

# Chapter 10

## Neural Network Dynamics and Audiovisual Integration



Julian Keil and Daniel Senkowski

**Abstract** Why does seeing a speaker's lip movements improve understanding speech in noisy environments? Why does simultaneous ringing and vibrating quicken answering a phone? These are questions of interest for researchers in the field of multisensory information processing. Electrophysiological approaches are suited to map the neural network dynamics underlying multisensory perception. Combining findings from behavioral, functional neuroimaging, and electrophysiological studies allows a comprehensive understanding of how information is integrated across the different senses. This chapter first provides an introduction on the relationships between neural network dynamics, as reflected in neural oscillations, and unisensory perception. Then, the relevance of neural network dynamics for multisensory perception is described, with a special focus on the auditory system. Moreover, the chapter provides an overview on how visual and auditory information can mutually influence each other and highlights the crucial role of ongoing neural network dynamics for upcoming multisensory perception. Finally, general principles of audiovisual integration are established, and open questions and future direction in the field of multisensory perception are discussed.

**Keywords** Alpha band · Amplitude · Beta band · Connectivity · Electroencephalography · Electrophysiology · Gamma band · Illusion · Magnetoencephalography · Multisensory integration · Oscillation · Phase · Speech

---

J. Keil (✉)

Biological Psychology, Christian-Albrechts-University Kiel, Kiel, Germany  
e-mail: [keil@psychologie.uni-kiel.de](mailto:keil@psychologie.uni-kiel.de)

D. Senkowski

Department of Psychiatry and Psychotherapy, Charité—Universitätsmedizin Berlin, Berlin, Germany  
e-mail: [daniel.senkowski@charite.de](mailto:daniel.senkowski@charite.de)

© Springer Nature Switzerland AG 2019

A. K. C. Lee et al. (eds.), *Multisensory Processes*, Springer Handbook of Auditory Research 68, [https://doi.org/10.1007/978-3-030-10461-0\\_10](https://doi.org/10.1007/978-3-030-10461-0_10)

201

## 10.1 Introduction

Why does seeing a speaker's lip movements improve understanding speech in noisy environments? Why does simultaneous ringing and vibrating quicken answering a phone? These are questions of interest for researchers in the field of multisensory information processing (Sumbly and Polack 1954; Pomper et al. 2014). The study of multisensory integration at the behavioral level can provide valuable information about the conditions under which information from different senses interact. Moreover, functional neuroimaging approaches are well suited to study which cortical regions are involved in the perception and processing of multisensory information. Electrophysiological approaches, in particular, are suited to map the neural network dynamics underlying multisensory perception. The combined knowledge from behavioral, functional neuroimaging, and electrophysiological studies allows a comprehensive understanding of how information is integrated across the different senses.

This chapter first provides an introduction on the relationships between neural network dynamics, as reflected in neural oscillations, and perception (Sect. 10.1.1). Then, the relevance of neural network dynamics for multisensory perception is described, with a special focus on the auditory system (Sects. 10.2 and 10.3). Subsequently, the chapter describes how visual and auditory information can mutually influence each other (Sects. 10.4 and 10.5). This chapter also highlights the crucial role of ongoing neural network dynamics for upcoming perception (Sect. 10.6). Finally, general principles of audiovisual integration based on presented findings are established (Sect. 10.7), and open questions and future direction in the field of multisensory perception are discussed (Sect. 10.7.4).

### 10.1.1 *Oscillatory Neural Activity Relates to Cognition and Perception*

“Clocks tick, bridges vibrate, and neural networks oscillate” (Buzsáki and Draguhn 2004). Oscillatory neural activity recorded through electroencephalography (EEG) or magnetoencephalography (MEG) can be understood as the synchronous waxing and waning of summed postsynaptic activity of large neural populations in circumscribed brain regions (Lopes da Silva 1991; Wang 2010). The resulting waveform can be dissected into different frequency components with distinct amplitudes (also called power) and phases (Herrmann et al. 1999; Mitra and Pesaran 1999). In the frequency components, two types of oscillatory responses, which reflect different aspects of neural synchronization, can be distinguished: evoked and induced oscillations (Tallon-Baudry and Bertrand 1999). The former are closely related to the onset of an external event and are strictly phase and time locked to the stimulus onset. The phase locking of oscillatory responses can be quantified by intertrial coherence (ITC; Cheron et al. 2007) and the summation over trials of identical

phase can result in event-related potentials (ERPs; Luck 2014). Induced oscillations can be elicited by stimulation but are also present independent of external stimulation. Induced oscillations do not have to be strictly phase and time locked to the onset of stimuli (Tallon-Baudry and Bertrand 1999). Evoked and induced oscillations can be modulated by cognitive processes. Moreover, functional connectivity, the interaction between oscillatory activities in different cortical regions, can be reflected in phase coherence. 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 (Fries 2005, 2015). Information processing as well as transfer and storage in the cortex has been hypothesized to rely on flexible cell assemblies, which are defined as transiently synchronized neural networks (Engel et al. 2001; Buzsáki and Draguhn 2004). The transient synchronization of cell assemblies by oscillatory activity depends on the coupling strength between neural populations as well as on the frequency distribution; as long as the frequencies of coupled cell assemblies are similar, the synchronization within the neural network can be sustained with weak synaptic links (Wang 2010). In general, the analysis of oscillatory cortical activity can provide valuable information on the temporal structure of local processes and network interactions underlying perception and cognition.

Neural networks in mammals exhibit oscillatory activity ranging between approximately 0.05 Hz and 350 Hz (Penttonen and Buzsáki 2003). In humans, oscillatory activity patterns were among the first signals recorded using EEG (Berger 1929; Bremer 1958). Within one neural network, neighboring frequency bands can compete with each other and can be associated with different cognitive and perceptual processes (Engel et al. 2001; Buzsáki and Draguhn 2004). Typically, multiple rhythms coexist at the same time, which result in complex waveforms consisting of high- and low-frequency oscillations (Steriade 2001). One way to organize the multiple rhythms is to divide the frequency spectrum into neighboring frequency bands (Buzsáki and Draguhn 2004). Oscillatory slow-wave activity (below 1 Hz) plays a prominent role in sleep and memory (Penttonen and Buzsáki 2003; Diekelmann and Born 2010), but also reflects changes in cortical excitability related to task performance (Birbaumer et al. 1990; Rockstroh et al. 1992). Above these slow-wave oscillations, Walter (1936) described the delta band, which comprises oscillatory activity below 4 Hz. In the frequency range of 4–7 Hz, Walter et al. (1966) identified the theta band. Both delta band and theta band activity have been related to memory processing (Klimesch 1999; Sauseng et al. 2005). More recently, theta band activity has been linked to cognitive control mechanisms such as attention and predictions (Cavanagh and Frank 2014). In his seminal article from the late 1920s, Berger described that the EEG is dominated by ongoing 8- to 12-Hz oscillations, which were later termed alpha band activity (Berger 1929). Of particular note was Berger's observation that alpha band activity changed with the participant's behavior: alpha band power increased when participants closed their eyes and decreased when they opened the eyes (Berger 1929). Ray and Cole (1985) proposed that oscillatory activity in different frequency bands reflects different cognitive processes. In two experiments, the authors established that alpha band activity relates to attentional processes and is increased if attention is not required. Additionally,

ongoing alpha band oscillations influence subsequent perception (Lange et al. 2014). Recently, the alpha band has been ascribed an important role in attention and the routing of information (Jensen and Mazaheri 2010; Klimesch 2012). Above the alpha band, Berger (1929) identified the beta band (13–30 Hz), but its functional significance has only been studied many years later (Pfurtscheller 1992; Engel and Fries 2010). Recent studies have provided evidence that, besides motor functions, beta band activity relates to cognitive and emotional processing and that it might reflect cortical feedback processing (Keil et al. 2016; Michalareas et al. 2016). Cortical activity in frequencies above the beta band (i.e., >30 Hz) has been coined gamma band activity (Adrian 1942; Bressler 1990). It has been proposed that oscillatory activity in the gamma band may form a mechanism for feature representation of a given stimulus (Lopes da Silva 1991). Findings from the auditory and visual domains support this notion. For instance, using intracranial recordings from the calcarine region of the visual cortex in epileptic patients, Chatrian et al. (1960) described a rhythmic response to visual stimulation at a frequency of around 50 Hz. Moreover, in response to auditory stimuli, Pantev et al. (1991) described a transient oscillatory response at around 40 Hz.

