## Neural Oscillations Orchestrate Multisensory Processing

The Neuroscientist 2018, Vol. 24(6) 609–626 © The Author(s) 2018 Article reuse guidelines: sagepub.com/journals-permissions DOI: 10.1177/1073858418755352 journals.sagepub.com/home/nro SAGE

## Julian Keil<sup>1,2</sup> and Daniel Senkowski<sup>2</sup>

#### Abstract

At any given moment, we receive input through our different sensory systems, and this information needs to be processed and integrated. Multisensory processing requires the coordinated activity of distinct cortical areas. Key mechanisms implicated in these processes include local neural oscillations and functional connectivity between distant cortical areas. Evidence is now emerging that neural oscillations in distinct frequency bands reflect different mechanisms of multisensory processing. Moreover, studies suggest that aberrant neural oscillations contribute to multisensory processing deficits in clinical populations, such as schizophrenia. In this article, we review recent literature on the neural mechanisms underlying multisensory processing, focusing on neural oscillations. We derive a framework that summarizes findings on (1) stimulus-driven multisensory processing, (2) the influence of top-down information on multisensory processing, and (3) the role of predictions for the formation of multisensory processing. These processes can act in parallel and are essential for multisensory processing.

#### **Keywords**

multisensory integration, crossmodal, audiovisual, top-down, bottom-up, attention, functional connectivity, electrophysiology

## Introduction

## **Oscillatory Processes and Perception**

Stimuli that reach our nervous system are initially processed in lower-order cortical areas specialized for detecting certain types of sensory signals. These stimuli can then be integrated across distributed cortical networks comprising lower-order and higher-order areas. One mechanism that has been proposed to underlie information integration across distributed cortical networks is transient synchronization of neural oscillations (Box 1; Fries 2015). Synchronization of neural oscillations has also been proposed as a mechanism for the integration of information across sensory modalities (Senkowski and others 2008; van Atteveldt and others 2014). Moreover, Siegel and others (2012) hypothesized that distinct spectral profiles index different local cortical nodes involved in information encoding and long-range integrative processing (Box 2). This hypothesis is based on the observation that bottom-up processing engages local networks in the gamma-band (>30 Hz), whereas top-down control through long-range integrative processing engages lower frequency bands (<30 Hz). It is likely that local networks can concurrently engage in bottom-up and top-down processes. When stimuli are simultaneously presented in

different sensory modalities, these stimuli can either be processed separately or integrated. Thus, multisensory paradigms are well suited for examining the role of neural oscillations in the transfer and integration of information across distributed cortical networks.

Box I. Basic Properties of Neural Oscillations.

Oscillatory neural activity recorded by electroencephalography (EEG) or magnetoencephalography (MEG) can be understood as the synchronous waxing and waning of summed postsynaptic activity of large neural populations (Wang 2010). The resulting waveform can be dissected into different frequency components with distinct amplitudes and phases (Mitra and Pesaran 1999). Different types of oscillatory responses, which reflect different aspects of neural synchronization, can be distinguished (Jones 2016; Tallon-Baudry and

(continued)

<sup>2</sup>Department of Psychiatry and Psychotherapy, St. Hedwig Hospital, Charité—Universitätsmedizin Berlin, Berlin, Germany

#### **Corresponding Author:**

Julian Keil, Biological Psychology, Christian-Albrechts-University Kiel, Olshausenstraße 62, 24118 Kiel, Germany. Email: Julian.keil@gmail.com

<sup>&</sup>lt;sup>1</sup>Biological Psychology, Christian-Albrechts-University Kiel, Kiel, Germany

#### Box I. (continued)

Bertrand 1999). Evoked oscillations are closely related to the onset of an external event and the summation over trials of identical phase can result in event-related potentials (ERPs). Induced oscillations can be elicited by stimulation but can also occur independent of external stimulation. Induced oscillations do not have to be strictly phase-locked and time-locked to the onset of stimuli, but can be modulated by cognitive processes (Tallon-Baudry and Bertrand 1999). Furthermore, stimulation can lead to a phase reset that is time-locked to stimulus onset. Phase-locking of oscillatory responses can be quantified as intertrial coherence (ITC; Delorme and Makeig 2004). In contrast to induced oscillations, pure phase reset is reflected in increased ITC. It also differs from evoked oscillations, in that pure phase reset is not accompanied by increased power. Importantly, phase reset plays a prominent role in multisensory processing (Kayser and others 2008; Lakatos and others 2007). Phase coherence can reflect functional connectivity, that is, the interaction between oscillatory activities in different cortical regions. 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 2015). The phase delay between oscillatory activities can indicate the direction of information flow between cortical regions (Maris and others 2016).

With respect to the role of neural oscillations for cognitive processes, Engel and Fries (2010) suggested that fast oscillations indicate states of high arousal, whereas states of low arousal are primarily characterized by slow oscillations. The authors also proposed that oscillations in different frequency bands enable interactions between neural populations. Similarly, Lakatos and others (2005) suggested a hierarchical organization, in which the phase of slow oscillations modulates the amplitude of faster oscillations. Moreover, the phase of slow oscillations likely reflects fluctuations in cortical excitability (Craddock and others 2016; Jensen and Mazaheri 2010). Relatedly, a number of findings suggest that low-frequency oscillations provide temporal windows for multisensory processing (Kayser and other 2008; Lakatos and others 2007; Simon and Wallace 2017).

**Box 2.** Frequency-Dependent Information Transfer in Unisensory Cortical Areas.

Recent studies have analyzed information transfer within unisensory cortical areas. In an intracranial EEG study using depth electrodes in primary auditory (A1) and associate auditory cortex (AAC), Fontolan and others (2014) examined directed functional connectivity. The authors found a top-down influence from AAC to A1 in the frequency range below 40 Hz, but a bottom-up influence from A1 to AAC in the frequency range above 40 Hz. In line with findings on crossmodal influence (Mercier and others 2015), the authors found evidence for a modulation of gamma-band amplitude in A1 by low-frequency phase in AAC. In a recent MEG study, Michalareas and others (2016) examined granger-causal

#### The Neuroscientist 24(6)

#### Box 2. (continued)

interactions along feed-forward and feedback projections within the visual cortex. In agreement with the findings of Fontolan and others (2014), the authors found gamma-band activity in feed-forward projections and alpha-and beta-band activity in feedback projections. Using laminar recordings from the macaque visual cortex, van Kerkoerle and others (2014) established gamma activity as a feed-forward rhythm, and alpha-band activity as a feedback rhythm. Whereas the former originates in the input layer 4 and propagates to superficial and deep layers, the latter is initiated in output layers 1, 2, and 5 and propagates toward layer 4. Taken together, recent findings underscore the role of low-frequency oscillations as a mechanism for top-down information flow, whereas high-frequency oscillations apparently act to propagate bottom-up information. These findings are widely in agreement with the results from studies on multisensory processing and perception reviewed in this article.

