

Crossmodal binding through neural coherence: implications for multisensory processing

Daniel Senkowski¹, Till R. Schneider¹, John J. Foxe^{2,3} and Andreas K. Engel¹

¹ Department of Neurophysiology and Pathophysiology, University Medical Center Hamburg-Eppendorf, Martinistrasse 52, 20246 Hamburg, Germany

² The Cognitive Neurophysiology Laboratory, Cognitive Neuroscience and Schizophrenia Research Program, Nathan S. Kline Institute for Psychiatric Research, 140 Old Orangeburg Road, Orangeburg, NY 10962, USA

³ Program in Cognitive Neuroscience, Department of Psychology, The City College of the City University of New York, 138th Street and Convent Avenue, New York, NY 10031, USA

Picture yourself on a crowded sidewalk with people milling about. The acoustic and visual signals generated by the crowd provide you with complementary information about their locations and motion which needs to be integrated. It is not well understood how such inputs from different sensory channels are combined into unified perceptual states. Coherence of oscillatory neural signals might be an essential mechanism supporting multisensory perception. Evidence is now emerging which indicates that coupled oscillatory activity might serve to link neural signals across uni- and multisensory regions and to express the degree of crossmodal matching of stimulus-related information. These results argue for a new view on multisensory processing which considers the dynamic interplay of neural populations as a key to crossmodal integration.

Introduction

The inputs delivered by different sensory organs provide us with both complementary and redundant information about the environment. Constantly, multisensory interactions occur in the brain to evaluate whether there is a matching of the information arriving through different channels, or whether the signals give rise to conflict and need to be processed separately. The outcome of these interactions is of critical importance for perception, cognitive processing and the control of action [1–4]. Recent studies have revealed that a vast number of cortical operations, including those carried out by primary regions, are shaped by inputs from multiple sensory modalities [4–6].

In recent years, an increasing number of studies have aimed at characterizing multisensory cortical regions, revealing multisensory processing in the superior temporal sulcus, the intraparietal sulcus and frontal regions, as well as insula and claustrum [4,6,7]. Interestingly, there is increasing evidence that neurons in areas formerly considered as unisensory like, such as auditory belt areas [3,4,6,8,9], can also exhibit multisensory characteristics. Furthermore, numerous subcortical structures are involved in multisensory processing. In addition to

the superior colliculus [1], these include the striatum [10], the cerebellum [11] and the amygdala [12], as well as nuclei of the thalamus [13].

Whereas the ubiquity and fundamental relevance of multisensory processing are becoming increasingly clear, the neural mechanisms underlying crossmodal inter-

Glossary

Electroencephalography (EEG): The standard noninvasive method to measure electrical potential changes arising from the brain. The EEG signal is recorded on the scalp surface and reflects the summed postsynaptic activity in the underlying cortical regions. A key advantage is the high temporal resolution of the method.

Evoked oscillatory activity: Oscillatory neuronal activity that is phase locked to the onset of eliciting events (e.g. sensory inputs). Evoked oscillatory activity can be quantified using spectral analysis if recorded single-trial responses are first averaged and only subsequently spectral power is calculated [19].

Functional connectivity: Functional coupling of neurons that emerges from, but is not exhaustively predicted by, anatomical connectivity. Phase synchrony of oscillatory signals establishes high functional connectivity, because interacting neurons can exert a stronger impact on one another if they are depolarized at the same time [20,23].

Magnetoencephalography (MEG): A noninvasive method for measuring the magnetic field associated with neuronal currents in the brain. The magnetic field is less disturbed by the characteristics of the tissues surrounding the brain than the electric signal components. Although comparable in terms of the temporal resolution, the MEG has a higher spatial resolution than the EEG.

Oscillation: Neurons or neuronal populations often show rhythmic activity patterns. Because they are band-limited processes characterized by a dominant frequency, oscillations can best be described using spectral analysis.

Oscillatory phase: Displacement of a certain feature of an oscillatory waveform (e.g. peak or trough) relative to a reference event such as stimulus onset or the same feature in a different oscillatory signal. Phase can be expressed as the fraction of an oscillatory cycle.

Oscillatory power: Measure for the size of an oscillatory signal, which scales with the square of the signal amplitude.

Phase coherence: Two oscillatory signals, for instance neural oscillations of two brain regions, are considered to be phase coherent when there is a constant relationship between the phases of the two signals over time.

Phase locking: An oscillatory signal is phase locked if a stable phase relationship to an external event (e.g. stimulus onset) is maintained across trials.

Phase resetting: Shift of the phase of an ongoing neural oscillation due to a sensory stimulus, which can lead to phase locking or to increased phase coherence with other oscillatory processes.

Phase synchrony: A special case of phase coherence with zero phase difference between the two signals under consideration.

Total oscillatory activity: Oscillatory neuronal activity that is related, but not necessarily phase locked to the onset of sensory or behavioral events. Total activity comprises both phase-locked and non-phase-locked oscillatory components and is obtained when spectral power is computed at the single-trial level prior to averaging [19].

Corresponding author: Engel, A.K. (ak.engel@uke.de).

actions are much less well understood. In this article, we review recent studies that cast new light on this issue. Whereas classical studies have postulated a feedforward convergence of unimodal signals as the primary mechanism for multisensory integration [1,2], there is now evidence that both feedback and lateral interactions are also relevant [4,6,14,15]. Beyond this changing view on the anatomical substrate of multisensory interactions, there is increasing awareness that complex dynamic interactions of cell populations, leading to coherent oscillatory firing patterns, might play a key role in mediating cross-systems integration in the brain [16–21]. Here we consider the hypothesis that synchronized oscillations (see Glossary) might provide a potential mechanism for crossmodal integration and for the selection of information that matches across different sensory channels.

Views on crossmodal integration

The classical view posits that multisensory integration occurs in a hierarchical manner by progressive convergence of pathways. In this view, sensory signals are merged only in higher association areas and specialized subcortical regions [1,2]. A core assumption of this approach is that perceptual information is primarily encoded in the firing rate of the cells involved. Multisensory integration, accordingly, is expressed by firing rate changes in neural populations receiving convergent inputs from different modalities. Response properties of such neurons are typically characterized by multisensory enhancement or suppression [2], that is, their responses to multisensory stimulation are larger or smaller, respectively, than those to the most effective unisensory stimuli. A large body of evidence demonstrates such multisensory response patterns in a wide set of brain regions [3,6,7].

In recent years, however, numerous authors have suggested that a pure convergence model might not be sufficient to account for all aspects of multisensory processing [4,6,14,15]. First, strong crossmodal interactions and modulation already occur in primary cortices, a result that is difficult to reconcile with the notion of hierarchical convergence. Second, a convergence scenario does not appear flexible enough to allow for rapid recombination of crossmodal signals into completely novel percepts. Third, a feedforward convergence model does not explain how low-level information about objects can remain accessible, because the high-level representation is noncompositional and does not explicitly make reference to elementary features.

