performed by stimulating the network with three objects simultaneously but assigning a different number of properties (from 1 to 4) to each object.

				STRONG							WEAK			
TRIALS	INPUTS						γ	INPUTS						γ
	Obj1	Obj2	Obj3	Obj1	Obj2	Obj3		Obj1	Obj2	Obj3	Obj1	Obj2	Obj3	
1	4	4	4	yes	yes	yes	0.3 (T>180ms), 0.4, 0.5	4	4	4	yes	yes	yes	0.5
2	4	4	3	yes	yes	yes	0.3, 0.4, 0.5	4	4	3	yes	yes	yes	0.4
3	4	3	4	yes	yes	yes	0.3, 0.4, 0.5	4	3	4	yes	yes	yes	0.3
4	3	4	4	yes	yes	yes	0.3, 0.5	3	4	4	yes	yes	yes	0.3
5	4	4	2	yes	yes	yes	(0.3, 0.4 T>200ms), 0.5	4	4	2	yes	yes	no	0.3 / 0.4
6	4	2	4	yes	yes	yes	0.3, 0.4, 0.5	4	2	4	yes	yes	yes	0.3
7	2	4	4	yes	yes	yes	0.3, 0.4 (T>200ms), 0.5	2	4	4	yes	yes	yes	0.3 / 0.5
8	1	4	4	no	yes	yes	0.3, 0.4 (T>300ms), 0.5	1	4	4	no	yes	yes	0.3 / 0.5
9	4	3	3	yes	yes	yes	0.3, 0.4, 0.5	4	3	3	yes	yes	yes	0.3
10	3	4	3	yes	yes	yes	0.3, 0.4 (T>300ms), 0.5	3	4	3	yes	yes	yes	0.3
11	3	3	4	yes	yes	yes	0.3, 0.4	3	3	4	yes	yes	yes	0.3
12	4	2	2	yes	yes	yes	0.4, 0.5	4	2	2	yes	yes	no	0.3 / 0.5
13	2	4	2	yes	yes	yes	0.3, 0.4, 0.5	2	4	2	yes	yes	no	0.3 / 0.5
14	2	2	4	yes	yes	yes	0.3, 0.4, 0.5	2	2	4	yes	yes	yes	0.3 / 0.5
15	3	3	3	yes	yes	yes	0.3, 0.4	3	3	3	yes	yes	yes	0.3
16	2	3	3	yes	yes	yes	0.3, 0.4, 0.5	2	3	3	yes	yes	yes	0.3 / 0.5
17	1	3	3	no	yes	yes	0.3, 0.4, 0.5	1	3	3	no	yes	yes	0.3

Objects received two different previous memorizations: a strong memorization (learning time: Obj1 = 1600 ms; Obj2 = 1400 ms; Obj3 = 1100 ms) and a weak memorization (Obj1 = 1300 ms; Obj2 = 1100 ms; Obj3 = 800 ms). In both cases, two input properties are sufficient for the reconstruction of Obj1 and Obj2; whereas Obj3 needs two properties in the first case and three properties in the second case. The number of input features for each object is specified in the second column. The third column describes whether the object is reconstructed (yes) or not (no). The fourth column signals the values of γ which allow a correct segmentation (the cases which requires a time, say T, longer than 150 ms to achieve segmentation are also indicated). Note that Obj1 in trials 8 and 17 (strong and weak memorization) and Obj3 in trials 5, 12 and 13 (weak memorization) are not reconstructed: these results are correct since these objects received too few attributes with respect to their memorization.

Table 2 summarizes the results of 17 simulations performed with the model, first using strong memorization and then weaker memorization (see the legend for explanations). Results can be described as follows: i) In simple conditions (for instance when all objects have strong memorization and most features are present in input) segmentation can be achieved with $\gamma = 0.5$ (which corresponds to an oscillation frequency as high as 55 Hz). However, when segmentation becomes more difficult (for instance, in case of weaker memorization and/or with a smaller number of input features) the value of γmust be reduced to 0.4 or 0.3 (i.e., the frequency is reduced to 40-45 Hz); ii) in case of strong memorization, the three objects can be correctly segmented and reconstructed, provided at least two of their attributes are given to the network. If only one attribute is assigned to one object, the object cannot be reconstructed, but the remaining objects can be correctly segmented as well. Only in two cases the correct reconstruction requires a lower value of γ(equal or less than 0.4, trials 11 and 15) iii) In case of weaker memorization, the network can still segment and memorize objects, provided at least two attributes are given for Obj1 and Obj2, and

three attributes are given to Obj3: however, in several cases the value of parameter γ must be decreased down to 0.3 to obtain a correct segmentation; iv) The downstream decision network provides a high output only when all four attributes of the object are simultaneously recollected (i.e., the corresponding neural groups oscillate in synchrony). In conditions where the four attributes are not retrieved, the output of the decision network remains at a low level, thus signaling that the object is not recognized.

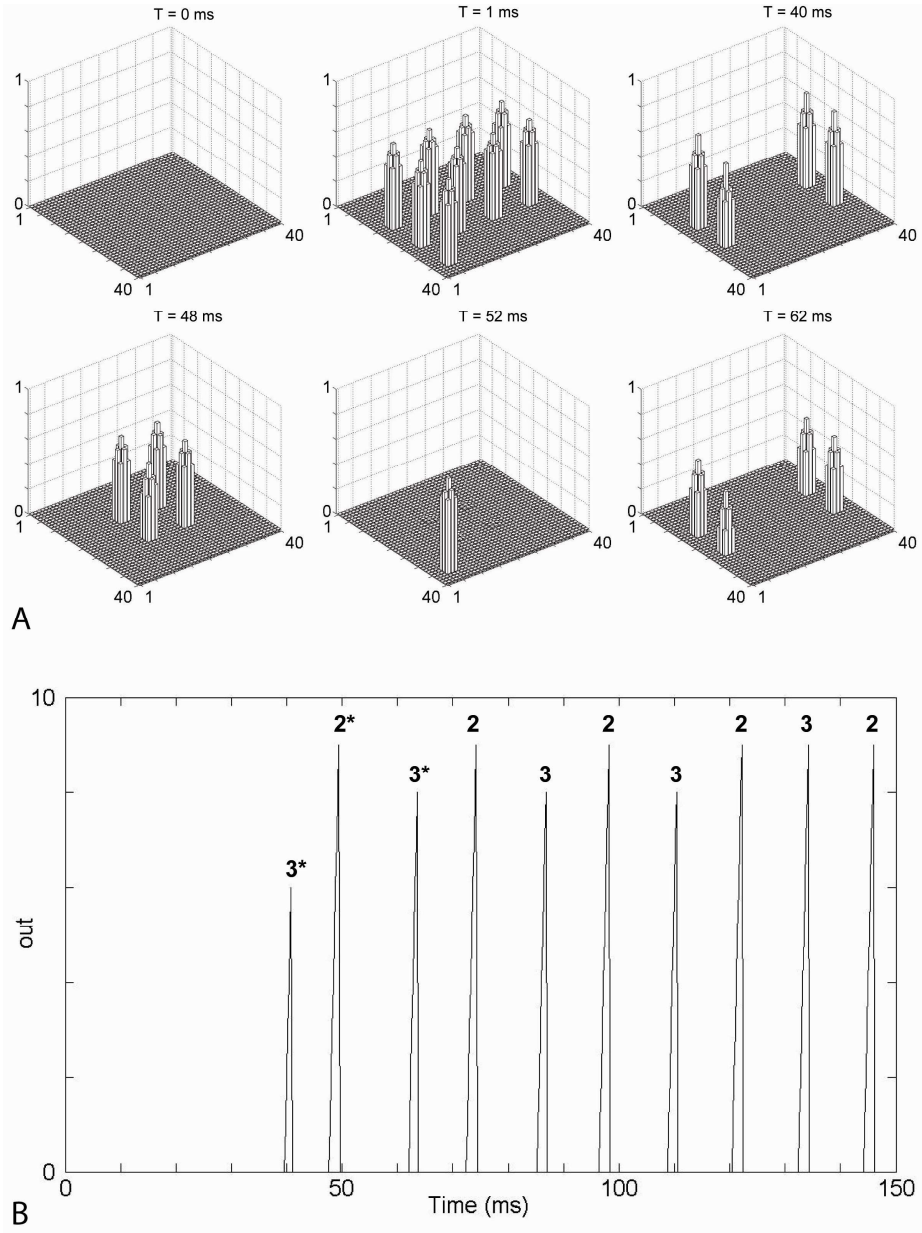


Figure 7 – *Panel A*: Network activity at different snapshots during a simulation with incomplete objects. Each pixel represents an oscillator. The emerging height is proportional to the corresponding oscillator's activity x_{ij}. The represented simulation refers to trial 8 in Table 2: object 1 receives only one input property, while objects 2 and 3 receive all properties. The external input for the stimulated properties is equal to 0.8. The network is able to reconstruct objects 2 and 3, whereas object 1 is not reconstructed (only the bubble corresponding to the stimulated property emerges). *Panel B*: Output of the decision network. During the initial transient (when no object reconstruction occurs), the output of the decision network remains at zero, then it reaches a value equal or grater than 6 during the correct reconstruction of object 2 and 3, whereas it remains at zero during the missed reconstruction of object 1. The number at the top indicates which object is recognized; the asterisks correspond to the snapshots of panel A where object recognition occurs.

In order to clarify the previous points, Figure 7 displays some snapshots of the network behavior at particular instants of the simulations (upper panels), and the output of the downstream

decision network (bottom panel), in an exemplary case: trial 8 in Table 2, when four attributes were given to the network for Objects 2 and 3, and the first object received just one property. It is worth noting that, in the snapshots, each attribute is represented by the excitation of an "activation bubble", characterized by the central excited neuron surrounded by a group of proximal active ones. In Figure 7 we can see that the network, after an initial transient period, is able to correctly reconstruct and segment Objects 2 and 3, but it is not able to reconstruct Object 1 (in fact, a single attribute is not enough to reconstruct this object). During the oscillatory phase of Object 1, only a single "activation bubble" is excited. Output of the downstream decision network reaches a high level (as high as 6) during the correct reconstruction of Objects 2 and 3, but it remains at zero during the oscillatory phase of Object 1.

				NOISE
TRIALS	INPUTS			γ = 0.3 - 0.5
	Obj1	Obj2	Obj3	
1	4	4	4	100%
2	4	4	3	60%
3	4	3	4	60%
4	3	4	4	80%
5	4	4	2	60%
6	4	2	4	100%
7	2	4	4	100%
8	1	4	4	100%
9	4	3	3	80%
10	3	4	3	100%
11	3	3	4	100%
12	4	2	2	60%
13	2	4	2	20%
14	2	2	4	100%
15	3	3	3	100%
16	2	3	3	100%
17	1	3	3	100%

Table 3 – The same simulations as in Table 2 (in case of strong memorization) were repeated starting from initial random conditions and adding Gaussian white noise to all neurons (zero mean value, standard deviation 0.02) both during the training phase and during the recovery phase. Each simulation was performed tenfold for each value of γ and the best performance is given. The third column reports the percentage of success obtained by varying γ between 0.3 and 0.5.obtained by varying γ between 0.3 and 0.5.

Simulations with Gaussian noise

The previous simulations, in case of strong memorization, were repeated starting from random initial conditions and adding Gaussian white noise (zero mean value, standard deviation 0.02) to all neurons both during the training section and during the subsequent simulations. We repeated every simulation tenfold to perform a statistical analysis, and all simulations were repeated for different values of γ.

Results are reported in Table 3, and can be summarized as follows: i) the presence of noise makes objects recognition more difficult. In some cases, the percentage of success decreases to 60% and the value of γ must be decreased to 0.3; ii) It is more difficult to find a clear relationship

between the value of γ which warrants the better performance and the complexity of the task.

Figure 8 displays some snapshots of the network behavior at particular instants of the simulations (upper panels), and the output of the downstream decision network (bottom panel), in an exemplary case: trial 1 in Table 3, when four attributes were given to the network for every object.

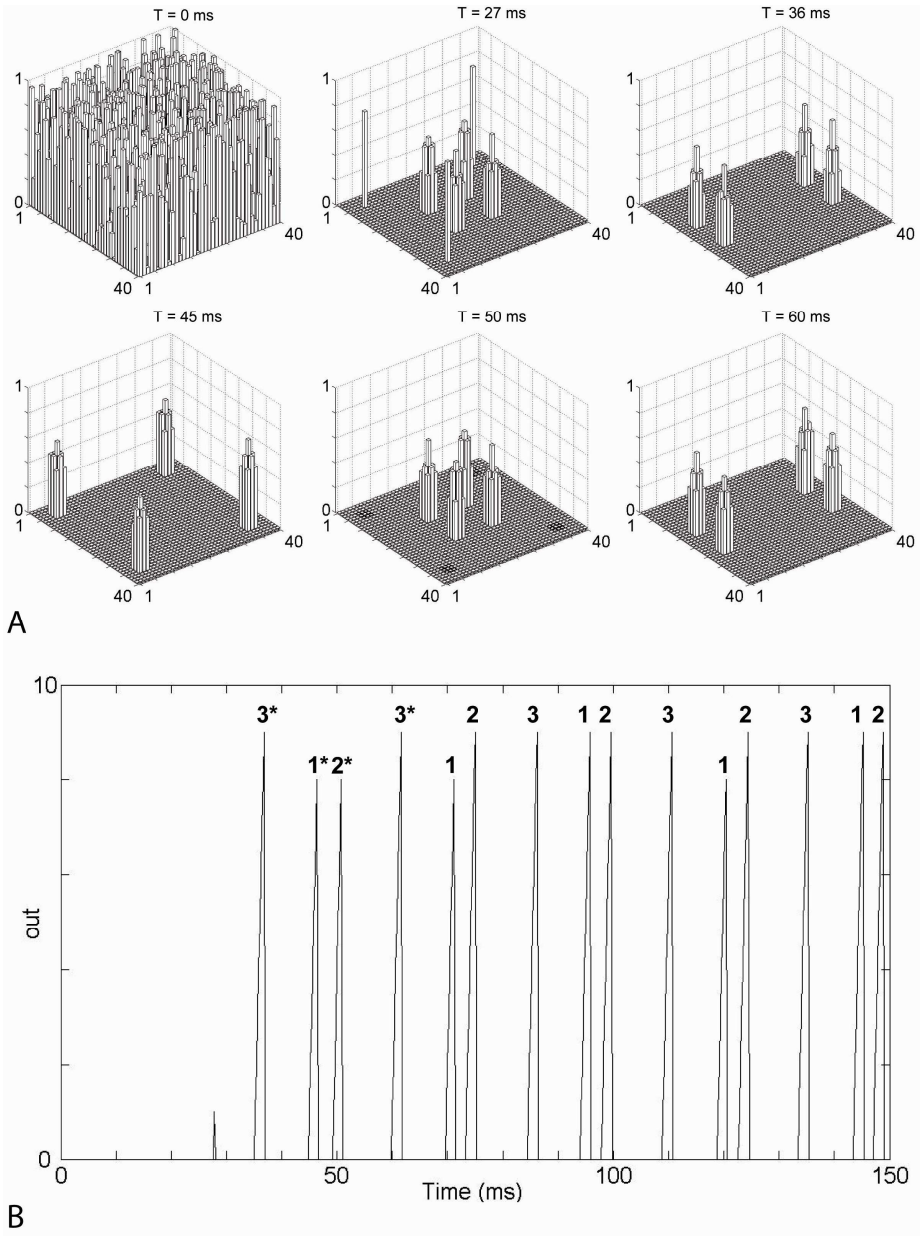


Figure 8 – *Panel A*: Network activity at different snapshots during a simulation with Gaussian noise. The represented simulation refers to trial 1 in Table 3: each object receives all four properties (attributes are stimulated with input 0.8). The network starts from random initial conditions and after a short transient it recognizes all the three objects (upper right panel, obj3; lower panels: obj1, obj2 and obj3 from left to right). *Panel B*: Output of the decision network. The output of the decision network reaches a value higher than 7 whenever an object is correctly recovered. The asterisks correspond to the four snapshots of panel A where object recognition occurs.

In Figure 8 the network starts from a random condition and after an initial transient period is able to correctly reconstruct and segment all the three objects. This is clearly evident in the output of the decision network, which always reaches a level higher than 7 in correspondence of a correct object reconstruction.

Finally, we repeated all simulations by using noise only during the reconstruction phase, but not during training phase. In these conditions, the network exhibits a percentage of success comparable to that of Table 2. Hence, we can conclude that noise has a detrimental effect only if applied during the learning period.

Simulations with modified objects

Table 4 shows results of 30 different simulations (strong memorization) performed both with some lacking features, and some features changed compared with the "exact" value (distance $d = 1$).

				STRONG d = 1			
TRIALS	INPUTS						γ
	Obj1	Obj2	Obj3	Obj1	Obj2	Obj3	
1	3+1	4	4	yes	yes	yes	0.3, 0.5
2	3+1	3	3	yes	yes	yes	0.3, 0.4, 0.5
3	2+1	4	4	yes	yes	yes	0.3, 0.5
4	2+1	3	3	yes	yes	yes	0.3, 0.4, 0.5
5	3+1	3+1	4	yes	yes	yes	0.3
6	3+1	3+1	3	yes	yes	yes	0.3
7	2+1	3+1	4	yes	yes	yes	0.3
8	2+1	3+1	3	yes	yes	yes	0.3, 0.5
9	2+1	2+1	4	yes	yes	yes	0.3, 0.4, 0.5
10	2+1	2+1	3	yes	yes	yes	0.5
11	3+1	4	3+1	yes	yes	yes	0.3, 0.4, 0.5
12	3+1	3	3+1	yes	yes	yes	0.3, 0.4, 0.5
13	2+1	4	2+1	yes	yes	yes	0.5
14	2+1	3	2+1	yes	yes	yes	0.3, 0.4, 0.5
15	3+1	3+1	3+1	yes	yes	yes	0.3, 0.4 T>180ms
16	2+1	2+1	2+1	yes	yes	yes	0.3, 0.5
17	1+1	4	4	no	yes	yes	no
18	1+1	3	4	no	yes	yes	no
19	1+1	3	2	no	yes	yes	no
20	1+1	3	3	yes	yes	yes	0.3, 0.4, 0.5
21	1+1	2	2	yes	yes	yes	0.3
22	1+1	2	3	yes	yes	yes	0.3
23	1+1	2	4	yes	yes	yes	0.3
24	1+1	1+1	3	no	no	yes	no
25	1+1	1+1	2	no	no	no	no
26	1+2	3	3	yes	yes	yes	0.3
27	1+2	4	4	yes	yes	yes	0.3
28	1+2	1+2	4	yes	yes	yes	0.3
29	1+2	1+2	3	yes	yes	yes	0.3
30	1+2	1+2	1+2	yes	yes	yes	0.3

Table 4 – Results of 30 different simulations performed using three objects as input to the network, but with some lacking features and some features shifted from the "exact value" (distance *d* = 1). All three objects previously received the same strong memorization as in Table 2. The number of exact features + the number of "not exact" features for each object is given in the second column. The meaning of all other symbols is the same as in Tables 2. In some trials (17-19, 24, 25), objects cannot be reconstructed. Errors are marked in grey in the third column.

Results can be summarized as follows: i) in most cases the three objects can be correctly segmented and reconstructed, also in presence of lacking and modified attributes. However, in some cases (for instance in all difficult conditions at the bottom of the table) the correct segmentation requires a value of γas low as 0.3. ii) When one object in the network is stimulated with one "exact" + one modified feature, its reconstruction succeeds only in some cases (see trial 20-23 in Table 4). In the other cases (see trials 17-19) the object cannot be reconstructed, whereas the remaining two objects can be segmented. iii) In trials 24-25 there are two simultaneous objects with one exact and one modified feature. In these conditions, neither of them can be reconstructed. iv) The downstream decision network provides a high output only when four

attributes of the object are simultaneously recollected (i.e., the corresponding neural groups oscillate in synchrony). However, in case of modified objects, this output is generally as low as 3-4. By contrast, in the reconstruction of objects without modifications, output of the decision network is always higher than 6. Hence, this output not only signals object reconstruction, but can be considered an index of object reliability, compared with previous experience.

Finally, we repeated these trials also increasing the distance in the modified objects (i.e. using $d = \sqrt{2}$); in this case the network has the capability to reconstruct objects only in simple cases (i.e., all the three objects needed that at least three of their "exact" attributes are given to the network).

Simulations with different number of objects

	4 objects				
TRIALS	INPUTS				γ
	Obj1	Obj2	Obj3	Obj4	
1	4	4	4	4	0.2
2	3	4	4	4	0.2
3	4	3	4	4	0.2
4	4	4	3	4	0.2
5	4	4	4	3	0.2
6	3	3	4	4	0.2
7	3	4	3	4	0.3
8	3	4	4	3	0.2
9	4	3	3	4	0.2
10	4	3	4	3	0.3
11	4	4	3	3	0.2
12	3	3	3	3	0.2
13	2	3	3	3	0.2
14	3	2	3	3	0.2
15	3	3	2	3	0.2

Table 5 – Results of 15 different simulations performed using four objects as input to the network, but with some lacking features. The first three objects received the same strong memorization as in Table 2 and they can be reconstructed starting from two input features; Obj4 was trained for 900 ms and requires three attributes to be reconstructed. The number of stimulated features for each object is given in the second column. In most trails, a correct segmentation requires a lower value of γ (γ = 0.2, see last column).

The network performance was tested by changing the number of objects simultaneously given as input.

First, we performed several trials (not shown for briefness) using only one or two objects as input, with either strong or weak memorization. In these simple cases segmentation can always be achieved with $\gamma = 0.5$. The cases with three objects were illustrated above (Tables 2-4). Then, the network was trained with one more object (named Obj4), represented by the following properties: Obj4 = [5,15 5,25 35,25 35,15] and trained for 900 ms. As in the previous exempla, the first three objects have strong memorization and can be reconstructed starting from only two attributes, whereas the last object has weaker memorization and requires at least three attributes to be correctly reconstructed.

The results, summarized in Table 5, show that the four objects can be correctly segmented and reconstructed, provided at least two attributes are given to the network for the first three objects, and at least three for the weak object. However, recognition requires a lower value of γ(i.e., γ= 0.2, that means a frequency as low as 30 Hz). In fact, only by oscillating at lower frequency, the network exhibits enough time to permit the appearance of an additional object after the emergence of the previous three. The same trials were repeated with the fourth object stored as strong as the others: also in this cases the network segments correctly with γ = 0.2.

Finally, we performed some trials by storing more than 4 objects simultaneously. Results show that the network can reconstruct and segment a maximum of four objects together, independently of the strength of synapses and number of properties stimulated for each object. If the value of γ is lowered to 0.1 (to further decrease frequency) the Wilson-Cowan equations stop oscillating.

	2 Objects with a common characteristic			
TRIALS	INPUTS			γ
	Obj1	Obj2	Obj3	
1	4	3+1c	3+1c	0.4
2	4	3+1c	2+1c	0.4
3	3	3+1c	3+1c	0.4
4	2	3+1c	3+1c	0.3*
5	1	3+1c	3+1c	0.4*
6	4	3	3	0.4
7	3	3+1c	2+1c	0.4
8	3	2+1c	3+1c	0.4
9	3	2+1c	2+1c	0.4
10	3	3	3	0.4
11	2	2+1c	2+1c	0.4*

Table 6 – Results of 11 different simulations performed using three objects as input to the network, with Obj2 and Obj3 sharing one common feature. All three objects received a strong memorization (learning time: Obj1 = 1600 ms, Obj2 and Obj3 = 1200ms). The number of features + the presence of the common feature (symbol 'c') for each object is given in the second column.. The best performance of the network is achieved using γ = 0.4; only in one case, object reconstruction is improved using γ = 0.3 (see last column). Asterisks marking trials 4, 5, and 11 indicate that Obj1 could not be reconstructed, but the other two objects were correctly segmented.

Simulations when two objects share one property

The network was trained with three objects (named Obj1, Obj2 and Obj3), represented by the following properties: Obj1 = [5,5 5,35 35,35 35,5], Obj2 = [15,15 15,25 25,25 25,15], Obj3 = [15,15 15,35 25,35 25,5]. Objects 2 and 3 share the first property. The three objects were separately presented for 1600 ms (Obj1), 1200 ms (Obj2) and 1200 ms (Obj3) (strong memorization).

The results of different simulations are summarized in Table 6. We found that the best performance of the network is achieved using γ = 0.4; with this value the oscillation frequency is about 45 Hz. However, in one case, we found that

some object reconstructions requires $\gamma = 0.3$ (i.e., the frequency is reduced to 40 Hz). It is worth noting that, in these conditions, Objects require at least three properties to be correctly reconstructed (i.e., they cannot be reconstruct starting from two properties only).

Figure 9 displays some snapshots of the network behavior at particular instants of the simulations in an exemplary case (trial 1 in Table 6, when four attributes were given to the network for each object). After an initial transient period, the network is able to correctly reconstruct and segment all the three objects.

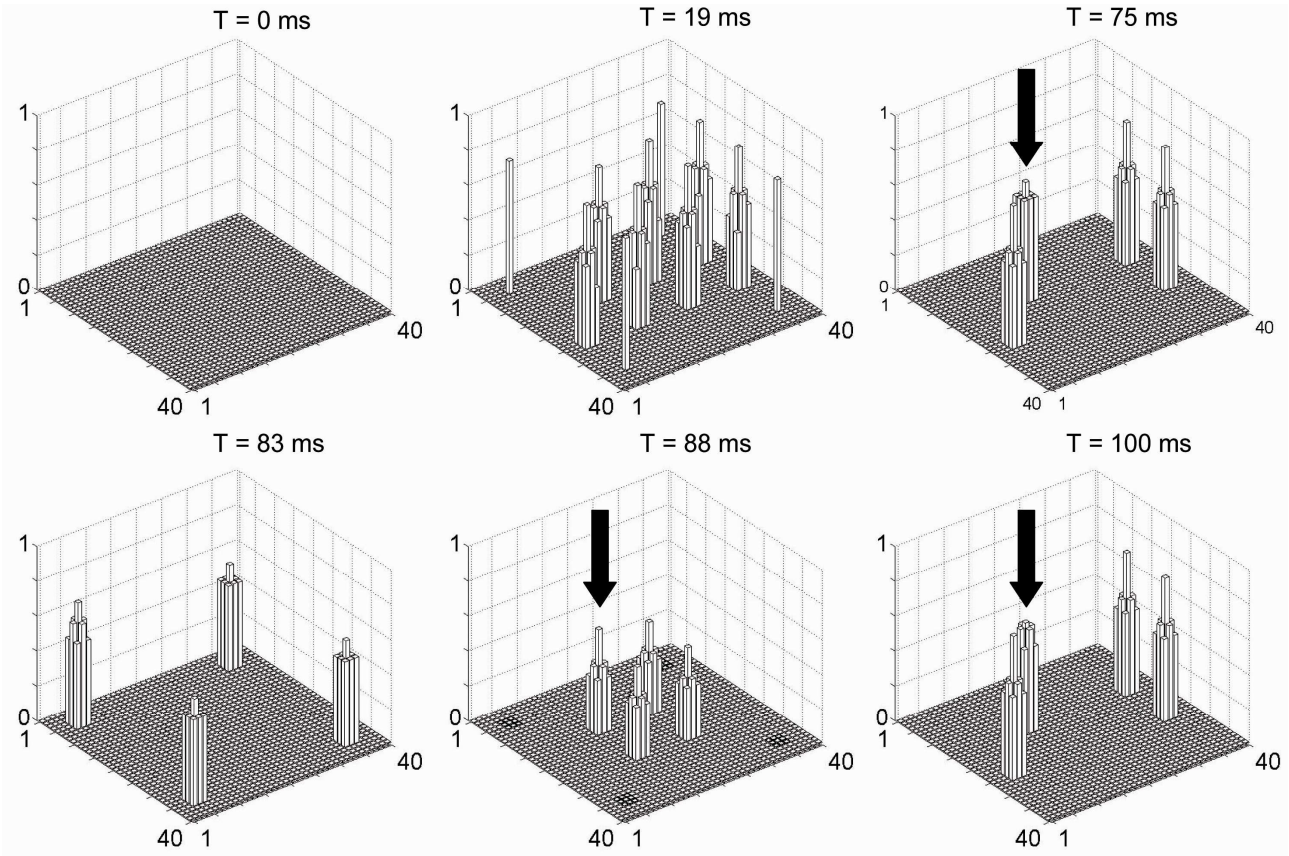


Figure 9 – Network activity at different snapshots during the simulation with two objects sharing an attribute. The represented simulation refers to trial 1 in Table 6: all the three objects receive four inputs property. The external input for the stimulated properties is equal to 0.8. After an initial transient, the network is able to reconstruct the objects; in particular objects 2 and 3, which share the property marked in figure with a black arrow, are well segregated.

Finally, we performed some simulations with objects sharing two features. In these conditions, the network fails to segment these objects.

Effect of the distance among features

Several of the previous simulations were repeated varying the distance between the attributes of two different objects. Results show that the network reaches a correct reconstruction, independently of the distance, as long as two "activation bubbles" belonging to the same area (i.e., bubbles of different objects) are not overlapping.

Simulations with variable γ

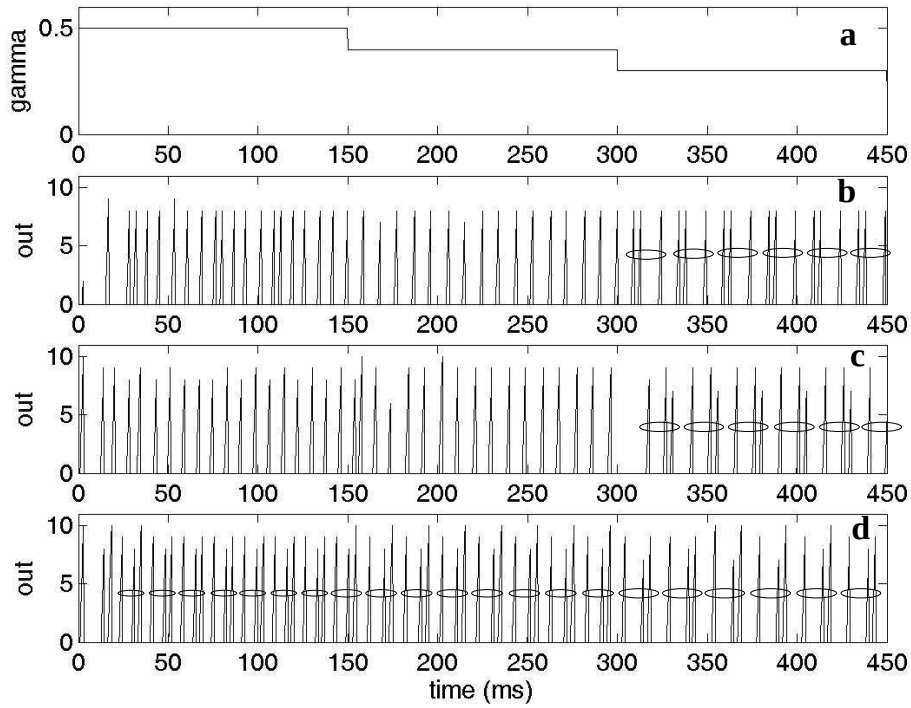


Figure 10 - Three examples of the output of the detection network, obtained by decreasing the frequency of oscillations. To this end, the value of parameter γ (panel a) has been decreased from 0.5 to 0.3 in three consecutive steps, with a 150 ms duration each. Panel b represents output of the detection network during a simulation of the trial 15 of Table 2 with strong memorisation. Panel c represents the same output during the simulation of trial 6 of Table 2, with weak memorisation. Panel d represents results obtained during the trial 6 of Table 2, with strong memorisation. The cases with correct segmentation of three objects are marked with ellipses. It is worth noting that segmentation in panels b and c is achieved only after 300 ms, when the value of γ has been decreased down to 0.3. In the bottom panel, segmentation is achieved almost immediately and is maintained even at lower values of γ.

The previous results show that the oscillation frequency (hence, the value of parameter γ) must be changed depending on the difficulty of the task; while simple tasks (like those with strong memorization in Table 2) can be solved with $\gamma = 0.5$, more difficult tasks (as those with weak

memorization, with four simultaneous objects, with lacking features, or superimposed features) often require a lower frequency. However, the use of a low γ may be detrimental in case of simple tasks, since some properties of one object may exploit the longer oscillation time available to appear separately from the others.

In order to make the system more robust, we implemented a simple method to manage γ automatically. It consists in starting the simulation with $\gamma = 0.5$, and then reducing its value in steps ($\Delta\gamma$ = -0.1 every 150 ms). Examples are shown in Figure 10.

This figure depicts the output of the decision network in three cases (trial 15 of Table 2, strong; trial 6 of Table 2, weak; and trial 6 of Table 2, strong). In the first two cases, the network fails to detect all three objects during the first 300 ms, but a correct segmentation is achieved after 300 ms, when γ is reduced to 0.3. In the last case, synchronization is almost immediately achieved with $\gamma = 0.5$, and is maintained throughout the simulation.

DISCUSSION

In this work, we present a mathematical model which exploits γ-band synchronization to simulate abstract representation of objects in associative-declarative memory. The term “abstract” representation is used here to denote an object represented as a collection of features, which may spread across different sensory modalities and can be independent of spatial relations. For this reason, we used similarity and previous knowledge rules to segment objects, without exploiting any spatial property. Hence, our model does not intend to represent image processing in the visual cortex, but rather processing in higher cortical associative areas. This aspect clearly distinguishes our model from other models (Kazanovich and Borisyuk, 2002; Ursino et al., 2003; Ursino and La Cara, 2004; Wang and Terman, 1997; Terman and Wang, 1995), which are explicitly devoted to image segmentation and in which different objects are separated on the basis of spatial properties

(such as proximity, smoothness, common fate). As summarized by Wang in a recent review paper (Wang, 2005), these models were especially focused on the figure-ground separation problem

Of course, the present model requires a pre-processing stage, which extracts features according to a topological organization. These features may vary depending on the kind of objects to deal with: they may include geometrical features (like shapes), colors, tones, flavours, etc. , each organized in a topological way. Analysis of these pre-processing algorithms is beyond the aim of the present work.

However, the ubiquitous role of γ-band activity suggests that it may also be involved in the formation of more "abstract" object representation in memory, in part independent on spatial position. In this case, synchronization may be driven by higher association rules, such as previous knowledge and similarity.

In the present work we assume that an object can be recovered and represented in memory even if some features are lacking in the external stimuli, or are moderately changed compared with a previous knowledge. Let us consider, for instance, the idea of a sea-shore (sun, flavor of fishes, waves and sand) or the idea of a flame (lightness, heating, sounds and smoke). In case of incomplete information, we assume that just a couple of these attributes is sufficient to retrieve the remaining properties in memory (for instance, the perception of smoke and lightness in a night can evoke the idea of a flame, with the expectation of burning wood noise and heating). Furthermore, these attributes can be a little different from those previously perceived (for instance, the lightness can have a different intensity, smoke can have a different smell). Similarly, one can recognize that a tree is a tree although the characteristics of the leaves and bark are different from one tree to another, and you can imagine lacking features of the tree (such the smell of flowers).

Although some previous models (summarized in the introduction) exhibit some common aspects with the present, there are also profound new aspects in our approach. In particular, none of these models combine oscillatory dynamics (synchronization and desynchronization), a topological organization of features, Hebbian learning and a final decision network into a single coherent

structure. In most models segmentation and memory are treated as separate problems. Most models lack a topological organization to describe the input. Others use unphysiological learning rules.

These originals aspects of our models are more exhaustively commented below.

i) The more recent models mentioned above (Borisyuk and Kazanovich, 2004; Knoblauch and Palm, 2001; Wang and Liu, 2002) perform segmentation and recognition at two different processing stages. Generally, a first layer of neurons segments a visual image on the basis of proximity and spatial connection laws. Subsequently, the information already segmented among different objects is sent to a feature extraction layer, and/or to an associative memory layer, which recognizes objects and implements prior knowledge rules. In other words, the spatial nature of visual images helps the solution of the problem, since it allows segmentation before object recognition by using different low-level rules (proximity, common fate, smoothness, etc.). Of course, previous knowledge can subsequently improve segmentation in ambiguous cases. By contrast, in the present model, segmentation and matching with previous memory occur at the same processing stage without the use of spatial information, i.e., without the aid of low-level spatial rules.

ii) The different features in our model are organized according to a topological map. We assumed that a previous processing step (not included here, but similar to a classic Kohonen's self-organizing map) extracts the main features of the input, and orders them in a topological way. Moreover, this topological organization of features is assumed as given (i.e., it is not subject to learning). This is reasonable, since the formation of topological maps require a long training period, which is much longer than the period required to store individual objects or facts in declarative memory (Rolls and Treves, 1998; Hertz et al., 1991).

Although there is no definite demonstration in neurophysiology that topographic maps, besides representing sensory information are also engaged in more complex mental operations, including object recognition, several recent papers stress this hypothesis.

Kohonen and Hari (1999) demonstrated that abstract feature maps arise spontaneously from the use of self-organizing algorithms. In a brilliant example reported in their paper (see also Ritter and

Kohonen, 1989) a semantic map of words is formed, in which words are automatically segregated into classes, and a finer structure can be found within classes.

Many other authors in more recent years discussed the possible role of topographic maps outside sensory areas. Pulvermuller (2005) hypothesizes that topographically organized regions of the cortex directly connect different features of a word. In this idea, a semantic process would engage many cortical areas. Simmons and Barsalou (2003) proposed a similarity-in-topography principle; this principle states that "the categorical structure of the word becomes instantiated in the topography of the brain's association areas". The role of topography in the organization of higher brain centers, and particularly in abstract cognitive representation, is stressed by Thivierge and Marcus in a recent review paper (2007).

Direct experimental demonstrations of these theories are still scanty. Tanaka (2003), by exploring cortical columns in the inferotemporal cortex, suggests that some form of topography is present in that region: most cells represent features of object images while cells within the same column respond to similar features. Maps which do not represent spatially ordered information can be found in the auditory cortex and in the olphactory system (Kaiser et al., 2002; Freeman, 1978). Recent neuroimaging studies suggest that maps can be found in cortical regions far downstream the primary sensory areas (such as in frontal and prefrontal regions) (Hagler and Sereno, 2006). Indirect evidence on the role of topological maps in cognition also comes from studies on patients with deficits in category recognition (see Damasio (1990) and Thivierge and Marcus (2007) for a review).

iii) Association among different features (i.e., linking features into a coherent representation of abstract objects) is realized using a Hebbian rule, which allows implementation of previous knowledge principle. By contrast, prior knowledge in some other networks (Borisyuk and Kazanovich, 2004; Wang and Liu, 2002) is implemented using more complex rules, which do not have a clear neurophysiological counterpart.

In our model the Hebb rule was implemented using the correlated moving average of the activity of presynaptic neurons, computed during the previous 10 ms, with the current activity of the post-synaptic neuron. Indeed, the required delay time for effective synaptic Hebbian modification is the order of less than ±10 ms (Abbott and Nelson, 2000). A further characteristic of our model is that the strength of object memorization can be easily varied, by increasing the training time and/or the learning rate. The consequence is that multiple objects can be stored in memory with a different pregnancy. In particular, the difference in the memorization strength in our model produces a difference in the balance between sensitivity and specificity for object recognition (sensitivity denotes the capacity to recognize true positives, specificity the capacity to reject true negatives). In case of strong memorization, a smaller number of features may be required to recover the object from partial information (see, for instance, objects 1 and 2 in Tables 2 and 3, which can be reconstructed starting from just two attributes over 4). Hence, the object can be recovered with very high sensitivity, but with a poor specificity. The opposite occurs in case of weak memorizations. With the present network it is easy to realize the best compromise between sensitivity and specificity, by acting on the learning rate β_0 in Eq. 11, and on the maximum level of synapses W_{max}. This compromise between sensitivity and specificity can be important in real life. For instance, if recognition of a given object is essential for survival, it is preferable to recover it even in a wrong condition (poor specificity), but avoiding any fault of recognition (high sensitivity). By contrast, other objects may be recognized only in evident cases, tolerating some omissions.

iv) The present model makes use of a “decision network”, placed downstream the topological maps of oscillating neurons. The function of this network is to reach a decision on whether, at a given instant, an acceptable representation of an object is present in the cortex. The output of this network can be sent out to other regions in the brain, to drive action and behavior. It is worth noting that our decision network makes use of neurons which detect the simultaneous presence of activity in different areas. The existence of coincidence-sensitive target neurons has been hypothesized by others in the context of assembly coding (Abeles, 1982; König et al., 1996). Experimental evidence

has been presented for neurons in the visual and somatosensory cortex which are triggered by synchronized thalamic afferents (Alonso et al., 1996; Roy and Allowey, 2001).

In all simulations, we tested the behavior of the decision network, and found that it performs well. Some errors may just occur during the initial transient period (the first 30-40 ms) which thus must be excluded from the analysis. The network provides a strong output only when four "activation bubbles" with correct dimension are present in the four areas (one per area). In all other cases the network does not recognize the object. Two aspects deserve comments. The output of the decision network may reach different levels: hence, it does not only recognize the existence of an object, but may also provide some information on the degree of reliability of the object itself. Second, in case of a modification of one property, the final object presents an activation bubble which is slightly translated compared with the original one (i.e., that used during the learning phase). This is an interesting property of our network: inter-area synapses allow restoration of lacking properties on the basis of previous knowledge, still preserving information on subtle differences between the actual object and the one originally memorized. It is interesting to observe that, in most of these cases, the decision network reaches lower levels, i.e., it gives a smaller confidence to object recognition.

We are aware that the decision layer has not a clear physiological evidence. However, we think that its function is important to provide a measure of the reliability of object recognition, i.e., a measure of how much the present information in the four topological areas can be exploited or not in the prosecution of the task. As pointed out by others (Osipova, 2006), the increase in neural synchronization may result in a drive to other areas of the brain participating to the same task. However, activities characterized by a failure in object detection, should not drive other areas. This requires the existence of a gating mechanism, which communicates information to other centers in case of correct recognition, warranting that poor information is not taken into account. In our model, this function is performed by the decision layer. The latter should not be considered as a "homunculus", but simply as a further step in a complex multilayer processing path. Furthermore,

output of the decision layer might be used as a feedback reinforcement signal, to improve learning of recognized objects, but inhibiting poorly recognized objects. A similar approach to classify and learn objects is used in the Adaptive Resonance theory (ART) (Carpenter and Grossberg, 2003). Finally, as discussed in the simulations of modified objects (Table 4), output of the decision network may be used to quantify the level of reliability of the recognition.

Furthermore, it is interesting to observe that the presence of a downstream decision network makes our model somewhat similar to the "match and utilization" qualitative model proposed by Herrmann et al (2004). In that model, a first processing step reflects a match between stimulus-related information and memory content. The result of this comparison is then read out by a second processing step, which is referred as "the utilization". In our model, the first step (matching) is performed by the network of Wilson-Cowan oscillators, in which inter-area synapses represent memory, and the network dynamics attempts to restore stored information. The second step (utilization) is realized by the downstream decision network.

An important aspect, emerging from our simulations, is the necessity to modify the frequency of oscillations (hence the parameter γ) depending on the complexity of the task. While simpler tasks can generally be solved with a value of γ as high as 0.5, more difficult tasks require a lower γ (down to 0.2 in case of four simultaneous objects). The dependence of the optimal oscillation frequency on the complexity of the task can be summarized as follows: If oscillation frequency is too high, the network has not enough time to process information, and so the objects have not enough time to appear separately in time division (for instance, in the case of four objects). However, if the oscillation frequency is too small, some properties may benefit of the dead time between one object and another, and appear as isolated properties instead of synchronizing with the other properties of the same object (for instance, this may be a consequence of noise in the learning phase).

