RIFL/SFL (2017): 87-102
DOI: 10.4396/SFL201707

The plurality of neurosemantics

Vivian M. De La Cruz
Dept. Of Educational Sciences
University of Catania
vmdelacr@unict.it

Alessio Plebe
Dept. of Cognitive Science
University of Messina
aplebe@unime.it

Abstract Neurosemantics is a relatively new approach to investigating the construction of linguistic meaning. In recent years, neurosemantics has been used in two different ways. One regards the meaning of the electrical and the chemical activities going on in neural circuits, or according to what we call the "the semantics of neurons" approach. The second, regards the type of semantics studied for years in philosophy: the meaning of language, but with the added intention of explaining in neural computational terms, what happens when people listen to and understand utterances. We think neurosemantics, understood as the construction of linguistic meaning in neural terms, requires an assumption of continuity. This is because the physiological strategies upon which language are based are no different in nature, to those by which neurons create non-linguistic conceptual systems. This continuity can subsist defending two different controversial notions, that of *representation* and *computation*.

In our work, we explore neurosemantics according to the second sense or approach mentioned above, but in doing so, we address much of the first sense or approach as well, in that we believe that the capacity of neural circuits in humans to support linguistic meaning, hinges on their peculiar role of coding experience. In this paper, we illustrate examples of linguistic phenomena that can be explained through the employment of these two concepts and briefly describe neurocomputational models that have simulated how this phenomena might be instantiated in the brain.

Keywords: Neurosemantics, neurocomputational models, language, representation, computation, cortex, population coding

Received 1 September 2017; accepted 6 December 2017.

0. Introduction[1]

Neurosemantics is one relatively new approach to investigating the construction of linguistic meaning. How does this approach intersect with the more traditional notions of semantics, and in particular those more commonly accepted in philosophy of language? Answering these questions is part of our ongoing research endeavor.

Our particular contribution to neurosemantics, which is the project the present paper is based, regards concrete neurocomputational proposals grounded on the joint concepts of representation and computation. It is an applied and practical one, based on a mathematical foundation, which in turn is grounded in how the brain computes. We do this through addressing aspects of linguistic semantics, or how linguistic meaning is constructed, first in purely neurocomputational terms, and second, through the identification of the principles upon which neurocomputational models of linguistic meaning can be constructed, and their corresponding neural bases.

1. Neurosemantics in two different senses

In recent years, in scientific contexts, neurosemantics has been used in two different ways. One regards the meaning of the electrical and the chemical activities taking place in neural circuits, or what we call the "the semantics of neurons" approach. This is the use Churchland (2001), Ryder (2004), or Breidbach (2007) make of the term. The second, regards the type of semantics that has been studied for years in philosophy: the meaning of linguistic expressions, but with the added intention of explaining in neural computational terms, what happens when people listen to and understand utterances. This is the main use Pulvermüller (2012) makes of the term neurosemantics, and it is not much different from that made by Feldman (2006), even if the latter never explicitly uses it. We call this the "semantics from neurons" approach.

Neurosemantics, in the first sense mentioned above, aims to investigate the meaning of the phenomena that takes place in the neural system such as electrical impulses, chemical reactions, and morphological modifications of cells. In this case semantics stands for those explanations that focus on factors that are external to the physical brain itself, in that they invoke or make use of notions such as that of neural representation. Neural representation is an idea according to which, neurons serve as devices for the mental construction of the external world. This notion only marginally concerns neuroscience, while it is central to neurophilosophy. In fact, this is the sense Paul Churchland refers to when he uses the term.

In the second case, neurosemantics instead regards the endeavor of accounting for language in terms of neural processes.

One of the first efforts to do so came from the connectionist movement of the 80's and 90's. This movement, however, avoided falling under the umbrella of the term. This was most likely due to the fact that its proponents were busy opposing the symbolic views of Fodor and Chomsky, by providing abstract non-symbolic computations in their stead.

In more recent years, the embracing of the biophysical behavior of neurons as the fundamental basis for understanding and explaining the semantic constructions of language, has instead been Friedmann Pulvermüller's primary objective. He has used

---

[1] Although the authors have equally contributed to the ideas and contents of this work, Author 1 was responsible for writing sections 0, 1, 3, 5, 6 and 8, and Author 2 for sections 2 and 4.

neurosemantics as the way of responding to the second question we touched upon above: what are the neuronal processes behind semantics, or how is linguistic meaning derived and/or constructed from the functioning of neurons?

Before we go further, we would like to underline that a terminological caveat is in order here. For us the expression "functioning of neurons" is a condensed way of addressing the complexity of factors involved in the computational properties of the brain, that instead have been formalized elsewhere (see PICCININI 2015), and that are not confined inside the physical boundaries of neural cells only.