A different account of crossmodal integration can be derived from data on the functional role of correlated neural activity, which is likely to play a key role in feature integration and response selection within the various sensory modalities [16–21]. Numerous studies in both animals and humans have shown that synchronized oscillatory activity in various frequency bands is related to a large set of cognitive and sensorimotor functions (Box 1). Based, in particular, on studies of gamma band (>30 Hz) oscillations, the hypothesis has been forwarded that synchronization of neural signals might be a key mechanism for integrating and selecting information in distributed networks. This so-called temporal correlation hypothesis

Box 1. Functional significance of oscillatory activity

Ongoing intrinsic and event-related oscillations are usually categorized into five frequency bands: delta (0.5–3.5 Hz), theta (4–7 Hz), alpha (8–12 Hz), beta (13–30 Hz) and gamma (>30 Hz). A large body of evidence suggests that oscillatory activity in these frequency bands is linked to a broad variety of perceptual, sensorimotor and cognitive operations [16–21,59–61]. Oscillatory activity in the delta band has been related to motivational processes and the brain reward system, and is the predominant brain rhythm during deep sleep [59,62]. Activity in the theta band has been linked to working memory functions, emotional arousal and fear conditioning [62,63]. Alpha band responses have been suggested to reflect cortical operations during the awake resting state in the absence of sensory inputs. More recent theories have proposed that alpha band oscillations might also relate to disengagement of task-irrelevant brain areas [60], as well as working memory function and short-term memory retention [61]. Neuronal responses in the beta band have been frequently linked to sensorimotor network processing [64]. Finally, synchronization in the gamma band has been related to a large number of cortical functions [16–21]. Based on studies in the visual modality, gamma band coherence of neural assemblies has been shown to be relevant for feature integration [65–67], surface segregation [68], perceptual stimulus selection [50,69] and attention [51,70]. Beyond the visual modality, gamma band synchrony has been implicated in auditory [36,71], somatosensory [72] and olfactory [73] processing. Moreover, gamma band synchrony has been shown to relate to sensorimotor integration [74], movement preparation [75] and memory formation [76,77]. Interestingly, phase synchrony and phase modulation of oscillations across the different frequency bands have been recently suggested to play a key role in the organization of networks engaged in complex cognitive functions such as speech processing [78] and memory encoding [61,63].

[17,18] predicts that coherence of neural signals allows highly specific patterns of functional connectivity (see Glossary) to be created, which enables flexible and context-dependent binding, the selection of relevant information and the efficient routing of signals through processing pathways [20,22,23].

The same principle of ‘integration through coherence’ [16–18] might serve to establish specific relationships across different sensory modalities, allowing the crossmodal binding of signals and the preferential routing of matched crossmodal information to downstream assemblies. Because it has been shown that phase coherence (see Glossary) predicts how strongly two interacting groups of neurons mutually influence each other’s response strength [23], this view does not contradict the notion that crossmodal interactions lead to changes of neuronal firing rates, but does shift emphasis toward considering a richer dynamic repertoire of neural interactions and a more flexible scenario for crossmodal communication in the brain.

Oscillatory activity in crossmodal processing

A variety of different paradigms have been used to study the role of oscillatory responses and neural coherence during multisensory processing (Table 1). Most studies have been performed in humans using EEG or MEG (see Glossary), with relatively few animal studies currently available. The approaches used address different aspects of multisensory processing, including (i) bottom-up processing of multisensory stimuli, (ii) crossmodally induced perceptual changes, (iii) modulation by top-down attention and (iv) crossmodal semantic matching. In all these

Table 1. Paradigms for the study of multisensory oscillations

First author	Refs	Method	Modality	Paradigm	Brain region					Analysis			Frequency					Window (ms)
					F	C	LT	OP	O	Tot	Evo	Coh	δ	θ	α	β	γ	
von Stein	[24]	EEG	AV	Passive			x	x				x				x	0–2000	
Kisley	[26]	EEG	AS	Passive		x					x					x	20–150	
Sakowitz	[27]	EEG	AV	Passive	x	x					x			x			50–300	
Sakowitz	[28]	EEG	AV	Passive		x		x	x		x					x	60–220	
Sakowitz	[29]	EEG	AV	Passive	x	x		x			x		x	x	x	x	50–300	
Lakatos	[30]	UA/LFP	AS	Passive			x			x	x		x		x	x	0–250	
Kayser	[31]	UA/LFP	AV	Passive			x			x				x			50–250	
Senkowski	[25]	EEG	AV	Speeded response	x	x			x		x				x		50–170	
Bhattacharya	[32]	EEG	AV	Crossmodal capture					x	x						x	150–500	
Mishra	[33]	EEG	AV	Crossmodal capture					x	x						x	110–240	
Kaiser	[34]	MEG	AV	McGurk effect	x			x	x	x						x	150–400	
Kaiser	[35]	MEG	AV	McGurk effect			x		x	x	x					x	200–350	
Kanayama	[37]	EEG	VS	Rubber hand illusion				x		x		x				x	200–250	
Senkowski	[40]	EEG	AV	Spatial attention	x						x					x	40–60	
Trenner	[42]	EEG	VS	Spatial attention	x	x			x	x				x	x	x	300–700	
Senkowski	[41]	EEG	AV	Nonspatial attention	x				x		x					x	30–120	
Foxe	[43]	EEG	AV	Nonspatial attention				x		x	x			x			750–1050	
Fu	[44]	EEG	AV	Nonspatial attention				x		x				x			700–1100	
Yuval-Greenberg	[45]	EEG	AV	Matching				x		x						x	230–290	
Schneider	[46]	EEG	AV	Matching			x			x						x	120–180	
Doesburg	[47]	EEG	AV	Matching	x			x				x				x	170–250	
Widmann	[48]	EEG	AV	Matching	x	x		x		x	x					x	20–220	
Hummel	[49]	EEG	VS	Matching		x			x	x		x		x			700–1700	

The colors indicate the type of paradigm used: blue, passive stimulation and simple response tasks; yellow, percept-related experiments; gray, attention paradigms; pink, crossmodal matching studies. The studies differ with respect to the recording methods (EEG, MEG or intracortical recordings), the species (humans or monkeys) and the brain regions and the frequency ranges where effects have been observed. Additionally, studies also differ in their focus on the analysis of evoked oscillatory activity, total oscillatory activity or coherence (see Glossary). Abbreviations: AS = audio-somatosensory; AV = audiovisual; C = central; Coh = coherence analysis; EEG = electroencephalography; Evo = evoked activity; F = frontal; LFP = local field potential recordings; LT = lateral-temporal; MEG = magnetoencephalography; O = occipital; OP = occipitoparietal; Tot = total activity; UA = unit recordings; VS = visual-somatosensory; α = alpha; β = beta; δ = delta; γ = gamma; θ = theta.

approaches, specific changes in oscillatory responses or coherence of neural activity have been observed, suggesting that temporally patterned neural signals might be relevant for more than just one type of multisensory interaction.