In order to deal with a variable oscillation frequency, in this work we proposed a simple strategy, which consists in decreasing γ in progressive steps. Each step should last enough to allow correct synchronization (approximately 100-150 ms). Although this strategy ensures good

segmentation for different tasks which require a different frequency (as shown in Figure 10), it is feasible that the brain implement a more sophisticate strategy to choice the best oscillation frequency. Maybe, the decision network output may be used to control the oscillation frequency in future versions of this model.

Several recent studies provide a physiological basis for a frequency change in γoscillations. The characteristics of γ rhythms are under the control of attention (Jensen et al., 2007; Börgers et al., 2005), and can be modulated by D4 dopamine receptor, and by cholinergic (acetylcholine and muscarinic) receptors (Rodriguez et al., 2004; Hentschke et al., 2007; Kuznetsova and Deth, 2008). In particular, D4 dopamine affects the spike duration and the interspike interval (Kuznetsova and Deth, 2008). Herculano-Huzel et al. (1999) recently observed that the frequency of gamma band oscillations depends on the level of central activation: at intermediate level of activation the frequency is in the range 70-105 Hz; at higher level of activation it is in the range 20-65 Hz, and the frequency decreases with further central activation, with an increase in the strength of response. Neuromodulation of γoscillations has been also investigated by Hasselmo and coworkers (2006).

The presence of a global inhibitor also requires a brief comment. Its presence is necessary to desynchronize oscillators in different objects, as pointed out by previous modeling papers (Ursino et al., 2003; Wang and Terman, 1997). However, it is at present difficult to find a clear biological counterpart for it. Its function may be related with that of attention. Indeed, some authors have shown that attention in visual cortical areas has an inhibitory influence and can induce a bias toward one of two competing stimuli (Treue and Maunsell, 1996). We are not aware of similar data in higher cortical areas.

Finally, it is worthwhile to discuss some possible limitations of the present work and delineate lines for future improvement.

Actually, the simulations presented in Table 3 show that model results are very sensitive to noise and that model's performance worsens if noise is added to the inputs. This result seems to contradict results of our previous study (see (Ursino et al., 2003)) and of studies by others (such as

(Horn and Opher, 1996)) showing that noise may have a beneficial effect on binding and segmentation. However, this contradiction is only apparent. In fact, in our trials noise has a detrimental effect solely if applied during the learning phase; if noise is applied merely during the segmentation phase (which is the condition tested in previous studies), the network performs very well. In order to avoid the detrimental effect of noise during training, it might be necessary to modify the learning rule for instance with the introduction of a decay term.

In the present trials, with strong memorization, two objects can be correctly segmented if they have just one common feature, but they cannot be segmented if they have two common features; hence, it seems that the network has very limited capacity to handle superimposed objects. However, this result was well expected, and depends on the limited number of features we used to represent objects. In particular, when one object is stored with strong inter-area synapses, it can be entirely recovered starting from two features. As a consequence, if two objects have two common features and strong memorization, presentation of two features of the first object activates its overall representation, and so unavoidably activates also the remaining two features of the second object. As a general rule, the maximum number of features that two objects can have in common, without showing interference, must be less than the number of features sufficient to reconstruct the entire object from incomplete information. Of course, representation of objects with a greater number of features may allow a larger superimposition to be handled with the same theoretical approach

The learning rule adopted in the present trial also deserves a few comments. First, in the present work we used a different learning time to memorize objects with a different strength, and so to break any initial symmetry in the network behaviour. Desynchronization, however, could also be obtained using a different learning rate for distinct objects (i.e., a different value of β_0), or starting with random initial conditions. Second, as specified above, the present learning rule is affected by the presence of random noise superimposed during the training period. This model limitation might be overcome in future works using a learning rule with a decay term, which converges to the average value of the input, by eliminating any noise with zero mean value. Finally, some authors in

recent years (Hasselmo, 2005; Manns et al., 2007) proposed that the theta rhythm may allow a rapid alternation in autoassociative memories (like the hippocampus) between conditions that promote memory encoding and conditions that promote memory retrieval. This aspect may also be the subject of a future model improvement.

REFERENCES

Tallon-Baudry C and Bertrand O. (1999) Oscillatory gamma activity in humans and its role in object representation. Trends Cogn.Sci. 3: 151-162.

Damasio AR. (1989) The brain binds entities and events by multiregional activation from convergence zones. Neural Computation. 1: 123-132.

Singer W and Gray CM. (1995) Visual Feature integration and the temporal correlation hypothesis. Ann.Rev.Neurosci. 18: 555-586.

Singer W. (1999) Neuronal synchrony: a versatile code for the definition of relations? Neuron. 24: 49-65.

Varela F, Lachaux JP, Rodriguez E, Martinerie J. (2001) The brainweb: phase synchronization and large-scale integration. Nat.Rev.Neurosci. 2: 229-239.

von der Malsburg C and Schneider W. (1986) A neural cocktail-party processor. Biol.Cybern. 54: 29-40.

Lebedev MA and Nelson RJ. (1995) Rhythmically firing 20-50 Hz neurons in monkey primary somatosensory cortex: activity patterns during initiation of vibratory-cued hand movements. J.Comput.Neurosci. 2: 313-334.

Brosch M, Budinger E, Scheich H. (2002) Stimulus-related gamma oscillations in primate auditory cortex. J.Neurophysiol. 87. 2715-2725.

Kaiser J, Lutzenberger W, Ackermann H, Birbaumer N. (2002) Dynamics of gamma-band activity induced by auditory pattern changes in humans. Cereb.Cortex. 12: 212-221.

Freeman WJ. (1978) Spatial properties of an EEG event in the olfactory bulb and cortex. Electroencephalogr.Clin.Neurophysiol. 44: 586-605.

Wehr M and Laurent G. (1996) Odour encoding by temporal sequences of firing in oscillating neural assemblies. Nature. 384: 162-166.

Rodriguez E, George N, Lachaux JP, Martinerie J, Renault B, Varela FJ. (1999) Perception's shadow: long-distance synchronization of human brain activity. Nature. 397: 430-433.

Bhattacharya J, Petsche H, Pereda E. (2001) Long-range synchrony in the gamma band: role in music perception. J.Neurosci. 21: 6329-6337.

Pulvermüller F, Preissl H, Lutzenberger W, Birbaumer N. (1996) Brain rhythms of language: nouns versus verbs. Eur.J.Neurosci. 8: 937-941.

Tallon-Baudry C, Bertrand O, Delpuech C, Pernier J. (1997) Oscillatory gamma band (30-70 Hz) activity induced by a visual search task in humans. J.Neurosci. 17: 722-734.

Tallon-Baudry C, Bertrand O, Peronnet F,Pernier J (1998) Induced gamma band activity during the dealy of a visual short-term memory task in humans. J.Neurosci. 18: 4244-4254.

Osipova D, Takashima A, Oostenvald R, Fernández G, Maris E, Jensen O. (2006) Theta and gamma oscillations predict encoding and retrieval of declarative memory. J.Neurosci. 26: 7523-7531.

Salinas E and Sejnowski TJ. (2001) Correlated neuronal activity and the flow of neural information. Nat.Rev.Neurosci. 2: 539-550.

Melloni L, Molina C, Pena M, Torres D, Singer W, Rodriguez E. (2007) Synchronization of neural activity accross cortical areas correlates with conscious perception. J.Neurosci. 27: 2858-2865.

Rolls ET and Treves A. (1998) Neural Networks and Brain Function Oxford: Oxford University Press.

Kohonen T and Hari R. (1999) Where the abstract feature maps of the brain might come from. Trends Neurosci. 22: 135-139.

Rolls ET and Deco G. (2002) Computation Neuroscience of Vision New York: Oxford University Press.

Markram H, Lübke J, Frotscher M, Sakmann B. (1997) Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSSs. Science. 275: 213-215.

Abbott LF and Nelson SB. (2000) Synaptic plasticity: taming the beast. Nat.Neurosci. 3: 1178-1183.

Zhang LI, Tao HW, Holt CE, Harrsi WA, Poo M. (1998) A critical window for cooperation and competition among developing retinotectal synapses. Nature. 395: 37-44.

Paulsen O and Sejnowski TJ. (2000) Natural patterns of activity and long-term synaptic plasticity. Curr.Opin.Neurobiol. 10: 172-179.

Borisyuk R and Kazanovich Y. (2004) Oscillatory model of attention-guided object selection and novelty detection. Neural Networks. 17: 899-915.

Campbell SR, Wang DL, Jayaprakash C. (1999) Synchrony and desynchrony in integrate-and-fire oscillators. Neural Computation. 11: 1595-1619.

Cesmeli E and Wang D. (2000) Motion Segmentation based on Motion/Brightness integration and Oscillatory Correlation. IEEE transactions on neural networks. 11: 935-947.

Chen K and Wang D. (2002) A dynamically coupled neural oscillator network for image segmentation. Neural Networks. 15: 423-439.

Eckhorn R, Reitboeck HJ, Arndt M, Dicke PW. (1990) Feature-linking via synchronization among distributed assemblies: simulation of results from cat cortex. Neural Computation. 2: 293-307.

Grossberg G and Grunewald A. (1997) Cortical synchronization and perceptual framing. J.Cogn.Neurosci. 9: 117-132.

Grossberg S and Somers D. (1991) Synchronized oscillations during cooperative feature linking in a cortical model of visual perception. Neural Networks. 4: 453-466.

Hendin O, Horn D, Tsodyks MV. (1998) Associative memory and segmentation in an oscillatory neural model of the olfactory bulb. J.Comput.Neurosci. 5: 157-169.

Horn D and Opher I. (1996) The importance of noise for segmentation and binding in dynamical neural systems. Int J Neural Syst. 7: 529-535.

Horn D and Opher I. (1996) Temporal segmentation in a neural dynamic system. Neural Computation. 8: 373-389.

Hoshino O, Kashimori Y, Kambara T. (1998) An olfactory recognition model based on spatio-temporal encoding of odour quality. Biol.Cybern., vol. 79 pp. 109-120, 1998.

Hummel JE and Biederman I. (1992) Dynamic binding in a neural network for shape recognition. Psychol.Rev. 99: 480-517.

Hummel JE. (2001). Complementary solutions to the binding problem in vision: Implications for shape perception and object recognition. Visual Cognition. 8: 489-517.

Kazanovich Y and Borisyuk R. (2002) Object selection by an oscillatory neural network. BioSystems. 67: 103-111.

Knoblauch A and Palm G. (2001) Pattern separation in spiking associative memories and visual areas and synchronisation. Neural Networks. 14: 763-780.

Kuntimad G and Ranganath HS. (1999) Perfect image segmentation using pulse coupled neural networks. IEEE Trans.Neural Networks. 10: 591-598.

Kuzmina M, Manykin E, Surina I. (2004) Oscillatory network with self-organized dynamical connections for synchronization-based image segmentation. BioSystems. 76: 43-53.

Levy N, Horn D, Meilijson I, Ruppin E. (2001) Distributed synchrony in a cell assembly of spiking neurons. Neural Networks. 14: 815-824.

Li Z. (1999) Visual segmentation by contextual influences via intra-cortical interactions in the primary visual cortex. Network. 10: 187-212.

Lourenço C, Babloyantz A, Hougardy M. (2000) Pattern segmentation in a binary/analog world: unsupervised learning versus memory storing. Neural Networks 13: 71-89.

Nakano H and Saito T. (2004) Grouping synchronization in a pulse-coupled network of chaotic spiking oscillators. IEEE Trans.Neural Networks. 15: 1018-1026.

Ranganath HS and Kuntimad G. (1999) Object detection using pulse coupled neural networks. IEEE Trans.Neural Networks. 10: 615-620.

Sompolinsky H, Golomb D, Kleinfeld D. (1990) Global processing of visual stimuli in a neural network of coupled oscillators. Proc.Natl.Acad.Sci.USA. 87: 7200-7204.

Ursino M, La Cara GE, Sarti A. (2003) Binding and segmentation of multiple objects through neural oscillators inhibited by contour information. Biol.Cybern. 89: 56-70.

Ursino M and La Cara GE (2004) Modeling segmentation of a visual scene via neural oscillators: fragmentation, discovery of details and attention. Network: Comput Neural Syst. 15: 69-89.

von der Malsburg C and Buhmann J. (1992) Sensory segmentation with coupled neural oscillators. Biol.Cybern. 67: 233-246.

Wang D, Buhmann J, von der Malsburg C. (1990) Pattern segmentation in associative memory," Neural Computation. 2: 94-106.

Wang D and Terman D. (1997) Image Segmentation based on oscillatory correlation. Neural Computation. 9: 805-836.

Wang D. and Liu X. (2002) Scene analysis by integrating primitive segmentation and associative memory. IEEE Transactions Man and Cybernetics - Part B: Cybernetics. 32: 254-268.

Yazdanbakhsh A and Grossberg S. (2004) Fast synchronization of perceptual grouping in laminar visual cortical circuits. Neural Networks. 17: 707-718.

Zhang X and Minai AA. (2004) Temporally sequenced intelligent block-matching and motion-segmentation using locally coupled networks. IEEE Trans.Neural Networks. 15: 1202-1214.

Zhang H, Xie Y, Wang Z, Zheng C. (2007) Adaptive synchronization between two different chaotic neural networks with time delay. IEEE Trans. Neural Netw. 18: 1841–1845.

Wu W and Chen T. (2008) Global synchronization criteria of linearly coupled neural network systems with time-varying coupling. IEEE Trans. Neural Netw. vol. 19: 319–332.

Rao AR, Cecchi GA, Peck CC, Kozsloski JR. (2008) Unsupervised segmentation with dynamical units. IEEE Trans. Neural Netw. 19: 168–182.

Ursino M, Magosso E, La Cara GE, Cuppini C. (2006) Object segmentation and recovery via neural oscillators implementing the similarity and prior knowledge gestalt rules. BioSystems. 85: 201-218.

Tanaka K. (2003) Columns for complex visual object features in the inferotemporal cortex: clustering of cells with similar but slightly different stimulus selectivity. Cereb.Cortex. 13: 90-99.

Angelucci A and Bressloff PC. (2006) Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1 neurons. Prog.Brain Res. 154: 93-120.

Schwabe L, Obermayer K, Angelucci A, Bressloff PC. (2006) The role of feedback in shaping the extra-classical receptive field of cortical neurons: a recurrent network model. J.Neurosci. vol. 26: 9117-9129.

Terman D and Wang D. (1995) Global competition and local cooperation in a network of oscillators. Physica D. 81: 148-176.

Wang D. (2005) The time dimension for scene analysis. IEEE Tr.Neural Networks. 16: 1401-1426.

Hertz J, Krogh A, Palmer RG. (1991) Introduction to the theory of neural computation. Redwood, CA: Addison-Wesley.

Ritter H and Kohonen T. (1989) Self-organizing semantic maps. Biol.Cybern. 61: 241-254.

Pulvermuller F. (2005) Brain mechanisms linking language and action. Nat.Rev.Neurosci. 6: 576-582.

Simmons WK and Barsalou LW. (2003) The similarity-in-topography principle: reconciling theories of conceptual deficits. Cogn.Neuropsychol. 20: 451-486.

Thivierge J and Marcus GF. (2007) The topographic brain: from neural connectivity to cognition. Trends Neurosci. 30: 251-259.

Hagler DJ and Sereno MI (2006) Spatial maps in frontal and prefrontal cortex. Neuroimage. 29: 567-577.

Damasio AR. (1990) Category-related recognition defects as a clue to the neural substrates of knowledge. Trends Neurosci. 13: 95-98.

Abeles M. (1982) Role of the cortical neuron: integrator or coincidence detector? Isr.J.Med.Sci. vol. 18: 83-92.

König P, Engel AK, Singer W. (1996) Integrator or coincidence detector? The role of the cortical neuron revisited. Trends Neurosci. 19: 130-137.

Alonso JM, Usrey WM, Reid RC. (1996) Precisely correlated firing in cells of the lateral geniculate nucleus. Nature. 383: 815-819.

Roy SA and Allowey KD. (2001) Coincidence detection or temporal integration? What the neurons in somatosensory cortex are doing. J.Neurosci. 21: 2462-2473.

Carpenter, G. A. and Grossberg, G. (2003) Adaptive resonance theory," in Arbib, M. A. (ed.) The Handbook of Brain Theory and Neural Networks Cambridge. MA: MIT Press 87-90.

Herrmann CS, Munch MHJ, Engel AK. (2004) Cognitive functions of gamma band activity:memory match and utilisation. Trends in Cognitive Sciences. 8: 347-355.

Jensen O, Kaiser J, Lachaux JP. (2007) Human gamma frequency oscillations associated with attention and memory. Trends Neurosci. 30: 317-324.

Börgers C, Epstein S, Kopell NJ. (2005) Background gamma rhythmicity and attention in cortical local circuits: a computational study. Proc.Natl.Acad.Sci.USA. 102: 7002-7007.

Rodriguez R, Kallenbach U, Singer W, Munk MH. (2004) Short- and long-term effects of cholinergic modulation of gamma oscillations and response synchronization in the visual cortex. J.Neurosci. 24: 10369-10378.

Hentschke H, Perkins MG, Pearce RA, Banks MI. (2007) Muscarinic blockade weakens interaction of gamma with theta rhythms in mouse hippocampus. Eur.J.Neurosci. 26: 1642-1656.

Kuznetsova AY and Deth RC. (2008) A model for modulation of neuronal synchronization by D4 dopamine receptor-mediated phospholipid methylation. J.Comput.Neurosci. 24: 314-329.

Herculano-Houzel S, Munk MH, Neuenschwander S, Singer W. (1999) Precisely synchronized oscillatory firing patterns require electroencephalographic activation. J.Neurosci. 19: 3992-4010.

Hasselmo ME and Giocomo LM. (2006) Cholinergic modulation of cortical function. J.Mol Neurosci. 30: 133-135.

Treue S and Maunsell JH. (1996) Attentional modulation of visual motion processing in cortical areas MT and MST. Nature. 382: 539-541.

Hasselmo ME. (2005) What is the function of hippocampal theta rhythm? Linking behavioral data to phasic properties of field potential and unit recording data. Hippocampus. 15: 936-949.

Manns JR, Zilli EA, Ong ME, Hasselmo ME, Eichenbaum H. (2007) Hippocampal CA1 spiking during encoding and retrieval: relation to theta phase. Neurobiol Learn Mem. 87: 9-20.

CHAPTER 1.3. A NEURAL NETWORK MODEL OF SEMANTIC MEMORY LINKING FEATURE-BASED OBJECT REPRESENTATION AND WORD

INTRODUCTION

In literature are present several theories of semantic memory, with a special focus on object representation. Despite of the differences in many aspects, most of them agree in considering the semantic memory as a distributed process, which engages many different cortical areas and exploits a multi-modal representation of objects. Damasio (1989) suggests that semantic representation is not a static store, but a dynamical one which recollects many fragmented motor and sensory features. A subsequent influential theory by Warrington et al. (Warrington and McCarthy, 1983;Warrington and Shallice, 1984) assumes the presence of multiple channels of processing, within both the sensory and motor systems. This theory accounts for the existence of distinct semantic sub-systems, specialized for sensory and functional attributes of objects. Extensions of this idea were formulated by Lauro-Grotto et al. (1997), Snowden et al. (2004), and Gainotti (Gainotti, 2000;Gainotti, 2006). All these authors assume that the semantic system is an integrated multimodal network, in which different areas store modality specific information. Similarly, the "Hierarchical Inference Theory" (Humphreys and Forde, 2001) assumes that semantic memory is organized by modality-specific stores for features, and that concepts derive from an interactive process between the semantic and perceptual levels of representation. Caramazza et al. (Caramazza et al., 1990;Caramazza and Shelton, 1998) proposed that classes (an essential aspect of semantics) are organized by conceptual features that are highly correlated, "neurally contiguous" and developed on an evolutionary basis. Similarly, Tyler and co-authors (2000) in a model known as "Conceptual

Structure Account", suggest that objects are represented as patterns of activation across features, and categories emerge from a distributed network among those items which exhibit shared or correlated features.

A recent conceptual model by Hart, Kraut and co-workers (Hart et al., 2002; Kraut et al., 2002) summarizes all these aspects clearly: the model assumes that Semantic Object Memory is based on cortical regions which encode both sensorimotor and higher order cognitive information (such as lexical, emotional, etc..); retrieval of objects from memory involves the activation of these distributed representations together, which are then integrated by means of synchronized activity modulated by the thalamus.

As we told above (see Introduction) a couple of problems have to be solved in order to realize a semantic neural network: i) information related to the same concept has to be linked together to realize the corresponding internal semantic representation; ii) create a relationship between representations and words; iii) segmentation problem for distinct internal representations.

A valuable support for the clarification of these problems and for the analysis of mechanisms formulated in these conceptual theories can derive from the use of computational models based on distributed neural networks, and from computer simulations. Among the others, Rogers et al. (2004) developed a computational model which contains three layers of units, coding for visual, verbal and semantic elements. The linguistic and visual representations communicate by means of the "hidden" semantic layer, which wishes to represents the "anterior temporal lobe". The model is then trained with the classical back-propagation algorithm, so that the linguistic representation can recall the visual one, and vice versa. Models which exploit features to study category-specific semantic impairment were also used by Small et al. (1995), McRae et al. (1997), Devlin et al. (1998) and Lambon Ralph et al. (2007). In the model by Devlin et al. (1998), in particular, features are topographically organized while specific categories are not. However, none of these models investigate how different words and their feature representation can coexist in the same semantic memory, i.e., they do not cope with the segmentation problem.

A model dealing with the retrieval of multiple patterns in semantic memory was proposed by Morelli et al. (2006) . In this model features are coded by neurons working in chaotic regimen, and the retrieval process consists of synchronization of neurons coding for the same pattern. The model is able to distinguish objects even in the presence of several common features.

As described in the previous chapters (see Chapt. 1.1 and 1.2), we developed an analogous model for object recognition (Ursino et al., 2006;Ursino et al., 2009) based on the following ideas: i) each object is described as a collection of abstract features (in the present simplified version of the model these features are just four); ii) each feature is encoded via a topological organization in a different cortical area. Similar features occupy proximal positions in the same area; iii) the relationship between features in the same object is learned from previous experience via a time-dependent Hebbian mechanism, which reinforces excitatory connections; iv) different features of the same object are linked together, but are separated from features in other simultaneous objects, via synchronization in the gamma band. To this end, the model uses Wilson-Cowan oscillators as the basic computation units.

The model incorporates both previous-knowledge (i.e., a long term memory) via the Hebbian training, and a similarity principle, thanks to the topological organisation of features. In the Chapter 1.2 we explain that the model can recognize multiple objects, and segment them from the other simultaneously present objects, even in case of deteriorated information, noise and moderate correlation among the inputs. Hence, it can represent a good candidate for building more complex models of semantic memory, in which multiple objects, represented as collection of features, engage a dynamical relationship with a linguistic area, maintaining their individuality via gamma-band synchronization.

The aim of the present work is to start building such a semantic memory model. To this end, the previous model has been expanded, including a lexical area, which represents words. This lexical area collects features (assumed to be extracted from a processing chain of sensory-motor information), and receives words (assumed to derive from a phonemic processing network); it thus

represents a converging zone, perhaps located in the anterior left temporal cortex (Rogers et al., 2004;Snowden et al., 2004;Ward, 2006). The overall network (feature + lexical) is then trained via a time-dependent Hebbian mechanism, so that words are associated with the corresponding objects, described as feature collections. Simulation results are presented, to show how the network can deal with the simultaneous presence of words and objects, and how objects can be evoked in the semantic memory from the presence of correlated information (for instance, a related word + one of the object's features).

The present model wishes to represent just a first version, portraying the basic mechanisms in act. Hence, in the last part of this chapter limitations of the current implementation are discussed and lines for further work delineated, aimed at building future more realistic and complete versions.

METHOD

Our model consists of two different layers: the first (named "feature network") is devoted to a description of objects represented as a collection of sensory-motor features. The second (named "lexical network") is devoted to the representation of words, from an upstream language process. The two networks communicate via trained synapses. Moreover, the lexical network also receives a signal from a "decision network", which recognizes whether a correct object information is present in the feature network, and avoids that a misleading representation can evoke a word. The two networks are separately described below; then, the algorithm for training synapses is presented and justified.

The bidimensional network of features

As described in the previous paper this network is composed of N neural oscillators, subdivided into F distinct cortical areas (see Figure1). Each area in the model is composed of $N_1 x N_2$ oscillators.

An oscillator may be silent, if it does not receive enough excitation, or may oscillate in the γ-frequency band, if excited by a sufficient input.

Each area is devoted to the representation of a specific attribute or feature of the object, according to a topological organization. Hence, one object is represented as the collection of *F* features (one feature per each area). We assume that each attribute has been extracted from a previous processing in the neocortex, which elaborates sensory-motor information.

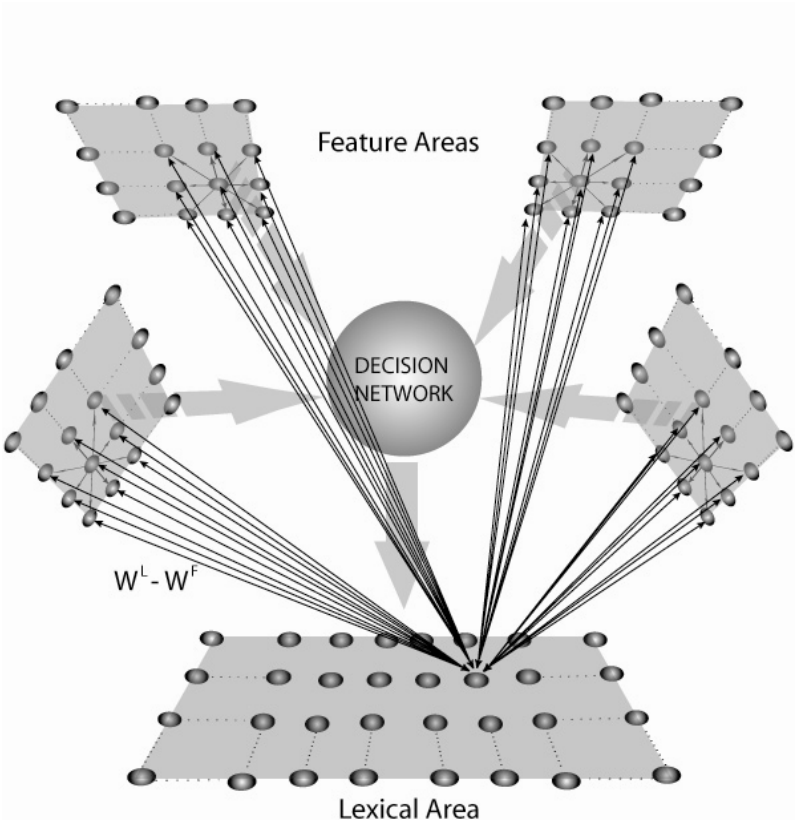


Figure 1– Schematic diagram describing the general structure of the network. The model presents 4 distinct Feature Areas (upper shadow squares) of 20x20 elements, which are described by means of Wilson-Cowan oscillators, and a Lexical Area of 40x40 elements (lower shadow square), which are represented by a first order dynamics and a sigmoidal relationship. In the Feature network, each oscillator is connected with other oscillators in the same area via lateral excitatory and inhibitory intra-area synapses, and with other oscillators in different areas via excitatory inter-area synapses. Moreover, elements of the feature and lexical networks are linked via excitatory recurrent synapses (W^F, W^L). Finally, the lexical area receives dishinibition from a decision network, which recognizes the presence of correctly-segmented objects.

Neural oscillators within the same area are connected via lateral excitatory and inhibitory synapses, according to a classical “Mexican hat” disposition, while neural oscillators belonging to different areas can be connected via excitatory synapses after training. These synapses are initially set to zero, but may assume a positive value through a learning phase, to memorize “prior knowledge” on attributes occurring together during the presentation of objects. Lateral synapses are not subjected to a training phase.

In the following, each oscillator will be denoted with the subscripts *ij* or *hk*. In the present study we adopted an exemplary network with 4 areas (F = 4) and 400 neural groups per area ($N_1 = N_2 =$ 20).

Each single oscillator consists of a feedback connection between an excitatory unit, x_{ij}, and an inhibitory unit, y_{ij} while the output of the network is the activity of all excitatory units. This is described with the following system of differential equations

$$\frac{d}{dt}x_{ij}(t) = -x_{ij}(t) + H\left(x_{ij}(t) - \beta \cdot y_{ij}(t) + E_{ij}(t) + V_{ij}^{L}(t) + I_{ij} - \varphi_x - z(t)\right) \quad (1)$$

$$\frac{d}{dt}y_{ij}(t) = -\gamma \cdot y_{ij}(t) + H\left(\alpha \cdot x_{ij}(t) - \varphi_y\right) + J_{ij}(t) \quad (2)$$

where H() represents a sigmoidal activation function defined as

$$H(\psi) = \frac{1}{1 + e^{-\frac{\psi}{T}}} \quad (3)$$

The other parameters in Eqs. (1) and (2) have the following meaning: α and β are positive parameters, defining the coupling from the excitatory to the inhibitory unit, and from the inhibitory to the excitatory unit of the same neural group, respectively. In particular, α significantly influences the amplitude of oscillations. Parameter γ is the reciprocal of a time constant and affects the oscillation frequency. The self-excitation of x_{ij} is set to 1, to establish a scale for the synaptic weights. Similarly, the time constant of x_{ij} is set to 1, and represents a scale for time *t*. φ_x and φ_y are offset terms for the sigmoidal functions in the excitatory and inhibitory units. I_{ij} represents the external stimulus for the oscillator in position *ij*, coming from the sensory-motor processing chain which extracts features. E_{ij} and J_{ij} represent coupling terms (respectively excitatory and inhibitory) from all other oscillators in the features network (see Eqs.4-7), while V_{ij}^{L} is the stimulus (excitatory) coming from the lexical area (Eq. 8). *z(t)* represents the activity of a global inhibitor whose role is to ensure separation among the objects simultaneously present. In particular, the

inhibitory signal prevents a subsequent object to pop up as long as a previous object is still active (see Chapters 1.1 and 1.2 for its description)

The coupling terms between elements in cortical areas, E_{ij} and J_{ij} in Eqs. (1) and (2) are computed as follows

$$E_{ij} = \sum_h \sum_k W_{ij,hk} \cdot x_{hk} + \sum_h \sum_k L^{EX}_{ij,hk} \cdot x_{hk} \tag{4}$$

$$J_{ij} = \sum_h \sum_k W_{ij,hk} \cdot x_{hk} + \sum_h \sum_k L^{IN}_{ij,hk} \cdot x_{hk} \tag{5}$$

where *ij* denotes the position of the postsynaptic (target) neuron, and *hk* the position of the presynaptic neuron, and the sums extend to all presynaptic neurons in the feature area. The symbols $W_{ij,hk}$ represent inter-area synapses, subjects to Hebbian learning (see next paragraph), which favour synchronization. The symbols $L^{EX}_{ij,hk}$ and $L^{IN}_{ij,hk}$ represent lateral excitatory and inhibitory synapses among neurons in the same area. It is worth noting that all terms $L^{EX}_{ij,hk}$ and $L^{IN}_{ij,hk}$ with neurons *ij* and *hk* belonging to *different* areas are set to zero. Conversely, all terms $W_{ij,hk}$, linking neurons *ij* and *hk* in the *same* area, are set to zero.

The Mexican hat disposition for the intra-area connections has been realized by means of two Gaussian functions, with excitation stronger but narrower than inhibition. Hence,

$$L^{EX}_{ij,hk} = \begin{cases} L^{EX}_0 e^{-\left[(i-h)^2+(j-k)^2\right]/(2\sigma^2_{ex})} & \text{if } ij \text{ and } hk \text{ are in the same area} \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

$$L^{IN}_{ij,hk} = \begin{cases} L^{IN}_0 e^{-\left[(i-h)^2+(j-k)^2\right]/(2\sigma^2_{in})} & \text{if } ij \text{ and } hk \text{ are in the same area} \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

where L^{EX}_0 and L^{IN}_0 are constant parameters, which establish the strength of lateral (excitatory and inhibitory) synapses, and σ_{ex} and σ_{in} determine the extension of these synapses.

Finally, the term V^L_{ij} coming from the lexical area is calculated as follows

$$V_{ij}^{L} = \sum_{h}\sum_{k} W_{ij,hk}^{L} \cdot x_{hk}^{L} \quad (8)$$

where x_{hk}^{L} represents the activity of the neuron *hk* in the lexical area and the symbols $W_{ij,hk}^{L}$ are the synapses from the lexical to the feature network (which are subject to Hebbian learning, see below).

The bidimensional lexical area

This area is made of M_1xM_2 units, described via a first order dynamic and a sigmoidal relationship. Each unit represents a specific "word". It can receive an input from a pre-processing stage which detects words from phonemes (see for instance Hopfield and Brody (Hopfield and Brody, 2001) for a possible model) or from groups of written letters, but it can also be stimulated through long-range synapses coming from the feature network; hence it represents an amodal convergence zone, as often hypothesized in the anterior temporal lobe (Damasio, 1989;Snowden et al., 2004;Ward, 2006). In this way, a "word" is linked with elements in feature areas representing specific properties of a stored object. All together, a "word" and its specific attributes are combined to embody the semantic meaning of that concept and the integrated network can indifferently be activated by language or sensory-motor information of an object.

The long range excitatory synapses between the lexical and the feature area are the result of a learning phase during which each "word" and the corresponding object are presented to the network to link an object and its name.

A problem with the lexical area is that its elements must be activated from the sensory-motor route only if an object is correctly recognized from its features, and correctly segmented from other objects (for exempla of incorrect segmentation in "difficult" conditions see the previous chapter). In case of incorrect object recognition or wrong segmentation, the corresponding word must not be evoked. To deal with this problem, we exploited a "decision network", developed in the previous chapter. This network received inputs from the feature areas and verified a certain number of requirements for the feature oscillators activity, to decide whether a correct object (i.e., an object

represented by exactly four simultaneous distinct features) is present or not: in particular, high values of its output (usually above 6) signified that object was correctly perceived. A detailed description of this decision network can be found in Chapter 1.2, hence it is not replicated here for briefness. In the present version of the model, we slightly modified this decision network so that its final output is ON/OFF in type. To this aim, we added a further layer to the previous decision network, which realizes a comparator with a threshold (φ^L) as high as 6. Accordingly, the final output of the decision network now assumes value 1 in case of correct object detection and 0 in case of incorrect detection.

In the present model we assume that the lexical network can be activated by the elements of the feature areas only if the decision network is in the ON state. This is realized sending sufficient inhibition to all elements of the lexical area. This inhibition is withdrawn by the decision network, as soon as a correct object is present.

In the following each element of the lexical area will be denoted with the subscripts *ij* or *hk* (*i*, h = 1, 2, …, M_1; *j*, k = 1,2,…, M_2) and with the superscript *L*. In the present study we adopted M_1 = M_2 = 40. Each single element exhibits a sigmoidal relationship (with lower threshold and upper saturation) and a first order dynamics (with a given time constant). This is described via the following differential equation:

$$\tau^L \cdot \frac{d}{dt} x_{ij}^L(t) = -x_{ij}^L(t) + H^L\left(u_{ij}^L(t)\right) \quad ; \tag{9}$$

τ^L is the time constant, which determines the speed of the answer to the stimulus, and $H^L\left(u^L(t)\right)$ is a sigmoidal function. The latter is described by the following equation:

$$H^L\left(u^L(t)\right) = \frac{1}{1+e^{-\left(u^L(t)-\vartheta^L\right)\cdot p^L}}; \tag{10}$$

where ϑ^L defines the input value at which neuron activity is half the maximum (central point) and p^L sets the slope at the central point. Eq. 10 conventionally sets the maximal neuron activity at 1 (i.e., all neuron activities are normalized to the maximum).

According to the previous description, the overall input, $u_{ij}^{L}(t)$, to a lexical neuron in the *ij*-position can be computed as follows

$$u_{ij}^{L}(t) = I_{ij}^{L}(t) + V_{ij}^{F} - G^{L} \cdot (1 - z^{L}(t)) \tag{11}$$

$I_{ij}^{L}(t)$ is the input produced by an external linguistic stimulation. V_{ij}^{F} represents the intensity of the input due to synaptic connections from the feature network; this synaptic input is computed as follows:

$$V_{ij}^{F} = \sum_{h} \sum_{k} W_{ij,hk}^{F} \cdot x_{hk} \tag{12}$$

where x_{hk} represents the activity of the neuron *hk* in the Feature Areas and $W_{ij,hk}^{F}$ the strength of synapses. Finally, the term $G^{L} \cdot (1 - z^{L}(t))$ accounts for the inhibition received by the lexical area, withdrawn by the decision network. In particular, $z^{L}(t)$ is a binary variable representing the output of the decision network (1 in case of correct detection, 0 in case of incorrect detection – see description above); hence, the strength of the inhibition sent to the Lexical Area is G^{L} when the decision network is in the OFF state, and becomes 0 when the decision network shifts to the ON state. It is worth noting that the external linguistic input $I_{ij}^{L}(t)$, when present, is set sufficiently high to overcome the inhibition received by the lexical area.

Synapses training

Phase 1: Training of inter-area synapses within the feature network

In a first phase, the network is trained to recognize objects without the presence of words. This means that we first learn an object as a collection of its properties, and only subsequently we associate a name to the object. To this end, objects are presented separately and the feature network is trained with the same algorithm used in the previous chapter (Chapter 1.2). This is a time dependent Hebbian learning, based on the correlation between the present activity in the post-

synaptic neuron, and the average activity of the pre-synaptic neuron in the previous 10 ms. After the training phase, the network is able to reconstruct objects from their features, as illustrated in (Ursino et al., 2009) (all mathematical details can also be found in Chapter 1.2).

Phase 2: Training of long-range synapses among the Lexical and the Feature Networks

As for inter-area synapses within the Feature network, also synapses linking the Lexical and the Feature Networks are realized by means of a Hebbian training phase. These synapses are trained in order to assign a "name" to a previously learnt object. In the following, an object stored in the Features Network will be represented with the notation:

$$obj = [i_1, j_1 \quad i_2, j_2 \quad i_3, j_3 \quad i_4, j_4]$$

where i_f, j_f represent the position of the neuron signaling the *f*-th attribute (*f* = 1,2,..,*F*, with *F* = 4 in our examples). A word in the Lexical Area will be represented with the notation:

$$wr = [i, j]$$

where *i, j* represent the position of the neuron signaling a word.

We trained simultaneously synapses from Feature Areas to Lexical Area, $W^F_{ij,hk}$, and connections from Lexical elements to Features network, $W^L_{ij,hk}$, assuming that these long-range synapses are initially set to zero, and that they are increased on the basis of the correlation between the activity of the presynaptic and postsynaptic neurons (time-dependent Hebbian learning).

During the second phase of the learning algorithm each object is presented alone, coupled with its corresponding lexical term. To this end, we present all properties of an object to the feature network, by providing inputs I_{ij}, high enough to activate the corresponding oscillators, and simultaneously excite the lexical area with an input high enough to lead the corresponding unit close to saturation.

As discussed in chapter 1.2 (Ursino et al., 2009), recent experimental data suggest that synaptic potentiation occurs if the pre-synaptic input precedes post-synaptic activity by 10 ms or less (Abbott

and Nelson, 2000;Markram et al., 1997). Accordingly, we assumed that the Hebbian rule depends on the present value of post-synaptic activity, $x_{ij}(t)$ or $x_{ij}^{L}(t)$, and on the moving average of the pre-synaptic activity (say $m_{hk}(t)$ or $m_{hk}^{L}(t)$) computed during the previous 10 ms. We calculate the moving average signal as follows

$$m_{hk}(t) = \frac{\sum_{m=0}^{N_s-1} x_{hk}(t - mT_s)}{N_s} \tag{13}$$

for signals in Feature Areas, while

$$m_{hk}^{L}(t) = \frac{\sum_{m=0}^{N_s-1} x_{hk}^{L}(t - mT_s)}{N_s} \tag{14}$$

for lexical activities.

T_S is the sampling period (in milliseconds), and N_S is the number of samples contained within 10 ms (i.e., $N_S = 10/T_S$). The synapses linking two neurons (say *ij* and *hk*) are then modified as follows during the learning phase

$$W_{ij,hk}^{F}(t + T_S) = W_{ij,hk}^{F}(t) + \beta_{ij,hk}^{F} \cdot x_{ij}^{L}(t) \cdot m_{hk}(t) \tag{15}$$

$$W_{ij,hk}^{L}(t + T_S) = W_{ij,hk}^{L}(t) + \beta_{ij,hk}^{L} \cdot x_{ij}(t) \cdot m_{hk}^{L}(t)$$

where $\beta_{ij,hk}^{F}$ and $\beta_{ij,hk}^{L}$ represent learning factors.

Moreover, we assumed that inter-area synapses cannot overcome a maximum saturation value. This is realized assuming that learning factors are progressively reduced to zero when synapses approach saturation. We have

$$\beta_{ij,hk}^{L} = \beta_0^{L}\left(W_{\max}^{L} - W_{ij,hk}^{L}\right) \tag{17}$$

$$\beta_{ij,hk}^{F} = \beta_0^{F}\left(W_{\max}^{F} - W_{ij,hk}^{F}\right) \tag{18}$$

where $W_{\max}^{L}$ and $W_{\max}^{F}$ are maximum values allowed for any synapse, and $\beta_o^L W_{\max}^{L}$ and $\beta_0^F W_{\max}^{F}$ are maximum learning factors (i.e., learning factors when synapses are zero).

According to the previous equations, arrays of inter-area synapses can be asymmetrical. Eqs. 17-18 imply that each synapse approximately increases according to a sigmoidal relationship, with upper saturation $W_{max.}$ The slope of this sigmoidal relationship (hence the increasing rate) is determined by parameter β_0.

The strength of synapses $W_{ij,hk}^{L}$ and $W_{ij,hk}^{F}$ at the end of the presentation of one object and its name, depends on two factors: parameter β_0 and the duration of the period along which the object and the word are presented to the network. The longer is this period, the higher is the value of synapses, and the strength of language correlation.

In the following, parameter β_0 is assumed to be the same for all synapses at a given instant. However, this parameter may be modified from one object to the next during the learning phase and from one connection to another. In this way, the model may account for objects with a different relevance (for instance, for the effect of attention, emotion, expectation and for all other factors which may affect storage) and for the presence of some features more relevant than others.

In the present work, we memorized 3 different objects and their correlated names:

Obj1 = [5,5 5,35 35,35 35,5] Wr1 = [5,5];

Obj2 = [15,15 15,25 25,25 25,15] Wr2 = [15,15];

Obj3 = [15,5 15,35 25,35 25,5] Wr3 = [15,35].

Subsequently, to analyze conditions characterized by correlated objects (i.e., objects having a common feature), training was repeated, starting from null long-range synapses, replacing Obj3 with another object (Obj4) defined as follows:

Obj4 = [15,15 15,35 25,35 25,5] Wr3 = [15,35]

An example of the second training phase is shown in Fig. 2A. This figure displays the temporal activity of the four attributes describing an object and of the element representing its “name”.

