Electrophysiological correlates of synchronous neural activity and attention: a short review

Ernst Niebur *

Department of Neuroscience, The Zanvyl Krieger Mind/Brain Institute, Johns Hopkins University, 3400 N Charles Street, Baltimore, MD 21218, USA

Accepted 22 August 2002

Abstract

Attentional selection implies preferential treatment of some sensory stimuli over others. This requires differential representation of attended and unattended stimuli. Most previous research has focused on pure rate codes for this representation but recent evidence indicates that a mixed code, involving both mean firing rate and temporal codes, may be employed. Of particular interest is a distinction of attended from unattended stimuli based on synchrony within neural populations. I review electrophysiological evidence at macroscopic, mesoscopic and microscopic spatial scales showing that the degree of synchronous activity varies with the attentional state of the perceiving organism.

Keywords: Attention; Synchrony; Neural coding

1. Introduction

In order to function in a complex and changing environment, organisms need to collect information from a multitude of sensors in multiple modalities. At a given time, not all of the information is, however, relevant for making behavioral decisions and it would be wasteful (and thus harmful from an evolutionary point of view) to process irrelevant information in detail. One of the most important parts of sensory perception is therefore the intelligent triage of relevant from irrelevant information (Niebur and Koch, 1998). This process is commonly called selective attention and it has been observed experimentally in a multitude of species, including those as low as drosophila (Wu et al., 2000). For a review of the psychophysics of selective attention in humans, see Egeth and Yantis (1997).

I will outline a functional problem for the representation of attention in the following section. Subsequently, Section 3 proposes a solution, namely that the behavioral and attentional state may be represented using a different code than that employed for the representation of stimulus properties or motor execution. Subsequently, I will review the evidence for a correlation of attentional state with synchrony at the macroscopic (EEG and MEG) level (Section 4), mesoscopic level (Section 5) and microscopic level (Section 6).
and at the level of individual neuron pairs (Section 6). The mechanisms underlying coding and decoding by synchrony are discussed in Section 7.

2. Attentional state of stimuli

If organisms can either attend to a given stimulus or not, there must be a mechanism to distinguish between these two cases. In other words, if a stimulus can be present at one occasion in which the context determines that it should be attended and at another occasion the same stimulus is not attended, then the neuronal representation of the stimulus cannot be determined completely by the physical properties of the stimulus. Instead, the difference in the attentional state of the stimulus (is it attended or not?) must be reflected in a difference of its representation (Niebur, 2000).

The nature of the representation of sensory stimuli is incompletely understood. One very successful model is rate coding, in which the representation is based on differential mean firing rates of neurons or populations of neurons (for a recent overview of the role of rate codes in attention, see Treue, 2001).

There is excellent evidence that many physical properties of sensory stimuli are represented in terms of a rate code (Adrian and Zotterman, 1926). It is also known that at least some sensory stimuli are represented not in terms of mean rates but in a temporal code. I will briefly come back to temporal coding of stimuli towards the end of this review.

3. Quasi-orthogonal coding of stimulus properties and attention

Whatever neural code is used centrally to represent the physical stimulus properties, the full neuronal representation of a stimulus must include additional properties that are essentially independent of, or orthogonal to, the physical stimulus. The best studied example of such properties, and the only one I consider here, is the attentional state attributed to it\(^1\). A pure mean rate code (or any other code based on one variable) for physical properties and attentional state appears problematic for representing two nearly orthogonal variables (physical properties and attentional state) because different combinations of these variables can be mapped onto the same rate.

For instance, assume that the contrast of a visual stimulus is represented by the firing rate of a certain neural population, with the highest contrast corresponding to the highest mean rate. Assume, furthermore, that the firing rates of neurons representing unattended stimuli are suppressed relative to those representing attended stimuli. An intermediate value of the mean firing rate then is ambiguous; it could represent either an unattended high-contrast stimulus or an attended low-contrast stimulus.

An alternative to a pure rate code for the physical properties of a stimulus and its attentional state is that these different variables are coded in different parameters of the underlying neural spike trains. It is conceivable that selective attention is represented in terms of the temporal fine structure of spike trains, which is, indeed, largely independent of the mean firing rate\(^2\).