Thus, oscillatory activity in different frequency bands relates to different perceptual and cognitive processes and reflects the functional states of neural networks (Lopes da Silva 1991). However, multiple neighboring frequency bands can be involved in a single process and multiple processes can relate to a single frequency. Moreover, the boundaries between different frequency bands can vary by task and recording technique (Buzsáki and Draguhn 2004). In summary, there is robust evidence that oscillatory activity in different frequency bands relates to various perceptual and cognitive functions (Table 10.1).

## 10.2 Role of Oscillatory Processes for Multisensory Integration and Perception

Multisensory perception requires processing in primary sensory areas as well as the formation of multimodal coherent percepts in distributed neural networks. In an early EEG study on oscillatory activity and multisensory processing, Sakowitz et al. (2001) found increased gamma band power in response to audiovisual stimuli

**Table 10.1** Overview of the classical frequency bands found in human electrophysiological data and examples of functions ascribed to the frequency bands

Name	Frequency range (Hz)	Exemplary function
Slow wave	<1	Sleep, Memory, Cortical excitability
Delta	1–3	Memory
Theta	4–7	Memory, Attention, Cognitive control
Alpha	8–12	Attention, Cognition, Routing of information
Beta	13–30	Attention, Cognition, Stimulus processing
Gamma	>30	Stimulus processing, Feature binding

compared with auditory or visual stimuli alone. A later EEG study extended this finding by showing that evoked gamma band power to audiovisual stimuli increases, in particular, for attended versus unattended stimuli (Senkowski et al. 2005). Interestingly, another study found increased occipital gamma band power following the presentation of incongruent audiovisual stimuli, but only if the audiovisual stimuli were integrated into a coherent perception (Bhattacharya et al. 2002). Whereas these studies demonstrate that multisensory processes or, at least, specific aspects of multisensory processes are presumably reflected in gamma band power, they did not examine the underlying cortical networks.

Traditionally, it has been assumed that multisensory integration is a higher order process that occurs after stimulus processing in unisensory cortical and subcortical areas (Driver and Noesselt 2008). However, a number of studies have challenged this idea by providing evidence for multisensory convergence in low-level sensory cortices (Schroeder and Foxe 2005; Ghazanfar and Schroeder 2006). Using intracranial recordings in monkeys, Lakatos et al. (2007) showed that somatosensory stimulation modulates activity in primary auditory areas. Interestingly, the authors found evidence for a theta band phase reset of ongoing oscillatory activity in the primary auditory cortex by concurrent somatosensory input. The authors suggested that stimulus responses are enhanced when their onset falls into a high-excitability phase and suppressed when the onset falls into a low-excitability phase. These observations are in-line with another study recording local field potentials and single-unit activity in monkeys, which highlights the role of oscillatory alpha band phase for the modulation of auditory evoked activity (Kayser et al. 2008). Analyzing ITC as a measure of transient phase synchronization in intracranial recordings from the visual cortex in humans, Mercier et al. (2013) found an influence of auditory stimulation on the processing of a concurrent visual stimulus in the theta band and low alpha band as well as in the beta band. Based on the finding of transient synchronization of delta and theta band oscillations during audiovisual stimulation in a follow-up intracranial study, Mercier et al. (2015) argued that optimally aligned low-frequency phases promote communication between cortical areas and that stimuli in one modality can reset the phase of an oscillation in a cortical area of the other modality. Taken together, these studies demonstrate cross-modal influences in primary sensory areas. Furthermore, it is likely that low-frequency oscillations mediate this cross-modal influence.

The finding that cross-modal processes influence primary sensory activity via low-frequency oscillatory activity implies a predictive process (Schroeder et al. 2008). In many natural settings, visual information precedes auditory information. For example, in audiovisual speech, the lip movements precede the articulation of phonemes (see Grant and Bernstein, Chap. 3). A mechanism that has been proposed for the transfer of information between cortical areas is neural coherence, as reflected in synchronized oscillatory activity (Fries 2005, 2015). For example, in audiovisual speech, the visual information can be transferred to the auditory cortex (Arnal et al. 2009). It has been proposed that audiovisual perception involves a network of primary visual and auditory areas as well as multisensory regions (Keil et al. 2012; Schepers et al. 2013). This network presumably reflects



reentrant bottom-up and top-down interactions between primary sensory and multisensory areas (Arnal and Giraud 2012).

In summary, there is robust evidence that multisensory integration can be reflected in increased gamma band power and that cross-modal processes can modulate cortical activity in primary sensory areas (van Atteveldt 2014). Furthermore, as hypothesized (Senkowski et al. 2008; Keil and Senkowski 2018), it is likely that information transfer in a network of primary sensory, multisensory, and frontal cortical areas is instantiated through synchronized oscillatory activity.

### 10.3 Principles of Multisensory Integration and Oscillatory Processes

The studies described in Sect. 10.2 suggest a relationship between oscillatory activity and multisensory perception. The current section focuses on the principles of multisensory perception and how these principles relate to oscillatory activity in the auditory system. Based on findings from a wide range of studies, three *principles of multisensory integration* have been established: the spatial principle, the temporal principle, and the principle of inverse effectiveness (Stein and Meredith 1993; Stein et al., 2014). In short, the principles state that multisensory integration is the strongest when the input modalities are (1) spatially concordant, (2) temporally aligned, and (3) when the neural responses to the presented stimuli are weak. In addition to the three principles of multisensory integration, the *modality appropriateness hypothesis* has been proposed (Welch and Warren 1980). The auditory system has a relatively low spatial acuity but high temporal resolution. In contrast, the visual system has a relatively low temporal resolution but a high spatial acuity. Therefore, it has been proposed that audiovisual integration will be governed by the auditory modality in tasks requiring high temporal resolution and by the visual modality in tasks requiring high spatial acuity. The modality appropriateness hypothesis has been extended in a maximum-likelihood-estimation framework, which puts forward the idea that information from each sensory modality is weighted based on its relative reliability (Ernst and Bühlhoff 2004). Therefore, it can be hypothesized that the auditory system will be especially affected by the visual system when a stimulus contains task-relevant spatial information. In turn, it can be hypothesized that the auditory system will prevail in tasks requiring high temporal resolution.

A previous EEG study examined the influence of audiovisual temporal synchrony on evoked gamma band oscillations (Senkowski et al. 2007). In line with the principle of temporal alignment, gamma band power following audiovisual stimulation was strongest when the auditory and visual inputs of an audiovisual stimulus were presented simultaneously. Interestingly, stimuli were perceived as being separated when the auditory input preceded the visual input by more than 100 ms. A later EEG study established the principle of inverse effectiveness for multisensory stimulus processing, as reflected in early event-related potentials (ERPs; Senkowski et al. 2011). In this study, ERP amplitudes were larger for bimodal audiovisual stimulation compared with combined ERPs following unimodal auditory or visual stimulation

but only when the stimuli were presented at a low intensity. Moreover, a local field-potential recording study in monkeys revealed that the principle of spatial alignment and the principle of inverse effectiveness were also reflected in neural oscillations (Lakatos et al. 2007). Interestingly, these authors found that a somatosensory stimulus shifted the neural oscillations in the ipsilateral auditory cortex to a low-excitatory phase and reduced event-related responses. In contrast, a contralateral somatosensory stimulus grouped oscillations around the ideal (i.e., high-excitatory) phase and increased event-related responses. Moreover, in agreement with the principle of inverse effectiveness, the event-related response in the auditory cortex was significantly enhanced when a somatosensory stimulus was added to an auditory stimulus that elicited only a weak response in isolation. Thus, multisensory integration requires flexible neuronal processing (van Atteveldt et al. 2014).