Moreover, the "moving average signals" (i.e., the quantities $m_{hk}(t)$ in Eq. 15) are also shown for the four neurons in the Feature Areas. These neurons received an input value as high as 0.8 and their activity is synchronized thanks to previous learning in the training phase one.

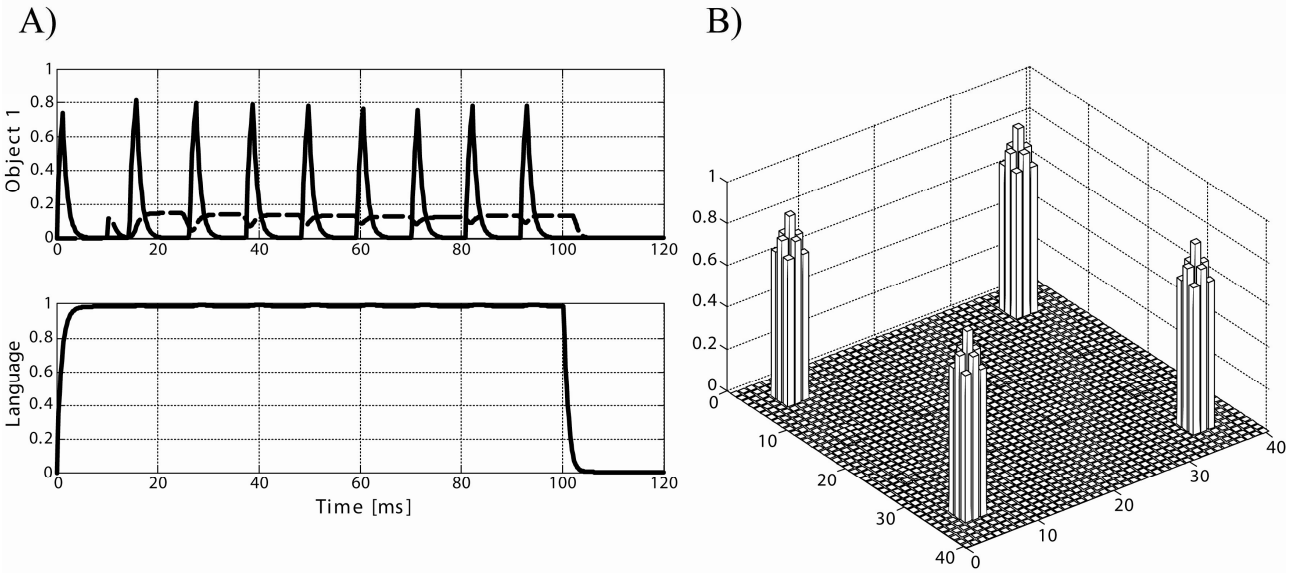


Figure 2 – Example of training of the synapses from the feature areas to the lexical area. The figure shows the training process of the synapses linking the four attributes of Obj1 ([5,5 5,35 35,35 35,5]) in the Feature Areas, with the corresponding word in the Lexical Area (codified by the neuron in position [5,5]). ***Panel A*****: *training of synapses from the feature to the lexical area* –The left upper panel shows the instantaneous activity (solid black line) and moving averaged activity (dashed black line) of the four oscillators representing the exact attributes of the object, during the training period. The left lower graph shows the instantaneous activity (solid black line) of the corresponding element in the lexical Area. Oscillators in the Feature Areas receive external input 0.8, while the lexical element receives an input as great as 20. The four oscillators synchronize due to the inter-area synapses between the Feature Areas. Inter-network synapses $W^F_{ij,hk}$ are created according to a Hebbian rule, thanks to the temporal superimposition of the moving average presynaptic signals in Feature Areas with the instantaneous activity of the post-synaptic neuron in the Lexical Area.** ***Panel B*****: $W^F_{ij,hk}$ *inter-network synapses after the learning phase*. The right plot shows the values of the synapses linking oscillators in the feature areas with the element [5,5] in the Lexical Area at the end of the learning process of Obj1 word. In particular, the figure displays the array $W^F_{55,hk}$ (representing the inter-area synapses directed to language element 5,5) by means of a three-dimensional graph: the *x,y* plane represents the coordinate *hk* within the feature network, and the height of the pixel in position *hk* represents the value of the synapse linking oscillator *hk* to element 5,5. Note that the element 5,5 receives the strongest synapses from the oscillators in the feature areas signalling the exact attributes, and weaker synapses from the other oscillators within the activation bubbles.**

Panel B in Fig. 2 displays the synapses linking the neuron 5,5 in the Lexical Area with the neurons in the Features Network after the learning phase (in particular, this figure shows the values of the array $W^F_{ij,hk}$ with h = 1, 2, … , 40; k = 1, 2, … , 40; i.e., , it represents the long-range synapses which *target* into the lexical neuron 5,5). As it is clear from this figure, after the learning phase the neuron receives synapses not only from the neurons describing the exact properties of the object (i.e., from neurons 5,5 5,35 35,35 35,5 in the Feature Areas) but also, although with smaller

strength, from the other proximal neurons in activation bubbles. This implements a similarity principle.

Table 1 – Values for parameters

Wilson-Cowan oscillators	
α	0.3
β	2.5
γ	0.3÷0.5 ms^{-1}
T	0.025
φ_x	0.7
φ_y	0.15
Lateral intra-area connections in Features Network	
L_0^{EX}	9
σ_{ex}	0.8
L_0^{IN}	3
σ_{in}	3.5
Hebbian rule for synapses between Features network and Lexical Area	
T_S	0.2 ms
N_S	50
β_0^L	0.125
$W_{\max}^L$	1
β_0^F	0.125
$W_{\max}^F$	1
Lexical Area	
ϑ^L	10
p^L	0.5
τ^L	1
G^L	20
Object attributes	
Obj1	[5,5 5,35 35,35 35,5]
Obj2	[15,15 15,25 25,25 25,15]
Obj3	[15,5 15,35 25,35 25,5]
Obj4	[15,15 15,35 25,35 25,5]
Lexical Elements	
Wr1	[5,5]
Wr2	[15,15]
Wr3	[15,35]

After the learning phase, the network can be used to recognize and reconstruct objects evoked by language or sensory inputs, even in the presence of lacking or modified information. In particular, the learning rates and training period were assigned so that each object can be reconstructed from two features, but cannot be correctly reconstructed starting from one feature only.

A list of parameters used in the present simulations is provided in Table 1.

RESULTS

The simulations portrayed in this section aim at illustrating some network abilities, such as the capacity to restore information from incomplete objects (i.e., objects which lack some features), the capacity to deal with multiple words, still maintaining a separate description of the individual objects, and the capacity to simultaneously retrieve information from words and perceived objects, also exploiting some correlation among objects.

Incomplete objects

Figure 3 shows the results of a simulation performed starting from incomplete information. The network has been trained with three different objects (Obj1, Obj2, Obj3), which do not share any feature.

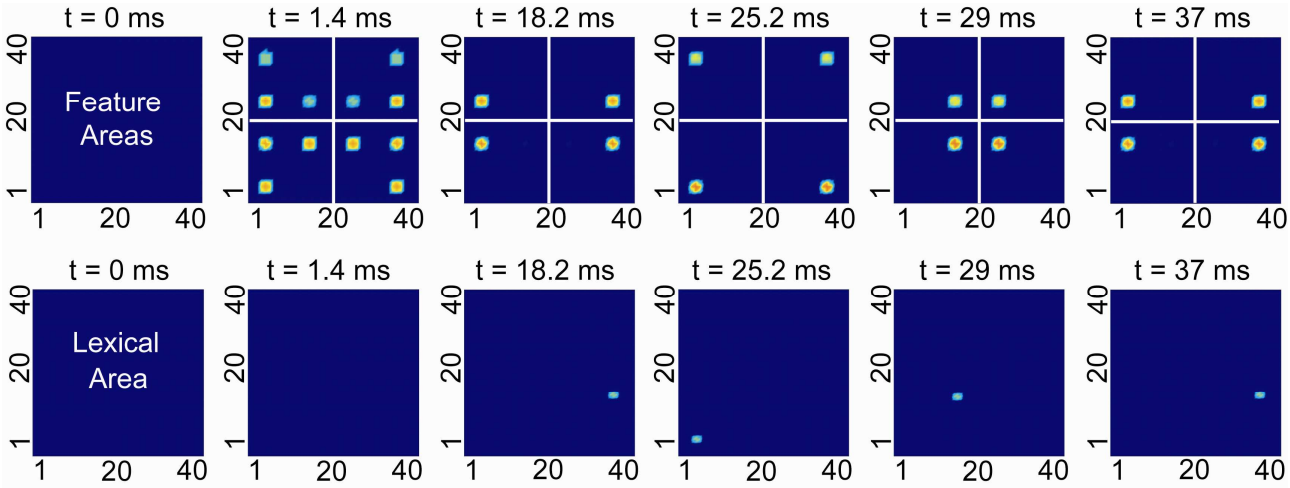


Figure 3 – Network activity at different snapshots during a simulation with incomplete objects. In the upper row each pixel represents an oscillator of the Feature Areas, while the lower row shows elements of the Lexical Area. The luminance (black means zero activity, white means maximum activity) is proportional to the corresponding oscillator's activity x_{ij}. During the simulation, object 1 and object 2 receive only two properties as input, while object 3 receives all properties. The external inputs to the Feature Areas for the stimulated properties are equal to 0.8. The Lexical Area does not receive any external stimulus. The network is able to reconstruct and recognize all the three presented objects in the Feature Areas and to evoke the right "word"(associated during the previous training phase) in the Lexical Area.

During the recovery phase, the feature area receives four features of the third object (i.e., the complete information) but only two features of the first and second objects. This condition corresponds to the case when a subject perceives objects from the external word, without any lexical input. After a short transient period (about 15 ms) during which the three objects appear together in the feature area, and no word is evoked in the lexical area, the network is able to segment the three objects correctly, by recovering the lacking features. During the appearance of individual objects (oscillating in time sharing in the γ-band) the decision network shifts into the ON state. As a consequence, the three objects evoke the corresponding words in the lexical area. It is worth noting that the representations of words in the lexical area oscillate in the γ-range too, i.e., with the same time-division as object representation in the feature area.

Multiple words

Fig. 4 shows an example of the network capacity to manage several simultaneous words, avoiding confusion in the object representation. In this simulation, the lexical area receives the three words (corresponding to the objects previously stored) whereas the feature area is unstimulated. This situation corresponds to the case of a subject who is listening to words, without any other external (sensory or motor) stimulation.

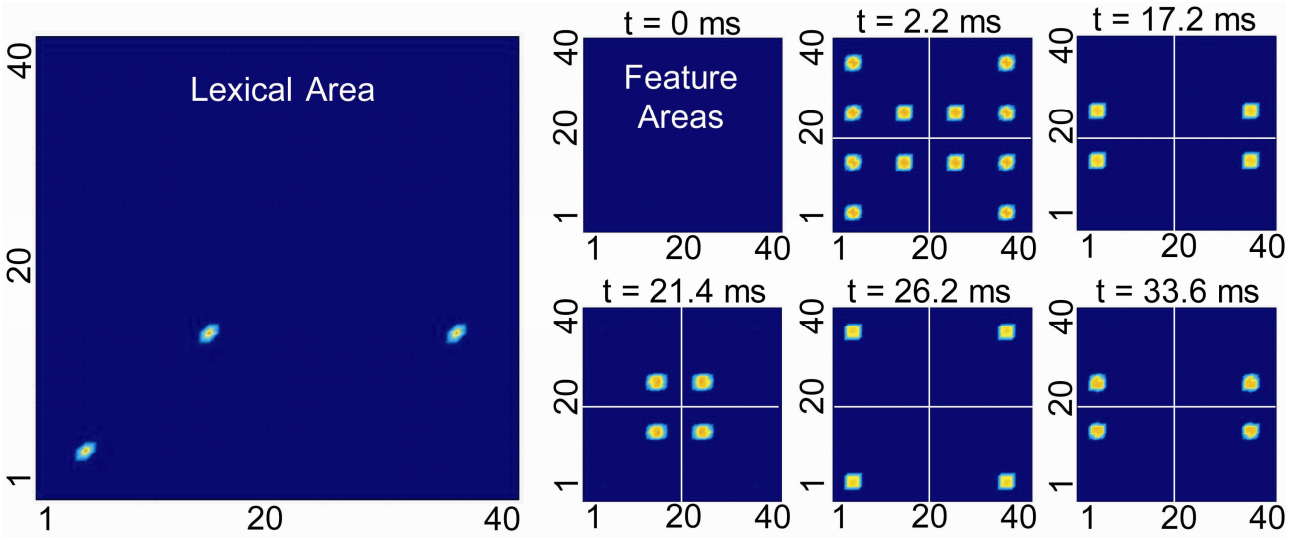


Figure 4 – Network activity at different snapshots during a simulation performed by giving 3 words simultaneously as input to the Lexical Area. In the left figure each pixel represents an element of the Lexical Area, while the right panels show different snapshots of the Feature Areas during the simulation. The external input to the Lexical Area for all stimulated words are as great as 20 to overcome inhibition, while Feature Areas do not receive any external stimulus. The network is able to evoke all the three objects in the Feature Areas and perform a correct segmentation in the γ-band, by associating the three objects to the corresponding stimulated "words" in the Lexical Area.

After a short transient period (about 15 ms) the network is able to evoke the representations of the objects in the feature areas, and to segment them correctly. Of course, the object representations in the feature area oscillate in the γ-range, whereas word representation is kept constant. This is simply a consequence of having used a non-oscillatory network to represent words.

Simultaneous word and feature inputs

Perhaps the most interesting characteristic of the network consists in the possibility to manage a mixed input condition, i.e., one in which the subject receives both words and sensory-motor information (the latter represented as a collection of features). An example is shown in Fig. 5. Here

the network receives one word (Obj2) and perceives two objects (Obj1 and Obj3, each of them with all features). After the usual transient period (15 ms), the word is able to evoke the corresponding object representation in the feature areas, while the two external objects evoke the corresponding words in the lexical area. The three objects are correctly segmented in the feature areas and oscillate in time division in the γ–range. In the lexical area only the word externally assigned exhibits a constant activity, while the other two words (evoked by the external features) oscillate in the γ-range.

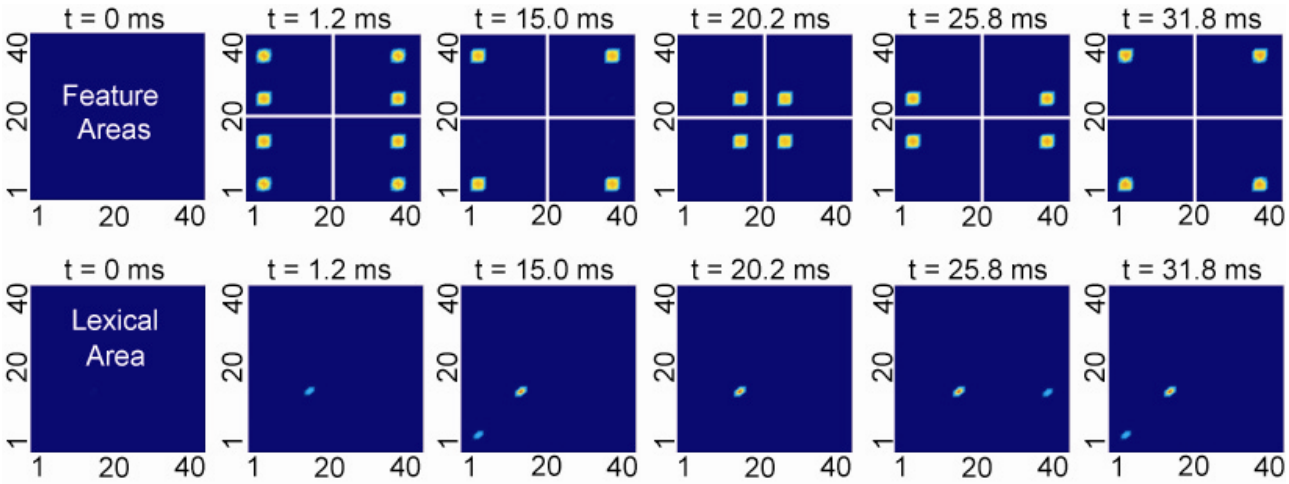


Figure 5 – Network activity at different snapshots during a simulation performed by giving one word (corresponding to Obj2) as input to the Lexical Area and two different objects (Obj1 and Obj3), completed with all their properties, to the Feature Areas. The network is able to recognize both the 2 objects directly presented to the Feature Areas and the one evoked by the word in the lexical Area. It is worth noting that the network constantly maintain the external word in the Lexical Area, whereas the other two words, evoked by stimulated features, oscillate in the γ-range.

A further interesting property of the present network is the possibility to establish a semantic link between words and features: this may occur when two objects, represented by different words, share some common features (i.e., they are correlated). In this situation, one may expect that hearing the first word may favour the recognition of the second object, even in presence of very incomplete perception. To illustrate this concept, the network was trained with three objects, two of them (Obj1 and Obj2) have no common features, while the third object (Obj4) shares one feature with the second. In a first simulation (not shown for briefness) the network received all features of the first object (Obj1) and just one feature of the second object (this feature was not shared by Obj4). As well expected, in this condition just Obj1 is recognized and evokes the corresponding word in the

lexical area, while the second object is not recognized (in fact, one feature is insufficient for object recognition; i.e., the other three features are not triggered by the inter-area connections and the corresponding word is not evoked). The same simulation was then repeated by stimulating the lexical area with the word corresponding to Obj4, and the feature areas with the same features described above (Obj1 with four features, Obj2 with just one not-shared feature). Fig. 6 shows that, in this case, the presence of the external word allows the complete reconstruction of Obj2, and the appearance of activity corresponding to its word in the lexical area. This signifies that the perception of a word can evoke a second correlated word, starting from a very incomplete perception.

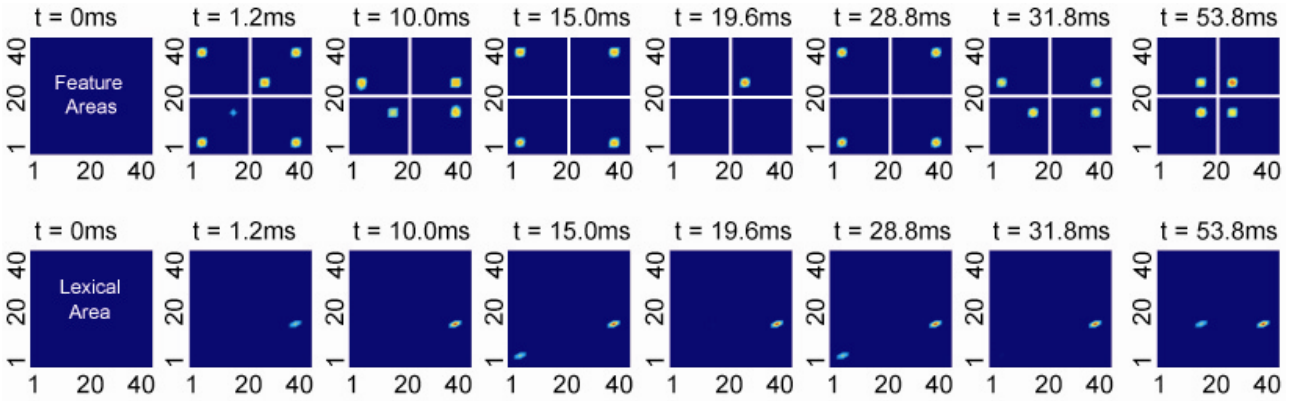


Figure 6 – Network activity during a simulation performed by giving all properties of an object to the feature areas (Obj1), one word as input to the Lexical Area (corresponding to Obj4) and just one property of a third object (Obj2) to the Feature Areas. It is worth noting that Obj2 shares one feature (not stimulated) with the object represented by the stimulated word. The network is able to recover the lacking features of Obj2, and to segment correctly all three objects in the feature areas. The external word is maintained constantly active in the lexical Area, whereas the other two words oscillate out of phase in the γ-band. It is noticeable that, in the absence of the external word, a single feature alone is insufficient to evoke the entire representation of Obj2 (unpublished simulation). Hence, a correlated word helps the recollection of an object from incomplete information.

Robustness vs. frequency changes

Finally, we performed some simulations to study whether synchronization is robust for what concerns moderate changes in the frequency of some oscillators. Our results suggest that moderate changes in frequency can be automatically corrected and oscillators are forced back to the gamma range by inputs coming from other oscillators in the network and from the global inhibitor. An example is shown in fig. 7.

In this simulation, twelve oscillators are simultaneously active, representing three different objects and words (but only eight of them are shown for briefness). However, two oscillators have a frequency quite different from the others (see figure legend for more details) while the other ten have similar frequency. The figure shows that, after a few cycles, oscillators in each object become synchronized (the first four for what concerns the first object and other four for what concerns the second object) despite the fact that two of them initially worked at a different frequency. More important, the frequency of the two "unstable" oscillators is quickly pulled back inside the gamma band.

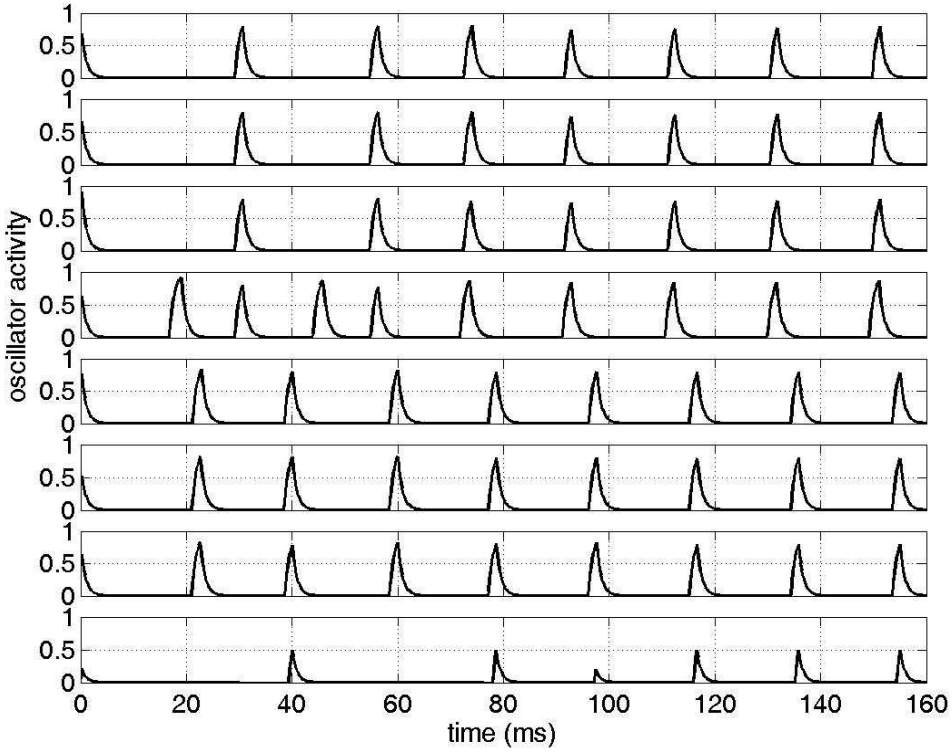


Figure 7 – Temporal activity of the eight oscillators corresponding to objects *Obj*1 and *Obj*2 (the four oscillators in the upper panel represent the features of the first object; the four oscillators in the bottom panels the feature of the second object). In this simulation all features of the three objects (*Obj*1, *Obj*2 and *Obj*3) were stimulated with an external input as high as 0.8, but some oscillators had different oscillation frequency (obtaining by varying parameter γ in Eq. (2)). In particular, the fourth oscillator in the figure had $\gamma = 0.45$ (corresponding to a frequency, in the absence of external connections, as high as 70–75 Hz) while the eighth oscillator had $\gamma = 0.15$ (corresponding to a frequency as low as 25 Hz). All other oscillators had $\gamma = 0.35$ (frequency about 45–50 Hz). Moreover, all oscillators started from a random initial state. As it is clear from this figure, after a transient period, the fourth and eighth oscillators become synchronized with the other features of the same object. In particular, the frequency of the fourth oscillator is reduced, while the frequency of the eighth is increased. Moreover, features of the third object (not shown here) are also correctly synchronized.

DISCUSSION

The idea that semantic memory involves a distributed representation of features, extending across different modalities and involving both sensory and motor information has been discussed by many authors in past years (Caramazza et al., 1990;Gainotti, 2000;Gainotti, 2006;HArt et al., 2007;Humphreys and Forde, 2001;Kraut et al., 2002;Lauro-Grotto et al., 1997;Tyler et al., 2000;Warrington and McCarthy, 1983;Warrington and Shallice, 1984) and is still the subject of active research by cognitive neuroscientists. This problem may benefit from computational approaches, which emphasize the virtues and limitations of present theories and help a clearer and rigorous conceptualisation of existing data.

In the present work, we propose a preliminary simple network for the simulation of semantic memory. Although extremely simplified compared with the reality, the network incorporates several characteristics, which should constitute the core of more sophisticate future models. The main characteristics of our network are summarized and critically commented hereafter:

i) abstract objects are represented as a collection of features; one object is recognized when all its features are in the active state. Furthermore, each feature is described by a zone of activation in an appropriate cortical area, and all features are topologically organized. This means that that similar features are coded by neural oscillators in proximal positions of the cortex, and are connected via reciprocal recurrent connections. This aspect implements a similarity principle in a plausible and straightforward manner: an object, similar to another object previously memorized, can still benefit from previous experience and can be recognized despite modest changes.

ii) multiple objects can be simultaneously represented in memory via synchronisation in the γ-band. Features which belong to the same object oscillate in phase, whereas features in different objects are out of phase. The idea that a time division can be exploited to represent mutiple objects was originally formulated by Milner (Milner, 1974) and Von der Marlsburg (von der Malsburg and Schneider, 1986), and has been recently applied to object segmentation especially in visual scenes (the so-called binding and segmentation problem, (Eckhorn, 1999;Singer and Gray, 1995;Wang and

Terman, 1997)). Our belief, already formulated in previous works (Ursino et al., 2006;Ursino et al., 2008), is that the same mechanism can be exploited to segment abstract objects too, represented as a collection of features. Several recent results support the idea that neural synchronization plays an essential role not only in low sensory perception but also in higher-cognitive tasks. A role of gamma activity has been demonstrated in recognition of music (Bhattacharya et al., 2001) words vs. non-words (where it seems to reflect association between words and meanings) (Pulvermüller et al., 1996) and recognition of black and white faces vs. meaningless figures (Rodriguez et al., 1999). Further studies suggest that theta and gamma oscillations play an important role in formation of declarative memory and retrieval (Osipova et al., 2006;Salinas and Sejnowski, 2001) and that synchronization increases with conscious perception compared with unconscious (subliminal) processes (Melloni et al., 2007).

iii) The object, represented by its features, is connected via re-entrant synapses to a lexical area devoted to the representation of words. In the present version of the model, words are described by means of “grand-mother cells” (one neuron per each word); of course, this aspect may be improved in future versions. The function of connections between the feature areas and the lexical area is to realize a sort of semantic relationship between the representation of objects and their words. This implicitly signifies that, after the training phase, the perception of objects may evoke the corresponding word, while listening to words may activate the multimodal object representation in cortical areas. Some results in the recent literature corroborate this viewpoint. Data from neurophysiological studies (Pulvermller et al., 2001;Pulvermüller et al., 2000), transcranial magnetic stimulation (Pulvermller et al., 2005b;Pulvermller et al., 2005a) and functional magnetic resonance (fMRI) (Hauk et al., 2004) suggest that the comprehension of words activates the motor and premotor cortex in a somatotopic manner. Martin et al. (Martin et al., 1995) found that colour words activate a region near the area involved in the perception of colour. Gonzales et al. (González et al., 2006), using fMRI, observed that subjects reading odour-related words display activation in

the olfactory regions of the brain. These results, taken together, support the idea that the neural representation of a word is associated with the corresponding perceptual information.

iv) Of course, the network needs a training phase in order to associate all features of the same object, and to associate objects with words. According to recent data, we assumed that training of synapses is based on the correlation between the instantaneous post-synaptic activity, and the pre-synaptic activity evaluated over a previous 10 ms interval. Indeed, various authors suggest that Hebbian reinforcement does not require a perfect correspondence between spikes, but operates over a temporal window (Abbott and Nelson, 2000;Markram et al., 1997). Buszacki (Buzsáki, 2006) hypothesized that this characteristic of synapse plasticity is one of the reasons for the ubiquitous role of gamma oscillations in the brain, since the γ-cycle approximately has the same temporal length as that required for the plasticity rules. In our model, a temporal window is essential to allow synchronisation among neural groups, which are initially out of phase (hence, whose activity is not perfectly superimposed) during the training period, but rapidly synchronize thanks to synaptic reinforcement.

In the present trials we assumed that network training, for each object, occurs in two distinct phases: in the first, an object is presented alone, i.e., without the associated word, and the different features of this object are linked together (for instance, a child is watching at a dog without listening any name). During a second phase the object and the corresponding word are presented together, and the object receives a name (the child is looking at the dog and is simultaneously listening the word “dog”). In other terms, we first trained synapses $W_{ij,hk}$ in Eqs. 4 and 5, and only subsequently we trained synapses $W^{L}_{ij,hk}$ and $W^{F}_{ij,hk}$ in Eqs. 8 and 12. This choice has been adopted since we claim it is more similar to natural learning. However, similar results may be obtained even by training all synapses together during a unique training phase.

An important aspect of our model, which may be exploited in future versions, is the possibility to realize simple semantic links between objects and words. In the present paper, we exploited the

possibility that two objects may have some common features in order to realize a simple association among the corresponding words. As illustrated in Fig. 6, perception of a single feature (which is insufficient to recognize one object and to evoke the corresponding word) may lead to object recognition in the presence of a correlated word. This simple example shows that memory is an integrated process, and that its content can be retrieved from the co-activation of different regions, coding for disparate properties (as in the classical test “desert” and “hump” allows retrieval of the word “camel”) (Kraut et al., 2002;Kraut et al., 2006). The idea that the correlation among features may be exploited to form classes of objects, and to detect category membership, without the need for a hierarchical representation of objects, has been formulated by some authors recently, and is at the basis of several conceptual theories of semantic memory (see (Gainotti, 2006;HArt et al., 2007) for a review). However, we are aware just of a handful of models which try to implement these concepts through neural networks and to simulate the main consequences by a computer.

A recent model by Morelli et al. (Morelli et al., 2006) shares some aspect with ours. In both models an object is represented as a collection of features, and the possibility to kept different objects simultaneously in memory is achieved via the synchronized firing activity of neurons which code for the same objects. Moreover, both models deal with the possibility that objects share some common features. Morelli et al. used a greater number of features (16) and analyzed the case of objects with 3 common features; while we used just 4 features in our exemplary simulations, with the possibility of one common feature. A significant difference is that we used Wilson-Cowan oscillators to synchronize features, hence in our simulations objects appear in “time division”, whereas Morelli et al. used chaotic dynamics. Other distinguishing aspects of our model concern the topological organisation of features (which implements a similarity principle), the presence of a decision network to recognize plausible objects, and the relationship between objects and words. Pulvermuller (Pulvermuller, 2005) and Simmons and Barsalou (Simmons and Barsalou, 2003) hypothesized that different features of a word are topographically organized in different regions of

the cortex. A model including a topological organization of features was developed by Devlin et al. (Devlin et al., 1998).

An important problem, that deserves some discussion, is how synchronisation among cognitive processes can be assessed in humans, and what may be the role of synchronisation in cognition.

In its broad sense, synchronisation represents the concurrence of events in time (Buzsáki, 2006). Different methods to quantify this concept have been proposed. The older ones used spectral coherence (Bressler et al., 1993) or detection of the temporal position of maxima values in filtered signals (Yordanova et al., 1997). More recently, a quantity used to quantify synchronism is the phase coherence (Lachaux et al., 1999;Melloni et al., 2007;Rodriguez et al., 1999). In brief, the method involves computing the phase difference between two signals in a time window, at a given frequency, and assessing the stability of such phase difference through all trials. Phase coherence between EEG electrodes has been used in neurophysiological studies, as a measure able to detect a functional relation, reflecting the co-activation of distant and task-relevant brain sites. A special case of phase coherence with zero phase difference is phase synchrony. Variants of phase synchronisation are phase coupling among different frequencies and phase-lock to an external stimulus [the interested reader can find more details in (Sauseng and Klimesch, 2008)]. Of course, the temporal resolution necessary to detect synchronisation with enough accuracy (order of tens or a few milliseconds) is not in the range of traditional neuroimaging techniques, such as fMRI or PET, and is only accessible by surface EEG or MEG. The latter techniques, however, measure electrical activity only in a large population of neurons with a poor spatial resolution. Finally, surface EEG and MEG should be processed with algorithms for electromagnetic source localisation, in order to detect activity in specific regions of interest in the cortex, involved in cognitive aspects of the task.

A further problem is why synchronization “especially in the gamma range” may be so important in cognitive processes. A first aspect is that neurons can fire only if they receive enough input excitation. Only coherent source activity, synchronized in a small temporal window, can be effective to trigger downstream neurons. Hence, synchronized activity is the most effective

mechanism to make information from spatially separated brain areas simultaneously available for further processing (binding) by maintaining it distinct from other information (segregation). Some authors suggested that binding of neural activity is necessary to trigger short-term memory (Crick and Koch, 1990), or to form a “global workspace” thus leading to the emergence of conscious states (Damasio, 1990;Grossberg, 1999). Another important point is that synchronization in the gamma-band can efficaciously trigger changes in synaptic efficacy (Buzsáki, 2006;Engel and Singer, 2001). Indeed, the temporal relation between pre-synaptic and post-synaptic activities, necessary to cause long term potentiation or long term depression, is within 10-20 ms (that is typical of the gamma range). This signifies that, synchronized oscillation in the gamma range can be crucial for memory formation.

Of course, in order to fully exploit phase synchronization in the gamma range, one needs that the frequency of oscillators is quite stable, and that possible frequency derangements outside the gamma range are rapidly corrected. There are various mechanisms (both in vivo and in the model) which may improve robustness against frequency changes. First, neural groups which participate to the representation of the same object are linked via excitatory synapses (in our model they target to both excitatory and inhibitory populations). These synapses rapidly induce synchronisation of the neural groups. In particular, neural groups which exhibit off-band activity are forced back into the gamma-band by inputs coming from the other populations. A second mechanism that may help robustness in the model is the activity of the global inhibitor. As soon as a single object terminates its cycle, all other neural groups are simultaneously uninhibited, and so may start their activation together. This mechanism forces the frequency of all oscillators (even in different objects) to the same frequency (see Fig. 7).

Finally, we wish to mention that, in vivo, gamma oscillations can be especially ascribed to connections between interneurons with fast synaptic kinetics (GABAa synapses) whose time constant is just a few milliseconds (Bartos et al., 2007;Jefferys et al., 1996). Hence, a rapid method to induce a gamma-band oscillation in real neural networks is to provide an excitatory input to these

interneurons. This makes the mechanism of gamma-band generation ubiquitous and easily controllable from external top-down influences.

Finally, it is important to point out the main limitations of the present model, which should become the subject of future improvements and extensions. A fundamental limitation is that each object is represented by an exact number of features (four in the present exempla); each feature must be active and synchronized with all the others for the decision network to recognize a plausible object. Of course, in the reality an object can be represented by a variable number of features, which, moreover, can be differently allocated in sensory or motor areas. For instance, differences in the perception of action vs. non-action words have been ascribed to the prevalence of motor features in the first class, and of sensory features in the second (Caramazza and Shelton, 1998;Crutch and Warrington, 2003). Similarly, the description of some objects (such as animals or plants) may require a greater number of features compared with those necessary for the description of simpler objects (Tyler et al., 2000;Tyler and Moss, 2001). For that reason, a fundamental forthcoming extension of our model should involve the possibility to describe objects with a variable number of features, distributed in alternative ways among cortical regions.

A second important limitation is that words are represented in the lexical area via “grand mother cells”, i.e., only one neuron is used to represent an individual word. This choice has been adopted since the aim of the present study was not that of providing a detailed description of the lexical area, but just using this area to establish a link between objects and words, so that words can evoke an object representation, and vice versa. Of course, future model versions may incorporate a more sophisticate description of the lexical area, in which, for instance, neurons code for different phonemes and a word is described by an ordered chain of phonemes (see for instance, (Marslen-Wilson, 1987)). Nevertheless, a representation of words as grand-mother cells has been adopted in recent models of word recognition from speech. For instance, Hopfield and Brody in a recent influential model (Hopfield and Brody, 2001) proposes that a single neuron with a small time constant can detect the occurrence of a particular spatio-temporal pattern (for instance a short word

or a phoneme) by revealing the transient synchronisation of previous events, occurring in a specific temporal sequence. The model by Hopfield, integrated with a working memory circuit to maintain words for the entire duration of a task, may represent the basis for a more sophisticate description of the relationship between the lexical area, speech and its auditory input.

A third future improvement may concern the decision network. In the present version, the decision network is necessary to avoid that a “false” object representation can evoke words. For instance, this is the case when more than four features are simultaneously active, i.e., the feature network failed to solve the segmentation problem. We are devising future versions in which the lexical area can recognize the occurrence of objects, even without the need for an upstream decision network.

A further point is that, in the present model, we have not included any connection among individual words in the lexical area. An important extension of our study may assume that words are reciprocally connected by excitatory weights, derived from a learning phase. For instance, two words occurring frequently together, may develop a recurrent link. In this manner, a semantic relationship between words and objects can be realized not only thanks to correlates features (which signal category membership) but also by direct connections among words.

Finally, future versions of the model may test the behaviour of alternative equations, such as relaxation oscillators, integrate and fire neurons or chaotic dynamics (as in the model by Morelli et al. (Morelli et al., 2006)) to synchronize neural activity. The virtues and limitations of individual neuron models in the field of semantic memory is an aspect which still deserves extensive analysis.

REFERENCES

Abbott LF and Nelson SB. (2000) Synaptic plasticity: taming the beast. Nat.Neurosci. 3: 1178-1183.

Bartos M, Vida I, Jonas P. (2007) Synaptic mechanisms of synchronized gamma oscillations in inhibitory interneuron networks. Nat Rev Neurosci 8: 45-56.

Bhattacharya J, Petsche H, Pereda E. (2001) Long-range synchrony in the gamma band: role in music perception. J.Neurosci. 21: 6329-6337.

Bressler SL, Coppola R, Nakamura R. (1993) Episodic multiregional cortical coherence at multiple frequencies during visual task performance. Nature 366: 153-156.

Buzsáki G. (2006) Rhythms of the brain. Oxford University Press, New York.

Caramazza A, Hillis A, Rapp B. (1990) The multiple semantics hypothesis: Multiple confusions? Cognitive Neuropsychology 7: 161-189.

Caramazza A and Shelton JR. (1998) Domain-specific knowledge systems in the brain the animate-inanimate distinction. J.Cogn.Neurosci. 10: 1-34.

Crick F and Koch C. (1990) "Towards a neurobiological theory of consciousness". Semin.Neurosci. 2: 263-275.

Crutch SJ and Warrington EK. (2003) The selective impairment of fruit and vegetable knowledge: A multiple processing channels account of fine-grain category specificity. Cognitive Neuropsychology 20: 355-372.

Damasio AR. (1989) Time-locked multiregional retro-activation: a systems level proposal for the neural substrates of recall and recognition. Cognition 33: 25-62.

Damasio AR. (1990) Synchronous activation in multiple cortical regions: a mechanism for recall. Semin.Neurosci. 2: 287-296.

Devlin JT, Gonnerman LM, Andersen ES, Seidenberg MS. (1998) Category-specific semantic deficits in focal and widespread brain damage: a computational account. J Cogn Neurosci 10: 77-94.

Eckhorn R. (1999) Neural Mechanisms of Scene Segmentation: Recordings from the Visual Cortex Suggest Basic Circuits for Linking Field Models. IEEE transactions on neural networks 10: 464-479.

Engel AK and Singer W. (2001) Temporal binding and the neural correlates of sensory awareness. Trends Cogn Sci 5: 16-25.

Gainotti G. (2000) What the locus of brain lesions tells us about the nature of the cognitive defect underlying category-specific disorders: a review. Cortex 36: 539-559.

Gainotti G. (2006) Anatomical functional and cognitive determinants of semantic memory disorders. Neuroscience and Behavioral Reviews 30: 577-594.

González J, Barros-Loscertales A, Pulvermuller F, Meseguer V, Sanjun A, Belloch V, Avila C. (2006) Reading cinnamon activates olfactory brain regions. Neuroimage 32: 906-912.

Grossberg S. (1999) The link between brain learning, attention, and consciousness. Conscious Cogn 8: 1-44.

Hart J, Anand R, Zoccoli S, Maguire M, Gamino J, Tillman G, King R, Kraut MA. (2007) Neural substrates of semantic memory. J Int Neuropsychol Soc 13: 865-880.

Hart J, Segal JB, Adkins E, Kraut M. (2002) Neural substrates of semantics. In: Hillis, A., ed., Handbook of language disorders. Psychology Press, Philadelphia.

Hauk O, Johnsrude I, Pulvermuller F. (2004) Somatotopic representation of action words in human motor and premotor cortex. Neuron 41: 301-307.

Hopfield JJ and Brody CD. (2001) What is a moment? Transient synchrony as a collective mechanism for spatiotemporal integration. Proc Natl Acad Sci U S A 98: 1282-1287.

Humphreys GW and Forde EME. (2001) Hierarchies, similarity, and interactivity in object recognition: "Category-specific" neurophysiological deficits. Behavioral and Brain Sciences 24: 453-509.

Jefferys JG, Traub RD, Whittington MA. (1996) Neuronal networks for induced 40 Hz rhythms. Trends Neurosci 19: 202-208.

Kraut MA, Cherry B, Pitcock JA, Vestal L, Henderson VW, Hart J. (2006) The Semantic Object Retrieval Test (SORT) in normal aging and Alzheimer disease. Cogn Behav Neurol 19: 177-184.

Kraut MA, Kremen S, Segal JB, Calhoun V, Moo LR, Hart J. (2002) Object activation from features in the semantic system. J Cogn Neurosci 14: 24-36.

Lachaux JP, Rodriguez E, Martinerie J, Varela FJ. (1999) Measuring phase synchrony in brain signals. Hum Brain Mapp 8: 194-208.

Lambon Ralph MA, Lowe C, Rogers TT. (2007) Neural basis of category-specific demantic deficits for living things: evidence from semantic dementia, HSVE and a neural network model. Brain 130: 1127-1137.

Lauro-Grotto R, Reich S, Visadoro M. (1997) The computational role of conscious processing in a model of semantic memory. In: Ito M, Miyashita S, Rolls E. (Eds.), Cognition, Computation and Consciousness. Oxford University Press, Oxford, pp. 249-263.

Markram H, Lübke J, Frotscher M, Sakmann B. (1997) Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSSs. Science 275: 213-215.

Marslen-Wilson WD. (1987) Functional parallelism in spoken word recognition. Cognition 8: 1-71.

Martin A, Haxby JV, Lalonde FM, Wiggs CL, Ungerleider LG. (1995) Discrete cortical regions associated with knowledge of colour and knowledge of action. Science 270 : 102-105.

McRae K, de Sa VR, Seidenberg MS. (1997) On the nature and scope of featural representations of word meaning. J Exp Psychol Gen 126: 99-130.

