Neuron, Vol. 24, 49–65, September, 1999, Copyright ©1999 by Cell Press

Neuronal Synchrony: A Versatile Code for the Definition of Relations?

Review

Wolf Singer*
Max-Planck-Institute for Brain Research
Deutschordenstrasse 46
60528 Frankfurt
Federal Republic of Germany

Most of our knowledge about the functional organization of neuronal systems is based on the analysis of the firing patterns of individual neurons that have been recorded one by one in succession. This approach permits assessment of event-related variations in discharge rate, but it precludes detection of any covariations in the amplitude or timing of distributed responses if these covariations result from internal neuronal interactions rather than from time locking to stimulus or motor events.

As it is likely that internal coordination of distributed responses is functionally as relevant as stimulus-induced coordination, multielectrode recordings are increasingly being used to analyze internally generated covariations of firing patterns. More than a decade ago, we used this method to reveal that neurons in the visual cortex tend to synchronize their discharges with a precision in the millisecond range when activated with a single contour (Gray and Singer, 1987, Soc. Neurosci., abstract; 1989), whereas they fail to do so when activated by different contours moving in different directions (Gray et al., 1989; Engel et al., 1991c). In addition, these stimulus-induced, context-dependent synchronization phenomena were found to be associated with a conspicuous oscillatory modulation of cell firing in a frequency range between 30 and 50 Hz, the so-called γ frequency range. Two aspects make this synchronization interesting. First, it results from internal coordination of spike timing and is not simply caused by stimulus-locked changes in discharge rate. Second, synchronization probability changes in a systematic way when the perceptual coherence of stimulus constellations is modified. Thus, this type of synchrony is not a trivial reflection of anatomical connectivity such as shared input through bifurcating axons, but instead results from context-dependent, dynamic interactions within the cortical network.

The evidence for an internal coordination of spike timing raises the question of whether it serves a function in cortical processing or whether it is merely an epiphenomenon. The goal of this paper is to review theoretical arguments and data that relate to this issue. The first part examines how well the nervous system, and in particular the cerebral cortex, can distinguish between synchronous and asynchronous responses, and whether any significance is attributed to precisely synchronized discharge patterns when these are coordinated by external events, e.g., by the synchronous onset of sensory stimuli. In the second part, data are reviewed from experiments that were designed to examine putative functions of internally generated synchronization. As the assessment of internally generated, non-stimulus-locked temporal relations requires the joint evaluation of responses from more than one neuron, only experiments that permit simultaneous measurements of responses from multiple units are considered. These include multielectrode recordings from multiple individual cells, but also measurements of local field potentials (LFPs) and electroencephalographic (EEG) or magnetoencephalographic (MEG) recordings. The signals of these latter methods reflect the average activity of large cell populations. Because this activity leads to measurable signal fluctuations only if it is sufficiently synchronized, these global recordings provide valuable information about the temporal relations between responses.

*E-mail: singer@mpih-frankfurt.mpg.de.

Two Complementary Binding Strategies

Discussions about the putative functional role of synchrony focus on the question of whether it can serve as a mechanism to bind distributed neuronal activity. To provide an adequate background for the examination of this question, I shall first deal with some general, implementation-independent aspects of binding operations.

As the Gestalt psychologists pointed out, our cognitive systems have the tendency to interpret objects and events as related if they are contiguous in space or time, or if they exhibit similarities in certain feature domains. Thus, contours that touch one another, have similar contrast, or move with the same speed in the same direction (common fate) are more likely to be perceived as components of the same object than spatially distant contours or contours that have no features in common. Likewise, events that coincide in time are interpreted with greater probability as related than events separated in time. At early stages of sensory processing, spatial relations and relations in feature space are represented by the amplitude and the topological relations of activation foci in ordered maps. Temporal relations, however, are represented by the relative timing of responses. In order to accomplish perceptual grouping, the distributed responses of feature-selective cells need to be bound together at some stage of processing. Evidence indicates that this is achieved in two complementary ways. One strategy is binding of responses by convergence of axonal projections. Axons of cells whose responses should be bound are made to converge onto a common target cell at the next-higher processing level. If the threshold of this binding unit is appropriately adjusted, its response signals the specific conjunction of features to which the feeding cells are tuned. We shall address this grouping strategy as "binding by convergence" or "binding by conjunction cells." This coding principle is also known as "labeled line coding" because the responses of a given unit have a fixed label attached to them; they always signal the same conjunction of input signals. The complementary strategy for response binding relies on dynamic selection and grouping of responses. Here, responses are bound by jointly enhancing their saliency relative to other, nonbound responses. Enhanced responses have a stronger impact on downstream processes than nonenhanced responses and

therefore dominate subsequent computations. Thus, the results of these computations will reflect the specific configuration of features to which cells with enhanced responses are tuned. We shall address this selection and grouping strategy as "dynamic binding" and the associated coding principle as "relational coding" or "assembly coding," because here the information about a particular conjunction is contained in the dynamically adjustable configuration of the enhanced responses of distributed neurons (for reviews of the extensive literature on labeled line and assembly coding, see Singer and Gray, 1995, and other contributions in this issue of *Neuron*).

The topology of connections in cortical networks reflects these two grouping strategies and accounts well for perceptual grouping according to contiguity in euclidian and feature space. The first strategy is implemented by the highly complex recombination of feedforward cortico-cortical connections. It leads to a large variety of conjunction-specific neurons, the complexity of which increases as one proceeds along the processing hierarchy. The second strategy appears to rely on two other classes of cortico-cortical connections (reviewed by Singer, 1995; Phillips and Singer, 1997): (1) reciprocal connections that link cells situated within the same cortical area, as well as cells distributed across different areas but occupying the same level in the processing hierarchy; and (2) feedback connections that reciprocate the feedforward connections. Together, these reciprocal cortico-cortical connections constitute the large majority of synaptic inputs to cortical cells. For the intraareal connections in primary visual cortex, it is established that they preferentially couple neurons that are nearby or that code for similar features (Gilbert and Wiesel, 1989; Schmidt et al., 1997a, 1997b). Hence, if these neurons are coactivated, they are more likely to have the saliency of their responses enhanced jointly by cooperative interactions than are neurons that are far apart or tuned to very dissimilar features. As a consequence, responses to contours that are contiguous in euclidian and/or feature space have an enhanced probability of being processed jointly and, thus, of being bound together.

Grouping According to Temporal Cues

In contrast to the numerous experimental and theoretical studies devoted to the analysis of grouping operations in the domain of spatial features, comparatively few studies have been devoted to the questions of (1) to what extent temporal contiguity of stimuli is exploited for perceptual grouping and (2) through which neuronal mechanisms such grouping could be achieved. Most of the evidence regarding perceptual grouping by temporal cues comes from recent psychophysical studies on vision. These studies indicate that spatially distributed contour elements are bound perceptually and interpreted as elements of a coherent figure if they appear or change synchronously, while elements that follow different time courses are perceived as unrelated (Leonards et al., 1996; Alais et al., 1998; Usher and Donnelly, 1998; Lee and Blake, 1999). The temporal resolution of this grouping mechanism is surprisingly high. Temporal offsets between the respective appearances of figure and ground elements of $<$10 ms still support perceptual grouping (Leonards et al., 1996). Because responses of neurons in the visual cortex follow the time course of the stimuli that evoke them, the results of these studies suggest that synchronous responses are bound perceptually while asynchronous responses are interpreted as unrelated.

It is noteworthy that the short offsets that support this segmentation are not perceptible, suggesting a dissociation between the perceptibility of small differences in the time course of stimuli on the one hand and the effect of such differences on perceptual grouping on the other. This dissociation may have to do with the fact that, in vision, the temporal cues supporting grouping as a function of stimulus synchrony are mediated mainly by the magnocellular pathway, while the other, nontemporal grouping cues are mediated by both the magno- and the parvocellular pathway (Leonards and Singer, 1998). These two pathways interact at multiple levels but subserve somewhat different functions. The magnocellular pathway is exquisitely sensitive to temporal features and can signal stimulus transients with high temporal resolution, while the parvocellular pathway operates with low temporal but high spatial and spectral resolution (for a review of the extensive literature, see Leonards and Singer, 1998). This functional dichotomy is relevant in the present context because it is a likely basis for the ability of the visual system to use both temporal and spatial cues in parallel for perceptual grouping. If within the same matrix of line elements one figure is defined by the synchronous onset of elements (temporal cue), and another spatially overlapping figure by differences in the orientation of the respective line elements (spatial cue), either the temporally or the spatially defined figure is perceived depending on the relative saliency of the two cues (temporal offset versus orientation difference) (Leonards and Singer, 1998). This suggests that spatial and temporal grouping cues are processed in parallel and, if they conflict, the less salient cue is disregarded.

This ability of the visual system to rely on either spatial or temporal cues, if the two cues are in conflict, is of considerable functional relevance. On the one hand, it permits binding of nontemporal features that are related but attached to temporally dispersed elements. On the other hand, it allows segregation of features that are unrelated but attached to temporally contiguous elements. The ability to base perceptual grouping on either spatial or temporal cues is also likely the cause of an apparent conflict between the psychophysical studies that support grouping based on temporal cues (see above) and a study that denies such a mechanism. Kiper et al. (1996) found that perception of figures defined by spatial cues is not impaired if false temporal conjunctions are introduced at random by presenting selected elements of the figure synchronously with elements of the background. Here, the temporal cues did not define a figure and hence may have been simply discarded through competition.

In summary, the results of psychophysical studies suggest the following conclusions. First, information about the temporal parameters of stimuli is transmitted over several processing stages with a precision in the

millisecond range. Second, asynchronies among spatially distributed responses of <10 ms are exploitable for perceptual grouping. Third, the mechanism that evaluates temporal relations among responses for perceptual grouping interprets synchronous responses as related and segregates them from responses that are temporally offset. Fourth, temporal and nontemporal grouping cues are evaluated in parallel, the former being conveyed mainly by the magnocellular pathway.

These findings raise the question of how temporal grouping cues are evaluated at the neuronal level. In analogy to grouping mechanisms for nontemporal features (see above), it would suffice that the synchronous responses to simultaneously appearing or simultaneously changing stimuli are more salient, i.e., have a stronger joint impact on cells at subsequent processing stages than the asynchronous responses to temporally dispersed stimuli. Two nonexclusive scenarios may be considered. First, synchronously active cells might cooperate particularly effectively through cortico-cortical connections and thereby increase their discharge rate. Second, synchronous responses might by themselves and without further amplification have a stronger impact on cells at subsequent processing stages than cells responding to temporally offset stimuli.

