

Neurodynamics of Consciousness

Diego Cosmelli, Jean-Philippe Lachaux, and Evan Thompson

Forthcoming in Philip David Zelazo, Morris Moscovitch, and Evan Thompson, eds., *The Cambridge Handbook of Consciousness*, 2007

Abstract

One of the main outstanding problems in the cognitive sciences is to understand how ongoing conscious experience is related to the workings of the brain and nervous system. Neurodynamics offers a powerful approach to this problem, because it provides a coherent framework for investigating change, variability, complex spatiotemporal patterns of activity, and multi-scale processes (among others). In this chapter, we advocate a neurodynamical approach to consciousness that integrates mathematical tools of analysis and modeling, sophisticated physiological data recordings, and detailed phenomenological descriptions. We begin by stating the basic intuition: Consciousness is an intrinsically dynamic phenomenon, and must therefore be studied within a framework that is capable of rendering its dynamics intelligible. We then discuss some of the formal, analytical features of dynamical systems theory, with particular reference to neurodynamics. We then review several neuroscientific proposals that make use of dynamical systems theory in characterizing the neurophysiological basis of consciousness. We continue by discussing the relation between spatiotemporal patterns of brain activity and consciousness, with particular attention to processes in the gamma frequency band. We then adopt a critical perspective and highlight a number of issues demanding further treatment. Finally, we close the chapter by discussing how phenomenological data can relate to and ultimately constrain neurodynamical descriptions, with the long-term aim being to go beyond a purely correlational strategy of research.

Keywords

Consciousness, neurodynamics, dynamical system theory, complex brain patterns, large-scale integration, gamma band, phase synchrony, phenomenological descriptions.

1. The Intuition

The central idea of this chapter is the notion of dynamics—the dynamics of neural activity, the dynamics of conscious experience, and the relation between them.

‘Dynamics’ is a multifarious concept. In a narrow sense, it refers to the change a circumscribed system undergoes in some time-dependent descriptive variable, for example a neuron’s membrane voltage (Abarbanel & Rabinovich, 2001). In a wide sense, ‘dynamics’ indicates a field of research concerned with nonlinear dynamical systems. Such systems range from mathematical models to experimental problems to actual concrete world systems (Van Gelder, 1999). Finally, in a more intimate sense, ‘dynamics’ refers back to the temporal nature of our observations themselves, and thus to our conscious experience and how it is deployed in time (Varela, 1999). The interplay of these different senses of dynamics is at the heart of this chapter.

What exactly are the properties of dynamical systems and why are they of interest in relation to consciousness? The entirety of this chapter will be concerned with this question, but to begin addressing it, we wish, in this introductory section, first to give an overview of the basic intuition underlying dynamical approaches to consciousness.

Briefly stated, a complex or nonlinear dynamical system can be described, at any time, by a position in a high-dimensional state space (Nicolis & Prigogine, 1989). The n coordinates of such a position are the values of the set of n variables that define the system. This position changes in time, and thus defines a trajectory, which will tend to explore a sub-region of the total state space. One can then measure the distance between any two points of the trajectory and show that under certain circumstances the trajectory can exhibit spontaneous recurrence: Small portions of the state space will be explored over and over, but never along the exact same path. When perturbed by external events, such a system will change its trajectory in a way that is never quite the same and that depends on its position in the state space at the time of the perturbation. Given this feature, plus the system’s extreme sensitivity to initial conditions, the system’s response to perturbations will be unpredictable in practice. It is therefore quite difficult to control such a system and constrain its movement along a pre-defined trajectory. For example, in the case of chaotic systems, such control involves applying a continuous succession of carefully chosen, delicate inputs, brute force usually being inefficient (think of what happens when you try to catch a fish out of water) (Garfinkel, Spano, Ditto, & Weiss, 1992; Schiff et al., 1994). In general terms, there will be a certain degree of dissociation between the observed behavior of these

systems and the patterns of external constraints that can be imposed on them. In other words, these systems exhibit a certain degree of autonomy: When external perturbations cease, the system goes on; when external perturbations become stationary, the system does not, but continues to move. Somehow we intuitively recognize such systems as ‘animated’ or ‘alive’, in contrast to simpler systems that respond in a linear and predictable way to external control (e.g., a stone that flies twice the distance when thrown twice as strong). Thus, such systems exhibit an intrinsic variability that cannot be attributed to noise, but appears to be constitutive of their functioning. Moreover, in the case of certain types of complex dynamical systems, one can reveal a characteristic spatiotemporal balance of functional segregation and cooperative integration.¹ This balance depends on the actual architecture of the system (its internal connectivity, for example), and is revealed in the transient establishment of distributed couplings among separated subsystems that in themselves present local encapsulated dynamics. Finally, some of these systems display what has been termed ‘self-organization’, that is, the emergence of collective coherent behavior starting from random initial conditions. This last feature, although not necessary for a system to be considered dynamical, has proven especially interesting when dealing with biological phenomena (Haken, 1983; Kelso, 1995).

The brain is a major case in point. The nervous system is a complex dynamical structure, in which individual neurons have intrinsic activity patterns and cooperate to produce coherent collective behavior (Llinas, 1988). The explosion of neuroimaging studies in the last fifteen years, as well as the substantial amount of data produced by electrophysiological techniques since the beginning of the twentieth century, have shown that the brain is never silent, but always in a state of ongoing functioning (Wicker, Ruby, Royet, & Fonlupt, 2003). The nervous system has a domain of viability, of allowed functioning, but within this domain it explores a multiplicity of possible states in a recurrent, yet always changing, manner (Palus, 1996). Incoming events are not sufficient to determine the system’s behavior, for any incoming event will change the system’s activity only as a result of how the system, given its current activity, responds to that event (Engel, Fries, & Singer, 2001b).

¹ We will define in more detail the notions of functional segregation and cooperative interaction in Section 4. Here we will only say that they can be considered analogous to local specialization and collective interaction, respectively.

If we now follow the thread of dynamics back to our own conscious experience, we can immediately notice that our consciousness manifests subjectively as a kind of continuously changing or flowing process of awareness, famously called the ‘stream of consciousness’ by William James (James, 1981). Our experience is made up of recurring perceptions, thoughts, images, and bodily sensations; and yet, however similar these events may be over time, there is always something new to each one, something ultimately unpredictable to every forthcoming moment. We can try to plan our day as strictly as we want, but the wanderings of our minds and how we react to the encounters we have in the actual world are things we cannot fully control. There seems to be an endogenous, spontaneous, ongoing flow to experience that is quite refractory to external constraints (Hanna & Thompson, 2003). Indeed, this dissociation can easily be made evident from the first-person perspective. If you sit down and close the windows, turn off the lights and close your eyes, so that external stimulation is greatly reduced *for* you, there is nevertheless still something going on subjectively *in* you, with an apparent temporal dynamics all its own.² Furthermore, at any moment, consciousness appears diverse and complex, rich with multiple, synchronous, and local contents (images, expectations, sounds, smells, kinesthetic feelings, etc.), yet it seems to hold together as a coherent and globally organized experience.

This intuitive convergence of complex dynamical patterns in experience and in brain activity is highly suggestive. It suggests that the framework of dynamical systems theory could offer a valuable way of bridging between the two domains of brain activity and subjective experience. If we wish to study the neurobiological processes related to consciousness, then we must provide a description of these processes that is (somehow) compatible with the dynamics of lived experience. On the other hand, dynamical aspects of experience might serve as a leading clue for uncovering and tracking the neurobiological processes crucial for consciousness.

In the rest of this chapter, we will explore this guiding intuition through a discussion of the following topics: formal dynamical systems, neurobiological theories based on dynamical systems principles, and the attempt to distinguish dynamical structures within experience that can constrain how we study the neurobiological basis of consciousness.

² This situation is, of course, only suggestive. A dream state might be a more rigorous case of a true sensory filter (Llinas & Pare, 1991).

2. Neurodynamics

2.1. Dynamical systems

Dynamical cognitive science has been defined as “a confederation of research efforts bound together by the idea that natural cognition is a dynamical phenomenon and best understood in dynamical terms” (Van Gelder, 1999, p. 243). Within this confederation, the job of the neurodynamicist is to model the neural basis of cognition using the tools of dynamical systems theory. Thus, the first thing we need to do is to define more precisely the notion of a dynamical system.

A dynamical system is a collection of interdependent variables that change in time. The state of the system at any time t is defined by the values of all the variables at that time; it can be represented by a position in an abstract ‘state space’, whose coordinates are the values of all the variables at t . The system’s behavior consists of transitions between states, and is described geometrically by a trajectory in the state space, which corresponds to the consecutive positions the system occupies as time passes.

At a first level of complexity, in the context of neurodynamics, we can think of the variables as being the membrane potentials of each individual neuron of the nervous system.³ These membrane potentials are obviously interdependent. Thus, at this level, the state of the nervous system at any time t would be defined by the value of all the membrane potentials at time t .

Although a dynamical system, in the most general terms, is any system that changes in time, dynamical systems theory gives special attention to *nonlinear* dynamical systems. The behavior of such systems is governed by nonlinear equations; in other words, some of the

³ One might rightly consider other variables, such as the local concentrations of certain neurotransmitters. Quantitative measures of the glial system should probably also be included. In fact, there is no one single way of choosing which variables to include in the system. This choice is largely driven by our current knowledge of the nervous system, which unfortunately remains quite limited. As a starting point, one needs to keep the following three elements in mind when choosing the variables of the system: (1) The time scale: if one candidate variable maintains a constant value during the time of observation of the system, then it does not need to be counted as a variable, but rather can be considered as a parameter (see below). (2) The spatial scale: the nervous system can be modeled at several scales—molecules, neurons, neural populations, etc. The variables should be meaningful at the spatial level of investigation. (3) The interdependence within the system: if the value of one candidate variable is fully determined by the values of the other variables, then it does not need to be included in the system.

mathematical functions used to derive the system's present state from its previous states and possible external inputs are nonlinear functions (for neurobiological examples, see Abarbanel & Rabinovich, 2001; Faure & Korn, 2001). Nonlinearity can endow the system with certain interesting properties. For example, when convection cells in a horizontal water layer are submitted to a thermal gradient above a critical value, the motions of billions of molecules spontaneously organize into long-range correlated macroscopic structures (Chandrasekhar 1961, cited in Le Van Quyen, 2003, p. 69; see also Kelso, 1995). Such properties of nonlinear systems led in the 1970s to an increased interest in the mathematics of dynamical system theory (Nicolis & Prigogine, 1977).

Several elements condition the behavior of a dynamical system during a given window of observation. Firstly, the system's behavior is conditioned by the values of a set of so-called *order parameters*. By definition, these parameters determine the exact mathematical equations that govern the system. This set of parameter values is a function of the architecture of the system (e.g., the synaptic weights between neurons), factors external to the system (e.g., outside temperature), and so on. These parameters cannot necessarily be controlled, and their dynamics is slower than that of the system itself. They can be considered as constant during the given window of observation, but are potentially variable across different observation periods. Their dynamics thus contrasts with that of the external inputs, which can have a dynamics as fast as that of the system. The governing equations of the system are also a function of these inputs, but the temporal evolution of the external inputs cannot be predicted from those equations (otherwise they could be considered as state variables of the system). Finally, all real systems include a noise component, which also counts as a factor in the governing equations, and thus affects the trajectory of the system.

2.2. Neurodynamics and the dynamical approach in neuroscience

Neurodynamics emerged from the proposal, which can be traced back to Ashby in the 1950s (Ashby, 1952), that the nervous system can be described as a nonlinear dynamical system. Although simple in appearance, this proposal deserves some attention: What does the nervous system look like from a dynamical point of view, and why is it nonlinear? The majority of the dynamical models of the nervous system describe the temporal evolution of the membrane potentials of neurons (Arbib, 2002). The behavior of any neuron of the system is a function of

both its own history of activation and the history of activation of every other neuron, thanks to the intrinsic connectivity of the nervous system. The precise influence of a given neuron on a second one is determined by the weight of the synapse that links them. Thus, the overall synaptic pattern in the nervous system provides the main set of order parameters in such models (see Arbib, 2002).⁴ To this description, it must be added that the system is not isolated, but under the constant influence of external sensory inputs that shape the behavior of peripheral sensory neurons.

There are many models available for the mathematical functions that link the membrane potentials of individual neurons to the history of the larger system and to the external inputs (Arbib, 2002). At this point, it is sufficient to state that these functions are nonlinear, and that this is the reason why the nervous system is described as a (spatially extended) nonlinear dynamical system (for reviews, see Faure & Korn, 2001; Korn & Faure, 2003).⁵

2.3. Chaos in the brain

As Le Van Quyen (2003, p. 69) notes, there was little echo to Ashby's original proposal to view the nervous system as a nonlinear dynamical system, mostly because the appropriate mathematical methods and computational tools to pursue this proposal were lacking at that time. The real boost to neurodynamics came later, in the 1980s and 90s, with the widespread emergence in the scientific community of interest in the properties of chaotic systems.

Chaotic systems are simply nonlinear systems, with their parameters set so that they possess an extreme sensitivity to initial conditions. Such sensitivity means that if one changes the initial position of the system in its state space, however slightly, the subsequent positions on the modified trajectory will diverge exponentially from what they would have been otherwise. Given this sensitive dependence on initial conditions, combined with the impossibility of determining the present state of the system with perfect precision, the future behavior of a chaotic

⁴ Alternative definitions of the nervous system as a dynamical system can include the synaptic weights themselves among the variables. At certain time-scales, the weights are a function of the evolution of the membrane potentials, for instance via Long-Term Potentiation (LTP) mechanisms. Models that include a changing connectivity can quickly become unmanageable, however, both mathematically and computationally (but see Ito & Kaneko, 2002).

⁵ For further details on this subject, we strongly recommend one of the original and most influential sources in neurodynamics, by Walter Freeman (Freeman, 1975a). See also Bressler & Kelso, 2001.

system is unpredictable. The system thus appears to have an inherent source of variability, for it will never react twice in the same way to identical external perturbations, even in the absence of noise.⁶

The possible existence of ‘chaos in the brain’ sparked much speculation and excitement. There were two related matters of debate: (i) whether the nervous system is actually chaotic (or whether there are subsystems in the brain that are) (Faure & Korn, 2001; Korn & Faure, 2003); and (ii) what use the nervous system could make of such chaotic behaviors (Faure & Korn, 2001; Korn & Faure, 2003; Skarda & Freeman, 1987; Tsuda, 2001).

The second question proved to be the easiest. One property of chaotic systems is that their dynamics can organize around the presence of ‘strange’ attractors. A strange attractor is a pattern of activity that captures nearby states (Arbib, Érdi, & Szentágothai, 1997): It occupies a subregion of the state space, a manifold, and if the trajectory of the system comes in contact with this manifold, the trajectory will stay on it subsequently, in the absence of external perturbations. The precise number and shapes of the attractors are determined by the parameters of the system, such as the intrinsic connectivity of the nervous system. Which particular attractor captures the system is determined by the system’s initial position. When the parameters define several strange attractors, then there can be associations between certain initial positions in the state space and certain attractors (Tsuda, 2001). This association is the basis for chaos-based perceptual systems: For example, in a common nonlinear model of olfactory processing (as reviewed in Korn & Faure, 2003), each odor is represented by a specific attractor, such that when confronted with slightly different olfactory stimuli, the trajectory of the system will converge onto the same attractor, if the stimuli actually correspond to the same odorant.⁷ Thus, in this model, perception is based on several coexisting attractors in a multistable system. An additional and important feature is that external perturbations to the system can make the system jump from one attractor to another; therefore, chaotic systems should not be thought of as ‘static’ and unreactive to the

⁶ For this reason, it would not necessarily be meaningful to repeat the same perturbation over and over to study the average reaction of a chaotic system, for there is no guarantee that this average reaction would have any meaning. Yet, such averaging procedures are the basis of almost all the imaging studies of the nervous system.

⁷ Note that the exact reaction of the system, that is, the precise trajectory that the system will follow to converge on the attractor, can be very different from one olfactory stimulation to another, even though the target attractor is the same for all. Therefore, the existence of attractors is compatible with the intrinsic variability of chaotic systems.

environment. Moreover, chaotic systems can be controlled, that is, they can be ‘forced’ to stay within specific portions of the state space via external perturbations. It must immediately be added, however, that the term ‘forced’ is misleading, for the external perturbations are nothing like brute force; rather, they must be thought of as like a series of subtle touches, carefully chosen to adapt to the system’s dynamic properties.

A sobering thought is that it is not clear whether the activity of the nervous system, considered as a whole, is chaotic. One requirement for a system to be chaotic is that its trajectory in the state space be constrained within geometrical structures that have a lower dimension than the space itself (this requirement is needed mainly to distinguish between chaotic and stochastic processes) (Wright & Liley, 1996). Unfortunately, the behavior of the nervous system cannot be observed directly in its actual state space, but only via limited sets of measurements that are crude projections of its actual state (like the electroencephalogram or EEG, that retains only the average activity of millions of neurons (Korn & Faure, 2003)).⁸

Nevertheless, the debate over the chaotic nature of brain activity proved productive and brought out of the shadows some ideas crucial to neurodynamics. For instance, central to neurodynamical thought is the idea that the variability of neural activity may be an integral part of the nervous system’s dynamics. This notion is orthogonal to a number of traditional (and still largely dominant) approaches in neuroscience that attribute this variability to ‘meaningless noise’. As a case in point, most brain imaging studies try to get rid of the variability of the neural activity by averaging brain recordings over multiple repetitions of the same process.⁹ These averaging procedures most likely give an oversimplified view of brain dynamics. In the near future, neuroscientists will undoubtedly have to go to the trouble of making sense of the neural variability by finding its experiential and behavioral correlates. Fortunately, new approaches

⁸ A recent review notes that “Incontrovertible proof that EEG reflects any simple chaotic process is generally lacking. There are grounds for reservation concerning reports of the dimensionality of EEG from direct measurement. Fundamental difficulties lie in the applicability of estimation algorithms to EEG data because of limitation in the size of data sets, noise contamination, and lack of signal stationarity” (Wright & Liley, 1996).