Melloni L, Molina C, Pena M, Torres D, Singer W, Rodriguez E. (2007) Synchronization of neural activity across cortical areas correlates with conscious perception. J.Neurosci. 27: 2858-2865.

Milner PM. (1974) A model for visual shape recognition. Psychol.Rev. 81: 521-535.

Morelli A, Lauro GR, Arecchi FT. (2006) Neural coding for the retrieval of multiple memory patterns. BioSystems 86: 100-109.

Osipova D, Takashima A, Oostenvald R, Fernández G, Maris E, Jensen O. (2006) Theta and gamma oscillations predict encoding and retrieval of declarative memory. J.Neurosci. 26: 7523-7531.

Pulvermüller F, Hauk O, Nikulin VV, Ilmoniemi RJ. (2005a) Functional links between motor and language systems. Eur J Neurosci 21: 793-797.

Pulvermüller F, Hrle M, Hummel,F. (2001) Walking or talking? Behavioral and neurophysiological correlates of action verb processing. Brain Lang 78: 143-168.

Pulvermüller F, Shtyrov Y, Ilmoniemi R. (2005b) Brain signatures of meaning access in action word recognition. J Cogn Neurosci 17: 884-892.

Pulvermüller F. (2005) Brain mechanisms linking language and action. Nat.Rev.Neurosci. 6: 576-582.

Pulvermüller F, Hrle M, Hummel F. (2000) Neurophysiological distinction of verb categories. NeuroReport 11: 2789-2793.

Pulvermüller F, Preissl H, Lutzenberger W, Birbaumer N. (1996) Brain rhythms of language: nouns versus verbs. Eur.J.Neurosci. 8: 937-941.

Rodriguez E, George N, Lachaux JP, Martinerie J, Renault B, Varela FJ. (1999) Perception's shadow: long-distance synchronization of human brain activity. Nature 397: 430-433.

Rogers TT, Lambon Ralph MA, Garrard P, Bozeat S, McClelland JL, Hodges JR, Patterson K. (2004) Structure and deterioration of semantic memory: a neuropsychological and computational investigation. Psychol Rev 111: 205-235.

Salinas E and Sejnowski TJ. (2001) Correlated neuronal activity and the flow of neural information. Nat.Rev.Neurosci. 2: 539-550.

Sauseng P and Klimesch W. (2008) What does phase information of oscillatory brain activity tell us about cognitive processes? Neurosci Biobehav Rev 32: 1001-1013.

Simmons WK and Barsalou LW. (2003) The similarity-in-topography principle: reconciling theories of conceptual deficits. Cogn.Neuropsychol. 20: 451-486.

Singer W and Gray CM. (1995) Visual Feature integration and the temporal correlation hypothesis. Ann.Rev.Neurosci. 18: 555-586.

Small SL, Hart J, Nguyen T, Gordon B. (1995) Distributed representations of semantic knowledge in the brain. Brain 118 (Pt 2): 441-453.

Snowden JS, Thompson JC, Neary D. (2004) Knowledge of famous faces and names in semantic dementia. Brain 127: 860-872.

Tulving E. (1983) Elements of episodic memory. Oxford Press, Oxford.

Tyler LK and Moss HE. (2001) Towards a distributed account of conceptual knowledge. Trends Cogn Sci 5: 244-252.

Tyler LK, Moss HE, Durrant-Peatfield MR, Levy JP. (2000) Conceptual structure and the structure of concepts: a distributed account of category-specific deficits. Brain Lang 75: 195-231.

Ursino M, Magosso E, Cuppini C. (2009) Recognition of abstract objects via neural oscillators: interaction among topological organization, associative memory and gamma-band synchronization. IEEE Tr.Neural Networks 20: 316-335.

Ursino M, Magosso E, La Cara GE, Cuppini C. (2006) Object segmentation and recovery via neural oscillators implementing the similarity and prior knowledge gestalt rules. BioSystems 85: 201-218.

von der Malsburg C and Schneider W. (1986) A neural cocktail-party processor. Biol.Cybern. 54: 29-40.

Wang D and Terman D. (1997) Image Segmentation based on oscillatory correlation. Neural Computation 9: 805-836.

Ward J. (2006) The student's guide to cognitive neuroscience. Psychology Press, Hove and New York.

Warrington EK and McCarthy R. (1983) Category specific access dysphasia. Brain 106: 859-878.

Warrington EK and Shallice T. (1984) Category specific semantic impairments. Brain 107: 829-854.

Yordanova J, Kolev V, Demiralp T. (1997) The phase-locking of auditory gamma band responses in humans is sensitive to task processing. NeuroReport 8: 3999-4004.

Part 2. MULTISENSORY INTEGRATION

In the natural environment, stimuli of different modalities occur at various locations in space and time. Among the brain's most important function, detecting, decoding and interpreting these information about external significant events are of primary importance and involve a host of neural circuitries, linking different specialized brain regions.

The ability to get information from different sensory sources and, more important, the skill to integrate them into a unitary internal object representation, is fundamental to determine the relationships among different sensory signals and to enhance detection and identification of external events to trigger the correct responses (Stein and Meredith, 1993; Welch and Warren, 1986).

This sensory interaction is usually referred as "multisensory integration", that denotes the capability of a neuron to produce a different response to a combination of various sensory-modality stimuli with respect to the single unisensory components.

Multisensory integration can result in either enhancement or depression of the neuron's response. This reflects the salience and the relationship between external sensory inputs: if they are related to the same event they will be integrated together to facilitate and speed up the detection, resulting in an higher and faster neuron's activity (*multisensory enhancement*); whereas if two or more stimuli derive from events competing for the brain's attention, the elicited response would be decreased (*multisensory depression*).

Evidence for multisensory convergence at neural level has been found in various structures of the mammalian brain outside the primary sensory areas (Stein and Meredith, 1993). The most studied locus of multisensory interaction is a layered midbrain structure, the superior colliculus (SC). This structure plays a critical role in the generation and control of orienting movements of the head and eyes towards external events (Sparks, 1986; Stein and Meredith, 1993). Many neurons in the deep layers of the SC receive converging visual, auditory and somatosensory afferents from various subcortical and extraprimary cortical sources (Edwards et al., 1979; Huerta and Harting, 1984; Wallace et al., 1993). Responses of such neurons to a combination of stimuli delivered in multiple sensory modalities can differ significantly from those evoked by any of their unisensory

inputs in a way that substantially facilitates the role of the SC in controlling attentive and orientation behaviour (Frens et al., 1995; Schroger and Widmann, 1998; Stein et al., 1989).

Several neurophysiological and behavioural studies were performed in order to identify the principles ruling multisensory integration in SC neurons. This process faces two different issues: the coincidence of sensory information in space, and the temporal coincidence of converging inputs. The combination of two different sensory stimuli (e.g., auditory and visual), present at close spatial and temporal proximity (as it occurs when they derive from the same event), is typically synergistic producing a neuron's response which is significantly grater than that evoked by the most effective of the two unimodal inputs individually (*multisensory enhancement*) (Bell et al., 2001; Kadunce et al., 2001; Meredith and Stein, 1986b; Meredith and Stein, 1996; Wallace et al., 1998). In some neurons, the response enhancement may even exceed the sum of their individual unimodal responses (*superadditivity*) (Kadunce et al., 2001; Meredith and Stein, 1986b; Perrault Jr et al., 2003; Perrault Jr et al., 2005; Wallace et al., 1998). On the other hand, when the two stimuli are presented at different locations or at different time (i.e. they likely derive from different events), no interaction occurs or the neuron's response is considerably depressed (*multisensory depression*) (Kadunce et al., 1997; Kadunce et al., 2001; Meredith and Stein, 1986a; Wallace et al., 1998). Globally, these properties of cross-modal integration are known as the *spatial rule*.

The spatial rule is accompanied by another well known integrative principle called *inverse effectiveness*, according to which the magnitude of the multisensory enhancement is inversely related to the effectiveness of the individual unimodal stimuli: combinations of weakly effective unisensory stimuli produce proportionally greater multisensory enhancement than more effective unisensory stimuli (Meredith and Stein, 1986b; Perrault Jr et al., 2003; Perrault Jr et al., 2005; Stanford et al., 2005; Stein and Meredith, 1993; Wallace et al., 1998). Overt behaviour responses have been found to follow these SC neurons' properties.

Many different studies have shown that the descending excitatory inputs from a specific association cortex (Anterior Ectosylvian Sulcus, AES) (Jiang et al., 2002, 2003) and a class of

membrane receptors (N-methyl-D-aspartate, *NMDA*, receptors) (Binns and Salt, 1996) strongly influence the multisensory integration capabilities of the SC multisensory neurons.

These principles provide a highly influential framework to explain the multisensory interactions observed in SC neurons and to predict how some characteristics of the stimuli (such as spatial location, time of occurrence and intensity) influence the integrative response, both at neuronal and behavioural level. However the underlying neural mechanisms and network architecture involved in these processes are not clearly identified yet. These properties of multisensory integration in the SC may depend both on the non-linear characteristics of neurons and on the particular arrangement of the neural circuitries which process the signals. In this regard, mathematical models inspired by physiological principles may play a critical role in providing a deeper understanding of the mechanisms which participate in multisensory integration, their possible importance in the origin of different phenomena, and the topology of the neural connections among different brain areas.

In this work we present some neural network models, based on neurophysiologically plausible mechanisms and on phenomenological results. The first (**chapter 2.1**) is a simple model of SC neurons focused on trying to investigate neuronal mechanisms underlying multisensory integration, with particular attention on the effect of the non-linear behaviour of SC neurons, the spatial information provided by different cortical inputs, and the relationships among neurons with different receptive fields. Moreover the model has been used to investigate which parameter modifications are able to explain the variability of multisensory neuron responses observed in-vivo.

The second model (**chapter 2.2**) goes over the physiological simplifications of the first one, including both cortical and non-cortical inputs to SC, to study and describe the effect of cortical deactivation on multisensory integration, and the role of NMDA receptors in driving SC behaviour.

CHAPTER 2.1. MULTISENSORY INTEGRATION IN THE SUPERIOR COLLICULUS: A NEURAL NETWORK MODEL

INTRODUCTION

Multisensory neurons the SC receive converging visual, auditory and somatosensory afferents from various subcortical and extraprimary cortical sources (Edwards et al., 1979; Huerta and Harting, 1984; Wallace et al., 1993). Responses of such neurons to a combination of stimuli of different sensory modalities can differ significantly from those evoked by any of their unisensory inputs in a way to facilitate the role of the SC in controlling attentive and orientation behaviour (Frens et al., 1995; Schroger and Widmann, 1998; Stein et al., 1989).

A consistent amount of physiological data has been gathered from studies performed in order to identify the principles ruling multisensory integration in SC neurons: the coincidence of sensory information in space, and the temporal coincidence of converging inputs. Modality-specific RFs of SC neurons are in spatial register, i.e. they represent similar regions of sensory space, so that they overlap. Both this receptive field alignment and the time coincidence of different stimuli are critical for normal multisensory processes in SC (Meredith and Stein, 1996). According to the temporal domain and the spatial organization, a combination of two different sensory stimuli (e.g., auditory and visual) could elicit a neuron's response which is significantly grater than that evoked by the most effective of the two unimodal inputs individually (*multisensory enhancement*) (Bell et al., 2001; Kadunce et al., 2001; Meredith and Stein, 1986b; Meredith and Stein, 1996; Wallace et al., 1998), when they fall within the overlapping receptive fields of the same neuron at the same time (both stimuli are likely produce from the same event). In some neurons, the response enhancement may even exceed the sum of their individual unimodal responses (*superadditivity*) (Kadunce et al.,

2001; Meredith and Stein, 1986b; Perrault Jr et al., 2003; Perrault Jr et al., 2005; Wallace et al., 1998). On the other hand, when the two stimuli are presented at different locations or at different time (i.e. they likely derive from different events), two alternative results can be observed: either no interaction occurs or the neuron's response to the within-field stimulus is considerably depressed (*multisensory depression*) (Kadunce et al., 1997; Kadunce et al., 2001; Meredith and Stein, 1986a; Wallace et al., 1998). Multisensory depression is assumed to derive from the presence of an inhibitory region which surrounds the excitatory receptive field (Kadunce et al., 1997). Globally, these properties of cross-modal integration are known as the *spatial rule*.

The spatial rule is accompanied by another well known integrative principle called *inverse effectiveness* (Meredith and Stein, 1986b; Perrault Jr et al., 2003; Perrault Jr et al., 2005; Stanford et al., 2005; Stein and Meredith, 1993; Wallace et al., 1998). Inverse effectiveness has functional sense in behavioural situations: the probability to detect a weak or ambiguous stimulus benefits more from multisensory enhancement than a high-intensity stimulus which is easily detected by a single modality alone (Jiang et al., 2002; Stein et al., 1989; Wilkinson et al., 1996).

These principles provide a highly influential framework to explain the multisensory interactions observed in SC neurons and to predict how some characteristics of the stimuli (such as spatial location, time of occurrence and intensity) influence the integrative response, both at neuronal and behavioural level.

A fundamental contribution to investigate the mechanisms by which modality-specific inputs achieve integration in SC can be obtained by using neural network models and computer simulations. Anastasio et al. (Anastasio, Patton, & Belkacem-Boussaid, 2000; Patton, Belkacem-Boussaid, & Anastasio, 2002) developed some models based on information theory, in which a SC neuron uses the Bayes' rule to compute the probability that a target is present in its receptive field. A similar approach was used by Colonius and Diederich (Colonius & Diederich, 2004). These models reproduce the major features of multisensory enhancement, suggesting that SC neurons can operate under the Bayes' rule. However, they did not provide insights into the neurobiological

mechanisms that may implement the needed computation. In subsequent papers, Anastasio and Patton (Anastasio & Patton, 2003; Patton & Anastasio, 2003) realized simple perceptron models implementing the Bayes'rule, to address issues of neurobiological mechanisms. With these models, the authors were able to account for crossmodal enhancement and, by using Hebbian learning mechanisms, to explain the existence of both multimodal and unimodal neurons. Furthermore, the neural implementation also accounted for within-modality suppression, given only the added hypothesis that inputs of the same modality have more spontaneous covariance than those of different modalities.

Although the latter models were somewhat inspired by physiological mechanisms, there are still some aspects which may benefit from a computational approach via neural networks and computer simulation techniques. In particular, previous models do not consider the spatial properties of multisensory integration, or the effect of the position and intensity of different inputs on the SC response. Moreover, these models do not include a competition among neurons with different receptive field to detect a target.

The aim of this work is to develop a neural network model, based on neurobiologically plausible mechanisms, which is able to reproduce and explain several in-vivo results in anesthetized animals. For simplicity, only the audio-visual integration is considered (i.e. the presence of somatosensory inputs is not considered). The model includes three neural networks, which communicate via synaptic connections. Two of them are unimodal and represent neurons coding visual and auditory stimuli, respectively; these networks represent unisensory cortical areas projecting afferents to SC, in particular two subregions of AES cortex (the Anterior Ectosylvian Visual area, AEV, and the Field Anterior Ectosylvian Sulcus, FAES); a downstream network, representing multimodal neurons in the SC, receives information from the upstream unisensory networks via feedforward synapses and integrates these information to produce the final response. Furthermore, neurons in each network are interconnected via lateral synapses. The activity of each neuron is described by a sigmoidal relationship with a threshold and a saturation. The present paper reports a quantitative

description of the mathematical model, including all equations. Through computer simulation, we show that the model is able to explain the main properties of multimodal neurons (cross-modality enhancement; inverse effectiveness; within-modality and cross-modality suppression). Furthermore, through a sensitivity analysis on the strength of SC synapses, we show that the existence of different typologies of SC neurons, described in the literature, can be reproduced quite well through changes in the synaptic strength. Finally, the possible role of feedback modulation on certain physio-pathological behaviour (such as facilitation or ventriloquism) is investigated, and its possible involvement in cognitive science discussed.

METHOD

In this section we will first describe the general structure of the model. Then all equations are presented and justified. Finally, parameter assignment is discussed, on the basis of previous neurophysiological data.

General model structure

– The model is composed of three areas. Elements of each area are organized in *NxM* dimension matrices, so that the structure keeps a spatial and geometrical similarity with the external world: neurons of each area respond only to stimuli coming from a limited zone of the space (see Fig. 1). Neurons normally are in a silent state (or exhibit just a mild basal activity) and can be activated if stimulated by a sufficiently strong input. Furthermore, each neuron exhibits a sigmoidal relationship (with lower threshold and upper saturation) and a first order dynamics (with a given time constant). The two upstream areas are unimodal, and respond to auditory and visual stimuli, respectively. A third downstream area represents neurons in the SC responsible for multisensory integration. These three areas have a topological organisation, i.e., proximal neurons respond to stimuli in proximal position of space.

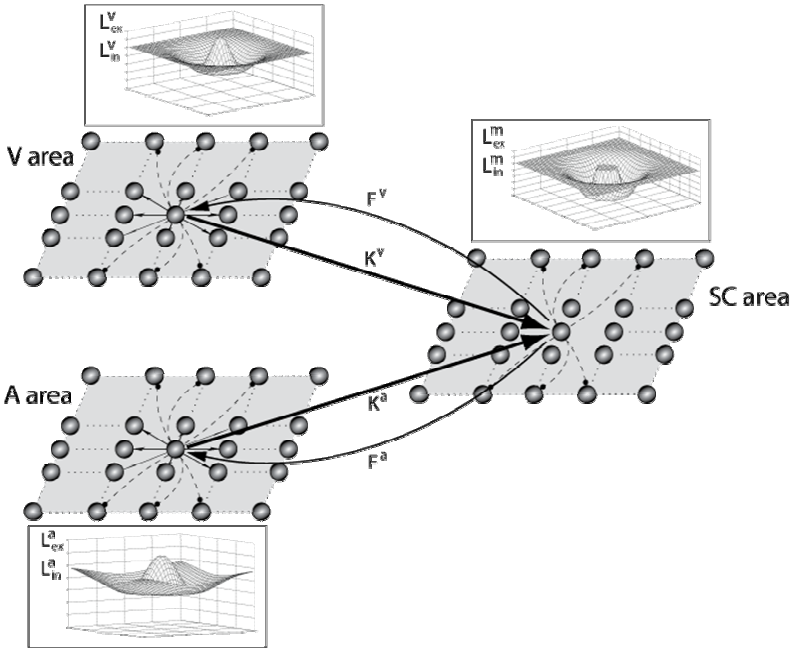


Figure 1 – Schematic diagram describing the general structure of the network. Each *grey circle* represents a neuron. Neurons are organized into 3 distinct areas of 40x40 elements (*V*, visual, *A*, auditory and *SC*, multimodal in the superior colliculus). Each neuron of these areas is connected with other elements in the same area via lateral excitatory and inhibitory intra-area synapses (*arrows* L_{ex} and L_{in} within the area). Neurons of the unimodal areas send feedforward excitatory inter-area synapses to multimodal neurons in the superior colliculus area located at the same position (*arrows k*): Multimodal neurons, in turn, send excitatory feedback inter-area connections to neurons of the unisensory areas (arrows *F;* see text for details).

– Each element of the unisensory areas has its own receptive field (*RF*) that can be partially superimposed on that of the other elements of the same area. The elements of the same unisensory area interact via lateral synapses, which can be both excitatory and inhibitory. These synapses are arranged according to a Mexican hat disposition (i.e., a circular excitatory region surrounded by a larger inhibitory annulus).

– The elements of the multisensory area in the superior colliculus receive inputs from the two neurons in the upstream areas (visual and auditory) whose RFs are located in the same spatial position. Moreover, elements in the SC are connected by lateral synapses, which also have a Mexican hat disposition.

– The multimodal neurons in the SC send a feedback excitatory input to the unimodal neurons whose RFs are located in the same spatial position; in this way, detection of a multimodal stimulus may help reinforcement of the unisensory stimuli in the upstream areas.

Mathematical description

In the following sections, quantities which refer to neurons in the auditory, visual or multisensory areas will be denoted with the superscripts *a*, *v* and *m*, respectively. The spatial

position of individual neurons will be described by the subscripts *ij* or *hk* (*i*, *h* = 1, 2, ..., *N* ; *j*, *k* = 1, 2, …, *M*).

The receptive fields of unisensory areas

In the present version we assume that each area is composed by 40x40 neurons (i.e.: *N* = 40; *M* = 40). Such a limited number of neurons was chosen to reduce the computational complexity for computer implementation. Neurons in each area differ in the position of their receptive field of 2.25 deg. Hence, each area covers 90 deg in the visual, acoustic or multisensory space. In the following, we will denote with x_i and y_j the center of the RF of a generic neuron *ij*. Hence, we can write:

$x_i = 2.25 \cdot i$ deg ($i = 1, 2, \ldots, 40$)

$y_j = 2.25 \cdot j$ deg ($j = 1, 2, \ldots, 40$)

The receptive field (say $R_{ij}^{s}(x, y)$) of neuron *ij* in the unisensory area *s* is described with a gaussian function. Hence,

$$R_{ij}^{s}(x,y) = R_0^s \cdot e^{-\frac{\left[(x_i - x)^2 + (y_j - y)^2\right]}{2\cdot(\sigma_R^s)^2}} \qquad s = a, v; \tag{1}$$

where the symbols *x,y* represent a generic coordinate in space, σ_R^s is the standard deviation of the Gaussian function (three standard deviations approximately cover the overall RF) and R_0^s is a parameter which sets the strength of the response.

According to Eq. (1), a stimulus presented at the position *x,y* excites not only the neuron centered in that zone, but also the proximal neurons whose receptive fields cover such position.

The input r_{ij}^{s}, that reaches the neuron *ij* in presence of a stimulus, is computed as the inner product of the stimulus and the receptive field. We can write:

$$r_{ij}^{s}(t) = \iint_{x,y} R_{ij}^{s}(x,y) \cdot i^{s}(x,y,t)\,dx\,dy \cong \sum_{x}\sum_{y} R_{ij}^{s}(x,y) \cdot i^{s}(x,y,t)\Delta x \Delta y \tag{2}$$

where, $i^s(x, y, t)$ is the external sensory stimulus presented at the coordinate *x*,*y* and at time *t*, and the right hand member of Eq. (2) means that the integral has been computed with the histogram rule (in this work, Δx=Δy= 2.25°).

The activity in the unisensory areas

Unisensory neurons can be stimulated not only by external inputs, but also through the connections with other elements in the same area. The input that a unisensory neuron gets from other elements of the same area is represented by the quantity l_{ij}^s, defined as:

$$l_{ij}^s(t) = \sum_{\mathrm{h,k}} L_{ij,hk}^s \cdot z_{hk}^s(\mathrm{t}) \qquad s = a, v; \tag{3}$$

where $z_{hk}^s(\mathrm{t})$ is the *hk*-neuron's activity (described below) and $L_{ij,hk}^s$ is the strength of the synaptic connection from the pre-synaptic neuron at the position *hk* to the post-synaptic neuron at the position *ij*.These synapses are symmetrical and arranged according to a "Mexican hat" function:

$$L_{ij,hk}^s = L_{ex}^s \cdot e^{-\frac{\left[d_x^2 + d_y^2\right]}{2\cdot\left(\sigma_{ex}^s\right)^2}} - L_{in}^s \cdot e^{-\frac{\left[d_x^2 + d_y^2\right]}{2\cdot\left(\sigma_{in}^s\right)^2}}. \qquad s = a, v \tag{4}$$

In this equation, L_{ex}^s and σ_{ex}^s define the excitatory Gaussian function, while L_{in}^s and σ_{in}^s the inhibitory one, and d_x, d_y represent the distance between the pre-synaptic and post-synaptic neurons in the horizontal and vertical coordinates. To avoid undesired border effects, synapses have been realized by a circular structure so that every neuron of each area receives the same number of side connections. This is realized assuming the following expressions for distances:

$$d_x = \begin{cases} |i-h| & \text{if } |i-h| \le N/2 \\ N - |i-h| & \text{if } |i-h| > N/2 \end{cases} \tag{5}$$

$$d_y = \begin{cases} |j-k| & \text{if } |j-k| \le M/2 \\ M - |j-k| & \text{if } |j-k| > M/2 \end{cases}$$

A further input to unisensory neurons is induced by the feedback, f_{ij}^{s}, from the Superior Colliculus. Such connections exclusively link neurons whose RFs are placed in the same *ij*-position in the Colliculus and the unisensory area. We have

$$f_{ij}^{s}(t) = F^{s} \cdot z_{ij}^{m}(t) \qquad s = a,v \tag{6}$$

where z_{ij}^{m} is the activity of the neuron in the SC at the same position *ij* and F^{s} is the strength of feedback synaptic connection.

According to the previous description, the total input (say $u_{ij}^{s}(t)$) received by a unisensory neuron at position *ij* is computed as follows,

$$u_{ij}^{s}(t) = r_{ij}^{s}(t) + l_{ij}^{s}(t) + f_{ij}^{s}(t) \; ; \qquad s = a,v \tag{7}$$

This is the sum of three components: r_{ij}^{s}, that represents the external sensory input; l_{ij}^{s}, coming from the intra-area synapses; and f_{ij}^{s}, the contribution of the feedback coming from the Superior Colliculus.

Finally, neuron activity is computed from its input, trough a static sigmoidal relationship and a first-order dynamics. This is described via the following differential equation:

$$\tau_{s} \cdot \frac{d}{dt} z_{ij}^{s}(t) = -z_{ij}^{s}(t) + \varphi\left(u_{ij}^{s}(t)\right) \; ; \qquad s = a,v \tag{8}$$

τ_{s} is the time constant, which determines the speed of the answer to the stimulus, and $\varphi\left(u^{s}(t)\right)$ is a sigmoidal function. The latter is described by the following equation:

$$\varphi\left(u^{s}(t)\right) = \frac{1}{1 + e^{-\left(u^{s}(t) - \vartheta^{s}\right) p^{s}}} \; ; \qquad \mathrm{s = a, v.} \tag{9}$$

where ϑ^{s} defines input value at which neuron activity is half the maximum (central point) and p^{s} sets the slope at the central point.

Such a function identifies three regions of work, depending on the intensity of the input: the sub-threshold behaviour of a neuron, a linear region (around ϑ^{s}), and a saturation region.

According to the previous equation, the maximal neuron activity is conventionally set at 1 (i.e., all neuron activities are normalized to the maximum).

The activity in the multisensory area

As said before, neurons of this area receive inputs from neurons in the two unisensory areas whose RFs are located in the same position, *ij*. This choice has been adopted since, according to experimental data, the auditory and visual RFs of a multisensory neuron are in spatial register (Kadunce et al. 2001; Meredith and Stein 1996), i.e., they represent similar regions in space. Furthermore, neurons in the superior colliculus also receive lateral synapses from other elements in the same area. Hence, the overall input, $u_{ij}^{m}(t)$, to a multisensory neuron can be computed as follows

$$u_{ij}^{m}(t) = k^{a} \cdot z_{ij}^{a}(t) + k^{v} \cdot z_{ij}^{v}(t) + l_{ij}^{m}(t) \tag{10}$$

k^{a} and k^{v} represent the intensities of the synaptic connections which link the acoustic and the visual areas with the superior colliculus. $z_{ij}^{a}(t)$ and $z_{ij}^{v}(t)$ are the activities of the acoustic and visual neurons in the *ij*-position (computed through Eq. (8)), and $l_{ij}^{m}(t)$ is the input due to the synapses between the elements of the Colliculus. The latter term is computed as follows

$$l_{ij}^{m}(t) = \sum_{h,k} L_{ij,hk}^{m} \cdot z_{hk}^{m}(t) \,. \tag{11}$$

We assumed that lateral synapses in the multisensory area have a Mexican hat disposition. The equation that outlines the synaptic links between colliculus elements is:

$$L_{ij,hk}^{m} = L_{ex}^{\mathrm{m}} \cdot e^{-\frac{\left[d_x^{\,2}+d_y^{\,2}\right]}{2\cdot\left(\sigma_{ex}^{m}\right)^2}} - L_{in}^{\mathrm{m}} \cdot e^{-\frac{\left[d_x^{\,2}+d_y^{\,2}\right]}{2\cdot\left(\sigma_{in}^{m}\right)^2}} \tag{12}$$

where the distances d_x and d_y have the same expression as in Eq. (5), and parameters L_{ex}^{m}, L_{in}^{m}, σ_{ex}^{m} and σ_{in}^{m} set the strength and spatial disposition of lateral synapses.

Then, the activity of a multisensory neuron is computed from its input by using equations similar to Eqs. (8) and (9).

Parameter assignment

Table 1 – Basal value of model parameters

Visual Area	
$N=M$	40
σ_R^v	1.5 (3.37°)
R_0^v	1
L_{ex}^v	1.6
σ_{ex}^v	3.5 (7.87°)
L_{in}^v	1.23
σ_{in}^v	6.3 (14.17°)
F^v	1
T_v	3 ms
ϕ^v	3
p^v	0.3
k^v	7
Acoustic Area	
$N=M$	40
σ_R^a	2 (4.5°)
R_0^a	1
L_{ex}^a	1
σ_{ex}^a	5.3 (11.92°)
L_{in}^a	0.8
σ_{in}^a	11.8 (26.55°)
F^a	1
T_a	3 ms
ϕ^a	3
p^a	0.3
k^a	6
Superior Colliculus Area	
$N=M$	40
L_{ex}^m	3.8
σ_{ex}^m	3.5 (8°)
L_{in}^m	3.3
σ_{in}^m	6.2 (13.95°)
T_m	3 ms
ϕ^m	3

The value of all model parameters is shown in Table 1. These parameters have been assigned starting from data in the literature according to the main criteria summarized below.

Receptive fields: σ_R^v and σ_R^a have been given so that the receptive fields of the visual neurons are approximately 10-15 deg in diameter, and those of acoustic neurons approximately 20-25 deg in diameter, according to data reported in (Kadunce et al., 2001). R_0^v and R_0^a are set to 1, to establish a scale for the inputs generated by the external stimuli.

Unimodal areas: Parameters which establish the extension and the strength of lateral synapses in the unimodal areas (i.e., L_{ex}^v L_{in}^v, σ_{ex}^v and σ_{in}^v for the visual area, and L_{ex}^a, L_{in}^a σ_{ex}^a and σ_{in}^a for the acoustic area) have been assigned to simultaneously satisfy several criteria: (1) the presence of an external stimulus produces an activation bubble of neurons which approximately coincide with the dimension of the receptive field; (2) according to data reported in Kadunce et al. (Kadunce et al. 1997) we

assumed that the surrounding inhibitory area is much larger than the activation bubble; (3) inhibition strength must be strong enough to avoid instability, i.e., an uncontrolled excitation which propagates to the overall area.

Superior colliculus area: Parameters which establish the extension and the strength of lateral synapses in the SC areas (i.e., L_{ex}^{m} L_{in}^{m}, σ_{ex}^{m} and σ_{in}^{m}) have been assigned to warrant cross-modal and within-modal integration in agreement with data reported in Kadunce et al. (Kadunce et al., 1997; Kadunce et al., 2001).

Parameters of the individual neurons (sigmoidal relationships and time constants): for the sake of simplicity, these parameters have been chosen equal for all neurons, independently of the respective area. The central abscissa, ϑ^{S}, has been assigned to have negligible neuron activity in basal condition (i.e., when the input is zero). The slope of the sigmoidal relationship, p^{S}, has been assigned to have a smooth transition from silence to saturation in response to unimodal and cross-modal inputs (Perrault Jr et al., 2005). The time constant agrees with values (a few milliseconds) normally used in deterministic mean-field equations (Ben-Yishai et al., 1995). In particular, this value can be chosen significant smaller than the membrane time constant (Treves, 1993).

Connections between the three areas The parameters of feedforward connections from the unisensory areas to the superior colliculus (i.e., k^{v} and k^{a}) have been set in order to: a) two visual and auditory stimuli of the same intensity produce similar effects on multimodal neurons, when acting separately. In this way, two stimuli of different modality can be directly compared in term of their effect. Since the auditory neurons have a greater RF, this criterion implies the use of greater feedforward synapses from the visual area; (b) have a significant multisensory enhancement; (c) have a greater dynamical range of multisensory neurons in response to cross-modal stimuli, compared with unimodal stimuli (i.e., a single stimulus cannot lead the SC neuron to saturation; Perrault et al. 2005). Furthermore, in this work we assumed that the effect of a visual stimulus on the SC neuron is moderately greater compared with the effect of an auditory stimulus. For the sake

of simplicity, the feedback connections from the SC neuron to the upstream visual and auditory neurons (i.e., parameters F^a and F^v) have been set to small values in this work, assuming that this effect is normally modest. These parameters may be the subject of a sensitivity analysis in future works.

All synapses in the multimodal area (L^m_{ex}, L^m_{in}, k^v, k^a, F^a and F^v) have then been the subject of a sensitivity analysis, as shown in section results.

RESULTS

Steady-state network behaviour

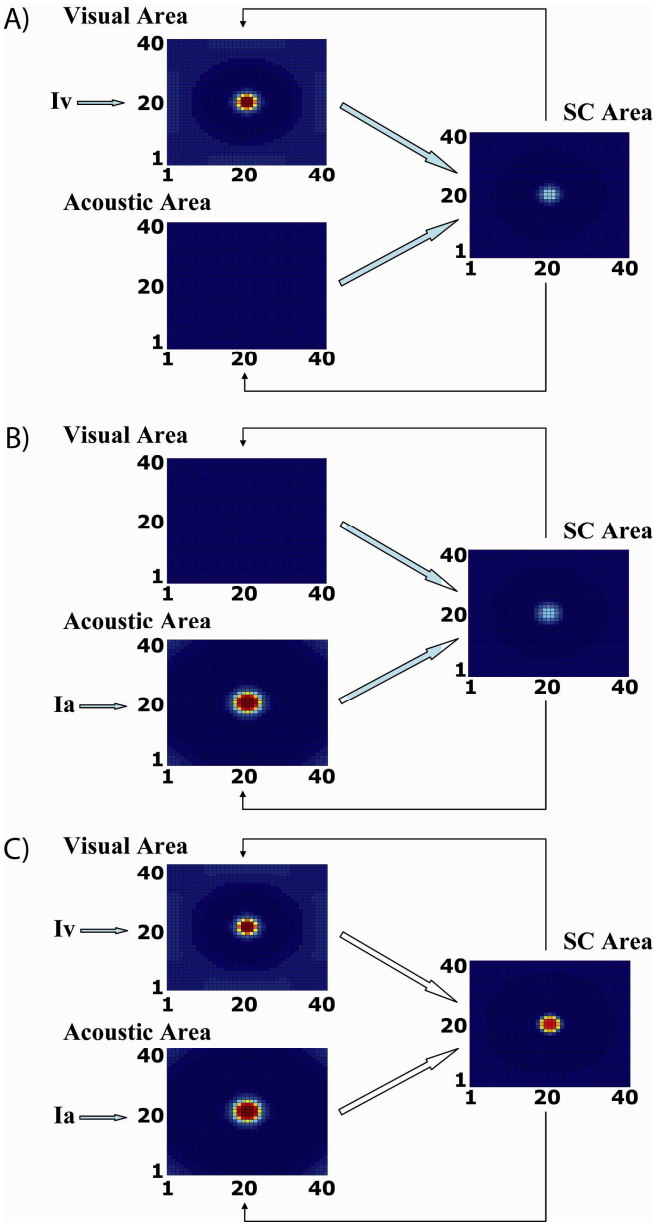


Figure 2 – Examples of network's responses to stimuli with different sensory modality. The intensity of each stimulus is the same (i^v=i^a= 22). Each pixel represents a neuron. The width of the activation bubble in the unisensory areas depends on the dimension of the modality-specific receptive field. The colour of each pixel is proportional to the corresponding neuron's activity z_{ij} (blue=0, red=1). The *arrows* indicate the inter-area connections (both feedforward and feedback). (a) The model is excited by an external visual stimulus, located at position 20,20. This stimulus produces the activation of a bubble of neurons both in the unisensory (visual) area and in the superior colliculus. (b) The network receives an acoustic stimulus at position 20,20 and an activation bubble occurs in the acoustic area and in the superior colliculus. (c) The model receives a multisensory input, represented by two cross-modal stimuli coming from the same location in space. Both unisensory areas present a bubble of activated-neurons; the superior colliculus shows an increased activation that is consistent with the phenomenon of multisensory enhancement.

All the results, except those reported in Fig. 18, have been calculated in steady-state conditions, by giving one or two simultaneous step stimuli to the model, with a duration 100

ms, starting from the basal resting condition (i.e., no stimulation). This time is sufficient for the model to reach a steady-state value. In particular, when computing model responses to different stimuli located at different points in the RF we did not use a moving stimulus, but we placed the stimulus at the assigned spatial position and waited enough time (100 ms), before putting it at the next position.

Figure 2 shows an example of the activation bubbles in the three areas, simulated by the model in presence of a unisensory visual input (upper panel), a unisensory auditory input (middle panel) and two simultaneous visual and auditory stimuli (bottom panel).

Figure 3 shows the response of a multisensory SC neuron evoked by a stimulus (or a pair of stimuli) located at different positions within and outside the receptive field. A bell shape is obtained, with the maximum located at the centre of the RFs. When the stimulus is placed at the border of the RF, neuron activity decreases down to ~9% of the bell peak. It is worth noting the strong enhancement of the multisensory response compared with the response to individual stimuli.

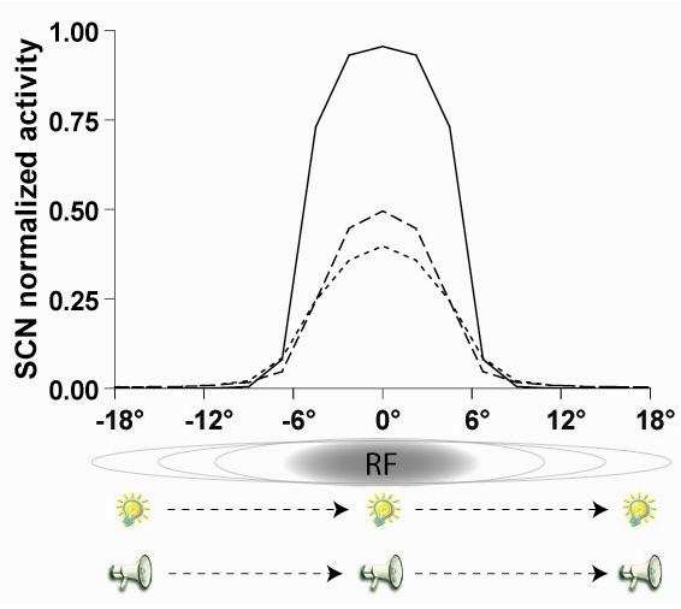


Figure 3 – The simulated normalized activity of a superior colliculus Neuron (*SCN*) evoked by an input whose application point shifts through the receptive field. The network is stimulated, first by an acoustic stimulus (*dotted line*), then by a visual stimulus (*dashed line*) and, finally, by two paired cross-modal stimuli (*solid line*) with the same intensity (i^v=i^a=22). Results have been calculated with the stimulus in the assigned spatial position and waiting for the network to reach the final steady-state condition. It is worth nothing the strong multisensory enhancement.

Results displayed in Fig. 3 resemble those experimentally obtained in SC cells by changing the stimulus locations in modality-specific and multisensory tests (see Fig. 5 in Kadunce et al. 2001). In these tests, a bell shape was obtained with the best point (both for unisensory and multisensory stimulation) approximately positioned at the center of the overlapping region between the auditory

and visual RFs. Multisensory response always overcame unisensory response. Neuron activity decreased down to approximately the 10% of the peak activity when the stimulus was applied at the border of the overlapping region, in agreement with model predictions.

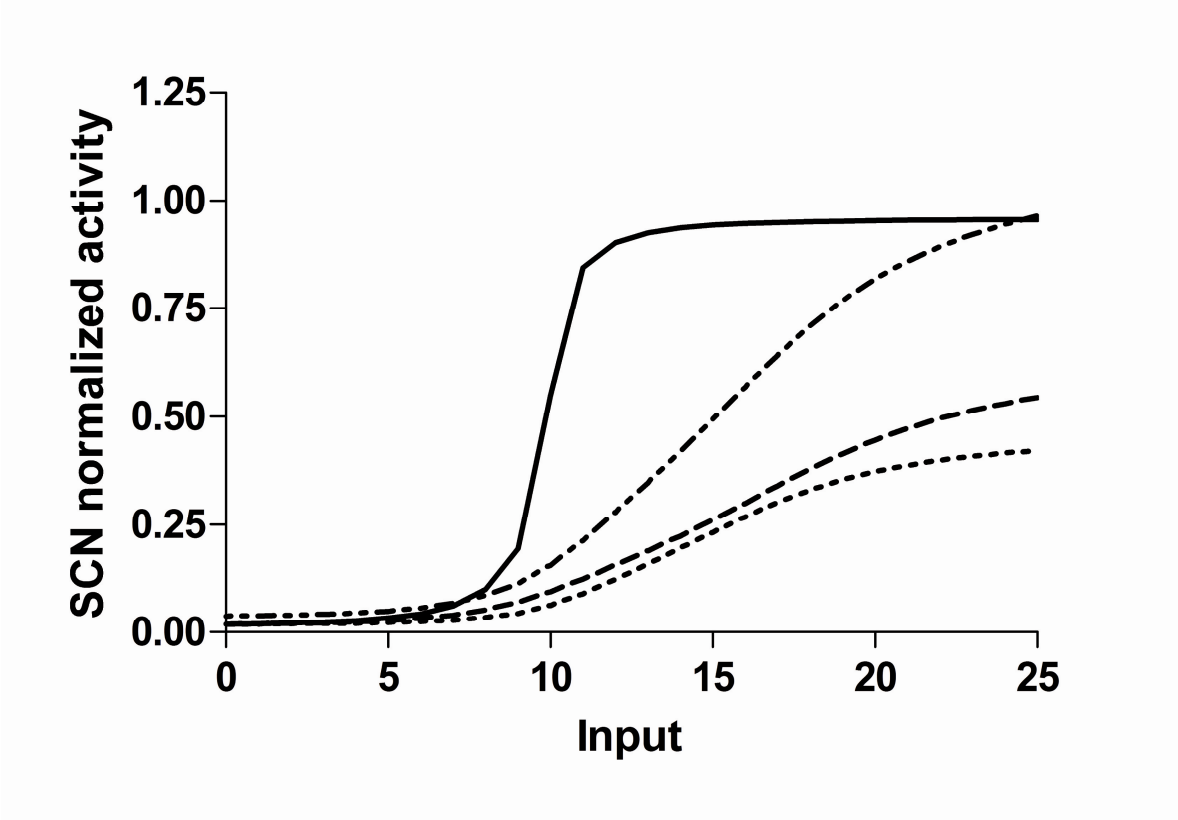


Figure 4 – Analysis of the response of a superior colliculus Neuron (SCN) to unimodal and crossmodal stimuli. The responses were assessed stimulating the model with an acoustic (*dotted line*), a visual (*dashed line*) and two paired multisensory (*solid line*) stimuli with different levels of intensity. Results have been computed in steady-state conditions. By way of comparison, the sum of the two unisensory responses (*dotted and dashed line*) is also presented in this figure. The stimulus was presented at the center of the RF of the observed SC neuron. The intensities of the stimuli of each sensory modality are shown in the x-axis, where we used normalized input. This means that the same value has been assigned to unisensory stimuli which produce comparable responses in the SC neurons. We chose low-intensity inputs (*i*=0 – 10) to reproduce the under-threshold behaviours, and the high-intensity ones (*i*>20) to obtain the saturation levels.