Bottom-up multisensory processing

One of the earliest EEG studies on crossmodal processing [24] compared coherence patterns during the presentation of auditory and visual object names, as well as pictures of objects, under conditions of passive stimulation. The authors observed an increase of coherence in the lower beta band between temporal and parietal electrode sites and suggested that coherent activity in this frequency range might play a role in the integration of meaningful semantic inputs in a modality-independent network of temporal and parietal areas.

Additional evidence for an involvement of beta oscillations in multisensory processing comes from a study in which subjects were instructed to respond by button-press to the appearance of auditory, visual and combined audiovisual stimuli [25]. In the crossmodal condition, an enhancement was observed for evoked (see Glossary) beta oscillations (cf. Table 1). The relative strength of this integration effect predicted the shortening of reaction times observed for multisensory audiovisual stimuli, suggesting an involvement of beta activity in multisensory processing.

Crossmodal effects on evoked beta responses have also been reported in a sensory gating paradigm [26], in which repetition of a stimulus after a short interval leads to a diminished second response as a result of a decrease in

saliency. In this study, auditory and somatosensory stimuli were presented at different interstimulus intervals. Stronger auditory and somatosensory evoked beta responses were found when the current and preceding stimulus differed in modality, suggesting an involvement of beta oscillations in signaling stimulus saliency. Further EEG studies of oscillatory activity in response to basic audiovisual stimuli [27–29] have shown that, in addition to the beta band, evoked oscillatory responses in the theta, alpha and gamma frequency range are also modulated by multisensory interactions.

Compelling evidence for an association between oscillatory responses and multisensory processing comes from a recent study on the modulation of auditory processing by somatosensory inputs in alert monkeys [30]. The authors investigated the effect of median nerve stimulation on auditory responses and observed a pronounced augmentation of delta, theta and gamma band oscillations (Figure 1a). Further analysis suggested that this effect was mainly due to a phase resetting (see Glossary) of auditory oscillations by the somatosensory inputs. Another intriguing observation in the same study was that systematic variation of the relative delay between somatosensory and auditory inputs led to multisensory response enhancements at intervals corresponding to the cycle length of gamma, theta and delta band oscillations, supporting the notion that oscillations are important in modulating communication between neural assemblies [20].

Additional support for phase resetting as a potential mechanism of crossmodal interaction comes from a recent study focusing on visual modulation of auditory processing in the monkey [31]. Using auditory and visual stimuli

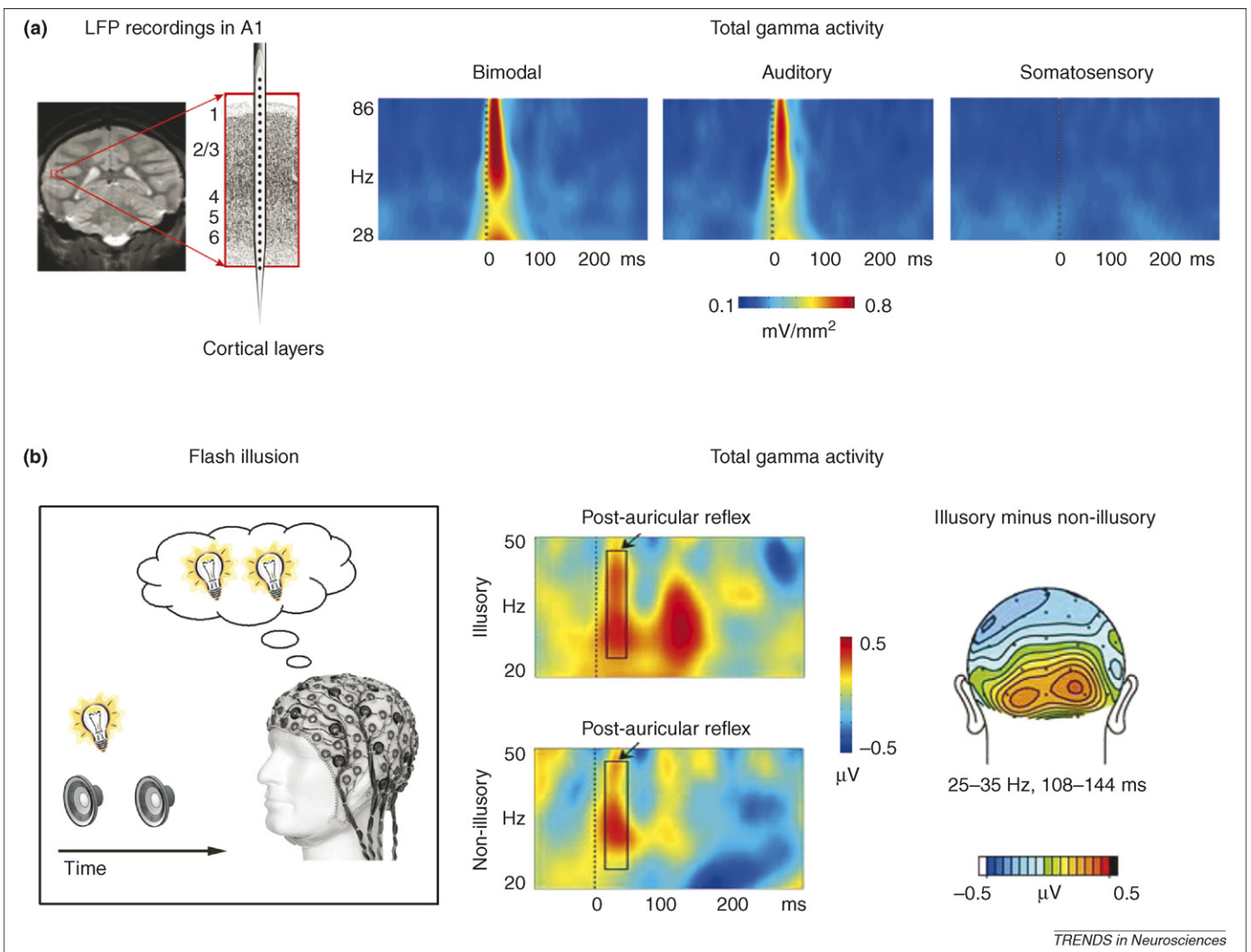


Figure 1. Multisensory processing influences gamma band oscillations. **(a)** Crossmodal effect on oscillatory responses during passive stimulation in the macaque monkey. Local field potentials (LFP) were recorded from primary auditory cortex during auditory, somatosensory and bimodal stimulation. Enhanced total GBA was found in supragranular layers (2/3) after bimodal compared to unimodal stimulation. Adapted, with permission, from Ref. [30]. **(b)** GBA during a flash illusion experiment. A single flash that is presented interposed between rapidly occurring auditory inputs is frequently, but not always, perceived as multiple flashes. GBA was enhanced over occipital scalp in a time window between 100 and 150 ms during illusory trials in which multiple flashes were perceived compared to nonillusory trials in which only a single flash was perceived. A very early burst of enhanced gamma power was attributed to the short latency reflex contraction of the post-auricular muscle. Adapted, with permission, from Ref. [33].

while recording in the auditory core and belt regions of awake behaving monkeys, the authors observed clear crossmodal effects. Importantly, visual stimuli could be shown to modulate the oscillatory phase (see Glossary) of auditory alpha and theta band activity. Taken together, the findings discussed above suggest that modulation of both the power (see Glossary) and the phase of oscillatory activity could be important mechanisms of crossmodal interaction.