After having used the dichotomy of the "semantics of neurons" and the "neurons of semantics" as an argumentative device, we now escape it, sustaining that it is difficult to fully embrace one without having to deal with questions regarding the other. We think neurosemantics understood as the construction of linguistic meaning in neural terms, requires an assumption of continuity. This is based on the consideration that, the physiological strategies upon which language are based are no different in nature, to those by which neurons create non-linguistic conceptual systems. This continuity can subsist defending two different controversial notions, that of representation and computation.

In our work, we explore neurosemantics according to the second sense or approach mentioned above, but in doing so, we address much of the first sense or approach as well, in that we believe that the capacity of neural circuits in humans to support linguistic meaning, hinges on their peculiar role in coding entities and facts of the world. Even if combining sequences of sounds for conveying meaning through words and sentences is a strategy that is exploited by humans only, its realization by way of neural circuits appears to be rooted in their disposition to code for things in the world, something that is shared in a wide variety of forms, with other animals.

2. Computation and representation

A road that has been particularly favored in cognitive science is that in which, mind semantics is explained by combining the two ideas of computation and representation. Whatever content is held in a mind, it is done through the manipulation of certain basic representative entities, and in accordance with a set of transformation rules. This is known as the Computational and Representational Theory of the Mind, championed by (FODOR 1975, 1987). Its original version argues that computation works on a language-like syntax and a compositional semantics basis over symbols and, to that extent, amounts to a language of thought. Even if computation and representation are often held as necessarily interwoven, as Fodor's famous slogan (1981) «no computation without representation» indicates, they deserve separate considerations.

The most commonly accepted and precise reference to computation, was originally proposed by Putnam (1960). Its mathematical foundation was within the realm of computer science (CHURCH 1941, KLEENE 1936; TURING 1936), and became widely accepted in cognitive science. When moving into the area of neurosemantics, the computer science framework remains important, but it is not sufficient in order to articulate what computation is and how it works. A basic tenet in neurophilosophy, since (P. S. CHURCHLAND and SEJNOWSKI 1994; SEJNOWSKI, KOCH, and CHURCHLAND 1988) has been that the essential activity of neurons is indeed that of computing, but it is a sort of computation that may significantly depart from the mathematical framework of computation developed in computer science. Neural computation has gradually been established as both a means to a better understanding of neural functioning by computational tools, and a way of assessing the genuine sort of computation that neural systems perform (DAYAN and ABBOTT 2001). At the same time, a great deal of philosophical effort has been made to refine notions of computation that are valid for both computers and minds (FRESCO 2014; MILKOWSKI 2013; PICCININI 2015). One of the best available accounts, provided by Piccinini (2015), makes use of the concept of vehicles upon which computation is performed, defined as entities or variables that can change state according to rules. Computation would then be a manipulation of vehicles according to those rules that involve vehicle properties and, specifically, differences between different portions of the vehicles, including their temporal changes. A fundamental constraint on vehicles is that they need to be "medium independent", their manipulating rules should obey differences between portions of the vehicles, along specified dimensions, but should be insensitive to any of the concrete physical properties of the vehicles. Thanks to the concept of vehicles, it is possible to give a more rigorous definition of digital computation, as a generic computation whose vehicles are strings of digits, macroscopic states whose type can be reliably and unambiguously distinguished by the systems of other macroscopic types, with the constraint that digit types should be finite in number. Neurons comply with the definition of computation based on vehicles, but not that of digital computation, because any attempt to reduce trains of action potentials to strings of digits, would fail to meet the finite constraint on the number of types, due to the continuous time of the events.

In addition to the ontological question, neural computation raises epistemological ones: can a computational description of neural behavior be explanatory of the phenomena associated with such behavior? More specifically, can computational descriptions of neural activities explain semantic phenomena? An answer to the general question has been sought inside the framework of mechanistic explanation, and is among the most accepted in biology in general (BRAILLARD and MALATERRE 2015), and in neuroscience in particular (CRAVER 2007).

In brief, mechanistic explanations require the identification of components of the system to be explained, the definition of the functions of each component, and the relations between functions in producing the set of behaviors of the system (BECHTEL & RICHARDSON 1993). One of the most elaborate attempts to reconcile neural computation with mechanistic explanation is attributed to Kaplan (2011) and Kaplan and Craver (2011). They proposed the 3M (model-mechanism-mapping) constraint. A model of a target phenomenon explains that phenomenon to the extent that:

1. The variables in the model correspond to identifiable components, activities, and organizational features of the target mechanism that produces, maintains, or underlies the phenomenon;
2. The dependencies posited among these variables in the model correspond to causal relations among the components of the target mechanism.

This constraint does not need to be strict, it is compatible with models where certain details are omitted, its guiding principle: computational models whose design principle disregards the structural correspondence between model components and external system components are merely predictive models, while in the case in which the design is based on such correspondence, they aim at being explanatory models. Models that here are suggested to explain certain neurosemantic phenomena attempt to meet the 3M criteria, paying close attention to the plausible correspondence of their components with brain components.