In particular, it was proposed in modeling work that neural responses to attended stimuli may be more synchronized between members of the neural population that represent the various physical properties in their mean rates. Based on conceptual work by Crick and Koch (1990a,b), detailed models were developed for the representation of attentional selection by either periodic (Niebur et al., 1993) or aperiodic (Niebur and Koch, 1994) synchrony.

\(^1\) It is likely that there is a need to represent, in addition, other state variables, for instance emotions associated with a stimulus. Much less is known about these states than about selective attention and I will not discuss them here.

\(^2\) Clearly, this independence is not absolute. If no action potentials are present (firing rate is 0), no temporal structure can be imposed on the (non-existent) spikes. The same is the case if the neuron fires at the maximum rate that is physiologically possible. Neither case in common.
It is important to note that this scheme necessarily leads to a mixed code that, at least at some stage of processing, requires a representation in terms of mean firing rate. The reason is that any behavior has eventually to lead to a change in muscle activity and that the force developed by a muscle seems to be determined essentially by the aggregate rate of discharge across all neurons presynaptic to this muscle. There are well-known neural mechanisms that convert synchrony codes into rate codes (the simplest being coincidence detection: a neuron fires only if a certain number of presynaptic neurons fires within a short time interval) and the aforementioned models specify that a rate code is introduced at a relatively early neuronal stage. In the cited modeling work it is assumed that this transformation from a temporal code to a rate code is only partial and that a detectable temporal code persists even at that stage and is available to distinguish attended and unattended signals at the subsequent neuronal level. Thus, the different types of features that characterize a stimulus—on the one hand its physical properties, on the other hand the attentional state (and possibly other internally generated states)—are available through many if not all processing stages beyond the level at which attentional control is introduced (e.g. they would not yet be present in the retina). These different types of signals then travel along the same information channels (i.e. axons carrying sequences of action potentials). Demultiplexing these composite signals yields either one or the other part of this information, or possibly combinations of both. Although the suggestions made for cortical microcircuitry that can process this kind of mixed code were not very detailed, it is clear that the ambiguity between weak, attended and strong, unattended stimuli can be resolved since the temporal signal is always available to distinguish between the two. The modeling studies showed that this mixed representation can then be processed repeatedly and the temporal signal used to further enhance the difference in rates between attended and unattended neurons, eventually increasing the difference to the extent that only the former will have behavioral consequences. Synchrony could persist until the very ‘output’ stage of the system, i.e. in the motor system. Interesting in this respect are the observations of synchronized activity in the motor system that are not only task dependent (Adams et al., 1989; Schmied et al., 1993) but specifically correlated with selective attention (Baker et al., 1997, 1999; Schmied et al., 2000).

The models could quantitatively explain neurophysiological key findings (e.g. Moran and Desimone, 1985) and they led to clear predictions, in particular that the degree of synchronous firing of a neural population representing a stimulus should vary with the attentional state of this stimulus. In the following three sections (Sections 4–6), I review experimental evidence supporting this hypothesis for different sizes of the probed neuronal populations.

4. Macroscopic patterns of neural activity

Although direct tests of the model predictions were lacking at the time they were made, a large body of evidence already hinted at a correlation between attention and coherent neural activity, manifested in the form of fast temporal variations (in the range of beta and gamma oscillations, i.e. about 15–80 Hz) of neural signals at multiple spatial scales. It should be noted that any systematic deviation from noise of the macroscopic electromagnetic fields generated by the brain is evidence for large-scale synchrony of the underlying neural activity since only synchronous activity of large neuronal populations can sum up and lead to macroscopic effects.