Taken together, a number of general principles for multisensory integration and cross-modal influence have been formulated. Recent electrophysiological studies suggested that these principles are reflected in cortical network dynamics involving neural oscillations.

## 10.4 Influence of Visual Input on Auditory Perception in Audiovisual Speech

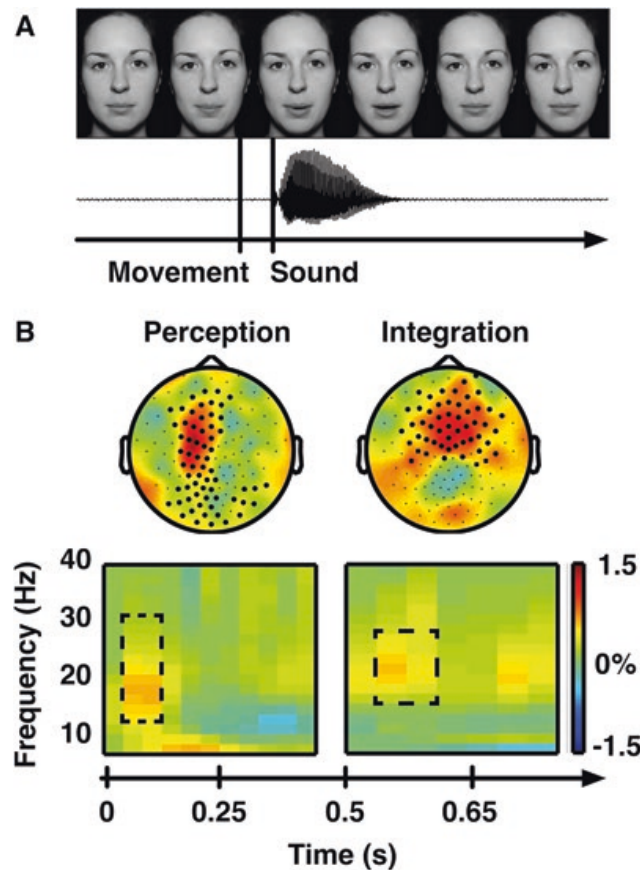
The auditory system, with its high temporal resolution, is very effective in the processing of temporal information. The visual system excels at spatial acuity. As described in Sect. 10.3, the auditory system might be especially affected by the visual system when a stimulus contains important spatial information. An example for the influence of visual information on auditory perception is observing the speaker's mouth movements during audiovisual speech processing. Here, the temporally complex auditory information is processed in the auditory cortex. Concurrently, the variations in the speaker's mouth movements before the utterance of a syllable are processed by the visual system. The lip movements can facilitate the processing of the auditory information. Moreover, rhythmic gestures provide coarse temporal cues for the onset of auditory stimuli (Biau et al. 2015; He et al. 2015). In an early MEG study comparing cortical activity evoked by auditory speech stimuli accompanied by either congruent or incongruent visual stimuli, Sams et al. (1991) showed that incongruent visual information from face movements influences syllable perception. Incongruent audiovisual stimuli elicited a mismatch response in the event-related field. A later MEG study used similar stimuli and found that incongruent audiovisual stimuli elicit stronger gamma band responses than congruent audiovisual stimuli (Kaiser et al. 2005). Interestingly, the study suggested a spatio-temporal hierarchy in the processing of audiovisual stimuli, which starts in posterior parietal cortical areas and spreads to occipital and frontal areas. More recently, Lange et al. (2013) compared neural oscillations to incongruent compared with congruent audiovisual speech stimuli and found increased gamma band and beta band power following congruent speech stimulation, although at a longer latency than Kaiser et al. (2005) found. Whereas the early gamma band power increase in the

study by Kaiser et al. (2005) might reflect the processing of the audiovisual mismatch, the later gamma band power increase following congruent speech found by Lange et al. (2013) might be related to audiovisual integration. The idea of a processing hierarchy has recently received support by an EEG study investigating oscillatory activity during the McGurk illusion (Roa Romero et al. 2015). The McGurk illusion involves incongruent audiovisual speech stimuli, which can be fused into an integrated, subjectively congruent audiovisual percept (McGurk and MacDonald 1976). In this study, incongruent audiovisual syllables, which were perceived as an illusory novel percept, were compared with congruent audiovisual stimuli. Again, incongruent stimuli were associated with increased stimulus processing, in this case reflected in beta band power reduction. These reductions were found at two temporal stages: initially over posterior scalp regions and then over frontal scalp regions (Fig. 10.1). With respect to the cortical areas critical to this process, multistage models of audiovisual integration involving the initial processing of auditory and visual information in primary sensory areas and the subsequent integration in parietal and frontal cortical areas have been recently suggested (Peelle and Sommers 2015; Bizley et al. 2016).

Support for the notion that audiovisual speech perception involves multiple processing stages comes from a MEG study (Arnal et al. 2011). The authors investigated the neuronal signatures of valid or invalid predictions that were based on congruent or incongruent visual speech information, respectively. By correlating the ERP with time-frequency-resolved ITC, the authors found that initial processing of audiovisual speech, independent of stimulus congruence, was reflected in increased delta band and theta band ITC around 100 ms after auditory stimulus onset. Furthermore, valid cross-modal predictions in congruent audiovisual speech stimuli were reflected in increased delta band ITC around 400 ms after auditory stimulus onset. In a case of invalid predictions in incongruent audiovisual stimuli, a later beta band component around 500 ms after auditory stimulus onset was identified. This beta band component presumably reflects the error of the prediction based on the visual syllable. These findings were discussed within the predictive coding framework (Rao and Ballard 1999) to describe cortical oscillatory activity as a mechanism for multisensory integration and temporal predictions. Arnal and Giraud (2012) suggested that temporal regularities induce low-frequency oscillations to align neuronal excitability with the predicted onset of upcoming stimuli (see also Van Wassenhove and Grzeczkowski 2015). Moreover, the authors proposed that top-down signals, which are based on previously available visual information, influence oscillatory activity in neural networks. The top-down signals are primarily conveyed in beta band oscillations, whereas the bottom-up sensory signals originating in primary sensory cortical areas are conveyed in gamma band activity.

In summary, audiovisual integration of speech, as reflected in beta band and gamma band oscillations, requires early and late processing stages. The early stage could reflect sensory processing, whereas the later stages could relate to the formation of a coherent percept. Moreover, it is likely that previously available information is relayed in a feedback manner to sensory cortical areas via beta band oscillations.





**Fig. 10.1** Visual input influences auditory perception in audiovisual speech. (A) In the McGurk illusion, a video of an actor pronouncing a syllable is dubbed with an incongruent auditory stream. The natural mouth movement before the auditory onset provides important information about the upcoming auditory stimulus. In the case of an incongruent auditory stimulus, these predictions are violated, and the nonmatching visual and auditory information are, for specific audiovisual syllable combinations, fused to a novel percept. (B) Formation of a coherent percept presumably occurs in two separate stages. In a first step, auditory and visual stimuli are perceived, processed, and fed forward. In conjunction with predictions based on the mouth movements, the information obtained in the first stage are integrated into a novel, fused percept. In the topographies (*top*), *solid dots* mark significant electrodes and *shadings* represent percentage of signal change from baseline. In the time-frequency plots (*bottom*), *dashed-line boxes* mark significant effects and *shadings* represent percentage of signal change from baseline. Adapted from Roa Romero et al. (2015)