⁹ In human electrophysiology, for instance, the dominant paradigm consists in recording the EEG of human subjects while presenting them with series of similar sensory stimulations. The signal studied is the evoked potential, that is the mean EEG response averaged over all the stimulations: the intertrial variability is considered as noise and disappears in the averaging procedure.

along these lines are emerging, such as trying to understand the brain response to sensory stimulation in the context of the brain's active state preceding the stimulation, and thus in relation to an active 'baseline' of neural activity that is far from neutral (Lutz, Lachaux, Martinerie, & Varela, 2002). This shift in the focus of brain imaging should not be underestimated: The brain's reaction is no longer viewed as a passive 'additive' response to the perturbation, but as an active 'integration' of the perturbation into the overall dynamics (Arieli, Sterkin, Grinvald, & Aertsen, 1996). In other words, the processing of an incoming sensory stimulation is no longer viewed as the simple triggering of a systematic, pre-specified, chain of neural operations that would unfold independently of the brain's current activity, as in a computer algorithm. For this reason, the neurodynamical approach is often presented as a sharp alternative to the computer metaphor of the brain (Freeman, 1999a; Kelso, 1995; van Gelder, 1998).

2.4 Self-organization and the emergence of spatiotemporal patterns

As Crutchfield and Kaneko (Crutchfield & Kaneko, 1987) note, dynamical systems theory has developed largely through the study of low-dimensional systems, with no spatial extension. To be useful for neuroscience, however, dynamical systems theory needs to consider the special properties conferred on the nervous system by its spatial extension.

Fortunately, there has been a recent coincidence between, on the theoretical side, the development of a theory of large-scale nonlinear systems, and on the experimental side, the advent of multi-electrode recordings and imaging techniques to map precisely the electrical activity of entire populations of neurons. This coincidence has led to renewed interest, in the biological community, in large-scale models of neural activity.

The study of large, spatially extended nonlinear systems is a field in itself, in which the interest in attractors shifts to the related one of spatiotemporal patterns. (We recommend the reader spend a few minutes looking for pictures of 'cellular automata' with the Google image search engine to see some beautiful examples of spatiotemporal structures.) As a result, the neurodynamical community is now becoming less focused on chaos and more focused on the properties of self-organization in nonlinear systems, and particularly the formation of transient

spatiotemporal structures in the brain.¹⁰ As noted by Freeman, in the preface to the second (electronic) printing of his seminal book, *Mass Action in the Nervous System* (Freeman, 1975): “The word ‘chaos’ has lost its value as a prescriptive label and should be dropped in the dustbin of history, but the phenomenon of organized disorder constantly changing with fluctuations across the edge of stability is not to be discarded” (Freeman, 2004).

Spatiotemporal structures are ubiquitous in the brain. Apart from the obvious physical construction of the system, they correspond to the emergence of transient functional couplings between distributed neurons. For a given period of time, the activity of a set of neurons shows an increased level of statistical dependency, as quantified, for example, by mutual information.¹¹ In pioneering work, Freeman (reviewed in Freeman, 2000a), observed spatiotemporal activity patterns in the olfactory bulb and interpreted them within the framework of dynamical systems theory. In an influential theoretical paper with Skarda, he proposed that sensory information was encoded in those patterns (Skarda & Freeman, 1987).

The classic example of spatiotemporal structures in the brain is the Hebbian reverberant cell assembly, which Hebb (1949) hypothesized as the basis for short-term memory (see Amit, 1994).¹² (This notion is also closely related to Varela’s (1995) idea of resonant cell assemblies, described below.) Reverberant cell assemblies are labile sets of neurons that transiently oscillate together at the same frequency, at the level of their membrane potential. They are the best studied spatiotemporal structures in the brain. Indeed, the cortex has sometimes been modeled as a lattice of coupled oscillators—in other words, as a juxtaposition of reverberant cell assemblies. One advantage of such models is that the behavior of oscillator lattices has been abundantly

¹⁰As Le Van Quyen (2003, p. 69) notes: “in physics, what is usually referred to as self-organization is the spontaneous formation of well organized structures, patterns, or behaviors, from random initial conditions. Typically, these systems possess a large number of elements or variables interacting in a complex way, and thus have very large state spaces. However, when started with some initial conditions, they tend to converge to small areas of this space which can be interpreted as a form of emergent eigenbehavior.”

¹¹ Mutual information quantifies the ability to predict the behavior of one element in the system from the behavior of one or several other elements (David, Cosmelli, & Friston, 2004). This measure is one of several tools used to quantify statistical dependence within a system. Note that what counts as a spatiotemporal structure will depend on which measure of statistical dependence is used.

¹² “It seems that short-term memory may be a reverberation in the closed loops of the cell assembly and between cell assemblies” (Hebb, as cited in Amit, 1994, p. 621).

investigated, mainly using numerical simulations (see Gutkin, Pinto, & Ermentrout, 2003; Kuramoto, 1984; Nunez, 2000; Wright & Liley, 1996).

The formation of spatiotemporal structures in such systems often takes the form of phase-synchronization patterns between the oscillators (Le Van Quyen, 2003).¹³ In the brain, phase-synchronization of large populations of neural oscillators can produce macroscale oscillations that can be picked up by mesoscale recordings (such as local field potentials) or macroscale recordings (such as EEG). For this reason, synchronous oscillations have been the easiest form of spatiotemporal structure to measure in the brain, and not surprisingly, the first one to be observed (Bressler & Freeman, 1980; Gray, Konig, Engel, & Singer, 1989) (see also the discussion of functional connectivity below).

One reason that resonant cell assemblies in particular, and spatiotemporal structures in general, are so appealing to neuroscience is because they provide a flexible and reversible way to bind together distributed neurons that may be primarily involved in very different functional processes. This type of binding has three fundamental features: (i) the ability to integrate distributed neural activities (*integration*); (ii) the ability to promote, and virtually extract, one particular set of neural activities above the rest of the brain activity (*segregation*); and (iii) the capacity to evolve easily through a succession of flexible and adaptive patterns (*metastability*). For example, one resonant assembly could transiently bind together the different populations of neurons involved in analyzing the shape, color, and motion of a visual object, and this temporary assembly would constitute a neural substrate for the transient perception of a visual object. This idea is the starting point of a very active stream of research that we will discuss later in this chapter (for an overview, see Roskies, 1999). Some authors have even proposed that every cognitive act corresponds to the formation of such a transient spatiotemporal pattern (Varela, 1995).

In summary, dynamical systems theory proposes a precise framework to analyze the spatiotemporal neural phenomena that occur at different levels of organization in the brain, such

¹³ A precise definition of synchronization between chaotic systems can be found in (Pikovsky, Zaks, Rosenblum, Osipov, & Kurths, 1997, p. 680): “The phase synchronization of a chaotic system can be defined as the occurrence of a certain relation between the phases of interacting systems or between the phase of a system and that of an external force, while the amplitudes can remain chaotic and are, in general, uncorrelated (see also Brown & Kocarev, 2000; Rosenblum, Pikovsky, & Kurths, 1996).

as the firing of individual neurons and the collective dynamics of synchronous oscillations within large networks. The future challenge is to relate these properties of self-organization to various aspects of mental life. This endeavor is still in its early phases, but the future looks promising. For example, recent analysis methods making use of the brain's dynamical properties have been proposed in order to anticipate epileptic seizures (Martinerie et al., 1998). There is now a strong general sense that the properties of metastable, neural spatiotemporal patterns match crucial aspects of conscious experience, and that neurodynamics may provide the tools and concepts to understand how the neural activity crucial for consciousness temporally unfolds. This trend is patent in a set of influential neuroscientific models of consciousness that we review in the next section of this chapter.

3. Examples of Neurodynamical Approaches to Consciousness

3.1. Introduction

Although neurodynamics is quite popular in the neurobiology of consciousness, it is still not widespread practice to formulate theories about the relation between consciousness and the brain in purely dynamical terms. Dynamical concepts are incorporated to varying degrees by different researchers and used alongside concepts from information theory or functionalist models of cognitive processing. In this section, we will review some models that make use of dynamical concepts in attempting to explain the phenomenon of consciousness. The list of models we cover is not meant to be exhaustive, but rather a small sample of a large spectrum of dynamical approaches to brain activity and consciousness. Furthermore, we do not intend to scrutinize these models in detail, but instead to highlight some common aspects, while providing an overview of their main proposals and hypotheses. The reader is referred to the original sources for more details.

3.2. Neurodynamical models of consciousness

3.2.1. Consciousness as order parameter and dynamical operator

Among the different approaches to neural dynamics, the pioneering work of Walter J. Freeman stands out as one of the most elaborate and truly dynamical theories of brain function (Freeman, 1975, 1999a, 1999b, 2000a, 2000b, 2004). His work is based mainly on animal studies, in particular electrophysiological recordings of the olfactory system of awake and behaving rabbits.

This approach can be summarized as follows: The point of departure is the neuronal population. A neuronal population is an aggregate of neurons, in which, through positive feedback, a state transition has occurred so that the ensemble presents steady-state, non-zero activity. When negative feedback is established between populations, where one is excitatory and the other inhibitory, oscillatory patterns of activity appear. This change implies a second state transition, where the resulting attractor is a limit cycle that reveals the steady-state oscillation of the mixed (excitatory-inhibitory) population. When three or more mixed populations combine among themselves by further negative and positive feedback, the resulting background activity becomes chaotic. This chaotic activity now distributed among the populations is the carrier of a spatial pattern of amplitude modulation that can be described by the local heights of the recorded waveform. When an input reaches the mixed population, an increase in the nonlinear feedback gain will produce a given amplitude-modulation pattern. The emergence of this pattern is considered to be the first step in perception: *Meaning* is embodied in these amplitude-modulation patterns of neural activity, whose structure is dependent on synaptic changes due to previous experience. Thus, the whole history of the animal sets the context in which the emerging spatiotemporal pattern is meaningful. Through the divergence and convergence of neural activity onto the entorhinal cortex, the pulse patterns coming from the bulb are smoothed, thereby enhancing the macroscopic amplitude-modulation pattern, while attenuating the sensory-driven microscopic activity. Thus, what the cortex ‘sees’ is a construction made by the bulb, not a mapping of the stimulus. Hence, in Freeman’s view, perception is an essentially active process, closer to hypothesis testing than to passive recovery of incoming information. This active stance is embodied in the process of ‘preference’ by which the limbic system (including the entorhinal cortex and the hippocampus in mammals), through corollary discharges to all sensory cortices, maintains an attentive expectancy of what is to come. The stimulus then confirms or disconfirms the hypothesis, through state transitions that generate the amplitude-modulation patterns described previously.¹⁴ The multisensory convergence onto the entorhinal cortex becomes the basis for the formation of Gestalts underlying the unitary character of perception. Finally, through

¹⁴ In Freeman’s words: “preference provides an order parameter that shapes the attractor landscapes, making it easier to capture expected or desired stimuli by enlarging or deepening the basins of their attractors. [...] corollary discharges do this by a macroscopic bias that tilts sensory attractor landscapes” (Freeman, 1999a, p.).

multiple feedback loops, global amplitude-modulation patterns of chaotic activity emerge throughout the entire hemisphere directing its subsequent activity. These loops comprise feedforward flow from the sensory systems to the entorhinal cortex and the motor systems, and feedback flow from the motor systems to the entorhinal cortex, and from the entorhinal cortex to the sensory systems. Such global brain states “emerge, persist for a small fraction of a second, then disappear and are replaced by other states” (Freeman, 1999b, p. 153).

For Freeman, it is this level of emergent and global cooperative activity that is crucial for consciousness, as these remarks indicate: “Consciousness... is a state variable that constrains the chaotic activities of the parts by quenching local fluctuations. It is an order parameter and an operator that comes into play in the action-perception cycle as an action is being concluded, and as the learning phase of perception begins” (Freeman, 1999a, p. 132). Furthermore: “[T]he globally coherent activity... may be an objective correlate of awareness... In this view, awareness is basically akin to the intervening state variable in a homeostatic mechanism, which is both a physical quantity, a dynamical operator, and the carrier of influence from the past into the future that supports the relation between a desired set point and an existing state” (Freeman, 1999b, p. 157).

3.2.2. Dynamic large-scale integration and radical embodiment

Another proposal that falls squarely within the neurodynamical framework is one formulated initially by Francisco J. Varela (Varela, 1995) and then developed with his collaborators, especially Evan Thompson (Thompson & Varela, 2001). Varela proposes to address the question of how neural mechanisms bring about “the flow of adapted and unified cognitive moments” (Varela, Lachaux, Rodriguez, and Martinerie, 2001, p. 229). The main working hypothesis is that a specific neuronal assembly underlies the operation of every unitary cognitive act. Here a neuronal assembly is understood as a distributed set of neurons in the brain that are linked through reciprocal and selective interactions, where the relevant variable is no longer single-neuron activity, but rather the dynamic nature of the links that are established between them. Varela and collaborators propose that such dynamical links are mediated by the transient establishment of phase relations (phase synchrony) across multiple frequency bands, especially in the beta (15-30 Hz) and gamma (30-80 Hz) range (Varela et al., 2001). Moreover, the transient nature of such dynamical links (and therefore of the neural assemblies themselves) is central to

the idea of large-scale integration, for it brings to the fore the notion that the system, rather than presenting a series of well defined states (attractors), shows metastable (self-limiting and recurrent) patterns of activity: “In the brain, there is no ‘settling down’ but an ongoing change marked only by transient coordination among populations, as the attractor itself changes owing to activity-dependent changes and modulations of synaptic connections” (Varela et al., 2001, p. 237). Large-scale integration through phase relations becomes fundamental for understanding brain dynamics as coordinated spatiotemporal patterns, and provides a plausible solution to the problem of how to relate the local specificity of activity in specialized cortical regions to the constraints imposed by the connectivity established with other distributed areas. We will see in Sections 4 and 5 below, as well as in the other dynamical approaches described in this section, that this balance of segregation and integration has been considered the hallmark of brain complexity, and as a plausible prerequisite for consciousness.

Thompson and Varela (2001) then qualify this view by placing it in a ‘radical embodiment’ framework. They propose that, although the neural processes relevant to consciousness are best mapped at the level of large-scale, transient spatiotemporal patterns, the processes crucial for consciousness are not brain-bound events, but comprise also the body embedded in the environment. By taking into account the notion of *emergent processes* as understood in complex systems theory (order parameters or collective variables and the boundary conditions they impose on local activities), they propose (i) that conscious awareness (as an order parameter or dynamical operator) is efficacious with respect to local neural events (see also Freeman 1999a, 1999b, and above); and (ii) that the processes crucial for consciousness so understood span at least three ‘cycles of operation’ that cut across brain-body-world divisions (Thompson & Varela, 2001, p. 424): (1) Regulatory organismic cycles, in which the maintenance of internal variables within a viable range is achieved “through sensors and effectors to and from the body that link neural activity to the basic homodynamic processes of internal organs.” This cycle is supposed to be the basis of the “inescapable affective backdrop of every conscious state,” also called ‘core consciousness’ (Damasio, 1998, 1999) or ‘primary-process consciousness’ (Panksepp, 1998). (2) Sensorimotor coupling between organism and environment, whereby what the organism senses is a function of how it moves and how it moves is a function of what it senses. Here, “transient neural assemblies mediate the coordination of sensory and motor surfaces, and sensorimotor coupling with the environment constrains and modulates this neural

dynamics.” (3) Cycles of intersubjective interaction, involving the recognition of the intentional meaning of actions and (in humans) linguistic communication. This last type of cycle depends on various levels of sensorimotor coupling, mediated in particular by the so-called mirror-neuron systems that show similar patterns of activation for both self generated, goal-directed actions and when one observes someone else performing the same action (Rizzolatti & Craighero, 2004)

As a final aspect of this proposal, Thompson and Varela hypothesize that “consciousness depends crucially on the manner in which brain dynamics are embedded in the somatic and environmental context of the animal’s life, and therefore there may be no such thing as a minimal internal neural correlate whose intrinsic properties are sufficient to produce conscious experience” (Thompson & Varela, 2001, p. 425; see also Noë & Thompson, 2004a).

3.2.3. Cortical coordination dynamics

Based on extensive work in human motor coordination, J.A. Scott Kelso has developed a detailed dynamical framework for understanding human cognition (Kelso, 1995). His main focus is the appearance of self-organized patterns, due to nonlinear interactions between system components, at both the neural and motor levels, as well as their role in human behavior. Kelso views the brain as fundamentally “a pattern forming self-organized system governed by potentially discoverable, nonlinear dynamic laws” (Kelso, 1995, p. 257). He proposes that cognitive processes “arise as metastable spatiotemporal patterns of brain activity that themselves are produced by cooperative interactions among neural clusters” (ibid). He then goes one step further, proposing that “an order parameter isomorphism connects mind and body, will and brain, mental and neural events. Mind itself is a spatiotemporal pattern that molds the metastable dynamic patterns of the brain” (Kelso, 1995, p. 288).