In both cases, two prerequisites need to be fulfilled. First, timing of discharges must be preserved across polysynaptic transmission chains with a precision in the millisecond range. Second, neurons must be able to differentiate between synchronous and asynchronous input. Synchronous excitatory postsynaptic potentials (EPSPs) must be more efficient than temporally dispersed EPSPs, and dispersions of <10 ms must already make a significant difference.

Temporal Precision in Neuronal Transmission

Contrary to what one should expect from the long time constants of synaptic integration in central neurons (see, e.g., Shadlen and Newsome, 1994), cortical networks can operate with amazing temporal precision. In the auditory cortex of mammals, the spiking patterns of single-cell responses to species-specific calls reproduce with millisecond precision from trial to trial (DeCharms et al., 1998; Kilgard and Merzenich, 1998). Comparable accuracy is found in song birds for auditory neurons responding to songs and for central motor neurons controlling the vocalization patterns (Yu and Margoliash, 1996; Doupe, 1997). In cat visual cortex, neurons faithfully follow flicker frequencies of up to 50 Hz and on occasion even up to 100 Hz (Rager and Singer, 1998). Highly synchronous oscillatory discharges of retinal responses that reach oscillation frequencies of up to 100 Hz are also transmitted reliably from the retina to primary visual cortex (Neuenschwander and Singer, 1996; Castelo-Branco et al., 1998a; Herculano et al., 1999; Neuenschwander et al., 1999) (Figure 1). Even neurons in the medial temporal cortex (MT/V5) of macaque monkeys, which are at least four synaptic stages away from the retina, signal the time structure of temporally modulated visual stimuli with a precision in the millisecond range (Buracas et al., 1998). Further indication of high temporal fidelity in neuronal transmission comes from evidence that neurons distributed within and across cortical areas, and even across the cerebral hemispheres and subcortical structures, can synchronize their spike discharges on the basis of oscillations in the γ frequency range (reviewed by Singer and Gray, 1995; for more recent findings, see Livingstone, 1996; Brecht et al., 1998; Friedman-Hill et al., 1999; Maldonado et al., 1999). This implies that dispersion of spike timing must have remained below the duration of a half cycle, i.e., below about 10 ms for cortical interactions and below 5 ms for retino-cortical transmission. The conspicuous spike patterns described by Prut et al. (1998) in the prefrontal cortex point in the same direction.

As proposed by Abeles (1991) and recently again by Shadlen and Newsome (1998), one possible way to achieve such high temporal precision in neuronal signaling despite "slow" neurons is synchronization of discharges across parallel channels, a special form of population coding. In the proposed models, this is achieved by cross-coupling parallel channels through diverging and converging axon collaterals. As demonstrated recently (Aertsen et al., 1996; Diesmann et al., 1997), such synfire chains (Abeles, 1991) have the interesting property that the synchronization of discharges across parallel channels does not decrease from one synaptic level to the next but may even increase if coupling is appropriately adjusted. The reason for this preservation of precision is that synchronized EPSP barrages are more effective in triggering postsynaptic spikes than temporally dispersed inputs (see below). The synchronous EPSP barrages generated in such synfire chains elicit postsynaptic spikes with minimal latency jitter and hence can transmit the temporal signature of stimuli with high precision over many synaptic stages. Note that what matters for this temporal precision in transmission is the rise time of the compound EPSPs rather than the passive membrane time constant of the integrating neurons.

Synchronization Enhances Saliency of Responses

If synchronized responses are grouped because they are more salient than nonsynchronized responses, neurons evaluating temporal grouping cues must respond differently to precisely synchronized and temporally dispersed barrages of EPSPs, and—as suggested by psychophysics—dispersions of <10 ms must be detectable. Again, at first glance, the long time constants of neuronal membranes seem incompatible with such differential sensitivity to coincident and dispersed input, but experimental observations suggest the contrary. In hippocampal cultures, most of the spontaneously occurring spikes are triggered by synchronously arriving EPSPs rather than by the smaller and more numerous temporally dispersed EPSPs, suggesting a privileged role of synchronized activity in synaptic transmission (Stevens and Zador, 1998). The same conclusion is suggested by in vivo intracellular recordings from pyramidal cells of the monkey motor cortex (Matsumura et al., 1996). Likewise, simultaneous recordings from coupled neuron triplets along thalamo-cortical (Alonso et al., 1996; Usrey and Reid, 1999) and intracortical pathways (Alonso and Martinez, 1998) in the visual system have revealed that EPSPs synchronized within intervals below 2 ms are more effective than EPSPs dispersed over longer intervals.

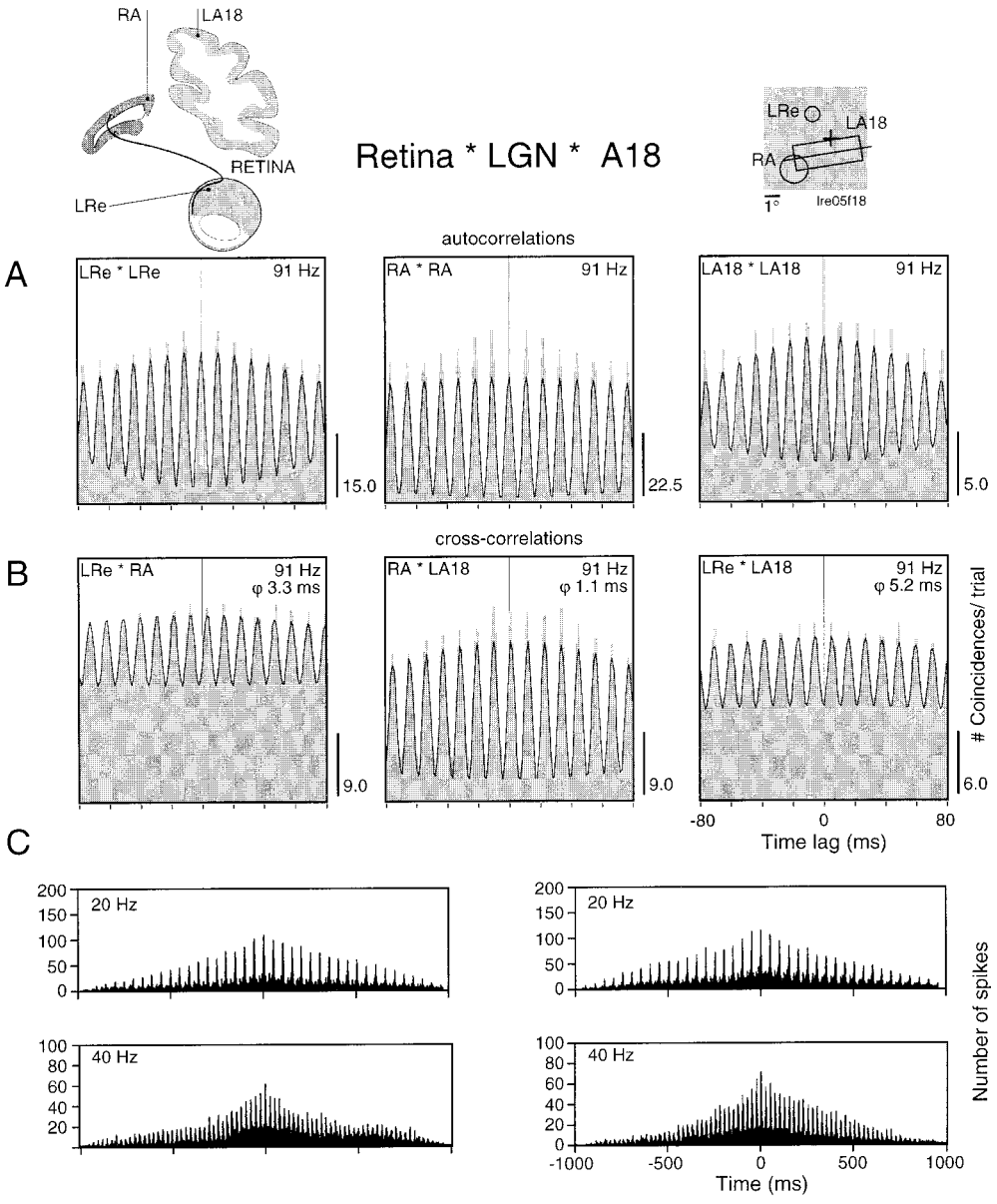

Figure 1. Precise Transmission of Temporal Signatures from the Retina to the Visual Cortex

(A and B) Synchronization between the retina, the lateral geniculate nucleus (LGN), and the cortex of oscillatory responses evoked by a stationary stimulus (top right, inset). Responses were recorded simultaneously from the left retina (LRe), right LGN lamina A (RA), and left area 18 (LA18; top left, inset).

(A) Autocorrelation functions. The onset of the stimulus evokes strong oscillatory patterning at all sites, at a frequency of 91 Hz.

(B) Cross-correlation functions. Responses are correlated between all recorded pairs.

(C) Correlated responses in area 17 following high-frequency flicker stimulation. Responses are recorded from two different sites in area 17. Responses are evoked by flicker stimuli at 20 Hz (top row) and 40 Hz (bottom row). Left columns show correlograms from simultaneously recorded responses; right columns show correlograms computed across nonoverlapping response epochs. Comparison of the original and the shifted correlograms indicates that the precise correlations among responses of distributed neurons are due to precise stimulus locking of the responses ([A] and [B] are modified from Castelo-Branco et al., 1998a, and [C] is modified from Rager and Singer, 1998).

Evidence on a more global level for the enhanced saliency of synchronized activity has been obtained with simultaneous recordings from several sites (area 18 and the posterior mediolateral suprasylvian sulcus [PMLS]) in cat visual cortex and retinotopically corresponding sites in the superior colliculus (Brecht et al., 1998). The impact of a particular group of cortical cells on target cells in the colliculus increased dramatically when the cortical cells' discharge was synchronized with other cortical cell groups projecting to the same collicular site (Figure 2). Enhanced saliency of synchronized responses can also be inferred from the tight correlation between perception and the strength of neuronal response synchronization observed in experiments on binocular rivalry in cats (Fries et al., 1997a) and human subjects (Tononi and Edelman, 1998; Tononi et al., 1998)

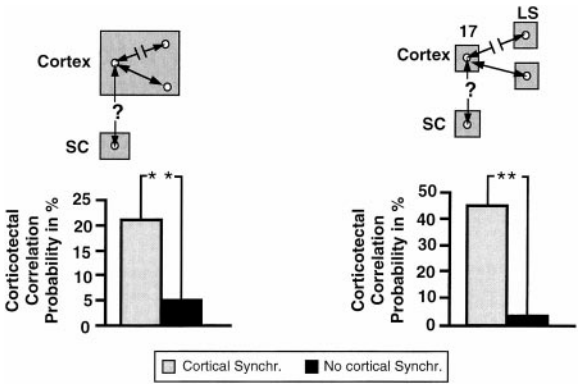

Figure 2. Dependence of Corticotectal Synchronization on Intracortical Synchronization within and across Cortical Areas

(Left) Percentage of significant cortico-tectal correlations if cortical cells recorded from two sites in the same cortical area synchronize (gray column) or do not synchronize (black column) their responses. (Right) Percentage of cortico-tectal interactions between A17 and the tectum if area 17 cells synchronize (gray column) or do not synchronize (black column) their responses with those of cells in the lateral suprasylvian sulcus (modified from Brecht et al., 1998).