## 10.5 Influence of Auditory Input on Visual Perception in Audiovisual Illusions

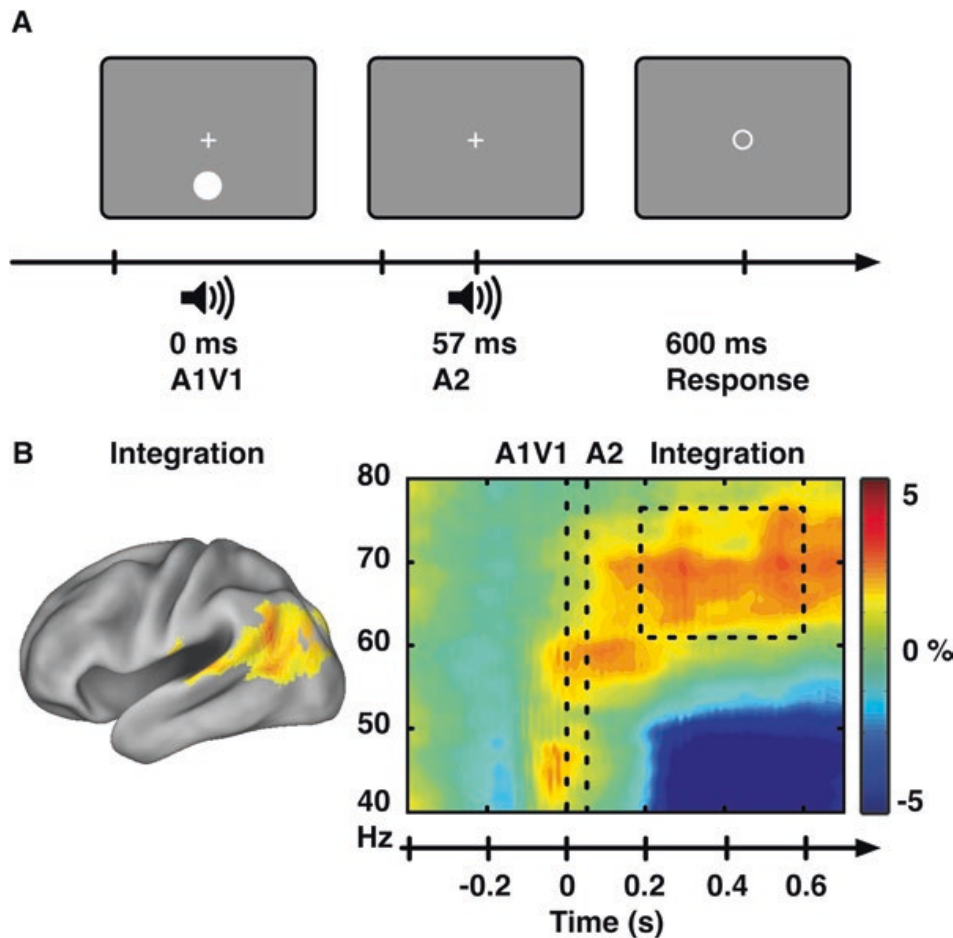
In Sect. 10.3, it was hypothesized that the auditory system prevails in tasks requiring high temporal resolution. Examples for the influence of auditory information on visual processing are visual illusions induced by concurrently presented auditory stimuli. For example, Shams et al. (2000) have shown that a short visual flash could be perceived as multiple flashes when it is accompanied by multiple short auditory noise bursts. The perception of the sound-induced flash illusion (SIFI) is

accompanied by increased ERP amplitudes over occipital electrodes (Shams et al. 2001). Notably, the SIFI only occurs if auditory and visual stimuli are presented above the detection threshold with a stimulus onset asynchrony of up to 115 ms. An interesting phenomenon in audiovisual illusions is alternating perception. In the McGurk illusion, as well as in the SIFI, individuals typically perceive the illusion in some but not all trials, even though the input is always the same. This allows for direct comparisons of the neural responses at varying perceptions but under identical audiovisual stimulation (Keil et al. 2012, 2014a). Analyzing trials in which the SIFI was perceived and trials in which no illusion occurred, Bhattacharya et al. (2002) found that the perception of the illusion is marked by a strong early gamma band power increase as well as a sustained cross-modal interaction in occipital electrodes. Mishra et al. (2007) replicated this finding in a direct comparison between oscillatory activity in illusion and nonillusion trials. Again, the perception of the illusion was marked by an increase in gamma band power in occipital electrodes. Moreover, the authors were able to distinguish an early and a late phase of audiovisual integration. A recent study replicated the role of gamma band power for the perception of the illusion and identified the left superior temporal gyrus (STG) as well as the extrastriatal cortex as putative cortical generators (Balz et al. 2016). In this study, multisensory integration of the audiovisual stimuli was marked by increased gamma band power. Importantly, the individual gamma band power was positively correlated to the SIFI rate, which represents an individual's likelihood to perceive the illusion. By additionally using magnetic resonance spectroscopy to measure neurotransmitter metabolite concentrations, it was observed that the GABA level in the STG modulated the relationship between gamma band power and the SIFI rate (Fig. 10.2). This finding points toward an influence of global cortical states on multisensory perception because the GABA concentration was recorded during rest.

Taken together, the cross-modal influence underlying the influence of auditory information on visual processing is reflected in increased induced gamma band power, which relates to the likelihood to perceive the sound induced flash illusion.

## 10.6 Anticipatory Activity Influences Cross-Modal Influence

In pioneering EEG research, Davis and Davis (1936) were early to suggest that the pattern and degree of cortical activity might be modified by various physiological and psychological states. In support of this idea, Lindsley (1952) demonstrated that the amplitude of auditory evoked potentials varies systematically with an underlying low-frequency phase. In the last decades, a number of studies have supported the assumption that cortical activity in response to a stimulus is influenced by the phase of ongoing oscillatory activity before the stimulus onset (Busch et al. 2009; Keil et al. 2014b). In addition to the phase, the power of oscillatory activity before stimulus onset also plays a role in perceptual processes (Van Dijk et al. 2008; Romei



**Fig. 10.2** Auditory input influences visual perception in audiovisual illusions. **(A)** In the sound-induced flash illusion, a single visual stimulus (V1) is paired with two consecutive auditory stimuli (A1 and A2). Subjects are asked to report the number of perceived visual stimuli. In approximately half of the trials, subjects reported an illusory perception of two visual stimuli. **(B)** After an initial perception of auditory and visual stimuli in primary sensory areas, incongruent information from both modalities is integrated to an illusory percept as reflected in gamma band power in the left superior temporal gyrus (STG). *Left: the shaded area on the cortical surface represents an increase relative to baseline for poststimulus gamma band power. Right: shadings represent percentage of signal change from baseline; vertical dashed lines indicate the onset of A1, V1, and A2; dashed-line box indicates the time-frequency window marking increased gamma band power during multisensory integration of the audiovisual stimuli.* Adapted from Balz et al. (2016)