What are the specific neural mechanisms underlying the establishment of such self-organized patterns? Kelso, in collaboration with Steven Bressler, proposes that the answer lies in the notion of ‘coordination dynamics’ (Bressler & Kelso, 2001). Coordination dynamics is presented as an integrative framework, in which the main issue is “to identify the key variables of coordination (defined as a functional ordering among interacting components) and their dynamics (rules that govern the stability and change of coordination patterns and the nonlinear coupling among components that give rise to them)” (Bressler & Kelso, 2001, p. 26). Using this framework, Bressler and Kelso address the question of how interacting, distributed cortical areas

allow the emergence of ongoing cognitive functions. On the basis of previous studies of bimanual coordination, they propose more specifically that the relevant collective variable is the relative phase (the continuous phase difference) among the given neural structures, which are themselves considered to be accurately described by nonlinear oscillators. They argue that this coordination variable is adequate because (i) it reveals the spatiotemporal ordering between interacting structures; (ii) changes in the relative phase occur more slowly than changes in the local component variables; and (iii) relative phase shows abrupt changes during phase transitions or bifurcations. When the two coordinated local neuronal populations have different intrinsic frequencies, the relative phase shows a metastable regime in the form of ‘attractiveness’ towards preferred modes of coordination, without settling into any unique one. Accordingly, Bressler and Kelso propose that “a crucial aspect of cognitive function, which can both integrate and segregate the activities of multiple distributed areas, is large-scale relative coordination governed by way of metastable dynamics” (Bressler & Kelso, 2001, p. 30).

3.2.4. The ‘Dynamic Core’ hypothesis

Gerald M. Edelman and Giulio Tononi have developed an account of the neural basis of consciousness that aims to explain two fundamental properties of conscious experience, which they call ‘integration’ and ‘differentiation’ (Edelman & Tononi, 2000; Tononi & Edelman, 1998). Integration refers to the unitary character of conscious experience, whereby the multiplicity of aspects, such as color, taste, audition, kinesthetic sense, etc., come together in a unique coherent experience. Differentiation is the capacity to experience any of a vast number of different possible conscious states. This capacity is intimately tied to what Edelman and Tononi call the informativeness of conscious experience, where each conscious state would be highly informative, given the reduction in uncertainty that is accomplished by the selection of one among a potentially infinite number of possible states.

Edelman and Tononi stress that consciousness is not a thing, but a process, and therefore should be explained in terms of neural processes and interactions, and not in terms of specific brain areas or local activities. More specifically, they postulate that to understand consciousness it is necessary to pinpoint neural processes that are themselves integrated, yet highly differentiated. Their answer to this problem is what they call ‘the Dynamic Core hypothesis’ (Edelman & Tononi, 2000; Tononi & Edelman, 1998). In their words: “1) a group of neurons can

contribute directly to conscious experience only if it is part of a distributed functional cluster that achieves high integration in hundreds of milliseconds. 2) To sustain conscious experience, it is essential that this functional cluster be highly differentiated, as indicated by high values of complexity. We call such a cluster of neuronal groups that are strongly interacting among themselves and that have distinct functional borders with the rest of the brain at the time scale of fractions of a second a ‘dynamic core,’ to emphasize both its integration and its constantly changing composition. A dynamic core is therefore a process, not a thing or a place, and it is defined in terms of neural interactions, rather than in terms of specific neural locations, connectivity or activity” (Edelman & Tononi, 2000, p. 144). In addition: “The dynamic core is a functional cluster: its participating neuronal groups are much more strongly interactive among themselves than with the rest of the brain. The dynamic core must also have high complexity: its global activity patterns must be selected within less than a second out of a very large repertoire” (Tononi & Edelman, 1998, p. 1849). They hypothesize that the dynamic core achieves integration on the basis of reentrant interactions among distributed neuronal groups, most likely mediated by the thalamocortical system. Specifically, for primary consciousness to arise, interactions are required between sensory cortices in different modalities and value-category and memory systems in frontal, temporal, and parietal areas.¹⁵

Edelman and Tononi claim that the dynamic core provides a “neural reference space for conscious experience” (Edelman & Tononi, 2000, p. 164). They depict this space as an n -dimensional neural space, where the number of dimensions is given by the number of neuronal groups that are part of the dynamic core at that moment. Such neuronal groups would be segregated into neural domains specialized for various functions, such as form, color, or orientation discrimination, proprioceptive or somatosensory inputs, and so on, and they would be brought together through reentrant interactions. The local activities of these groups would therefore need to be understood in relation to the unified process constituted by the functional cluster, that is, the entire dynamic core: “The pure sensation of red is a particular neural state identified by a point within the N -dimensional neural space defined by the integrated activity of

¹⁵ Edelman and Tononi distinguish primary consciousness from higher-order consciousness. The former involves the capacity to construct a mental scene to guide behavior without the semantic, linguistic, and self-reflective capacities unique to the latter.

all the groups of neurons that constitute the dynamic core... The conscious discrimination corresponding to the quale of seeing red acquires its full meaning only when considered in the appropriate, much larger, neural reference space” (Edelman & Tononi, 2000, p. 167).

3.3. Related models

Several other authors have advanced models in which dynamical-system concepts are present, yet appear less explicitly. Nevertheless, these approaches also aim to describe the formation of spatiotemporal patterns of brain activity that are crucial for action, perception, and consciousness. Therefore, we believe that it is important to keep these models in mind as part and parcel of the wider research program of neurodynamics.

3.3.1. The cortico-thalamic dialogue

Rodolfo Llinas and his collaborators have proposed a model of how consciousness is related to brain activity, in which the notion of emergent collective activity plays a central role (Llinas & Pare, 1991; Llinas & Ribary, 2001; Llinas, Ribary, Contreras, & Pedroarena, 1998). In particular, Llinas postulates that consciousness arises from the ongoing dialogue between the cortex and the thalamus (Llinas & D. Pare, 1991). He calls attention to the fact that most of the input to the thalamus comes from the cortex, rather than from peripheral sensory systems. On this basis, he proposes that the brain be considered as a ‘closed system’ that can generate and sustain its own activity thanks to the intrinsic electrical properties of neurons (Llinas, 1988) and the connectivity they establish. The interplay of these two main characteristics underlies the establishment of “global resonant states which we know as cognition” (Llinas & Ribary, 2001, p. 167).

A crucial feature of this proposal is the precise temporal relations established by neurons in the cortico-thalamic loop. This temporal mapping is viewed as a ‘functional geometry’, and involves oscillatory activity at different spatial scales, ranging from individual neurons to the cortical mantle. In particular, 40-Hz oscillations that traverse the cortex in a highly spatially structured manner are considered as candidates for the production of a “temporal conjunction of rhythmic activity over large ensemble of neurons” (Llinas & Ribary, 2001, p. 168). Such gamma oscillations are believed to be sustained by a thalamo-cortical resonant circuit involving pyramidal neurons in layer IV of the neocortex, relay-thalamic neurons, and reticular-nucleus neurons. In particular, temporal binding is supposed to be generated by the conjunction of a

specific circuit involving specific sensory and motor nuclei projecting to layer IV and the feedback via the reticular nucleus, and a non-specific circuit involving non-specific intralaminar nuclei projecting to the most superficial layer of the cortex and collaterals to the reticular and non-specific thalamic nuclei. Thus, the ‘specific’ system is supposed to supply the *content* that relates to the external world, and the nonspecific system is supposed to give rise to the temporal conjunction or the *context* (on the basis of a more interoceptive context concerned with alertness). Together they generate a single cognitive experience (Llinas & Ribary, 2001, p. 173).

3.3.1. Timing and binding

Wolf Singer and collaborators have extensively investigated the issue of temporal correlations between cortical neurons and the role this phenomenon could play in solving what has been called ‘the binding problem’ (Engel & Singer, 2001; Gray et al., 1989; Singer & Gray, 1995). This is the problem of how the signals from the separate neuronal populations concerned with distinct object features (color, shape, motion, etc.) are bound together into a unified perceptual representation. The main idea behind their approach is that there is a “temporary association of neurons into functionally coherent assemblies that as a whole represent a particular content whereby each individual neuron is tuned to one of the elementary features of composite perceptual objects” (Singer, 1998, p. 1831). The specific hypothesis is neurons become members of such coherent assemblies through the precise synchronization of their discharges; in other words, such synchronization establishes a “code for relatedness” (Singer, 1998, p. 1837). Recently, Singer and colleagues (Engel & Singer, 2001, Engel et al., 1999) have extended this framework to the address the issue of phenomenal awareness. Their argument can be summarized as follows: (1) Brains capable of phenomenal awareness should be able to generate meta-representations of their own cognitive processes. (2) Meta-representations are realized by an iterative process, in which higher order cortical areas read low order (sensory) areas. (3) Combinatorial flexibility of meta-representations is obtained via dynamical cell assemblies. (4) Binding of such assemblies is effected by transient synchrony that establishes a code for relatedness among features and facilitates downstream evaluation and impact. (5) Such

assemblies need desynchronized EEG¹⁶ (which correlates with phenomenal awareness in the waking state and REM dream state, and shows high frequency beta and gamma oscillatory activity), and are facilitated by attention (Singer, 1998).

Engel, Fries, and Singer (Engel, Fries, & Singer, 2001; Engel & Singer, 2001) also explicitly espouse a ‘dynamicist view’ of brain function. According to this view, brain processes are not passive, stimulus-driven, and hierarchical, but active, context-dependent, endogenously driven, and distributed. In particular, “spatio-temporal patterns of ongoing activity... translate the functional architecture of the system and its pre-stimulation history into dynamic states of anticipation” (Engel et al., 2001, p. 705). In this dynamicist account of top-down influences, relevant patterns are generated as a result of continuous large-scale interactions, and these patterns can bias the saliency of sensory signals by changes in their temporal correlations. Endogenous, self-generated activity displays distinct spatio-temporal patterns, and these patterns bias the self-organizing process that leads to the temporal coordination of input-triggered responses and their binding into functionally coherent assemblies. This dynamicist approach thus stresses the importance of top-down influence in the form of large-scale dynamics that express contextual influences and stored knowledge in the system, and that can modulate local processing and hence the downstream effect of the impinging event (Engel et al., 2001).

3.3.2. The neural correlates of consciousness

Francis Crick and Christof Koch have employed dynamical concepts in a series of proposals regarding the relation between neural activity and conscious perception (Crick & Koch, 1990, 1998, 2003). In their view, the best way for the neuroscience of consciousness to proceed is first to uncover the neural correlates of consciousness (NCCs), in particular the neural contents of visual consciousness. They define a neural correlate of consciousness as a minimal set of neuronal events necessary and sufficient for a given state of phenomenal consciousness (see also

¹⁶ Here we are referring to desynchronized EEG in the classical sleep/wake cycle sense, in which desynchronized gamma and beta frequencies dominate the EEG of the waking state, by contrast with the synchronized slow wave delta frequency EEG of sleep. This notion of desynchronized and synchronized EEG as a whole should not be confused with the synchronization and desynchronization of particular EEG signals, which is more accurately termed phase-locking and phase-scattering respectively (see Section 4.4 below).

Chalmers, 2000). Here we summarize the version of their theory presented in one of their last joint articles on consciousness (Crick & Koch, 2003).

They begin with the notion of an ‘unconscious homunculus’, which is a system consisting of frontal regions of the brain ‘looking at the back, mostly sensory region’ (Crick & Koch, 2003, p. 120). Crick and Koch propose that we are not conscious of our thoughts, but only of sensory representations of them in imagination. The brain presents multiple unconscious processing modules, mostly feedforward, that act as ‘zombie’ modes. These modules present stereotyped responses in a sort of ‘cortical reflex’, whereas conscious modes are necessary only to deal with time-consuming, less stereotyped situations that need planning and decision-making. The most important point, however, with regard to the NCC issue, is the existence of dynamic coalitions of neurons, in the form of neural assemblies whose sustained activity embodies the contents of consciousness. Explicit representations of particular aspects of the (visual) scene are present in special brain regions (‘critical nodes’), and these representations are bound together in the dynamic neural coalitions. Additionally, Crick and Koch suggest that higher levels of cortical processing are first reached by the feedforward sensory sweep, and that only through back-propagation of activity from higher to lower levels do the lower levels gain access to this information. They distinguish ‘driving’ from ‘modulating’ connections, and suggest that the feedforward sweep is mostly driving activity in the frontal regions, whereas the feedback return onto sensory cortices is mainly modulating. In the specific case of conscious perception, they propose that it is not a continuous phenomenon, but rather that it works on the basis of a series of ‘snapshots’. Such snapshots are possibly related to alpha and theta rhythms, and are the reflection of a certain threshold that has been overcome (for a certain amount of time) by neural activity, enabling it to become conscious. Conscious coalitions would therefore be continually ‘forming, growing or disappearing’ (Crick & Koch, 2003, p. 122). Crick and Koch propose that attention is fundamental in biasing the competition among coalitions that share critical nodes. Attention produces the effective binding of different attributes of the given conscious content by means of shared ‘membership in a particular coalition’ (Crick & Koch, 2003, p. 123).

Although Crick and Koch recognize that the mechanism to establish such coalitions probably involves synchronous firing between distributed populations, they explicitly state that they do not longer believe that 40 Hz oscillatory activity is a sufficient condition for consciousness. Finally, they propose that there is a set of neural processes that, although not part

of the NCC, is affected by the NCC, both with respect to its actual firing and with respect to synaptic modifications due to previous experience. This ‘penumbra’ (Crick and Koch, 2003, p. 124), could eventually become conscious, if incorporated into the NCC.

3.3.3. Consciousness as global workspace

Stanislas Dehaene, Jean-Pierre Changeux and collaborators have explored an alternative model of the brain functioning that underlies the accessibility to verbal report of conscious experience (Dehaene, Kerszberg, & Changeux, 1998; Dehaene & Naccache, 2001; Dehaene, Sergent, & Changeux, 2003). The main proposal of their model is the existence of “two main computational spaces within the brain” (Dehaene, Kerszberg, & Changeux, 1998, p. 14529). The first computational space consists of a series of functionally segregated and specialized modules or processors that constitute a parallel distributed network (examples of modular processors would be primary visual cortex (V1) or the mirror neuron system in area F5 of the premotor cortex). The second computational space is not confined to a series of brain areas, but rather distributed among multiple cortical regions. The main property of this second space is massive reciprocal connectivity on the basis of horizontal projections (long-range cortico-cortical connections). Through descending connections, this ‘global workspace’ determines the contributions of the modular processors of the first computational space by selecting a specific set while suppressing another. Through this selective mobilization of the specialized processors into the global workspace, a ‘brain scale’ state can be reached, in which a group of workspace neurons are spontaneously coactivated, while the rest is suppressed. As a result, an exclusive ‘representation’ invades the workspace and “may remain active in an autonomous manner and resist changes in peripheral activity. If it is negatively evaluated, or if attention fails, it may however be spontaneously and randomly replaced by another discrete combination of workspace neurons. Functionally, this neural property implements an active ‘generator of diversity,’ which constantly projects and tests hypotheses (or pre-representations) on the outside world. The dynamics of workspace neuron activity is thus characterized by a constant flow of individual coherent episodes of variable duration” (Dehaene, Kerszberg, & Changeux, 1998, p. 14530). This postulated workspace has access to the world through ‘perceptual circuits’; ‘motor programming circuits’ enable action guidance; ‘long-term memory circuits’ enable access to past experiences; ‘evaluation circuits’ allow negative-positive judgments; and ‘attention circuits’ endow the

workspace with the capacity to alter its own activity separately from the influence of external inputs. Through connections with motor and language centers, the workspace makes its resident representation available for verbal report by the subject. Thus, Dehaene and Changeux see consciousness as a selective global pattern: “when a piece of information such as the identity of a stimulus accesses a sufficient subset of workspace neurons, their activity becomes self-sustained and can be broadcasted via long-distance connections to a vast set of defined areas, thus creating a global and exclusive availability for a given stimulus, which is then subjectively experienced as conscious” (Dehaene, Kersberg, & Changeux, 2003).

3.3.4. Summary

The majority of the approaches reviewed above stress the importance of a certain type of distributed, spatiotemporal pattern of neural activity that ‘demarcates’ itself from the background activity of the brain. Such patterns are described as ongoing, transient, metastable coordination processes among separate neurons, and they are considered to be crucial for the moment-to-moment emergence and formation of conscious experience. Another related feature crucial to several of the above approaches is that these spatiotemporal patterns reveal the interplay of two apparently fundamental principles of brain organization and function, namely, functional segregation and cooperative interaction or integration. This interplay and the dynamical properties of the brain’s spatiotemporal activity patterns will be the focus of the following sections.

4. The Search for Meaningful Spatiotemporal Patterns in the Brain

4.1. Introduction

Despite their significant differences, all the above models agree that the constitution of dynamic spatiotemporal patterns of neural activity plays a central role in the emergence of consciousness. This section discusses the practical aspects of the search for such patterns. After a short review of the connectivity of the brain, we discuss the detection of such patterns in real brain data. A short mathematical presentation leads us to the concept of synchrony, which is the preferred candidate to-date for such patterns.

4.2. Connectivity in the brain

The organization of the brain's connectivity is what ultimately determines the form of the neural spatiotemporal patterns. For this reason, it is useful to start with a review of some basic facts about this architecture.

The brain is probably one of the most complex biological systems we know (Edelman & Tononi, 2000). Its complexity is certainly due in great part to its histological and morphological structure, and one of the most striking aspects of the brain as a system is the connectivity pattern it exhibits. This pattern is that of a compact but distributed tissue, with local clusters of highly connected neurons that establish long-range interactions. In general, two neurons in the brain are always in interaction either directly or via a certain number of intermediate cells. It is useful to distinguish two levels of connectivity in the brain:

Local connections: Several types of neurons coexist in the neocortex. Within a given portion of neocortex, a complex arrangement of pyramidal, spiny stellate, and smooth stellate cells can be found. This arrangement of collateral axons, dendritic trees, and cell bodies gives rise to clusters of interconnected neurons that extend over a fraction of a millimeter. Neurons tend to organize into radial clusters that share functional characteristics, known as functional columns. These structures are particularly evident in somatic sensory cortex and visual cortex, and are believed to play a fundamental role in basic discriminative capacities.