The dynamical ranges of a SC neuron are reported in Fig. 4, for an auditory (dotted line), a visual (dashed line) and a multisensory (continuous line) stimulation. These responses have been obtained by using either a single stimulus (auditory or visual) of increasing strength or two paired stimuli (visual + auditory) located at the centre of the RF. The dynamical range is defined as the difference in neuron activity at saturation and at threshold. Furthermore, by way of comparison, the sum of the two unisensory responses is also presented in the same figure (dash-dotted line). Two aspects of these curves are of interest: first, the dynamical range to multisensory stimulation is much greater than that to a single stimulus. Second, the neuron exhibits a superadditive response at low values of

the input stimuli (just above threshold), while the response becomes additive/subadditive at high stimulation levels.

Perrault et al. (2005) used two different metrics to quantitatively evaluate multisensory integration. The first, i.e. the interactive index, is a measure of response increase induced by two cross-modal stimuli compared to a single stimulus, and is defined as follows:

$$\text{Interactive Index} = \left[\frac{Mr - Ur_{\max}}{Ur_{\max}} \right] \cdot 100;$$

where Mr (multisensory response) is the response evoked by the combined-modality stimulus, and $Ur_{\max}$ (unisensory response) is the response evoked by the most effective unisensory stimulus.

The second, named the multisensory contrast, is used to quantify additivity:

$$\text{Multisensory Contrast} = [(Mr + BA) - (Vr + Ar)];$$

where Mr, as described above, is the response evoked by the multisensory stimulus, BA is the basal activity, Vr and Ar are the responses evoked by a unisensory visual and an auditory stimulus, respectively.

Figure 5 displays the interactive index computed at different values of the input stimuli. In particular, the top panel shows the response of a SC neuron to an individual auditory or visual stimulus (ranging from above threshold to saturation) located at the center of the RF, and to a combination of both paired stimuli. The multisensory response is always higher than the individual responses. From these data, the interactive index has been computed (bottom panel). According to the principle of inverse effectiveness, this index decreases from about 500% in case of weak unisensory stimuli (just above the threshold) down to 60–50% in case of strong stimuli (input values = 30 and 40 in the figure). It is worth noting that for input values above 25 neuron behaviour becomes subadditive (see Fig. 4). Values of interactive index obtained with the model fall in the ranges reported in the literature. Perrault et al. (2005) found values of interactive index between 1,000% and 300% in case of superadditivity, and between 100% and 20% in case of additivity/subadditivity. In Kadunce et al. (2001), the interactive index was 94% and 55% in case of

subadditivity, and above 200% in case of superadditivity. Wallace et al. (1996), reports an average enhancement of 122% (with values even up 270%) for poorly effective stimuli, and of 31% (range between –6% and 67%) for highly effective stimuli.

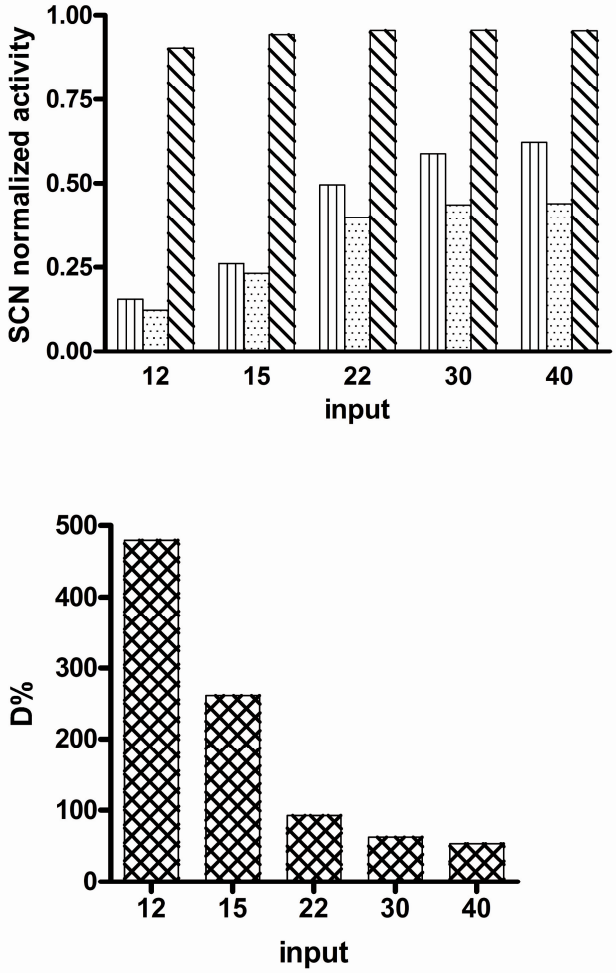


Figure 6 – Analysis of the multisensory contrast, defined as the difference between the activity of a SCN stimulated by two cross-modal stimuli coming from the same location in space and the sum of the two unisensory responses. In each simulation we have paired a constant visual input with an acoustic stimulus of different levels of intensity, to evaluate the inverse effectiveness principle. The intensities of the acoustic stimulus are presented in the x-axis. Results were obtained in steady-state conditions. *Upper panel* The constant visual stimulus was chosen close to the threshold for unisensory stimulation (i^v=12). In this case neuron's behaviour is "superadditive" for all values of the acoustic stimulus. *Bottom panel* The constant visual stimulus was chosen close to the saturation region for unisensory stimulation (i^v= 30). In this condition, neuron's behaviour changes from superadditive to additive/subadditive by increasing the second (auditory) stimulus.

Figure 5 – Analysis of the interactive index that underlines the phenomenon of the enhancement and the inverse effectiveness principle. *Upper panel* Value of the steady-state response of a SCN to an acoustic (*dotted bar*), visual (*vertical-lined bar*) and multisensory (*diagonal-lined bar*) stimulus, located at the center of the neuron's RF, at various intensities (x-axes), ranging from above threshold (*i*=12) to saturation (*i*>30). *Bottom panel* Interactive index (*D%*) computed from the data in the *upper panel*, as the per cent increase of the multisensory response compared to the maximum unisensory response.

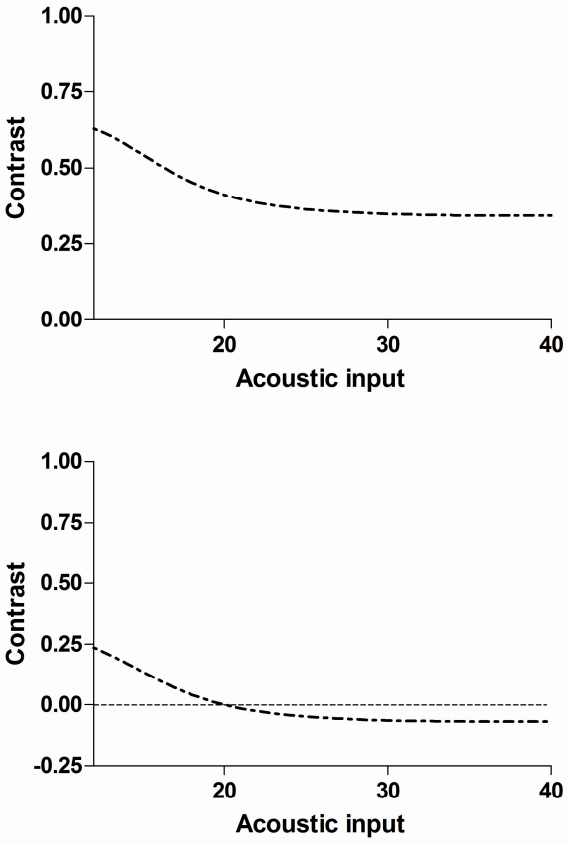


Figure 6 analyzes the inverse effectiveness principle by using the multisensory contrast. The upper panel shows contrast as a function of the auditory stimulus, computed when the visual stimulus is set to a small value (i^v = 12; just above threshold). Neuron behaviour is superadditive (i.e., contrast is greater than zero) for all values of the acoustic input. The bottom panel shows the same figure, computed using a high value for the visual input (i^v = 30; close to saturation). In this condition, neuron behaviour changes from superadditive to additive (contrast almost zero) or to subadditive by increasing the second (auditory) stimulus.

Figure 7 shows the role of the distance between two stimuli on the integrated response. The panels (a) show the case in which a visual stimulus is located at the center of the RF, and either a second visual stimulus (within-modality interaction) or a second auditory stimulus (cross-modality interaction) is moved from the center to the periphery. The activity in the visual unimodal area, at the central position of the RF, is shown in the left panel, while the middle panel shows activity in the corresponding multimodal neuron. The panels (b) consider a central auditory stimulus (again with within-modality and cross-modality interaction with a second stimulus). The activity in the unimodal auditory area at the central position of the RF is shown in the left panel, while activity in the same position of the SC is shown in the middle panel. The results confirm that a second stimulus of a different modality located within the receptive field causes significant cross-modal enhancement, whereas in the case of a within-modality stimulus the enhancement is mild (i.e., a second stimulus of the same modality, located inside the RF does not evoke a significantly greater response). The absence of significant within-modality enhancement agrees with experimental data (Stein and Meredith 1993; Wallace et al. 1996). If the second stimulus is moved away from the RF, one can observe significant within-modality suppression as well as significant cross-modality suppression. Within modality suppression is strong in both modalities (auditory and visual) leading to almost 70% reduction in the SC response. This model result is consistent with the study by Kadunce et al., reporting a magnitude of within-modality depression greater than 50% for the majority of visual and auditory responsive neurons. Moreover, in the same study, no significant

differences in the magnitude of response suppression were observed between within- and cross-modality suppression, as predicted by the model. Finally, the suppressive regions are quite large (25–30°) in the model, in accordance with physiological data (Kadunce et al. 1997).

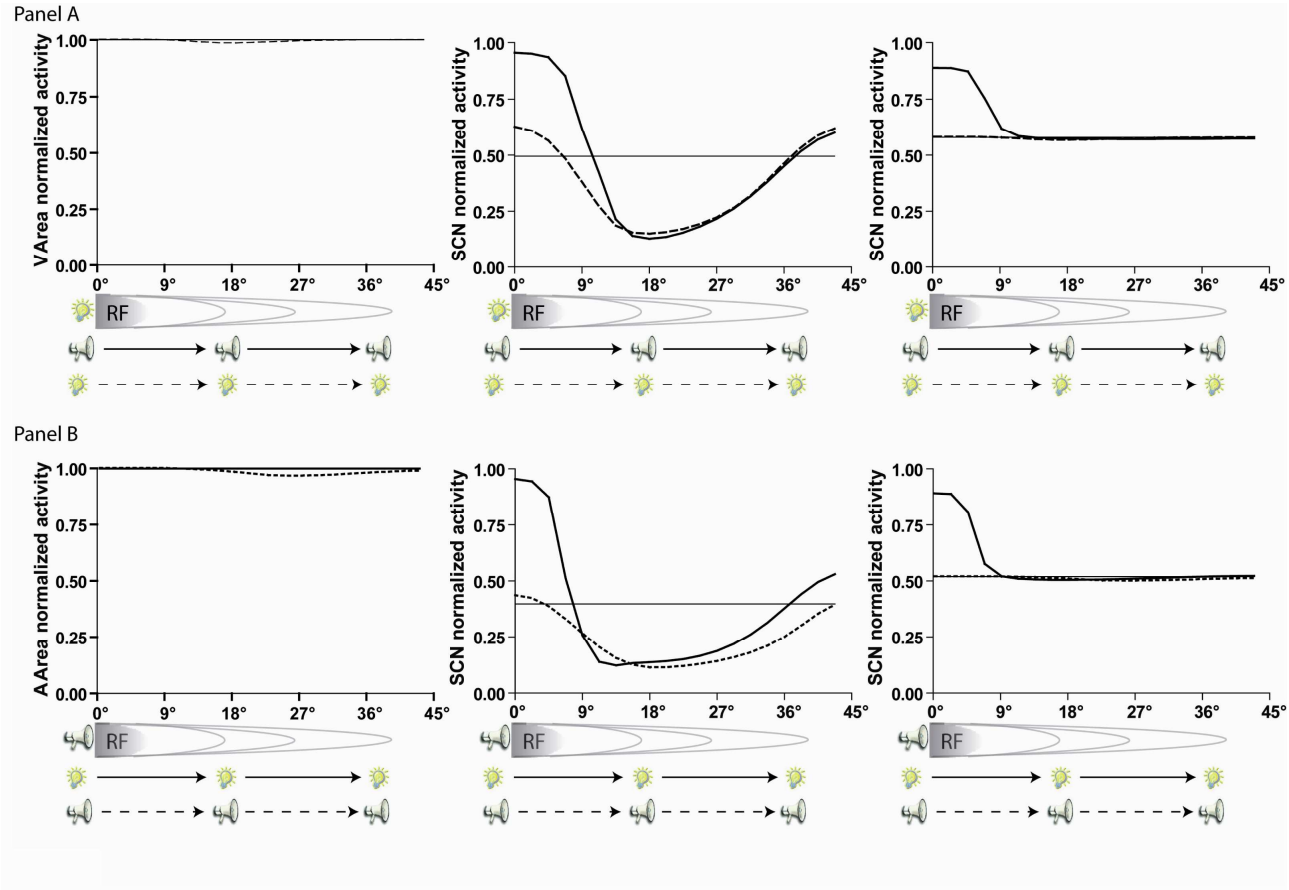


Figure 7 – Effect of the distance between two stimuli on the integrative response of the multisensory SC neurons. A first stimulus was located at the center of the RF of the observed neuron. A second stimulus, either of the same modality (*dashed line*) or of a different modality (*solid line*), is applied at different distances from the center of the RF. All simulations refer to steady-state conditions. The distance between the two stimuli is shown in the x-axis. The intensity of both stimuli was $i_a = i_v = 22$. *Panel A*: a visual stimulus is fixed at the center of the RF, and either a second visual stimulus (within-modality interaction) or an auditory stimulus (cross-modality interaction) is placed at a different distance from the center of the RF. *Panel B*: the same experiment described above, repeated by maintaining the auditory stimulus fixed, at the center of the RF, and paired either by a second auditory stimulus or a visual stimulus. Results in the *left column* show activity in the visual (a) or auditory (b) neuron in the unimodal area, located at the center of the RF. Results in the *middle column* show activity of the multimodal SC neuron at the same position, obtained with basal parameter values. Results in the *right column* show activity of the same multimodal SC neuron obtained by setting all lateral synapses in the multimodal area at zero. In this condition, in order to avoid excessive basal excitation for multimodal neurons, the central abscissa of the sigmoidal relationships (i.e., parameter ϑ in Eq. 8) has been increased from 3 to 6. Elimination of lateral multimodal synapses abolishes both cross-modality and within-modality suppression. In fact, as shown in the left column, within-modality suppression in the unimodal areas is quite negligible. The horizontal line in the middle and right columns represents SC neuron activity in response to a central stimulus alone.

In the model, cross-modality suppression is a consequence of long-range lateral inhibition within the SC area. By contrast, within-modality suppression may depend on the concurrent action of two mechanisms: lateral inhibition within the unimodal (visual and auditory) areas and lateral inhibition

within the SC area. In order to identify the specific role of these two mechanisms, the previous simulations have been repeated by assuming the absence of lateral interactions within the multimodal area (right columns). In these conditions, both cross-modality suppression and within-modality suppression completely disappear. This result suggests that, with the basal values of parameters (see Table 1), lateral unimodal synapses play a negligible role in generating within-modality suppression, and lateral inhibition among SC neurons is fully responsible for both within-modality and cross-modality suppression. This assumption is confirmed by the activity in unimodal areas shown in the left panels of Fig. 7, showing that lateral inhibition in unimodal areas, with basal parameter values, is negligible.

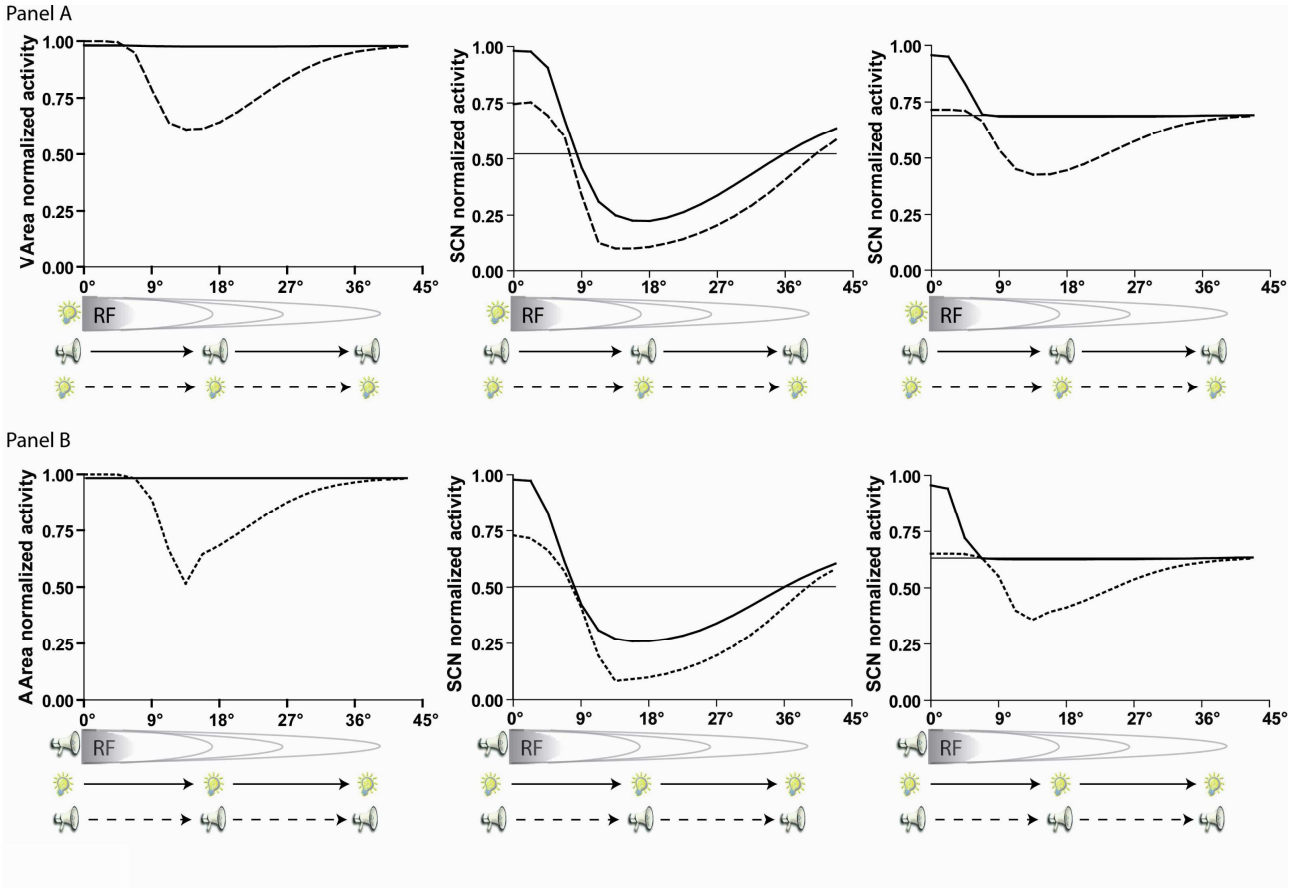


Figure 8 – Same simulations as in figure 7, but using a different set of parameters. In particular, in these simulations we increased the strength of lateral synapses in the unimodal areas, and the strength of feedforward synapses from unimodal areas to the multimodal area. The new parameters are: L_{ex}^{v} **= 5.4;** L_{in}^{v} **= 4.7;** L_{ex}^{a} **= 4.2;** L_{in}^{a} **= 3.5;** $\sigma_{ex}^{v} = \sigma_{ex}^{a}$ **= 2.8 (6.3°);** $\sigma_{in}^{v} = \sigma_{in}^{a}$ **=7.4 (16.65°);** k^{v} **= 9;** k^{a} **= 8. The intensity of both stimuli was** $i^{a} = i^{v} =$ **17. In these conditions, elimination of lateral synapses in the multi-modal area abolishes cross-modal suppression but does not abolish within-modality suppression (*right column*). In fact, as shown in the left column, within-modality suppression is already evident in the unimodal areas.**

Figure 8 displays the same simulations as in Fig. 7 assuming stronger lateral inhibition within unimodal areas (see the legend to the figure for the modified values of parameters). In this case, when all mechanisms are included, within-modality suppression results slightly greater than cross-modality suppression (middle column). If lateral interactions in SC area are set to zero (right column), cross-modality suppression vanishes, while within modality suppression still survives. In fact, as shown in the left panels, a significant within-modality suppression is now evident in the unimodal areas. Hence, a different balance between lateral inhibition in the unimodal areas and in the multimodal area may explain the existence of within-modality suppression without cross-modality suppression as documented in the literature (Kadunce et al. 1997).

Results in Figs. 7 and 8 were obtained by performing the same parameter changes in all locations of the network, i.e., all neurons in the network behave in the same manner. However, in a real set up, neurons with different behavior can be observed within the same network. In order to show this possibility, we repeated a few simulations concerning within- and cross-modality suppression, starting from the basal network (that is the network with the same parameters as in Table 1) but we modified the local value of synapses only at some specific positions.

The results are illustrated in Figure 9, where three exemplary positions are considered: (a) position 29.25°, 29.25°: we assume that all neurons in this position (multimodal and unimodal) receive the basal value of synapses. The SC neuron exhibits both within-modality and cross-modality suppression (top panel). (b) Position 58.5°, 45°: we assume that the multimodal neuron in this position does not receive any lateral connection within the SC, but the unimodal neurons at the same position receives basal values of synapses. The SC neuron exhibits neither within nor cross-modality suppression (middle panel). (c) Position 29.25°, 58.5°: the multimodal neuron does not receive any lateral connection within the SC, while the neuron in the acoustic unimodal area at the same position receives lateral inhibitory synapses stronger than basal (see legend for details). In this case, the SC neuron exhibits within-modality suppression (using acoustic stimulation) without

cross-modality suppression (bottom panel). These results demonstrate that the present model can easily be extended to build networks in which all types of behavior co-exist.

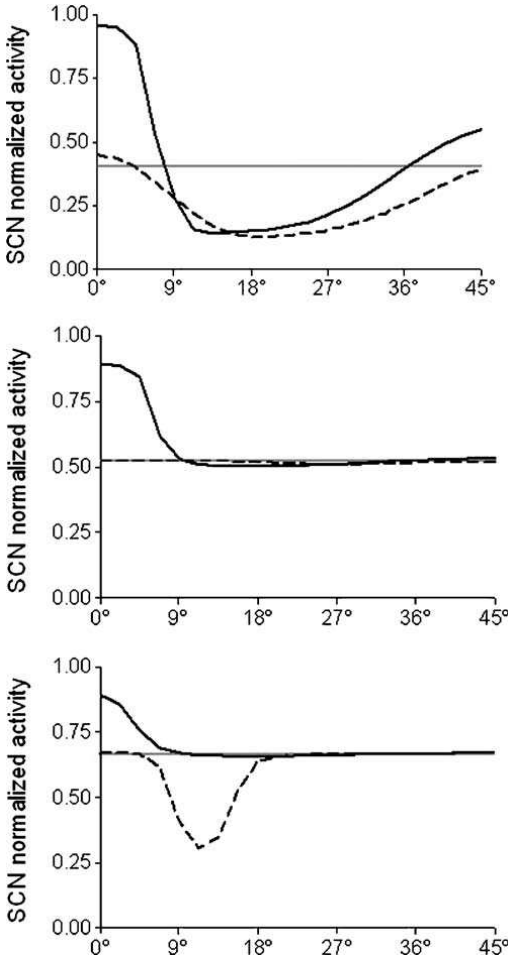


Figure 9 – Same simulations as in *panels* (b) of Figs. 7 and 8 (an acoustic stimulus is fixed at the center of the RF, and within- and cross-modality interactions are tested) but using a non-homogeneous network in which the local value of synapses can change from one position to another. Three exemplary positions are shown. *Upper panel* Position 29.25°, 29.25° — All neurons in this position receive basal synapses. An acoustic stimulus (i^a=22) is fixed at the centre of the RF at that position, and a either a second visual stimulus (i^v=22, cross-modality interaction, *solid line*) or a second auditory stimulus (i^a=22, within-modality interaction, *dashed line*) is placed at different distances from the centre of the RF. The SC neuron exhibits both within- and cross-modality suppression. *Middle panel* Position 58.5°, 45° — The SC neuron in this position does not receive lateral synapses (and its threshold parameter ϑ^n has been increased from 3 to 6), whereas the neurons in the unimodal areas receive basal lateral synapses. The same experiment as in *upper panel* has been repeated in this position. The SC neuron exhibits neither cross- nor within-modality suppression. *Bottom panel* Position 29.25°, 58.5° — The SC neuron in this position does not receive lateral synapses (and its threshold parameter ϑ^n has been increased from 3 to 6), whereas the neuron at the same position in the acoustic area receives lateral synapses stronger than basal (the parameters used are: L^a_{ex} = 4.2; L^a_{in} = 3.5; σ^a_{ex} =2.8 (6.3°); σ^a_{in} =7.4 (16.65°); ϑ^a=−10; k^a =8). Furthermore, in order to avoid excessive inhibition to this acoustic neuron, also its neighbouring neurons (three for each side) receive stronger lateral synapses (same parameters as above). An acoustic stimulus (i^a = 22) is fixed at the centre of the RF at that position, and a either a second visual stimulus (i^v = 9, cross-modality interaction, *solid line*) or a second auditory stimulus (i^a = 9, within-modality interaction, *dashed line*) is placed at different distances from the center of the RF. In these conditions, the SC neuron exhibits within-modality suppression without cross-modality suppression.

To go into more depth in the mechanisms underlying the SC multisensory integration and its response behaviour, we performed some more computer simulations. All results refer to steady-state conditions; to this end, the stimuli were maintained at their spatial position for a time interval (100 ms) sufficient for the exhaustion of all transient responses, before moving them to a different position.

Simulation with basal parameter values – A first set of simulations has been performed with all parameters at their basal values, as reported in Tab. 1.

Figure 10 describes the response to two visual stimuli of the same intensity. The stimuli are initially placed at a distance which avoids any interference; the distance is then progressively

reduced to zero. The first three rows (panel A) represent the activity in the visual, auditory and multimodal areas in five exemplary cases. The bottom panel (panel B) represents a profile of the response in the multimodal area.

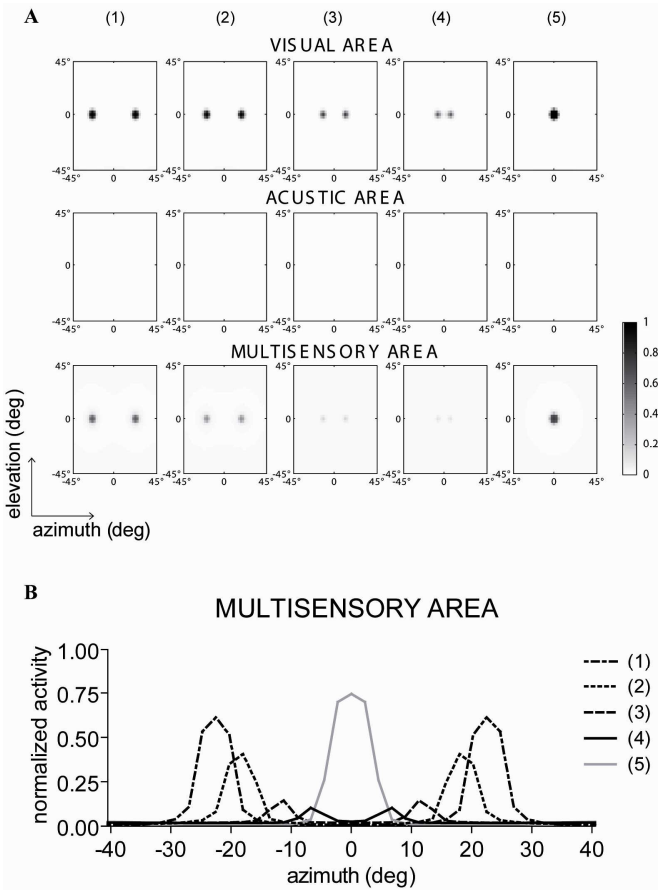
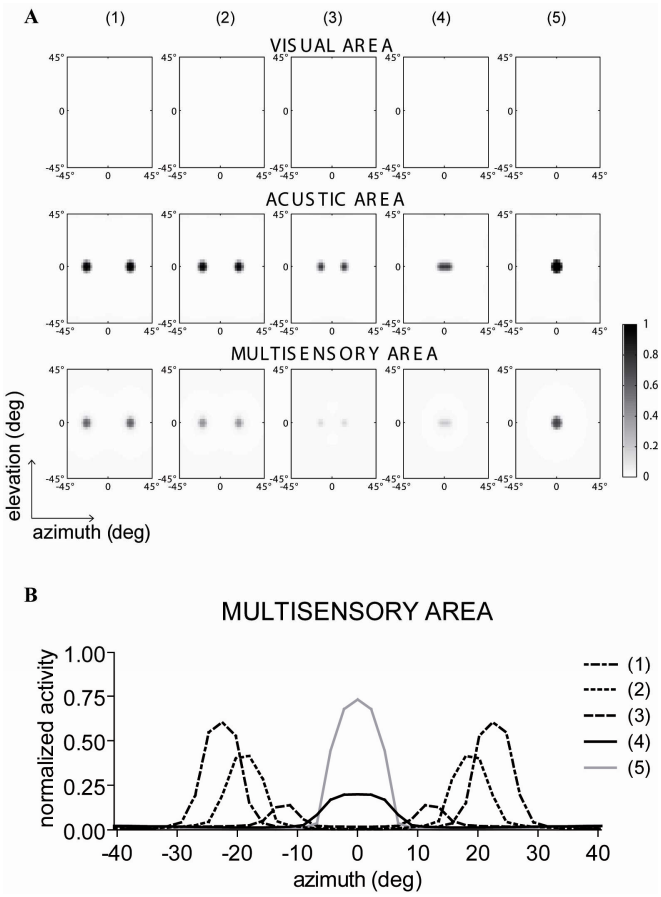


Figure 10 – Model response to two simultaneous visual stimuli (intensity i^v = 17) placed at five different positions in space. Stimuli are punctual and are always applied at the same elevation (vertical coordinate 0°) but at different distances along the azimuth. Stimuli location: (1) [−22.5° 0°] [22.5° 0°], azimuth distance = 45°; (2) [−18° 0°] [18° 0°], azimuth distance = 36°; (3) [−11.25° 0°] [11.25° 0°], azimuth distance = 22.5°; (4) [−6.75° 0°] [6.75° 0°], azimuth distance = 12.5°; (5) overlapped stimuli at [0° 0°]. Fig. 10A: Each column depicts the activity of all neurons in the three areas of the model (visual, auditory and multisensory) in steady state conditions, after application of two stimuli at a specific position (see column label). The darkness of the colour represents the magnitude of neuron activity. A strong within modality suppression in the multisensory area is evident in the third and fourth columns. Fig. 10B: profile showing the response of the neurons in the multisensory area having RF centre at the vertical coordinate 0° and at the azimuth coordinate from −40.5° to +40.5° (that is, 37 neurons positioned along the middle line of the vertical field) during the five simulations depicted in Fig. 10A. Each line pattern corresponds to one of the five simulations (see line label). Within modality suppression is evident (greater than 80%). Within modality enhancement of two superimposed stimuli (simulation 5) is mild.

Figure 11 describes results of similar simulations, but obtained by using two simultaneous auditory stimuli.

The following considerations can be drawn from the results, as shown in Figs. 10 and 11:

i) each stimulus induces an activation bubble in the corresponding unimodal area. When two stimuli of the same modality are separated by 10-15 deg, one can observe a significant attenuation in the activation bubbles. This is a consequence of the lateral inhibition within the unimodal areas, which implements a competitive mechanism. While two stimuli are very close, a single larger bubble is evident in the central position, as a consequence of the short-distance lateral excitation. ii) The activity of multimodal neurons exhibits an evident reduction when two stimuli are approached to each other. This suppression is greater than that observed in the unimodal area, and may even reach

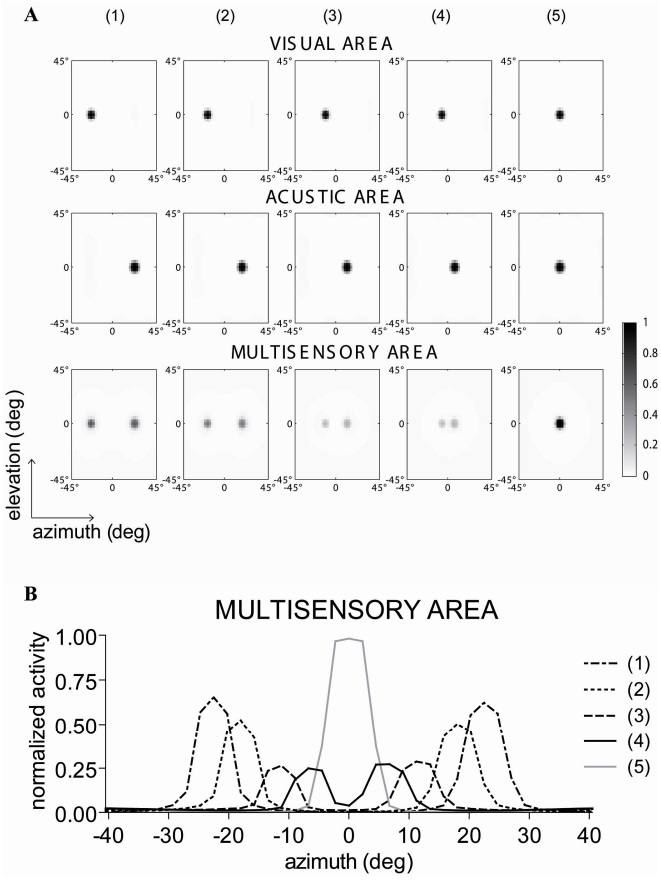


more than 80% of the original level (i.e., the level which is evident when the two stimuli are distant, and no suppression occurs). Finally, when the two stimuli are almost superimposed, one can observe just a mild enhancement (the increase in the response of multimodal neurons is less than 10%).

Figure 11 – Model response to two simultaneous auditory stimuli (intensity i^a = 17) placed at five different positions in space. The meaning of symbols is the same as in Fig. 10. Within modality suppression is evident (greater than 70%). Within modality enhancement of two superimposed stimuli (simulation 5) is mild. Worth noting is the smaller resolution in case of proximal stimuli (simulation 4) compared with Fig. 10, as a consequence of the larger acoustic RFs.

Figure 12 illustrates the case of cross-modal integration, that is the results obtained by applying a visual and an auditory stimulus of the same intensity at different positions.

Figure 12 – Model response to two simultaneous cross-modal stimuli (intensity i^v = 17, i^a = 17) placed at five different positions in space. The meaning of symbols is the same as in Fig. 10. Cross modality suppression is evident (greater than 60%). Cross-modal enhancement of two superimposed stimuli (simulation 5) is also evident (about +50%).



In this case, no suppression in the unimodal areas is evident. This is the consequence of the absence of any direct connections among the two unimodal areas, and of the small values assigned to feedback synapses from the multimodal area to the unimodal areas (Tab. I). In the present model, indeed, two unimodal

areas can interact only via feedback links from multimodal areas, since no direct connection between unisensory areas is provided. In fact, the only way the two unimodal areas can communicate in our model is via the presence of these feedback links. By contrast, a strong crossmodal interaction is evident in the multimodal area. Cross-modal suppression (about 60% of the basal level) occurs when the distance between the two stimuli is in the range 10-15 deg. Two proximal stimuli cause a significant cross-modal enhancement: the response of the central multimodal neuron is greater (+ 50-60%) than the response to each individual stimulus.

In order to simulate the inverse effectiveness property of SC neurons, we repeated the same simulations as in Fig. 12, using different levels for the two input signals. From these simulations, cross-modality enhancement has been evaluated by means of the Interactive Index. Results are displayed in Figure 13 as a function of the intensity of the input stimuli. This figure shows that enhancement computed with input stimuli of low intensity may be as great as 500% and decreases monotonically with the intensity of the inputs. At strong intensity, enhancement declines to 30-40% or less.

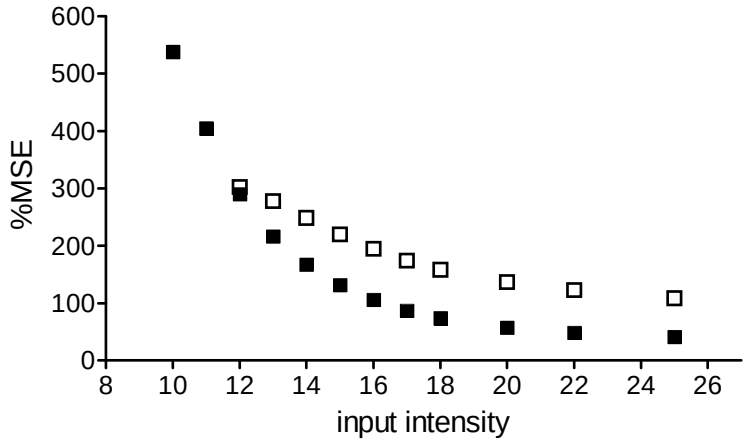


Figure 13 – Multisensory enhancement computed with the model in steady state conditions, in response to two superimposed cross-modal stimuli of the same intensity, placed at the centre (solid symbols) or at the periphery (open symbols) of the RF. The intensity of the stimuli is plotted in the x-axis. Stimuli with intensity greater than 12 have been applied at the periphery since lower inputs do not produce any significant response. Enhancement decreases with stimulus strength, according to the inverse effectiveness principle. Moreover, stimuli placed at a weakly effective location (at the periphery of RF) induce greater enhancement than do stimuli at a more effective location (at the centre of RF).

In conclusion, with basal parameter values the model exhibits strong cross-modality enhancement, but negligible within-modality enhancement. Cross-modality enhancement satisfies the inverse-effectiveness principle. Furthermore, the model exhibits both within-modality and cross-modality suppression in response to distal stimuli, the first phenomenon being a little stronger than

the second. All these results agree, both qualitatively and quantitatively, with data reported in the literature.

Sensitivity analysis – A subsequent set of simulations has been performed to unmask the role of the synapses in the multimodal area (i.e., lateral synapses; feedforward synapses from the unimodal areas to the multimodal areas; feedback synapses from the multimodal area to the unimodal areas).

The aim of these simulations is to show how different responses of SC cells, described by neurophysiological studies, can emerge in the model as a simple consequence of differences in synaptic strength.

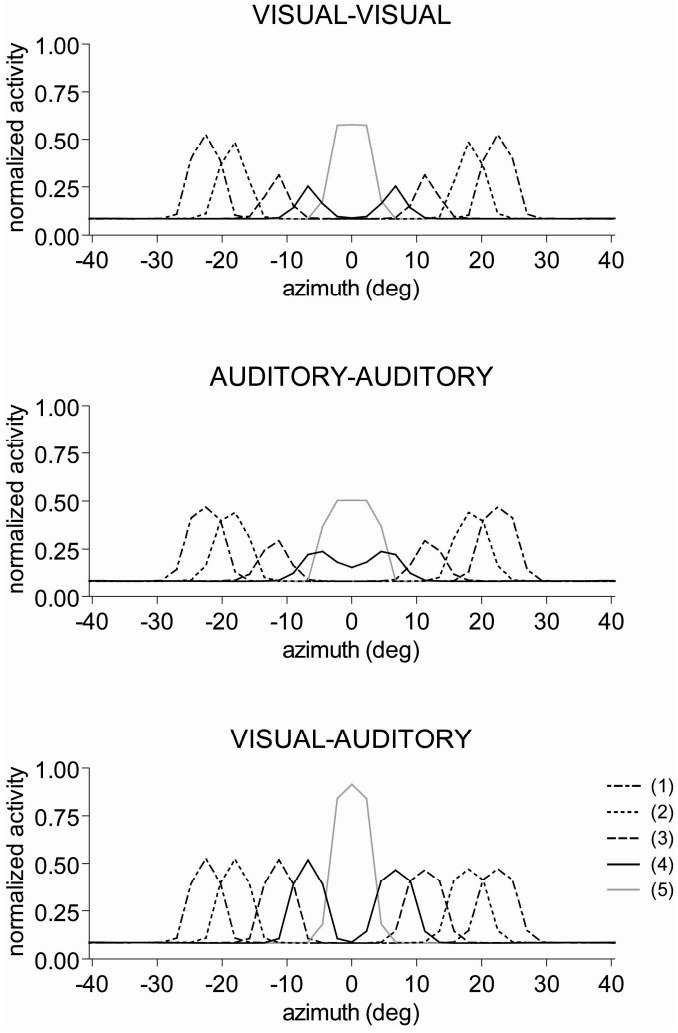


Figure 14 – *Sensitivity analysis on the role of lateral synapses in the multisensory area*. The three panels show the response of neurons in the multisensory area whose RF is centred at the vertical coordinate 0° and at the azimuth coordinate from −40.5° to +40.5° (that is, 37 neurons positioned along the middle line of the vertical field) during the five simulations depicted in Fig. 10 (visual–visual, upper panel), in Fig. 11 (auditory–auditory, middle panel) and in Fig. 12 (visual–auditory, lower panel). The number labelling each line pattern identifies the simulation, as in the Figs. 10–12 (see legend of Fig. 10). Results differ from those in Figs. 10–12, since the strength of all lateral synapses in the multisensory areas has been set at zero. Moreover, to maintain a mild neuron activity despite the absence of lateral inhibition, the position of the neuron sigmoid function has been translated ($\vartheta = 8$ in Eq. (9)). In these conditions, within-modality suppression is still present (upper and middle panels) whereas cross-modality suppression is abolished (bottom panel). However, cross-modal enhancement is evident yet (bottom panel).

Figure 14 shows the profile of multimodal neuron response, in the same conditions as in Figs. 10-12, after the total elimination of lateral synapses in the multimodal area. Results show that deleting lateral multimodal synapses does not eliminate either within-modality suppression or cross-modality enhancement. By contrast, cross modality suppression disappears. Hence, in these conditions the model can explain the presence of cells which exhibit within-

modality suppression without cross-modality suppression, as observed in many SC neurons (Kadunce et al., 1997) .

Figure 15 shows the effect of a change in feedforward synapses from the unimodal areas to the multimodal area.

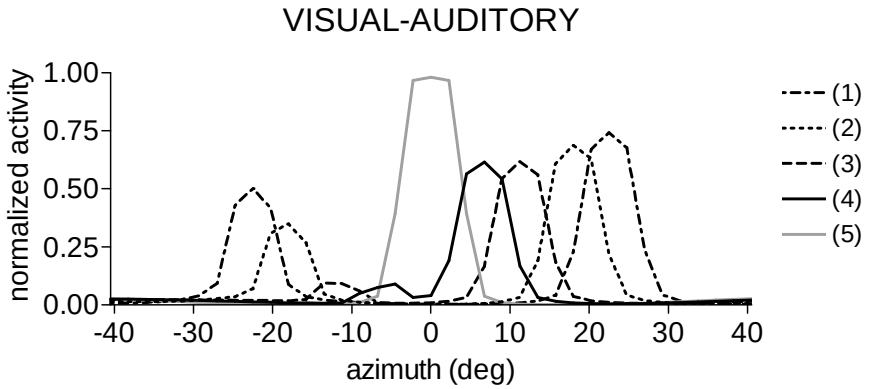


Figure 15 – *Sensitivity analysis on the role of feed forward synapses*. The panel shows the response of neurons in the multisensory area whose RF is centred at the vertical coordinate 0° and at the azimuth coordinate from −40.5° to +40.5° (that is, 37 neurons positioned along the middle line of the vertical field) during the five simulations depicted in Fig. 12 (visual–auditory). The number labelling each line pattern identifies the simulation, as in the Figs. 10–12. Results differ from those in Fig. 12, since we decreased the feedforward synapses coming from the visual unimodal area ($k^v = 8$) and increased the feedforward synapses coming from the auditory unimodal area ($k^a = 9$). In these conditions, one can observe cross-modal suppression of the auditory stimulus on the response to the visual stimulus, but not viceversa.