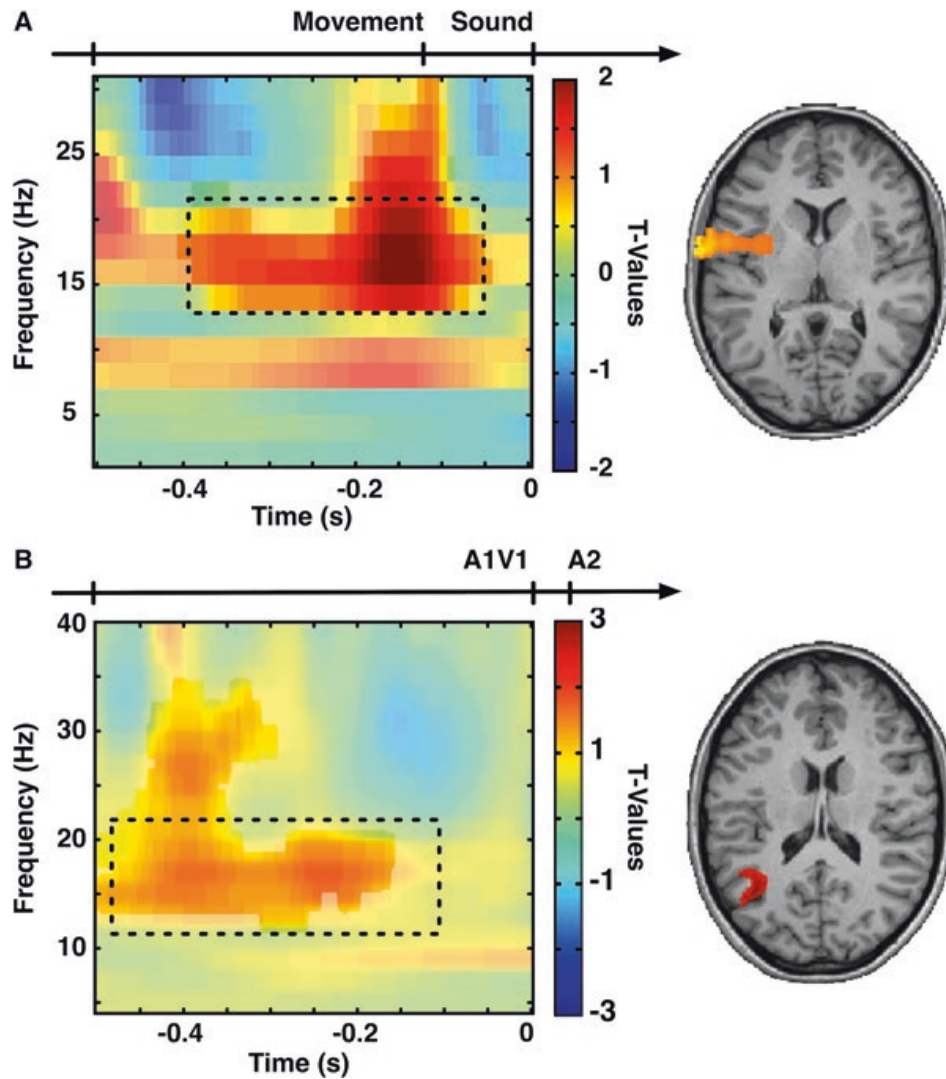
et al. 2010). Moreover, network processes, as reflected in functional connectivity, also influence perception (Weisz et al. 2014; Leske et al. 2015). Thus far, the vast majority of studies have investigated the influence of prestimulus activity on unisensory processing.

More recently, a number of studies have started to suggest that oscillatory activity before the stimulus onset also influences the processing and perception of multisensory stimuli (Pomper et al. 2015; Keil et al. 2016). For instance, predictions based on visual information before auditory stimulus onset can modulate audiovisual integration (Arnal et al. 2011). In a similar vein, expectations based on auditory

cues modulate ongoing oscillatory activity in the visual and somatosensory cortices (Pomper et al. 2015) as well as functional connectivity networks comprising frontal, parietal, and primary sensory areas (Leonardelli et al. 2015; Keil et al. 2016). Ongoing fluctuations of local cortical oscillations and functional connectivity networks have been found to also influence multisensory perception when there are no specific predictions and expectations (Lange et al. 2011; Keil et al. 2012). For example, one study compared oscillatory neural activity before stimulus onset between trials in which incongruent audiovisual speech stimuli were perceived as the McGurk illusion with trials in which either the auditory or the visual input dominated the percept (Keil et al. 2012). A main finding of this study was that prestimulus beta band power and functional connectivity influenced upcoming perception (Fig. 10.3A). More specifically, beta band power was increased in the left STG, precuneus, and middle frontal gyrus before stimulus onset, in trials in which the illusion was perceived. Interestingly, before the perception of the illusion, the left STG was decoupled from cortical areas associated with face (i.e., fusiform gyrus) or voice (i.e., Brodmann area 22) processing. Similar results were obtained in a study comparing incongruent audiovisual trials in which the SIFI was perceived and trials where the SIFI was not perceived (Keil et al. 2014a). The study revealed increased beta band power before the perception of the illusion (Fig. 10.3B). In addition, the left STG was coupled to left auditory cortical areas but decoupled from visual cortical areas before the illusion. Furthermore, the stronger the functional connectivity between the left STG and the left auditory cortex, the higher the likelihood of an illusion. These data provide strong evidence for a role of the left STG in audiovisual integration. In case of degraded bottom-up input, the formation of a fused percept is supported by strong beta band power (see also Schepers et al. 2013). In case of imbalanced reliability of the bottom-up input of various modalities, information from one modality can dominate the subjective percept.

Two recent studies using the SIFI further highlighted the role of low-frequency oscillations for audiovisual perception (Cecere et al. 2015; Keil and Senkowski 2017). Cecere et al. (2015) found a negative correlation between the participants' individual alpha band frequency (IAF) and their illusion rate, which indicates that alpha band oscillations provide a temporal window in which the cross-modal influence could induce an illusion. Underscoring the role of low-frequency oscillations for cross-modal influence, the authors also found that increasing the IAF using transcranial direct current stimulation reduces the probability of an illusion perception, whereas a reduction of the IAF had the opposite effect. Recently, Keil and Senkowski (2017) corroborated the relationship between the IAF and the SIFI perception rate and localized this effect to the occipital cortex.

Taken together, local cortical activity and the information transfer between cortical network nodes critically influence the processing and perception of multisensory stimuli. Furthermore, functional connectivity networks seem to mediate how information is relayed between unisensory, multisensory, and higher order cortical areas. Hence, there is strong evidence that ongoing oscillatory activity influences unisensory and multisensory perception.



**Fig. 10.3** Anticipatory activity influences cross-modal influence. Cortical activity before the onset of audiovisual stimulation influences perception of the McGurk illusion (A) and perception of the sound-induced flash illusion (B). The cross-modal influence at the behavioral level is opposite between the two illusions. However, empirical data show that similar cortical processes (i.e., increased beta band power in the left STG) influence upcoming perception in both illusions. *Left: shadings* represent results ( $T$  values) of the statistical comparison via  $t$ -tests between trials with and without the illusion; *dashed-line boxes* indicate the time-frequency window marking increased beta band power prior to multisensory integration of the audiovisual stimuli. *Right: shaded areas* in the brains represent results ( $T$  values) of the statistical comparison via  $t$ -tests between trials with and without the illusion. Adapted from Keil et al. (2012, 2014a)

## 10.7 Summary and Open Questions

This chapter reviewed empirical findings on the neural mechanisms underlying multisensory processing, with a focus on oscillatory activity. Based on the available findings, it can be postulated that multisensory processing and perception rely on a complex and dynamic cross-frequency interaction pattern within widespread neural



networks (Keil and Senkowski 2018). Currently available evidence suggests a hierarchical interplay between low-frequency phase and high-frequency power during multisensory processing; low-frequency oscillations presumably provide temporal windows of integration for the cross-modal influence. Successful multisensory integration is subsequently reflected in increased high-frequency power.

### ***10.7.1 Low-Frequency Oscillations Transfer Feedback Information and Cross-Modal Influence***

An increasing number of studies suggest that low-frequency oscillations (delta, theta, and alpha bands), might serve as a mechanism to control local cortical activity. The phase of these oscillations has been shown to modulate stimulus evoked activity and perception (Busch et al. 2009; Keil et al. 2014b). Moreover, as demonstrated in a number of studies, low-frequency oscillations seem to reflect cross-modal influences (Lakatos et al. 2007; Mercier et al. 2015). In addition, prior information based on stimulus properties influences local cortical activity (Roa Romero et al. 2016). The modulating influence can be found in primary sensory areas (e.g., visual cortex) as well as higher order areas (e.g., frontal cortex). Thus, information that is transferred from frontal cortical areas to multisensory and unisensory cortical areas can represent abstract top-down processes, such as attention (Keil et al. 2016). Additionally, information that is transferred between these cortical areas can also represent stimulus properties, such as timing, rhythmicity, or space (Lakatos et al. 2007; Mercier et al. 2015).

### ***10.7.2 High-Frequency Oscillations Reflect Perception and Integration***

Oscillatory activity above 12 Hz (i.e., in the frequency of the beta band and gamma band) has been implied to reflect perception and stimulus integration (Senkowski et al. 2008). Furthermore, it has been shown that multisensory integration is reflected in increased gamma band power in traditional multisensory cortical areas, such as the STG (Balz et al. 2016). The analyses of beta band and gamma band power modulations during multisensory stimulus processing revealed different stages of multisensory integration (Roa Romero et al. 2015; Bizley et al. 2016). In audiovisual speech perception, stimuli are processed and different input streams are compared for congruence at a putative early stage. In a later stage, the different input streams are combined and, in case of incongruence, resolved to a subjectively congruent percept (Peelle and Sommers 2015).