In the Computational and Representational Theory of the Mind, much less theorizing can be found for representation than that seen for computation. Fodor (1990) pioneered the view that has come to be known as informational atomism, that is, that basic mental representations are simple unstructured "atomic" symbols. An opposing view, one that has found more consensus among psychologists, and is consistent with most twentieth century philosophy, is that representations are structured entities (CUMMINS 1989), in two different senses: first, in possessing conditions of every content of representation, other contents are involved, second, every content of representation is composed by parts, which in turn can be content of other representations. For example, in the domain of ANIMALS the representation of DOG involves sharing some similarities, but having differences with CAT, and both representations are composed of parts, like LEG and TAIL. In addition, structural representations retain a longstanding tradition that dates back to antiquity, and found clear articulation in Hume (1748), representations should bear some similarity with what is represented.

The foundation of structural representations has had an independent history, one far from the domain of computer science. The first attempt to formalize similarity in terms of structural relations was thanks to Russell (1927), who introduced the concept of relation-number, defined as the class of all relations similar to the given relation. For Russell, similarity in terms of relation-number was an abstract theory of how a mathematical system can represent aspects of the physical world, but one crucial application of this theory was in sensorial perception, therefore, targeting mental representations. In the 70s, a new mathematical domain began to be developed that followed in the footsteps of Russell's foundational work on structural representations, that of measurement, by Krantz, Luce, Suppes, and Tversky (1971). The purpose of this theory was to systematize all kinds of measures, i.e. mappings from domains of physical (or economic, social, psychological) stuff, into the codomain of numbers. The theory of measurement grew considerably for twenty years (LUCE, KRANTZ, SUPPES, and TVERSKY 1990; SUPPES, KRANTZ, LUCE and TVERSKY 1989), but, contrary to Russell, had no interest in applying the theory to mental representations. Swoyer (1991) returned to the original aim of Russell, and tried to widen the scope of measurement theory to include the domain of mental representations. He adopted the same basic definition of representations, that requires the establishment of a homomorphism between relations that can be identified in the world domain (the represented system), and corresponding relations in the (mental) representing codomain.

The account of structural mental representations based on the theory of measurement is affected by several problems, the most serious being the excess of liberalism: there can be many trivial domains satisfying the homomorphism requirements, this affliction is called panrepresentationalism. Still, this account is defended as being the best one available in cognitive science, and for Ramsey (2007) a failure of neural representation is exactly the lack of structures adhering to Swoyer's criteria. Not only do we not see particular reasons why neural representation should lack structure, as already remarked by others (MORGAN 2014; SHAGRIR 2012; SPREVAK 2011), we argue that structural properties are key features that turn neural responses into genuine representations. In particular, we will argue that inside the cerebral cortex there are known structural organizations that are suitable for neurosemantics.

3. Representations in the cortex

There are reasons to restrict our analysis of neurosemantics to the cerebral cortex, even if we are convinced that representational mechanisms not too far from those in place in the cortex may be present in other structures of the brain. One reason is that the most common and fundamental use of semantics is in behaviors such as reasoning, remembering, planning, conversing, which typically take place mainly in the cerebral cortex. Moreover, the cortex expresses such a wide range of high level functions, from vision (CHALUPA AND WERNER 2003), somatic perception (NELSON 2002), language (PULVERMÜLLER 2002), planning and decision making (FUSTER 2008), to consciousness (DEHAENE 2014), and moral cognition (VERPLAETSE, SCHRIJVER, VANNESTE and BRAECKMAN 2009), all with a remarkable uniformity in structure. This suggests a common basic representational mechanism. In addition, the cortex is one of the most plastic parts of the brain, a fundamental feature in sustaining the formation of semantic representations. Finally, there are known organizational principles of the cortex, that provide convenient starting points for an explanation of structural representation, that is able to bridge the gap between the empirical domain and the representing codomain, left in Swoyer’s account.

Among the distinguishing features of the mammalian cortex there is columnar organization, made by cords of cells oriented radially, often supposed to be a sort of basic processing circuit (DOUGLAS, MARTIN and WHITTERIDGE 1989; HUBEL and WIESEL 1959; MOUNTCASTLE 1957). Columnar organization is coupled with the concept of “cortical map”, an organization orthogonal to the columns, along the 2-dimensional surface of the cortex (MOUNTCASTLE 1957). In sensory and motor areas of the cortex, maps exhibit a topological mapping between the cortical space and the distal – sensorial or motor – space (FELLEMAN and VAN ESSEN 1991).

In cases where the topological mapping is directly of a spatial nature, like in retinotopy, or in somatosensory maps, this relation may well qualify as representation by virtue of its structural similarity property. In the primary visual area an impressive number of overlapping features have been discovered: retinotopy, ocular dominance (Wiesel & Hubel, 1965), orientation selectivity (HUBEL and WIESEL 1968), color (TOOTELL, SWITKES, SILVERMAN and HAMILTON 1988), and spatial frequency (TOOTELL, SILVERMAN, HAMILTON, SWITKES AND DE VALOIS 1988). Maps often contain modular repetitions of small segments of receptor areas, within a topography at larger scale (KRUBITZER 1995). Even if topological order has long been supposed to exist in the homotypical cortex as well (KOHONEN and HARI 2000; THIVIERGE and MARCUS 2007), the interpretation of contents in higher-level maps seems to involve more complex relations.