## Do Neural Oscillations in Different Frequencies Reflect Distinct Mechanisms in Multisensory Processing?

In our environment, we are constantly confronted with an abundance of information. Much of this information originates from identical sources and is redundant or complementary. Multisensory processing allows us to organize this information. To date, it is not well understood how ongoing neural oscillations, network architecture, and functional connectivity influence the processing and perception of multisensory stimuli. Rapid progress in electrophysiological research has extended our knowledge of the cortical mechanisms underlying bottom-up, stimulusdriven information processing. In addition, recent research has shed light on the influence of top-down processes, such as attention and expectations, on multisensory perception and the underlying neural processes. Here, we will review recent studies suggesting that neural oscillations in distinct frequency ranges reflect different mechanisms of multisensory processing.

## Study Selection and Structure of the Review

We will summarize electrophysiological research from the last decade on multisensory processing in human adults. Using the PubMed<sup>1</sup> search engine, we found 275 research papers, 93 of which we selected according to our focus on neural oscillations. Based on the abstracts, we identified 53 studies recording electrophysiological data during a multisensory task (Table 1). These studies feature various paradigms and research questions. We assigned the 53 studies to four main categories: (1) stimulus-driven mechanisms (section "Stimulus-Driven Mechanisms of Multisensory Processing"); (2) multisensory illusions (section "Multisensory Illusions"); (3) top-down control

First Author	Year	Meth.	Mod.	Category	Parameter					Frequency				
					Ev.	Ind.	Tot.	ITC	Conn.	δ	θ	α	β	γ
Bauer	2012	MEG	ΤV	Attention		*						*	*	
Barutchu	2013	EEG	AV	Attention								*	*	
Pomper	2015	EEG	ΤV	Attention			*	*		*		*	*	
Keil	2016	EEG	TV	Attention			*		*	*	*	*	*	
Kumagai	2016	EEG	AV	Attention			*	*			*	*		
Keller	2017	EEG	AV	Attention			*	*			*	*		
Yuval-Greenberg	2007	EEG	AV	Congruence	*		*							*
Arnal	2011	MEG	AV	Congruence			*	*		*	*		*	*
Kanayama	2012	EEG	ΤV	Congruence	*			*			*			*
Lange	2013	MEG	AV	Congruence			*		*			*	*	*
Quinn	2013	ECoG	ΤV	Congruence	*	*		*						*
Schelenz	2013	EEG	AV	Congruence		*						*		
Gleiss	2014	EEG	AV	Congruence			*			*	*	*		
Göschl	2015	EEG	ΤV	Congruence			*				*	*	*	
He	2015	EEG	AV	Congruence			*					*	*	
Ohki	2016	MEG	AV	Congruence			*		*	*	*	*	*	
Krebber	2015	EEG	ΤV	Congruence			*							*
Kayser	2017	EEG	AV	Congruence			*					*		
Jessen	2012	EEG	AV	Crossmodal			*						*	
Mercier	2013	ECoG	AV	Crossmodal				*			*	*	*	*
Schepers	2013	EEG	AV	Crossmodal			*						*	
Gleiss	2014	EEG	AV	Crossmodal			*		*			*	*	
Biau	2015	EEG	AV	Crossmodal				*			*	*		
Kanayama	2015	EEG	ΤV	Crossmodal	*			*			*	*		
Lin	2015	EEG	AV	Crossmodal	*		*	*			*	*	*	*
Mercier	2015	ECoG	AV	Crossmodal	*			*		*	*			
Rhone	2015	ECoG	AV	Crossmodal	*								*	*
Mishra	2007	EEG	AV	Illusion			*							*
Kanayama	2007	EEG	ΤV	Illusion			*		*					*
, Kanayama	2009	EEG	ΤV	Illusion			*		*					*
, Hipp	2011	EEG	AV	Illusion			*		*				*	*
Lange	2011	MEG	ΤV	Illusion			*				*	*	*	*
Keil	2012	MEG	AV	Illusion		*			*		*		*	
Lange	2013	MEG	ΤV	Illusion		*						*		*
Evans	2013	EEG	ΤV	Illusion		*			*			*	*	
Keil	2014	MEG	AV	Illusion		*						*		
Cecere	2015	EEG	AV	Illusion			*		*			*		
Leonardelli	2015	MEG	ΤV	Illusion			*					*	*	
Roa Romero	2015	EEG	AV	Illusion			*							*
Balz	2016	EEG	AV	Illusion			*		*			*	*	*
Kumar	2016	EEG	AV	Illusion			*	*				*		
Grabot	2017	MEG	AV	Illusion										
Keil	2017	FEG	AV	Illusion			*					*		
Rao	2017	FEG	τv	Illusion			*					*	*	
Morís Fernández	2017	FEG	AV	Illusion			*				*			
Senkowski	2007	EFG	AV	Synchrony	*									*
l enggenhager	2007	FFG	TV	Synchrony			*					*		*
Kösem	2014	MFG	AV	Synchrony				*		*				
Kambe	2015	EEG	AV	Synchrony				*					*	

## Table I. Overview of the Empirical Reports Considered in the Current Review.

First Author	Year	Meth.	Mod.	Category		Pa	aramet	Frequency				
					Ev.	Ind.	Tot.	ITC Conn.	δ	θ	α	β
Yuan	2016	EEG	AV	Synchrony		*						*

Synchrony

Synchrony

Synchrony

## Table I. (continued)

2017

2017

2017

EEG

EEG

MEG

A٧

A٧

AV

TV = tactile-visual; AV = audiovisual; Ev. = evoked power; Ind. = induced power; Tot. = total power; TTC = intertrial coherence; Conn. = functional connectivity;  $\delta$  = delta-band (3–4 Hz);  $\theta$  = theta-band (4–8 Hz);  $\alpha$  = alpha-band (8–12 Hz);  $\beta$  = beta-band (13–30 Hz);  $\gamma$  = gamma-band (>30 Hz).

The category indicates the type of experimental paradigm (from top to bottom): attention tasks; stimulus-congruence experiments; crossmodal experiments; illusion paradigms; temporal synchrony. The studies differ with respect to the recording methods (EEG, MEG, or electrocorticography), and the frequency ranges where effects have been observed. Additionally, studies also differ in their focus, which variously included the analysis of evoked oscillatory power, induced oscillatory power, total oscillatory power, intertrial coherence, or functional connectivity (Box 1).

and ongoing oscillations (section "Top-Down Control of Multisensory Processing"); and (4) speech perception (section "Audiovisual Speech Processing"). Studies that cover more than one topic will be featured whenever suitable. To summarize the outcomes of the reviewed studies, we derived an integrative framework for the role of neural oscillations in multisensory processing (section "An Integrative Framework for the Role of Neural Oscillations in Multisensory Processing"). Briefly put, we propose that neural oscillations in different frequency bands reflect key mechanisms, that is, feed-forward-feedback processing, attention modulation, and predictive coding in multisensory processing. We also propose that these mechanisms can act in parallel. Open questions and clinical implications of research on neural oscillations and multisensory processing are discussed in the final section (section "Open Questions and Future Directions").