### ***10.7.3 Functional Connectivity Guides Integration***

Whereas perception and multisensory integration are reflected in high-frequency power, both processes are modulated by a low-frequency oscillatory phase. Therefore, modulatory information has to be transferred within functional connectivity networks encompassing primary sensory areas, traditional multisensory areas, and higher order frontal areas (Senkowski et al. 2008; Keil and Senkowski 2018). A number of studies have shown that feedback information is conveyed in alpha band and beta band functional connectivity. Furthermore, cue-induced attention or expectations also modulate low-frequency functional connectivity (Keil et al. 2016).

### ***10.7.4 Open Questions***

In the last decade, research on the neural mechanisms underlying the integration and perception of multisensory information as well as on the role of oscillatory processes therein has made tremendous progress. It has been found that the effects in neural oscillations go along with the principles of multisensory integration, but several open questions remain to be answered. For instance, the temporal evolution of multisensory perception and integration is still not well understood. A number of studies have shown that multisensory perception, as reflected in oscillatory activity, requires multiple processing stages. However, it is so far unknown which cortical nodes are active at a given latency. Future studies could integrate recent progress in technical methods and analytical approaches to analyze time-frequency-resolved oscillatory activity on the level of cortical sources. This will help to elucidate the progression of multisensory stimulus processing. Another open question pertains to the role of attention, predictions, and expectations for multisensory perception and the underlying neural oscillatory patterns. Recent studies have highlighted the role of prior expectations and attention for prestimulus oscillations in multisensory paradigms. Yet it remains to be elucidated how cognitive processes influence multisensory processing, how they influence network architecture, and in which oscillatory signatures they are reflected. Future studies should exploit the full spectrum of information available from electrophysiological data to capture the complex network processes underlying the integration of multisensory information as well as how cognitive processes modulate neural oscillations in these networks. The functional significance of the separate network nodes for multisensory perception also has not been fully clarified. Electrophysiological as well as functional imaging studies have identified a number of cortical regions involved in multisensory perception, but these studies have mostly used correlation approaches. Future studies should therefore turn to more causal approaches in which stimulation can be used to directly test the functional role of cortical areas. For instance, transcranial magnetic stimulation could be used to apply a so-called virtual lesion to selectively interrupt activity

within a cortical area to study how cortical activity, multisensory integration, and perception are influenced. In addition, entrainment of cortical networks via transcranial direct/alternating current stimulation could be used to obtain information on the role of specific oscillatory frequencies for multisensory integration and perception. In conclusion, the studies reviewed above suggest that multisensory perception relies on dynamic neural networks in which information is transferred through oscillatory activity. An important endeavor will be to more precisely study the functional roles of the different frequency bands and their interplay for multisensory integrative processing.

**Compliance with Ethics Requirements** Julian Keil declares that he has no conflict of interest.

Daniel Senkowski declares that he has no conflict of interest.

## References

- Adrian, E. D. (1942). Olfactory reactions in the brain of the hedgehog. *The Journal of Physiology*, 100(4), 459–473.
- Arnal, L. H., & Giraud, A.-L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, 16(7), 390–398. <https://doi.org/10.1016/j.tics.2012.05.003>.
- Arnal, L. H., Morillon, B., Kell, C. A., & Giraud, A.-L. (2009). Dual neural routing of visual facilitation in speech processing. *The Journal of Neuroscience*, 29(43), 13445–13453. <https://doi.org/10.1523/JNEUROSCI.3194-09.2009>.
- Arnal, L. H., Wyart, V., & Giraud, A.-L. (2011). Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nature Neuroscience*, 14(6), 797–801. <https://doi.org/10.1038/nn.2810>.
- Balz, J., Keil, J., Roa Romero, Y., Meckle, R., Schubert, F., Aydin, S., Ittermann, B., Gallinat, J., & Senkowski, D. (2016). GABA concentration in superior temporal sulcus predicts gamma power and perception in the sound-induced flash illusion. *NeuroImage*, 125, 724–730. <https://doi.org/10.1016/j.neuroimage.2015.10.087>.
- Berger, H. (1929). Über das elektroenkephalogramm des menschen. *Archiv für Psychiatrie und Nervenkrankheiten*, 87, 527–570.
- Bhattacharya, J., Shams, L., & Shimojo, S. (2002). Sound-induced illusory flash perception: Role of gamma band responses. *NeuroReport*, 13(14), 1727–1730.
- Biau, E., Torralba, M., Fuentemilla, L., de Diego Balaguer, R., & Soto-Faraco, S. (2015). Speaker's hand gestures modulate speech perception through phase resetting of ongoing neural oscillations. *Cortex*, 68, 76–85. <https://doi.org/10.1016/j.cortex.2014.11.018>.
- Birbaumer, N., Elbert, T., Canavan, A. G., & Rockstroh, B. (1990). Slow potentials of the cerebral cortex and behavior. *Physiological Reviews*, 70(1), 1–41.
- Bizley, J. K., Maddox, R. K., & Lee, A. K. C. (2016). Defining auditory-visual objects: Behavioral tests and physiological mechanisms. *Trends in Neurosciences*, 39(2), 74–85. <https://doi.org/10.1016/j.tins.2015.12.007>.
- Bremer, F. (1958). Cerebral and cerebellar potentials. *Physiological Reviews*, 38(3), 357–388.
- Bressler, S. L. (1990). The gamma wave: A cortical information carrier? *Trends in Neurosciences*, 13(5), 161–162.
- Busch, N. A., Dubois, J., & Vanrullen, R. (2009). The phase of ongoing EEG oscillations predicts visual perception. *The Journal of Neuroscience*, 29(24), 7869–7876. <https://doi.org/10.1523/JNEUROSCI.0113-09.2009>.

- Buzsáki, G., & Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science*, 304(5679), 1926–1929. <https://doi.org/10.1126/science.1099745>.
- Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, 18(8), 414–421. <https://doi.org/10.1016/j.tics.2014.04.012>.
- Cecere, R., Rees, G., & Romei, V. (2015). Individual differences in alpha frequency drive crossmodal illusory perception. *Current Biology*, 25(2), 231–235. <https://doi.org/10.1016/j.cub.2014.11.034>.
- Chatrian, G. E., Bickford, R. G., & Uihlein, A. (1960). Depth electrographic study of a fast rhythm evoked from the human calcarine region by steady illumination. *Electroencephalography and Clinical Neurophysiology*, 12, 167–176.
- Cheron, G., Cebolla, A., De Saedeleer, C., Bengoetxea, A., Leurs, F., Leroy, A., & Dan, B. (2007). Pure phase-locking of beta/gamma oscillation contributes to the N30 frontal component of somatosensory evoked potentials. *BMC Neuroscience*, 8(1), 75. <https://doi.org/10.1186/1471-2202-8-75>.
- Davis, H., & Davis, P. A. (1936). Action potentials of the brain: In normal persons and in normal states of cerebral activity. *Archives of Neurology and Psychiatry*, 36(6), 1214–1224.
- Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience*, 11(2), 114–126. <https://doi.org/10.1038/nrn2762>.
- Driver, J., & Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on “sensory-specific” brain regions, neural responses, and judgments. *Neuron*, 57(1), 11–23. <https://doi.org/10.1016/j.neuron.2007.12.013>.
- Engel, A. K., & Fries, P. (2010). Beta-band oscillations—signalling the status quo? *Current Opinion in Neurobiology*, 20(2), 156–165. <https://doi.org/10.1016/j.conb.2010.02.015>.
- Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top-down processing. *Nature Reviews Neuroscience*, 2(10), 704–716. <https://doi.org/10.1038/35094565>.
- Ernst, M. O., & Bühlhoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, 8(4), 162–169. <https://doi.org/10.1016/j.tics.2004.02.002>.
- Fries, P. (2005). A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, 9(10), 474–480. <https://doi.org/10.1016/j.tics.2005.08.011>.
- Fries, P. (2015). Rhythms for cognition: Communication through coherence. *Neuron*, 88(1), 220–235. <https://doi.org/10.1016/j.neuron.2015.09.034>.
- Ghazanfar, A., & Schroeder, C. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences*, 10(6), 278–285. <https://doi.org/10.1016/j.tics.2006.04.008>.
- He, Y., Gebhardt, H., Steines, M., Sammer, G., Kircher, T., Nagels, A., & Straube, B. (2015). The EEG and fMRI signatures of neural integration: An investigation of meaningful gestures and corresponding speech. *Neuropsychologia*, 72, 27–42. <https://doi.org/10.1016/j.neuropsychologia.2015.04.018>.
- Herrmann, C. S., Mecklinger, A., & Pfeifer, E. (1999). Gamma responses and ERPs in a visual classification task. *Clinical Neurophysiology*, 110(4), 636–642.
- Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: Gating by inhibition. *Frontiers in Human Neuroscience*, 4, 186. <https://doi.org/10.3389/fnhum.2010.00186>.
- Kaiser, J., Hertrich, I., Ackermann, H., Mathiak, K., & Lutzenberger, W. (2005). Hearing lips: Gamma-band activity during audiovisual speech perception. *Cerebral Cortex*, 15(5), 646–653. <https://doi.org/10.1093/cercor/bhh166>.
- Kayser, C., Petkov, C. I., & Logothetis, N. K. (2008). Visual modulation of neurons in auditory cortex. *Cerebral Cortex*, 18(7), 1560–1574. <https://doi.org/10.1093/cercor/bhm187>.
- Keil, J., & Senkowski, D. (2017). Individual alpha frequency relates to the sound-induced flash illusion. *Multisensory Research*, 30(6), 565–578. <https://doi.org/10.1163/22134808-00002572>.
- Keil, J., & Senkowski, D. (2018). Neural oscillations orchestrate multisensory processing. *Neuroscientist*, 24(6), 609–626. <https://journals.sagepub.com/doi/10.1177/10738584187553522>.