A fruitful concept is that the power of representing information in cortical circuits lies in the combination of the activities of many columnar units, a concept often named “distributed coding” (HINTON, MCCLELLAND and RUMELHART 1986), also known as “population coding”, “vector coding” and close to the “state space representation” idea (P. M. CHURCHLAND 1989). We deem this idea to currently be the best account of representation at the level of mental concepts and linguistic meaning, with populations of cells in high level maps coding for different kinds of objects or facts by their concurrent level of firing.

By using population coding, the structural representation of concepts in the cortex can be formalized in a more precise way than the general use of homomorphism in the abstract model of Swoyer. Mathematical details of how population coding can work as homomorphic mapping between domains of the external world and the codomain of cortical maps can be found in Plebe and De La Cruz (2016: 105).

4. Neurosemantics of visual objects

In our neurosemantics endeavor we concentrated on investigating the semantics of objects whose essential features are captured by their visual appearance, or in other words, how the representation (and eventual meaning) of object names could be constructed in the brain, from a neural point of view. We created a series of neurocomputational models, the first of which was a sort of prelude to a full blown semantics, with a simulation of the full visual pathway that transforms the perception of light into the recognition of object categories, together with the auditory pathway, in a simulation of the emergence of a first lexicon, that in infants begins with naming the objects they most often see. Most of the components of this model, the methods used for its development and subsequent analyses, were shared by later models we developed. These subsequent models, tapped into a range of semantic phenomena typically observed in the early stages of language development in children, such as the change in the speed of learning, and the so called "fast-mapping" phenomenon. Due to space limitations, in this section, we will limit ourselves to briefly describing work done with our initial object recognition model, further details on this as well as subsequent models developed can be found in Plebe and De La Cruz (2016). A neurocomputational modeling of color semantics will be discussed below.

4.1 Object recognition model

The object recognition model simulated the development of the ability to segment and recognize objects on one side, and segment and recognize words on the other, and the ability to join the two represented entities in linguistic meaning. In real infants, these processes take place in the brain at the crossroad between the ventral visual and auditory streams of processing, and the portions where these streams of processing happen in the brain, are the portions we simulated. We also employed the mechanism of coincidence detection because, in our opinion, it is the fundamental principle of cortical representation. We took it as the core mechanism for explaining the emergence of the semantics of visual objects.

Coincidence detection is a multilevel mechanism. At the lower level, it is implemented in local synaptic connections by their sensitivity to the occurrence of simultaneous activation of neighboring units. At an intermediate level, it is responsible for building selectivity in units to recurrent patterns, such as oriented lines in the visual scene, or classes of phonemes. At the highest level, it captures the coincidence of seeing certain objects while hearing the same sound, which then becomes associated with the category of similar objects named by the sound, which is a word.

Briefly stated, this initial neurosemantic model was built upon a number of Topographica maps, simulating several cortical areas, as well as on a few thalamic maps. Initially, the model lacked functional processing, with all synaptic connections initialized to small random values. During the experiments the model was exposed to a series of stimuli, which reproduced at a small and essential scale, early human brain development relevant for its cortical areas. Fig. 1 shows the maps that compose the model.

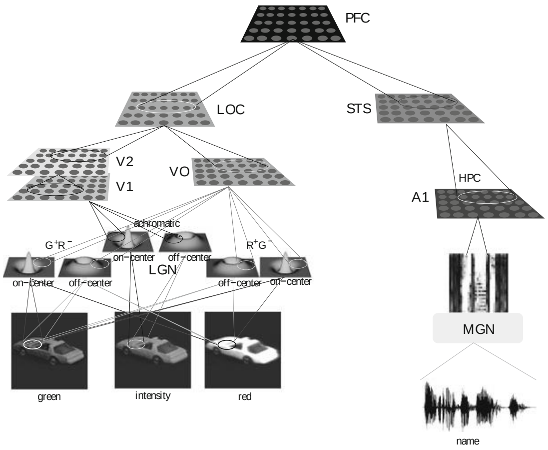

Fig 1. Overall scheme of the object recognition model. LGN-Lateral geniculated nucleus, MGN-Medial geniculated nucleus, V1-Primary visual cortex, V2-Secondary visual cortex, VO-Ventral occipital, A1-Auditory primary cortex, LOC-Lateral occipital complex, STS-Superior temporal sulcus, PFC-Pre-frontal cortex

In a first phase only maps V1, VO, and A1 are plastic, and adapt their synaptic connections. The visual stimuli are synthetic random blobs that mimic waves of spontaneous retinal activity that are known to play a fundamental role in the ontogenesis of the visual system (MASTRONARDE 1983). Those presented to V1 are elongated along random directions, to stimulate orientation selectivity, while blobs to VO are circular, with constant hues, and random size, position, and intensity, in order to induce color constancy. The A1 map was exposed to short trains of waves sweeping linearly around a central frequency, with time duration, central frequencies and sweeping intervals varied randomly.