Long-range connections: In addition to the bodies and dendritic trees of the local neurons, axons from deep structures and other cortical regions terminate at different points in the six-layered structure of the neocortex. Likewise, pyramidal neurons in a given region of the neocortex have axons that extend into the white matter, and reach both deep structures and other cortical regions. At least four patterns of long distance connectivity can be distinguished in the brain (Abeles, 1991): (i) between cortical neurons within one hemisphere; (ii) between cortical neurons of different hemispheres; (iii) between cortical neurons and deep nuclei; and (iv) between brainstem modulatory systems and extended areas of the cortex. In general, long-range connectivity obeys a reciprocity rule (Varela, 1995): If A projects to B, then B projects to A. This rule clearly favors the establishment of recursive loops. Nevertheless, some basal ganglia nuclei present a slightly different connectivity structure: Although they receive axons from cortical neurons, they project only through the thalamus into frontal lobe regions (Edelman & Tononi, 2000).

Interestingly, however, no one zone in the brain can be distinguished as the ultimate highest level, at least in terms of the connectivity patterns. Indeed, the massively interconnected nature of the brain suggests that dynamic relations between local and distant activities will *necessarily* be established whatever the observed origin of a given activation is. On the other hand, it is true that clusters of more strongly interconnected regions are evident. Stephan and collaborators recognize at least three main clusters in the primate cortex: (i) visual (occipito-temporal); (ii) somatomotor (mainly pre and post-central, but extending into parietal regions); and (iii) orbito-temporopolar-insular (Stephan et al., 2000). It is interesting that the overall structural connectivity (Hilgetag & Kaiser, 2004) and functional connectivity (Stephan et al., 2000) show a ‘small world’ architecture. Networks having such an architecture display remarkable properties, such as reduced average length path (reaching any node from any other node is accomplished in a minimal number of steps), high synchronizability, enhanced signal propagation speed, and stability (one can randomly eliminate links without affecting substantially the network properties) (Watts & Strogatz, 1998).

One of the most important conclusions of the study of the brain as a system is that despite its massive interconnectedness, the brain shows a strong segregation into clusters at both structural and functional levels. The interplay of these two characteristic features of the brain lies at the basis of one of the most interesting issues in contemporary neuroscience—the large-scale integration of brain activity and its role in the unified nature of experience (James, 1890; Varela, 1995; Von der Malsburg, 1981).

The combination of extensive neuropsychological studies since Broca and the explosive use of imaging techniques in the last fifteen years have highlighted two main principles of brain functioning (Edelman & Tononi, 2000; Friston, 2002a, 2002b, 2005). On the one hand, a functional encapsulation is evident: Distinct regions in the brain contribute differentially to different aspects of adaptive behavior; for instance, bilateral damage to the human homologue of V5/MT (in the middle temporal area) can lead to a restricted impairment in the capacity to discriminate movement (akinetopsia) (Zihl, Von Cramon, & Mai, 1983). On the other hand, for a given cognitive task, it is rarely the case that *only* one isolated region shows significant activation. For example, directing attention to a particular location of the visual field correlates with the concomitant activation of several cortical regions, preferentially right parietal, anterior cingulate, and occipital cortices (Mesulam, 1999).

Indeed, as we saw above, the connectivity pattern of the mammalian brain reveals a complex structure of recursively connected distant areas (Hilgetag & Kaiser, 2004; Stephan et al., 2000). Although the cortico-cortical connectivity pattern is paradigmatic, the structure of recursive connections is most prominently reflected in the thalamocortical matrix (Edelman & Tononi, 2000; Llinas & Ribary, 2001). For instance, the lateral geniculate nucleus (LGN) of the thalamus receives only around 5-10% of its inputs (not more than 20%) from the retina, while the remaining connections come from local inhibitory networks, descending inputs from layer VI of the visual cortex, and ascending inputs from the brainstem (Sherman & Guillery, 2002). Yet the LGN is the major relay in the visual pathway from the retina to the cortex. Such a complex structure of recursive, re-entrant, and interconnected networks that pervade the mammalian brain (Edelman & Tononi, 2000) strongly suggests the existence of constitutive cooperative interactions, and therefore integrative activity, among different regions.

Nevertheless, the presence of anatomical connectivity is not enough to explain effective interactions between separate regions (Friston, 2002a, 2002b). Indeed, besides being connected, it is necessary that such regions establish interdependent activation, in order to account adequately for the integration of functionally separate activity (Bressler, 1995). Thus, one task facing the neurodynamicist is to detect such interdependent activation from the brain recordings available today.

4.3 Detecting interdependent activations from real brain recordings

4.3.1. The data

The activity of the brain can be recorded at several different spatial and temporal scales. The neurodynamicist will be primarily interested in those techniques fast enough to follow the formation of spatiotemporal patterns in the hundreds of milliseconds time-scale. Because the construction of such patterns often involves activities at the millisecond time scale (e.g., in the case of fast neural oscillations), the desired temporal resolution is on this order of milliseconds. In practice, this excludes the neuroimaging techniques based on slow metabolic measures, such as functional Magnetic Resonance Imaging (fMRI) or Positron Emission Topography (PET).¹⁷

¹⁷ Mathematical methods are available to detect correlations between localized metabolic activations as measured by fMRI and PET. The advantage of these methods is that they provide maps of functional connectivity with a high

Millisecond temporal resolution is accessible through (i) direct intracellular and extracellular measurements of individual neurons; (ii) recordings of local field potentials (LFP), or of the electromagnetic fields of large neural populations that produce the electroencephalographic (EEG) and magnetoencephalographic (MEG) signals. LFP are the summation of the membrane potentials of populations of neurons. The size of the populations depends on the site and precision of the recordings: Local microelectrodes can record small populations, extending over less than a square millimeter of tissue, whereas scalp EEG electrodes or MEG sensors (and optical imaging) record the average activity of several square centimeters of cortex. At an intermediate level, intracranial recordings from human patients can record from a couple of square millimeters of cortex (Lachaux, Rudrauf, & Kahane, 2003). Except in those rare situations justified by therapeutical reasons, human recordings are almost exclusively non-invasive, and performed therefore at the centimeter-wide spatial resolution of MEG or EEG.

With this panel of recording techniques, spatiotemporal patterns can in principle be observed at three levels: (i) as interactions between simultaneous recordings of multiple individual neurons; (ii) as interactions between simultaneous recordings of multiple individual LFPs, and (iii) in single LFP recordings. The third level is intermediate between the first two: Because an individual LFP records from a single neural population, the average activity of the LFP is sensitive to the spatiotemporal organization of activity within this population. For example, if all the neurons are co-active periodically, the average activity in the LFP will be a massive oscillation, much stronger than if the neurons are not synchronous.

4.3.2. Some simple mathematical considerations

As we have seen, the organization of the brain suggests that interactions between distributed neuronal groups are bound to occur, given their massive interconnectedness. We also mentioned that recordings of neuronal activity can be obtained by a diversity of approaches and at several levels of spatial resolution. With these points in mind, let us return to the central question of this section: How does one detect neural spatiotemporal patterns from real brain data? The definition

spatial resolution in normal human subjects. These methods, however, measure interactions that occur on the time scale of a couple of seconds at best. This temporal resolution may be sufficient when studying the neural correlates of slow experiential patterns, such as the evolution of certain emotions (Buchel & Friston, 1997).

of the ‘dynamic core’ by Tononi and Edelman (1998) provides a useful starting point: “The dynamic core is a functional cluster: its participating neuronal groups are much more strongly interactive among themselves than with the rest of the brain” (Tononi & Edelman, 1998). The challenge for the neurodynamicist is therefore to find neurons or groups of neurons with particularly strong (but transient) interactions.

In keeping with the dynamical approach, we can usefully consider this question in geometrical terms. Consider the n simultaneous measures of brain activity that one can record in a typical electrophysiological setting (e.g., 64 measures from 64 EEG channels). At any time t , the n simultaneous measurements define a position in an n -dimensional state space, and the evolution of this position in time defines a trajectory. If the measurements are independent from each other, then the trajectory will progressively completely fill a hypercubic portion of the state space, leaving no hole. In contrast, if there are interactions between the measured neuronal populations, then the trajectory will fit into a restricted portion of the full space, and be constrained onto a manifold with a (fractal) dimension less than that of the state space.¹⁸ What this means in informational terms is that, for at least one pair of the measured neural populations, measuring the activity of the first population provides some information about the activity of the second one. The probability distribution of the activity of Population 2 (the probability $p(y)$ that this activity is y) *given that* the activity of Population 1 is x , is different from what it would be if the activity of Population 1 were x' . A metaphor: If we know where John will spend the afternoon, we can predict with some accuracy that his wife Ann will spend the afternoon in the same city, but we cannot predict where Jane, unrelated to John, will be. Certain measures, such as mutual information (David, Cosmelli, & Friston, 2004), quantify exactly this sharpening of the probability distribution.

The transient nature of neural interactions, however, makes general measures based on such geometrical formulations difficult to apply. The main problem is that, to know whether or not the trajectory fills up the whole space, the experimenter needs to observe it during time windows typically order of magnitudes longer than the typical life-span of cell assemblies. This

¹⁸ The fractal dimension d of a trajectory can be envisioned as follows: Imagine that each point along the trajectory is in fact a small ball of lead. Then, the total mass of lead contained in a sphere centered on the trajectory will increase as a function of the sphere radius r proportionally to rd . If the trajectory occupies all the space, then d is equal to the dimension of the space. If the trajectory is a straight line, or a circle, d equals 1.

difficulty can be avoided if the researcher assumes a priori what the shape of the manifold will be, onto which the trajectory is constrained. Because spatiotemporal patterns can potentially take an infinite number of shapes in the state space, a possible solution is to assume a specific shape and to build a special detector for this shape (surfing on the advances of signal processing). It is easy to see that it takes fewer measurements to test whether the trajectory stays on a circle, or whether it follows some general, unknown, geometrical structure.

In the absence of noise, three successive measurements are sufficient to know whether the trajectory stays on a straight line, which would correspond to a linear relationship between the recorded activities. Synchrony between quasi-periodic oscillators, that is, the transient phase-locking of their oscillations, is a good example of such an interaction. But linear relationships can be extended to other trajectories constrained on simple manifolds of dimension 1¹⁹, such as a circle. This is the case for two oscillators rotating at the same frequency with a constant phase lag.

4.4. Synchrony: Perhaps not the best candidate, but at least the simplest

The practical reason that synchrony has so far been the best-studied (if not the only) type of transient interaction between neural populations is the ease with which it can be detected. Furthermore, since the development of the EEG, it has been evident that oscillations are ubiquitous in the brain. This fact, combined with the relation between coordinated oscillatory activity and several important cognitive functions (discussed below), has also contributed to the development of approaches that seek to detect the occurrence of synchrony from real neurobiological signals.

In its original neurophysiological formulation, ‘synchrony’ refers to a positive correlation between the spike timing of a set of neurons. In other words, if we consider two neurons within a synchronous population, the probability of the first neuron to fire a spike is significantly higher at specific delays relative to the spikes of the other neuron.²⁰ In the simplest case, this delay is zero,

¹⁹ See the previous note.

²⁰ Note that synchrony can occur between two neurons without an actual direct relation between them, if they are driven by a common driver. This fact reveals one of the limitations of the synchrony measure so far. The three neuron system that includes the driver, however, can be seen as a larger spatiotemporal pattern revealed by the synchrony measure.

which means that the neurons have a high probability of firing simultaneously. In general, this probability, as well as the eventual delay, is quantified by the cross-correlogram between the spike trains of the two neurons (Perkel, Gerstein, & Moore, 1967). One speaks of oscillatory synchrony, or synchronous oscillations, if the neurons tend to fire at periodic latencies. Numerous animal studies, over the past twenty years, have now established that synchrony is ubiquitous in virtually all sensory and motor modalities. It has often been found to be related to perception, memory, and motor programming (see Roskies, 1999, for a group of excellent reviews that summarize these results).

Synchrony has also been studied in humans as an instance of spatiotemporal patterns of interdependent neural activity, although at a different level from animal studies. Here an important distinction needs to be made between the local recordings of individual neurons, almost always accessible only in animals, and the more global recordings of entire neural populations, accessible in humans through scalp EEG or MEG. EEG and MEG average across large neuronal assemblies, and hence oscillatory synchrony between neurons shows up as changes of power in particular frequency bands. The reason this happens is that groups of synchronously firing neural oscillators can be modeled as oscillators themselves, with the amplitude of the oscillations depending on the number of individual oscillators in the group, and on the precision of the synchrony between them. This point entails a further distinction: On the one hand, oscillatory activity as recorded by an individual EEG electrode or MEG sensor implies already a certain amount of local synchronous activity. On the other hand, one can choose to consider synchronization between oscillations produced by distant neuronal populations (separated by several centimeters) in order to describe distributed spatiotemporal patterns that occur at a more global level. In any case, when dealing with EEG and MEG recording one is always in the presence of noisy data. Hence, any interdependence measure must be understood in a statistical sense, throughout a given temporal window (Lachaux et al., 2002; Lachaux, Rodriguez, Martinerie, & Varela, 1999; Le Van Quyen et al., 2001).

Synchrony at the more regional or local level has been demonstrated repeatedly in humans, in relation to integrative mechanisms in language, memory, attention, and motor tasks, in virtually all the sensory modalities. For example, the perception of coherent objects in humans is specifically associated with synchronous oscillations in the gamma range (above 30 Hz), the so-called 'induced gamma response' (for a review, see Tallon-Baudry & Bertrand, 1999). This

response, although not completely time-locked to the stimulus presentation, typically starts in posterior brain areas (over the occipital cortex) around 200 milliseconds after the stimulus, and then returns gradually to the pre-stimulus level, when the stimulus does not require further analysis (Lachaux et al., 2000, 2005; Tallon-Baudry & Bertrand, 1999).

As we mentioned above, oscillations produced by two neural populations can also be synchronous, within larger cell assemblies. This synchrony can be detected by a transient phase-locking between the oscillations of the two local fields (Lachaux et al., 1999, 2002; Rodriguez et al., 1999). Such long-range synchrony between distant neural populations has been suggested as a plausible candidate to mediate the integration of activity in functionally specialized and distinct brain regions (Bressler, 1995; Varela et al., 2001). For example, Tallon-Baudry and colleagues have shown in humans that during the maintenance of a complex shape in visual short-term memory, two functionally distinct regions within the ventral visual pathway, the fusiform gyrus and the lateral occipital sulcus, produce synchronous oscillations around 20 Hz (Tallon-Baudry, Bertrand, & Fischer, 2001).

Synchronization is a complex concept that can cover several possibly distinct types of temporal relations such as coherence, frequency-synchronization, phase-synchronization, generalized synchronization, as well as others (Brown & Kocarev, 2000; Friston, 1997; Pikovsky et al., 1997). Here we have focused on synchrony as either occurring between stochastic point-processes, such as spike trains, or in terms of phase-relations between oscillatory processes (phase-locking synchrony). As mentioned above, this focus is mainly due to technical limitations in the estimation of generalized measures of synchronous activation. With this point in mind it is possible to say that synchrony, as presented here, appears as a simple measure of precise temporal relations between neural processes that can enable one to follow the formation of spatiotemporal brain patterns relevant for consciousness. Not surprisingly, this mechanism is referred to in several of the dynamical models reviewed above. In the next section, we will review more specifically a set of results concerning the relation between consciousness and the current ‘crowd’s favorite’ among neurodynamicists—synchrony in the gamma range.

5. The Crowd's Favorite: The Gamma Band

5.1. Evidence for a relation between gamma synchrony and consciousness

We have mentioned that synchrony among oscillating neural populations is a plausible candidate to mediate functional connectivity, and therefore to allow the formation of spatiotemporal structures, such as those reviewed in the previous sections. In this section, we return to this hypothesis in more detail, with a particular focus on gamma band oscillations, which have been repeatedly associated with consciousness in the last fifteen years.

The putative role of gamma band oscillations in the formation of conscious experience was proposed by Crick and Koch (Crick & Koch, 1990), shortly after Singer and colleagues (Gray et al., 1989) had completed a series of observations in the cat visual cortex showing that neurons tend to synchronize their spiking activity when stimulated with parts of the same visual object, such as a moving bar (whereas they do not synchronize when stimulated with features that cannot be part of the same object). Those observations matched theoretical predictions by Von der Malsburg (Von der Malsburg, 1981) that synchrony could be used to achieve figure/ground segmentation during perception of the visual scene. Thus, synchrony was assumed to provide a solution to the visual binding problem (the problem, discussed above, of integrating distinct visual features into a unified and coherent perception) (see Roskies, 1999).

In their 1990 paper, Crick and Koch pushed this idea further by stating that visual consciousness of the object occurs only when its features are bound together as a result of this type of synchronous activity. This hypothesis was in good agreement with the feature-integration-theory, proposed by Anne Treisman, suggesting that attention is necessary to bind together the features of objects (Treisman & Gelade, 1980).²¹ Henceforth, a close relation between gamma synchrony and attention and consciousness has ensued (Fell et al., 2003; Varela, 1995). This association has been very appealing for neurodynamicists addressing the consciousness issue, because it provides ground material for the neural spatiotemporal patterns they associate with consciousness on the appropriate time-scale. Indeed, several dynamic models among those reviewed above, specifically those by Singer, Llinas, Varela, and their respective collaborators,

²¹ This theory was based, among other things, on the observation of false conjunctions in the absence of attention: When presented briefly with a red square and a blue circle outside of the scope of attention, a subject would sometimes report having seen a red circle and a blue square. Such perception is typically an incorrect binding of the color and shape attributes.

consider synchronous activity in high frequency ranges, most preferentially the gamma range, as crucial for conscious experience.