Percept-related multisensory oscillations

A powerful approach to studying crossmodal integration is the use of physically identical multisensory events that can lead to different percepts across trials. A well-known example is the sound-induced visual flash illusion wherein a single flash of light accompanied by a pair of rapidly presented auditory beeps is often perceived as two flashes [32,33]. This paradigm allows the direct comparison of neural responses to illusory trials (i.e. when more than one flash is perceived) with nonillusory trials (i.e. when a

single flash is perceived), while keeping the physical parameters of the presented stimuli constant. Two recent studies have investigated gamma band activity (GBA) during this crossmodal capture effect [32,33]. In both, a GBA increase was observed over occipital areas for illusory trials (Figure 1b).

Using a modified version of the McGurk effect, the link between GBA and illusory perception during audiovisual speech processing has been addressed in MEG investigations [34,35]. In the McGurk illusion, an auditory phoneme is dubbed onto a video showing an incongruent articulatory lip movement, which often leads to an illusory auditory percept. For instance, when the syllable /ba/ is presented in the auditory modality together with visual lip movements corresponding to the syllable /ga/, most subjects report hearing the fused percept /da/. Exploiting this crossmodal effect, an enhanced GBA was observed in epochs where an illusory auditory percept was induced by a visual deviant within a continuous stream of multisensory audiovisual speech stimuli [34,35]. Remarkably, the topo-

graphy of this effect was comparable to the frontal topography of a GBA enhancement obtained in an auditory mismatch study [36], suggesting that the GBA effect in the McGurk illusion study might represent a perceived auditory pattern change caused by the visual lip movement.

Further evidence for a link between gamma band oscillations and illusory crossmodal perception comes from a study using a crossmodal rubber hand illusion paradigm [37]. In this study, a rubber hand was positioned on top of a box in which the subject's own hand was placed and visual inputs attached to the rubber hand were presented either at the same or a different stimulation site (forefinger or thumb) as temporally coincident tactile inputs were presented to the subject's hand. Temporally coincident and spatially congruent stimulation produced the illusory perception in participants that the rubber hand formed part of their own body. Closely related to the strength of the perceptual effect, stronger GBA and phase synchrony (see Glossary) occurred in trials with crossmodal congruence [37], indicating that gamma band oscillations might be involved in mediating the visuo-tactile interaction underlying this illusion. In sum, the findings discussed in this section suggest that oscillatory activity,

in particular at gamma band frequencies, can reflect perceptual changes resulting from crossmodal interactions.

Attentional modulation of multisensory oscillations

Behavioral, electrophysiological and functional imaging studies have shown that attention plays an important role in multisensory processing [38,39]. The effect of spatial attention on GBA in a multisensory setting has recently been investigated [40]. Subjects attended to streams of auditory, visual and combined audiovisual stimuli in the left and right hemifields. An enhancement of the phase locking (see Glossary) for the short-latency GBA was found for attended multisensory stimuli (Figure 2a). This suggests that early GBA enhancement might reflect an amplification mechanism which facilitates the integration of task-relevant multisensory inputs.

The effects of nonspatial intersensory attention shifts, and the temporal relation between auditory and visual inputs on the early evoked GBA, have also been investigated using EEG [41]. In this study, subjects attended either auditory or visual targets within a continuous stream of centrally presented stimuli. Whereas the shift