(see also below). Finally, simulation studies also indicate that neurons with conventional integrate-and-fire properties can be quite sensitive to the temporal dispersion of synaptic input. Large-scale simulations of biologically inspired thalamocortical networks revealed that neurons exhibited a strong tendency to engage in oscillatory firing patterns and to synchronize their responses. When synchrony was artificially disrupted by introducing a jitter in spike timing, transmission across polysynaptic pathways was drastically reduced (Lumer et al., 1997a, 1997b).

There are several mechanisms, some of which have been identified only recently, that make synchronously arriving EPSPs more efficient than temporally dispersed EPSPs. First, because of their exponential decay, simultaneous EPSPs summate more effectively than temporally dispersed EPSPs, and there is some evidence for supralinear summation due to voltage-gated dendritic conductances. Second, firing threshold is sensitive to the rising slope of the depolarization and lowers for fast-rising depolarizations (C. M. Gray, personal communication). Third, the effect of EPSPs is dramatically enhanced when these coincide with a back-propagating dendritic spike and hence with the input that generated this spike (Larkum et al., 1999). All three mechanisms are sensitive to dispersions of EPSPs in the range of a few milliseconds.

In conclusion, both psychophysical and physiological evidence indicates that neuronal networks are exquisitely sensitive to temporal relations among discharges in input connections, assigning particular significance to coincident, i.e., synchronous input. Synchronicity serves as a tag of relatedness most likely because it causes an increase in the saliency of the synchronized responses, which in turn favors their joint evaluation (binding) at subsequent processing stages. Conditions are thus comparable to those in which figure elements have higher contrast than do elements of the background: in that case, too, responses become grouped according to saliency. The only difference is that, in the case of enhanced contrast, saliency is increased because of higher discharge rates rather than synchronization.

A Role for Internally Generated Synchrony

Internally generated response synchrony closely resembles that induced by synchronously presented stimuli with respect to both its temporal precision and its magnitude (Rager and Singer, 1998). This raises the question of whether the internally synchronized discharges affect processing in the same way as externally induced synchrony. If so, internal synchronization could serve to bind distributed activity according to grouping criteria set by the brain itself, and the results of internal grouping processes could be evaluated by the same neuronal mechanisms that have evolved for the evaluation of externally induced timing relations. Responses synchronized by internal interactions would undergo joint enhancement of their saliency and be treated as related by subsequent processing stages. Due to the high temporal precision with which cortical networks can distinguish synchronous from nonsynchronous events (see above), internal synchronization could serve to define relations between distributed responses with high temporal resolution, and could thus ideally complement selection and grouping operations based on sustained enhancement of discharge rate. In principle, synchronization can enhance the impact of particular response constellations on the basis of individual discharges. If, in addition, internally generated synchronization patterns can change on a fast time scale, different relations could be defined in rapid succession for successive segments of sustained, temporally overlapping responses of distributed neurons (Figure 3). Thus, if synchrony of discharges served as a signature of relatedness not only when induced from outside but also when generated internally, it could be utilized in all processes where flexible and context-dependent selection and grouping of responses are required on a fast time scale.

The Nature and Detectability of Internally Generated Synchrony

Synchronization is considered to be of internal origin if two neurons exhibit a statistically significant covariation in firing probability that cannot be attributed to stimulus-locked covariations in discharge rate. Such episodes of synchronous firing can occur spontaneously, e.g., during the various sleep stages (Steriade, 1999); they can appear during responses to sensory stimuli; or they can be associated with cognitive processes such as focusing attention, analyzing complex patterns, storing contents in short term memory, and preparing movements (see below). These internally generated covariations of firing probability manifest themselves at different time scales ranging from a few milliseconds to hundreds of milliseconds, as indicated by the variable widths of the peaks in cross-correlograms. Often, synchronization is associated with an oscillatory patterning of the discharges, the frequency of these oscillations covering a broad range and exhibiting a marked state dependence. Typically, synchronized EEG states such as those that occur during drowsiness, deep sleep, and

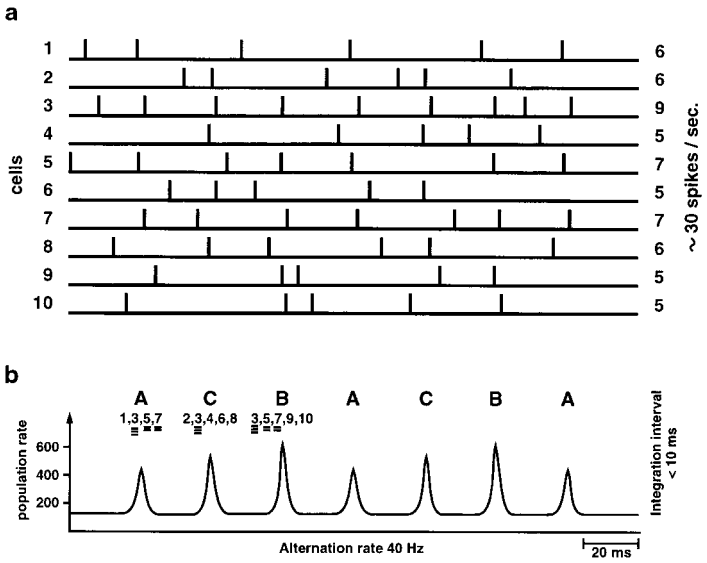

Figure 3. Disambiguation of Temporally Overlapping Assemblies by Synchronization of Discharges in Selected Response Segments

Despite the sustained and constant discharge rate (~30 Hz) of the ten depicted cells, the saliency of responses is transiently and repeatedly increased for three different, spatially overlapping assemblies (A–C) by synchronization of the discharges of the subpopulations of cells constituting the three assemblies. Note that the individual spike trains appear nonoscillatory, while the spike density of the population response fluctuates periodically (lower continuous line). Note also that cells 5 and 7 are shared by two and cell 3 is shared by all three assemblies. Because cell 3 contributes spikes to more assemblies than the others, its discharge frequency is slightly higher. In this case, the different assemblies are interleaved with a rate of 40 Hz, and the integration interval for the evaluation of the population response is assumed to be around 10 ms.

anesthesia are associated with broad correlation peaks and, if present, low oscillation frequencies (2–10 Hz); activated, desynchronized EEG states that are characteristic of the awake, performing brain but also of paradoxical sleep are associated with sharp correlation peaks (<15 ms) and high oscillation frequencies (30–60 Hz). In this state, synchronization exhibits maximal topological specificity and can occur over particularly large distances (Munk et al., 1996; Herculano et al., 1999).

Relations between synchronization and changes in discharge rate are variable. When synchronization occurs on a coarse time scale—i.e., when correlation peaks are broader than the average interspike interval of the respective cells' discharges—synchronization can be considered to be the result of internally produced covariations in discharge rate. However, when synchronization occurs on a fast time scale with correlation peaks narrower than the average interspike intervals, it can be entirely independent of discharge rate. In this case, changes in spike timing that leave average discharge rate unaffected can lead to drastic changes in synchrony (see Figure 3 and König et al., 1996). Most of the experimental results reviewed below show such independence. Note, however, that one could determine discharge rates on a spike-by-spike basis, defining instant frequencies as the inverse of the respective interspike intervals. In that case, some of the synchronized events might be interpreted as the result of coherent, ultra-rapid rate fluctuations; but even then relations between rate terms and synchrony can break down, because synchrony can be achieved by advancing a spike in one cell (rate increase) and delaying a spike in the other (rate decrease).

This potential independence of synchrony from rate fluctuations is one reason why internally generated synchrony can only be assessed by direct correlation analysis of simultaneously recorded neurons and not by comparing rate fluctuations of successively recorded cells. Another reason is that the internal coordination increases the covariance of rate changes above the level expected from comparison of successively recorded event-related rate changes. Thus, even the coarser correlation patterns that can be described in terms of coherent rate fluctuations can only be assessed from correlations of simultaneously recorded responses if coherence is due to internal interactions. An example of such a case is the recent observation that the latencies of stimulus-induced rate increases of simultaneously recorded cortical neurons covary due to internal coordination to a much higher degree than expected from the latency scatter of successively recorded responses (Fries et al., 1998, Eur. J. Neurosci., abstract). This coordination in response timing could not have been disclosed by studying the respective neurons one by one. Finally, it is to be expected that there are processes in the brain that are based on self-paced coordination of both the timing and the amplitude of distributed neuronal activity and are only loosely, if at all, time locked to externally measurable events. These hidden but potentially important covariations of spike timing will also be detectable only by simultaneous recordings. Hence, internally generated covariations in both the timing and the amplitude of neuronal activity can only be disclosed with multicell recordings.

Why Is There a Need for Dynamic Grouping of Responses?

There are numerous instances in cortical processing where dynamic selection and grouping of responses for further joint processing are required. Such is the case whenever conjunctions have to be signaled for which there are no specific conjunction units. But there is a deeper argument that suggests that dynamic grouping is indispensable at all processing levels, even despite the implementation of highly complex conjunction units. This argument is derived from facts about the way cortical networks encode features in general. It appears that cortex exploits the option of coarse coding in order to economize on neuron numbers. The essence of this coding strategy is that contents are not encoded by the responses of individual sharply tuned neurons but by

specific constellations of graded responses of numerous broadly tuned neurons. The advantage is that the number of different activity patterns that can be configured from a given set of neurons is much larger than the number of neurons in the set. Hence, coarse coding is parsimonious with respect to neuronal numbers because a given neuron can participate in many different representations when active in various constellations. The constraint is, however, that neurons participating in coarse coding must be broadly tuned. This, in turn, has a price. Coarse codes may become ambiguous when stimuli that overlap in euclidian or feature space, or in both, need to be represented simultaneously (see also Gray, 1999; von der Malsburg, 1999 [both in this issue of *Neuron*]). Because of the broad tuning of cells participating in coarse coding, place codes alone may not be sufficient to permit subsequent processing stages to detect which of the many distributed responses ought to be associated with which stimulus. Since coarse coding appears to characterize all levels of processing, this superposition problem is of a very general nature. Strategies to solve it can therefore be expected to be similar across the various stages of the cortical processing hierarchy.