The second phase involved V2 and STS maps as well. The visual stimuli comprised pairs of elongated blobs, the same previously used for V1, with a coinciding end point. These sort of patterns stimulate the selectivity of units to patterns that are slightly more complex than oriented lines, like corners. The auditory stimuli were synthesized waves of the 7200 most common English words with length in range from 3 to 10 characters, generated using the Festival software (BLACK AND TAYLOR 1997).

In a third phase corresponding to the phase just after eye opening in the newborn, natural images were used. In order to include the identification of an object under different perspectives, the COIL-100 collection was used, where for each of the 100 real childhood related objects, 72 different views were available.

The last phase corresponded with early language acquisition, and the model was exposed to events in which, an object is viewed and a label corresponding to its basic category was heard simultaneously. The 100 objects were grouped manually into 38 categories. Certain categories, such as cup or medicine had five exemplars in the object collection, while others, such as telephone, had only one exemplar. Each category word was converted from text to waves using an American male voice and female voice at standard and slow speeds in the Festival software.

All models were composed by a complex yet extremely simplified hierarchy of neural sheets, with respect to the real brain. In each, the emergence of functions consistent with their physiological counterpart areas was expected, it was also expected that they would contribute to subsequent sheets, up to the higher level. For this purpose, the range of stimuli used at the earlier stages of development were a subset of all the possible stimuli that could achieve similar functions, in order to simplify the overall experiment.

At the end of the simulated development, several types of topological organization were found in the maps of the model, which are consistent with the known role of those maps in cortical hierarchy, such as orientation selectivity in the V1 map, response to specific hues regardless of intensity in VO, and responsiveness to angles in V2 (for more on the functions that emerge in these lower areas see PLEBE and DE LA CRUZ 2016).

It is what happened in the higher maps that was most interesting to our discussion on the construction of linguistic meaning in the brain. Our analysis of the higher maps was made under the assumption of population coding (discussed in Section 4), considered by many to be a powerful representational strategy in the cortex.

In brief, analysis of the population coding in the higher maps of our model, LOC, STS and PFC, as a result of our experiments we found what follows (for full details on all the experiments done and the results see PLEBE and DE LA CRUZ 2016).

The LOC map responds to different views of the same object. It assists in the recognition of objects under different viewing conditions. Results of the experiments done with the model show the population coding of the LOC map and indicate that it is relatively accurate in discriminating objects it was exposed to even when the objects were shown according to different angles.

The STS map codes words for their phonological form. Population coding for some of the words used as object labels showed the accuracy of STS in discriminating between all the different words learned by the model.

In the PFC map the coding is amodal in kind and associates information of a visual and acoustic nature. Representations in PFC are the result of the tension between the normativity in naming things, and the natural similarities of their appearance. Some examples resulting from this analysis show separated coding of visual and auditory information compared with the full amodal population coding. It was interesting to detect a global partitioning in map PFC, with linguistic information more clustered on the right side and visual on the left, still with large overlaps. Often the amodal coding looks like a combination of the visual and linguistic representations. In other cases, one modality seems to be prevalent, maybe because more influenced by the linguistic representation, while in others the visual information is predominant.

What the results with our neurosemantics computational model exploring object recognition and early word learning processes indicate, is that from a neural point of view, linguistic meaning is initially constructed in the brain thanks to coincidence detection mechanisms linked to associative learning processes, along with population coding strategies.

There are those that might not be satisfied with this claim, and that might counter argue stating that while neural representations are internal properties of an individual, linguistic meaning, on the other hand, is constructed within a community of speakers. This line of reasoning might thus lead to asking how we get from the internal individual representation of meaning at the neural level, to the social and inter-subjective one.

We address this apparent gap in the next section, where we will describe an extension of our model and results obtained, that indicate how internal representations develop in individuals that quite possibly reflect their social environment and as well as cultural differentiations.

5. Neurosemantics of color terms

A long-standing question in language development research has concerned the contrast between early word learning and the learning of adjectives. Color terminology is an interesting and particular case of adjective learning. In our neurosemantics research we extended our neurocomputational model to simulate neural processes behind the learning of adjectives in linguistic development, and in particular color terms. We went further, investigating how both exposure to language as well as the range of colors dominant in a particular natural environment might influence not only how color terms are learned, but how colors are perceived, a topic that is at the center of the historic linguistic relativity debate (for details see PLEBE & DE LA CRUZ 2014, 2016: 166-176).