In basal conditions (Tab. 1) these synapses have been set so that the effects of two identical visual and auditory stimuli on multimodal neurons were comparable. In order to break this balance, in the simulations depicted in Fig. 15, we slightly reduced the feedforward synapses from the visual area, and we slightly increased those from the auditory area. In these conditions, an auditory stimulus has a stronger effect on multimodal neurons than an identical visual stimulus. As it is evident in Fig. 15, as a consequence of the competitive mechanism implemented in the multimodal area, the auditory stimulus now causes a strong cross-modal suppression, whereas the cross-modal suppression caused by a visual stimulus is almost negligible. Hence, the model can easily explain the presence of multimodal neurons characterized by the suppression of one-modality on the other one (in our example, auditory on visual), but not vice-versa (Kadunce et al., 1997) .

A last set of simulations has been performed to point out the potential influence of the feedback synapses on model behaviour. In basal conditions, the strength of these synapses has been set to

quite low values, in order to induce just a mild effect on the unimodal areas. However, an interesting behaviour emerges, if the feedback synapses are reinforced.

A first example is illustrated in Figure 16.

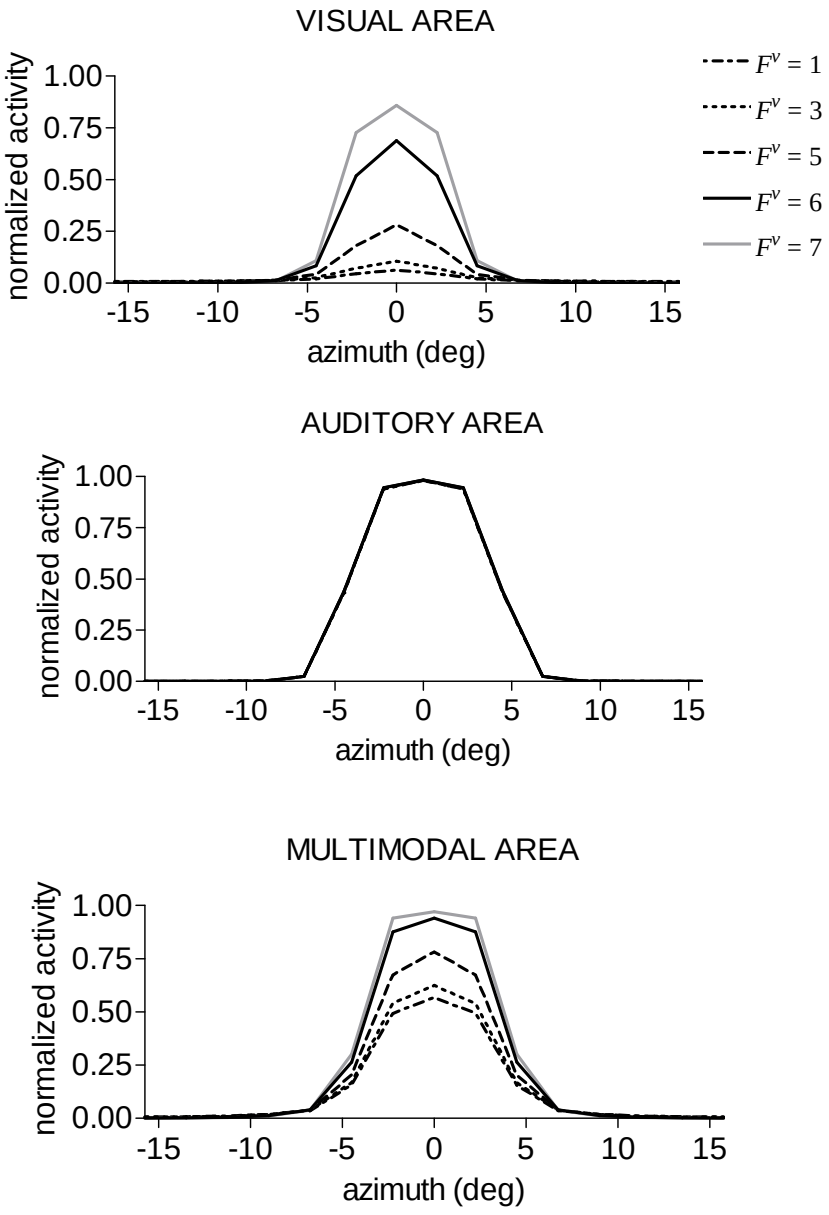


Figure 16 – ***Role of feedback synapses on the reinforcement of poor perception*** **— The three panels show the response of neurons whose RF is centred at the vertical coordinate 0° and at the azimuth coordinate from −15.75° to +15.75° (that is, 15 neurons positioned along the middle line of the vertical field) in the visual (upper panel), auditory (middle panel) and multisensory (bottom panel) areas. The curves have been computed in steady state conditions in response to a mild visual stimulus (i^v = 4) and a simultaneous strong auditory stimulus (i^a = 17) positioned at the central point of space (coordinates 0°, 0°). Five simulations were performed using increasing values of the feedback synapses from the multisensory area to the visual unisensory area (F^v = 1, 3, 5, 6 and 7). In basal conditions, the visual stimulus evokes just negligible activity in its unisensory area. Increasing the feedback strength causes the appearance of strong activity in the visual area, evoked by the paired auditory stimulus.**

Here we applied a mild visual stimulus at the central spatial position, and a simultaneous strong auditory stimulus in the same position. This condition may simulate, for instance, what occurs in a patient with a visual deficit (here simulated by the administration of a mild visual stimulus), with only poor response to visual stimulation. It is known that such patient may benefit by application of a concurrent strong auditory stimulus in the same spatial location to enhance the detection of a

visual stimulus (Bolognini et al., 2005; Frassinetti et al., 2005) . With the basal value of the feedback synapses, the visual stimulus can evoke only a negligible response in the unimodal visual area. By contrast, the auditory stimulus evokes a strong response in the auditory unimodal area which, in turn, can trigger a moderate response of multimodal neurons. The simulation has then been repeated with a progressive increase in feedback synapses from the multimodal area to unimodal visual area: an increase in these synapses can induce a reinforcement of the activity in the unimodal visual area. For values of the synaptic strength greater than 5, a large activation bubble appears in the unimodal visual area, similar to that evoked by a strong visual stimulus. The result is that a poor visual stimulus, which cannot evoke *per se* a significant activity in the unimodal area if presented alone, can be reinforced and fully perceived thanks to the occurrence of an auditory stimulus placed in the same position (Bolognini et al., 2005; Ladavas, 2008).

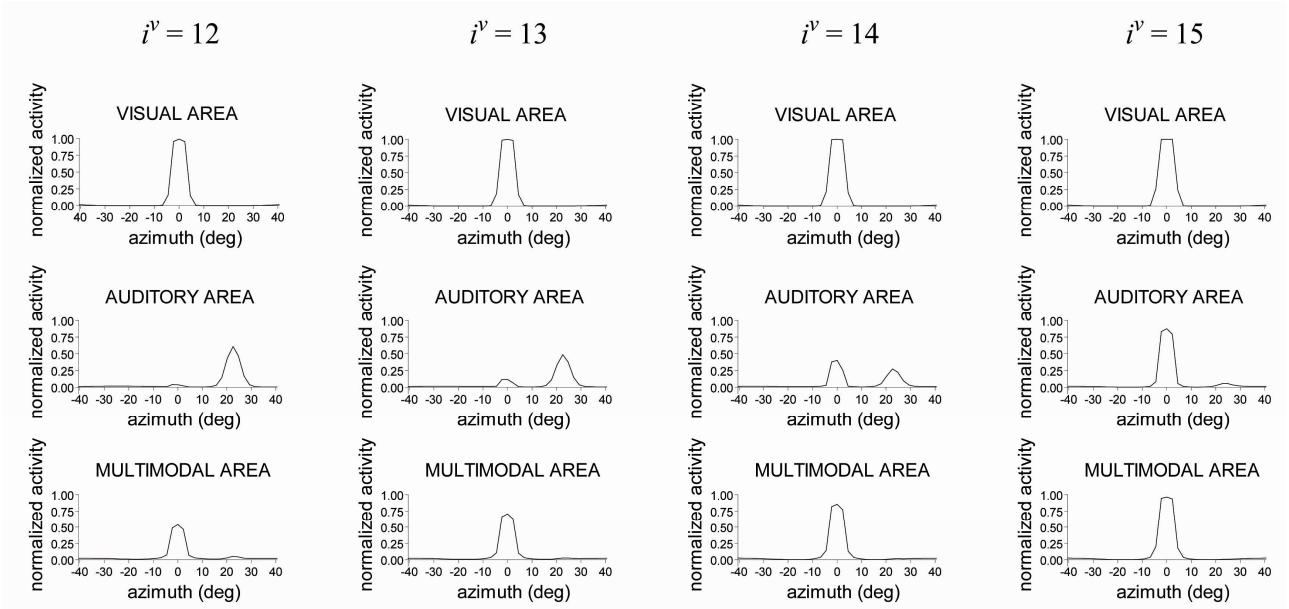


Figure 17 – *Role of feedback synapses on ventriloquism* — Each row shows the response profile of neurons whose RF is centred at the vertical coordinate 0° and at the azimuth coordinate from −40.5° to +40.5° (that is, 37 neurons positioned along the middle line of the vertical field) in the visual (upper panel), auditory (middle panel) and multisensory (bottom panel) areas. The curves have been computed in steady state conditions in response to a moderate auditory stimulus ($i^a = 9$) positioned at the coordinate (22.5°, 0°) and to a simultaneous stronger visual stimulus positioned at the central point of space (coordinates 0°, 0°). In these simulations we assumed high values of feedback synapses ($F^v = F^a = 15$). The four columns differ for what concerns the intensity of the visual stimulus (from left to right: $i^v = 12, 13, 14, 15$). In the left column, the auditory stimulus is perceived at the correct position. In the two middle columns, one can observe a conflict between two auditory activities (located at the original position and at the position of the visual stimulus). Finally, in the right column, the visual stimulus ‘‘captures’’ the auditory one at its position, and suppresses acoustic activity at the original position (ventriloquism).

Figure 17 displays the simulation results obtained by presenting a stronger visual stimulus and a weaker auditory stimulus at different positions (distance 22 deg). Both stimuli are strong enough to evoke a significant activity in the corresponding unimodal areas, although the visual activity is higher. Furthermore, in these simulations we assumed the existence of strong feedback synapses from the multimodal to both unimodal areas. In these conditions, if the two input stimuli are no too different in intensity (i^v = 12; i^a = 9) the visual and the auditory stimuli are perceived separately at the correct position, although only the stronger stimulus (i.e., the visual one) can evoke a consistent activity in the multimodal area. If the visual stimulus is slightly increased (i^v = 13 or 14, second and third columns), we can observe the presence of two zones of activity in the auditory area: i.e., a competition occurs on the position of the auditory stimulus. Finally, if the visual stimulus is further increased (i^v = 15), an activation bubble occurs in the auditory area at the same position of the visual stimulus, as a consequence of the feedback from the multimodal area; the auditory activity in the original position is almost completely suppressed by the presence of competitive lateral inhibition. This result simulates the ventriloquism (Pick et al., 1969).

Network dynamical response

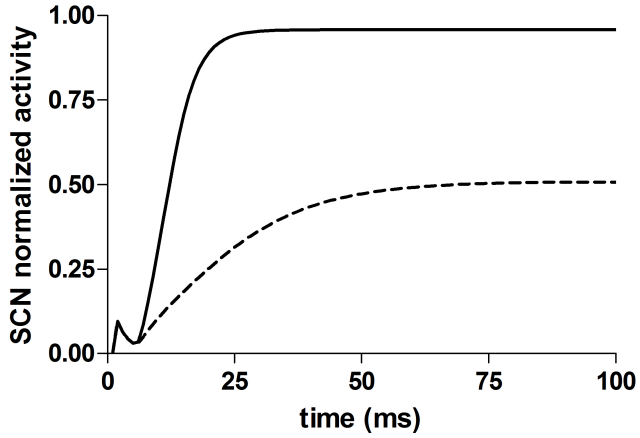


Figure 18 – The temporal response of a SC neuron evoked by a visual stimulus alone (*dashed line*) and by a multi-sensory stimulation (visual + auditory; *solid line*). In both cases, the stimuli were applied at the time *t=0* ms, starting from an initial basal condition (no stimulation). *Vertical lines* denote the settling time (i.e., the time required to approach 90% of the final steady-state level). The settling time of the multisensory response is less than one-half the settling time of the unisensory response (17 ms vs. 43 ms).

Figure 18 shows the temporal response of a SC neuron evoked by a visual stimulus alone, and by a multi-sensory stimulation (visual + auditory). In both cases, the stimuli were applied at time t=0 ms, starting from an initial baseline condition (no stimulation). The settling time of the responses

(i.e., the time required to reach 90% of the final steady-state level) is indicated in the figure. This is as high as 43 ms for the unimodal response, and is reduced down to 17 ms for the multisensory one. Hence, multisensory integration does not only reinforce the final response, but also significantly reduces the settling time. The value of settling time shown in Fig. 18 is much higher than the value expected on the basis of the time constant used in Eq. (8). Indeed, in a feedback network the settling time is significantly affected by the time required for the feedback mechanisms to reach an equilibrium. In order to gain a deeper understanding on the mechanisms responsible for the slow transient response, we performed a sensitivity analysis on the strength of lateral synapses in both unimodal and multimodal areas.

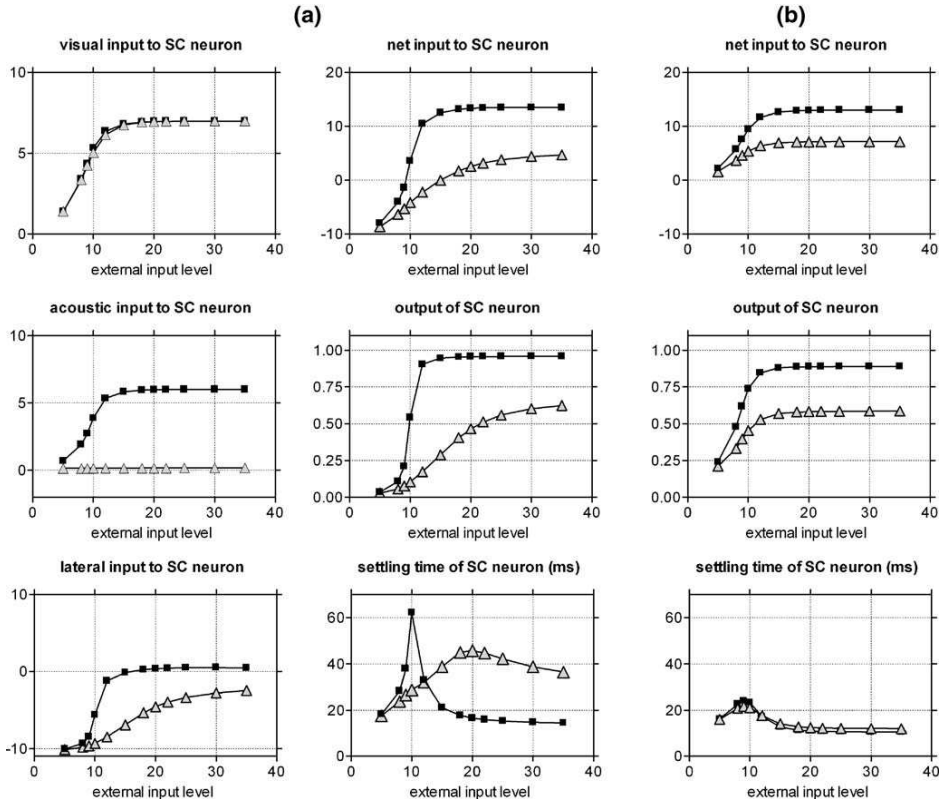


Figure 19 – Settling time of the SC neuron response to a step unimodal visual input (*triangle*) and to two step cross-modal stimuli with the same strength (*square*) plotted as a function of stimulus strength. Results in the (a**) have been obtained with basal parameter values. Those in (**b**) have been obtained after elimination of lateral synapses in the multimodal area. In the latter case, in order to avoid excessive excitation of multimodal neurons in basal conditions, the central abscissa of the sigmoidal relationships [i.e., parameter ϑ^m in Eq. (9)] has been increased from 3 to 6. (**a**) The left column shows the visual input, the auditory input, and the lateral input to the SC neuron (the latter is equal to the difference of lateral excitation and lateral inhibition). The right column shows the net input (that is the sum of the three previous contributions), the output of the SC neuron and the settling time. It is worth noting that the settling time rapidly decreases when lateral excitation balances lateral inhibition. (**b**) The same figures as in (**a**), computed without lateral synapses in SC. Visual and acoustic inputs are the same, while lateral input is zero. Hence, only the last three panels are redrawn.**

Moreover, we computed the settling times for different (multimodal and unimodal) values of the input stimuli (Figure 19). In particular, we first computed the settling times with basal parameter values; then, we set lateral synapses in the multimodal area to zero, by maintaining the unimodal area synapses; finally, we also suppressed lateral synapses in the unimodal areas. Figure 19a shows

the results obtained with basal parameter values and all mechanisms intact. The settling times for unimodal and bimodal stimulation are comparable at very low levels of the input (range 10 – 30 ms). If the input is moderately increased, the settling time becomes much higher in case of bimodal stimulation than in case of unimodal stimulation (60 vs. 30 ms). Conversely, at high input level, settling time to bimodal stimulation decreases below 20 ms, while settling time to unimodal stimulation increases to more than 40 ms. Suppression of lateral synapses in the multimodal area [Fig. 19(b)] causes a dramatic reduction of settling time (down to about 12 ms for both unimodal and bimodal stimulation). Finally, if also unimodal lateral synapses are suppressed (results not shown for briefness) the settling time displays a further minor decrease (down to 8–9 ms). In the latter condition, the only feedback mechanism is from multimodal to unimodal areas. We can thus conclude that the settling time is affected by the feedback mechanism in the multimodal area, and this time is reduced by high multimodal stimulation.

In order to provide a deeper description of the different factors affecting the settling time, in Fig. 19(a) we also show the output of the SC neuron and the input to the SC neuron, distinguishing between the visual input, the acoustic input, the lateral input (i.e., that due to lateral synapses in the SC) and the net input (i.e., the sum of the previous factors). Results show that the settling time is particularly high when the neuron works close to the central region and its net input (due to the sum of the input coming from the unimodal areas and the lateral inputs from other neurons in the SC) is close to zero, or just a little positive. When excitation becomes much higher than inhibition, the settling time decreases rapidly. Substantially, settling time dramatically falls when the overall lateral excitation becomes comparable to the lateral inhibition, and lateral excitation and inhibition balance.

DISCUSSION

Many studies in the last two decades described the physiological properties of neurons in the superior colliculus which integrate stimuli from different sensory modalities, in order to produce efferent motor commands, namely head and eyes orientation (Kadunce et al. 1997, 2001; King and Palmer 1985; Meredith and Stein 1986a, b; Perrault et al. 2003, 2005; Populin and Yin 2002; Stanford et al. 2005; Wallace et al. 1998). Results from these studies contributed to individuate some general principles ruling the integrative properties of these neurons. First, two stimuli are strongly integrated when they occur in close spatial and temporal register. On the other hand, stimuli which do not overlap in space and time may exert a reciprocal inhibitory influence (cross-modal and within-modal suppression). Finally, multisensory integration is much stronger for stimuli individually less efficient in inducing a unisensory response (principle of inverse effectiveness).

All these experimental data provide a clear and coherent scenario on the properties of multimodal neurons. This scenario is further supported by behavioural experiments both on animals (Stein et al. 1988, 1989) and humans (Frassinetti et al. 2002), showing facilitation or suppression of attentive/orientation responses in the presence of multimodal stimuli compared with unimodal stimuli (Amlot et al. 2003; Bermant and Welch 1976; Frens et al. 1995; Hughes et al. 1994; Perrott et al. 1990).

These properties of multisensory integration depend not only on neuronal individual characteristics, but also on the organization of the circuitry that processes unimodal stimuli and conveys these stimuli toward multi-sensory neurons. A deeper insight into these mechanisms and into the possible topology of the neural network involved can be provided by mathematical models and computer simulation techniques. Mathematical models allow the formulation of hypotheses in rigorous quantitative terms, the validation/rejection of these hypotheses on the basis of available experimental data and the synthesis of multiple knowledge into a coherent structure. In particular, in the case of multisensory SC neurons, there are presently enough data to attempt an accurate analysis

and validation with mathematical models, and the synthesis of these data into a comprehensive, although simplified, theoretical scenario.

In recent years, Patton et al. (2002) and Anastasio et al. (2000) developed mathematical models, based on a Bayesian approach, to study the properties of neurons in the deep superior colliculus. They postulated that these neurons compute the posterior probability that a target is present in their receptive field, and showed that this hypothesis can explain cross-modal enhancement. In a subsequent work (Patton and Anastasio 2003), the same authors proposed some neural implementations, based on modified perceptron models, and showed that these can explain cross-modal enhancement and within-modality suppression. Although the latter models share some aspects with our (especially in the use of sigmoidal non-linearities) there are also fundamental differences. First, the authors do not explicitly consider the spatial arrangement of the input stimuli, not their intensity, but modify the covariance of the input channels (assuming that inputs of the same modality have greater spontaneous covariance than inputs of different modality). This assumption may be true if the two unimodal stimuli come from neurons with overlapping receptive fields. Furthermore, the authors use a multiplicative interaction in their models. Our model adds several new aspects: it considers the spatial position and the intensity of the input stimuli explicitly, and simulates the effect of competitive interactions involved in target detection. It analyzes dynamical ranges to stimuli of increasing amplitude. Finally, it explains enhancement and suppression without assuming multiplicative interaction at the synaptic level, but considering a Mexican-hat disposition of synapses between adjacent neurons.

In a further version of their model, Anastasio and Patton (2003) separately considered the ascending and descending inputs to SC neurons, and trained the connection weights from these inputs with different rules, to have both unimodal and multimodal neurons in the same theoretical model. This differentiation is not considered in our work, but may be the subject of future extensions (see also discussion below).

The present work was designed to elucidate possible neural mechanisms involved in multisensory integration in SC by using a mathematical model. To this aim, we developed a model of a simple neural circuit which encompasses several mechanisms, still maintaining a moderate level of complexity. Actually, the model aspires to represent a good compromise between completeness, on one hand, and conceptual (and computational) simplicity on the other.

The basic idea of this model is that multimodal neurons in the superior colliculus receive their inputs from two upstream unimodal areas, one devoted to a topological organisation of visual stimuli and another devoted to a topological organisation of auditory stimuli. For the sake of simplicity, in this model somatosensory stimuli are neglected, i.e., we consider only the problem of audiovisual integration. Moreover, the exact location of these areas is not established in our model, i.e., we did not look for a definite anatomical counterpart. Experimental data suggest that multisensory neurons are created by the convergence of modality-specific afferents coming from different sources (Edwards et al. 1979; Huerta and Harting 1984; Wallace et al. 1993). Moreover, results of recent experiments (Jiang et al. 2001; Jiang and Stein 2003) indicate that SC neurons respond to these inputs in different ways: the nature of multisensory integration is altered depending on the considered input sources and cortical deactivation. Limitations of our model in fitting these experiments, and lines for future improvements, will be discussed at the end.

Several mechanisms have been included in this simple basal circuit, each with a specific significance and a possible role in affecting final responses: (1) non-linearities in the activation function of single neurons (i.e., a lower threshold and upper saturation, expressed with a sigmoidal relationship). As will be commented below, these nonlinearities are essential to understand some important properties of multisensory integration, such as the inverse effectiveness, and the possibility of superadditive, additive or subadditive integration. (2) Lateral synapses (excitatory or inhibitory) among neurons in the same unimodal area. They have been modelled with a classical “Mexican hat” disposition, i.e., a close facilitatory area surrounded by an inhibitory annulus. These synapses play a fundamental role in producing the receptive field of multimodal neurons. Moreover,

they contribute to the within-modality suppression documented in many experiments in the absence of cross-modal suppression (Kadunce et al. 1997, 2001). (3) Feedforward connections from unimodal to multimodal neurons. The strength of these synapses affects the sensitivity of multimodal neurons and their unisensory dynamical range (i.e., the maximum response to a single stimulus of a given modality). (4) Lateral synapses among multisensory neurons. These synapses are necessary to obtain a significant cross-modality suppression between spatially separated auditory and visual stimuli, as documented in recent experiments (Kadunce et al. 2001). Moreover, they significantly affect the settling time of the response. (5) Excitatory backward connections from multimodal neurons to unimodal neurons at the same spatial position. Inclusion of these connections considers the possibility that the response by a multimodal neuron reinforces the response at an earlier unimodal area (for instance, that a strong visual stimulus may help perception of a weak auditory stimulus in the same position, and vice versa). In the present simulations the strength of these backward connections has been maintained quite low, hence they do not play a major role in simulation results. However, it may be interesting to investigate the possible effect of a reinforcement of these synapses in further studies.

In summary, although some important properties in the model (for instances, cross-modal enhancement and inverse effectiveness) derive from sigmoidal non-linearities, lateral synapses in the unimodal and multi-modal areas also play a significant role, explaining the suppression (either within-modality or cross-modality) between two distal stimuli and affecting the settling time of the response. All elements included in the model are necessary to account for the variability of in vivo SC cell behaviour.

By incorporating the previous mechanisms, and using a single set of parameters (see Table 1), the model was able to make several predictions, which can be compared with experimental data. In the following, the main simulation results are critically commented:

1. Inverse effectiveness—As it is evident in Figs. 5 and 6, the capacity of multisensory neurons to integrate crossmodal stimuli strongly depends on the intensity of unisensory inputs. As in Perrault

et al. (2005), in the present work the facilitatory interaction has been quantified using two alternative metrics, namely interactive index and contrast. The first relates the multisensory response to the stronger unisensory response. The second relates the multisensory response to the predicted sum of the two unisensory responses. Both metrics are affected by the intensity of the unisensory inputs, with cross-modal response exhibiting a significant decrease if stimulus intensity is progressively raised. This behaviour, which is know as “inverse effectiveness”, is a consequence of the non-linear property of neurons, and depends on the position on the sigmoidal relationship after application of the more effective input. To explain this mechanism, let us consider the case in which, after application of the more effective unisensory stimulus, the SC neuron works at the lower portion of its sigmoidal relationship, close to the threshold. Here, application of a second stimulus may move the working point into the linear portion of the curve, thus causing a disproportionate increase in the response compared with that evoked by the first input (superadditivity, enhancement greater than 100%). By contrast, if the neuron works in the central (quasi-linear) region, the effect of a second stimulus is simply additive. Finally, if the upper saturation region is approached one can have subadditivity, since a second stimulus can induce only a minor increase in neuron activity. The last case is not simulated in this work since, with the present value of feedforward synapses, a single stimulus cannot move the working point close to the upper saturation region. Sub-additivity, however, can be mimicked by increasing the feedforward synapses.

Furthermore, recent experiments (Carriere, Royal, & Wallace, 2008) on multisensory neurons in the cortex (whose properties closely resemble those of multisensory neurons in the SC (Carriere et al., 2008; Wallace et al., 1993)), have shown that multisensory enhancement is stronger for input stimuli applied at weakly effective positions within the RF than at more effective locations. As shown in Fig. 13, model predictions agree with this in-vivo observation: stimuli applied at the border of RF produce greater enhancement than stimuli applied at RF centre. Hence, the model is

able to account for the principle of inverse effectiveness both as a function of the intensity properties of input stimuli and of the spatial properties of stimuli location within the RF.

2. Dynamic range—The multisensory dynamic range of multimodal neurons is greater than the unisensory dynamical range (Perrault et al. 2005). This means that the maximal response evoked by a combination of auditory and visual stimuli in close spatial and temporal register is greater than the maximal response evoked by a single stimulus of either modality (see Fig. 5 in Perrault et al. 2005). Such a property is explained in our model by the presence of two sigmoidal relationships, disposed in a series arrangement. Let us consider a single stimulus and progressively increase its intensity: in our model, the maximal response in the SC (see Fig. 4) is determined by the upper saturation of neurons in the upstream unimodal area, and by the strength of the feedforward synapses linking this unimodal neuron to the downstream (multimodal) neuron [synapses k^a or k^v in Eq. 10]. This input does not lead multimodal neurons to saturation. Consequently, if we apply a combination of a visual and an auditory stimulus, and progressively increase their intensity (multisensory dynamic range), the downstream multimodal neuron can be driven closer to its upper saturation and exhibits a greater response.

3. Cross-modality vs. within modality integration—According to the literature (Stein and Meredith 1993) in our model a combination of two cross-modal stimuli within the RF results in significant enhancement of the SC response, but the same effect is not visible when the two stimuli are presented as a within-modality pair. A second within-modality stimulus applied within the RF causes just a mild enhancement (Fig. 7).

4. Spatial relationship between two (within-modal or cross-modal) stimuli—In agreement with experimental data (Kadunce et al. 1997, 2001), our model shows that, as the spatial distance between two stimuli increases, multisensory integration in SC layer shifts from enhancement to suppression both using within-modality and cross-modality stimuli. However with our choice of basal parameter values (Table 1), synapses in unimodal areas do not play a relevant role and do not affect suppression properties of superior colliculus: in these conditions both within-modality and

cross-modality suppression depend mainly on the presence of lateral inhibition in the multimodal area (see Fig. 7). Hence, within-modality and cross-modality suppression cannot be decoupled. Previous results in the literature (Kadunce et al. 1997) show the existence of different types of SC multimodal neurons: some exhibit both cross-modality suppression and within-modality suppression (as in the exemplum in Fig. 7); others exhibit within modality suppression without cross-modality suppression (as in Fig. 8 right panels). The model suggests that these differences can be ascribed to a different balance between lateral inhibition in the unimodal and multimodal areas. In fact, increasing lateral inhibition in the unimodal area with poor lateral inhibition in the multimodal area may explain within-modality suppression without cross-modality suppression.

5. Temporal dynamics—As illustrated in Figs. 18 and 19 the temporal response to a combination of two stimuli in different sensory modalities is much faster than the temporal response to a single stimulus. Looking at Fig. 18, we can say that the settling time evoked by two large stimuli is less than half the settling time evoked by a single stimulus. Furthermore, sensitivity analysis (Fig. 19) demonstrates that the observed settling time is affected not only by the inputs and time constants, but also by the lateral feedback in the multimodal area. Hence its value is an emergent property of the network which depends on the time required for feedback mechanisms to reach a steady state level. This time is significantly decreased by two large cross-modal stimuli compared with a unimodal stimulus. In particular, as shown in Fig. 19, the settling time falls dramatically when lateral inhibition is overcome by lateral excitation. However, it is important to stress that the settling time in our model does not replicate the temporal pattern of neuron response to real (auditory and visual) input stimuli, as measured in vivo, but only represents the network dynamics. Actually, the temporal pattern of neuron response during in vivo experiments may depend on additional factors, such as the time dynamics of the peripheral receptors (such as the retina and the cochlea), as well the latency of the neural pathways from the receptor to the SC. Analysis of these factors is well beyond the aim of the present work.

However, we think that the 20–25 ms difference in settling times, evident in Figs. 18 and 19a may be of interest, and may in part explain the difference in overt behaviour between multimodal and unimodal stimulation observed during behavioural experiments (Frens et al. 1995; Perrott et al. 1990).

An interesting aspect of our simulations (see Fig. 5) is that the behaviour of a neuron in response to a second stimulus can shift from entirely superadditive to superadditive–additive depending on the intensity of the first stimulus applied. This signifies that, contrarily to what frequently claimed in the literature (Perrault et al. 2005), the behaviour of a neuron in terms of its multisensory contrast is not an intrinsic property, but depends on its particular operative conditions.

In summary, the results in Figs. 3, 4, 5, 6, 7, 10, 11, 12, 13 and in Fig. 18 have been obtained only by changing input stimuli, (i.e., modality, intensity and spatial position), without altering any parameter of the model. Therefore, they mimic results which can be obtained on a single neuron, not on a group of different neurons. The purpose was to show that the model, designed on the basis of few principles, with a single set of parameters (hence, representing a single case) can explain and summarize several data, characterized by different properties of the input.

An example of possible individual variability among neurons and classes of neurons, and the importance of synaptic connections in the multisensory area, has been investigated in Figs. 8, 9, 14, 15, 16 and 17, by changing the value of lateral and feedforward synapses in multimodal and unimodal areas. These simulations emphasize the possibility to have within- modality suppression without cross-modality suppression, as observed in some SC neurons (Kadunce et al. 1997).

Finally, we wish to comment on possible model limitations, which may the subject of future improvements. First, the structure of the model is drastically simplified compared with the reality. In our model, we assumed that the receptive fields of acoustic and visual neurons, converging to the same multimodal neuron, have a circular shape and are exactly centered at the same position in space, with the visual neuron having a smaller RF compared with the acoustic neuron (hence, the RF of visual neurons is entirely contained inside the RF of acoustic neurons). By contrast, as clearly

documented in the literature, the two RFs do not exhibit a 100% overlap (although, as reported in Kadunce et al. 2001, in many multisensory neurons more than 70% of the visual RF is contained within the area of the auditory RF). Moreover, in most simulations we assumed that synapses within a single area are perfectly symmetrical and that there is no difference in the properties of neurons within the same area, except for the position of their RF. Of course, in real networks, neurons and synapses exhibit a random variability, and two proximal neurons in the same area can exhibit different properties and disparate responses. Of course, these simplifications have been adopted to have a more straightforward model, and to make the analysis of results easier.

A first step to overcome this limitation was presented in Fig. 9, where we locally modified the values of synapses and displayed the activity of three neurons in the same network with different properties. These results show that the network can be easily changed, to account for the simultaneous existence of neurons with different properties, as experimentally observed.

Another limitation of our model is the absence of direct synapses between neurons in the two unimodal areas. Indeed, a unimodal auditory neuron (area A) and a unimodal visual neuron (area V) communicate only indirectly, through the backward synapses coming from the multimodal SC neuron (area SC). This choice has been adopted according to a principle of parsimony, i.e., to limit the number of mechanisms in the model. It is possible, however, that neurons in unimodal areas communicate also directly via lateral synapses. The possible effect of these links on model response may be the subject of future extensions.

Jiang et al. (2001) and Jiang and Stein (2003) demonstrated that the capacity of SC neurons to integrate crossmodal sensory stimuli is strongly dependent on influences from two cortical areas [*the anterior ectosylvian sulcus (AES) and the rostral lateral suprasylvian sulcus (rLS)*]. However, the response to unimodal stimuli remains largely intact even if these cortices are temporarily deactivated (Jiang et al. 2001; Jiang and Stein 2003). This aspect raises additional problems for a mathematical model, which might be solved including more unimodal areas and/or a more complex topology for the network. For instance, the two unimodal areas in the present model might represent

the visual and auditory inputs from the AES (and perhaps rLS), allowing multisensory integration (either enhancement or suppression). In particular, the AES contains distinct sensory representations (a somatosensory, a visual and an auditory region) and also has many multisensory neurons. Yet, only unimodal neurons in the AES send inputs to the SC (Stein 1998; Wallace et al. 1993). Additional inputs from other cortical and subcortical structures might be responsible for the responses observed after deactivation of the AES. Hence, a more complete model should include at least four distinct unimodal inputs (two visual and two auditory) to SC neurons, reflecting descending inputs from cortico-collicular regions (responsible for multisensory integration) and ascending inputs coming from a variety of other sources (which do not produce multisensory integration). The present model considers only the first (descending) inputs, hence cannot simulate SC behaviour after AES and rLS deactivation. Of course, development of the more complex model might be the subject of future refinements and extensions.

Finally, it is important to stress that most of the data mentioned in this work have been obtained in anesthetized animals, hence the model presented here aspires to simulate these conditions. In recent years, controversial results have been published on the possible effect of anaesthetics on bimodal enhancement in the superior colliculus. While some authors report multisensory integration and cross-modal enhancement in alert untrained cats (Wallace et al. 1998) others observed depressed enhancement in behaving cats compared with anesthetized animals (Frens and Van Opstal 1998; Peck 1996; Populin 2005; Populin and Yin 2002). There are two main aspects which can in part explain these differences. First, some authors (Populin and Yin 2002) used a different metrics to quantify multimodal enhancement in alert animals. By this measure, the authors consider multimodal enhancement only in case of superadditivity. In our model, superadditivity may be converted to simple additivity, or even to subadditivity by changing some model parameters. Second, the SC receives a vast intrinsic inhibitory network (Mize et al. 1994), and receives both ascending subcortical inputs and descending inputs from cortical regions (such as AES and rLS). It is thus possible that anaesthesia modifies some parameters in the model, alters the balance among

the inputs, and/or the balance between inhibition and excitation. These effects may be studied in future versions of the model, including additional inputs and using a sensitivity analysis on parameter changes and/or on input changes.

REFERENCES

Amlot R, Walken R, Driver J, Spence C. (2003) Multimodal Visual-somatosensory Integration in Saccade Generation. Neuropsychology 41: 1-15.

Anastasio TJ and Patton PE. (2003) A two-stage unsupervised learning algorithm reproduces multisensory enhancement in a neural network model of the corticotectal system. J. Neurosci. 23: 6713-6727.

Anastasio TJ, Patton PE, Belkacem-Boussaid K. (2000) Using Bayes rule to model multisensory enhancement in the superior colliculus. Neural. Comput. 12: 1165-1187.

Bell AH, Corneil BD, Meredith MA, Munoz DP. (2001) The influence of stimulus properties on multisensory processing in the awake primate superior colliculus. Can. J. Exp. Psychol. 55: 123-132.

Ben-Yishai R, Bar O, Sompolinsky H. (1995) Theory of orientation tuning in visual cortex. Proc Natl Acad Sci U S A 92: 3844-3848.

Bermant RI and Welch RB. (1976) Effect of Degree of Separation of Visual-auditory Stimulus and Eye Position upon Spatial Interaction of Vision and Audition. Percept Mot Skills 43: 487-493.

Edwards SB, Ginsburgh CL, Henkel CK, Stein BE. (1979) Sources of Subcortical Projections to the Superior Colliculus in the Cat. J Comp Neurol 184: 309-330.

Frassinetti F, Bolognini N, Ladavas E. (2002) Enhancement of visual perception by crossmodal visuo-auditory interaction. Exp Brain Res 147: 332-343.

Frens MA and Van Opstal AJ. (1998) Visual-auditory interactions modulate saccade-related activity in monkey superior colliculus. Brain Res Bull 46: 211-224.

Frens MA, Van Opstal AJ, Van der Willigen RF. (1995) Spatial and Temporal Factors Determine Auditory-visual Interactions in Human Saccadic Eye Movements. Percept Psychophys 57: 802-816.

Huerta MF and Harting JK. (1984) The Mammalian Superior Colliculus: Studies of its Morphology and Connections. In: H.Vanegas, ed. Comparative Neurology of the Optic Tectum. New York. pp. 687-773.

Hughes HC, Reuter-Lorenz PA, Nozawa G, Fendrich R. (1994) Visual-auditory Interactions in Sensorimotor Processing: Saccades Versus Manual Responses. J Exp Psychol Hum Percept Perform 20: 131-153.

Jiang W, Jiang H, Stein BE. (2002) Two Corticotectal Areas Facilitate Multisensory Orientation Behavior. J Cogn Neurosci 14: 1240-1255.

Jiang W and Stein BE. (2003) Cortex Controls Multisensory Depression in Superior Colliculus. J Neurophysiol 90: 2123-2135.

Jiang W, Wallace MT, Jiang H, Vaughan JW, Stein BE. (2001) Two Cortical Areas Mediate Multisensory Integration in Superior Colliculus Neurons. J Neurophysiol 85: 506-522.

Kadunce DC, Vaughan JW, Wallace MT, Benedek G, Stein BE. (1997) Mechanisms of Within- and Cross-Modality Suppression in the Superior Colliculus. J Neurophysiol 78: 2834-2847.

Kadunce DC, Vaughan JW, Wallace MT, Stein BE. (2001) The Influence of Visual and Auditory Receptive Field Organization on Multisensory Integration in the Superior Colliculus. Exp Brain Res 139: 303-310.

King AJ and Palmer AR. (1985) Integration of Visual and Auditory Information in Bimodal Neurons in the Guinea Pig Superior Colliculus. Exp Brain Res 60: 492-500.

Meredith MA and Stein BE. (1986a) Spatial Factors Determine the Activity of Multisensory Neurons in Cat Superior Colliculus. Brain Res 365: 350-354.

Meredith MA and Stein BE. (1986b) Visual, Auditory, and Somatosensory Convergence on Cells in Superior Colliculus Results in Multisensory Integration. J Neurophysiol 56: 640-662.

Meredith MA and Stein BE. (1996) Spatial determinants of multisensory integration in cat superior colliculus neurons. J Neurophysiol 75: 1843-1857.

Mize RR, Withworth H, Nunes-Cardozo B, van der Want J. (1994) Ultrastructural organization of GABA in the rabbit superior colliculus revealed by quantitative postembedding immunocytochemistry. J Comp Neurol 341: 273-287.

Patton PE and Anastasio TJ. (2003) Modelling cross-modal enhancement and modality-specific suppression in multisensory neurons. Neural Comput 15: 783-810.

Patton PE, Belkacem-Boussaid K, Anastasio TJ. (2002) Multimodality in the superior colliculus: an information theoretic analysis. Brain Res Cogn Brain Res 14: 10-19.

Peck CK. (1996) Visual-auditory integration in cat superior colliculus: implications for neuronal control of the orienting response. Prog Brain Res. 112: 167-177.

Perrault Jr TJ, Vaughan JW, Stein BE, Wallace MT. (2003) Neuron-Specific Response Characteristics Predict the Magnitude of Multisensory Integration. J Neurophysiol 90: 4022-4026.

Perrault Jr TJ, Vaughan JW, Stein BE, Wallace MT. (2005) Superior Colliculus Neurons Use Distinct Operational Modes in the Integration of Multisensory Stimuli. J Neurophysiol 93: 2575-2586.

Perrott DR, Saberi K, Brown K, Strybel TZ. (1990) Auditory Psychomotor Coordination and Visual Search Performance. Percept Psychophys 48: 214-226.

Populin LC. (2005) Anesthetics change the excitation/inhibition balance that governs sensory processing in the cat superior colliculus. J Neurosci 25: 5903-5914.

Populin LC and Yin TCT. (2002) Bimodal Interactions in the Superior Colliculus of the Behaving Cat. J Neurosci 22: 2826-2834.

Schroger E and Widmann A. (1998) Speeded responses to audiovisual signal changes result from bimodal integration. Psychophysiology 35: 755-759.

Sparks DL. (1986) Translation of sensory signals into commands for control of saccadic eye movements: role of primate superior colliculus. Physiol Rev 66: 118-171.