To illustrate, here is a concrete example of this general problem. Due to broad tuning (coarse coding) of neurons in primary visual cortex for position, orientation, length, width, contrast, and a few other features, a single elongated contour evokes graded responses in a very large number of neurons. Because the response amplitude of any one of these neurons is influenced by variations of the stimulus along any of these feature dimensions, individual responses convey little specific information. The precise configuration of the stimulus can only be deduced if a large number of the graded responses evoked by the stimulus are evaluated jointly. If only a single contour is present on a homogeneous background, this poses no difficulty. Superposition problems arise, however, if several nearby or spatially overlapping contours are embedded in a textured background. In that case, joint evaluation of responses needs to be restricted to responses evoked by the same contours.

A classical assumption is that there are neurons at subsequent processing levels that receive convergent input in various constellations from subsets of broadly tuned low-level neurons, and that thereby acquire selectivity for only the particular constellation of features that characterizes one of the objects (binding by convergence). In principle, population codes could be fully disambiguated by such conjunction-specific binding units; but this solution is very expensive in terms of the number of required binding units if all possible conjunctions were to be represented in this explicit way. One would require as many binding units as there are distinguishable population states. This is clearly not an attractive strategy because it sacrifices the main advantage of coarse coding—the parsimonious use of neurons.

It has been proposed, therefore, that the cerebral cortex uses two coding strategies in parallel: first, partial disambiguation of coarse population codes by the implementation of conjunction units, and second, disambiguation of population codes by dynamic, context-dependent binding. Note that in both cases the essence of the process is the representation of relations. In the first case, this is achieved through binding by convergence; in the second case, through dynamic binding (Singer, 1995; 1999a; Phillips and Singer, 1997). Stereotyped, frequently occurring, and particularly behaviorally relevant conjunctions would be represented by specific binding units because this strategy is faster and less susceptible to binding errors. However, because conjunction-specific neurons cannot exhaust the full combinatorial space and cannot represent unanticipated conjunctions, they should be recruitable into dynamically configured population codes that represent meta-conjunctions for which there are no binding units. Thus, even conjunction units ought to be broadly tuned, in order to allow them to participate in coarse coding. The idea is that this dual coding strategy is iterated throughout the processing hierarchy, whereby some feature conjunctions get represented explicitly by individual neurons and others by population codes. Which conjunctions are represented explicitly by the binding units at the various processing stages can be deduced from the rate-coded response properties of the respective neurons. Evidence indicates that these conjunctions become increasingly complex, abstract, and multimodal as one proceeds along the processing hierarchy, and that at the interface between sensory and motor systems their responses shift progressively from stimulus-related to movement-related features.

At each processing level, novel, unanticipated conjunctions or conjunctions for which specific binding units have not (yet) been implemented can then be represented by additional population codes. Thus, despite the implementation of ever more sophisticated binding units, flexible and context-dependent grouping (binding) operations remain necessary at each processing level in order to represent relations that are not encoded by conjunction-specific neurons and in order to resolve superposition problems if they occur.

In addition to their flexibility and virtually inexhaustible coding capacity, dynamically bound population codes have the further advantage that computational results can be represented at all stages of processing by population codes that have the same format. The only changes concern the nature and complexity of the feature conjunctions that are represented explicitly by the neurons forming the respective populations. Thus, distributed representations can be mapped directly onto other distributed representations, which should facilitate polymodal integration and sensory–motor coordination, thus circumventing bottleneck problems.

In sensory–motor coordination, a vast number of different associations must be realized between sensory representations and motor programs (Roelfsema et al., 1996). As the latter are in all likelihood distributed in nature, it would be quite uneconomical to first convert distributed sensory representations into single-cell codes, to then relay their output onto command neurons that represent a particular movement, and then to recreate a distributed code from the output of these cells. Direct conversion of one distributed code into another appears as the more elegant solution, but again, it requires dynamic binding.

In conclusion, if relations are encoded both by conjunction-specific neurons and by dynamically associated assemblies of such neurons, rapid and flexible

grouping operations have to be accomplished at all levels of cortical processing. Note, however, that the two strategies, although available at all stages, need not always be used together. One can imagine conditions where the elementary features of a scene are all explicitly represented by conjunction-specific units, and where the arrangement of contours poses no superposition problem, but where the objects defined by these features are novel and are not covered by conjunction-specific cells of a higher order. In this case, dynamic grouping may only be necessary at higher levels of processing. Conversely, a highly familiar object could be explicitly represented by a set of object-specific binding units—but this object could be embedded in a complex background allowing for many unfamiliar conjunctions between elements of the figure and elements of the background. In that case, extensive grouping operations would be required at peripheral levels of processing to associate the correct set of elementary conjunction units, but no dynamic grouping would be necessary at higher levels provided there is no superposition problem at these levels. This possibility needs to be considered when designing experiments in search of binding mechanisms. It should be noted that not every task requires dynamic binding at all stages of processing. In some studies on figure–ground discrimination, for example, animals are not forced to identify a particular figure and may respond if they simply detect an inhomogeneity in a pattern. In such cases, it is likely that they opt for the fastest strategy and rely on response changes of some conjunction neurons, rather than go through the time-consuming process of dynamic feature binding.

Dynamic Grouping Mechanisms

One proposal is that mechanisms of selective attention may subserve dynamic binding of distributed responses (see Ghose and Maunsell, 1999; Shadlen and Movshon, 1999; Reynolds and Desimone, 1999; Wolfe, 1999 [all in this issue of *Neuron*]). In vision it is assumed, for example, that object-centered attention selectively enhances the discharge rate of neurons responding to features of the same contour or object, via top-down projections, and that the resulting joint increase of saliency leads to joint processing of the selected responses at subsequent stages (see, e.g., Treisman, 1996). This mechanism encounters several difficulties if not supplemented by additional grouping operations. First, it presupposes that higher centers "know" which of the peripheral responses code for the contours of a particular object. However, this information is only available once an object is identified, and this is often possible only after the population codes at lower levels have been disambiguated. Thus, it must first be clarified whether the many simultaneous responses in a low-level area code for coherent or disjunct contours before it can be decided whether one has to do with one or another object. Second, because these top-down projections fan out over large cortical domains, it is difficult to see how the projections could enhance the responses of the cells coding for one out of several overlapping contours with the required topological specificity (Salin and Bullier, 1995).

These problems can be alleviated if responses within the respective processing stages undergo a first, preliminary grouping according to criteria that are adapted to the grouping problems encountered at each stage. In primary visual cortex, these would be rather simple criteria because neurons at this stage code for simple features. All that needs to be accomplished here is the grouping of responses that are likely to be generated by the same contour. This would be the case for responses evoked by a continuous contour, by contour segments that are collinear, or by contours that share similarities in any of the simple feature domains for which neurons in primary visual cortex are responsive. In areas specialized for the analysis of other features—such as, for example, motion—grouping would have to occur according to motion parameters and so on. All these grouping operations could in principle occur in parallel in the various cortical areas. Since all relevant areas are reciprocally coupled, preliminary grouping results could themselves bias the grouping operations in other areas until a solution is found that is most consistent with the set of low- and high-level grouping criteria that reside in the architecture of the various visual areas. Such dynamic, distributed grouping through iterative reentry has been proposed as one of the core functions of cortical networks (Edelman, 1987, 1989). Analogous to the proposed mechanism of attentional grouping, these distributed grouping operations could be based on joint enhancement of the firing rate of the selected neurons.

However, grouping of responses solely by joint rate enhancement may engender problems. First, it can lead to ambiguities. It may not always be easy to distinguish whether rate increases are due to grouping or to variations in stimulus properties, such as, for example, contrast changes by uneven illumination. Second, in cases where objects overlap in euclidian or feature space, only responses to a single object can be grouped at any one moment. Otherwise, it would be again unclear which of the selected responses belong to which population code. Because evaluation of nonsynchronized rate changes requires integration of a minimal number of EPSPs arriving successively from the selected cells, the pace at which different populations can be defined by multiplexing is slow.

Both problems could be alleviated by introducing internal synchronization as an additional grouping mechanism. First, synchronization can bias the saliency of responses independently of rate fluctuations. Second, because it relies on coincidence detection rather than temporal summation, synchronization can define relations with sufficiently high temporal precision to permit multiplexing.

Complementarity of Rate Modulation and Synchronization

Because grouping of responses by synchronization is bound to enhance discharge rates of selected target cells in areas receiving synchronized input, fast synchronization codes and more sustained rate codes could ultimately coexist and perhaps even optimally complement one another. Sustained rate-coded input from a given processing stage could be rapidly disambiguated and bound through synchronization at other processing stages; this grouped activity would in turn lead to a specific pattern of sustained, rate-modulated responses at the following stage where these activity patterns can

again be subject to disambiguation by selective synchronization and so on. Once grouping operations have converged, the respective solutions may have to be stabilized. If groups need to be kept separate, they could each engage in sustained synchronized oscillations with high intra- and little intergroup synchrony, but if there are no superposition problems left, rate enhancement alone would be sufficient. As discussed below, there is evidence for such coexistence of synchronization and rate codes.

Predictions

If internally generated synchronization is to serve as a signature of relatedness in the same way as stimulus-locked synchronization appears to (see above), it needs to meet several criteria. First, to be compatible with known processing speed, synchronization must be achieved rapidly within maximally a few tens of milliseconds (Rolls and Tovee, 1994; Thorpe et al., 1996). Second, changes in synchrony and changes in discharge rate should be independently adjustable; i.e., there should be cases where synchrony among two neurons increases (or decreases) without a concomitant increase (or decrease) in discharge rate. Third, internal synchronization must be sufficiently precise so that synchronous activity is more effective than nonsynchronized activity in driving cells in target structures. Fourth, there must be relations between the occurrence, the dynamics, and the topological distribution of synchronization patterns on the one hand and specific perceptual or motor processes on the other hand. These relations must be sufficiently consistent to permit predictions of behavior from measurements of synchronization patterns. Finally, the connections responsible for the internal generation of synchrony must be susceptible to use-dependent modifications of synaptic gain, so that synchronization probability can be increased for groups of cells that have often been synchronously active in the past. This is required to install grouping criteria by learning and to form stable assemblies of neurons for the representation of perceptual objects in memory. The coincidence-detecting mechanism that mediates these use-dependent gain changes must operate with the same temporal precision as the synchronizing mechanism. In the following paragraphs, experimental data addressing these predictions are reviewed.