We adapted the basic architecture of our model as already described in Section 4.1, to attempt a simplified, yet biologically plausible simulation of human color processing, and a reasonable account of the interaction between color perception and language. The precise reconstruction of human color physiology was essential for this simulation, therefore, all visual inputs were converted into the three components according to the spectral responses of the short, medium, and long wavelength retinal receptors. The conversion was performed using empirical data of spectral sensitivities of cone receptors gathered by Stockman and Sharpe (2000). The set of opponent receptive fields collected in the model under the label LGN is more complex than in previous models as well, accounting for all color opponent combinations in human vision (STOCKMAN AND BRAINARD 2010).

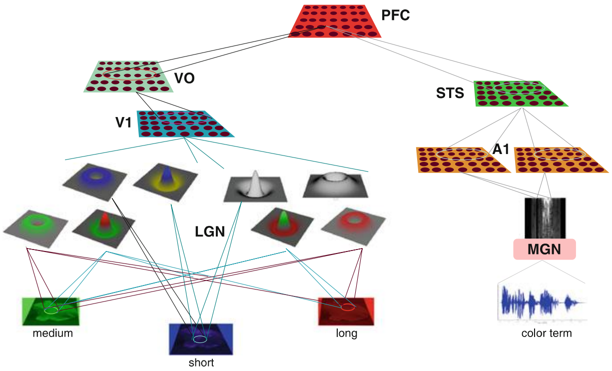

Fig. 2. Scheme of the model simulating the semantics of color terms. The components are mostly the same as those in the model in Fig. 2, except for the subcortical color components, which here correspond to the short, medium, and long wavelength retinal receptors

While the early developmental phase was common to all the models, in the linguistic phase three different models were developed, corresponding to the separate effects of three different environments, Berinmo, Himba (cultures whose color terminology and ability to perceive color was studied by Debi Roberson and colleagues) as well as the typical, more varied yet undifferentiated visual world of western cultures. The rationale behind these experiments is that we sustain that the ability humans have of seeing colors is not only strongly influenced by the biological constraints acting upon the visual system of our species, by the color terms our linguistic community teaches us to use, but also by the "visual diet" provided to us by our natural environment. We therefore, thought it interesting to investigate the possible impact the environment might have in the cases of Berinmo and Himba, since the landscapes in which the two groups live are drastically different. In the pre-linguistic phase of development of the model, natural images were used as stimuli for the visual pathway, with three variations. A neutral one lacks dominant hues, and is typical of many urban environments in modern cultures, where the most common objects and scenes seen by infants are made by humans, with a wide range of colors, prevailing over the natural hue biases of the natural environment, if any. We used random pictures from the Flowers and Landscape collections of the McGill Calibrated Color Image Database (OLMOS and KINGDOM 2004).

The other two environments were those typical of Berinmo and Himba that contrary to the neutral urban environment, are dominated by specific ranges of hues. The Berinmo environment is the luxurious vegetation of Papua New Guinea, along the large Sepik river, with villages found under the shadow of tall trees. The Himba people live in the open rocky desert lands of Northern Namibia, dominated by warm earth-colored hues.

All pictures of the Berinmo and Himba environments were courtesy of Debi Roberson, and were shot during her color terms investigations of these people (see Fig. 3 below). The analysis of the spectral differences between the three sets is in (PLEBE and DE LA CRUZ 2014).

Fig. 3. Examples of the pictures used to train three different versions of the model, adapted for different environments. On the left, neutral pictures from the Flowers and Landscape collections, in the middle, the New Guinea – Sepik river environment (Courtesy of Debi Roberson), on the right, the Namibia – Kunene district environment (Courtesy of Debi Roberson)

Results showed a clear categorical perception effect for all languages, with the blue-green boundary, in particular, strongly affecting the English model, while having an indifferent effect on Himba and Berinmo. The neurosemantics of color terms simulated by the models briefly discussed, seem to corroborate the relativist view (for more details see PLEBE & DE LA CRUZ 2014; 2016: 166-176). We certainly believe that the neural mechanisms of human vision, place important constraints on the construction of a lexical system of color terms, but these constraints would allow a large variety of color categories, which would depend on the history of our languages as well as our cultures, and perhaps partly, on our physical environments.

6. Conclusions

The neurosemantic enterprise can be considered a child of our neuroscientific era. It tackles long standing questions regarding the relationship between language and the mind, but it does so, not only in a brain-inspired but brain-informed manner. Only in recent years has it become feasible to construct neurocomputational models that reflect brain representational mechanisms, as those briefly touched upon here, with components corresponding, at least in part, to relevant cortical and subcortical areas of the brain. Future progress in the field will have to address further the neuromechanics of the brain and provide explanations on the why and how different areas of the brain (as well as the body through sensorimotor processes) are involved in the construction of linguistic meaning and the representation and processing of that meaning. Of course, a number of challenges are already present and even more are waiting, but the future for a consolidated and mature neurosemantics looks promising.