Fortunately, the association between the gamma band and attention, vigilance, and consciousness is not just based on its theoretical appeal, but also on sound experimental evidence. For instance, it is well known that the precise synchronization of neuronal discharges is more prevalent during states characterized by arousal, and moreover that gamma oscillations are particularly prominent during epochs of higher vigilance (Herculano-Houzel, Munk, Neuenschwander, & Singer, 1999; Rodriguez, Kallenbach, Singer, & Munk, 2004). In cats, for example, gamma synchrony is stronger after the stimulation of the mesencephalic reticular formation (Munk, Roelfsema, Konig, Engel, & Singer, 1996). Furthermore, EEG/MEG gamma-band activity is present both during REM sleep and awake states, with a much stronger amplitude than during deep sleep (reviewed in Engel et al., 1999).²²

Several studies have also demonstrated that the presentation of sensory stimuli elicits stronger gamma synchrony when attention is focused on the stimulus, than when attention is diverted away. This finding was observed in monkeys for somatosensory stimulations, and found again recently for neurons in area V4 of monkeys presented with small visual gratings (Fries et al., 2001).

There is also evidence for a more direct relation between gamma activity and consciousness. Lachaux and colleagues have recently shown that the perception of faces is associated with strong gamma oscillations in face-specific regions along the ventral visual stream (Lachaux et al., 2005). Epileptic patients with intracranial electrodes that record directly from the fusiform face area (a region along the ventral visual pathway particularly associated with the perception of faces) were presented with high-contrast ‘Mooney figures’ representing faces. Because the figures were presented briefly, for 200 milliseconds, they were consciously perceived as faces only half of the time. The authors reported that the gamma band response to the images was significantly stronger when the figures were actually consciously perceived as faces, than when they were not. This high-resolution study followed a previous one (Rodriguez et al., 1999), using the same protocol in normal subjects with non-invasive scalp EEG recordings;

²² This has led to the suggestion that, gamma-band synchrony is the trace of similar processes in the emergence of dreaming consciousness in REM sleep and waking consciousness (Engel et al., 1999).

this study showed that gamma oscillations tend to synchronize across widely separated brain areas (typically frontal versus occipital) only when the figures are perceived as faces.

Fries and colleagues (Fries, Roelfsema, Engel, Konig, & Singer, 1997) have shown an even more direct relation between gamma synchrony and consciousness. They showed that during binocular rivalry in cats, the level of synchrony between visual neurons follows in time the shift of perceptual dominance. Cats were presented with two visual patterns moving simultaneously in different directions: One pattern was presented to the left eye and the other to the right eye. Under such circumstances, the visual percept cannot encompass the two contradictory patterns and instead alternates between them (hence the term ‘binocular rivalry’). The results of this study showed that neurons stimulated by the perceived stimulus were strongly synchronized, with strong gamma oscillations, whereas cells stimulated by the suppressed visual pattern showed only weak synchrony. This experiment is highly relevant to the study of visual consciousness, because conscious perception is decoupled from the drive of the sensory inputs (the physical stimulus remains constant while perception does not), and gamma synchrony is used as an indicator of which pattern is being consciously perceived by the cat.

Gamma synchrony has been further associated with consciousness in the context of the attentional blink effect. The attentional blink occurs when a subject must detect two targets in a series of rapidly presented pictures (at a rate of about 10 per second). Typically, the second target is detected (and consciously perceived) less frequently when it comes within 500 milliseconds of the first target, as if the subject had ‘blinked’. Fell and colleagues have argued that the blink could be due to the suppression of gamma synchronization shortly after the response to the first target (Fell, Klaver, Elger, & Fernandez, 2002). Once again, gamma synchrony would be necessary for conscious perception.

This proposal is consistent with a recent observation from Lachaux and colleagues, in the face perception paradigm detailed above (Lachaux et al., 2005), that parts of the primary visual cortex shut-down, with respect to gamma activity, after the presentation of a Mooney figure: There is a drop of energy in the gamma band, below the baseline level, which lasts a couple of hundreds of milliseconds, and is simultaneous with the induced gamma increase in the fusiform face area. This drop in gamma could be the trace of a transient deactivation of the primary visual cortex that could cause the transient attentional blink after a meaningful visual stimulus. The

visual cortex would be transiently ‘unavailable’ while processing particularly meaningful stimuli, as in a reflex protective mode.

Further hints about the role of gamma synchrony come, albeit indirectly, from the experimental contributions of Benjamin Libet (Gomes, 1998; Libet, 2002). In a series of classic experiments in patients, mixing direct intracranial electric stimulations and peripheral somatosensory stimulations, Libet revealed a number of interesting properties of somatosensory awareness. These include: (1) An electrical cortical or thalamic stimulus requires a duration of more than 250 milliseconds to be felt, whereas a skin stimulus of 20 milliseconds is sufficient. (2) If a direct cortical (electrical) stimulus occurs within 250 milliseconds after a skin stimulus, it can suppress or enhance the felt perception of the latter stimulus. (3) For a skin stimulus to be felt as synchronous with a non-overlapping cortical stimulus, the skin stimulus must be delayed about 250 milliseconds relative to the latter stimulus. Interestingly, all three properties match quite closely the known temporal dynamics of the cortical gamma response induced by sensory stimuli. This match is particularly intriguing considering the fact that Libet used rhythmic electrical stimulations in the gamma range (typically 60 Hz trains of electric pulses). If the induced gamma response is involved in the conscious perception of a sensory stimulus, then one would indeed expect that (i) a rhythmic train of electrical stimulations in the gamma range could mimic the effect of the induced gamma response, if it possesses the same temporal properties; that is, if it starts roughly 250 milliseconds after the mimicked stimulus onset and lasts for at least 250 milliseconds, then (ii) it should be felt as synchronous with a corresponding skin stimulus, and (iii) possibly interfere with perception of that latter stimulus. In brief, Libet’s observations can readily be interpreted via the involvement of the sensory-induced gamma response in sensory awareness, at least in the case of somatosensory awareness.

In summary, the previous studies certainly build a strong case for the role of resonant assemblies, oscillating in the gamma range, as neural correlates of sensory awareness. Nevertheless, this assessment is not the end of the story, for a number of arguments make it difficult to equate gamma synchrony and consciousness.

5.2. Problems concerning the link between gamma synchrony and consciousness

The first problem to mention is that gamma synchrony can be observed in unconscious, anesthetized animals, although it is stronger when animals are awake (see Searwards & Searwards,

2001, for arguments against the role of gamma synchrony in consciousness). Searwards and Searwards further argue that gamma oscillatory activities have been detected in structures that most likely do not participate in the generation of sensory awareness, such as the hippocampal formation: “Obviously hippocampal activities could not contribute to sensory awareness since lesions to that structure do not result in purely sensory deficits of any kind” (Searwards & Searwards, 2001, p. 492). This argument, as well as others, lead them to conclude that “while synchronization and oscillatory patterning may be necessary conditions for activities to participate in generating awareness, they are certainly not sufficient” (Searwards & Searwards, 2001, p. 492).²³

This point echoes the conclusions from a study by Revonsuo and colleagues (Revonsuo, Wilenius-Emet, Kuusela, & Lehto, 1997). In this study, they recorded the gamma-band response of normal subjects during the fusion of random-dot stereograms, and observed that while 40-Hz synchronized oscillations seemed to participate in the construction of the unified percept, they were not maintained during the continuous viewing (and conscious perception) of the same stimulus once it had been constructed. Lachaux (unpublished findings) repeatedly confirmed this observation with human intracranial recordings: The gamma response induced by durable visual stimuli in the visual system often stops before the end of the stimulus presentation, despite the fact that the subjects still fixate the images and consciously perceive them.²⁴

These considerations indicate that other spatiotemporal structures may participate in the emergence and the stabilization of the conscious percept. The presence of such structures is especially the case for short-term memory, which has been proposed as a central component of consciousness (Baars & Franklin, 2003). In visual short-term memory, when an individual has to maintain a conscious representation of a complex visual shape, using mental imagery, for a couple of seconds, synchrony happens not in the gamma range, but in the lower beta range (between 15 and 20 Hz), between distributed sites of the ventral visual pathway (Tallon-Baudry et al., 2001). Therefore, resonant cell assemblies in the beta range may also subserve continuous visual perception (if only in its imagery aspect).

²³ We do not wish at this point to step into the debate about which brain areas actually participate in the generation of sensory awareness (see Rees, Kreiman, & Koch, 2002).

²⁴ Letter strings presented to the subject for 1 second, for instance, generate an induced gamma response that lasts roughly only for the first 500 milliseconds (Lachaux, unpublished observations).

The above studies serve to stress the point that gamma synchrony may be necessary in the emergence of a conscious perception, but perhaps only in this emergence. Once formed, the percept could then continue via other cell mechanisms, in the form of other types of spatiotemporal structures.

Nevertheless, even at the initial level of this emergence, the role of gamma synchrony needs to be clarified. As we have seen, gamma synchrony occurs in anesthetized animals, and is therefore not sufficient for consciousness (Sewards & Sewards, 2001). One interesting possibility, in the case of the visual system, is that gamma synchrony could be involved in the formation of visual objects. Visual objects are the preferred targets of visual attention, and yet they present themselves to us only via conscious perception. Furthermore, as argued by the Feature Integration Theory (Treisman & Gelade, 1980), visual objects seem to require visual attention to form. The question thus arises of which comes first, objects or attention? One solution to this problem is that in the absence of attention there are only ‘pre-objects’, that is, bundles of features that are object candidates and that are sufficient to attract attention, which would then finish the construction and remain grabbed by them (Wolfe & Bennett, 1997). Engel and Singer (Engel et al., 1999) propose that gamma synchrony may mediate this mechanism. According to this proposal, proto-objects, based on their physical features and Gestalt properties, assemble in the form of nascent cell assemblies via gamma synchrony. This synchrony corresponds to the kind of ‘automatic’ synchrony observed in anesthetized animals. This nascent synchrony is reinforced in awake animals, such that there is a formation of the visual object. This process corresponds to the grabbing of attention by the object, and is simultaneous with the object’s actual formation for perception. In this model, attention and gamma synchrony become two sides of the same coin, as long as one is ready to extend the concept of attention (usually associated with conscious perception) to a general selection mechanism that includes an unconscious pre-selection mechanism. This pre-selection mechanism is the one observed in anesthetized animals. Attention, in its classic ‘conscious’ sense, is thus envisioned as the tip of the selection iceberg.

Can we therefore relate the full formation of resonant gamma assemblies to the emergence of consciousness? The answer would seem to be yes, in a certain sense, namely, that the content of the resonant gamma assembly is accessible to verbal report, working memory, and so on. On this view, gamma synchrony is necessary for any kind of sensory awareness. This view

gains support from Engel and Singer's observation that synchrony is related to all of the four presumed component processes of awareness, namely, arousal, segmentation, selection, and working memory (Engel et al., 1999). In the next section, however, we will examine certain problems with this idea that will lead us to qualify it.

6. Consciousness and Dynamical Structures: Some Qualifications

6.1. Introduction

Throughout this chapter we have explored the view that consciousness seems to require the formation of distinct, dynamic spatiotemporal structures in the brain. This view is, after all, one of the main points of agreement among the different neurodynamical proposals we reviewed in Section 3. In this section, we will take a more critical stance regarding this central issue and put forth some qualifications we believe are important to keep in mind.

In several of the neurodynamical theories we have discussed, the notion of a distributed neuronal assembly, understood as some kind of synchronous pattern of activation, is central to explaining the neuronal basis of consciousness. As we saw in Section 5, the gamma band has been a preferred region of the frequency domain, in which such assemblies have been studied. Whether restricted to this frequency band or spanning multiple frequencies, an emergent and stabilized spatiotemporal pattern is seen as a prerequisite for conscious experience to happen.

This viewpoint, however, raises at least two related questions. On the one hand, if such patterns are necessary for consciousness, and if we can distinguish them as having a certain spatiotemporal unity, what happens between patterns? Are we conscious during such transitions? Or is consciousness a sequence of snapshots, in which the apparently seamless fusion of successive moments into the ongoing flow of experience is achieved by some additional mechanism?

On the other hand, can we define a stable conscious moment within the flow, and are we therefore entitled to suppose that during such a moment, the assembly will 'hold' or 'contain' a certain unity, even though during that moment one can distinguish a change (or changes) in one's experience? Recall that dynamic assemblies are supposed to last for several hundreds of milliseconds, but our sensory experience can change within that duration. Suppose, for example, you are sitting in a train, staring out of the window, and as you look out into the countryside, trees, electricity poles, and other objects swiftly cross your visual field, without your being able

fully and stably to grasp them. Yet you know they are trees, electricity poles, and other objects. Does your rapid experience of each of these objects correspond to a distinct assembly? Or is it rather a matter of one global assembly, in which various local assemblies ‘ride’? In several neurodynamical proposals, as we have seen, an experience of an object is supposed to depend on the formation of distinct, coherent brain patterns. But a conscious moment can include full-fledged objects as well as less definite visual patterns that, although conscious to a certain extent, cannot be completely described as stable entities.

These two interrelated features—ongoing flow and fleeting experiences—need to be addressed by any neurodynamical approach to consciousness. In the remainder of this section, we discuss both features and propose a simple distinction that may help to clarify the issues at hand.

6.2. Ongoing flow and fleeting experiences

The issue of the ongoing, fluid nature of conscious experience is certainly not new.²⁵ William James, in his famous chapter on “The Stream of Thought” (James, 1890, Chapter IX), provides a detailed description of the structure of this flow. He distinguishes at least two fundamental aspects—‘substantive’, stable moments, in which one is actually conscious of something, and ‘transitive’, fleeting moments, in which one passes from one content to another. He describes consciousness as like a bird’s life, for it seems to be made up of an alternation of flights and perchings. James remarks that substantive moments can be recognized as such, whereas transitive moments are quite difficult to pinpoint accurately. They present themselves as tendencies and changes between states, and not as distinct contents immediately definable in themselves, save by some retrospective exercise.

How do these phenomenological observations relate to the neurodynamical picture of the brain and its relation to consciousness? As we have seen, most neurodynamical proposals stress that each conscious state depends on a specific neural assembly or emerging dynamic pattern, but the issue of how transitions between states take place and what they mean in terms of the experiencing subject is less frequently addressed. With regard to this issue, the proposals of

²⁵ For an extensive presentation of questions concerning the experience of time, we refer the reader to the notable work by Charles Sherover (Sherover, 1991).

Varela and Kelso are the most explicit and developed.²⁶ These authors stress the *metastable* nature of such patterns, so that successive moments of distributed neural coherence combine in a continuous and ongoing fashion, in contrast to a sequence of clear-cut states.²⁷ These approaches present attractive alternatives that seem to fit nicely with James's intuitions. They also allow for a different interpretation of what counts as a meaningful dynamic pattern. Rather than seeing these patterns as individual assemblies that arise, maintain themselves for a brief period, and then subside, they can be viewed as one itinerant trajectory, and thus as one pattern (Friston, 1997, 2000; Varela, 1999), in which the rate of change is the only internal definition of the stability of a given moment. In any case, neurodynamical approaches must deal explicitly with this issue of the apparent unity of the flow of consciousness,²⁸ as opposed to the unity of moment-to-moment experience.

The second question to which we wish to draw attention is related to the stability of actual perceived objects during a conscious moment. As we mentioned above, the notion of an assembly implicitly incorporates a notion of stability during the life-span of the pattern in question. Our sensory environment, however, can be subject to rapid change in time windows less than several hundred milliseconds, and yet we are, to a certain extent, aware of the change as taking place. This fact would seem to pose a difficulty for any theory that postulates a neural assembly, organized on a slower time-scale, as necessary for conscious experience. On the other hand, not every object of the visual scene is perceived as stably as one might naively think. This fact is especially clear in inattention blindness experiments (Simons, 2000). In such experiments, subjects are asked to focus on a particular task and set of stimuli in a visual scene. If an additional stimulus appears unexpectedly in that scene, the subjects are often unable to report it afterwards. What is particularly striking with such 'inattention blindness' is that it can happen even for very distinctive and salient objects. In one famous example (described in Simons, 2000), subjects watch people passing basketballs. Three people wearing white T-shirts pass a ball to each other, while three other people wearing black T-shirts pass another ball to each other. The subjects have

²⁶ Varela in particular proposed a neurodynamical account of Husserl's phenomenological account of time consciousness (see Varela, 1999, and for further extensive discussion, Thompson, 2007).

²⁷ Tononi and Edelman (1998) do mention that their dynamic core is constantly changing, but they do not further develop this point.

²⁸ The question of whether this unity is illusory or real remains an unresolved problem (VanRullen & Koch, 2003).

to count the number of passes between the white players, which occur at a fast enough rate to require the full attention of the viewer. After 45 seconds of the display, a man in a gorilla suit walks across the scene, stops for a moment in between the players, waves his hands in the air and then exits through the other side five seconds later. It is well documented that a high portion of the viewers fail to report seeing this gorilla.