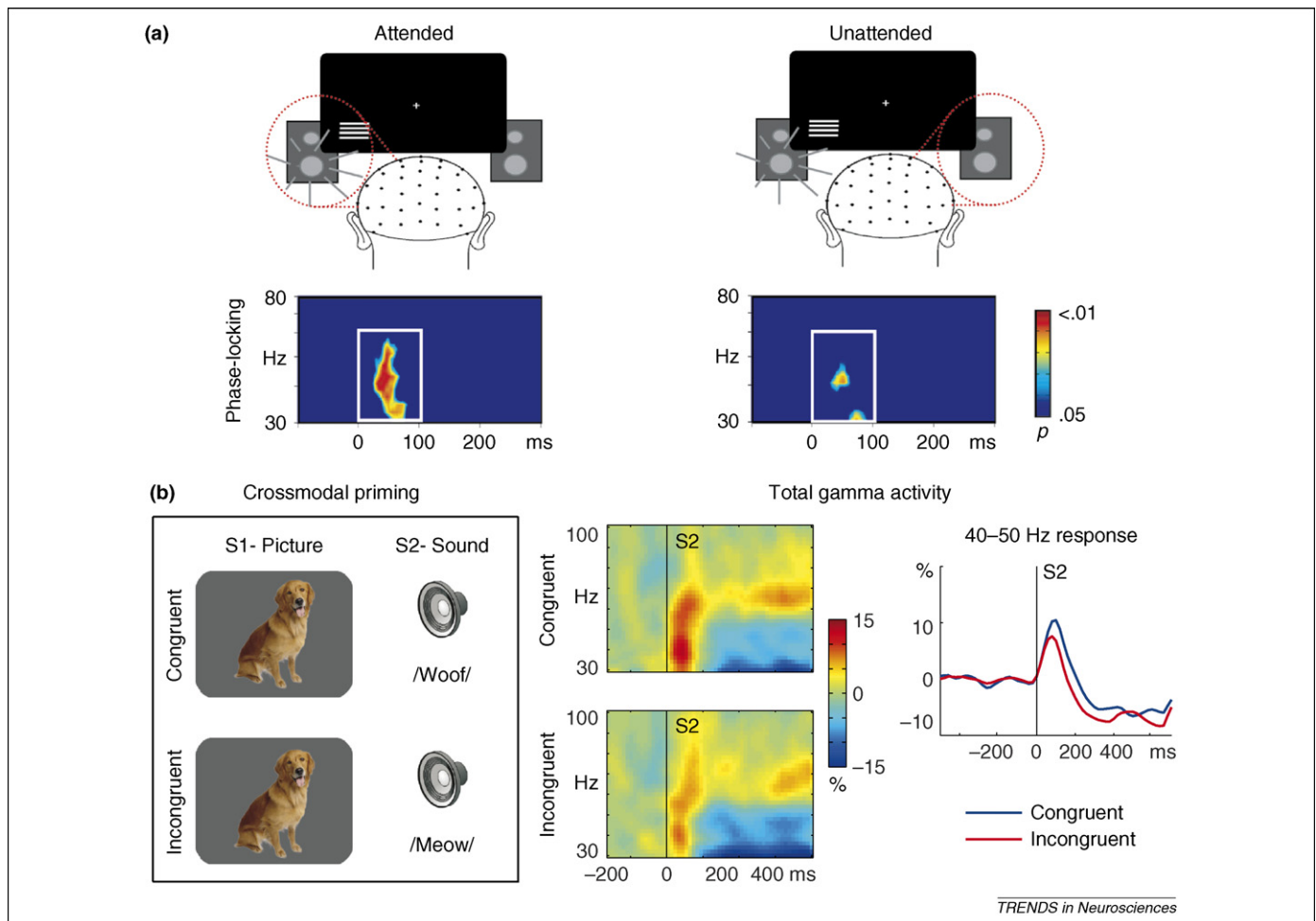


Figure 2. Relation of gamma oscillations to crossmodal attention and matching. (a) GBA and crossmodal attention. Gamma band phase locking was compared for attended (left panel) and unattended (right panel) audiovisual inputs. When attended, physically identical stimuli caused a stronger phase locking of frontal gamma activity at a latency of ~50 ms compared to when they were unattended. Adapted, with permission of Springer Science and Business Media, from Ref. [40]. (b) Total GBA during crossmodal semantic priming. Subjects had to categorize auditory stimuli (S2) that were presented following a semantically congruent or incongruent visual prime (S1). For the congruent condition, enhanced GBA was observed. Adapted, with permission, from Ref. [46].

of attention to either the visual or the auditory stimulus component did not influence GBA in this paradigm, a robust enhancement of GBA was observed when auditory and visual inputs were presented with the closest temporal synchrony (0 ± 25 ms). This indicates that the integration of auditory and visual signals, as reflected in high-frequency oscillatory activity, is sensitive to the relative onset timing of the sensory inputs [1,2,14]. This finding agrees with earlier studies showing that temporal contiguity enhances the saliency of multisensory stimuli and facilitates crossmodal integration [1,2].

Another series of studies has investigated the role of alpha band oscillatory activity during intersensory attention shifts. In these studies, rather than continuously integrating multisensory inputs, subjects had to disengage attention from one or the other constituent unisensory elements so that a difficult task could be solved in just a single sensory stream [42–44]. Using tactile, auditory or visual inputs as cues, these investigations have shown that suppression of ongoing alpha oscillations is likely to reflect an endogenous shift of attention to the visual modality and this occurs independent of the modality of the stimulus used to cue attention. Critically, when subjects were cued to ignore concurrent visual inputs, an increase in alpha oscillatory activity was seen over visual attentional control regions. This suggests that when visual and nonvisual inputs are to be segregated from each other, oscillations in the alpha band might cause a suppression of integrative functioning, perhaps by interrupting integrative mechanisms in other frequency bands. Taken together, the available evidence suggests that oscillatory responses in multiple frequency bands are modulated by spatial attention shifts in multisensory settings. However, more studies are needed to elaborate a consistent view on the attentional modulation of dynamic crossmodal interactions.

Oscillations and crossmodal matching

An important factor influencing crossmodal integration is the semantic matching of information across sensory channels. A recent study has addressed this issue during audiovisual processing in an object recognition task [45], in which sounds of animals were presented in combination with a picture of either the same or a different animal. Larger GBA was observed for semantically congruent compared to semantically incongruent audiovisual stimuli. We have recently obtained similar results using an audiovisual priming paradigm [46], in which semantically congruent and incongruent pairs of visual and auditory stimuli were presented in sequence, and the auditory stimuli had to be categorized (Figure 2b). Gamma band responses to auditory inputs were enhanced when visual and auditory inputs were semantically congruent compared to when they were incongruent. Source localization using the method of 'linear beamforming' revealed that the GBA enhancement occurred mainly in multisensory regions in the left lateral temporal cortex [46]. This suggests that the GBA enhancement might reflect a matching operation between multisensory temporal cortex and auditory cortex that is carried out upon appearance of the auditory stimulus.

The effect of multisensory matching of meaningful stimuli on oscillatory activity has also been the subject of studies that have used verbal or abstract symbolic materials. Exploiting the interesting case of synchronous versus asynchronous audiovisual speech [47], changes in phase coherence were shown to occur in a transiently activated gamma oscillatory network. GBA phase coherence was increased for asynchronous as compared to synchronous speech between frontal and left posterior sensors, whereas GBA amplitude showed an enhancement for synchronous compared to asynchronous speech at long latencies after stimulus onset.

The relationship between early evoked auditory GBA and multisensory processing has also been investigated in an audiovisual symbol-to-sound matching paradigm [48]. An enhanced evoked GBA was found for auditory stimuli that matched the elements of a visual pattern compared to auditory inputs that did not match the visual pattern. In another study, the role of neural synchronization between visual and sensorimotor cortex was examined in a multisensory matching task in which tactile Braille stimuli and visual dot patterns had to be compared [49]. In trials in which subjects performed well compared to trials in which they performed poorly, an enhancement of phase coherence in the alpha band between occipital and lateral central regions was observed. In summary, the available studies suggest that crossmodal matching might be reflected in both local and long-range changes of neural coherence.

Functional role of neural synchrony for crossmodal interactions

The data available so far support the hypothesis that coherence of oscillatory responses might play a crucial role in multisensory processing. They show that multisensory interactions can be accompanied by condition-specific changes in oscillatory responses which often, albeit not always, occur in the gamma band [28,30,32–35,37,40,41,45–48]. The effects observed in EEG or MEG signals are likely to result not only from changes in oscillatory power but also in the phase coherence within the underlying neuronal populations. Several of the studies reviewed above have tried to address this question directly, and provide evidence that coherence of neural signals across cortical areas might be a crucial mechanism involved in multisensory processing [24,37,47,49]. Moreover, the available studies demonstrate effects of multisensory interactions on oscillatory responses at multiple levels, including primary sensory areas [30,31,35] as well as higher-order multisensory and frontal areas [25,34], suggesting that coherent neural activity might play a role for both 'early' and 'late' integration of multisensory signals.