## Stimulus-Driven Mechanisms of Multisensory Processing

In speeded response tasks, individuals usually respond faster to multisensory stimuli than to unisensory stimuli (Miller 1986; Pomper and others 2014). There also are qualitative differences between the processing of concurrent stimuli from multiple modalities and that of the respective single modalities presented alone. Specifically, stimulus-driven processing, as reflected through gamma-band power, is increased in response to multisensory compared with unisensory stimuli, especially when both stimuli are attended (Senkowski and others 2005) and integrated (Bhattacharya and others 2002). Recent studies elucidated various factors that influence stimulus-driven processing of multisensory stimuli. Analyzing ITC in intracranial recordings, Mercier and others (2013) showed an influence of auditory stimulation on the processing of a concurrent visual stimulus, reflected in increased ITC in the visual cortex in the thetaband (5-8 Hz) and alpha-band (8-12 Hz), as well as in the beta-band (13-30 Hz). Based on increased phase reset of delta-band (3-4 Hz) and theta-band (5-8 Hz) oscillations in the auditory cortex during audiovisual stimulation in a follow-up intracranial study, Mercier and others (2015) proposed a mechanism for the crossmodal influence between different senses (Fig. 1): Optimally aligned phases promote communication between cortical areas, and stimuli in one modality can reset the phase of an oscillation in a cortical area of another modality. Furthermore, increased crossmodal phase alignment in this study correlated with faster behavioral responses. These findings are in line with previous results from animal studies, which indicate that enhanced crossmodal processing relates to a phase reset in slow oscillations (Kayser and others 2008; Lakatos and others 2007). Similarly, using surface EEG during visuotactile stimulation, Kanayama and others (2015) found increased thetaband ITC in somatosensory and anterior cingulate cortices following multisensory stimulation, relative to unisensory stimulation. In summary, it is likely that crossmodal influence involves low-frequency phase alignment. Moreover, stimulus-driven multisensory processing can enhance perception, which is reflected in increased gamma-band power.

## The Role of Stimulus Congruence in Multisensory Processing

Whereas stimulus-driven multisensory processing can enhance perception, the type of concurrent information also has a decisive influence on stimulus processing. Yuval-Greenberg and Deouell (2007) showed that semantically congruent audiovisual stimulation, in comparison with incongruent stimulation, is associated with enhanced occipitoparietal gamma-band power.<sup>2</sup> Similarly, Kanayama and others (2012) reported that visuotactile congruence is reflected in increased parietal evoked gamma-band power. Using intracranial recordings, Quinn and others (2013) analyzed the spatiotemporal profile of visuotactile integration, as reflected in high gamma-band power (70–190 Hz) and

Covic

Simon

Giordano

local field potentials (LFPs). Interestingly, this study revealed that multisensory processing can be expressed through subadditive effects, that is, suppression of high gamma-band power in temporoparietal and dorsolateral prefrontal cortices following multisensory compared to unisensory stimulation. Further evidence for the role of gamma-band power for the processing of congruent multisensory stimulation comes from an EEG study by Krebber and others (2015). The authors examined the role of motion congruence for the processing of visuotactile stimuli and found increased gamma-band power during congruent versus incongruent visuotactile motion stimulation, which was localized in visual and somatosensory cortices (Fig. 2A). Underscoring the behavioral relevance of this observation, increased gamma-band power was correlated with faster responses to visual and tactile targets. A behavioral study, in which subjects were asked to detect visuotactile congruence, also found enhanced performance for congruent compared with incongruent multisensory stimuli (Göschl and others 2014). A follow-up EEG study revealed increased late beta-band power in the right somatosensory cortex and supramarginal gyrus during congruent compared with incongruent stimulation (Göschl and others 2015). In an EEG study, Gleiss and Kayser (2014a) analyzed the influence of task-irrelevant auditory stimuli on visual perception. In line with previous results, they found that concurrently presented spatially and temporally aligned sounds reduce visual detection thresholds. Interestingly, the improved perception correlated with reduced occipital alpha-and beta-band power. Investigating the crossmodal influence of auditory stimuli on visual motion stimuli in another EEG study, Gleiss and Kayser (2014b) found increased occipital alpha-band power during congruent compared with incongruent audiovisual motion stimulation. This finding was corroborated in a recent follow-up study (Kayser and others 2017). Taken together, the data available suggest that congruent multisensory stimulation facilitates stimulus processing, as reflected in reduced alpha- and beta-band power and increased gamma-band power. Notably, the type of stimulus congruence, for example, semantic or motion congruence, appears to have different effects on multisensory processing.

## The Impact of Temporal Alignment on Multisensory Processing

Comparable with the above-described effects of stimulus congruence on neural oscillations, Senkowski and others (2007) demonstrated enhanced evoked occipital gammaband power following temporally aligned, that is, synchronous, compared with asynchronous audiovisual stimulation. Interestingly, studies examining temporal alignment and neural oscillations have revealed additional effects. Analyzing visual-somatosensory stimulus processing in a virtual-reality environment, Lenggenhager and others (2011) showed reduced somatosensory alphaband power during asynchronous compared with synchronous stimulation. In another study using rhythmic audiovisual stimulation, Covic and others (2017) found increased occipitoparietal steady-state power during synchronous compared with asynchronous stimulation. Moreover, Kösem and others (2014) demonstrated that phase shifts in entrained oscillations in auditory areas map the perceived synchrony of audiovisual stimuli. Findings by Kambe and others (2015) further support the role of oscillatory phase for multisensory processing. The authors found increased beta-band ITC in central and occipital electrodes only when subjects perceived audiovisual stimuli as synchronous but not when they perceived them as asynchronous (Fig. 2B). More recently, Simon and Wallace (2017) showed that the effect of a visual stimulus on frontal theta-band power depends on the phase of low-frequency oscillations that were entrained by an auditory stimulus. This finding is in agreement with previous intracranial studies in monkeys (Kayser and others 2008; Lakatos and others 2007) (Box 1). To summarize, enhancement of gamma- and alphaband power, as well as increases in beta-band ITC, reflects the temporal alignment of stimuli from multiple sensory modalities. Moreover, phase reset of low-frequency oscillations presumably plays an important role in multisensory processing.

## **Multisensory Illusions**

Multisensory illusions, such as the rubber hand illusion (RHI), the sound-induced flash illusion (SIFI), or the McGurk illusion, allow the comparison of variable perception following physically identical stimulation. This can reveal important information on the way we perceive our environment. The RHI is a visuotactile illusion, in which the apparent visual location of a body part induces an illusory shift in proprioception. In two experiments Kanayama and others (2007, 2009) showed that congruent visuotactile stimuli that induced the RHI are associated with increased parietal gamma-band power and global gamma-band functional connectivity. In addition, two further studies revealed reduced central alpha- and beta-band power during the perception of the RHI (Evans and Blanke 2013; Rao and Kayser 2017).