Stanford TR, Quessy S, Stein BE. (2005) Evaluating the Operations Underlying Multisensory Integration in the Cat Superior Colliculus. J Neurosci 25: 6499-6508.

Stein BE. (1998) Neural Mechanisms for Synthesizing Sensory Information and Producing Adaptive Behaviors. Exp Brain Res 123: 124-135.

Stein BE, Huneycutt WS, Meredith MA. (1988) Neurons and behavior: the same rules of multisensory integration apply. Brain Res 448: 355-358.

Stein BE and Meredith MA. (1993) The Merging of the Senses. MIT Press, Cambridge, MA.

Stein BE, Meredith MA, Huneycutt WS, McDade L. (1989) Behavioral indices of multisensory integration: orientation to visual cues is affected by auditory stimuli. J Cogn Neurosci 1: 12-24.

Treves A. (1993) Mean-field analisys of neuronal spike dynamics. Network 4: 259-284.

Wallace MT, Meredith MA, Stein BE. (1993) Converging Influences from Visual, Auditory, and Somatosensory Cortices onto Output Neurons of the Superior Colliculus. J Neurophysiol 69: 1797-1809.

Wallace MT, Meredith MA, Stein BE. (1998) Multisensory Integration in the Superior Colliculus of the Alert Cat. J Neurophysiol 80: 1006-1010.

Welch RB and Warren DH. (1986) Intersensory Interaction. In: KR Boff, L Kaufman, JP Thomas, eds. Handbook of perception and human performance. Wiley, New York. pp. 1-36.

Wilkinson LK, Meredith MA, Stein BE. (1996) The role of anterior ectosylvian cortex in cross-modality orientation and approach behavior. Exp Brain Res 112: 1-10.

CHAPTER 2.2. A NEURAL NETWORK MODEL DESCRIBING SENSORY INTEGRATION IN THE SUPERIOR COLLICULUS

INTRODUCTION

SC neurons receive converging visual, auditory and somatosensory inputs from a variety of cortical and subcortical areas (Edwards et al., 1979; Huerta and Harting, 1984; Wallace et al., 1993). In the cat, some inputs to the SC descend from neurons in different regions of the associative cortex, in particular the anterior ectosylvian sulcus (AES), and the rostral lateral suprasylvian sulcus (rLS), while other ascending projections come from subcortical sources. These inputs are topographically organized in sensory maps, with each individual map in spatial register with the others.

Several behavioural and neurophysiological results (Jiang et al., 2001 and Jiang and Stein, 2003) demonstrated that the capability of SC neurons to integrate crossmodal sensory stimuli is strongly dependent on influences from two cortical areas (*the anterior ectosylvian sulcus (AES) and the rostral lateral suprasylvian sulcus (rLS)*). Despite the response to unimodal stimuli remains largely intact, if these cortices are temporarily deactivated, the response to cross-modal stimuli simply resembles that to the most effective of their unisensory component (Jiang et al. 2001; Jiang and Stein 2003).

Other experiments (Binns and Salt, 1996) stress the role of a class of membrane receptors (N-methyl-D-aspartate, *NMDA*, receptors) in eliciting multisensory enhancement in the Superior Colliculus neurons. These receptors are voltage dependent and could work as detectors of spatial and temporal coincident stimuli to the SC. After application of AP5, an NMDA receptor antagonist, the response to multimodal stimuli is greatly reduced, and the response to stimuli of increasing

amplitude is converted from non-linear to almost-linear (Binns and Salt, 1996). These results suggest that NMDA receptors are of great importance in multisensory enhancement, and that non-linearities of the SC neuron response, mediated by these receptors, are important for multisensory integration (Rowland et al., 2007).

This great number of results gathered in recent years on multimodal integration in the SC may now allow a theoretical formalisation via mathematical models. In particular, a model may help understanding the role of the different mechanisms involved in the SC response, and the relationships among the different inputs (cortical and non-cortical) which target to SC neurons. However, just a few models have been proposed until now.

Anastasio, Patton et al. (Anastasio, Patton, & Belkacem-Boussaid, 2000; Patton, Belkacem-Boussaid, & Anastasio, 2002; Anastasio and Patton, 2003; ; Patton & Anastasio, 2003), Colonius and Diederich (2004), Knill and Pouget (2004) recently developed some models based on the information theory, in which neurons implement the Bayes rule. These models suggest that SC neurons guarantee optimal performance for target detection, by computing the posterior probability that a target is present in their receptive field. Moreover, the models by Anastasio and Patton can account for a variety of behaviour, including the existence of cross-modal enhancement, the existence of both multimodal and unimodal neurons as well as within-modality suppression, given only the added hypothesis that inputs of the same modality have more spontaneous covariance than those of different modalities (Anastasio, Patton et al., 2003 ; Patton & Anastasio, 2003). However, these models are based on information theory rather than on neurophysiological concepts.

Recently, Rowland et al. (2007) developed a computational model in which SC neurons receive ascending and descending inputs. The model stresses the role of NMDA receptors in producing multisensory enhancement via a non-linear behaviour at the receptor level. However, it does not include spatial information on the inputs nor account for mutual relationship among neurons with different receptive fields.

The first version of our model of multisensory integration in the SC (see for details Chapter 2.1), includes a topographical organization of neurons RFs, and it can reproduce different experimental results in the literature, such as the inverse effectiveness rule, multisensory enhancement in response to cross-modal stimuli in spatial register, cross-modal and within-modal suppression in response to stimuli with disparate spatial properties (Ursino et al., 2009). That version of the model, however, accounted only for the presence of the cortical descending inputs, and so provided only a partial description of the present knowledge on the SC behaviour. Due to this limitation, the model was unable to explain SC behaviour after cortical deactivation nor it could be used to analyze SC changes during the development period in the early life, when descending inputs are probably weak and still not-organized. A more sophisticate model is required to study the relationships between non-AES and AES sources, to describe the effect of cortical deactivation on multisensory integration and formulate hypotheses on possible development changes.

Aim of this work is to present an improved version of the previous model, in which the interaction between cortical AES and not-AES inputs on multisensory SC neurons is described. The model hypothesizes the existence of a competitive mechanism between descending and ascending sources, to explain results in the neurophysiological literature. Moreover, the role of non-linearities in neuron response is emphasized, in the same line as in previous modelling works (Anastasio, Patton, & Belkacem-Boussaid, 2000; Patton, Belkacem-Boussaid, & Anastasio, 2002; Anastasio and Patton, 2003; ; Patton & Anastasio, 2003; Colonius and Diederich, 2004; Knill and Pouget, 2004; Rowland et al., 2007). Results demonstrate that the model is able to incorporate many additional experimental results (including partial or total cortical deactivation and NMDA blockade) into a comprehensive schema and to provide indications for future experimental and/or theoretical work. In perspective, it may be used to study the mechanisms at the basis of development of multisensory integration in the early life.

METHOD

General model structure

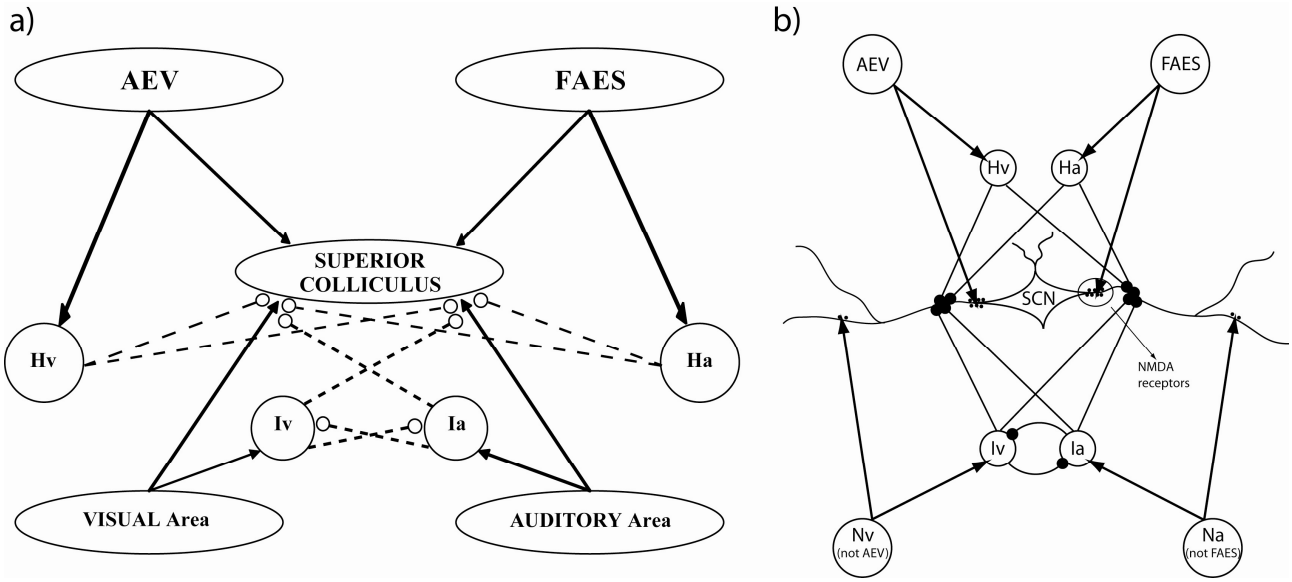


Figure 7 – *The general structure of the network (fig.1a) and its physiological counterpart (fig.1b)*. The four projection areas (AES and non-AES) make excitatory connections (arrows) with the SC and with interneurons. The interneurons work in concert to provide two competitive mechanisms based on their inhibitory synapses (black dots). Ha and Hv = interneurons receiving auditory (a) or visual (v) input from cortex; Ia, Iv = interneurons receiving auditory and visual inputs from non-AES areas.

– The model involves 4 regions of sensory input (see Fig.1), each projecting topographically to an area which represents the SC; two of these input areas represent unisensory projections from AES, specifically AEV (visual) and FAES (auditory) regions; the other represent unisensory visual and auditory inputs from non-AES (e.g. ascending) sources. Neurons in the SC area are responsible for multisensory integration. Each region is a chain of 100 neurons and its structure is arranged to keep a spatial and geometrical similarity with the external world: neurons of each area respond only to stimuli coming from a limited zone of space. Neurons normally are in a silent state (or exhibit just a mild basal activity) and can be activated if stimulated by a sufficiently strong input. Furthermore, each neuron exhibits a sigmoidal relationship (with lower threshold and upper saturation) and a first order dynamics (with a given time constant). All these areas have a topological organisation, i.e., proximal neurons respond to stimuli in proximal position of space.

– Each element of the unisensory areas has its own receptive field (RF) that can be partially superimposed on that of the other elements of the same area. Elements of the same

unisensory area interact via lateral synapses, which can be both excitatory and inhibitory. These synapses are arranged according to a Mexican hat disposition (i.e., a circular excitatory region surrounded by a larger inhibitory annulus).

– Elements in the superior colliculus receive inputs from neurons in the 4 unisensory areas (visual and auditory) whose RFs are located in the same spatial position. Moreover, elements in the SC are connected by lateral synapses, which also have a Mexican hat disposition.

– The model also includes two different competitive mechanisms realized by means of 4 different populations of inhibitory interneurons. These interneurons do not act directly on the input of the multimodal neurons in the colliculus, but modulate the strength of ascending excitatory synapses.

    1. The first mechanism aims at to mimicking the effect of the cortex on non-AES sources, assuming that, when the cortex is active, it dominates on non-cortical inputs. This is realized through two populations of cortical inhibitory interneurons (H_v and H_a in Fig. 1), which receive inputs from FAES neurons and AEV neurons respectively, and, if stimulated, inhibit stimuli coming from non-AES (e.g. ascending) sources.
    2. the second competitive mechanism is realized through the interaction between the inhibitory interneurons I_v and I_a, so that the stronger unisensory input converging from non-AES regions overwhelms the weaker.

Note that these regions of interneurons are arranged as a chain of 100 elements: each interneuron receives its input only from one excitatory neuron belonging to the corresponding unisensory input regions.

Mathematical description

The following notations will be used throughout the manuscript.

A quantity which refers to a single neuron will be denoted with one superscript (say *h*) and one subscript (say *i*). The superscript represents the region the neuron belongs to. The subscript ($i = 1,2, \ldots, N$) denotes the spatial position of its receptive field. In particular, the symbol u_i^h will be used to represent neuron input, and z_i^h to represent the neuron activity (normalized between 0 and 1),

The following symbols will be used as superscripts to denote the nine different regions of the model (see Fig. 1):

Ca (cortical auditory): FAES neurons:

Cv (cortical visual): AEV neurons;

Na: not-FAES auditory neurons;

Nv: (not-cortical visual): not-AEV visual neurons;

Hv: cortical inhibitory interneurons which receive inputs from AEV;

Ha: cortical inhibitory interneurons which receive inputs from FAES;

Ia: inhibitory interneurons which receive inputs from not-FAES;

Iv: inhibitory interneurons which receive inputs from not-AEV;

Sm: (superior colliculus multimodal): multimodal neurons in the superior colliculus

Each excitatory synapse linking two neurons in different regions, both at the same position *i*, will be denoted with the symbol $W_i^{h,k}$, where the first superscript (*h*) represents the target region and the second superscript (*k*) the region from where synapses originate. Each inhibitory synapse originating from an interneuron at position *i* will be denoted with the symbol $K_i^{h,k}$, where the meaning of symbols is the same as explained above. The lateral (excitatory or inhibitory) synapses linking two neurons in the same region (but with different spatial position) will be denoted with $L_{i,j}^h$, where *h* is the region and the subscripts *i* and *j* represent the position of the post-synaptic neuron and of the pre-synaptic neuron, respectively.

The receptive fields of unisensory areas.

In the present version we assume that each area is composed by a chain of *N* neurons (with *N* = 100). Such a limited number of neurons was chosen to reduce the computational complexity for computer implementation. The use of a monodimensional arrangement has a physiological reliability since SC neurons are much more specific in the azimuthal direction than in the vertical one (Stein, 1993; Stein, 1976). Neurons in each area differ in the position of their receptive field by 1.8°. Hence, each area covers 180° in the visual, acoustic or multisensory space. In the following, we will denote with x_i the center of the RF of a generic neuron *i*. Hence, we can write:

$x_i = 1.8 \cdot i$ deg (i = 1, 2,…, 100)

The receptive field (say $R_i^s(x)$) of neuron *i* in the unisensory area *s* is described with a gaussian function. Hence,

$$R_i^s(x) = R_0^s \cdot e^{-\frac{\left[(x_i - x)^2\right]}{2\cdot(\sigma_R^s)^2}} \qquad s = \text{Ca, Cv, Na, Nv;} \tag{1}$$

where the symbols *x* represent a generic coordinate in space, σ_R^S is the standard deviation of the Gaussian function (three standard deviations approximately cover the overall RF) and R_0^S is a parameter which sets the strength of the response.

According to Eq. (1), a stimulus presented at the position *x* excites not only the neuron centered in that zone, but also the proximal neurons whose receptive fields cover such position.

The sensory input r_i^s, that reaches the neuron *i* in presence of a stimulus, is computed as the inner product of the stimulus and the receptive field. We can write:

$$r_i^s(t) = \int_x R_i^s(x) \cdot i^s(x,t) dx \cong \sum_x R_i^s(x) \cdot i^s(x,t) \Delta x \tag{2}$$

where, $i^s(x,t)$ is the external sensory stimulus presented at the coordinate *x* and at time *t*, and the right hand member of Eq. (2) means that the integral has been computed with the histogram rule (in this work, Δx = 1.8 deg).

The activity in the unisensory areas

Unisensory neurons can be stimulated not only by external inputs, but also through the connections with other elements in the same area.

The input that a unisensory neuron gets from other elements of the same area is represented by the quantity l_i^s, defined as:

$$l_i^s(t) = \sum_{\mathrm{j}} L_{i,j}^s \cdot z_j^s(\mathrm{t}) \qquad \mathrm{s} = \mathrm{Ca, Cv, Na, Nv}; \tag{3}$$

where $z_j^s(\mathrm{t})$ is the *j*-neuron's activity (described below) and $L_{i,j}^s$ is the strength of the synaptic connection from the pre-synaptic neuron at the position *j* to the post-synaptic neuron at the position *i*. These synapses are symmetrical and arranged according to a "Mexican hat" function:

$$L_{i,j}^s = L_{ex}^s \cdot e^{-\frac{\left[d_x{}^2\right]}{2\cdot\left(\sigma_{ex}^s\right)^2}} - L_{in}^s \cdot e^{-\frac{\left[d_x{}^2\right]}{2\cdot\left(\sigma_{in}^s\right)^2}}. \qquad \mathrm{s} = \mathrm{Ca, Cv, Na, Nv}; \tag{4}$$

In this equation, L_{ex}^s and σ_{ex}^s define the excitatory Gaussian function, while L_{in}^s and σ_{in}^s the inhibitory one, and d_x represents the distance between the pre-synaptic and post-synaptic neurons. To avoid undesired border effects, synapses have been realized with a circular structure so that every neuron of each area receives the same number of side connections. This is realized assuming the following expression for the distance:

$$d_x = \begin{cases} |i-j| & \text{if } |i-j| \le N/2 \\ N - |i-j| & \text{if } |i-j| > N/2 \end{cases} \tag{5}$$

According to the previous description, the total input (say $u_i^s(t)$) received by a unisensory neuron at position *i* is computed as follows,

$$u_i^s(t) = r_i^s(t) + l_i^s(t); \qquad \mathrm{s} = \mathrm{Ca, Cv, Na, Nv}. \tag{6}$$

This is the sum of two components: r_i^s, that represents the external sensory input; and l_i^s, coming from the intra-area synapses.

Finally, neuron activity is computed from its input, through a static sigmoidal relationship and a first-order dynamic. This is described via the following differential equation:

$$\tau_s \cdot \frac{d}{dt} z_i^s(t) = -z_i^s(t) + \varphi\left(u_i^s(t)\right) \quad ; \qquad s = Ca,\ Cv,\ Na,\ Nv \tag{7}$$

τ_s is the time constant, which determines the speed of the answer to the stimulus, and $\varphi\left(u^s(t)\right)$ is a sigmoidal function. The latter is described by the following equation:

$$\varphi\left(u^s(t)\right) = \frac{1}{1+e^{-\left(u^s(t)-\vartheta^s\right)p^s}} \quad ; \qquad s = Ca,\ Cv,\ Na,\ Nv \tag{8}$$

where ϑ^s defines the input value at which neuron activity is half the maximum (central point) and p^s sets the slope at the central point.

Such function identifies three regions of work, depending on the intensity of the input: the sub-threshold behaviour of a neuron, a linear region (around ϑ^s), and a saturation region. According to the previous equation, the maximal neuron activity is conventionally set at 1 (i.e., all neuron activities are normalized to the maximum).

The competitive mechanisms

As said before, the network includes two competitive mechanisms realized by means of four areas of inhibitory interneurons. These populations are unisensory, i.e. each area of interneurons is stimulated only by inputs coming from one of the four unisensory input regions.

The first mechanism reproduces the effect of AES cortex on the ascending paths. It involves the populations *Hv* and *Ha* stimulated by AEV cortex or FAES cortex, respectively (accordingly to V. Fuentes-Santamaria et al., 2007).

Inputs received by these interneurons are computed as follows:

$$u_i^{Ha}(t) = W_i^{Ha,Ca} \cdot z_i^{Ca}(t) \ ; \tag{11}$$

$$u_i^{Hv}(t) = W_i^{Hv,Cv} \cdot z_i^{Cv}(t) \ ; \tag{12}$$

where $W_i^{Ha,Ca}$ and $W_i^{Hv,Cv}$ represent intensities of synaptic connections between a pre-synaptic neuron in the cortex and the corresponding element in the relative interneurons area, and z_i^{Ca}, z_i^{Cv} are the activities of the unisensory neuron in FAES cortex and in AEV cortex.

The second mechanism involves the other two areas of interneurons, *Iv* and *Ia*. Elements of these populations are stimulated by neurons belonging or to the visual not-AEV area, or to the auditory not-FAES area. The purpose of this mechanism is to realize a competition between the two inputs coming from not-AES sources so that the stronger overwhelms the weaker. To this end, interneurons of these regions receive not only a stimulus from the corresponding unisensory region, but also a reciprocally inhibitory input from the ascending interneuron of the other sensory modality.

Inputs received by these interneurons are computed as follows:

$$u_i^{Ia}(t) = W_i^{Ia,Na} \cdot z_i^{Na}(t) - K_i^{Ia,Iv} \cdot z_i^{Iv}; \quad (13)$$

$$u_i^{Iv}(t) = W_i^{Iv,Nv} \cdot z_i^{Nv}(t) - K_i^{Iv,Ia} \cdot z_i^{Ia}; \quad (14)$$

where symbols $W_i^{h,k}$ and $K_i^{h,k}$ represent excitatory and inhibitory synapses linking different regions, and $z_i^{Na}, z_i^{Nv}, z_i^{Ia}$ and z_i^{Iv} are neuron activities at position *i.*

The activities are computed from the inputs through a static sigmoidal relationship and a first-order dynamics analogous to Eqs. (7) and (8) with superscripts *Ha*, *Hv*, *Ia* and *Iv*.

The activity in the multisensory area

Multisensory neurons in the superior colliculus receive inputs from neurons in the unisensory regions whose RFs are located in the same position, *i*. Furthermore, excitation from unisensory areas is modulated by inhibitory interneurons in the same spatial position. This choice has been adopted since, according to experimental data, the auditory and visual RFs of a multisensory neuron are in spatial register (Kadunce et al., 2001; Meredith and Stein, 1996), i.e., they represent similar

regions in space. Finally, neurons in the superior colliculus also receive lateral synapses form other elements in the same area.

The auditory and visual inputs coming from FAES and AEV regions are not inhibited, hence they are simply computed as the product of the synaptic weight and the activity of the upstream unisensory neuron. We can write

$$u_i^{Sm,Ca} = W_i^{Sm,Ca} \cdot z_i^{Ca} \tag{15}$$

$$u_i^{Sm,Cv} = W_i^{Sm,Cv} \cdot z_i^{Cv} \tag{16}$$

The overall auditory and visual inputs, coming from not-FAES and not-AEV region, are computed with a more complex equation, since the excitatory activity coming from the upstream unimodal neuron is modulated by inhibitory synapses. In particular, the excitatory activity is reduced by activity from both cortical interneurons (areas *Ha* and *Hv*) and from the activity of the not-cortical interneuron with different sensory modality (either *Ia* or *Iv*). We have

$$u_i^{Sm,Na}(t) = W_i^{Sm,Na} \cdot z_i^{Na}(t) \cdot \left(1\text{-}K_i^{Sm,Ha} \cdot z_i^{Ha}(t)\right) \cdot \left(1\text{-}K_i^{Sm,Hv} \cdot z_i^{Hv}(t)\right) \cdot \left(1\text{-}K_i^{Sm,Iv} \cdot z_i^{Iv}(t)\right) \tag{17}$$

$$u_i^{Sm,Nv}(t) = W_i^{Sm,Nv} \cdot z_i^{Nv}(t) \cdot \left(1\text{-}K_i^{Sm,Ha} \cdot z_i^{Ha}(t)\right) \cdot \left(1\text{-}K_i^{Sm,Hv} \cdot z_i^{Hv}(t)\right) \cdot \left(1\text{-}K_i^{Sm,Ia} \cdot z_i^{Ia}(t)\right) \tag{18}$$

where the meaning of symbols is the same as explained above.

Finally, a multimodal neuron also receives lateral input (say $l_i^{Sm}(t)$), form other neurons in the same area. The latter term is computed as follows

$$l_i^{Sm}(t) = \sum_j L_{i,j}^{Sm} \cdot z_j^{Sm}(t) \,. \tag{19}$$

We assumed that lateral synapses in the multisensory area have a Mexican hat disposition. The equation is:

$$L_{i,j}^{Sm} = L_{ex}^{Sm} \cdot e^{-\frac{\left[d_x^2\right]}{2\cdot\left(\sigma_{ex}^{Sm}\right)^2}} - L_{in}^{Sm} \cdot e^{-\frac{\left[d_x^2\right]}{2\cdot\left(\sigma_{in}^{Sm}\right)^2}} \tag{20}$$

where the distance d_x has the same expression as in Eq. (5), and parameters L_{ex}^{Sm}, L_{in}^{Sm}, σ_{ex}^{Sm} and σ_{in}^{Sm} set the strength and spatial disposition of lateral synapses.

Finally, the overall input, $u_i^m(t)$, to a multisensory neuron can be computed as follows

$$u_i^{Sm}(t) = u_i^{Sm,Ca}(t) + u_i^{Sm,Cv}(t) + u_i^{Sm,Na}(t) + u_i^{Sm,Nv}(t) + I_i^{Sm}(t) \tag{21}$$

The activity of a multisensory neuron is computed from its input by using equations similar to Eqs. (7) and (8).

Parameter assignment

The value of all model parameters is shown in Tab. I.

Table 5 – Parameter values used in the present model.

AEV Area	
N	100
σ_R^{Cv}	1 (1.8°)
R_0^{Cv}	1
L_{ex}^{Cv}	5.4
σ_{ex}^{Cv}	2.8 (5.04°)
L_{in}^{Cv}	4.72
σ_{in}^{Cv}	7.4 (13.32°)
τ_{Cv}	3 ms
$\vartheta_{\cdot}^{Cv}$	6
p^{Cv}	0.3
$W_i^{Hv,Cv}$	15 (0 without NMDA)
$W_i^{Sm,Cv}$	7.7 (1 without NMDA)

FAES Area	
N	100
σ_R^{Ca}	1.5 (2.7°)
R_0^{Ca}	1
L_{ex}^{Ca}	4.2
σ_{ex}^{Ca}	2.8 (5.04°)
L_{in}^{Ca}	3.55
σ_{in}^{Ca}	7.4 (13.32°)
τ_{Ca}	3 ms
ϑ^{Ca}	6
p^{Ca}	0.3
$W_i^{Ha,Ca}$	14
$W_i^{Sm,Ca}$	5.9

Superior Colliculus Area	
N	100
L_{in}^{Sm}	3.3
L_{ex}^{Sm}	3.8
σ_{ex}^{Sm}	3.5 (6.3°)
σ_{in}^{Sm}	6.2 (11.16°)
τ_{Sm}	3 ms
ϑ^{Sm}	6
p^{Sm}	0.3

Not-AEV Area	
N	100
σ_R^{Nv}	1 (1.8°)
R_0^{Nv}	1
L_{ex}^{Nv}	5.4
σ_{ex}^{Nv}	2.8 (5.04°)
L_{in}^{Nv}	4.72
σ_{in}^{Nv}	7.4 (13.32°)
τ_{Nv}	3 ms
$\vartheta_{\cdot}^{Nv}$	6
P^{Nv}	0.3
$W_i^{Iv,Nv}$	15
$W_i^{Sm,Nv}$	5

Not-FAES Area	
N	100
σ_R^{Na}	1.5 (2.7°)
R_0^{Na}	1
L_{ex}^{Na}	4.2
σ_{ex}^{Na}	2.8 (5.04°)
L_{in}^{Na}	3.55
σ_{in}^{Na}	7.4 (13.32°)
τ_{Na}	3 ms
ϑ^{Na}	6
P^{Na}	0.3
$W_i^{Ia,Na}$	14
$W_i^{Sm,Na}$	4

Visual Cortical Interneurons (H_v)	
N	100
τ_{Hv}	3 ms
$\vartheta_{\cdot}^{Hv}$	3
p^{Hv}	1
$K_i^{Sm,Hv}$	1

Visual Interneurons (I_v)	
N	100
τ_{Iv}	3 ms
$\vartheta_{\cdot}^{Iv}$	3
p^{Iv}	1
$K_i^{Ia,Iv}$	33
$K_i^{Sm,Iv}$	1

Auditory Cortical Interneurons (H_a)	
N	100
τ_{Ha}	3 ms
ϑ^{Ha}	3
p^{Ha}	1
$K_i^{Sm,Ha}$	1

Auditory Interneurons (I_a)	
N	100
τ_{Ia}	3 ms
ϑ^{Ia}	3
p^{Ia}	1
$K_i^{Iv,Ia}$	33
$K_i^{Sm,Ia}$	1

These parameters have been assigned starting from data in the literature according to the main criteria summarized below.

Receptive fields: σ_R^{Cv} , σ_R^{Nv} and σ_R^{Ca} , σ_R^{Na} have been given so that the receptive fields of the visual are approximately 10 deg in diameter, and those of acoustic neurons approximately 10-15 deg in diameter, according to data reported in (Kadunce et al., 2001). R_0^{Cv}, R_0^{Nv} and R_0^{Ca} , R_0^{Na} are set to 1, to establish a scale for the inputs generated by the external stimuli.

Parameters of the excitatory neurons (sigmoidal relationships and time constants): for the sake of simplicity, these parameters have been chosen equal for all neurons, independently of the respective area. All together they are responsible for eliciting the activity in SC neurons. The central abscissa of neurons and receptors, ϑ^S, has been assigned to have small neuron activity in basal condition (i.e., without any external stimulus). The slope of the sigmoidal relationships, p^S, has been assigned to have a smooth transition from silence to saturation in response to unimodal and cross-modal inputs (Perrault Jr et al., 2005). The time constant, describing neurons dynamics, agrees with values (a few milliseconds) normally used in deterministic mean-field equations (Ben-Yishai et al., 1995). In particular, this value can be chosen significant smaller than the membrane time constant (Treves, 1993).

Parameters of the inhibitory interneurons (sigmoidal relationships and time constants): Also in this case parameters have been given the same values for all inhibitory interneurons, independently of their area. The slope, p^S, and the central abscissa, ϑ^S, of the sigmoidal relationships have been assigned to have a fast transition from silence to saturation in response to inputs coming from unisensory areas, but a small basal activity in the absence of any external stimulation. This allows the implementation of a strong competitive mechanism even in the presence of a moderate stimulation. The time constant is the same as for the excitatory interneurons.

Lateral synapses in unimodal areas: Parameters which establish the extension and the strength of lateral synapses in the unimodal areas (i.e., L_{ex}^s L_{in}^s, σ_{ex}^s and σ_{in}^s) have been assigned to simultaneously satisfy several criteria: i) the presence of an external stimulus produces an activation bubble of neurons which approximately coincide with the dimension of the receptive

field; ii) according to data reported in Kadunce et al. (Kadunce et al., 1997) we assumed that an inhibitory area surrounds the activation bubble; iii) inhibition strength must be strong enough to avoid instability, i.e., an uncontrolled excitation which propagates to the overall area; iv) stimulating the suppressive region with a second stimulus can induce a within-modality suppression as high as 50% (Kadunce et al., 1997).

Lateral synapses in the superior colliculus: Parameters which establish the extension and the strength of lateral synapses in the SC areas (i.e., L_{ex}^{Sm} L_{in}^{Sm}, σ_{ex}^{Sm} and σ_{in}^{Sm}) have been assigned to warrant that two cross-modal stimuli inside the receptive field cause enhancement (Stein and Meredith, 1993), two unimodal stimuli inside the RF cause no enhancement (or even a marginal suppression at the boundary, according to data reported in Alvarado et al., 2007a,b), and two cross-modal or unimodal stimuli placed at far positions (i.e., the first inside the RF, the second outside the RF) cause a significant suppression, in agreement with data reported in Kadunce et al. (Kadunce et al., 1997; Kadunce et al., 2001).

Connections between AES subregions and SC: The parameters of feedforward connections from the unisensory AES areas to the superior colliculus (i.e., $W^{Sm,Cv}$ and $W^{Sm,Ca}$) have been set in order to: a) have a significant multisensory enhancement; b) have a greater dynamical range (i.e., the excursion from sub-threshold to saturation activity) of multisensory neurons in response to cross-modal stimuli, compared with unimodal stimuli (i.e., a single stimulus cannot lead the SC neuron to saturation) (Perrault et al., 2005). Furthermore, in this work we assumed that the effect of a visual stimulus on the SC neuron is moderately greater compared with the effect of an auditory stimulus (Perrault et al., 2005).

Connections between not-AES subregions and SC: The parameters of feedforward connections from the unisensory not-AES areas to the superior colliculus (i.e., $W^{Sm,Nv}$ and $W^{Sm,Na}$) have been set to have unisensory responses 50% depressed, when AES is deactivated, compared with those

elicited by AES stimuli (a normal condition), in agreement with data reported in Alvarado et al. (Alvarado et al., 2007a,b).

Competitive mechanisms: Connections from neurons in the unisensory AES areas to cortical interneurons (i.e., $W^{Hv,Cv}$ and $W^{Ha,Ca}$) are given so that even a moderate activity in the AES cortex can lead the interneuron close to saturation. Similarly, connections from unisensory neurons in not-AES regions and interneurons (i.e., $W^{Iv,Nv}$ and $W^{Ia,Na}$) are given so that the corresponding interneuron is lead close to saturation if the corresponding not-AES region is moderately active. Parameters which establish the strength of inhibitory influence on the excitatory synapses (i.e., parameters K^s in Eqs. 17 and 18) have been set to 1; in this manner, when the interneuron is in saturation, the target excitatory synapses is completely inhibited. Finally, inhibitory synapses $K^{Iv,Ia}$ and $K^{Ia,Iv}$ in Eqs. 13 and 14, have been set to have a strong winner takes all competition between the two interneurons in the ascending route; hence, during cortical deactivation, the stronger ascending unisensory input overwhelms the weaker.

RESULTS

Effect of cortical deactivation

A first set of simulations has been performed to compare the behavior of the SC in response to unimodal and crossmodal stimuli of different intensities, first with the intact cortex, and then after a total or partial deactivation of AES. Cortex deactivation has been simulated by assigning a value equal to zero to all output signals exiting from the deactivated area.

Results obtained with the intact model are shown in Fig. 2a. In response to unimodal auditory or visual stimulation of increasing intensities, SC neurons exhibit a progressive augmentation in their response, with saturation at 0.3-0.4 (i.e., about 30-40% of the maximal SC response). Model parameters were set so that the visual response is moderately higher than the auditory one, in

agreement with data reported in (Perrault et al., 2005). Saturation in the SC depends on saturation in neurons of the unimodal areas and on the values used for the synapses which link unimodal areas to the SC.

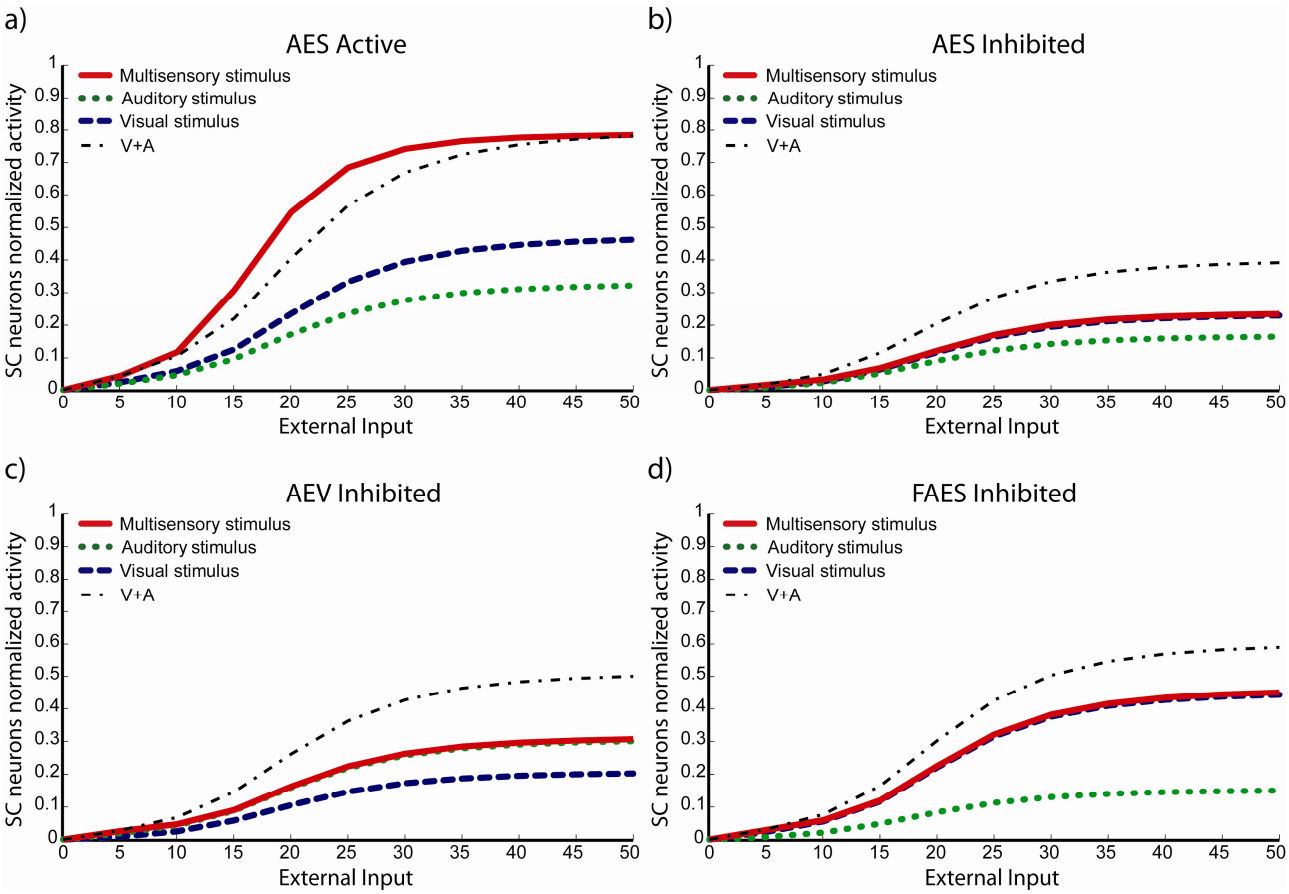


Figure 8 – *Behavior of the network as function of AES cortex.* These figures show the activity of SC neurons in response to different inputs with AES cortex active (fig.2a) or inhibited, fully (fig.2b) or only partially (AEV inhibited, fig.2c, FAES inhibited, fig.2d). If the AES is totally inhibited (Fig. 2b) the SC shows no multisensory integration, the unisensory responses are reduced by about 50% and the response to two cross-modal stimuli looks like the stronger unisensory one. If just the AEV cortex is inhibited, the SC presents a normal response to an auditory stimulation, but the response to a unimodal visual stimulation is reduced by about 50% compared to that produced when the AEV cortex is active. The multisensory response looks like the stronger one (in this case the auditory one). In fig.2d FAES is inhibited: the SC response to a visual stimulation is unaffected whereas the response to an auditory stimulus is depressed compared with the intact case; multisensory stimulation elicits a response similar to the visual one. In all simulations the activity was assessed by stimulating the model with auditory (dotted line), visual (dashed line) and multisensory (solid line) inputs, at various intensities. The stimuli were presented in the center of the RF of the observed SC neuron. Note the loss of multisensory integration when AES is deactivated even partially. Multisensory integration capability needs both AES subregions active.

Cross-modal stimulation evokes a much greater response of SC neurons, which is superadditive at small intensities of the input stimuli, and becomes just additive at high intensities. The latter result agrees with the inverse effectiveness rule (Meredith and Stein 1986b; Perrault et al. 2003, 2005; Stanford et al. 2005; Stein and Meredith 1993; Wallace et al. 1998).

Two relevant changes are evident in the SC behavior after total deactivation of cortical areas (Fig. 2b). First, the unimodal responses are smaller than in the intact case, with upper saturation at about 0.1-0.2 of the maximum activity (i.e., the reduction in the unimodal response is greater than 50%, accordingly with data presented in Alvarado et al., 2007b, see Figure 1 and 2). Second, cross-modal integration is completely lost: the response to paired auditory and visual stimuli is just trivially greater than the stronger of the two unimodal responses: the highest Multisensory Enhancement computed by the model with AES deactivated is equal to 6.3%. These results agree with those reported in Alvarado et al. (2007a,b) and in Jiang et al. (2001): after a complete deactivation of AES the response to a cross-modal stimulation is comparable to the response elicits by the most effective unisensory component (see for example fig. 8 in Jiang et al., 2001, in which the Response Enhancement % is between +9% and -14%).

Further simulations have then been performed assuming deactivation of the visual cortex (AEV) only (Fig. 2c) and of auditory cortex (FAES) only (Fig. 2d). Results show a significant reduction of the unimodal response (more than 50%) in the modality affected by the deactivation procedure, whereas the other unimodal response is almost unchanged. In these conditions too the multimodal integration is lost and the multimodal response is undistinguishable from the stronger unimodal one.

Results reported in Fig. 2b (total deactivation of the cortex) have been obtained assuming a strong competition between the two ascending paths, so that the weaker stimulus is inhibited and only the stronger can significantly affect the SC. A different behaviour can be obtained assuming a weaker competition.

Fig.3 shows results of a sensitivity analysis, in which weaker competition is simulated by progressively reducing the strength of the inhibitory synapses between the interneurons in the ascending path (i.e., parameters $K_i^{Ia,Iv}$ and $K_i^{Iv,Ia}$ in Eqs. 13 and 14) during total cortical deactivation. In case of strong competition, the SC response to cross modal stimuli resembles the response to the stronger unisensory stimulus. Conversely, assuming weak competition, both interneurons I_a and I_v display non-zero activity and inhibit excitation from the ascending path of the

different sensory modality. As a consequence, the SC response to cross-modal stimulation becomes even smaller than the stronger individual unisensory response. This apparently paradoxical result, which is a consequence of the supposed competition in the ascending routes, finds several experimental confirms. Cases of multisensory responses weaker than the dominant unisensory response are reported, for instance, in Jiang et al. (2001, see Fig. 9) after AES deactivation.