For several reasons, coherent oscillatory signals might be well suited to serve crossmodal integration. It has been argued that synchronization of neural activity might help to cope with binding problems that occur in distributed architectures [16–18]. Clearly, multisensory processing poses binding problems in at least two respects [15]: first, information must be integrated across different cortical and subcortical regions; second, real-world scenes comprise multiple objects, creating the need to segregate

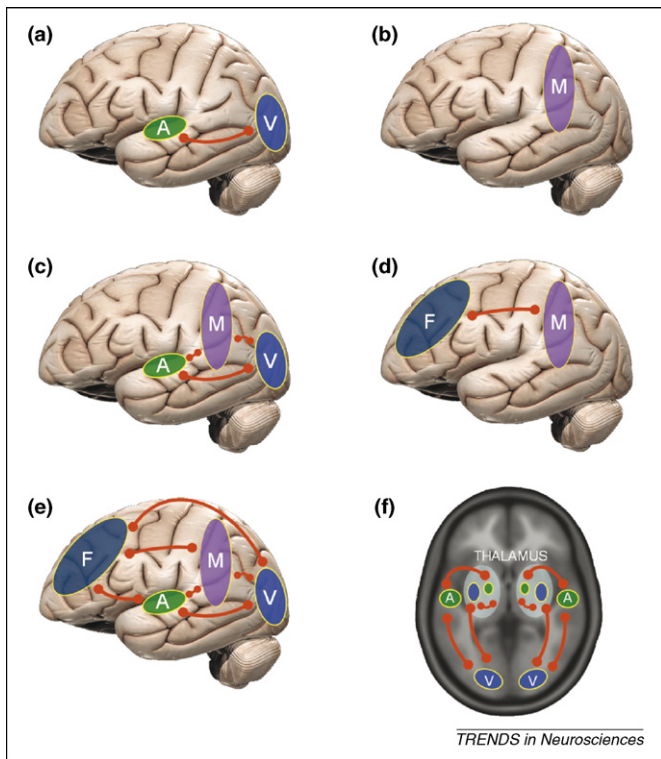


Figure 3. Hypothetical scenarios for crossmodal binding through neural coherence. Using the case of audiovisual interactions, several possible scenarios are schematically depicted in the figure. **(a)** The simplest scenario predicts that during multisensory interactions, neural synchronization could change between early sensory areas. **(b)** An alternative possibility is that changes in neural coherence or power occur mainly within cell assemblies of multisensory association cortices such as, for instance, parietal or superior temporal regions. **(c)** Changes in neural synchrony among unimodal regions could also be associated with enhanced oscillatory activity in multisensory areas. This could reflect reentrant bottom-up and top-down interactions between uni- and multimodal cortices. **(d)** In addition, changes in multisensory perception will often also involve frontal regions, which might exert a modulatory influence on temporal patterns in multisensory parieto-temporal regions through oscillatory coupling. **(e)** Most likely, at least for multisensory processing in natural environments, these interactions will combine into a highly complex pattern involving frontal cortex, temporo-parietal regions as well as unimodal cortices. **(f)** Very likely, the network of regions engaged in coherent activity will also involve subcortical structures such as thalamic nuclei. Abbreviations: A = auditory cortex; V = visual cortex; M = higher-order multisensory regions; F = prefrontal cortex. Brain image reproduced with permission from <http://www.joelertola.com>.

unrelated neural signals within processing modules, while at the same time selectively coordinating signals across channels. It seems unlikely that such complex coordination could be achieved by selective anatomical connections, because these would not provide sufficient flexibility to cope with a fast-changing multisensory world. By contrast, establishment of relations through neural coherence could provide both the required flexibility and selectivity, as transient phase synchrony of oscillatory signals allows the functional connectivity between spatially distributed neuronal populations to be dynamically modulated [20,22,23]. Furthermore, there is evidence to suggest that assemblies of coherently active neurons are endowed with a competitive advantage, leading to selection of strongly synchronized populations and suppression of decorrelated activity in situations involving competing or ambiguous stimuli [50,51]. Because synchronized signals are particularly efficient in driving downstream cell populations [23] and in modulating synaptic weights [52], such a mechanism

would allow the preferential processing and storage of the respective information and, thus, strongly facilitate attentional selection, memory formation, decision making and action generation [18–20].

If neural coherence does indeed support multisensory integration, several scenarios seem possible regarding the interaction of ‘early’ and higher-order regions (Figure 3). Matching or congruent multisensory information would very likely lead to enhanced coherence among neurons that primarily respond to unisensory inputs. This in turn would lead to stronger activation of cells in multisensory regions that receive input from the respective modalities. Top-down influences from higher-order multisensory regions to cell populations in primary or secondary cortices might also be mediated by coherence of oscillatory signals [18,19]. Presumably, during perception of natural multimodal scenes both bottom-up and top-down interactions come into play, and reentrant processing is very likely to occur.

Outlook

Clearly, the hypothesis that neural synchrony might play a role in multisensory processing requires further experimental testing (Box 2). Thus far, only a relatively small number of multisensory studies have used coherence measures to explicitly address interactions across different neural systems. Substantial progress will require studies in humans with approaches suitable to capture dynamic cross-systems interactions in source space [53,54]. Furthermore, to characterize the role of correlated activity for multisensory processing at the cellular level, additional microelectrode studies in higher mammals will be indispensable.

The view discussed here also has potential implications for explaining conditions of altered crossmodal perception. Abnormal synchronization across brain regions might play a role in conditions such as synesthesia or in disorders such as schizophrenia or autism [55,56]. In synesthesia, excessively strong multisensory coherence might occur which then would not just modulate processing in unimodal

Box 2. Outstanding questions

- o How can the temporal correlation hypothesis account for the various types of multisensory interactions?
- o Will current or future approaches allow the quantification of multisensory coherence in source space?
- o What is the role of different frequency bands in multisensory integration? Why does the gamma frequency range seem to play a prominent role?
- o What is the role of crossfrequency interactions in multisensory processing?
- o What is the role of neural synchrony for crossmodal processing in clinical populations in which deficits in multisensory integration have been shown such as schizophrenia or autism?
- o Can abnormalities in neural coherence account for variants of multisensory interactions such as those found in synesthesia?
- o By which mechanisms do top-down factors such as attention or memory influence multisensory integration that is mediated by neural coherence?
- o Which types of connections and which cortical regions are critical in controlling multisensory integration through neural coherence?

regions but actually drive sensory neurons even in the absence of a proper stimulus. By contrast, abnormal weakness or instability of crossmodal coupling might account for the impairment of multisensory integration that is observed in patients with schizophrenia [57] or autism [58].

We believe that the study of synchronization phenomena can lead to a new view on multisensory processing which considers the dynamic interplay of neural populations as a key to crossmodal integration and stipulates the development of new research approaches and experimental strategies. Conversely, the investigation of multisensory interactions might also provide a crucial test bed for further validation of the temporal correlation hypothesis [16–18], because task- or percept-related changes of coherence between independent neural sources have hardly been shown in humans so far. In this context, the role of oscillations of different frequencies in crossmodal integration is yet another unexplored issue that future studies will need to address.