Another prominent illusion is the SIFI. Here, a single flash that is accompanied by two rapid sounds is often perceived as two flashes (Shams and others 2000). Mishra and others (2007) showed that an increase in poststimulus occipital gamma-band power reflects the perception of the SIFI (Fig. 3A, left column). In a similar vein, increased occipital gamma-band power also reflects the touchinduced flash illusion, where two rapid tactile stimuli



**Figure 1.** Bottom-up multisensory integration likely involves crossmodal influence between cortical areas, as reflected in lowfrequency power and phase reset. Here, electrodes in auditory cortices, highlighted by the red circle, of an epileptic patient implanted with intracranial electrodes show increased ITC (labeled PCI in the color bar) to audiovisual stimulation, relative to auditory and visual stimulation alone. The time-frequency representation illustrates the change of ITC over time (*x*-axis) in different frequencies (*y*-axis). Different colors indicate the difference in ITC following audiovisual and the combined auditory and visual stimulation for each time-frequency combination. Lower ITC following audiovisual stimulation is marked in blue, increased ITC following audiovisual stimulation is marked in red (adapted from Mercier and others 2015). The right panel illustrates the presumed effect of a single visual stimulus (V) on the phase of ongoing neural oscillations in auditory cortical areas (A): A visual stimulus leads to a phase alignment of ongoing neural oscillations in auditory cortical areas across trials.

induce the double flash illusion (Lange and others 2011). More recently, Balz and others (2016a) investigated the role of gamma-band power in the superior temporal gyrus (STG) for the SIFI (Fig. 3A, middle column). Using EEG and magnetic resonance spectroscopy, the authors showed that the perception of the SIFI is correlated with gammaband power in the STG. Moreover, the study revealed a three-way relationship between the gamma-aminobutyric acid (GABA) level in the STG, gamma-band power in the STG, and the SIFI illusion rate. It is possible that GABA mediates the relationship between gamma-power in the STG and the SIFI illusion rate.

Finally, the McGurk illusion involves incongruent audiovisual speech, which can be fused into a subjectively congruent audiovisual percept (McGurk and MacDonald 1976). By comparing the illusion with trials in which the illusion was not perceived, Keil and others (2012) found reduced frontal theta-band power following perception of the illusion. In other words, increased thetaband power indexed perceived audiovisual incongruence. Morís Fernández and others (2017b) recently showed increased central theta-band power following the McGurk illusion, as well as incongruent audiovisual stimuli compared with congruent audiovisual stimuli. The authors argue that theta-band power is linked to a general-purpose conflict detection mechanism. Similarly, Roa Romero and others (2015) compared McGurk illusion trials with congruent audiovisual syllables. The authors identified an early and a late stage of audiovisual integration, which were both marked by stronger poststimulus beta-band power decreases in frontoparietal electrodes following the illusion. In addition to local power, largescale functional connectivity networks appear relevant for illusion perception. For example, Kumar and others (2016) found increased beta-band global coherence during the McGurk illusion.

Taken together, the studies reviewed above show that multisensory integration resulting in illusions is associated with increased perceptual processing, as reflected in enhanced gamma-band power and reduced theta-, alpha-, and beta-band power. The studies also suggest that changes in neural oscillations relate to the varying perception and are not merely caused by invariant stimulation. The involvement of both high- and low-frequency bands further indicates that multisensory illusions rely on the dynamic interplay of stimulus-driven processing and topdown control.

## Top-Down Control of Multisensory Processing

There is ample evidence that top-down control modulates neural oscillations and multisensory processing, as well as multisensory perception. Recent studies have focused on the role of attention and expectations on multisensory processing (Macaluso and others 2016). Task demands also affect multisensory processing (Auksztulewicz and others 2017; Mégevand and others 2013). Moreover, fluctuations in ongoing neural oscillations have been shown to influence top-down processing and perception (Lange and others 2014).

## Top-Down Control and Expectations Affect Multisensory Processing

In multisensory paradigms, changes in local alpha-band power can reflect shifting of attention in space, to one



**Figure 2.** (A) Congruent visuotactile motion stimulation is accompanied by enhanced gamma-band power in sensory cortical areas. Time-frequency plots depict neural activity in the visual cortex (left column) and right somatosensory cortex (middle column), that is, contralateral to the stimulation site, for congruent visuotactile motion stimulation (top row) and incongruent motion stimulation (bottom row) (adapted from Krebber and others 2015). The diagram of neural oscillations in the right column illustrates the increased gamma-band power in visual and somatosensory cortical areas following congruent (top) and incongruent (bottom) visuotactile (VT) stimulation. (B) Perceived synchrony (top row) during audiovisual stimulation is accompanied by increased beta-band ITC in visual (left column) and auditory (middle column) electrodes, reflecting activity from auditory and visual cortical areas (adapted from Kambe and others 2015). The diagram of neural oscillations in the right column illustrates the presumed phase alignment in visual and auditory cortical areas following an audiovisual (AV) stimulus that is perceived as synchronous ([AV], top row). By contrast, no phase alignment occurs if the AV stimulus is not perceived as asynchronous ([A] [V], bottom row).



**Figure 3.** (A) The perception of the flash illusion is indexed by enhanced gamma-band power. Time-frequency representations illustrate the relative change in oscillatory power from baseline for the comparison between illusion and no-illusion trials (adapted from Mishra and others 2007, left column; Balz and others 2016a, middle column). The diagram of ongoing neural oscillations illustrates the way the perception of the flash illusion is reflected in increased gamma-band power following audiovisual (AV) stimulation (right column). (B) Multisensory integration, as examined in the flash illusion, is influenced by alpha- and beta-band power. Prior to stimulus onset, beta-band power is increased (left column), and alpha-band power is decreased (middle column) when an illusion is subsequently perceived. The time-frequency representations show *t*-values for the comparison between illusion and no-illusion trials (adapted from Keil and others 2014, left column; Lange and others 2013a, middle column). The diagram of ongoing neural oscillations illustrates increased beta-band and decreased alpha-band power prior to the perception of the illusion (right column).