Bibliography

Bechtel, W., & Richardson, R. C. (1993), *Discovering complexity: Decomposition and localization as scientific research strategies*, Princeton University Press, Princeton (NJ).

BLACK, A. W., & TAYLOR, P. A. (1997), *The festival speech synthesis system: System documenta- tion*. Technical report HCRC/TR-83, Human Communication Research Centre, University of Edinburgh, Edinburgh.

BRAILLARD, P.-A., & MALATERRE, C. (2015), *Explanation in biology: An introduction*, in P.-A. BRAILLARD & C. MALATERRE (eds.), *Explanation in biology, history, philosophy and theory of the life sciences*, Springer-Verlag, Berlin, pp. 1-28.

BREIDBACH, O. (2007), «Neurosemantics, neurons and system theory», in *Theory Bioscience*, 126, pp. 23-33.

CHALUPA, L., & WERNER, J. (eds.), (2003), *The visual neurosciences*, MIT Press, Cambridge (MA).

CHURCH, A. (1941), *The calculi of lambda conversion*, Princeton (NJ), Princeton University Press.

CHURCHLAND, P. M. (1989), *A neurocomputational perspective: The nature of mind and the structure of science*, MIT Press, Cambridge (MA).

CHURCHLAND, P. S., & SEJNOWSKI, T. (1994), *The computational brain*, MIT Press, Cambridge (MA).

CHURCHLAND, P. M. (2001), *Neurosemantics: On the mappings of minds and the portrayal of worlds*, in K. WHITE (ed.), *The emergence of mind*, Fondazione Carlo Erba, Milan.

CRAVER, C. F. (2007), *Explaining the brain: mechanisms and the mosaic unity of neuroscience*, Oxford University Press, Oxford (UK).

CUMMINS, R. (1989), *Meaning and mental representation*, MIT Press, Cambridge (MA).

DAYAN, P., & ABBOTT, L. F. (2001), *Theoretical neuroscience*, MIT Press, Cambridge (MA).

DEHAENE, S. (2014), *Consciousness and the brain: Deciphering how the brain codes our thoughts*, Viking Adult, New York.

DOUGLAS, R. J., MARTIN, K. A., & WHITTERIDGE, D. (1989), «A canonical microcircuit for neocortex», in *Neural Computation*, 1, pp. 480-488.

FELDMAN, J. A. (2006), *From molecule to metaphor: A neural theory of language*, MIT Press, Cambridge.

FELLEMAN, D. J., & VAN ESSEN, D. C. (1991), «Distributed hierarchical processing in the primate cerebral cortex», in *Cerebral Cortex*, 1, pp. 1-47.

FODOR, J. (1975), *The language of thought*, Harvard University Press, Cambridge (MA).

FODOR, J. (1981), *Representations: Philosofical essay on the foundation of cognitive science*, MIT Press, Cambridge (MA).

FODOR, J. (1987), *Psychosemantics: The problem of meaning in the philosophy of mind*, MIT Press, Cambridge (MA).

FODOR, J. (1990), *A theory of content and other essays*, Cambridge University Press, Cambridge (UK).

FRESCO, N. (2014), *Physical computation and cognitive science*, Springer-Verlag, Berlin.

FUSTER, J. M. (2008), *The prefrontal cortex*, Academic Press, New York 2008^4.

HINTON, G. E., MCCLELLAND, J. L., & RUMELHART, D. E. (1986), *Distributed representations*, in D. E. RUMELHART & J. L. MCCLELLAND (eds.), *Parallel distributed processing: Explorations in the microstructure of cognition*, vol. 1, pp. 77-109, MIT Press, Cambridge (MA).

HUBEL, D., & WIESEL, T. (1959), «Receptive fields of single neurones in the cat's striate cortex», in *Journal of Physiology*, 148, pp. 574-591.

HUBEL, D., & WIESEL, T. (1968), «Receptive fields and functional architecture of mokey striate cortex», in *Journal of Physiology*, 195, pp. 215-243.

HUME, D. (1748), *An enquiry concerning human understanding*, A. Millar, London.

KAPLAN, D. M. (2011), «Explanation and description in computational neuroscience», in *Synthese*, 183, pp. 339-373.

KAPLAN, D. M., & CRAVER, C. F. (2011), *Towards a mechanistic philosophy of neuroscience*, in S. FRENCH & J. SAATSI (eds.), *Continuum companion to the philosophy of science*, pp. 268-292, Continuum Press, London.

Kleene, S. C. (1936), «General recursive functions of natural numbers», in *Mathematische Annalen*, 112, pp. 727-742.

KOHONEN, T., & HARI, R. (2000), «Where the abstract feature maps of the brain might come from» in *Trends in Neurosciences*, 22, pp. 135-139.

KRANTZ, D., LUCE, D., SUPPES, P., & TVERSKY, A. (1971), *Foundations of measurement – volume i additive and polynomial representations*, Academic Press, New York.