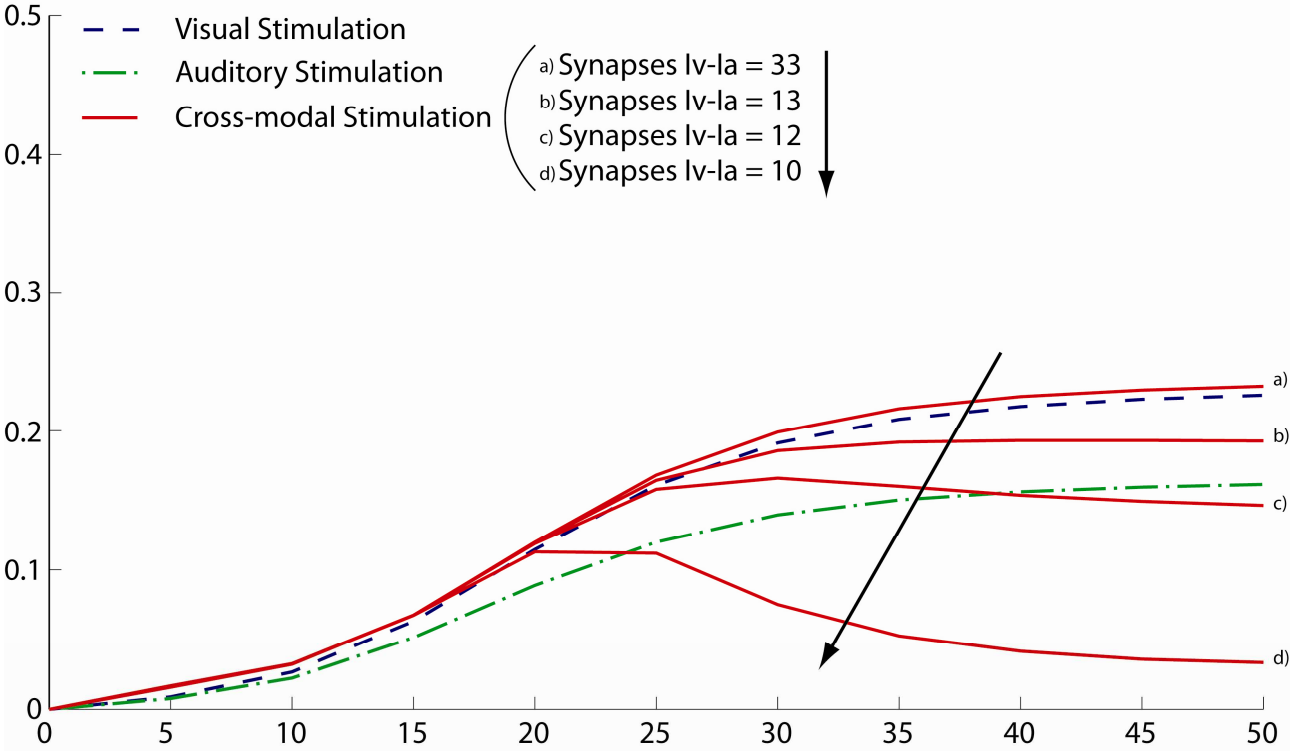


Figure 9 – *Sensitivity analysis of the strength of inhibitory competition in the ascending path* – The figure shows the activity of SC neurons (continuous lines) in response to different cross-modal inputs during total AES deactivation (the same case as in fig.2b) and assuming a different strength for the inhibitory synapses between the interneurons in the ascending path (i.e., parameters $K_i^{Ia,Iv}$ and $K_i^{Iv,Ia}$ in Eqs. 13 and 14). The responses to unimodal (auditory or visual) stimulation are also shown for comparison (dashed lines). In case of strong competition ($K_i^{Ia,Iv}$ and $K_i^{Iv,Ia}$ greater than 15), the SC response to cross modal stimuli resembles the response to the stronger unisensory stimulus. Conversely, assuming weak competition ($K_i^{Ia,Iv}$ and $K_i^{Iv,Ia}$ smaller than 12-13) the SC response to cross-modal stimulation becomes smaller than the stronger individual unisensory response.

Effect of NMDA deactivation

A few experimental results suggest that multisensory integration in the SC depends on the presence of NMDA receptors (Binns and Salt, 1996). In particular, the responses to cross-modal stimuli are consistently reduced during application of AP5, an NMDA receptor antagonist. Moreover, the response to unimodal visual stimuli are greatly reduced, whereas inconsistent results

are reported for what concerns the response to unimodal auditory stimuli. These results may reflect a different functional role of NMDA receptors in synaptic transmission of auditory and visual response in the SC (Binns and Salt, 1996).

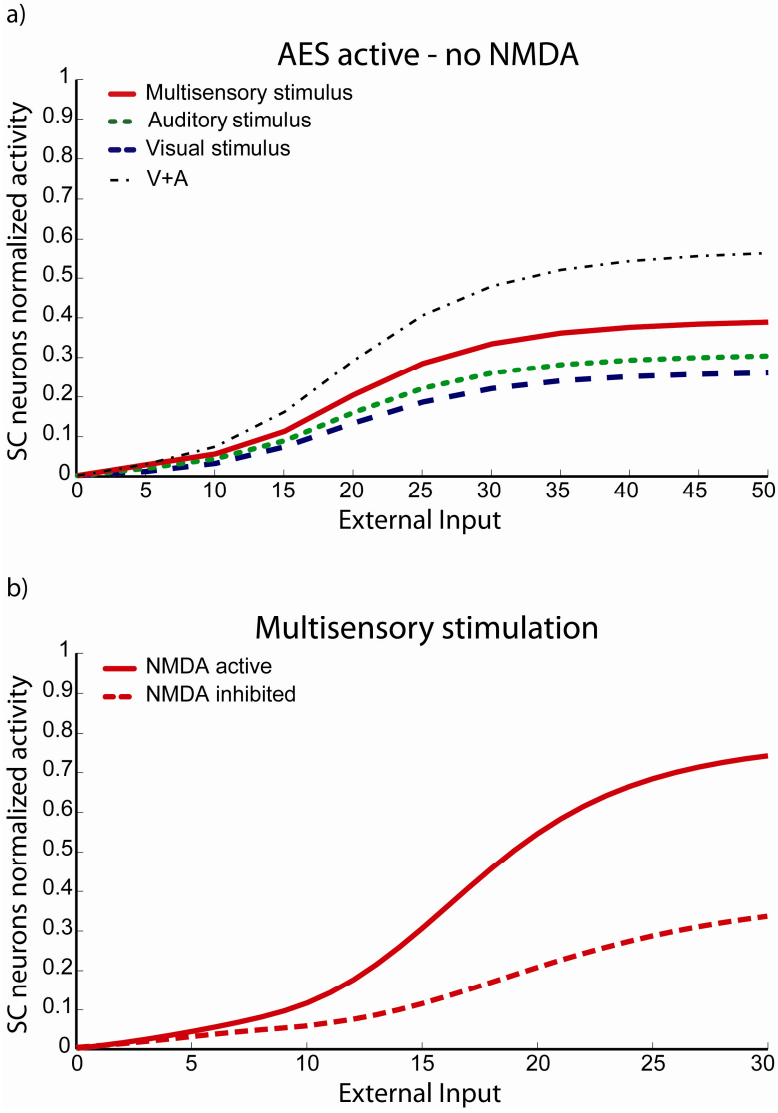


Figure 10 – *Behavior of the network with NMDA receptors deactivated*. The upper panel shows the activity of SC neurons after deactivation of NMDA receptors, in the same simulations as in Fig. 2. Deactivation of NMDA receptors causes a 43% decrease in the unimodal response to visual stimuli, whereas it barely influences auditory response (-7%). The multisensory response is significantly reduced too and results lower than the sum of unisensory responses at every input intensity (subadditivity). The lower panel compares the multimodal responses in the intact case and after NMDA deactivation. It is worth noting that the characteristic becomes quite linear after deactivation.

In order to simulate data by Binns and Salt (1996) we assumed that deactivation of NMDA receptors by AP5 greatly reduces the strength of all synapses which exit from cortical area AEV (the reduction is reported in Tab. 1) but does not significantly affect synapses exiting from area FAES. Results are shown in Fig. 4a as a function of the input intensity. By comparing this figure with the unaffected case (i.e., Fig. 2a) one can observe that simulation of NMDA inactivation causes a significant reduction in the visual response, a moderate reduction in the auditory response, and a strong reduction in the multisensory response. In particular, a mild cross-modal enhancement still occurs, but it is now sub-additive. The cross-modal response to two paired auditory and visual stimuli is just scarcely greater than the unimodal auditory response. These results agree fairly well with those reported by Binns and Salt (1996).

Fig. 4b compares the cross-modal response of the SC before and after NMDA deactivation. Results clearly show that the characteristic "neuron activity vs. stimulus intensity" is strongly non-linear in the intact condition, but becomes quite linear after NMDA deactivation. This results agree with data shown in (Binns and Salt, 1996; Rowland et al., 2007).

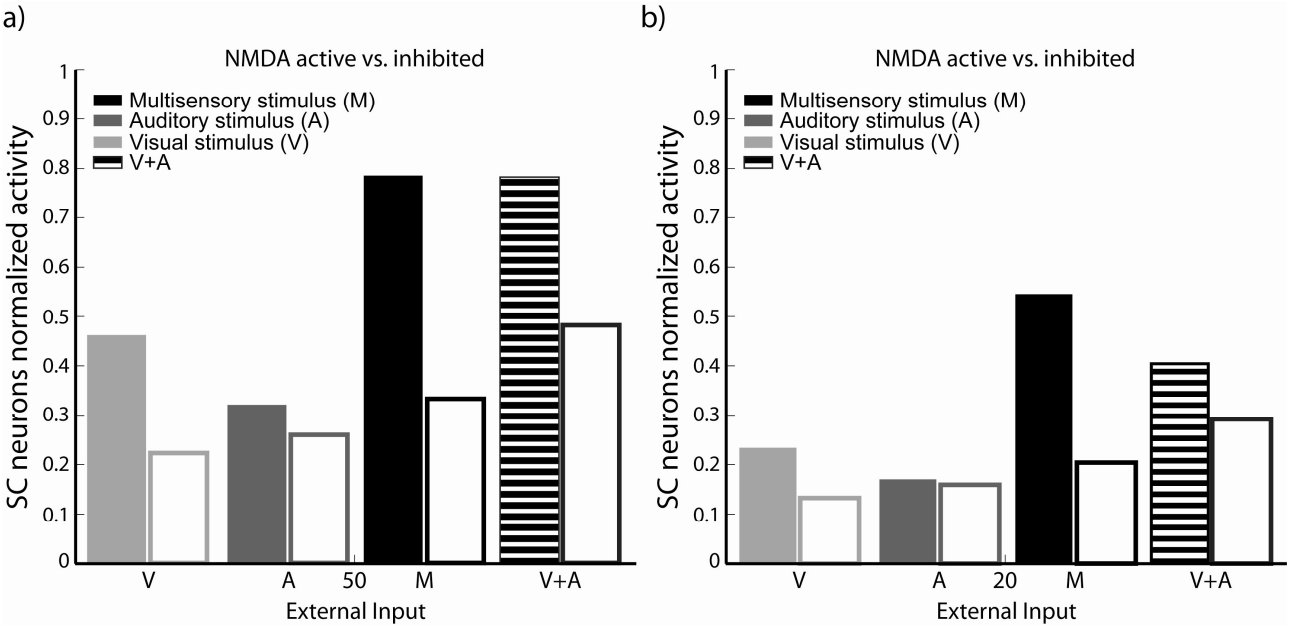


Figure 11 – ***Unisensory and multisensory responses with NMDA receptors active or inhibited.*** **Activities were assessed by presenting to the network auditory (dark-grey bars), visual (light-grey bars) and multisensory (black bars) inputs at two different intensities (I=50, fig.5.a, i.e., close to saturation; I=20, fig.5b, i.e. just above threshold) at the center of the RF both with NMDA receptors active (filled bars) and inhibited (empty bars). It is worth noting that the visual response is more affected (50%) by the NMDA inhibition than the auditory one, and the cross-modal response is reduced more than the sum of the unimodal stimuli.**

In order to allow a more direct comparison between model results and real data, Fig. 5 shows the reduction in the SC response after NMDA deactivation in the three cases of visual unimodal, auditory unimodal and cross-modal stimulation. The decrease in the sum of the two unimodal responses is also reported. Model results are given at two different levels of the input stimuli: high intensity close to saturation of the unimodal neurons (Fig. 5a); moderate intensity, when the unimodal neurons work just above threshold (Fig. 5b). The data by Binns and Salt agree quite well with those obtained with the model at high input intensities. In particular, the simulated effect of AP5 reduces the visual response by 43.4% in the model (45 ± 9% in Binns and Salt). The response to the auditory stimulus was reduced by 6.7% in the model (-4 ± 21% in Binns and Salt). The response to the combined stimuli was reduced by 62.6% in the model (59 ± 7% in Binns and Salt)

while the sum of the single modality response was reduced by 27.9% in the model (26 ± 10% in Binns and Salt).

Inputs with different spatial position

An important characteristic of SC neurons is the way they integrate stimuli coming from different spatial positions. If a second stimulus is applied outside the RF of an SC neuron, it fails to produce enhancement or may even cause a depression of the SC response (Meredith and Stein 1986b, 1996; Meredith et al. 1987; Stein et al. 1993; Stein and Wallace 1996; Wallace et al. 1996, 1998). This phenomenon occurs both in case of two unimodal stimuli with disparate spatial position (unimodal depression) and in case of spatially disparate cross-modal stimuli (cross-modal depression).

In order to study this phenomenon, the model was stimulated with two simultaneous stimuli, the first located at the center of the RF of the target SC neuron, the second at a given distance from the center. Both stimuli have a large intensity, i.e., they can lead unimodal neurons close to saturation.

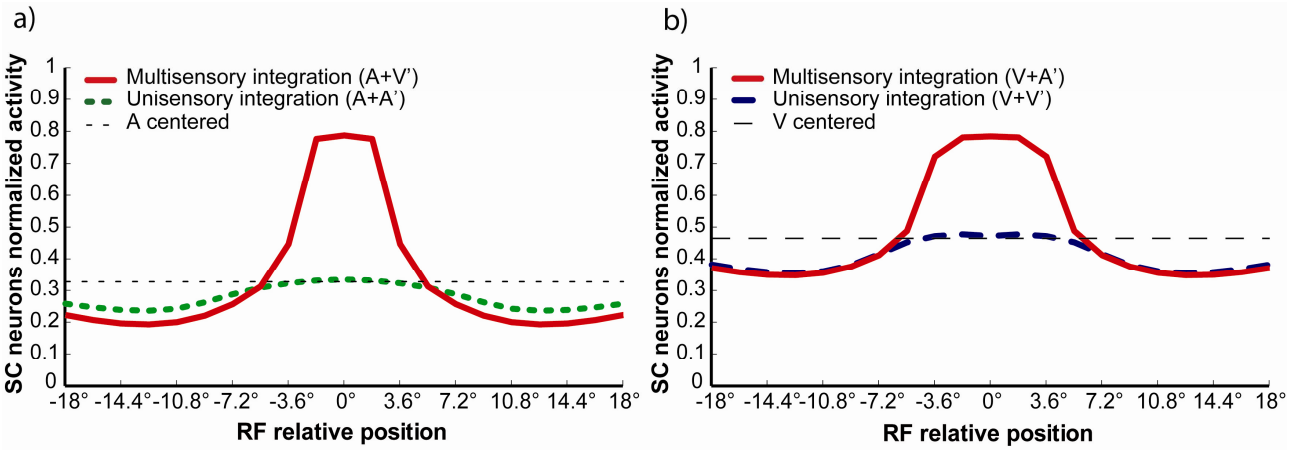


Figure 12 – *Integration as a function of the position of two stimuli.* The figures show the response of the network to paired stimuli in different spatial configurations. Simulations are made by stimulating the model with an auditory (fig.6a) or a visual (fig. 6b) stimulus at the center of the RF of the observed SC neuron. The response elicited by this unimodal stimulus (dotted thin line, in case of auditory stimulation; dashed thin line, for visual stimulation) is then compared with those produced by coupling either a second stimulus of the same sensory modality (dotted thick line, for auditory input; dashed thick line, for visual input) or a stimulus of different sensory modality (solid lines) in different positions. The x axis displays the relative position of the second stimulus relative to the center of the RF. x=0° means that both stimuli are at the center of the RF; increasing x means that the position of the second stimulus is more and more far from the RF. Results show: multisensory enhancement in case of cross-modal stimulation inside the RF irrespective of the position of the two stimuli; no unisensory enhancement within the RF; multisensory and unisensory inhibition in case of two stimuli far in space.

The effect of various distances was tested (see Figs. 6 a and b). In any case, we left enough time after application of the second stimulus for the model to reach a final steady-state condition (i.e., only stationary states are examined here). Results, obtained with the intact cortex, are summarized in Fig. 6a in the case of a central auditory stimulus, and in Fig. 6b in the case of a central visual stimulus.

The following conclusions can be drawn from the results. A second cross-modal stimulus positioned at the center of the RF causes a significant enhancement (about 100-150% in the exempla in Fig.6 relative position = 0°). Conversely, no significant enhancement occurs when a second unimodal stimulus is positioned at the RF center. If the second cross modal stimulus is located at the margin of the RF (Fig.6 relative position = 3.6°) one can still observe a cross-modal enhancement (especially in the case of a central visual stimulus), whereas a second unimodal stimulus located at the periphery of the RF causes moderate depression. The latter result agrees with data reported in Alvarado et al. (Alvarado et al., 2007a,b). Finally, a second stimulus located far from the RF (i.e., not in spatial register with the first) induces a significant depression of the SC response (Fig. 6 relative position = 8-18°), both in case of unimodal and cross-modal stimulation.

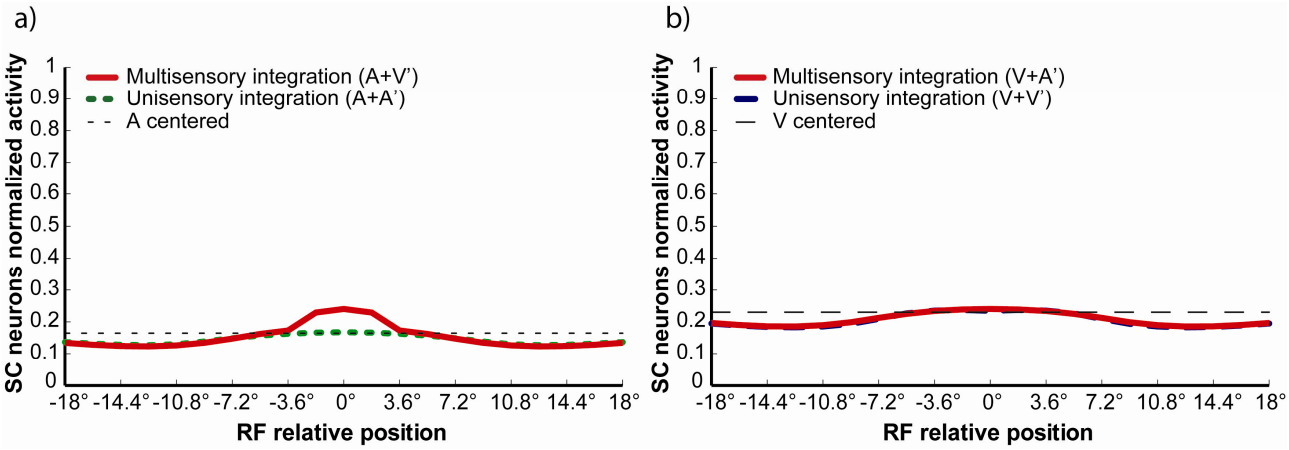


Figure 13 – *The effect of AES cortex on the integration*. The same simulations as in Fig. 6 performed after inactivation of the AES cortex. Results show: 1) a reduction in the SC response both to a unisensory and to a multisensory stimulation; 2) the loss of multisensory enhancement in case of cross-modal stimulation inside the RF: the response of the network looks like the one elicited by the strongest unisensory input; 3) a slight inhibition in case of two stimuli of the same or different sensory modality far in space.

Finally, Fig. 7 shows the effect of two stimuli at different spatial position (these are the same simulations as in Fig. 6) after deactivation of the AES. In this condition, as already shown in Fig. 2, cross-modal enhancement disappears, and the overall response is much weaker than in the intact

case. Moreover, one can observe that cross-modal and unimodal depression are still evident in case of distant stimuli, although the cross-modal depression is weaker than in the intact case (With AES deactivated the maximum cross-modal depression with A fixed is 25% and the unisensory is 22.6% while with AES active cross-modal depression is 41.3%, while the maximum unisensory depression is 28.1%; in the case of V fixed with AES active we have a cross-modal depression of 25.2% and a unisensory one of 23.7%, while with AES deactivated we have 19.4% for the cross-modal depression and 20.6% for the unisensory one).

It's worth noting, in Fig. 7a, that a second visual stimulus placed at the center of the RF causes an increase in the overall response. However, this cannot be considered as a cross-modal enhancement since, in agreement with Fig. 7b, the final response is almost indistinguishable from the response produced by a visual stimulus alone.

DISCUSSION

In the present work we developed and validated a new model of multisensory integration in the cat superior colliculus, to embed disparate information from the present literature into a coherent comprehensive framework. The model aspires to account for several results, including the presence of cross-modal enhancement, cross modal and unimodal suppression, the inverse effectiveness, the effect of selective cortical deactivation and the effect of NMDA blockade. The model was developed using only a few basic mechanisms, which respect neurophysiological principles, although some of them are still hypothetical. More particularly, the fundamental aspects of the model are: i) SC neurons receive inputs from four different unisensory sources; two of them descending from the cortex (AES) and two ascending from subcortical structures. Moreover, these unisensory inputs are in spatial register and obey to a topological organisation. The presence of descending and ascending inputs is well documented in the literature. Subcortical inputs may reach the SC directly from the eye and spinal cord, or from a variety of subcortical nuclei (Stein and

Meredith, 1993). Descending inputs originate from AES and rLS (McHaffie et al. 1988; Meredith and Clemo 1989; Stein et al. 1983; Wallace et al. 1993). An interesting aspect, which seems to contradict the present assumptions, is that the AES also includes multisensory neurons, able to associate cross modal inputs. However, these neurons do not target to the SC but constitute a circuit independent form the SC (Wallace et al., 1993). The model considers only unisensory AES neurons and their projections to the SC, by neglecting the association role of the AES cortex. ii) All neurons exhibit a non-linear characteristic, with lower threshold and upper saturation. This is a well-known behaviour of neurons. This non-linear characteristic is essential to permit the passage from superadditive enhancement in case of mild stimuli to additive or subadditive enhancement in case of large stimuli, according to the inverse effectiveness principle. Moreover, the model predicts that, in case of a depression in the target synapses (as in case of NMDA blockade), the response in the SC becomes more linear, since the neuron does not reach its saturation region. iii) The model hypothesizes the presence of a Mexican hat disposition for the synapses within the unimodal regions. This assumption allows the formation of large “activation bubbles” in all areas and, above all, can explain the occurrence of cross-modal and unimodal suppression in case of stimuli with spatial disparity. The presence of Mexican hat disposition of synapses is well documented in the cortex, and is strictly related with the formation of topological maps. This assumption is perhaps less documented for what concerns subcortical structures. A consequence of having assumed a Mexican hat arrangement for synapses in the subcortical regions is that the model still exhibits cross-modal and within-modal suppression to stimuli not in spatial register even after AES deactivation (see Fig. 6). This result may be the subject of future experimental validation. iv) The most hypothetical aspect of the model concerns inhibition. In particular, two inhibitory mechanisms have been incorporated. The first assumes that the descending inputs, when present, can inhibit the ascending ones. This mechanism is necessary to simulate the loss of multisensory enhancement occurring after selective cortical deactivation (for instance, several authors observed that, after selective deactivation of the AEV, cross-modal enhancement is lost, i.e. a visual stimulus does not

enhance the auditory response. This implies a block of the ascending visual input to the SC by the descending auditory stimulus). The second mechanism assumes a strong competition between the two ascending sources, so that the dominant ascending inputs causes the almost complete inhibition of the other (winner takes all dynamics, WTA). Strong competition was necessary to mimic the observation that, during total AES deactivation, SC loses the capacity to integrate multisensory cues, and neuron response in most cases resembles that to the dominant unisensory stimulus (Jiang et al., 2001). Although the presence of a WTA dynamics in the ascending path is just hypothetical at the present stage, this kind of interaction is frequently met in the brain, especially for the selection of sensory inputs. For instance, WTA dynamics may be essential in some kinds of selective attention mechanisms, to avoid the interference from less relevant inputs which are treated as "distractors". The model suggests that selection of the stronger sensory input via a WTA mechanism, may be the original choice implemented in the subcortical ascending path, whereas a more sophisticate "sensory fusion strategy", able to exploit the presence of different sensory inputs in spatial and temporal register, may have evolved subsequently from the cortex. In the cat, this second mechanism maturates only in the early months of life under the influence of correlated sensory stimuli from the environment (Wallace and Stein, 2007)

It is worth noting that, using the previous assumptions and a single set of parameters, the model is able to mimic several different kinds of behaviour reported in the literature, not only in qualitative but also in acceptable quantitative agreement. The changes in SC response to unimodal and cross-modal stimuli after total or selective deactivation of the cortex (realized by eliminating all input afferents from the deactivated area) are in the range reported in the literature (see Alvarado et al., 2007a,b; Jiang et al., 2001) as we described in the "Results" section.

In order to simulate experiments consisting of NMDA blockade, we included an additional hypothesis. The simpler way to simulate the absence of NMDA receptor was to reduce the synaptic strength from AES to the SC. This hypothesis is reasonable since synaptic efficacy is related with the numbers of ion channels opened by the stimuli which, in turn, is receptor-dependent. Binns and

Salt (1996) observed that NMDA blockade strongly reduces the SC response to visual stimuli presented alone, whereas the response to auditory stimuli exhibited only minor changes. Accordingly, in the model we assumed that NMDA blockade strongly reduces synapses targeting from AEV to the SC, whereas synapses from FAES to SC are only moderately affected. Using this simple hypothesis the model can explain not only the percentage changes in SC response observed after unimodal (auditory and visual) stimulation, but also the changes in the multimodal responses and the shift from non-linear to linear characteristic of the neuron. Possible reasons why NMDA blockade can affect the visual descending path more than the auditory one are discussed in Binns and Salt (1996): the authors mention the possibility that there are different types of NMDA receptors with different pharmacological properties in visual, auditory or somatosensory inputs, or that NMDA vs. AMPA receptors may have a relatively different importance in the mediation of visual vs. the auditory response.

Of course, the present model exhibits important simplification, which must be recognized and justified, and may become the target of future improvements.

A first simplification is that we neglected the role of the rLS. This simplification is justified by the observation that AES appears to be a more important mediator of multisensory integration than rLS (Jiang et al., 2001). In many SC neurons the multisensory enhancement depends only on influences from AES: enhancement is eliminated during AES deactivation whereas deactivating the rLS has almost no effect on neuron response. However, in some SC neurons multisensory enhancement depends on an intact rLS (see Fig. 5 in Stein, 2005) while, in some additional cases, both AES and rLS seem able to induce multisensory enhancement even in the absence of the other area (see Fig. 6 in Stein, 2005).

In the present work we decided to include only the region AES (further subdivided into the AEV and FAES) according to a parsimony principle. Inclusion of another cortical region (rLS) could be done by simply adding two additional unisensory areas in the model with the corresponding descending connections to the SC; however, this choice would have increased the computational

complexity of the model and the number of parameters without evident conceptual benefits. We expect that, using two different cortical regions, and different (perhaps random) weights for the descending synapses, one may conceive a model in which multisensory integration in some SC neurons may depend on the AES only, on the rLS only or on both. However, these aspects are well beyond the aim of the present work.

A further simplification in the present model is that neurons in each area are arranged according to a unidimensional chain, whereas a two-dimensional lattice was used in previous works (Magosso et al., 2008; Ursino et al., 2009). This simplification dramatically reduces the computational cost (especially for what concerns the number of synapses) still maintaining all main conceptual properties of the model. Furthermore, this simplification is justified since SC neurons are more specific in the azimuthal direction than in the vertical one. Hence, even the use of a uniform bi-dimensional lattice would represent a simplification of reality.

Finally, it is important to discuss what may be the role of the present model in future research. The fundamental role of mathematical models in neurophysiology is to propose possible mechanisms able to explain existing data, and to suggest further experiments to validate or reject the proposed hypotheses. In this regard, the present work exhibits important aspects of novelty. It suggests the presence of competitive mechanisms between the different sources which target into the SC neurons. There is at present no anatomical evidence that GABAergic SC interneurons have the same disposition assumed in the present model, i.e., that interneurons receiving descending inputs act on the ascending paths, while those receiving ascending inputs work on the ascending path of the other sensory modality via a competitive mechanism. A consequence of this hypothesis is that, in case of winner takes all dynamics (i.e., strong competition), the response of the SC neuron after AES deactivation equals the response to the stronger input. However, in case of weaker competitive mechanism (which may occur due to natural individual variability among neurons and synapses), the response of the SC neuron to a cross-modal stimulation after AES deactivation

becomes weaker than the stronger unisensory response (this is the exemplum presented in Fig. 3). There are some experimental confirms of this result. (Jiang et al., 2001).

Another hypothesis, which may drive future experimental strategies, is that NMDA receptors are important on the visual descending paths, but have a less evident role on the auditory descending path as well as on both ascending paths. In particular, in case of NMDA blockade, the model predicts that the descending visual path is drastically reduced, and so the ascending visual path becomes relevant in producing the final response to visual unisensory stimulation.

In our opinion, however, the most important contribution of the present model in the future research may be in the study of multisensory maturation during the early months of life. Wallace and Stein (1997) observed that the number of multisensory neurons in the SC gradually increases during the first postnatal months, and that the foremost multisensory neurons do not exhibit significant enhancement. Subsequently, Wallace and Stein (2000 and 2007) observed that the onset of multisensory integration depends on cortical influences, as well as on the presence of correlated input stimuli (for instance auditory and visual cues in spatial register) from the environment. Hence, the present model, with the addition of some learning rules, may be used to formulate hypothesis on synaptic plasticity during the early months life, and verify the effect of these hypotheses on SC multisensory integration characteristics. This may be of great value to gain a deepen understanding on the mechanisms of multisensory maturation and on the basic organization which drives the formation of an adult SC.

REFERENCES

Alvarado JC, Vaughan JW, Stanford TR, Stein BE (2007a) Multisensory versus unisensory integration: contrasting modes in the superior colliculus. J Neurophysiol 97:3193–3205.

Alvarado JC, Stanford TR, Vaughan JW, Stein BE (2007b) Cortex mediates multisensory but not unisensory integration in superior colliculus. J Neurosci 27:12775–12786.

Anastasio TJ and Patton PE. (2003) A two-stage unsupervised learning algorithm reproduces multisensory enhancement in a neural network model of the corticotectal system. J. Neurosci. 23: 6713-6727.

Anastasio TJ, Patton PE, Belkacem-Boussaid K. (2000) Using Bayes rule to model multisensory enhancement in the superior colliculus. Neural. Comput. 12: 1165-1187.

Binns KE and Salt TE (1996) Importance of NMDA Receptors for Multimodal Integration in the Deep Layers of the Cat Superior Colliculus. J Neurophysiol. 75:920-30

Ben-Yishai R, Bar O, Sompolinsky H. (1995) Theory of orientation tuning in visual cortex. Proc Natl Acad Sci U S A 92: 3844-3848.

Colonius H and Diederich A. (2004) Multisensory interaction in saccadic reaction time: a time-window-of-integration model. J Cogn Neurosci. 16:1000-9

Edwards SB, Ginsburgh CL, Henkel CK, Stein BE. (1979) Sources of Subcortical Projections to the Superior Colliculus in the Cat. J Comp Neurol 184: 309-330.

Fuentes-Santamaria V, Alvarado JC, Stein BE, McHaffie JG (2007) Cortex Contacts both Output Neurons and Nitrergic Interneurons in the Superior Colliculus: Direct and Indirect Routes for Multisensory Integration Cerebral Cortex doi:10.1093/cercor/bhm192

Huerta MF and Harting JK. (1984) The Mammalian Superior Colliculus: Studies of its Morphology and Connections. In: H.Vanegas, ed. Comparative Neurology of the Optic Tectum. New York. pp. 687-773.

Jiang W, Wallace MT, Jiang H, Vaughan JW, Stein BE. (2001) Two cortical areas mediate multisensory integration in superior colliculus neurons. J Neurophysiol 85: 506–522.

Jiang W, Jiang H, Stein BE. (2002) Two Corticotectal Areas Facilitate Multisensory Orientation Behavior. J Cogn Neurosci 14: 1240-1255.

Jiang W and Stein BE. (2003) Cortex Controls Multisensory Depression in Superior Colliculus. J Neurophysiol 90: 2123-2135.

Jiang W, Wallace MT, Jiang H, Vaughan JW, Stein BE. (2001) Two Cortical Areas Mediate Multisensory Integration in Superior Colliculus Neurons. J Neurophysiol 85: 506-522.

Kadunce DC, Vaughan JW, Wallace MT, Benedek G, Stein BE. (1997) Mechanisms of Within- and Cross-Modality Suppression in the Superior Colliculus. J Neurophysiol 78: 2834-2847.

Kadunce DC, Vaughan JW, Wallace MT, Stein BE. (2001) The Influence of Visual and Auditory Receptive Field Organization on Multisensory Integration in the Superior Colliculus. Exp Brain Res 139: 303-310.

Knill DC and Pouget A. (2004) The Bayesian brain: the role of uncertainty in neural coding and computation. Trends Neurosci. 27: 712-9.

Magosso E, Cuppini C, Serino A, Di Pellegrino G, Ursino M. (2008) A theoretical study of multisensory integration in the superior colliculus by a neural network model. Neural Netw. 21:817-29.

McHaffie JG, Kruger L, Clemo HR, Stein BE. (1988) Corticothalamic and corticotectal somatosensory projections from the anterior ectosylvian sulcus (SIV cortex) in neonatal cats: an anatomical demonstration with HRP and 3H-leucine. J Comp Neurol. 274: 115-26.

Meredith MA and Clemo HR. (1989) Auditory cortical projection from the anterior ectosylvian sulcus (Field AES) to the superior colliculus in the cat: an anatomical and electrophysiological study. J Comp Neurol. 289:687-707.

Meredith MA and Stein BE. (1986b) Visual, Auditory, and Somatosensory Convergence on Cells in Superior Colliculus Results in Multisensory Integration. J Neurophysiol 56: 640-662.

Meredith MA and Stein BE. (1996) Spatial determinants of multisensory integration in cat superior colliculus neurons. J Neurophysiol 75: 1843-1857.

Meredith MA, Nemitz JW, Stein BE. (1987) Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. J Neurosci. 7: 3215-29.

Patton PE and Anastasio TJ. (2003) Modelling cross-modal enhancement and modality-specific suppression in multisensory neurons. Neural Comput 15: 783-810.

Patton PE, Belkacem-Boussaid K, Anastasio TJ. (2002) Multimodality in the superior colliculus: an information theoretic analysis. Brain Res Cogn Brain Res 14: 10-19.

Perrault Jr TJ, Vaughan JW, Stein BE, Wallace MT. (2003) Neuron-Specific Response Characteristics Predict the Magnitude of Multisensory Integration. J Neurophysiol 90: 4022-4026.

Perrault Jr TJ, Vaughan JW, Stein BE, Wallace MT. (2005) Superior Colliculus Neurons Use Distinct Operational Modes in the Integration of Multisensory Stimuli. J Neurophysiol 93: 2575-2586.

Rowland BA, Stanford TR, Stein BE. (2007) A model of the neural mechanisms underlying multisensory integration in the superior colliculus. Perception 36: 1431-1443.

Stanford TR, Quessy S, Stein BE. (2005) Evaluating the Operations Underlying Multisensory Integration in the Cat Superior Colliculus. J Neurosci 25: 6499-6508.

Stein BE. (2005) The development of a dialogue between cortex and midbrain to integrate multisensory information. Exp Brain Res. 166: 305–315.

Stein BE, Magalhaes-Castro B, Kruger L. (1976) Relationship Between Visual and Tactile Representations in Cat Superior Colliculus. J Neurophysiol 39: 401-419.

Stein BE and Meredith MA. (1993) The Merging of the Senses. MIT Press, Cambridge, MA.

Stein BE, Meredith MA, Wallace MT. (1993) The visually responsive neuron and beyond: multisensory integration in cat and monkey. Prog Brain Res 95: 79–90.

Stein BE, Spencer RF, Edwards SB. (1983) Corticotectal and corticothalamic efferent projections of SIV somatosensory cortex in cat. J. Neurophysiol. 50: 896–909.

Stein BE and Wallace MT. (1996) Comparisons of crossmodality integration in midbrain and cortex. Prog. Brain Res. 112: 289–299.

Ursino M, Cuppini C, Magosso E, Serino A, di Pellegrino G. (2009) Multisensory integration in the superior colliculus: a neural network model. J Comput Neurosci. 26: 55–73.

Treves A. (1993) Mean-field analisys of neuronal spike dynamics. Network 4: 259-284.

Wallace MT, Meredith MA, Stein BE. (1993) Converging Influences from Visual, Auditory, and Somatosensory Cortices onto Output Neurons of the Superior Colliculus. J Neurophysiol 69: 1797-1809.

Wallace MT, Meredith MA, Stein BE. (1998) Multisensory Integration in the Superior Colliculus of the Alert Cat. J Neurophysiol 80: 1006-1010.

Wallace MT and Stein BE. (1997) Development of Multisensory Neurons and Multisensory Integration in Cat Superior Colliculus. J. Neurosci. 17: 2429–2444

Wallace MT and Stein BE. (2000) The role of experience in the development of multisensory integration. Soc Neurosci Abstr 26: 1220.

Wallace MT and Stein BE. (2007) Early experience determines how the senses will interact. J. Neurophysiol. 97: 921–926.

Wallace MT, Wilkinson LK, Stein BE (1996) Representation and integration of multiple sensory inputs in primate superior colliculus. J. Neurophysiol. 76: 1246–1266.

CONCLUSIONS

The models proposed in this work provide several adjuncts to the experimental knowledge and research: on one hand, our models can be of value to clarify understanding of in-vivo data available in literature; on the other hand, they can suggest novel and fruitful routes for further investigation of the examined cognitive processes.

In particular, the models point out plausible scenarios about the functioning of some neural structures (in particular the SC), difficult to analyze only by means of classical experimental methods, and identify, with neurobiological consistency, possible neural mechanisms underlying high level cognitive processes (object recognition, memorization, and semantic memory).

Within these theoretical frameworks, several experimental data are coherently synthesized, and the variability of in-vivo data is explained, ascribing it to few parameters modifications. Experimental results acquire more value and can be better exploited when their relationships are evidenced in rigorous quantitative terms.

In addition, model predictions can suggest the existence of additional mechanisms in the neural circuits (e.g. top-down attentive mechanisms to resolve perceptual conflicts or new language learning mechanisms), that may become the subjects of future novel and fruitful neurophysiologic experiments.

In the following, the main contributions of the models in each area of this research activity, are briefly highlighted:

Object representation and semantic memory – These models, by using simple Gestalt rules of similarity and previous knowledge, and time division to detect multiple contemporary objects:
(i) give a plausible interpretation of neural mechanisms, in particular the γ-band synchronization, involved in high level cognitive processes such as object recognition and semantic memory;

(ii) are able to implement a theoretical structure that incorporates the last conceptual theories present in literature and reproduces neurophysiological data;

(iii) realize, via a neurobiological plausible Hebbian learning algorithm, a bidirectional relationship between cortical representations of abstract objects and a lessical area, devoted to words representation;

(iv) can be exploited to form classes of objects, and to detect category membership, without the need for a hierarchical representation of objects.

SC multisensory integration capabilities – The proposed models provide a potential scenario of neural circuitries and mechanisms underling multisensory integration in the SC. The main strengths of the model can be resumed as follows:

(i) it is entirely based on neurobiologically plausible mechanisms;

(ii) with a single set of parameters it is able to simulate several characteristics of SC neurons in quantitative agreement with experimental findings (multisensory enhancement, inverse effectiveness, within-modality and cross-modality suppression, effect of cortical deactivation, role of NMDA receptors);

(iii) changes in some model parameters, still maintaining the same network topology, can explain the variability of in-vivo SC cell behaviour and account for some perception illusions;

(iv) it maintains a moderate level of computational complexity;

(v) it can be useful to investigate the mechanisms involved in maturation of multisensory capabilities during early life.

Finally, in subsequent studies these mathematical models, with suitable modifications and extensions, may be used to address further cognitive problems, such as investigate the developmental changes responsible for the formation of a multimodal space representation, shed light on neurological deficits in audio-visual perception, assess the potential effects of rehabilitation paradigms, explore the learning paradigms in language acquisition.

PUBLICATIONS

- **International Journals**

  M. Ursino, E. Magosso, G.E. La Cara, C. Cuppini. "Object segmentation and recovery via neural oscillators implementing the similarity and previous knowledge Gestalt rules", *Biosystems* 85(3): 201-218, Sep 2006.
  ISSN 0303-2647 (IF 1.646)

  M. Ursino, E. Magosso, C. Cuppini "Possible mechanisms underlying tilt aftereffect in the primary visual cortex: a critical analysis with the aid of simple computational models", *Vision Res* 48(13): 1456-1470, Jun 2008.
  ISSN 0042-6989 (IF 2.055)

  E. Magosso, C. Cuppini, A. Serino, G. di Pellegrino, M. Ursino. "A theoretical study of multisensory integration in the superior colliculus by a neural network model", *Neural Netw* 21(6): 817-829, Aug 2008.
  ISSN 0893-6080 (IF 1.951)

  M. Ursino, C. Cuppini, E. Magosso, A. Serino, G. Di Pellegrino. "Multisensory integration in the superior colliculus: a neural network model", *J Comput Neurosci* 26: 55-73, 2009
  ISSN 0929-5313 (IF 1.928)

  M. Ursino, E. Magosso, C. Cuppini. "Recognition of abstract objects via neural oscillators: interaction among topological organization, associative memory and gamma-band synchronization", *IEEE Tran Neural Netw* Feb. 20(2) 316-335, 2009
  ISSN 1045-9227 (IF 2.7769)

  C. Cuppini, E. Magosso, M. Ursino: A Neural Network Model of Semantic Memory Linking Feature-Based Object Representation and Words. *BioSystems*, (in press) doi:10.1016/j.biosystems.2009.01.006
  ISSN 0303-2647 (IF 1.646)

- **International Books**

  - C. Cuppini, E. Magosso, A. Serino, G. di Pellegrino, M. Ursino, "A neural network for the analysis of multisensory integration in the Superior Colliculus", In: *Artificial Neural Networks-ICANN'07 Proceedings of the 17th International Conference* 9-13 September, Porto, Portugal (Series: *Lecture Notes in Computer Science*) Springer Berlin/Heidelberg (ISBN: 978-3-540-74693-5).

- **Extended relations at International Conferences**

  E. Magosso, G.E. La Cara, C. Cuppini, M. Ursino. "Simulation of similarity and previous knowledge Gestalt rules with coupled neural oscillators". *Proceedings of the 3rd European Medical and Biological Engineering Conference EMBEC'05* 20-25 novembre 2005, Praga, Repubblica Ceca, paper n. 1491, pp. 1-5, su CD-ROM (ISSN: 1727-1983).

  E. Magosso, C. Cuppini, M. Ursino. "Object segmentation and reconstruction via an oscillatory neural network: interaction among learning, memory, topological organization and γ-band synchronization", *Proceedings of the 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society 2006 IEEE-EMBC 2006*, 30 agosto-3 settembre 2006, New York City (USA), pp. 4953-4956, The Printing House, Inc. (CD-ROM, ISBN 14244-0033-03).

- **Abstract at International Conferences**

  E. Magosso, M. Ursino, G. E. La Cara, C. Cuppini. "An oscillatory neural network implementing the similarity and prior knowledge gestalt rules: object segmentation and retrieval", *12th Annual Meeting Human Brain Mapping*, June 11-15, 2006, Florence, Italy. Pubblicato su *Neuroimage*, Vol. 31, Suppl. 1, Abstract n. 452 M-PM

  C. Cuppini, E. Magosso, M. Ursino, "A neural network model for the study of the relationship between language processing, topographical rapresentation and γ-band synchronization" *Proceedings of the X International Conference on Cognitive Neuroscience* September 1-5, 2008, Bodrum, Turkey, p 202. [Cuppini C, Magosso E and Ursino M (2008). A neural network model for the study of the relationship between language processing, topographical representation and γ-band synchronization. *Frontiers in Human Neuroscience. Conference Abstract: 10th International Conference on Cognitive Neuroscience*. doi: 10.3389/conf.neuro.09.2009.01.120]

  C. Cuppini, M. Ursino, E. Magosso, B. A. Rowland, B. E. Stein "A neural network model of multisensory maturation in superior colliculus (SC) neurons" *Abstracts Neuroscience 2008***,** November 15-19, 2008, Washington DC (in press)