sensory modality, or toward specific stimulus features. Using visuotactile stimulation, Bauer and others (2012) found that top-down attention modulates modality-specific frequencies, with alpha-band power decreases, over occipital cortex and beta-band power decreases over somatosensory cortex. Another study, using an audiovisual discrimination task, showed that stimulus relevance modulates stimulus-driven decreases in central and occipital alpha- and beta-band power (Barutchu and others 2013). Moreover, in an audiovisual task, spatial attention reduced contralateral temporal alpha-band power and concurrently increased contralateral frontal alpha-band power (Kumagai and others 2016). Göschl and others (2015) compared neural oscillations in a visuotactile congruence evaluation task and a more demanding target detection task. The study revealed that alpha- and beta-band power are decreased in the premotor cortex, somatosensory cortex, and supramarginal gyrus, in the detection compared with the evaluation task. The authors suggested that this observation reflects the higher engagement of cortical areas associated with visuotactile integration in the more demanding detection task. Together,

these findings are in agreement with the "gating-by-inhibition" hypothesis, which posits that oscillatory activity in the alpha-band provides an inhibitory mechanism, which reduces the processing capacities of a task-irrelevant cortical area to optimize performance in active cortical areas (Jensen and Mazaheri 2010). Interestingly, the phase of alpha-band oscillations can act to temporally organize incoming information to prevent information overload (Gips and others 2016; Ronconi and others 2017). In another EEG study, in which intersensory attention and temporal expectations were concurrently manipulated, Pomper and others (2015) observed modulations of alpha- and beta-band power. The authors found that intersensory attention reduces alpha- and beta-band power in the visual cortex when subjects attended to visual stimuli. Additionally, beta-band power in the somatosensory cortex was reduced when attending to tactile stimuli. Furthermore, beta- and delta-band modulations in the motor and somatosensory cortex reflected temporal expectations. In a follow-up examination, Keil and others (2016) found that intersensory attention modulates an alphaband functional connectivity network, encompassing visual, somatosensory, and inferior parietal areas. Moreover, temporal expectations modulate a beta-band functional connectivity network, involving visual, parietal, and frontal areas. Interestingly, intersensory attention and temporal expectation both modulated a common theta-band functional connectivity network, encompassing frontal and parietal cortical areas. This provides evidence for two distinct modes of attention in multisensory processing: a bottom-up driven, automatic mode, and a deliberate mode requiring top-down control. In accordance with this idea, Keller and others (2017) recently showed that alpha-band power is critical for selectively ignoring task-irrelevant information, whereas theta-band power modulations are associated with divided attention, independent of task difficulty. In summary, selectively attending to one constituent of a multisensory stimulus primarily modulates local alpha- and beta-band activity, depending on the attended stimulus modality. Moreover, divided attention and expectations primarily modulate thetaand beta-band oscillations in a more widespread cortical network, presumably reflecting top-down control.

## Ongoing Oscillations Shape the Top-Down Control on Multisensory Processing

Ongoing fluctuations of neural oscillations, which presumably reflect changes in cortical excitability (Lange and others 2013a), can affect the top-down influence on multisensory processing and perception. In an influential EEG study, Hipp and others (2011) investigated neural coherence in the audiovisual bounce-pass paradigm. Therein, two moving vertical bars approach each other, overlap, and diverge again. At the point of overlap, a sound is presented, resulting in a bistable percept in which the bars either appear to pass each other or bounce off each other. The authors found that increased coherence in two cortical networks predicted an integrated audiovisual percept. Beta-band coherence characterized a network spanning frontal, parietal, occipitotemporal, and occipital cortical areas. Gamma-band coherence marked a second network encompassing parietal and temporal cortical areas. Hence, this suggests that ongoing fluctuations in cortical networks predict multisensory processing.

In an MEG study using audiovisual speech stimuli eliciting the McGurk illusion, Keil and others (2012) found increased beta-band power in the STG, precuneus, and right frontal cortex prior to an integrated audiovisual percept. Also, beta-band functional connectivity, involving temporal, occipital, and frontal areas, was increased prior to the perception of the illusion. In another MEG study, Keil and others (2014) obtained similar results for incongruent audiovisual stimuli eliciting the SIFI. Again, increased beta-band power in the STG preceded a multisensory illusion (Fig. 3B, left column). Moreover, alphaand beta-band functional connectivity in a network spanning temporal, parietal, and frontal areas

differentiated between illusory and nonillusory percepts. Interestingly, increased beta-band functional connectivity between STG and primary auditory cortex was related to illusion perception on a single trial level. Using a visuotactile adaptation of the SIFI, Lange and others (2013b) found that cortical activity prior to stimulus onset predicts perception and reflects cortical excitability (Box 1). In this study, reduced alpha-band power in visual cortical areas and increased gamma-band power in parietal and temporal cortical areas preceded the illusion (Fig. 3B, middle column). In contrast, Leonardelli and others (2015) found increased frontocentral alpha-band power as well as increased functional connectivity between frontal, parietal, and occipital cortex prior to an integrated, albeit not illusory, visuotactile percept. Moreover, a study comparing the perceived simultaneity of asynchronous audiovisual stimuli found increased occipital beta- and gamma-band power prior to a simultaneous perception (Yuan and others 2016). Recently, Grabot and others (2017) showed that alpha-band power in the right supramarginal gyrus prior to audiovisual temporal order judgments influences the temporal sequencing of events. Two studies using the SIFI further highlighted the role of low-frequency oscillations for the shaping of audiovisual perception (Cecere and others 2015; Keil and Senkowski 2017). Cecere and others (2015) found a correlation between the individual alpha-band frequency and illusion rate, which indicates that alpha-band oscillations provide a temporal window in which the crossmodal influence could induce an illusion. Recently, Keil and Senkowski (2017) could replicate the relationship between the individual alpha-band frequency and the SIFI perception rate and localized this effect to the occipital cortex. Thus, these studies suggest that neural oscillations influence upcoming multisensory processing. More specifically, alpha-band oscillations might indicate excitability in primary sensory areas and provide windows of opportunity for crossmodal influence. Furthermore, increased betaand gamma-band power in multisensory cortical areas might indicate increased readiness to integrate information.

Another approach to examine the role of ongoing oscillations for multisensory processing is the modulation of neural activity through electric brain stimulation. Underscoring the crucial role of low-frequency oscillations for crossmodal influence, Cecere and others (2015) found that modulating the individual alpha-band frequency using transcranial direct current stimulation modulated the probability of an illusion perception. Moreover, using intracranial electric stimulation, Beauchamp and others (2012) found the illusory perception of a phosphene only when gamma-band power in the temporoparietal junction was present. Interestingly, perturbing this area using TMS impaired the SIFI perception (Hamilton and others 2013; Kamke and others 2012). Romei and others (2012) showed that TMS-evoked phosphene perception varied with an auditory evoked alpha-band phase, thus highlighting the role of multisensory phase reset. Hence, these studies suggest a causal relationship between ongoing neural oscillations and crossmodal processing, as well as perception.

Taken together, top-down processes, such as attention and expectations, modulate local cortical activity and functional connectivity networks, thereby orchestrating the integration of multisensory stimuli. Local cortical activity and activity within functional connectivity networks fluctuate, and the cortical state prior to the presentation of a stimulus also predicts multisensory processing and perception. The studies reviewed above suggest that multisensory processing is fostered by (1) increased excitability in primary sensory areas, (2) enhanced local activity in multisensory cortical areas, and (3) increased functional connectivity within distributed cortical networks.