The Precision of Internal Synchrony

Internally generated synchronization can be as precise as externally induced synchronization, in particular when the global EEG is in an activated desynchronized state, as is characteristic for the aroused attentive brain. In that case, the widths of the correlation peaks at half height are typically in the range of $<$10 ms (reviewed by Singer et al., 1997), and synchronization is often associated with an oscillatory patterning of the respective responses in the high β or the γ frequency range (from 30 to 60 Hz) (Munk et al., 1996; Herculano et al., 1999). This periodic patterning reduces the probability of spurious correlations and enhances substantially the precision with which the discharges of different neurons are synchronized (Maldonado et al., 1999). To date, few computational models of the neocortex have taken these characteristic features of population dynamics into consideration. This and the usually conservative assumptions about time constants of dendritic integration may be reasons for the persisting skepticism concerning the functional significance of synchronization.

In this context, it is also important to note that analysis of single-cell responses often fails to reveal that a cell participates in a synchronously oscillating assembly. The reason is that individual cells usually do not fire with every cycle. Interspike interval distributions may appear Poissonian, even though the cell's discharges are precisely time locked to an oscillatory process and synchronized with other cells in the population. Such cycle skipping is a well-known phenomenon of oscillatory processes in the hippocampus (Buzsaki and Chrobak, 1995; Buzsaki, 1996) and may be of functional significance rather than a reflection of noise fluctuations. If cycles are skipped in a systematic way, the constellations of cells discharging in synchrony can change from cycle to cycle. In the case of 40 Hz oscillations, different contents could then be encoded in time slices following one another at intervals of 25 ms (Jensen and Lisman, 1998; see also Figure 3). Indications that such a coding strategy may actually be used have been obtained in the olfactory system of insects (Wehr and Laurent, 1996). As discussed above, cortical networks should be able to operate with the required temporal resolution, because otherwise they would not be able to maintain synchrony at 40 Hz to begin with.

Rapid Synchronization

Synchronization can be established very rapidly. Simulations revealed that networks of reciprocally coupled spiking neurons can undergo very rapid transitions from uncorrelated to synchronized states (Bauer and Pawelzik, 1993; Deppisch et al., 1993; Gerstner and van Hemmen, 1993; van Vreeswijk et al., 1994; Hopfield and Hertz, 1995; Gerstner et al., 1996). This agrees with the observation that neurons in the visual system engage in synchronous activity, often with additional oscillatory patterning, at the very same time they increase their discharge rate in response to the light stimulus (Eckhorn et al., 1988; Gray et al., 1992; Neuenschwander and Singer, 1996; Castelo-Branco et al., 1998a).

A cortical mechanism that can achieve such nearly instantaneous synchronization has been recently identified. It exploits two properties: first, the ability of oscillating cells to delay their output relative to incoming EPSPs and, second, the oscillatory patterning of ongoing cortical activity. When the membrane potential of a cell undergoes an oscillatory modulation, EPSPs with an NMDA receptor–mediated component evoke spikes not necessarily at the time of their occurrence but only when the cell reaches the peak of the next depolarizing cycle. The reason is that NMDA receptors still occupied by glutamate are reactivated by the cyclic depolarizations that remove the voltage-dependent magnesium block (Volgushev et al., 1998). With such a mechanism, responses to temporally dispersed EPSPs can become synchronized within less than an oscillation cycle by appropriate shifting of spike latencies.

Recent correlation studies have shown that fluctuations in response latency to stationary flashed stimuli

can covary with a precision in the millisecond range for subsets of neurons located in different columns of the visual cortex and even for neurons in different hemispheres (Fries et al., 1997b, Soc. Neurosci., abstract). This correlation between the very first discharges of a response appears to be based on the delay mechanism identified in the slice experiments. Comparison between actual response latencies and immediately preceding fluctuations of the local field potential revealed that the response latency shifted as a function of the polarity of the preceding field potential fluctuation. This effect was confined to episodes in which the field potential oscillations exhibited high power in the γ frequency range, emphasizing the role of γ oscillations in the coordination of response timing. In all likelihood, the spatiotemporal patterns of these fluctuations reflect the architecture and the actual functional state of intra- and interareal association connections (Arieli et al., 1996). Thus, grouping by synchronization can be extremely fast and still occur as a function of both the prewired associational dispositions and the current functional state of the cortical network.

External versus Internal Synchronization

Latency shifting by internal synchronization could compromise precise signaling of temporal stimulus features. Thus, responses that convey information about the precise timing of stimuli and that support grouping through externally imposed synchrony should be exempted from internal latency adjustments. In the visual system, the magno- and parvocellular pathways are kept separate over the first few processing stages, and grouping cues provided by the two systems appear to be processed independently and in parallel (see above). This supports the possibility that internal synchronization mechanisms act differentially on magno- and parvocellular pathways and do not affect the former at those levels of processing where grouping is achieved according to external timing and where precisely timed motor responses are programmed. Since all sensory systems have developed parallel pathways with differential sensitivity to temporal (phasic) and nontemporal (sustained) stimulus features, it may be a general strategy to use external, stimulus-induced and internal, self-generated timing relations in parallel for the grouping of responses. The advantage would be that ultimately the relatedness of temporal and nontemporal stimulus features could be expressed in the same format—namely, in the degree of synchrony—and this should facilitate joint evaluation of temporal and nontemporal grouping cues at processing stages where both stimulus attributes have to be bound together. However, to the best of our knowledge, this possibility has not yet been investigated.

Relations between Response Synchronization and Gestalt Rules for Perceptual Grouping

Based on the hypothesis that internal synchronization of discharges could serve to group responses for joint processing, a series of experiments has been performed in the search for a correlation between dynamic changes of response synchronization and particular stimulus configurations. One prediction tested was that in early visual areas, synchronization probability should reflect some of the basic Gestalt criteria according to which the visual system groups related features during scene segmentation. A consistent finding was that neurons distributed across different columns within the same or different visual areas, and even across hemispheres, synchronized their responses with near-zero phase lag when activated with a single contour but fired independently when stimulated simultaneously with two different contours (Gray et al., 1989; Engel et al., 1991a, 1991b, 1991c; Freiwald et al., 1995; Kreiter and Singer, 1996). This suggested that synchronization was the result of a context-dependent selection and grouping process. Analysis of the dependence of synchrony on receptive field and stimulus configurations revealed that the probability and strength of response synchronization reflected elementary Gestalt criteria for perceptual grouping such as continuity, proximity, similarity in the orientation domain, collinearity, and common fate (reviewed by Singer et al., 1997; Gray, 1999 [this issue of *Neuron*]). These early experiments were performed in anesthetized animals, but more recent multielectrode recordings from awake cats and monkeys indicate that these synchronization phenomena are not artifacts of anesthesia but are even more pronounced when the animals are awake and attentive (Kreiter and Singer, 1992, 1996; Frien et al., 1994; Fries et al., 1997a; Gray and Viana Di Prisco, 1997; Friedman-Hill et al., 1999; Maldonado et al., 1999) (see Figure 4). In none of these experiments have systematic changes in synchronization probability been associated with systematic changes of the neurons' discharge rate.

A particularly close correlation between neuronal synchrony and perceptual grouping has recently been observed in experiments with plaid stimuli. These stimuli are well suited for the study of dynamic binding mechanisms because minor changes of the stimulus cause a binary switch in perceptual grouping. Two superimposed gratings moving in different directions (plaid stimuli) may be perceived either as two surfaces, one being transparent and sliding on top of the other (component motion), or as a single surface, consisting of crossed bars, that moves in a direction intermediate to the component vectors (pattern motion) (Adelson and Movshon, 1982; Stoner et al., 1990). Which percept dominates depends on the luminance of grating intersections, because this variable defines the degree of transparency (Albright and Stoner, 1995). Component (or pattern) motion is perceived when luminance conditions are compatible (or incompatible) with transparency (Figure 5A). Thus, this is a case in which local changes in stimulus properties cause global changes in perceptual grouping. In the case of component motion, responses evoked by the two gratings must be segregated, and only responses evoked by the contours of the same grating must be grouped to represent one of the two surfaces; in the case of pattern motion, responses to all contours must be bound together to represent a single surface. If this grouping of responses is initiated by selective synchronization, three predictions must hold (see Figure 5B). First, neurons that prefer the direction of motion of one of the two gratings and have collinearly aligned receptive fields should always synchronize their responses because they respond always to contours that belong to the same surface. Second, two neurons that

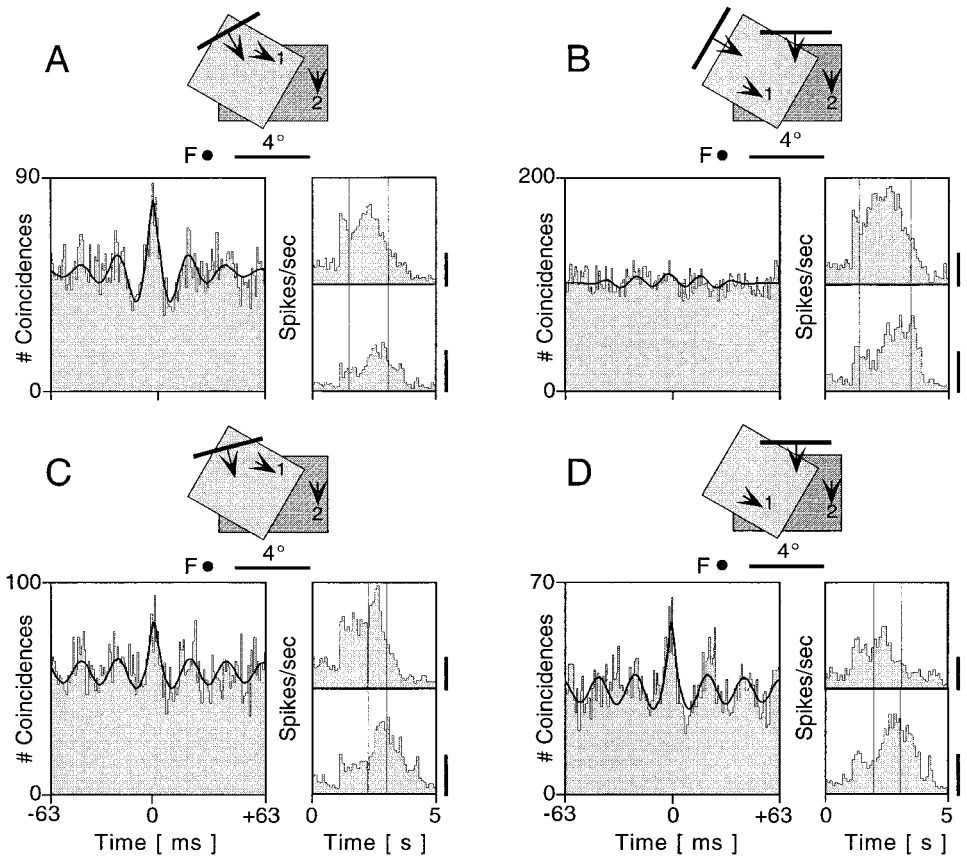

Figure 4. Stimulus Dependence of Neuronal Synchronization in Area MT of the Visual Cortex of a Macaque Monkey Carrying Out a Fixation Task

Neuronal responses were obtained from two cell groups with different directional preferences. The figure shows cross-correlograms and peri-stimulus time histograms for four different stimulation conditions. The small insets indicate the receptive field locations (1 and 2) with respect to the fixation point (F) and the directional preference of the neurons (small arrows).