KRUBITZER, L. (1995), «The organization of neocortex in mammals: are species differences really so different?» in *Trends in Neuroscience*, 8, pp. 408-417.

LUCE, D., KRANTZ, D., SUPPES, P., & TVERSKY, A. (1990), *Foundations of measurement – volume iii representation, axiomatization, and invariance*, Academic Press, New York.

MACHAMER, P., DARDEN, L., & CRAVER, C. F. (2000), «Thinking about mechanisms», in *Philosophy of Science*, 67, pp. 1-84.

MASTRONARDE, D. N. (1983), «Correlated firing of retinal ganglion cells: I. Spontaneously active inputs in X- and Y-cells», in *Journal of Neuroscience*, 14, pp. 409-441.

MIŁKOWSKI, M. (2013), *Explaining the computational mind*, MIT Press, Cambridge (MA).

MORGAN, A. (2014), «Representations gone mental», in *Synthese*, 191, pp. 213-244.

MOUNTCASTLE, V. (1957), «Modality and topographic properties of single neurons in cats somatic sensory cortex», in *Journal of Neurophysiology*, 20, pp. 408-434.

NELSON, R. J. (2002), *The somatosensory system: deciphering the brain's own body image*, CRC Press, Boca Raton (FL).

PICCININI, G. (2015), *Physical computation: A mechanistic account*, Oxford University Press, Oxford (UK).

PLEBE, A., & DE LA CRUZ, V. M. (2014), *Color seeing and speaking – effects of biology, environment and language*, in ANDERSON W., BIGGAM, C. P., HOUGH, C., & KAY, C. (eds.), *Colour studies. A broad spectrum*, John Benjamins, Amsterdam, pp. 291-306.

PLEBE, A., & DE LA CRUZ, V. M. (2016), *Neurosemantics – neural processes and the construction of linguistic meaning*, Springer, Berlin.

PULVERMÜLLER, F. (2002), *The neuroscience of language: On brain circuits of words and serial order*, Cambridge University Press, Cambridge (UK).

PULVERMÜLLER, F. (2012), «Meaning and the brain: The neurosemantics of referential, interactive, and combinatorial knowledge», in *Journal of Neurolinguistics*, 25, pp. 423-459.

PUTNAM, H. (1960), *Minds and machines*, in HOOK S. (ed.), *Dimensions of mind*, New York University Press, New York.

RAMSEY, W. M. (2007), *Representation reconsidered*, Cambridge University Press, Cambridge (UK).

RUSSELL, B. (1927), *The analysis of matter*, Harcourt, London.

RYDER, D. (2004), «Sinbad neurosemantics: A theory of mental representation», in *Minds and Machines*, 19, pp. 211-240.

SEJNOWSKI, T. J., Koch, C., & Churchland, P. S. (1988), «Computational neuroscience», in *Science*, 241, p. 1299.

SHAGRIR, O. (2012), «Structural representations and the brain», in *British Journal for the Philosophy of Science*, 63, pp. 519-545.

SPREVAK, M. (2011), «William m. ramsey, representation reconsidered», in *British Journal for the Philosophy of Science*, 62, pp. 669-675.

STOCKMAN, A., & BRAINARD, D. H. (2010), *Color vision mechanisms*, in BASS M. (ed.), *OSA handbook of optics*, McGraw Hill, New York, Sections 11.1–11.104.

SUPPES, P., KRANTZ, D., LUCE, D., & TVERSKY, A. (1989), *Foundations of measurement – volume ii geometrical, threshold, and probabilistic representations*, Academic Press, New York.

SWOYER, C. (1991), «Structural representation and surrogative reasoning», in *Synthese*, 87, pp. 449-508.

THIVIERGE, J.-P., & MARCUS, G. F. (2007), «The topographic brain: from neural connectivity to cognition», in *Trends in Neuroscience*, 30, pp. 251-259.

TOOTELL, R. B., SILVERMAN, M. S., HAMILTON, S. L., SWITKES, E., & DE VALOIS, R. (1988), «Functional anatomy of the macaque striate cortex. V. spatial frequency», in *Journal of Neuroscience*, 8, pp. 1610-1624.

TOOTELL, R. B., SWITKES, E., SILVERMAN, M. S., & HAMILTON, S. L. (1988), «Functional anatomy of the macaque striate cortex. III. color», in *Journal of Neuroscience*, 8, pp. 1531-1568.

TURING, A. (1936), «On computable numbers, with an application to the Entscheidungsproblem», in *Proceedings of the London Mathematical Society*, 42, pp. 230-265.

VERPLAETSE, J., SCHRIJVER, J. D., VANNESTE, S., & BRAECKMAN, J. (eds.). (2009), *The moral brain essays on the evolutionary and neuroscientific aspects of morality*, Springer-Verlag, Berlin.

WIESEL, T., & HUBEL, D. (1965), «Binocular interaction in striate cortex of kittens reared with artificial squint», in *Journal of Neurophysiology*, 28, pp. 1041-1059.