## **Audiovisual Speech Processing**

Recent studies have underlined the importance of rhythmic stimulus properties for auditory perception, especially speech perception (Henry and Obleser 2012). Giraud and Poeppel (2012) proposed that the phase of ongoing cortical activity aligns with slow oscillations of the speech amplitude envelope, and that this mechanism facilitates audiovisual speech processing. Similarly, gestures appear to influence speech perception. Two studies found an influence of concurrent gestures on central ongoing low-frequency oscillations during audiovisual speech perception (Biau and others 2015; He and others 2015). This influence appears to be augmented by emotional content (Jessen and others 2012). Furthermore, in a study comparing temporally aligned and nonaligned speech stimuli, Ohki and others (2016) found that the coupling between temporal delta-band phase and betaband power reflects audiovisual stimulus encoding. Another study compared the processing of bisensory audiovisual versus unisensory auditory speech stimuli (Schepers and others 2013). The authors found suppression of central beta-band power, which was modulated by concurrent noise stimulation. A further study, which examined noise during audiovisual speech perception, has also highlighted the role of early gamma-band oscillations (Lin and others 2015). The study revealed that auditory noise increases early induced and evoked gamma power, as well as ITC at frontal electrodes. In addition, a more recent MEG study showed that speech encoding, as reflected in low-frequency oscillations, varied with the signal-to-noise ratio (Giordano and others 2017). Importantly, the benefit of seeing a speaker's mouth movements was related to enhanced delta- and theta-band functional connectivity between frontal and temporal areas. Taken together, the various findings from these studies suggest that different frequency bands subserve different mechanisms of audiovisual speech processing and perception.

The predictive value of input from one sensory modality, for example, visual input, for the input from another modality, for example, auditory input, plays an important role in audiovisual speech integration (Fig. 4, upper row). The predictive coding hypothesis posits that only unpredicted aspects of a signal are fed forward between cortical areas, which reduces redundancy and facilitates stimulus processing (Huang and Rao 2011). Hence, predicting input from one sensory modality by top-down information or concurrent information from another modality releases cognitive resources, because only violations of the predictions will be processed in further detail. With respect to the underlying neural mechanisms, Arnal and Giraud (2012) proposed that gamma-band oscillations transfer feed-forward information across sensory cortices, whereas beta-band oscillations are involved in topdown signaling and control.

Similar to nonspeech audiovisual stimuli (section "The Role of Stimulus Congruence in Multisensory Processing"), the congruence of auditory and visual inputs has a major impact on the processing and integration of audiovisual speech stimuli. Analyzing congruent and incongruent audiovisual speech stimuli, Arnal and others (2011) found that incongruent stimuli lead to increased correlations between gamma-band ITC, power, and evoked brain activity, which were absent following congruent stimuli. However, the authors did not directly compare neural oscillations between congruent and incongruent speech. Comparing congruent and incongruent audiovisual speech stimuli, Lange and others (2013a) found increased left temporoparietal gamma- and beta-band power following congruent speech. The authors hypothesized that the enhanced gamma- and beta-band power following congruent stimulation reflects evaluation of matching audiovisual information. Interestingly, incongruent compared with congruent stimulation evoked stronger ERPs and increased alpha-band power. This suggests that stimulus processing beyond the initial processing of unexpected information is gated by alpha-band activity. In another study, Roa Romero and others (2015) compared congruent and incongruent audiovisual syllables and identified an early and a late stage of audiovisual integration. Both processing stages were marked by stronger poststimulus frontocentral betaband power decreases following the integration of incongruent stimuli to the McGurk illusion (Fig. 4, middle row). This finding is in line with a recent multistage model of speech perception (Peelle and Sommers 2015). According to this model, in a first stage, visual information increases the sensitivity to auditory information. In a second stage, specific information on the speech content is extracted (Fig. 4, bottom row). In a similar vein, Roa Romero and others (2015) proposed that auditory and visual information are compared at an early stage and integrated at a later stage. Tentative support for this model comes from a human intracranial recording study in which Rhone and others (2015) investigated the processing stages at which visual and auditory speech interact. The authors found that high gamma-band power in Heschl's gyrus did not differentiate between audiovisual speech and nonspeech signals, whereas gamma-band power in the STG was enhanced following meaningful audiovisual speech stimuli. Using emotional audiovisual speech stimuli, a combined EEG-fMRI study found increased alpha-band power in frontal areas following emotionally congruent stimulation (Schelenz and others 2013). Interestingly, a recent study examining the McGurk illusion also reported stronger central thetaband power following incongruent compared with congruent audiovisual speech (Morís Fernández and others 2017b). Frontal theta-band power has been frequently associated with cognitive control and mismatch or error related-processing (Cavanagh and Frank 2014). Hence, enhanced frontal theta-band power for incongruent versus congruent audiovisual speech likely reflects crossmodal prediction error processing and conflict resolution. Supporting the notion of multistage processing, a recent fMRI study provides evidence for initial audiovisual conflict detection and subsequent resolution of this conflict involving frontal cortical areas (Morís Fernández and others 2017a).

In summary, recent findings suggest that low- and high-frequency oscillations in audiovisual speech perception subserve distinct functional roles. Whereas violations of visual predictions or incongruence between auditory and visual stimuli are reflected in gamma-band power, the transfer of predictions and the integration of mismatching information into a coherent percept seems to be primarily reflected in beta-band functional connectivity. Moreover, enhancement of frontal theta- and alpha-band power could reflect the resolution of audiovisual mismatch and processing of audiovisual prediction errors.

## An Integrative Framework for the Role of Neural Oscillations in Multisensory Processing

Based on the outcome of the studies reviewed above, we attempt to structure the converging findings into an integrative framework (Fig. 5). Within this framework, we propose three complementary mechanisms that shape multisensory processing: (i) feed-forward–feedback processing, (ii) attention modulation, and (iii) predictive coding. In the following we will provide some predictions about the role of neural oscillations for the three mechanisms. We propose that the three mechanisms can concurrently affect

multisensory processing and that the relative involvement of each mechanism depends on task demands.

Mechanism (i): The studies reviewed in sections "Stimulus-Driven Mechanisms of Multisensory Processing" and "Multisensory Illusions" suggest that dynamic feed-forward-feedback processing is primarily reflected in gammaand beta-band power, respectively. During multisensory processing, information that is initially processed in unisensory cortical areas is fed forward to higher-order cortical areas. In these areas, the information can be integrated with sensory input from other modalities. This corresponds to the classical feed-forward model of multisensory integration (Stein and Meredith 1993). Information on stimulus congruence is then fed back between higher-order and lower-order cortical areas. In addition, information processed within one cortical area can modulate ongoing activity in another area, for example, through delta- and theta-band phase resetting (Fig. 5, left panel).