Acknowledgements

We thank Markus Siegel, Peter König, Jörg Hipp and Gernot Supp for helpful comments on the manuscript. D.S. received support from a NARSAD 2006 young investigator award and the German Research Foundation (SE 1859/1–1). J.J.F. received support from a U.S. National Institute of Mental Health grant (NIMH–RO1 MH65350). A.K.E. acknowledges support by the European Union (IST-027268, NEST-043457), the German Research Foundation (GRK 1247/1), the German Federal Ministry of Education and Research (01GW0561) and the Volkswagen Foundation (II/80609).

References

- Stein, B.E. and Meredith, M.A. (1993) *The Merging of the Senses*, MIT Press
- Meredith, M.A. (2002) On the neuronal basis for multisensory convergence: a brief overview. *Cogn. Brain Res.* 14, 31–40
- Macaluso, E. and Driver, J. (2005) Multisensory spatial interactions: a window onto functional integration in the human brain. *Trends Neurosci.* 28, 264–271
- Kayser, C. and Logothetis, N.K. (2007) Do early sensory cortices integrate crossmodal information? *Brain Struct. Funct.* 212, 121–132
- Amedi, A. *et al.* (2005) Functional imaging of human crossmodal identification and object recognition. *Exp. Brain Res.* 166, 559–571
- Ghazanfar, A.A. and Schroeder, C.E. (2006) Is neocortex essentially multisensory? *Trends Cogn. Sci.* 10, 278–285
- Calvert, G.A. (2001) Crossmodal processing in the human brain: insights from functional neuroimaging studies. *Cereb. Cortex* 11, 1110–1123
- Foxe, J.J. *et al.* (2002) Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. *J. Neurophysiol.* 88, 540–543
- Kayser, C. *et al.* (2005) Integration of touch and sound in auditory cortex. *Neuron* 48, 373–384
- Nagy, A. *et al.* (2006) Multisensory integration in the basal ganglia. *Eur. J. Neurosci.* 24, 917–924
- Baumann, O. and Greenlee, M.W. (2007) Neural correlates of coherent audiovisual motion perception. *Cereb. Cortex* 17, 1433–1443
- Nishijo, H. *et al.* (1988) Topographic distribution of modality-specific amygdalar neurons in alert monkey. *J. Neurosci.* 8, 3556–3569
- Komura, Y. *et al.* (2005) Auditory thalamus integrates visual inputs into behavioral gains. *Nat. Neurosci.* 8, 1203–1209
- Driver, J. and Spence, C. (2000) Multisensory perception: beyond modularity and convergence. *Curr. Biol.* 10, R731–R735
- Foxe, J.J. and Schroeder, C.E. (2005) The case for feedforward multisensory convergence during early cortical processing. *Neuroreport* 16, 419–423
- Engel, A.K. *et al.* (1992) Temporal coding in the visual cortex: new vistas on integration in the nervous system. *Trends Neurosci.* 15, 218–226
- Singer, W. and Gray, C.M. (1995) Visual feature integration and the temporal correlation hypothesis. *Annu. Rev. Neurosci.* 18, 555–586
- Engel, A.K. *et al.* (2001) Dynamic predictions: oscillations and synchrony in top-down processing. *Nat. Rev. Neurosci.* 2, 704–716
- Herrmann, C.S. *et al.* (2004) Cognitive functions of gamma-band activity: memory match and utilization. *Trends Cogn. Sci.* 8, 347–355
- Fries, P. (2005) A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn. Sci.* 9, 474–480
- Jensen, O. *et al.* (2007) Human gamma-frequency oscillations associated with attention and memory. *Trends Neurosci.* 30, 317–324
- Salinas, E. and Sejnowski, T.J. (2001) Correlated neuronal activity and the flow of neural information. *Nat. Rev. Neurosci.* 2, 539–550
- Womelsdorf, T. *et al.* (2007) Modulation of neuronal interactions through neuronal synchronization. *Science* 316, 1609–1612
- von Stein, A. *et al.* (1999) Synchronization between temporal and parietal cortex during multimodal object processing in man. *Cereb. Cortex* 9, 137–150
- Senkowski, D. *et al.* (2006) Oscillatory beta activity predicts response speed during a multisensory audiovisual reaction time task: a high-density electrical mapping study. *Cereb. Cortex* 16, 1556–1565
- Kisley, M.A. and Cornwell, Z.M. (2006) Gamma and beta neural activity evoked during a sensory gating paradigm: effects of auditory, somatosensory and cross-modal stimulation. *Clin. Neurophysiol.* 117, 2549–2563
- Sakowitz, O.W. *et al.* (2000) Oscillatory frontal theta responses are increased upon bisensory stimulation. *Clin. Neurophysiol.* 111, 884–893
- Sakowitz, O.W. *et al.* (2001) Bisensory stimulation increases gamma-responses over multiple cortical regions. *Brain Res. Cogn. Brain Res.* 11, 267–279
- Sakowitz, O.W. *et al.* (2005) Spatio-temporal frequency characteristics of intersensory components in audiovisual evoked potentials. *Brain Res. Cogn. Brain Res.* 23, 316–326
- Lakatos, P. *et al.* (2007) Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron* 53, 279–292
- Kayser, C. *et al.* (2008) Visual modulation of neurons in auditory cortex. *Cereb. Cortex* 10.1093/cercor/bhm187 (<http://cercor.oxfordjournals.org>)
- Bhattacharya, J. *et al.* (2002) Sound-induced illusory flash perception: role of gamma band responses. *Neuroreport* 13, 1727–1730
- Mishra, J. *et al.* (2007) Early cross-modal interactions in auditory and visual cortex underlie a sound-induced visual illusion. *J. Neurosci.* 27, 4120–4131
- Kaiser, J. *et al.* (2005) Hearing lips: gamma-band activity during audiovisual speech perception. *Cereb. Cortex* 15, 646–653
- Kaiser, J. *et al.* (2006) Gamma-band activity over early sensory areas predicts detection of changes in audiovisual speech stimuli. *Neuroimage* 30, 1376–1382
- Kaiser, J. *et al.* (2002) Dynamics of gamma-band activity induced by auditory pattern changes in humans. *Cereb. Cortex* 12, 212–221
- Kanayama, N. *et al.* (2007) Crossmodal effect with rubber hand illusion and gamma-band activity. *Psychophysiology* 44, 392–402
- Macaluso, E. *et al.* (2000) Modulation of human visual cortex by crossmodal spatial attention. *Science* 289, 1206–1208
- Foxe, J.J. *et al.* (2005) Biasing the brain's attentional set: I. Cue driven deployments of intersensory selective attention. *Exp. Brain Res.* 166, 370–392
- Senkowski, D. *et al.* (2005) Multisensory processing and oscillatory gamma responses: effects of spatial selective attention. *Exp. Brain Res.* 166, 411–426
- Senkowski, D. *et al.* (2007) Good times for multisensory integration: effects of the precision of temporal synchrony as revealed by gamma-band oscillations. *Neuropsychologia* 45, 561–571
- Trenner, M.U. *et al.* (2008) What happens in between? Human oscillatory brain activity related to crossmodal spatial cueing. *PLoS ONE* 3, e1467
- Foxe, J.J. *et al.* (1998) Parieto-occipital approximately 10 Hz activity reflects anticipatory state of visual attention mechanisms. *Neuroreport* 9, 3929–3933