(A) A single moving stimulus bar, whose direction of motion was intermediate between the neurons' preferences, led to a pronounced synchronization of the two cell groups, as indicated by the central maximum in the cross-correlogram.

(B) Presentation of two stimuli moving in the respective preferred directions of cell groups 1 and 2 abolishes synchronization.

(C and D) The synchronization observed with a single stimulus does not depend on its particular orientation.

(C) Changing orientation and direction of motion by 15° or (D) using one of the bars from the configuration in (B) had little influence on synchronization. Scale bars for the peri-stimulus time histograms correspond to 40 spikes/s. The continuous line superimposed on the correlograms represents a damped cosine function that was fitted to the data to assess the significance of the correlogram modulation (modified from Kreiter and Singer, 1996).

are tuned to the motion directions of the two gratings should synchronize their responses in the case of pattern motion because they then respond to contours of the same surface, but they should not synchronize in the case of component motion because their responses are then evoked by contours belonging to different surfaces. Third, neurons preferring the direction of pattern motion should also synchronize only in the pattern and not in the component motion condition.

An important aspect of these predictions is that the expected changes in synchrony differ for different cell pairs, depending on the configuration of their receptive fields. Thus, when searching for relations between synchrony and cognitive functions, it is not only crucial to identify the processing stage where one assumes a particular binding function to be accomplished but also to select the appropriate cell pairs. Averaging data across cell pairs with different receptive field configurations can mask dynamic changes in synchrony and is likely to reveal only the static anisotropies in the network of synchronizing connections. Such a problem may have contributed to the negative results of a recent study that failed to show a relation between perceptual grouping and internal synchronization in monkey striate cortex (Lamme and Spekreijse, 1999).

In the case of the plaid stimuli, predictions were tested with multielectrode recordings from areas 18 and PMLS of the visual cortex of lightly anesthetized cats, after we had confirmed with eye movement recordings in awake cats that the animals distinguished between component and pattern motion. Cross-correlation analysis of responses from cell pairs distributed either within or across areas 18 and PMLS confirmed all three predictions. Cells synchronized their activity if they responded to contours that are perceived as belonging to the same surface (Castelo-Branco et al., 1998b, Eur. J. Neurosci., abstract) (Figure 5C). Analysis of the neurons' discharge rate confirmed that most of the cells in these visual

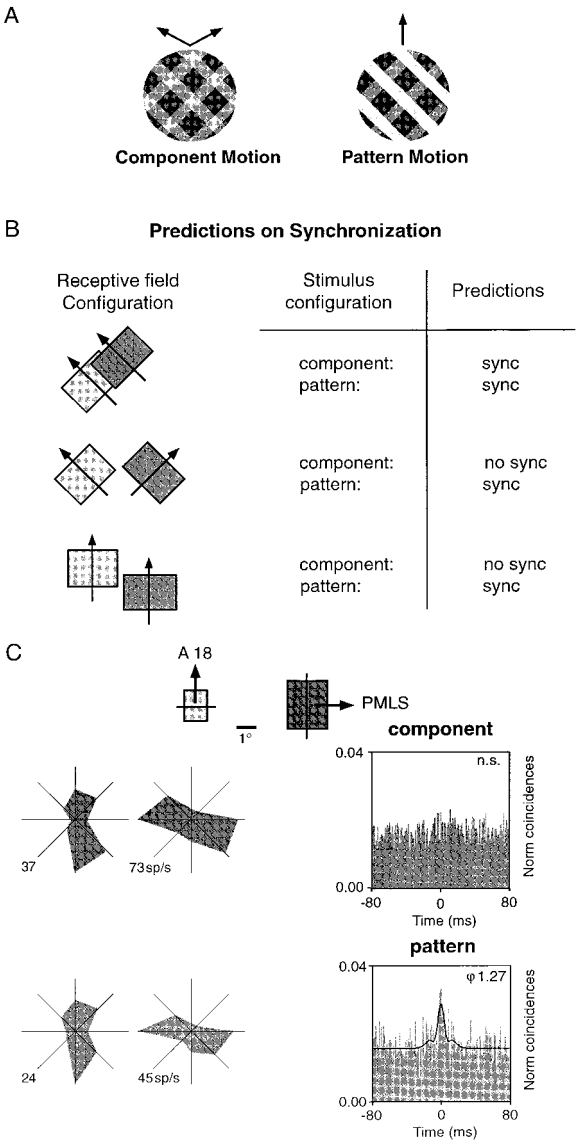

Figure 5. Neuronal Sychronization and the Bistable Perception of Plaids

(A) Two superimposed gratings that differ in orientation and drift in different directions are perceived either as two independently moving gratings (component motion) or as a single pattern drifting in the intermediate direction (pattern motion), depending on whether the luminance conditions at the intersections are compatible with transparency.

(B) Predictions on the synchronization behavior of neurons as a function of their receptive field configuration (left) and stimulation conditions (right).

(C) Changes in synchronization behavior of two neurons recorded simultaneously from areas 18 and PMLS that were activated with a plaid stimulus under component (top graph) and pattern motion (bottom graph) conditions. The two neurons preferred gratings with orthogonal orientation (see receptive field configuration [top] and tuning curves obtained with component and pattern, respectively) and synchronized their responses only when activated with the pattern stimulus (compare cross-correlograms on the right) (courtesy of M. Castelo-Branco and S. Neuenschwander).

areas respond preferentially to the component gratings of the plaids (component-specific cells; Gizzi et al., 1990) and not to the pattern as a whole. However, in contrast to synchrony, variations in response amplitude failed to reflect the transition from component to pattern motion induced by transparency manipulation. Dynamic changes in synchronization could, thus, serve to encode in a context-dependent way the relations among the simultaneous responses to spatially superimposed contours and thereby bias their association with distinct surfaces. Future investigations have to clarify whether the populations of differentially synchronized neurons already serve as the final representations of the perceived surfaces or whether, in the case of pattern motion, additional assemblies are formed. These would then have to consist of conjunction units tuned to the specific constellations of superimposed gratings, and their responses would have to be bound together to signal that they code for the same surface.

In all experiments referred to in this review, trials with different stimulus configurations were interleaved randomly to counterbalance possible effects of slow drifts in central state that may cause covariations of discharge rate at a slow time scale. Moreover, phasic response components were excluded from correlation analysis to avoid artifactual correlations due to dynamic changes in discharge rate (Brody, 1999a, 1999b). Accordingly, the shift predictors (correlograms between responses to the same stimulus configuration selected from different trials) were always flat, indicating that the changes in the correlations were caused by context-dependent changes of internal interactions and not by changes in stimulus-locked or state-dependent covariations of discharge rate.

The Anatomical Substrate

Studies involving lesions (Engel et al., 1991a) and developmental manipulations (Löwel and Singer, 1992; König et al., 1993) indicate that the interactions responsible for these stimulus-specific synchronization phenomena are mediated at least in part by cortico-cortical connections that reciprocally link cells in the same cortical area, as well as cells distributed across different areas, and even across the two hemispheres. The elementary grouping criteria applied in early vision could thus reside in the architecture of these association connections. The evidence that intrinsic excitatory connections preferentially link neurons which code for features that tend to be grouped is consistent with this possibility (Ts'o and Gilbert, 1988; Gilbert and Wiesel, 1989; Malach et al., 1993; Schmidt et al., 1997a, 1997b).

Temporal Constraints of Use-Dependent Plasticity

If synchronization of responses serves as a mechanism to group responses, synchronizing connections must be susceptible to use-dependent modifications. This is required in order to implement new grouping criteria by learning and to stabilize assemblies representing previously experienced conjunctions. Synchronization probability must remain increased for groups of cells that have been forced previously to engage repeatedly in highly synchronous firing, and the mechanism mediating these use-dependent changes in synchronization probability must be capable of distinguishing between synchronous and nonsynchronous firing with a temporal resolution that is in the same range as the temporal precision of observed synchrony. Both postulates have

recently been confirmed. Herculano et al. (1997, Soc. Neurosci., abstract) have shown that synchronization probability increases between groups of neurons if these engage repeatedly in synchronous oscillatory firing in the γ frequency range. This enhanced tendency of neurons to synchronize their responses to coherent stimuli can be reduced again by having the same cells engage repeatedly in oscillatory firing patterns that are decorrelated. These increases and decreases of synchronization probability appear to depend upon the precise phase relations of the oscillatory discharges during conditioning, because on a coarser time scale ($>$50 ms) the neurons' activity overlapped completely in both conditioning paradigms.

Recent data from cortical slices support such a dependency of synaptic plasticity on precise timing between pre- and postsynaptic discharges. Varying the temporal relations between presynaptic and postsynaptic responses in simultaneously recorded coupled cortical cells revealed that long-term potentiation (LTP) results when the EPSP precedes the postsynaptic spike within intervals of 10 ms or less, while the polarity of the modification reverses to long-term depression (LTD) as soon as the EPSP follows the spike (Markram et al., 1997; Zhang et al., 1998). Thus, shifts of a few milliseconds in the timing relations between pre- and postsynaptic discharges suffice to invert the polarity of use-dependent synaptic modifications. The mechanism permitting such precise evaluation of the temporal contiguity of pre- and postsynaptic responses is, with all likelihood, the active dendritic response associated with the back-propagating spike (Magee and Johnston, 1997).