*Mechanism (ii):* In tasks requiring top-down control, intersensory attention is reflected in local alpha- and betaband power, as well as alpha- and beta-band functional connectivity networks. The studies reviewed in section "Top-Down Control of Multisensory Processing" suggest that attention modulates the integration of cortical areas into distributed networks via theta-band functional connectivity. Furthermore, ongoing fluctuations of cortical activity in the alpha- and beta-band likely reflect fluctuations of cortical excitability and influence stimulus processing (Fig. 5, middle panel).

*Mechanism (iii):* Information from one sensory modality generates predictions, which influence information processing in another modality. The studies reviewed in section "Audiovisual Speech Processing" suggest that predictive processing across sensory modalities is reflected in low-frequency power in frontal cortical areas, and beta-band functional connectivity. Violations of multisensory predictions are primarily reflected in high-frequency power in sensory cortices. Our proposed third mechanism relates to the model for the transfer of topdown and bottom-up information within the predictive coding framework introduced by Arnal and Giraud (2012) (Fig. 5, right panel).

# Audiovisual Speech Perception as an Example Scenario

An example scenario for the information-processing cascade is the generation of predictions based on attended visual information for an upcoming auditory stimulus (Fig. 6). Visual input is processed in primary visual cortex as reflected in alpha-, beta-, and gamma-band power. Bottom-up information in then transferred to frontal cortical areas as reflected in beta- and gamma-band oscillations. Based on this information, predictions on upcoming



**Figure 4.** Audiovisual speech perception involves the coordinated interplay between bottom-up and top-town influences. In naturalistic audiovisual speech, visual information precedes the onset of the auditory signal (upper row). Thus, predictions regarding the auditory signal can be formed based on visual information. Here, perception of incongruent audiovisual speech stimuli resulting in the McGurk illusion was contrasted with perception of congruent audiovisual speech stimuli (middle row). Time-frequency plots illustrate the difference in relative change from baseline between congruent and McGurk illusion trials for the perception (left column) and integration (right column) stages. The dashed lines mark the time-frequency windows in which a significant effect was found (adapted from Roa Romero and others 2015). Audiovisual speech perception and integration presumably involves separate processing stages, as illustrated in the bottom row. First, auditory and visual information are processed in primary sensory areas and compared. Subsequently, information from both sensory modalities is integrated. In the case of incongruent information, both processes require enhanced neural processing resources.



Figure 5. Summary of the framework of multisensory processing. We propose that neural oscillations in different frequency bands reflect different aspects of multisensory processing. Here we propose three complementary key mechanisms that shape multisensory processing. We hypothesize that the mechanisms can act in parallel and that the relative involvement of each these mechanisms depends on task demands. Our framework comprises primary sensory areas (e.g., auditory [A] and visual [V] cortical areas), multisensory areas (e.g., the superior temporal gyrus or angular gyrus [M]), and higher order cortical areas (e.g., frontal cortical areas [F]). These areas are crucial for the information-processing cascade in multisensory integration and perception. Mechanism (i): Feed-forward-feedback multisensory integration. Sensory information is processed in primary sensory cortical areas, as reflected in increased theta-, alpha-, and beta-band ITC (grey letters) as well as reduced alpha- and beta-band power, and increased local gamma-band power (black letters). The crossmodal influence between primary cortical areas is reflected in low-frequency oscillations, such as delta- and theta-band oscillations. Mechanism (ii): Attention modulated multisensory integration. Intersensory attention is reflected in local alpha-, beta-, and gamma-band power, as well as alpha- and beta-band functional connectivity networks. Top-down attention is reflected in theta-band functional connectivity between frontal and sensory areas. Ongoing fluctuations of alpha- and beta-band power presumably reflect fluctuations of cortical excitability, which additionally influences multisensory processing. Mechanism (iii): Predictive coding and multisensory integration. Information from one sensory modality generates predictions, which influence information processing in other sensory modalities. The top-down influence of these predictions from frontal cortical areas to multisensory areas and primary sensory areas is reflected in frontal gamma-band ITC (grey letters), theta- and gamma-band power (black letters), and beta-band functional connectivity. Violations of predictions are reflected in enhancement of gamma-band power and functional connectivity.  $\delta$  = delta-band (3-4 Hz);  $\theta$  = thetaband (4–8 Hz);  $\alpha$  = alpha-band (8–12 Hz);  $\beta$  = beta-band (13–30 Hz);  $\gamma$  = gamma-band (>30 Hz).

auditory information are generated involving low-frequency oscillations (Fig. 6, left column). These top-down predictions are transferred from frontal cortical areas to primary sensory cortices. Therein, theta-band power and beta-band functional connectivity presumably play a major role. Then, a crossmodal influence from visual to auditory cortical regions involving delta- and theta-band oscillations can be established to optimally align neural oscillations within these regions (Fig. 6, middle column). When the auditory stimulus impinges on the system, it is processed in the auditory cortex and information is transferred to higher-order cortical areas, as reflected in gamma-band power. There, it is compared to the visual stimulus. This comparison presumably involves betaband oscillations. In the case of audiovisual congruence, processing in primary sensory areas is facilitated. However, in the case of audiovisual incongruence, the violation of a prediction is fed forward to frontal cortical areas. Under specific conditions, such as those eliciting the McGurk illusion, processing in frontal cortical areas, as reflected in theta-band power, can resolve the incongruence and lead to a fused multisensory percept (Fig. 6, right column). The example scenario underlines the complementary roles of the three major mechanisms depicted in Figure 5.

## Open Questions and Future Directions

#### Predictions for Multisensory Processing

In recent years, the hypothesis that synchronous cortical oscillations play a prominent role in unisensory and multisensory processing has received substantial support (Box 2). Current findings further corroborate the idea that local oscillations and functional connectivity reflect bottom-up as well as top-down influence during multisensory processing. Our integrative framework allows the formulation of predictions for various multisensory processing scenarios. First, multisensory perception requires joint activity of primary sensory and higherorder cortical areas, whereas crossmodal influence can occur directly between sensory cortices. Second, the various stages of the information-processing cascade underlying multisensory processing are reflected in distinct spectral signatures. Third, the processing of congruent and incongruent multisensory information requires different processing stages, especially when incongruent information needs to be resolved into a coherent percept. Fourth, low-frequency neural oscillations should not only influence audiovisual integration but also the



**Figure 6.** Audiovisual speech perception as an example of the information processing cascade in multisensory perception. The top row illustrates the auditory and visual information, as well as the temporal delay between visual movement information and the auditory stream. The bottom row illustrates three different postulated stages of information processing. Left column: Predictions on the upcoming auditory stimulus are formed based on visual information on the mouth movements, which are processed in primary visual cortex and transferred to frontal cortical areas. Information processing involves alpha-, beta-, and gamma-band power, and the information transfer is presumably reflected in beta- and gamma-band oscillations. The formation of predictions likely involves theta-band oscillations. Middle column: These predictions are transferred from frontal cortical areas to primary sensory cortices, as reflected in theta- and beta-band oscillations. A crossmodal influence from visual to auditory stimulus impinges on the system, it is processed in the auditory cortex and information is transferred to higher-order cortical areas, as reflected in gamma-band oscillations. There it is compared to the visual stimulus, which presumably involves beta-band oscillations. In the case of audiovisual incongruence, the violation of a prediction is fed forward to frontal cortical areas and the incongruence can be resolved to a fused multisensory percept, as reflected in theta-band power.

perception of audiovisual synchrony, as well as the detection of stimulus asynchronies. Fifth, modulating cognitive or attentional load should influence multisensory processing and integration similar to the top-down influence in the theta-band.