- Keil, J., Müller, N., Ihssen, N., & Weisz, N. (2012). On the variability of the McGurk effect: Audiovisual integration depends on prestimulus brain states. *Cerebral Cortex*, 22(1), 221–231. <https://doi.org/10.1093/cercor/bhr125>.
- Keil, J., Müller, N., Hartmann, T., & Weisz, N. (2014a). Prestimulus beta power and phase synchrony influence the sound-induced flash illusion. *Cerebral Cortex*, 24(5), 1278–1288. <https://doi.org/10.1093/cercor/bhs409>.
- Keil, J., Timm, J., SanMiguel, I., Schulz, H., Obleser, J., & Schönwiesner, M. (2014b). Cortical brain states and corticospinal synchronization influence TMS-evoked motor potentials. *Journal of Neurophysiology*, 111(3), 513–519. <https://doi.org/10.1152/jn.00387.2013>.
- Keil, J., Pomper, U., & Senkowski, D. (2016). Distinct patterns of local oscillatory activity and functional connectivity underlie intersensory attention and temporal prediction. *Cortex*, 74, 277–288. <https://doi.org/10.1016/j.cortex.2015.10.023>.
- Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews*, 29, 169–195.
- Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences*, 16(12), 606–617. <https://doi.org/10.1016/j.tics.2012.10.007>.
- Lakatos, P., Chen, C.-M., O'Connell, M. N., Mills, A., & Schroeder, C. E. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron*, 53(2), 279–292. <https://doi.org/10.1016/j.neuron.2006.12.011>.
- Lange, J., Oostenveld, R., & Fries, P. (2011). Perception of the touch-induced visual double-flash illusion correlates with changes of rhythmic neuronal activity in human visual and somatosensory areas. *NeuroImage*, 54(2), 1395–1405. <https://doi.org/10.1016/j.neuroimage.2010.09.031>.
- Lange, J., Christian, N., & Schnitzler, A. (2013). Audio-visual congruency alters power and coherence of oscillatory activity within and between cortical areas. *NeuroImage*, 79, 111–120. <https://doi.org/10.1016/j.neuroimage.2013.04.064>.
- Lange, J., Keil, J., Schnitzler, A., Van Dijk, H., & Weisz, N. (2014). The role of alpha oscillations for illusory perception. *Behavioural Brain Research*, 271, 294–301. <https://doi.org/10.1016/j.bbr.2014.06.015>.
- Leonardelli, E., Braun, C., Weisz, N., Lithari, C., Occelli, V., & Zampini, M. (2015). Prestimulus oscillatory alpha power and connectivity patterns predispose perceptual integration of an audio and a tactile stimulus. *Human Brain Mapping*, 36(9), 3486–3498. <https://doi.org/10.1002/hbm.22857>.
- Leske, S., Ruhnau, P., Frey, J., Lithari, C., Müller, N., Hartmann, T., & Weisz, N. (2015). Prestimulus network integration of auditory cortex predisposes near-threshold perception independently of local excitability. *Cerebral Cortex*, 25, 4898–4907. <https://doi.org/10.1093/cercor/bhv212>.
- Lindsley, D. B. (1952). Psychological phenomena and the electroencephalogram. *Electroencephalography and Clinical Neurophysiology*, 4(4), 443–456.
- Lopes da Silva, F. (1991). Neural mechanisms underlying brain waves: From neural membranes to networks. *Electroencephalography and Clinical Neurophysiology*, 79(2), 81–93.
- Luck, S. J. (2014). *An introduction to the event-related potential technique*. Cambridge, MA: The MIT Press.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264(5588), 746–748.
- Mercier, M. R., Foxe, J. J., Fiebelkorn, I. C., Butler, J. S., Schwartz, T. H., & Molholm, S. (2013). Auditory-driven phase reset in visual cortex: Human electrocorticography reveals mechanisms of early multisensory integration. *NeuroImage*, 79, 19–29. <https://doi.org/10.1016/j.neuroimage.2013.04.060>.
- Mercier, M. R., Molholm, S., Fiebelkorn, I. C., Butler, J. S., Schwartz, T. H., & Foxe, J. J. (2015). Neuro-oscillatory phase alignment drives speeded multisensory response times: An electrocorticographic investigation. *The Journal of Neuroscience*, 35(22), 8546–8557. <https://doi.org/10.1523/JNEUROSCI.4527-14.2015>.
- Michalareas, G., Vezoli, J., van Pelt, S., Schoffelen, J.-M., Kennedy, H., & Fries, P. (2016). Alpha-beta and gamma rhythms subserve feedback and feedforward influences among human visual cortical areas. *Neuron*, 89(2), 384–397. <https://doi.org/10.1016/j.neuron.2015.12.018>.