Under the sustained activation conditions applied in the in vivo experiments of Herculano et al. (1997, Soc. Neurosci., abstract), the high temporal selectivity of synaptic modifications resulted most likely from the oscillatory modulation of the neuronal responses. That an oscillatory patterning of activity can effectively narrow the temporal window for coincidence detection has been shown in vitro. Experiments in slices of the hippocampus indicate that the polarity of synaptic gain changes depends critically on the phase relations between pre- and postsynaptic discharges if these overlap on a coarse time scale but exhibit an oscillatory modulation (Huerta and Lisman, 1996). The same holds for the visual cortex. Pyramidal cells of rat visual cortex slices were made to discharge tonically at 20 Hz by injecting sinusoidally modulated current through a patch pipette. Simultaneously, EPSPs were evoked, also at 20 Hz, by electrical stimulation of excitatory afferents. Changing the phase relations between pre- and postsynaptic activity revealed that the stimulated input tended to undergo LTP when the EPSPs were coincident with the spikes, while afferents consistently underwent LTD when the EPSPs fell in the troughs of the membrane potential oscillations. Thus, although pre- and postsynaptic activation overlapped completely on a coarse time scale, phase shifts of $<$20 ms between individual EPSPs and spikes reversed the polarity of the synaptic modifications (Wespatat et al., 1999).

In conclusion, the synchronization mechanism identified at early levels of visual processing shares essential features with the processes postulated for the dynamic binding of responses. First, synchronization is precise and rapidly established when the brain is in an activated state. Second, synchronization is selective and context dependent, reflecting some of the elementary Gestalt criteria that determine perceptual grouping. Third, synchrony is due to intracortical interactions that are based on the network of reciprocal association connections. Fourth, the adaptive mechanisms supporting use-dependent modifications of synchronization probability operate with the required temporal resolution.

Relations between Synchronization and Perception

While the experiments reviewed above had the goal to identify basic characteristics of response synchronization, the studies described in the following paragraphs were aimed at establishing links between synchronization phenomena and behavior. A close relation between response synchronization and perception has been found in cats who suffered from strabismic amblyopia, a developmental impairment of vision, which results in suppression of the amblyopic eye, reduced visual acuity, and crowding. Crowding refers to the inability of amblyopic subjects to identify target stimuli if these are surrounded by nearby contours and is thought to result from false binding of responses evoked by the target and the embedding background. Quite unexpectedly, the light responses of individual neurons in the primary visual cortex of amblyopic cats were normal, and the neurons continued to respond vigorously to stimuli that the animals were unable to perceive through the amblyopic eye. The only significant correlate of amblyopia detected in this study was a reduction of response synchronization among neurons driven by the amblyopic eye. This reduction in synchrony became particularly pronounced when responses were evoked with gratings that had been identified in previous behavioral experiments as too fine to be resolvable by the animal through the amblyopic eye (Roelfsema et al., 1994). Thus, impaired synchrony is likely to account for at least some of the perceptual deficits: by reducing the saliency of responses, it could explain why signals from the amblyopic eye cannot compete successfully with the well-synchronized responses from the normal eye and are excluded from supporting perception when both eyes are open. Recent recordings from area 21, a higher visual area, have indeed shown that the poorly synchronized responses evoked from the amblyopic eye drive neurons in area 21 less efficiently than the well-synchronized responses from the normal eye (Schröder et al., 1998, Eur. J. Neurosci., abstract). Poor synchronization could also be responsible for the reduction in visual acuity and for crowding, because it is expected to impair disambiguation of responses evoked by closely apposed stimuli.

A correlation between response synchronization and perception has also been documented in experiments on binocular rivalry that were again performed in strabismic animals (Fries et al., 1997a). Perception in strabismic subjects always alternates between the two eyes. This can be exploited to investigate how neuronal responses to constant stimuli change if they pass from being selected and perceived to being suppressed and excluded from perception and vice versa (Figure 6). The outcome of these experiments was surprising, because the responses of neurons in areas 17 and 18 were not attenuated when they were excluded from supporting perception. A close and highly significant correlation existed,

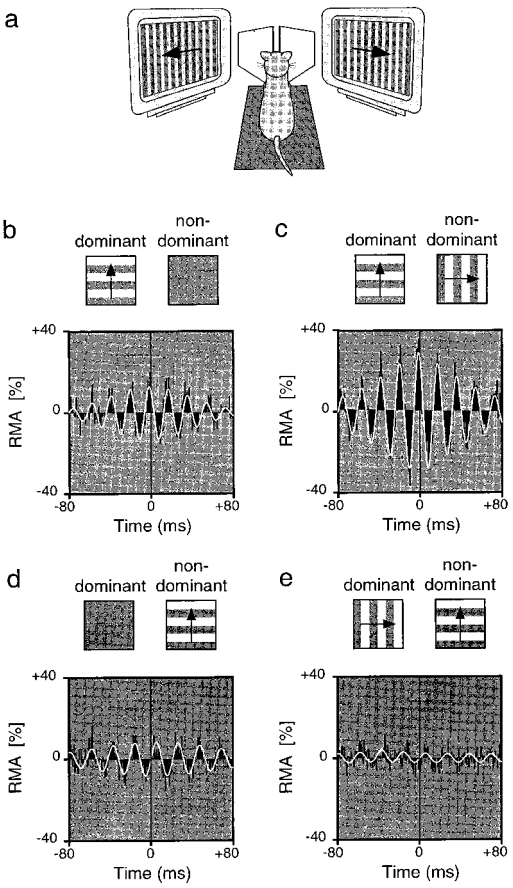

Figure 6. Neuronal Synchronization under Conditions of Binocular Rivalry

(A) Using two mirrors, different patterns were presented to the two eyes of strabismic cats. (B) through (E) show normalized cross-correlograms for two pairs of recording sites activated by the eye that won (B and C) and lost (D and E) in interocular competition, respectively. Insets above the correlograms indicate stimulation conditions. Under monocular stimulation (B), cells driven by the winning eye show a significant correlation, which is enhanced after introduction of the rivalrous stimulus to the other eye (C). The reverse is the case for cells driven by the losing eye (compare conditions [D] and [E]). The white continuous line superimposed on the correlograms represents a damped cosine function fitted to the data. RMA is the relative modulation amplitude of the center peak in the correlogram, computed as the ratio of peak amplitude over the offset of correlogram modulation. This measure reflects the strength of synchrony (modified from Fries et al., 1997a).

however, between changes in the strength of response synchronization and the outcome of rivalry. Cells mediating responses of the eye that won in interocular competition increased the synchronization of their responses upon presentation of the rivalrous stimulus to the other, losing eye, while the reverse was true for cells driven by the eye that became suppressed. In addition, there was a marked increase in the oscillatory modulation of the synchronized discharges that supported perception. This agrees with the hypothesis (see above) that the outcome of a selection process can be stabilized by having the selected group of neurons engage in stable, synchronized oscillatory response patterns. Thus, in primary visual cortex, there are instances where internal selection of responses for further processing is associated with enhanced synchronization rather than increased firing. This agrees with rivalry experiments in awake, behaving monkeys, which showed no systematic relation between the strength of visual responses and perception in early visual areas but a clear correlation between perceptual suppression and loss of neuronal responses in higher visual areas (Logothetis and Schall, 1989b; Leopold and Logothetis, 1996; Sheinberg and Logothetis, 1997). This is what one expects if the saliency of the responses from the two eyes is adjusted at early processing stages by modulating synchronization rather than discharge rates. These results, of course, do not exclude additional response selection by rate modulation. Thus, the outcome of rivalry can also be biased by changing the contrast of the patterns presented to the two eyes, and this manifests itself in rate changes.

Global Interareal Synchronization

More global synchronization phenomena, involving task-dependent coordination of the oscillatory activity of whole cortical areas, have been observed in cats trained to perform a visually triggered motor response. The visual, association, somatosensory, and motor areas involved in the execution of the task synchronized their activity in the γ frequency range as soon as the animals focused their attention on the relevant stimulus; the strength of synchronization among areas reflected precisely the coupling of these areas by cortico-cortical connections. Immediately after the appearance of the visual stimulus, synchronization increased further, and these coordinated activation patterns were maintained until the task was completed. However, once the reward was available and the animals engaged in consummatory behavior, these coherent patterns collapsed and gave way to low-frequency oscillatory activity that did not exhibit any consistent relations with regard to phase and areal topology (Roelfsema et al., 1997). These results suggest that an attention-related process had imposed a coherent temporal pattern on the activity of cortical areas required for the execution of the task. As discussed above, coordinated subthreshold modulation of membrane potential fluctuations acts like a temporal filter that permits rapid but selective synchronization (Fries et al., 1997b, Soc. Neurosci., abstract). Here, attentional mechanisms, involving perhaps nonspecific thalamic projections (Ribary et al., 1991; Steriade, 1999), seem to exploit this option to prepare cortical areas for rapid synchronization, thereby accelerating subsequent selection and grouping of responses. If these anticipatory synchronization patterns exhibit columnar specificity, they could not only accelerate but also bias subsequent grouping as a function of expectancy. This then could provide a mechanism for the well-known phenomenon that the outcome of grouping depends also on what one expects to see.

Neuronal Synchrony and Motor Performance

In the motor cortex of primates, oscillatory modulation of activity in the γ frequency range and synchronization of responses both within and across hemispheres have been observed in association with self-paced grasping

movements or visually guided reaching tasks (Murthy and Fetz, 1996a, 1996b). Another study found recurring phases of heightened discharge synchrony among simultaneously recorded single cells in primate motor cortex that were precisely correlated with the animal's expectancy of a "go" signal (Riehle et al., 1997). Remarkably, these epochs of enhanced synchronization were not associated with measurable changes in the cells' discharge rates, further evidence that synchrony and discharge rate can be independent. A close relation between precisely timed discharge patterns and the type of anticipated motor responses has also been described by Prut et al. (1998) for simultaneously recorded neurons in the frontal cortex of monkeys trained to perform visually cued targeting movements. Finally, synchronous oscillatory discharge patterns were observed in motion-sensitive visual areas of the monkey (MT/V5) in association with the preparation of visually guided eye movements. In this case, synchrony broke down, however, upon presentation of the stimulus (De Oliveira et al., 1997). While in these cases synchronization was mainly associated with movement preparation, recent studies in monkeys trained to perform visually guided targeting movements showed also a close correlation between patterns of discharge synchrony and the actual direction of the arm movements (Ojakangas et al., 1997, Soc. Neurosci., abstract; Donoghue et al., 1998; Hatsopoulos et al., 1998). In these recordings, the strength of the correlations varied with the direction of motion and provided additional information about movement parameters.