While there is much evidence for a correlation between spatially coherent activity and cognitive states in general, recently reviewed by Pulvermüller et al. (1997) and by Varela et al. (2001), fewer studies have explicitly focused on the attentional state. An early example, more than 25 years old, is a study by Schvent et al. (1976) demonstrating that event related (or evoked) potentials differ significantly between attended and unattended situations. More recently, Desmedt and Tomberg (1994) showed that activity in prefrontal and parietal cortex synchronizes in a state of somatic selective attention. For a recent review of this literature see Herrmann and Knight (2001).
Evoked potentials were shown to be influenced not only by attentional selection in a single modality but also by cross-modal attention (reviewed by Eimer and Driver, 2001). If no temporal locking of stimulus and macroscopic electromagnetic activity is assumed, modulations of the frequency contents (which reflect underlying neuronal synchrony in the respective frequency bands) is studied for both electrical (electroencephalogram (EEG)) and magnetic (magnetoencephalogram (MEG)) signals. Of particular importance are modulations in fast frequency bands, notably gamma (25–80 Hz) but also beta (15–25 Hz). Recent examples of research showing the influence of selective attention on large-scale gamma activity in the human brain are studies by Shibata et al. (1999), Gruber et al. (1999), Müller et al. (2000) in EEG, and by Sokolov et al. (1999) in MEG. Conversely, it was shown by Munk et al. (1996) that electrical activation of the mesencephalic reticular formation, known to lead to EEG states corresponding to higher states of vigilance, increases stimulus-induced synchronization in the visual cortex. Wrobel (2000) and by Keil et al. (2001) review further evidence suggesting that coherent high frequency activity in the beta and gamma correlates with attention.

5. Mesoscopic patterns of neural activity

At the scale of smaller neural populations, experimental evidence for a correlation of synchrony with attention was provided by simultaneous recordings of the local field potential (LFP) in primate sensorimotor cortex of behaving primates. Murthy and Fetz (1996a) and by Fetz et al. (2000) used simultaneously recorded LFPs in sensorimotor cortex. They found that episodes of 20–40 Hz oscillations occurred frequently and became synchronized over a large cortical area during exploratory arm movements while the animal performed a task involving complex motor actions and sensory feedback (retrieving a treat from a Kluvver board). The oscillations were found to have no reliable relation to particular components of the movement and therefore, seem unlikely to be involved directly in movement execution. Fetz and collaborators concluded that, instead, they may represent a neural correlate of attention during demanding sensorimotor behaviors. These results confirmed earlier findings showing that neuronal synchronization is enhanced during periods of focused attention in primates (Rougeul et al., 1978).

Fast coherent oscillations were likewise observed in recordings of LFP and multi-unit activity in primate motor cortex where it again was found that they were altered by behavioral condition and, consequently, believed to be related to global processes and not the details of motor action that is encoded in the firing rate (Donoghue et al., 1998).

Most work so far has been focused on the primate but correlations between the internal perceptive state of the animal and the synchrony between cortical cells have also be observed in the awake behaving cat. One striking example are the observations of strong and consistent gamma activity in the electro-corticogram recorded in sensorimotor cortex of cats attending intensely to a behaviorally highly important stimulus (a living mouse; Bouyer et al., 1981). Roelfsema et al. (1997) showed that tight synchrony between widely separated cortical areas was obtained when the behaving cat responded to novel situations but not during reward or inter-trial periods. The support for the basic hypothesis put forward in this review is somewhat circumstantial in this case, relying on the conjecture that novel situations will be met by the animal with a higher level of attention than more familiar ones. More recently, von Stein et al. (2000) found that corticocortical synchronization in perceiving and behaving cats reflects the internal state of the animals (expectancy and behavioral goals), particularly in the theta and alpha frequency ranges.

6. Synchronous firing of individual cell pairs

In the motor system, evidence implicating tight neural synchrony with the expectation of a stimulus was provided by Riehle et al. (2000). Again, it is not a priori clear that the period when the animal is expecting a stimulus (that the animal
then has to respond to in a differentiated way) is necessarily characterized by heightened attention, compared with periods when the animal is certain that no action is required and no relevant information provided. But it seems at least possible, and maybe plausible, that this is the case. More specific control of selective attention would be preferable but experiments in which the internal state of the animal is controlled and in which neurophysiological recordings of individual pairs of neurons or small populations of neurons are performed that allow to draw conclusions about the synchrony of the population are few and far between. In the absence of a tighter control of the attentional state (see below for some examples where this is the case) we will take those results as, at least, circumstantial evidence.

In the preparation studied by Richle et al. (2000), multiple single-neuron activity was recorded in primate motor cortex during the performance of delayed pointing tasks. Significant synchronous activity was observed during the expectation period, with synchrony typically becoming tighter towards the end of this period. In a different task (center-out reaching), Hatsopoulos et al. (1998) likewise found synchrony in primate motor cortex, strongest just before movement initiation. A slightly different time course was described by Baker et al. (2001) who found synchrony between individual neurons (as well as between groups of neurons) in monkey primate motor cortex to be task-dependent. In a precision-grip task, synchrony was maximal during the ‘hold’ period (while the monkey was attentively waiting) and much smaller in the movement part (at which time the perceptual decision has been completed), in contradistinction to the firing rate that showed the opposite effect. This result is consistent with earlier reports (Baker et al., 1997; Donoghue et al., 1998) of oscillations in motor cortex being encountered mainly during steady holding and not during the execution phase of a motor task, in particular when taken in combination with reports by Murthy and Fetz (1996b) that synchrony between individual units in sensorimotor cortex is strongest during periods of LFP oscillations.