- Mishra, J., Martinez, A., Sejnowski, T. J., & Hillyard, S. A. (2007). Early cross-modal interactions in auditory and visual cortex underlie a sound-induced visual illusion. *The Journal of Neuroscience*, 27(15), 4120–4131. <https://doi.org/10.1523/JNEUROSCI.4912-06.2007>.
- Mitra, P. P., & Pesaran, B. (1999). Analysis of dynamic brain imaging data. *Biophysical Journal*, 76(2), 691–708. [https://doi.org/10.1016/S0006-3495\(99\)77236-X](https://doi.org/10.1016/S0006-3495(99)77236-X).
- Pantev, C., Makeig, S., Hoke, M., Galambos, R., Hampson, S., & Gallen, C. (1991). Human auditory evoked gamma-band magnetic fields. *Proceedings of the National Academy of Sciences of the United States of America*, 88(20), 8996–9000.
- Peelle, J. E., & Sommers, M. S. (2015). Prediction and constraint in audiovisual speech perception. *Cortex*, 68, 169–181. <https://doi.org/10.1016/j.cortex.2015.03.006>.
- Penttonen, M., & Buzsáki, G. (2003). Natural logarithmic relationship between brain oscillators. *Thalamus & Related Systems*, 2(2), 145–152. [https://doi.org/10.1016/S1472-9288\(03\)00007-4](https://doi.org/10.1016/S1472-9288(03)00007-4).
- Pfurtscheller, G. (1992). Event-related synchronization (ERS): An electrophysiological correlate of cortical areas at rest. *Electroencephalography and Clinical Neurophysiology*, 83(1), 62–69.
- Pomper, U., Brincker, J., Harwood, J., Prikhodko, I., & Senkowski, D. (2014). Taking a call is facilitated by the multisensory processing of smartphone vibrations, sounds, and flashes. *PLoS One*, 9(8), e103238. <https://doi.org/10.1371/journal.pone.0103238.s002>.
- Pomper, U., Keil, J., Foxe, J. J., & Senkowski, D. (2015). Intersensory selective attention and temporal orienting operate in parallel and are instantiated in spatially distinct sensory and motor cortices. *Human Brain Mapping*, 36(8), 3246–3259. <https://doi.org/10.1002/hbm.22845>.
- Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1), 79–87. <https://doi.org/10.1038/4580>.
- Ray, W. J., & Cole, H. W. (1985). EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes. *Science*, 228(4700), 750–752.
- Roa Romero, Y., Senkowski, D., & Keil, J. (2015). Early and late beta-band power reflect audiovisual perception in the McGurk illusion. *Journal of Neurophysiology*, 113(7), 2342–2350. <https://doi.org/10.1152/jn.00783.2014>.
- Roa Romero, Y., Keil, J., Balz, J., Gallinat, J., & Senkowski, D. (2016). Reduced frontal theta oscillations indicate altered crossmodal prediction error processing in schizophrenia. *Journal of Neurophysiology*, 116(3), 1396–1407. <https://doi.org/10.1152/jn.00096.2016>.
- Rockstroh, B., Müller, M., Cohen, R., & Elbert, T. (1992). Probing the functional brain state during P300-evocation. *Journal of Psychophysiology*, 6, 175–184.
- Romei, V., Gross, J., & Thut, G. (2010). On the role of prestimulus alpha rhythms over occipitoparietal areas in visual input regulation: Correlation or causation? *The Journal of Neuroscience*, 30(25), 8692–8697. <https://doi.org/10.1523/JNEUROSCI.0160-10.2010>.
- Sakowitz, O. W., Quiroga, R. Q., Schürmann, M., & Başar, E. (2001). Bisensory stimulation increases gamma-responses over multiple cortical regions. *Cognitive Brain Research*, 11(2), 267–279. [https://doi.org/10.1016/S0926-6410\(00\)00081-1](https://doi.org/10.1016/S0926-6410(00)00081-1).
- Sams, M., Aulanko, R., Hämäläinen, M., Hari, R., Lounasmaa, O. V., Lu, S. T., & Simola, J. (1991). Seeing speech: Visual information from lip movements modifies activity in the human auditory cortex. *Neuroscience Letters*, 127(1), 141–145.
- Sauseng, P., Klimesch, W., Schabus, M., & Doppelmayr, M. (2005). Fronto-parietal EEG coherence in theta and upper alpha reflect central executive functions of working memory. *International Journal of Psychophysiology*, 57(2), 97–103. <https://doi.org/10.1016/j.ijpsycho.2005.03.018>.
- Schepers, I. M., Schneider, T. R., Hipp, J. F., Engel, A. K., & Senkowski, D. (2013). Noise alters beta-band activity in superior temporal cortex during audiovisual speech processing. *NeuroImage*, 70, 101–112. <https://doi.org/10.1016/j.neuroimage.2012.11.066>.
- Schroeder, C. E., & Foxe, J. (2005). Multisensory contributions to low-level, “unisensory” processing. *Current Opinion in Neurobiology*, 15(4), 454–458. <https://doi.org/10.1016/j.conb.2005.06.008>.
- Schroeder, C. E., Lakatos, P., Kajikawa, Y., Partan, S., & Puce, A. (2008). Neuronal oscillations and visual amplification of speech. *Trends in Cognitive Sciences*, 12(3), 106–113. <https://doi.org/10.1016/j.tics.2008.01.002>.

- Senkowski, D., Talsma, D., Herrmann, C. S., & Woldorff, M. G. (2005). Multisensory processing and oscillatory gamma responses: Effects of spatial selective attention. *Experimental Brain Research*, 166(3-4), 411–426. <https://doi.org/10.1007/s00221-005-2381-z>.
- Senkowski, D., Talsma, D., Grigutsch, M., Herrmann, C. S., & Woldorff, M. G. (2007). Good times for multisensory integration: Effects of the precision of temporal synchrony as revealed by gamma-band oscillations. *Neuropsychologia*, 45(3), 561–571. <https://doi.org/10.1016/j.neuropsychologia.2006.01.013>.
- Senkowski, D., Schneider, T., Foxe, J., & Engel, A. (2008). Crossmodal binding through neural coherence: Implications for multisensory processing. *Trends in Neurosciences*, 31(8), 401–409. <https://doi.org/10.1016/j.tins.2008.05.002>.
- Senkowski, D., Saint-Amour, D., Höfle, M., & Foxe, J. J. (2011). Multisensory interactions in early evoked brain activity follow the principle of inverse effectiveness. *NeuroImage*, 56(4), 2200–2208. <https://doi.org/10.1016/j.neuroimage.2011.03.075>.
- Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions. What you see is what you hear. *Nature*, 408(6814), 788. <https://doi.org/10.1038/35048669>.
- Shams, L., Kamitani, Y., Thompson, S., & Shimojo, S. (2001). Sound alters visual evoked potentials in humans. *NeuroReport*, 12(17), 3849–3852.
- Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. Cambridge, MA: The MIT Press.
- Stein, B. E., Stanford, T. R., & Rowland, B. A. (2014). Development of multisensory integration from the perspective of the individual neuron. *Nature Reviews Neuroscience*, 15(8), 520–535. <https://doi.org/10.1038/nrn3742>.
- Steriade, M. (2001). Impact of network activities on neuronal properties in corticothalamic systems. *Journal of Neurophysiology*, 86(1), 1–39.
- Sumby, W. H., & Polack, I. (1954). Perceptual amplification of speech sounds by visual cues. *The Journal of the Acoustical Society of America*, 26, 212–215.
- Tallon-Baudry, C., & Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends in Cognitive Sciences*, 3(4), 151–162.
- van Atteveldt, N., Murray, M. M., Thut, G., & Schroeder, C. E. (2014). Multisensory integration: Flexible use of general operations. *Neuron*, 81(6), 1240–1253. <https://doi.org/10.1016/j.neuron.2014.02.044>.
- Van Dijk, H., Schoffelen, J.-M., Oostenveld, R., & Jensen, O. (2008). Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. *The Journal of Neuroscience*, 28(8), 1816–1823. <https://doi.org/10.1523/JNEUROSCI.1853-07.2008>.
- van Wassenhove, V., & Grzeczkowski, L. (2015). Visual-induced expectations modulate auditory cortical responses. *Frontiers in Neuroscience*, 9, 11. <https://doi.org/10.3389/fnins.2015.00011>.
- Walter, W. G. (1936). The location of cerebral tumours by electro-encephalography. *The Lancet*, 228(5893), 305–308.
- Walter, D. O., Rhodes, J. M., Brown, D., & Adey, W. R. (1966). Comprehensive spectral analysis of human EEG generators in posterior cerebral regions. *Electroencephalography and Clinical Neurophysiology*, 20(3), 224–237.
- Wang, X.-J. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. *Physiological Reviews*, 90, 1195–1268. <https://doi.org/10.1152/physrev.00035.2008>.
- Weisz, N., Wühle, A., Monittola, G., Demarchi, G., Frey, J., Popov, T., & Braun, C. (2014). Prestimulus oscillatory power and connectivity patterns predispose conscious somatosensory perception. *Proceedings of the National Academy of Sciences of the United States of America*, 111(4), E417–E425. <https://doi.org/10.1073/pnas.1317267111>.
- Welch, R. B., & Warren, D. H. (1980). Immediate perceptual responses to intersensory discrepancy. *Psychological Bulletin*, 88(3), 638–667.