Studies in Humans

Correlates between the occurrence of synchronized oscillatory activity in the γ frequency range and cognitive processes have recently been observed in human subjects with EEG and MEG recordings. In a series of pioneering EEG studies, evidence has been obtained that the engagement in cognitive tasks requiring figure–ground distinctions and feature binding (Tallon-Baudry et al., 1996, 1997) or perceptual switching during perception of bistable figures (Keil et al., 1999) or the retention of visual patterns in short-term memory (Tallon-Baudry et al., 1998, 1999) is associated with transient increases of γ oscillations in cortical areas involved in the respective tasks. Evidence is also available that oscillatory activity in the γ frequency range is contingent with states of focused attention and with language processing (reviewed by Pulvermüller et al., 1997). These findings have recently been extended by the demonstration that during a face recognition task, populations of neurons do not only synchronize locally on the basis of γ oscillations, but also phase lock across large distances with zero phase lag. A particularly interesting finding was that these large-scale synchronization patterns dissolve and settle into new configurations at exactly the time when subjects had recognized the pattern and prepared the execution of the motor response (Rodriguez et al., 1999).

A close correlation has also been found between the establishment of coherent γ oscillations across cortical areas and the acquisition of a visual–tactile association in classical conditioning. Once subjects have learned about the association between a visual and a tactile stimulus, the visual stimulus evokes γ activity over the visual cortex that is coherent with γ activity over the somatosensory cortex contralateral to the stimulated finger, and this coherence breaks down again upon extinction (Miltner et al., 1999). Finally, evidence has been obtained in an MEG study on binocular rivalry that activity evoked from the eye that dominates perception during rivalry is better synchronized than activity evoked by the suppressed eye (Tononi and Edelman, 1998; Tononi et al., 1998). This agrees with the rivalry experiments in cats (Fries et al., 1997a).

Also consistent with a functional role for synchronized γ oscillations is the observation that these are closely related to arousal and attention. Synchronization of neuronal responses in the γ frequency range does not occur when the EEG exhibits low-frequency oscillations, as is the case in deep anesthesia or slow-wave sleep, but is particularly pronounced during activated "desynchronized" states of the EEG (Munk et al., 1996; Herculano et al., 1999). In human subjects, γ oscillations are enhanced in responses to attended as compared to nonattended stimuli (Tiitinen et al., 1993); they disappear in deep anesthesia (Madler and Pöppel, 1987) but are present during paradoxical sleep, during which dreaming—i.e., reactivation of memories—is thought to occur (Llinas and Ribary, 1993; Steriade, 1999).

Manipulations of Synchrony

The results reviewed so far provide correlative evidence for a role of internally generated response synchronization in neuronal processing, but do not prove that the nervous system attributes significance to the precise temporal relations among discharges if these are generated by intrinsic interactions rather than by external stimuli. To prove a causal relation requires the demonstration that selective manipulation of synchrony alters behavior. Recently, Brecht et al. (1997, Soc. Neurosci., abstract) obtained evidence from microstimulation experiments in the optic tectum of awake cats that the network responsible for the control of eye movements is exquisitely sensitive to the synchronicity of discharges of neurons in the superior colliculus. Microstimulation of tectal neurons with a 50 Hz train of pulses evokes a targeting eye movement whose amplitude and direction (vector) depend on the site of stimulation. If two spatially distant sites are stimulated with two synchronous train stimuli, the vector of the resulting eye movement is the average of the vectors corresponding to the two sites. However, if the two trains are phase shifted by $>$5 ms, the vector of the resulting eye movement switches from the average to the sum of the individual vectors. This indicates that changes in the synchronicity among two cell populations that discharge periodically with constant frequency are detected by downstream networks and are converted into different eye movement programs. Since tectal cell populations get entrained in synchronous oscillations by oscillating cell groups in the cortex (Brecht et al., 1998), this phase sensitivity of the eye movement generator could contribute to the selection of targeting movements in the presence of multiple overlapping objects. Again, this possibility has to be tested with natural stimuli and multielectrode recordings in the awake animal.

So far, the most direct evidence for a functional role of precise timing relations among discharges in parallel channels comes from the insect olfactory system. In a sequence of studies on olfactory coding in locusts and bees, Laurent and coworkers demonstrated, in analogy to similar conditions in the mammalian olfactory bulb (Freeman and Skarda, 1985), that neurons along the olfactory pathway engage in highly synchronous oscillatory activity (around 20 Hz) when stimulated with odors. Further analysis revealed that (1) information about the mixture of odors is contained in the combination of spikes that are precisely synchronized across parallel channels—information that cannot be retrieved by considering only the changes in discharge rate in the respective channels (Laurent, 1996; Laurent et al., 1996; MacLeod and Laurent, 1996; Wehr and Laurent, 1996); (2) odor discrimination deteriorates if these synchronization patterns are disturbed but rate modulation is unaltered (Stopfer et al., 1997); and (3) target cells at higher processing stages lose some of their odor specificity if the responses of their inputs are desynchronized but otherwise unchanged (MacLeod et al., 1998). Interestingly, this oscillatory patterning of responses develops only with repeated exposure to the same mixture of odors, suggesting that it is the result of an active reorganization of network interactions (M. Stopfer, personal communication). These studies provide the first evidence, albeit not for the mammalian brain, that precise synchronization of discharges in conjunction with an oscillatory patterning of responses carries information that is used for the segregation and identification of sensory stimuli, and that the information conveyed by these oscillations goes beyond that contained in the discharge rate of individual neurons.

Conclusions

The data reviewed in this paper indicate that neuronal networks including the cerebral cortex (1) transmit and process temporal signatures of spike trains with a precision in the millisecond range; (2) respond more readily to synchronous than to asynchronous input; (3) treat synchronous discharges as related, at least when synchrony is induced by external stimulation; (4) synchronize responses with a precision in the millisecond range independently of external timing; and (5) possess mechanisms for the use-dependent modification of synaptic efficacy that operate with the same high temporal resolution as the synchronizing mechanism. Together with the numerous correlations between internally generated neuronal synchronization on the one hand, and perceptual grouping rules, cognitive processes, and motor performance on the other, these data are consistent with the hypothesis that neuronal networks exploit the option to encode information about the relatedness of responses by modulating not only the rate of individual, conjunction-specific neurons but also by temporal coordination of distributed responses. By adjusting the temporal relations among discharges of distributed neurons, response segments could be selected with high temporal resolution for further joint processing, thereby establishing temporary bonds in a highly dynamic and flexible way. This possibility is supported by the fact that measurements of internally generated synchrony permit inferences about stimulus configurations, perception, and motor performance that could not be drawn from measurements of discharge rate alone.

As discussed above, grouping of responses through synchronization is likely to influence the discharge rate of neurons receiving input from the synchronously active cells. Thus, the results of dynamic binding operations are to a certain extent reflected by rate changes at higher levels of the processing hierarchy. However, because there is no single cortical area that can be considered the final stage of processing, and because coarse codes and distributed representations are common at all levels, the rate variations of individual cells can only reflect partial solutions that most often require further specification by dynamic grouping. Since analysis of these supervening grouping operations requires measurements of temporal relations among distributed responses, it seems unavoidable in the search for neuronal codes to complement single-cell analysis of rate changes with multicellular analysis of temporal relations. Because of the distributed nature of cortical processing, this analysis should probably not be restricted to individual, cytoarchitectonically defined areas but may have to be extended at least across all those areas that occupy the same processing level. Moreover, this analysis will eventually have to be based on more sophisticated statistical procedures than pairwise correlations, because the dynamics of neuronal network interactions are likely to evolve in a high-dimensional state space. This space of dynamic relations will have to be explored experimentally in order to find out whether—and, if so, to what extent—cortical networks exploit the option to define relations among distributed activities by generating self-paced, precisely timed covariations in firing patterns.

Experiments exploring the limits of such a dynamic coding strategy will have to be pursued at many different levels, including both in vitro and in vivo studies. In vitro preparations are probably suited best to investigate the temporal limits of synaptic integration and plasticity, to identify the nature of dynamic interactions among reciprocally coupled neurons, and to analyze the mechanisms underlying oscillatory patterning and synchronization. Irrespective of whether the correlation hypothesis resists the test of time, these questions are important in their own right, as little is known to date about the temporal dynamics of neuronal networks.

However, the currently debated question of whether synchronization serves binding can only be approached with multicellular recordings in behaving animals that signal what they perceive or intend to do. In vision, the best tests are probably those that rely on patterns such as plaids or other bistable figures that can be perceived in different ways depending on how their elements get bound, but other tasks requiring the establishment of flexible, temporary conjunctions are suitable as well. One attractive example is storage of ever-changing, novel conjunctions in working memory. In these in vivo studies, it will be crucial to make educated guesses as to which neurons at which processing stage are likely to be involved in solving the binding problem posed by a particular task; and, as in the experiments with plaids, it needs to be considered that neuron pairs with different receptive field constellations may exhibit very different

synchronization behavior, excluding indiscriminate averaging across pairs. Furthermore, as discussed elsewhere (Singer et al., 1997), epochs of synchronized activity may be very brief in behaving animals and difficult to detect. However, this problem may be overcome by more sophisticated techniques than simple cross-correlation, and by designing tasks that require more sustained binding.

The question remains as to what form of evidence we are prepared to accept as conclusive. Most of the evidence for the functional significance of changes in discharge rates in single-cell studies is correlative in nature: if changes in discharge rate correlate with other measurable events, they are commonly considered to be meaningful in the respective context, and if a particular behavior can be predicted from changes in activity, a causal relation is assumed. If the same criteria are accepted for synchrony, then it should be possible to establish evidence for or against the synchronization hypothesis at the same level of confidence as is found acceptable for single cell studies—and some of the more recent correlation studies come close. Much more challenging are attempts to prove causal relations, because this requires manipulations that increase or disrupt synchrony without altering discharge rates. Electrical stimulation has been successfully applied to increase discharge rates (see Shadlen and Movshon, 1999 [this issue of *Neuron*]), but this strategy is inherently problematic because in addition to enhancing discharge rate electrical stimulation also synchronizes the activated cell populations. Still, this strategy seems worth pursuing, and multisite stimulation with phase-adjusted train stimuli is one option. Also, it is to be expected that better understanding of the mechanisms leading to synchronization will eventually provide finer tools to interfere specifically with temporal coordination.

In conclusion, I believe that the theoretical implications of the synchronization hypothesis and the data available to date are of sufficient interest to motivate further examination. The application of the new methods required to test the hypothesis will undoubtedly provide new insights into the dynamics of neuronal interactions. If it then turns out that the hypothesis falls short of the real complexity—which is bound to be the case—we will have learned something about the role of time in neuronal processing that we would not have learned otherwise. Based on current evidence, I consider it highly unlikely that self-paced temporal coordination of distributed activity will prove irrelevant for cortical processing; if it does, we nonetheless shall have made a great step forward, because it is the unexpected result that contains maximal information.

References

A comprehensive reference list for all reviews can be found on pages 111–125.