In models like the one advocated by Singer and collaborators (see Section 3.3.1) there is a strong correspondence between a figure-ground distinction (and therefore an object) and the formation of a synchronous assembly. This correspondence would seem to imply that only fully formed assemblies can ‘support’ some type of perceptual recognition of the object in question. As discussed above, however, both phenomenological observation of one’s own experience and experiments such as the unnoticed gorilla suggest that a great deal of experience may be unstable and fleeting. Where would such fleeting experiences of quasi-objects fall in the framework of dynamic neural assemblies? Lamme (2003, 2004) has proposed that such fleeting experiences belong to ‘phenomenal consciousness’ (i.e., are subjectively experienced, but not necessarily accessible to verbal report), whereas more stable experiences belong also to ‘access consciousness’ (i.e., are available to verbal report and rational action guidance) (see Block, 1997, 2000, for this distinction between phenomenal consciousness and access consciousness).²⁹ Neurodynamical models need to be able to account for this evanescent aspect of conscious experience in a more explicit way.

More precisely, we propose that the stable/fleeting duality be considered a structural feature of consciousness experience (see also Section 7) and dealt with accordingly. In a certain sense, this duality mirrors the access/phenomenal distinction, but without assuming that there can be fleeting phenomenally conscious experiences that are inaccessible in principle to verbal report. In endorsing the need to make this stable/fleeting distinction, we also stress the need to consider the possibility of the more ephemeral aspects of experience as being accessible to verbal report, if

²⁹ Lamme’s distinction, however, is not completely equivalent to Block’s initial proposal (Block, 1995). In its original formulation, phenomenal consciousness is subjective experience, in the sense that there is something it is like for the subject to be in the state. Access consciousness, on the other hand, is an information-theoretical concept that is supposed to account for the availability of conscious information for further rational guidance of behavior, including reportability. The conceptual and empirical validity of this distinction is a matter of lively debate in the science of consciousness (see Block, 1997, 2000; see also the discussion in Thompson, Lutz, and Cosmelli, 2004).

approached with the appropriate first-person and second-person phenomenological methods (Varela and Shear, 1999; Petitmengin, in press).

Given this structural distinction between stable and fleeting aspects of experience, it would be interesting to see how a neurodynamical theory that relates the formation of well defined spatiotemporal patterns in brain activity to conscious experience would deal with the intrinsic *mobility* to any given perceptual act. For example, the feed-forward stream (or sweep, FFS) is defined as the earliest activation of cells in successive areas of the cortical hierarchy. In the visual modality, it starts with the retina, the LGN, V1, and then the extrastriate visual areas and the parietal and temporal cortex. Thorpe and colleagues (Thorpe, Fize, & Marlot, 1996) have shown that the FFS is sufficient to carry out complex visual processing, such as detecting whether a natural scene presented for 20 milliseconds contains an animal. It is tempting to relate the more stable aspect of experience to the formation of spatiotemporal patterns, in the sense of dynamic neural assemblies mediated by recurrent neural interactions, whereas the fleeting, unstable awareness could be embodied through the rapid FFS that *modulates* and continuously affects the formation of such assemblies, while not being fully excluded from a certain level of perceptual experience. This proposal is highly speculative, but is intended simply as a way to highlight the necessity of dealing with the stable/fleeting structure that appears to be inherent in each and every conscious moment.

To conclude this section on qualifications to the dynamic approach, we would like briefly to draw the reader's attention to another aspect of consciousness that is significant in light of the preceding discussion and the overall topic of this chapter. This aspect is the subjectivity or subjective character of consciousness. For example, Damasio (1999) has stressed that, in addition to understanding the neurobiological basis for the stream of object-directed conscious experiences, it is also necessary to understand the neurobiological basis for "the sense of self in the act of knowing" (Parvizi & Damasio, 2001; see also Panksepp, 1999, for a convergent argument, and Wicker et al., 2003). The sense of self with which Damasio is concerned is a primitive kind of conscious self-awareness that does not depend on reflection, introspection, or possession of the concept of a self. In phenomenological terms, it corresponds to the fundamental 'ipseity' (I-ness or selfhood, by contrast with otherness or alterity) belonging to subjective experience (see Lutz, Dunne, & Davidson, this volume; Thompson & Zahavi, this volume). In a related line of argument, Searle (2000) has suggested that a major drawback of current attempts

to uncover the neural correlates of consciousness in human beings is that they begin with already conscious subjects. He advocates a ‘field of consciousness’³⁰ viewpoint, in which the perceptual experience of an object arises as a modification of a pre-existing conscious ‘ground-state’ that is unified, subjective, and qualitative. In this context, the transition between conscious states need not be punctuated by a radical gap in consciousness, but can rather be a modulation of a more basic state of background consciousness, which accounts for the fact that even such transitive moments are felt as belonging to oneself. Here dynamic patterns in the form of transient and distributed co-active assemblies would mainly reflect the nervous system’s own homeodynamic activity, that is, its maintenance of a range of internal regularities in the face of its ongoing compensation for the systematic perturbations to which it is exposed from both the sensory environment and the internal bodily milieu (Damasio, 1999; Maturana and Varela, 1980).

Nevertheless, it remains difficult to see how metastable assemblies of co-active neurons could by themselves account for this crucial aspect of the subjectivity of consciousness. This crucial feature is often put to the side as something to deal with once the issue of the neural correlates of perceptual consciousness has been resolved (e.g., Crick and Koch, 2003). Our view, however, is that unless the subjectivity of consciousness is adequately confronted, and its biological basis understood, proposals about the neural correlates of perceptual consciousness will provide limited insight into consciousness overall. Thus, the issue of subjectivity is a nontrivial matter that any neurodynamical approach must confront sooner or later, if it is to become a cogent theory of consciousness. We will briefly pick up this thread in the next section when discussing how to relate phenomenological descriptions to neurodynamical accounts.

7. The Future: Beyond Correlation?

7.1. Introduction

So far we have dealt primarily with the issue of meaningful spatiotemporal patterns in the brain and their relevance to the study of conscious experience. It may have become increasingly evident to the reader, however, that the issue of *how* to relate such patterns to *experience* as a first-person phenomenon has been left untouched. Indeed, one of the major challenges facing the cognitive sciences is precisely how to relate these two domains—the domain of third-person,

³⁰ The notion that consciousness has a unified field structure goes back to A. Gurwitsch (Gurwitsch, 1964).

biobehavioral processes and the domain of first-person, subjective experience. What is the right way to conceptualize this relation, and what is the best way to approach it methodologically? These questions have not yet received anything near a satisfactory answer from the neuroscientific community. We do not intend to propose an answer to them here. Rather, we wish to highlight some conceptual and practical issues in the quest to understand the relation between these two domains, while keeping in mind the dynamical insights we have gained from the previous exposition.

7.2. Correlation and emergence

The first question that comes to mind is the extent to which the entire neurodynamical approach rests on a merely correlational strategy. In coarse terms, one isolates a given target experience, say the perception of a figure; one determines the neural patterns that correlate with the moment the subject sees the figure; and one then concludes that the conscious experience depends on such neural patterns.³¹ In the last decade or so this correlational approach, in the form of the search for the neural correlates of consciousness, has undergone important developments and become more sophisticated with regard to its conceptual formulation, methodological commitments, and empirical results (Block, 1996; Crick & Koch, 1990, 1998; Rees, Kreiman, & Koch, 2002). Here the central idea is that rather than formulating explanatory principles about the relation between neural activity and experience, what has to be done first is to determine those neural processes that can count as a “specific system in the brain whose activity correlates directly with states of consciousness” (according to the Association for the Scientific Study of Consciousness, cited by Chalmers, 2000, pp. 17-18). Once such processes have been found, then one can turn to the issue of how they are causally related to experience itself.³²

³¹ We will see below that this general characterization needs some important qualifications, in particular at the level of determining what counts as a valid conscious experience, and how to contrast such a conscious experience with possibly unconscious processing in similar situations.

³² The theoretical validity and empirical plausibility of this approach remains a matter of extensive discussion. Rather than endorse or reject it, we wish to highlight it as an influential approach that can serve as a reference for further discussion. The interested reader is referred to several interesting publications (and references therein) on this controversial and interesting question (Crick & Koch, 2003; Metzinger, 2000; Noë & Thompson, 2004a, 2004b; Pessoa, Thompson, & Noë, 1998).

Neurodynamics as a research program is devoted, at least methodologically, to this correlational strategy, and in this sense remains closely linked to the NCC program. Of course, this commitment is due to the fact that, in the scientific tradition, establishing a relation between two target events or phenomena is mainly approached by establishing a correlation in their occurrence. Causal relations can then be assessed on the basis of altering one of the target events and observing whether and how the other changes. This ‘interventionist’ strategy can be employed in the case of brain functioning and consciousness by using microstimulation during surgery or transcranial magnetic stimulation (TMS). Nevertheless, by itself this strategy does not guarantee the elucidation of the underlying causal mechanisms.

Several of the proposals reviewed above, however, formulate explicit links between the neural and the experiential in terms of the notion of *emergence* or *emergent phenomena*, and thus can be considered as attempts to go beyond a purely correlational description. Although ‘emergence’ is a complex concept subject to multiple interpretations (see Keslo, 1995; Thompson, 2007; Thompson & Varela, 2001), in simple terms it can be defined as follows: A process is emergent when (i) it belongs to an ensemble or network of elements; (ii) it does not belong to any single element; and (iii) it happens spontaneously given both the way the elements interact locally and the way those interactions are globally constrained and regulated. Thus, an emergent process cannot be understood at the level of local components taken individually, but depends rather on the relations established between them. Furthermore, an emergent process not only depends on the local components, but also constrains their degrees of freedom, a two-way process that has been termed ‘circular causality’ (Haken, 1983). Especially in Freeman’s and Varela’s approaches, conscious experience is considered to be an emergent process. The difference between their views is that whereas Freeman (1999a, 1999b) proposes that consciousness is a global brain state, Varela proposes that consciousness may encompass multiple cycles of organismic regulation that are not fully restricted to the brain (Thompson & Varela, 2001). Nevertheless, although principles of emergence have been clearly formulated at the level of physical processes and molecular interactions (Nicolis & Prigogine, 1989), in the case of conscious experience, such principles still need to be understood and formulated in a more formal way. We believe that the study of complex systems offers a promising approach in this direction (Le Van Quyen, 2003; Thompson, 2007).

7.3. Gaining access to experience

As mentioned earlier, any neurodynamical approach to consciousness must eventually deal with the issue of how to describe experience itself. In the previous sections of this chapter we have discussed mainly spatiotemporal brain patterns in relation to consciousness, but we now turn to consider the other side of the issue, namely, how to gain access to experience itself and render it accessible to scientific description.

On the more operational side of this question, one can ask how it is possible to set up an experimental paradigm that addresses the issue of gaining access to experience in a way that allows us to study the underlying neuronal processes. Not all that is going on in the brain is necessarily related to what the subject is consciously experiencing. It is known that, during our conscious engagement with the world, a non-negligible part of our adapted behavior depends on nonconscious processes that are carried on without us being aware of their functioning. For example, do you have any feeling whatsoever of the oxygen level in your blood right now? Yet this bodily state of affairs can be crucial to your capacity to be here right now reading this text. Although this example is extreme, carefully crafted experiments reveal that even perceptual information can be used to guide behavior in a nonconscious way. For example, when a subject is presented with a small circle surrounded by larger circles, the small circle appears smaller than if it is presented in isolation. Yet if the subject is asked to reach for it, his fingers adopt a grip size that is consistent with the true size of the circle and not with its illusory dimension (Milner & Goodale, 1995). Another classic example is known as blindsight (Danckert & Goodale, 2000; Weiskrantz, 1990). In this neurological condition, conscious visual experience is impaired due to damage in primary visual cortex, yet subjects can produce quite accurate motor actions, such as introducing an envelope through a horizontal slot or pointing to a target they claim not to see. Thus, the problem arises of how to determine those neural processes that show some kind of direct relation to the actual conscious experience of the subject, in contrast to those that sustain ongoing and nonconscious adaptive behavior in the world.

Several experimental approaches have become paradigmatic in this endeavor. In general, the rationale behind these experiments is to dissociate what is presented to the subject from what the subject sees, in order to distinguish the neural patterns that are specific for conscious perception. Among these approaches, three stand out as the most well studied and influential. The first is visual masking (Dehaene & Naccache, 2001; Hollender, 1986), in which short-lived visual

stimuli flanked by meaningless masks are not perceived consciously, yet can alter future behavior (an index of nonconscious processing). The second is inattentional blindness and change blindness (O'Regan & Noë, 2001; Simons, 2000), in which diverting the subject's attention can render major changes in the scene unnoticed. The third is binocular rivalry (Blake, 1989; Blake & Logothetis, 2002), in which the presentation to each eye of a different image induces an alternation in conscious perception among the two alternatives, despite the fact that both are always present.

This last experimental paradigm is particularly relevant to the issue of gaining access to experience, because it provides an ongoing, slow phenomenon that can be described by the subject. In virtue of its alternating character, the experience lends itself to repetitive scrutiny, in order to better characterize 'what it is like' subjectively to undergo it. Finally, because both stimuli do not change, yet perception changes dramatically, binocular rivalry evidences the endogenous and ongoing character of experience, and therefore calls for attending to those neural processes that share this fundamentally dynamical structure.³³

These considerations suggest that, besides experimental paradigms for dissociating unconscious and conscious processes, we need to be able to capture the dynamics of experience itself. Hence it is necessary for the experimenter to take measurements of each phenomenon—the dynamics of the brain and the dynamics of experience. Measurements should provide public data, that is, information that can be shared with another observer. One recurrent problem with consciousness is that the direct observation of experience is accessible only to the subject, and such observation is not a public measurement in itself. The experience therefore needs to be transcribed into public data in a subsequent step, in order to provide so-called 'first-person data'. What the status of first-person data is, and to what extent the subject can play an active role in describing his or her experience, are matters of active debate in the science of consciousness (Jack & Roepstorff, 2003; Varela & Shear, 1999). We cannot review these debates here. Rather, in the remainder of this section, we wish to explore two complementary lines of investigation that are relevant to the issue of making experience more scientifically accessible.

³³ This feature pertains to multistable and ambiguous perception in general (see Leopold & Logothetis, 1999).

7.4. A topological approach to first-person data

In a general sense, one expects measures to be somehow organized in a universe, the measurement universe, that is the set of all the possible ‘values’ that measure can take. The term ‘universe’ must be understood in its statistical sense, and simply refers to the set of possible values, states, or items that can be valid measurements. For instance, a single word is one particular item among all the possible words. The universe may be discrete (as for words or sentences) or continuous (as for magnetic fields). In any case, a measurement will be the selection of one particular value allowed in a given universe, based on the present state of the observed phenomenon. For a given subjective experience, this may correspond to the selection of one description, among all the possible written descriptions that can be produced (say) in a couple of minutes.³⁴

We mentioned that the measurement should be ‘organized’. This means that it should be provided with some sort of topology: It should be possible to estimate a distance between two measures. Indeed, it should be possible to say whether measure A is closer to measure B than it is to measure C. Without any kind of topology, it would be difficult to compare the dynamics of the two phenomena. For instance, the notion of *stability* requires a distance: Stability means that the phenomenon remains somewhat constant during a certain time interval; this further means that the *distance* between consecutive measures is shorter now than what it was in earlier observation windows. The notion of distance is also central to the concept of *recurrence*: If we find a certain neural pattern that correlates with a conscious experience, we expect this neural pattern to repeat when the same experience repeats. Because neither neural patterns nor experiences repeat in a perfectly reproducible way, we also need a way to know whether a certain neural pattern or experience looks like one that occurred in the past. This requires a quantification of resemblance between two measures, that is, a *distance*.

³⁴ This way of defining measures of subjective experience should be sufficiently general to include all the measures used in psychophysics: Choosing to press one button among two, or to press at a particular time, for example, fits into that definition. Psychophysics is indeed partly about first-person data. For example, the experimenter shows a shape to a subject and asks him to press button A if what he sees looks more like a circle, and button B if it looks more like a square. The subject’s answer is based on one particular element of his subjective experience (he selects one particular action in the universe of allowed responses, based on the observation of his conscious visual experience). The button press can therefore be seen as a (very crude) description of a conscious content.

Note that this first definition is large enough to include many possible measures. In fact, a dance could be considered as a measure or a series of measures if each successive body configuration constitutes by itself a measure. A drawing could also be a measure. But to be actually useful, we insist that the subject and the experimenter should agree on a measure of distance, which enables anybody to evaluate the degree of similarity between two measurements. The Basic Requirement (so called in the following) is that the distance should be consistent with the experience of the subject (as only the subject can tell), that is: If measure A is closer to measure B than to measure C, then the elements of experience that led the subject to select measure A should appear to him as closer to the elements that led him to choose measure B than to the elements associated with measure C.³⁵ This requirement directly implies, for instance, that recurrences in the subject's experience should translate into recurrences in the measure.

Once provided with measures of (some elements of) the subjective experience, and with measures of neural phenomena, it should be possible to establish a relationship between the two phenomena by comparing the dynamics of those measures: Related phenomena should provide sets of measures with compatible dynamics. That is, once again, stability in experience should be associated with stability (or stationarity) in the neural dynamics, while moments of change should be correlated with changing (or non-stationary) neural processes.³⁶

³⁵ In other words, what is needed here is first a possible one-to-one *monotonic* correspondence between the phenomena under investigation and the measurements. 'Monotonic' is to be understood in its usual mathematical sense: For three phenomena pa , pb and pc and their corresponding measurements $m(pa)$, $m(pb)$ and $m(pc)$, it would be desirable that if $D(pa, pb) > D(pa, pc)$ (D being a subjective distance between experiential phenomena), then $d(m(pa), m(pb)) > d(m(pa), m(pc))$ (d being the distance defined by the experimenter and the subject in the universe of measures) (for a convergent perspective see also Fell, 2004).