In the primate visual system, de Oliveira et al. (1997) studied the influence of attentive expectation in the dorsal pathway which is known to be involved in motion detection. Previously, Newsome and colleagues (Newsome et al., 1989) had demonstrated that the physical stimulus properties (direction of motion) are represented in terms of the discharge rates of neuronal populations. In a similar motion discrimination task, de Oliveira et al. (1997) studied the temporal correlations between simultaneously recorded pairs of units in the dorsal pathway. They found these pairs to fire synchronously during the expectation period before stimulus presentation and a contrast-dependent reduction of correlation strength on stimulus onset. However, correlation strength did not vary systematically with stimulus direction. Therefore, synchrony does not convey specific information about the physical stimulus properties (direction of motion) and they suggested that it, instead, could be related to attentive expectation.

A study designed as a direct test of the hypothesis from the Niebur and Koch (1994) model (see end of Section 3) was performed by Steinmetz et al. (2000). They investigated the synchronous firing of neuron pairs in area SII macaque somatosensory cortex, an area whose spike rates are known to be influenced by selective attention (Hsiao et al., 1993). The animals were trained to switch attention between a visual task and a tactile discrimination task. A majority of neuron pairs in SII cortex fired synchronously and, consistent with the model predictions, the degree of synchrony was affected by the monkey’s attentional state. Synchrony increased in 80% and decreased in 20% of neuron pairs when the monkey attended to the tactile task.

Further evidence for the correlation between synchrony and attention was found in the primate visual system by Fries et al. (2001). These authors recorded both multi-unit activity and LFP in monkey extrastriate area V4 while the animals switched attention between behaviorally relevant stimuli and distracters. Neurons activated by the attended stimulus showed increased synchronization in the gamma range but reduced low-frequency (<17 Hz) synchronization compared with neurons at nearby V4 sites activated by distracters.
7. Causes and effects

In this section, I will briefly discuss three issues. The first is a brief presentation of a suggested alternative role for synchronous firing. The second is the question of possible mechanisms that can generate the observed synchrony between neurons representing attended stimuli. In other words, who creates the synchrony, and how? The third is the question of possible mechanisms that make use of neural synchrony. In other words, who will decode the synchrony, and how?

Regarding the first question, representation of the attentional state is not the only role that has been attributed to the synchrony structure of neural spike trains. Excellent arguments have been made in favor of a view in which synchrony is used to ‘bind together’ different parts of the representation of objects (von der Malsburg, 1981; Eckhorn et al., 1988; Gray et al., 1989). In this view, synchrony is generated by means of lateral coupling between neurons (possibly in different cortical areas) that all participate in the representation of one object (Sompolinsky et al., 1990; König and Schillen, 1991). The discussion of this idea is lively and summarized in a recent Special Issue of Neuron (vol. 24, 1999). At present, it is not clear whether synchrony has a role, if any, in the ‘binding’ of features, as the mechanism discussed in this paragraph is commonly referred to, or in selective attention, or in a combination of both. It should be noted that the latter was suggested very early on by Treisman and Gelade (1980) and remains a strong possibility.

The details of attentional control are not known at present although most models incorporate at least some degree of competition between simultaneously present stimuli (Desimone and Duncan, 1995). One framework, at least for bottom-up attention, is based on the proposed existence of a saliency map (Koch and Ullman, 1985), a topographic representation of the sensory space in which a winner-take-all mechanism selects the (unique) attended location as the only active location. Control of attention, in this framework, is reduced to selecting that location which is active at a given time. Detailed computational models have been developed for the underlying bottom-up selection process (Niebur and Koch, 1996; Itti et al., 1998; Parkhurst et al., 2002). No detailed mechanism was proposed how activity in the saliency map translates into attentional control but a very natural mechanism would require nothing but a roughly topographical projection from the saliency map to early cortices in the pattern-processing pathway (e.g. V1 in vision or S1 in somesthesia), onto all neurons representing the attended stimulus. Under the premise of activity at the attended location only, these neurons, and only these, would receive additional input from the saliency map and therefore, tend to fire synchronously. The mechanism underlying the observed synchrony is common input from the structures controlling which stimuli are selected.