- 44 Fu, K.M. *et al.* (2001) Attention-dependent suppression of distracter visual input can be cross-modally cued as indexed by anticipatory parieto-occipital alpha-band oscillations. *Brain Res. Cogn. Brain Res.* 12, 145–152
- 45 Yuval-Greenberg, S. and Deouell, L.Y. (2007) What you see is not (always) what you hear: induced gamma band responses reflect cross-modal interactions in familiar object recognition. *J. Neurosci.* 27, 1090–1096
- 46 Schneider, T.R. *et al.* (2008) Enhanced EEG gamma-band activity reflects multisensory semantic matching in visual-to-auditory object priming. *Neuroimage*, DOI: 10.1016/j.neuroimage.2008.05.033 In: <http://www.sciencedirect.com/science/journal/10538119>
- 47 Doesburg, S.M. *et al.* (2008) Asynchrony from synchrony: long-range gamma-band neural synchrony accompanies perception of audiovisual speech asynchrony. *Exp. Brain Res.* 185, 11–20
- 48 Widmann, A. *et al.* (2007) Binding symbols and sounds: evidence from event-related oscillatory gamma-band activity. *Cereb. Cortex* 17, 2696–2702
- 49 Hummel, F. and Gerloff, C. (2005) Larger interregional synchrony is associated with greater behavioral success in a complex sensory integration task in humans. *Cereb. Cortex* 15, 670–678
- 50 Fries, P. *et al.* (1997) Synchronization of oscillatory responses in visual cortex correlates with perception in interocular rivalry. *Proc. Natl. Acad. Sci. U. S. A.* 94, 12699–12704
- 51 Fries, P. *et al.* (2001) Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* 291, 1560–1563
- 52 Bi, G. and Poo, M. (2001) Synaptic modification by correlated activity: Hebb's postulate revisited. *Annu. Rev. Neurosci.* 24, 139–166
- 53 Gross, J. *et al.* (2001) Dynamic imaging of coherent sources: studying neural interactions in the human brain. *Proc. Natl. Acad. Sci. U. S. A.* 98, 694–699
- 54 Supp, G.G. *et al.* (2007) Directed cortical information flow during human object recognition: analyzing induced EEG gamma-band responses in brain's source space. *PLoS ONE* 2, e684
- 55 Schnitzler, A. and Gross, J. (2005) Normal and pathological oscillatory communication in the brain. *Nat. Rev. Neurosci.* 6, 285–296
- 56 Uhlhaas, P.J. and Singer, W. (2006) Neural synchrony in brain disorders: relevance for cognitive dysfunctions and pathophysiology. *Neuron* 52, 155–168
- 57 Ross, L.A. *et al.* (2007) Impaired multisensory processing in schizophrenia: deficits in the visual enhancement of speech comprehension under noisy environmental conditions. *Schizophr. Res.* 97, 173–183
- 58 Iarocci, G. and McDonald, J. (2006) Sensory integration and the perceptual experience of persons with autism. *J. Autism Dev. Disord.* 36, 77–90
- 59 Basar, E. *et al.* (2000) Brain oscillations in perception and memory. *Int. J. Psychophysiol.* 35, 95–124
- 60 Klimesch, W. *et al.* (2006) Upper alpha ERD and absolute power: their meaning for memory performance. *Prog. Brain Res.* 159, 151–165
- 61 Palva, S. and Palva, J.M. (2007) New vistas for α -frequency band oscillations. *Trends Neurosci.* 30, 150–158
- 62 Knyazev, G.G. (2007) Motivation, emotion, and their inhibitory control mirrored in brain oscillations. *Neurosci. Biobehav. Rev.* 31, 377–395
- 63 Jensen, O. and Lisman, J.E. (2005) Hippocampal sequence-encoding driven by a cortical multi-item working memory buffer. *Trends Neurosci.* 28, 67–72
- 64 Brovelli, A. *et al.* (2004) Beta oscillations in a large-scale sensorimotor cortical network: directional influences revealed by Granger causality. *Proc. Natl. Acad. Sci. U. S. A.* 101, 9849–9854
- 65 Gray, C.M. *et al.* (1989) Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature* 338, 334–337
- 66 Engel, A.K. *et al.* (1991) Direct physiological evidence for scene segmentation by temporal coding. *Proc. Natl. Acad. Sci. U. S. A.* 88, 9136–9140
- 67 Tallon-Baudry, C. *et al.* (1996) Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *J. Neurosci.* 16, 4240–4249
- 68 Castelo-Branco, M. *et al.* (2000) Neural synchrony correlates with surface segregation rules. *Nature* 405, 685–689
- 69 Siegel, M. *et al.* (2007) High-frequency activity in human visual cortex is modulated by visual motion strength. *Cereb. Cortex* 17, 732–741
- 70 Müller, M.M. *et al.* (2000) Modulation of induced gamma band activity in the human EEG by attention and visual information processing. *Int. J. Psychophysiol.* 38, 283–299
- 71 Debener, S. *et al.* (2003) Top-down attentional processing enhances auditory evoked gamma band activity. *Neuroreport* 14, 683–686
- 72 Bauer, M. *et al.* (2006) Tactile spatial attention enhances gamma-band activity in somatosensory cortex and reduces low-frequency activity in parieto-occipital areas. *J. Neurosci.* 26, 490–501
- 73 Wehr, M. and Laurent, G. (1996) Odour encoding by temporal sequences of firing in oscillating neural assemblies. *Nature* 384, 162–166
- 74 Roelfsema, P.R. *et al.* (1997) Visuomotor integration is associated with zero time-lag synchronization among cortical areas. *Nature* 385, 157–161
- 75 Sanes, J.N. and Donoghue, J.P. (1993) Oscillations in local field potentials of the primate motor cortex during voluntary movement. *Proc. Natl. Acad. Sci. U. S. A.* 90, 4470–4474
- 76 Tallon-Baudry, C. *et al.* (1998) Induced γ -band activity during the delay of a visual short-term memory task in humans. *J. Neurosci.* 18, 4244–4254
- 77 Osipova, D. *et al.* (2006) Theta and gamma oscillations predict encoding and retrieval of declarative memory. *J. Neurosci.* 26, 7523–7531
- 78 Schroeder, C.E. *et al.* (2008) Neuronal oscillations and visual amplification of speech. *Trends Cogn. Sci.* 12, 106–113