³⁶ This relation implies an additional requirement for the measures of the subjective experience: They should be timed. Indeed, the dynamics of experiential phenomena can only be accessed through series of consecutive timed measures (as simple as a series of button presses, for instance, or the time course of the pressure applied on a joystick). Therefore, to establish a strong relation between the dynamics of an experience and the formation of certain patterns of neural activity, one should be able to say that the experience started at time $t=2s$ and fully developed between $t=5s$ and $t=10s$ (this is easy to understand in the case of an emotional reaction to a sound, for example). It does not follow, however, that the *time as experienced* must correspond precisely to the *timing of neural processes*. The former is a matter of the *content* of experience and the latter of the neural *vehicles* that (in ways we do not fully understand) embody or encode those contents. Within certain small temporal windows, a given neural

7.5. A ‘structural invariants’ approach to first-person data

As a complement to the fine-grained topological description presented above, it seems possible to adopt what can be termed a ‘structural invariant’ strategy. Here the main aim is to obtain, through descriptions of the target experience, an account of that which is invariant (or stable) as a feature of the experience, regardless of whether it is one or another subject that undergoes it. The roots of this approach go back to the method adopted in phenomenological philosophy (see Thompson & Zahavi, this volume). Here, through several repetitions of the same experience in different contexts, one can arrive first at a certain subjective invariant, and then, through contrast with other subjects, intersubjective invariants that are present in the original experience, no matter how many versions of it one tries and no matter how many different subjects engage in it. A traditional example is the structure of the visual field, in which what one sees focally always appears as relatively detailed center surrounded by an increasingly less detailed region, which, at the limit, fades into an ungraspable indeterminacy. In the particular context of the neurodynamics of consciousness, the relevance of this type approach can be illustrated by recent work on the experience of binocular rivalry (Cosmelli et al., 2004).

As we briefly described above, binocular rivalry occurs whenever one is presented with dissimilar images, one to each eye. The subjective experience is that of an ongoing alternation between both possible images, with only one of them consciously perceived at a time. If the images are large, then during the transition from one to the other, one can distinguish a mosaic, patchwork pattern composed of both images, but as a rule, if the adequate contrast and luminance conditions are met, at any given point of the visual image only one of the images (or part of it) will be seen (will dominate) in an exclusive way. In general, binocular rivalry is considered a clear-cut alternation between two states, and average measures of the brain state during one or the other dominance period are contrasted. Most commonly, the subject’s indication via a button press of the moment when the alternation takes place is used to fix a rigid temporal reference around which the average brain responses are defined.

vehicle could encode one *event* as happening *before* another *event*, even though that *neural vehicle* occurs *after* the *neural vehicle* encoding the second event (see Dennett & Kinsbourne, 1992).

We recently used this experimental protocol to investigate the underlying neural patterns, but with the specific objective of describing their spatiotemporal evolution throughout extended periods, and without presupposing a rigid two-state structure (Cosmelli et al., 2004). To do so, we worked with a group of subjects who were extensively exposed to the experience, and produced free, ongoing descriptions of what they were seeing and how they were experiencing it. As conflicting stimuli we used a human face and a moving pattern with an intrinsic frequency (a frequency tag, see (Brown & Norcia, 1997; Tononi & Edelman, 1998). This intrinsic frequency was incorporated in order to tag a neural evoked response that could be followed by magnetoencephalography (MEG).

The descriptions produced by the subjects showed some interesting features: In addition to experiencing the well-known alternation between both images, the subjects repeatedly described this alternation as extremely variable in the way it occurred. Although sometimes the alternation from one image to the other started in the center of the field and progressed towards the outer limits, in other occasions it began on one side, from the top or the bottom, or even from the external borders, and then progressively invaded the pre-existing image. Most subjects claimed that it was difficult to give a stable description of how these transitions took place, because at each time they developed in a different way. Nevertheless, all subjects invariantly stated that dominance periods would alternate and recur, no matter what the subjects did or how much they tried to prevent it from happening.

At a first coarse level, these descriptions already provide us with some crucial aspects of the experience of rivalry: This experience is one of an ongoing flow of *recurrent* dominant periods, in which alternations are extremely *variable* in the way they develop. This feature is indeed a hallmark of binocular rivalry that will be experienced by any normal observer, and is thus a structural invariant in the sense described above. Although this descriptive feature is not particularly novel, it nevertheless points towards a concrete restriction in the *methods* we need to choose to analyze the underlying neural processes (and consequently what we understand as the neural underpinnings of consciousness). If we wish to reveal neural patterns that are meaningful in the context of this specific experience, then we cannot impose a rigid temporal grid and suppose that there is such a thing as an average transition from one image to the other. This point, however, is rarely acknowledged. We therefore developed a statistical framework that considered significant any neural activity that is recurrent in time, without any restrictions on the temporal

pattern of activation. The result was an original description of a network of distributed cortical regions that showed synchronous activation modulated in concert with conscious dominance periods. Moreover, the dynamics of modulation of these brain patterns showed a striking similarity to the bell-type pattern that William James had predicted (more than a century ago) would underlie the occurrence of any given conscious moment (James, 1981).

An important contribution of the structural invariant approach is thus that it can serve as an effective constraint on how we study the dynamic brain patterns. Basic phenomenological observation shows that experience (or the stream of consciousness) is at least (i) dynamic and ongoing; (ii) continuous;³⁷ (iii) can be parsed, so one can distinguish in a given subjective experience components or aspects that are more visual, or more auditory, for instance, and eventually segment it along such dimensions; and (iv) is recurrent, in the sense that we recognize objects, feelings, thoughts, memories, etc., as seen or felt before, even though they are never experienced in the same way. These properties, although certainly not exhaustive of our conscious lives, do suggest that methods that allow for processes of compatible dynamics should be preferred if we want to advance in our understanding of the neural underpinnings of consciousness.

In addition to this methodological constraint, however, the structural invariants approach can potentially make a further contribution. As we mentioned above, one of the most prominent structural invariants of consciousness is precisely its subjective character, in the sense of its fundamental, pre-reflective and pre-conceptual ‘ipsiety’ (see Lutz, Dunne, and Davidson, this volume; Thompson & Zahavi, this volume; see also Zahavi, 2005, for an extended discussion). This backdrop of consciousness pervades the occurrence of specific states of perceptual consciousness. It would appear to call for an explanation not so much in terms of the dynamic behavior of the system (e.g., only in terms of the dynamical properties of the nervous system’s patterns of activity), but rather in terms of how a certain *self-referring perspective can emerge from a certain dynamical organization* (Rudrauf et al., 2003; Thompson, 2007). Whether this type of account is beyond the domain of neurodynamics as we have defined it here is an empirical issue. The crucial point is that if, through some enriched neurodynamical plus

³⁷ ‘Continuous’ here is not meant as the opposite of discrete, but rather is used to mean that consciousness does not jump around with no connection whatsoever from one sort of experience to another.

organismic plus biological approach (e.g., Damasio, 1999; Varela, 1979), one could account for the *conditions of possibility of a minimally subjective system*, then transcending a purely correlational strategy would become a real possibility.

7.6. Can we avoid the pitfalls of Introspectionism ?

One recurrent question, when discussing the use of first-person data, is how to avoid the pitfalls of Introspectionism. Introspectionism was an attempt to use introspection as a scientific method to elaborate psychological theories. It was the main scientific approach to mental phenomena at the beginning of psychology, but was later dismissed by the scientific community in favor of Behaviorism (reviewed by Vermersch, in Depraz, Varela, & Vermersch, 2003). The main problem with introspection, as used at that time, was that it provided conflicting theories. The root of the problem was in fact methodological: It was never possible to ascertain whether the introspective reports met the Basic Requirement mentioned above, and there were serious doubts about the correspondence between the descriptions of the experiences and the experiences themselves. On the other hand, there was little explicit statement of the introspective method by which to proceed to explore and describe experience, and hence the actual testing and refinement of the research method, as opposed to the content of its descriptions, remained underdeveloped (Varela, 1996). Consequently, an important part of the cognitive science community is generally reluctant to use first-person data. Is it therefore possible to build a neurodynamics of consciousness, given that it must rely on first-person data?

Our position is that the whole issue is a technical one: If the measure providing first-person data meets the Basic Requirement described above, then the measure is useful. Alternatively, in the structural invariant approach, if a given invariant is stable across all subjects for a given experimental paradigm, it should be considered valid. In fact, the real question is not whether cognitive scientists should ‘trust the subject’, but in which conditions they can trust the subject, and what they should ask. First-person data, defined as measures of the subjective experience, are continuously being used in psychophysics: When a subject presses a button to indicate that he saw a blue square, and not a red circle, he provides a measure of his immediate perceptual experience, in its simplest form. In this extremely simple form, first-person reports are considered as perfectly valid and trustworthy. At the other extreme, first-person data about the precise dynamic of subtle variations of emotions would probably be considered less reliable (this

means that they would not meet the Basic Requirement—the same subtle variations would not lead to the *same* first-person data, if repeated).³⁸ So, in fact, the real question concerning first-person data is “where shall we draw the line between what is acceptable, perfectly good data, and what is not?”³⁹ A related question of equal importance is whether this line is the same for all individuals, and whether it is fixed within a single individual or whether training can move the line (see Lutz, Dunne, & Davidson, this volume).⁴⁰ We believe that this question should become central in cognitive neuroscience in the near future; especially in view of the advent of new fields, such as the neuroscience of emotions or the neuroscience of consciousness itself. Such emergent fields heavily rely on trustworthy measures of subjective experience.

8. Conclusion

In summary, the neurodynamics of consciousness is an attempt to relate two dynamical phenomena that take place in a subject—the formation of metastable patterns in the subject’s neural activity, and the transient emergence of dissociable elements or aspects of his conscious experience. To establish such a relation, cognitive neuroscientists need to observe systematic similarities between the dynamical properties of these two phenomena. In this sense, the neurodynamical approach works at the level of correlations, albeit refined ones. On the experiential side, this approach requires the subject to provide first-person descriptions that can serve as ‘public’ measures of experience, with at least two objectives. The first objective is to capture reliably the degree of similarity (or disparity) between different subjective phenomena, and produce timings that can be compared to the timing of neural measurements. The second objective is to produce descriptions of the structural invariants of the experience in question, in

³⁸ Consider, however, the possibility of working with individuals that can produce and stabilize mental states more reliably (see Lutz, Dunne, and Davidson, this volume). The issue of working with ‘experts’ or trained subjects is important and controversial (Jack & Roepstorff, 2003; Lutz, Dunne, & Davidson, this volume; Lutz & Thompson, 2003; Varela & Shear, 1999).

³⁹ Cognitive psychologists sometimes ask subjects very difficult questions, so how can they trust their answers? Why shall we trust the button presses of a subject during a binocular rivalry experiment? The subject is asked to press the button as soon as one pattern dominates completely, but how can one be sure that the subject can actually do this task reliably, or that he has this sort of fine capacity to attend to his own visual experience and its dynamics in time?

⁴⁰ There is a similar problem with the measure of neural events. For instance, with EEG, the noise level is sometimes simply so strong that measures of gamma activity cannot be made: The Basic Requirement is not met.

order to constrain the methods that are chosen to determine which neural activity is to be considered significant. It is not yet clear how much of the complexity of consciousness can be revealed in this way, and this question constitutes an important field of investigation for the future.

References

- Abarbanel, H. D., & Rabinovich, M. I. (2001). Neurodynamics: nonlinear dynamics and neurobiology. *Curr Opin Neurobiol*, *11*(4), 423-430.
- Abeles, M. (1991). *Corticonics*: Cambridge University Press.
- Amit, D. J. (1994). The Hebbian paradigm reintegrated: Local reverberations as internal representations. *Behavioral and Brain sciences*, *18*, 617-626.
- Arbib, M. A. (2002). *Handbook of Brain Theory and Neural Networks (2nd ed.)*. Cambridge: MIT Press.
- Arbib, M. A., Érdi, P., & Szentágothai, J. (1997). *Neural Organization: Structure, Function, and Dynamics*. Cambridge, MA: MIT press.
- Arieli, A., Sterkin, A., Grinvald, A., & Aertsen, A. (1996). Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. *Science*, *273*(5283), 1868-1871.
- Ashby, R. (1952). *Design for a Brain*. London.: Chapman-Hall.
- Baars, B. J., & Franklin, S. (2003). How conscious experience and working memory interact. *Trends Cogn Sci*, *7*(4), 166-172.
- Blake, R. A. (1989). A neural theory of binocular rivalry. *Psychol. Rev.*(96), 145-167.
- Blake, R., & Logothetis, N. K. (2002). Visual competition. *Nat Rev Neurosci*, *3*(1), 13-21.
- Block, N. (1996). How can we find the neural correlate of consciousness? *Trends Neurosci*, *19*(11), 456-459.
- Block, N. (1997). On a confusion about a function of consciousness. In N. Block, O. Flanagan, and G. Güzeldere (eds.), *The Nature of Consciousness: Philosophical Debates*, pp. 375-416. Cambridge, MA: The MIT Press/A Bradford Book.
- Block, N. (2001). Paradox and cross purposes in recent work on consciousness. *Cognition* *79*: 197-219.
- Bressler, S. L. (1995). Large-scale cortical networks and cognition. *Brain Res Brain Res Rev*, *20*(3), 288-304.
- Bressler, S. L., & Freeman, W. J. (1980). Frequency analysis of olfactory system EEG in cat, rabbit, and rat. *Electroencephalogr Clin Neurophysiol*, *50*(1-2), 19-24.

- Bressler, S. L., & Kelso, J. A. (2001). Cortical coordination dynamics and cognition. *Trends Cogn Sci*, 5(1), 26-36.
- Brown, R., & Kocarev, L. (2000). A unifying definition of synchronization for dynamical systems. *Chaos*, 10(2), 344-349.
- Brown, R. J., & Norcia, A. M. (1997). A method for investigating binocular rivalry in real-time with steady-state VEP. *Vision Res*, 37, 2401-2408.
- Buchel, C., & Friston, K. J. (1997). Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fMRI. *Cereb Cortex*, 7(8), 768-778.
- Chalmers, D. J. (2000). What is a neural correlate of consciousness. In T. Metzinger (Ed.), *Neural Correlates of Consciousness: Empirical and Conceptual Questions*. Cambridge, MA: MIT Press.
- Chandrasekhar, S. (1961). *Hydrodynamic and Hydromagnetic stability*. Oxford: Oxford University Press.
- Cosmelli, D., David, O., Lachaux, J. P., Martinerie, J., Garnero, L., Renault, B., et al. (2004). Waves of consciousness: ongoing cortical patterns during binocular rivalry. *Neuroimage*, 23(1), 128-140.
- Crick, F. C., & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in Neuroscience*, 2, 263-275.
- Crick, F.C., & Koch, C. (1998). Consciousness and neuroscience. *Cereb Cortex*, 8(2), 97-107.
- Crick, F.C., & Koch, C. (2003). A framework for consciousness. *Nat Neurosci*, 6(2), 119-126.
- Crutchfield, J. P., & Kaneko, K. (1987). Phenomenology of Spatiotemporal Chaos. In *Directions in Chaos* (pp. 272-353): World Scientific.
- Damasio, A. R. (1998). Investigating the biology of consciousness. *Philos Trans R Soc Lond B Biol Sci*, 353(1377), 1879-1882.
- Damasio, A. (1999). *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*: Harcourt.
- Danckert, J., & Goodale, M. A. (2000). A conscious route to unconscious vision. *Curr Biol*, 10(2), 64-67.
- David, O., Cosmelli, D., & Friston, K. J. (2004). Evaluation of different measures of functional connectivity using a neural mass model. *Neuroimage*, 21(2), 659-673.

- Dehaene, S., Kerszberg, M., & Changeux, J. P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proc Natl Acad Sci U S A*, 95(24), 14529-14534.
- Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*, 79(1-2), 1-37.
- Dehaene, S., Sergent, C., & Changeux, J. P. (2003). A neuronal network model linking subjective reports and objective physiological data during conscious perception. *Proc Natl Acad Sci U S A*, 100(14), 8520-8525.
- Dennett, D.C. and Kinsbourne, M. (1992). Time and the observer: the where and when of consciousness in the brain. *Behavioral and Brain Sciences* 15: 183-247.
- Depraz, N., Varela, F., & Vermersch, P. (2003). *On becoming aware. A pragmatics of experiencing*. Amsterdam: Benjamin.
- Edelman, G. M., & Tononi, G. (2000). *A Universe of Consciousness: how matter becomes imagination*. New York: Basic Books.
- Engel, A. K., Fries, P., Konig, P., Brecht, M., & Singer, W. (1999). Temporal binding, binocular rivalry, and consciousness. *Conscious Cogn*, 8(2), 128-151.
- Engel, A. K., Fries, P., & Singer, W. (2001a). Dynamic predictions: oscillations and synchrony in top-down processing. *Nat Rev Neurosci*, 2(10), 704-716.
- Engel, A. K., & Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends Cogn Sci*, 5(1), 16-25.
- Faure, P., & Korn, H. (2001). Is there chaos in the brain? I. Concepts of nonlinear dynamics and methods of investigation. *C R Acad Sci III*, 324(9), 773-793.
- Fell, J. (2004). Identifying neural correlates of consciousness: the state space approach. *Consciousness and Cognition* 13: 709-29.
- Fell, J., Fernandez, G., Klaver, P., Elger, C. E., & Fries, P. (2003). Is synchronized neuronal gamma activity relevant for selective attention? *Brain Res Brain Res Rev*, 42(3), 265-272.
- Fell, J., Klaver, P., Elger, C. E., & Fernandez, G. (2002). Suppression of EEG gamma activity may cause the attentional blink. *Conscious Cogn*, 11(1), 114-122.
- Freeman, W. J. (1975). *Mass Action in the Nervous System*: Elsevier Science & Technology Books.
- Freeman, W.J. (1999a). *How brains make up their minds*. London: Weidenfeld and Nicolson.