The third question, the effect of neuronal synchrony, has been discussed in great detail in the literature. Johnson (1980) discussed the effect of correlated activity in sensory discrimination tasks and showed that, in general, correlation can either enhance or suppress discrimination. Specifically for the representation of selective attention, I have discussed in Section 3 that the physical properties and the attentional state of a perceived stimulus may be represented in the same information channel. There is a multitude of biophysical mechanisms whose very existence suggests that a pure rate code—i.e. a representation in which information is entirely coded by the number of spikes within a certain time interval—is unlikely to describe biological information processing completely. Such mechanisms include the interactions between spikes in different spike trains that were discussed above but also interactions within the same spike train, e.g. paired-pulse facilitation (Thomson et al., 1993; Usrey et al., 1998) and paired-pulse depression (Davies et al., 1990; Nelson, 1991; Thomson et al., 1993; Rozov and Burnashev, 1999; Matveev and Wang, 2000; Tsoodyks and Markram, 1997). This short review is not the place to discuss the complexities of this field in any detail, so I will proceed with the commonly made simplification that physiological situations will be such that correlation increases postsynaptic impact (but see Bernander et al., 1994, for a more sophisticated analysis). This would have the desired effect that attended stimuli will be enhanced.
over non-attended stimuli. A mechanism based on increased postsynaptic efficacy of synchronous over asynchronous activity has received much experimental support both in vitro (e.g. Euler and Denk, 2001) and in vivo (e.g. Usrey et al., 2000). The increased synaptic efficacy of synchronous activity was demonstrated particularly convincingly by Roy and Alloway (2001) who showed that thalamocortical efficacy after synchronous thalamic activity was nearly twice as large as the efficacy rate obtained when pairs of thalamic neurons discharged asynchronously. Their results indicate that presynaptic synchrony has a significant impact on cortical responsiveness and suggest that neuronal synchronization may play a critical role in the transmission of sensory information from one brain region to another.

For the sake of simplicity, I have so far assumed that physical properties of sensory stimuli are represented by firing rates. It is well-known that this is not necessarily the case. While there is excellent evidence that many neuronal representations are based at least partially on rate codes, it is unlikely that they are based entirely on them. For instance, synchrony has been shown to play a decisive role in the representation of olfactory stimuli in insects. Stopfer et al. (1997) showed that locust olfactory discrimination is impaired if synchrony between neurons is abolished, even if the firing rates are unchanged. Although similar spatiotemporal patterns are observed in the olfactory bulb and piriform cortex in vertebrates, an analogous experiment has not been performed in this system. In vertebrates it is known, however, that there are a number of instances in which physical properties of stimuli are represented in a temporal code, not in a rate code. This has been observed both in somesthesia (Mountcastle et al., 1990) and in audition (Frisina, 2001), although it appears that the more central neural representations of these stimuli employ a rate code. A perceived quality that is defined by its temporal characteristics is tactile flutter, known to be represented in the central nervous system by a temporal code (Mountcastle et al., 1969). Salinas et al. (2000) showed in primary somatosensory cortical areas of primates that attention to the flutter stimulus significantly tightens the locking of individual spikes to the tactile stimuli.

8. Conclusions

Evidence in both sensory (visual and somatosensory) and motor parts of the brain is accumulating that the attentional state of the organism is correlated with synchronous activity of neural populations. Since the mean activity is known to be used in the representation of stimuli, a mixed code is proposed here in which both a temporal code and a rate code is employed to represent all properties (both physical and those internal to the organism) of a stimulus. This leads to the prediction that both the temporal structure of spike trains and their mean rate should change in systematic ways when the stimulus properties (either internal or external) are changed, although, of course, not necessarily together. These predictions are amenable to experimental testing.

Acknowledgements

This work was supported by the NSF through a CAREER grant to the author.

References


Murthy, V.N., Fetz, E.E., 1996b. Synchronization of neurons during local field potential oscillations in sensorimotor