- Freeman, W. J. (1999b). Consciousness, intentionality and causality. *Journal of Consciousness Studies*, 6(11-12), 143-172.
- Freeman, W. J. (2000a). *Neurodynamics: An Exploration of Mesoscopic Brain Dynamics*. London: Springer-Verlag.
- Freeman, W. J. (2000b). Mesoscopic neurodynamics: from neuron to brain. *J Physiol Paris*, 94(5-6), 303-322.
- Freeman, W. J. (2004). *Mass Action in the Nervous System. Second (electronic) printing*. URL: <http://sulcus.berkeley.edu/>
- Fries, P., Roelfsema, P. R., Engel, A. K., Konig, P., & Singer, W. (1997). Synchronization of oscillatory responses in visual cortex correlates with perception in interocular rivalry. *Proc Natl Acad Sci U S A*, 94(23), 12699-12704.
- Fries, P., Reynolds, J.H., Rorie, A.E., Desimone, R. (2001) Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* 291:1560-1563.
- Friston, K. J. (1997). Transients, metastability, and neuronal dynamics. *Neuroimage*, 5(2), 164-171.
- Friston, K. J. (2000). The labile brain. I. Neuronal transients and nonlinear coupling. *Philos Trans R Soc Lond B Biol Sci*, 355(1394), 215-236.
- Friston, K. J. (2002a). Beyond Phrenology: What can Neuroimaging Tell Us About Distributed Circuitry? *Annu Rev Neurosci*, 25, 221-250.
- Friston, K. (2002b). Functional integration and inference in the brain. *Prog Neurobiol*, 68(2), 113-143.
- Friston, K. J. (2005). Models of brain function in neuroimaging. *Annu Rev Psychol.*, 56, 57-87.
- Garfinkel, A., Spano, M.L., Ditto, W.L., & Weiss, J.N. (1992). Controlling cardiac chaos. *Science* 257: 1230-1235.
- Gomes, G. (1998). The timing of conscious experience: a critical review and reinterpretation of Libet's research. *Conscious Cogn*, 7(4), 559-595.
- Gray, C. M., Konig, P., Engel, A. K., & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, 338(6213), 334-337.
- Gurwitsch, A. (1964). *The field of consciousness*. Pittsburgh: Duquesne University Press.

- Gutkin, B., Pinto, D., & Ermentrout, B. (2003). Mathematical neuroscience: from neurons to circuits to systems. *J Physiol Paris*, 97(2-3), 209-219.
- Haken, H. (1983). *Synergetics: An Introduction: Nonequilibrium Phase Transitions and Self-Organization in Physics, Chemistry, and Biology*. New York: Springer-Verlag.
- Hanna, R., & Thompson, E. (2003). The Spontaneity of Consciousness. In E. Thompson (Ed.), *New Essays in Phenomenological Philosophy of Mind. Canadian Journal of Philosophy, Supplementary Volume 29* (pp. 133-162).
- Herculano-Houzel, S., Munk, M. H., Neuenschwander, S., & Singer, W. (1999). Precisely synchronized oscillatory firing patterns require electroencephalographic activation. *J Neurosci*, 19(10), 3992-4010.
- Hilgetag, C. C., & Kaiser, M. (2004). Clustered organization of cortical connectivity. *Neuroinformatics*, 2(3), 353-360.
- Hollender, D. (1986). Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: A survey and appraisal. *Behavioral and Brain Sciences*, 9, 1-66.
- Ito, J., & Kaneko, K. (2002). Spontaneous Structure Formation in a Network of Chaotic Units with Variable Connection Strengths. *Physical Review Letters*, 88(2).
- Jack, A., & Roepstorff, A. (2003). Trusting the subject (volume 1). *Journal of Consciousness Studies*, 10(9-10).
- James, W. (1981). *The Principles of Psychology*. Cambridge, MA: Harvard University Press.
- Kelso, J. A. S. (1995). *Dynamic Patterns: The Self-Organization of Brain and Behavior*. Cambridge, MA: MIT Press.
- Korn, H., & Faure, P. (2003). Is there chaos in the brain? II. Experimental evidence and related models. *C R Biol*, 326(9), 787-840.
- Kuramoto, Y. (1984). Cooperative dynamics of oscillator community: A study based on lattice of rings. *Progress of Theoretical Physics Suppl*, 79, 223-240.
- Lachaux, J.-P., George, N., Tallon-Baudry, C., Martinerie, J., Hugueville, L., Minotti, L., et al. (2005). The many faces of the gamma band response to complex visual stimuli. *NeuroImage*, 25(2), 491-501.

- Lachaux, J. P., Lutz, A., Rudrauf, D., Cosmelli, D., Le Van Quyen, M., Martinerie, J., et al. (2002). Estimating the time-course of coherence between single-trial brain signals: an introduction to wavelet coherence. *Neurophysiol Clin*, 32(3), 157-174.
- Lachaux, J. P., Rodriguez, E., Martinerie, J., Adam, C., Hasboun, D., & Varela, F. J. (2000). A quantitative study of gamma-band activity in human intracranial recordings triggered by visual stimuli. *Eur J Neurosci*, 12(7), 2608-2622.
- Lachaux, J. P., Rodriguez, E., Martinerie, J., & Varela, F. J. (1999). Measuring phase synchrony in brain signals. *Hum Brain Mapp*, 8(4), 194-208.
- Lachaux, J. P., Rudrauf, D., & Kahane, P. (2003). Intracranial EEG and human brain mapping. *J Physiol Paris*, 97(4-6), 613-628.
- Le Van Quyen, M. (2003). Disentangling the dynamic core: a research program for a neurodynamics at the large-scale. *Biol Res*, 36(1), 67-88.
- Le Van Quyen, M., Foucher, J., Lachaux, J.-P., Rodriguez, E., Lutz, A., Martinerie, J., et al. (2001). Comparison of Hilbert transform and wavelet methods for the analysis of neuronal synchrony. *J Neurosci Methods*, 111(2), 83-98.
- Leopold, D. A., & Logothetis, N. K. (1999). Multistable phenomena: changing views in perception. *Trends Cogn Sci*, 3(7), 254-264.
- Libet, B. (2002). The timing of mental events: Libet's experimental findings and their implications. *Conscious Cogn*, 11(2), 291-299; discussion 304-233.
- Llinas, R. (1988). The intrinsic electrophysiological properties of mammalian neurons: insights into central nervous system function. *Science*, 242(4886), 1654-1664.
- Llinas, R., & Pare, D. (1991). Of dreaming and wakefulness. *Neuroscience*, 44(3), 521-535.
- Llinas, R., & Ribary, U. (2001). Consciousness and the brain. The thalamocortical dialogue in health and disease. *Ann NY Acad Sci.*, 929, 166-175.
- Llinas, R., Ribary, U., Contreras, D., & Pedroarena, C. (1998). The neuronal basis for consciousness. *Philos Trans R Soc Lond B Biol Sci*, 353(1377), 1841-1849.
- Lutz, A., Lachaux, J. P., Martinerie, J., & Varela, F. J. (2002). Guiding the study of brain dynamics by using first-person data: synchrony patterns correlate with ongoing conscious states during a simple visual task. *Proc Natl Acad Sci U S A*, 99(3), 1586-1591.

- Lutz, A. and Thompson, E. (2003). Neurophenomenology: Integrating subjective experience and brain dynamics in the neuroscience of consciousness. *Journal of Consciousness Studies* 10: 31-52.
- Martinerie, J., Adam, C., Le Van Quyen, M., Baulac, M., Clemenceau, S., Renault, B., et al. (1998). Epileptic seizures can be anticipated by non-linear analysis. *Nat Med*, 4(10), 1173-1176.
- Maturana, H. R. and Varela, F J. (1980). *Autopoiesis and Cognition: The Realization of the Living*. Boston Studies in the Philosophy of Science, vol. 42. Dordrecht: D. Reidel.
- Mesulam, M. M. (1999). Spatial attention and neglect: parietal, frontal and cingulate contributions to the mental representation and attentional targeting of salient extrapersonal events. *Philos Trans R Soc Lond B Biol Sci.*, 354(1387), 1325-1346.
- Metzinger, T. (Ed.). (2000). *Neural Correlates of Consciousness: Empirical and Conceptual Questions*: MIT Press.
- Milner, D., & Goodale, M. (1995). *The visual brain in action*. Oxford: Oxford University Press.
- Munk, M. H., Roelfsema, P. R., König, P., Engel, A. K., & Singer, W. (1996). Role of reticular activation in the modulation of intracortical synchronization. *Science*, 272(5259), 271-274.
- Nicolis, G., & Prigogine, G. (1977). *Self-organization in nonequilibrium systems*. New-York: Wiley.
- Nicolis, G., & Prigogine, I. (1989). *Exploring Complexity: An Introduction*. New York: W.H. Freeman & Company.
- Noë, A., & Thompson, E. (2004a). Are there neural correlates of consciousness? *Journal of Consciousness Studies* 11: 3-28.
- Noë, A. and Thompson, E. (2004b). Sorting out the neural basis of consciousness. Authors' reply to commentators. *Journal of Consciousness Studies* 11: 87-98.
- Nunez, P. L. (2000). Toward a Quantitative Description of Large Scale Neocortical Dynamic Function and EEG. *Behavioral and Brain Sciences*, 23(3).
- O'Regan, K., & Noë, A. (2001). sensorimotor account of vision and visual consciousness. *Behav Brain Sci*, 24(5), 939-973.
- Palus, M. (1996). Nonlinearity in normal human EEG: cycles, temporal asymmetry, nonstationarity and randomness, not chaos. *Biol. Cybern.*, 75, 389-396.

- Panksepp, J. (1998). The periconscious substrates of consciousness: Affective states and the evolutionary origins of self. *Journal of Consciousness Studies* 5: 566-582.
- Parvizi, J. and Damasio, A. (2001). Consciousness and the brainstem. *Cognition* 79: 135-159.
- Perkel, D. H., Gerstein, G. L., & Moore, G. P. (1967). Neuronal spike trains and stochastic point processes. I. The single spike train. *Biophys J.*, 7(4), 391-418.
- Petitmengin, C. (in press). Describing one's subjective experience in the second-person. *Phenomenology and the Cognitive Sciences*.
- Pessoa, L., Thompson, E., & Noë, A. (1998). Finding out about filling-in: a guide to perceptual completion for visual science and the philosophy of perception. *Behav Brain Sci.*, 21(6), 723-748.
- Pikovsky, A., Zaks, M., Rosenblum, M., Osipov, G., & Kurths, J. (1997). Phase synchronization of chaotic oscillations in terms of periodic orbits. *Chaos*, 7(4), 680-687.
- Rees, G., Kreiman, G., & Koch, C. (2002). Neural correlates of consciousness in humans. *Nat Rev Neurosci*, 3(4), 261-270.
- Revonsuo, A., Wilenius-Emet, M., Kuusela, J., & Lehto, M. (1997). The neural generation of a unified illusion in human vision. *Neuroreport*, 8(18), 3867-3870.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.*, 27, 169-192.
- Rodriguez, E., George, N., Lachaux, J. P., Martinerie, J., Renault, B., & Varela, F. J. (1999). Perception's shadow: long-distance synchronization of human brain activity. *Nature*, 397(6718), 430-433.
- Rodriguez, R., Kallenbach, U., Singer, W., & Munk, M. H. (2004). Short- and long-term effects of cholinergic modulation on gamma oscillations and response synchronization in the visual cortex. *J Neurosci*, 24(46), 10369-10378.
- Rosenblum, M., Pikovsky, A., & Kurths, J. (1996). Phase Synchronization of Chaotic Oscillators. *Physical Review Letters*, 76(11), 1804-1807.
- Roskies, A. L. (1999). The Binding Problem : Special Issue. *Neuron*, 24(7-9).
- Rudrauf, D., Lutz, A., Cosmelli, D., Lachaux, J.-P., and Le Van Quyen, M. (2003). From autopoiesis to neurophenomenology: Francisco Varela's exploration of the biophysics of being. *Biological Research* 36: 27-66.

- Schiff, S.J., Jerger, K., Duong, D.H., Chang, T., & Spano, M.L. (1994). Controlling chaos in the brain. *Nature* 8: 615-620.
- Searle, J. (2000). Consciousness. *Annual Review of Neuroscience*, 23.
- Sewards, T. V., & Sewards, M. A. (2001). On the correlation between synchronized oscillatory activities and consciousness. *Conscious Cogn*, 10(4), 485-495.
- Sherman, S. M., & Guillery, R. W. (2002). The role of the thalamus in the flow of information to the cortex. *Philos Trans R Soc Lond B Biol Sci*, 357(1428), 1695-1708.
- Sherover, C. (1991). *The Human Experience of Time*: Northwestern University Press.
- Simons, D. J. (2000). Attentional capture and inattention blindness. *Trends Cogn Sci*, 4(4), 147-155.
- Singer, W. (1998). Consciousness and the structure of neuronal representations. *Philos Trans R Soc Lond B Biol Sci*, 353(1377), 1829-1840.
- Singer, W., & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annu Rev Neurosci*, 18, 555-586.
- Skarda, C. A., & Freeman, W. J. (1987). How brains make chaos in order to make sense of the world. *Behavioral and Brain Sciences*, 10(2), 161-195.
- Stephan, K. E., Hilgetag, C. C., Burns, G. A., O'Neill, M. A., Young, M. P., & Kotter, R. (2000). Computational analysis of functional connectivity between areas of primate cerebral cortex. *Philos Trans R Soc Lond B Biol Sci*, 355(1393), 111-126.
- Tallon-Baudry, C., & Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends Cogn Sci*, 3(4), 151-162.
- Tallon-Baudry, C., Bertrand, O., & Fischer, C. (2001). Oscillatory synchrony between human extrastriate areas during visual short-term memory maintenance. *J Neurosci*, 21(20), RC177.
- Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambridge, MA: Harvard University Press.
- Thompson, E., Lutz, A., and Cosmelli, D. (2005). Neurophenomenology: an introduction for neurophilosophers. In A. Brook and K. Akins, eds., *Cognition and the Brain: The Philosophy and Neuroscience Movement*, pp. 40-97. New York and Cambridge: Cambridge University Press.

- Thompson, E., & Varela, F. J. (2001). Radical embodiment: neural dynamics and consciousness. *Trends Cogn Sci*, 5(10), 418-425.
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381(6582), 520-522.
- Tononi, G., & Edelman, G. M. (1998). Consciousness and complexity. *Science*, 282(5395), 1846-1851.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognit Psychol*, 12(1), 97-136.
- Tsuda, I. (2001). Toward an interpretation of dynamic neural activity in terms of chaotic dynamical systems. *Behav Brain Sci*, 24(5), 793-810; discussion 810-748.
- van Gelder, T. J. (1998). The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences*, 21, 1-14.
- Van Gelder, T. J. (1999). Dynamic approaches to cognition. In R. Wilson & F. Keil (Eds.), *The MIT Encyclopedia of Cognitive Sciences* (pp. 244-246). Cambridge MA: MIT Press.
- VanRullen, R., & Koch, C. (2003). Is perception discrete or continuous? *Trends Cogn Sci.*, 7(5), 207-213.
- Varela, F. (1995). Resonant cell assemblies: A new approach to cognitive functioning and neuronal synchrony. *Biol.Research*, 28(81-95).
- Varela, F. J. (1996). Neurophenomenology : A Methodological Remedy for the Hard Problem. *Journal of Consciousness Studies*, 3(4), 330-349.
- Varela, F. (1999). The specious present: The neurophenomenology of time consciousness. In J. Petitot, F. J. Varela, B. Pachoud & J. M. Roy (Eds.), *Naturalizing Phenomenology* (pp. 266-314). Stanford: Stanford University Press.
- Varela, F., Lachaux, J. P., Rodriguez, E., & Martinerie, J. (2001). The brainweb: phase synchronization and large-scale integration. *Nat Rev Neurosci*, 2(4), 229-239.
- Varela, F. J., & Shear, J. (1999). First-person methodologies: why, when and how. *J.Consc.Studies*, 6(2-3), 1-14.
- Von der Malsburg, C. (1981). *The correlation theory of brain function*. Göttingen: Germany: Dept. of Neurobiology, Max-Plank-Institute for Biophysical Chemistry.
- Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684), 440-442.

- Weiskrantz, L. (1990). The Ferrier lecture, 1989. Outlooks for blindsight: explicit methodologies for implicit processes. *Proc R Soc Lond B Biol Sci.*, 239(1296), 247-278.
- Wicker, B., Ruby, P., Royet, J. P., & Fonlupt, P. (2003). A relation between rest and the self in the brain? *Brain Res Brain Res Rev*, 43(2), 224-230.
- Wolfe, J. M., & Bennett, S. C. (1997). Preattentive object files: shapeless bundles of basic features. *Vision Res*, 37(1), 25-43.
- Wright, J. J., & Liley, D. T. J. (1996). Dynamics of the brain at global and microscopic scales: Neural networks and the EEG. *Behavioral and Brain Sciences*, 19, 285-320.
- Zahavi, D. (2005). *Subjectivity and Selfhood: Investigating the First-Person Perspective*. Cambridge, MA: The MIT Press/A Bradford Book.
- Zihl, J., Von Cramon, D., & Mai, N. (1983). Selective disturbance of movement vision after bilateral brain damage. *Brain*, 106(2), 313-340.