In this review, we have presented evidence for the relevance of different frequency band oscillations for the different mechanisms underlying multisensory processing. However, our understanding of the putative role of cross-frequency coupling during multisensory processing is, thus far, relatively poor. Recent evidence indicates an influence of cross-frequency coupling for delta-band phase and beta-band amplitude (Ohki and others 2016). For instance, future studies may address the relationships between local theta-band power in frontal cortical areas and theta- and beta-band functional connectivity between frontal and sensory cortical areas, especially with respect to attention and cognitive control. Furthermore, the role of cross-frequency coupling for the relationships between stimulus-based predictions and attention remains unclear. The precise order of activity in primary sensory and higher-order cortical areas during multisensory processing should be investigated in further detail. Recent findings lend support to the idea of temporally orchestrated processing stages during multisensory processing, including predictions and crossmodal comparisons (Lange and others 2014; Morís Fernández and others 2017b; Roa Romero and others 2015). In future research, data from intracranial recordings in humans, with high temporal resolution and broad spatial coverage, may help shed new light on the precise role of the different frequency band activities and their interplay for multisensory processing.

## Implications for the Understanding of Psychiatric Disorders

Our integrative framework also has implications for the understanding of psychiatric disorders in which aberrant multisensory processing has been observed. For example, Roa Romero and others (2016a) compared processing of the McGurk illusion in patients with schizophrenia and healthy control participants. While no group difference was found in the early processing stage of the McGurk illusion, the late processing stage, which was reflected in frontal alpha-band modulations, was impaired in the patient group. Supporting the role of alpha-band suppression for audiovisual integration, the authors found a negative correlation between frontal alpha-band modulations and the McGurk illusion rate. Another study focused on crossmodal predictions in audiovisual speech processing in schizophrenia (Roa Romero and others 2016b). This study indicates that audiovisual incongruence detection is reflected in enhanced early evoked responses. Moreover, the authors found that crossmodal prediction error processing involves frontal theta-band oscillations, and that this processing is impaired in schizophrenia. These findings suggest processing deficits at multiple stages of the multisensory processing cascade. Based on our framework, we predict that patients with schizophrenia suffer from deficits in the generation of crossmodal predictions, the evaluation of stimulus congruence, and the resolution of possible incongruence. The deficits in prediction generation might be reflected in aberrant gamma-band oscillations, which signal impaired feed-forward processing of sensory information. The deficits in stimulus evaluation and integration could be indexed by aberrant frontal theta-band oscillations. Further analyses beyond abnormal top-down control as a mechanism for altered multisensory perception in psychiatric disorders should also be explored. For example, it would be interesting to examine whether the bottom-up driven processing of congruent multisensory stimuli, as reflected in increased gammaband power, is impaired in schizophrenia. In accord with the findings of aberrant crossmodal generation of predictions, we would expect that patients with schizophrenia exhibit lower gamma-band power following congruent stimulation, which again could reflect impaired feed-forward processing of sensory information. First evidence supporting this prediction comes from a recent study by Balz and others (2016b). The authors found that aberrant beta- and gamma-band oscillations reflect multisensory processing deficits in schizophrenia. Taken together, new developments in integrative research approaches focusing on local cortical processes and functional connectivity networks in healthy individuals and in clinical populations will contribute to a deeper understanding of the interplay between bottom-up and top-down processes during multisensory perception. The studies reviewed here suggest that different frequency band oscillations subserve distinct but complementary mechanisms that orchestrate multisensory processing.

### Glossary

In accordance with previous work (Stein and others 2010), we use the following operational definitions

throughout this manuscript. Please note that these definitions are not intended to be final.

*Bottom-up:* Any process that involves mainly stimulus-driven processing.

*Top-down:* Any process that involves primarily cognitive processes, such as attention, expectations, and predictions.

*Unisensory stimuli:* Stimuli from a single sensory modality.

*Multisensory stimuli:* Two or more modality-specific stimuli from different sensory modalities.

*Modality-specific stimuli processing:* Processing of unisensory stimuli in the respective sensory cortical areas.

*Crossmodal:* Any process involving the influence from one sensory modality on processing or perception of another sensory modality.

*Multisensory processing:* Any process involving the simultaneous processing of one or more modality-specific stimuli from different sensory modalities.

*Multisensory integration:* Any process involving multisensory processing, in which the neural response is different from responses following the modality-specific responses.

*Crossmodal influence:* Any process in which processing of a modality-specific stimulus affects neural responses in cortical areas processing a different sensory modality, or the processing of a different modality specific stimulus.

*Multisensory perception:* Any process involving multisensory integration, in which the neural response is different to responses following the modality-specific responses and which results in a coherent percept.

#### Acknowledgments

We thank James Moran for help with preparation of this work. We also thank the anonymous reviewers for helpful suggestions.

#### **Declaration of Conflicting Interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

#### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by grants from the German Research Foundation (DFG: SE1859/3-1 and SE1859/4-1 to DS; KE1828/2-1 and KE1828/4-1 to JK), and the European Union (ERC-2010-StG-20091209 to DS).

#### Notes

- Studies included in this review were identified based on a PubMed search conducted on October 11, 2017, using the search term "(EEG OR MEG) AND (multisensory or crossmodal) AND ("2007/01/01"[Date - Publication]: "2017/12/31"[Date - Publication]) NOT review NOT (infant OR adolescent) NOT (mouse OR rat)".
- However, the authors questioned their own results due to a possible confounding influence of micro-saccades (Yuval-Greenberg and others 2008).

#### **ORCID** iD

Julian Keil (D) https://orcid.org/0000-0003-4195-7397

#### References

- Arnal LH, Giraud AL. 2012. Cortical oscillations and sensory predictions. Trends Cogn Sci 16:390–8.
- Arnal LH, Wyart V, Giraud AL. 2011. Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. Nat Neurosci 14:797–801.
- Auksztulewicz R, Friston KJ, Nobre AC. 2017. Task relevance modulates the behavioural and neural effects of sensory predictions. PLoS Biol 15:e2003143.
- Balz J, Keil J, Romero YR, Mekle R, Schubert F, Aydin S, and others. 2016a. GABA concentration in superior temporal sulcus predicts gamma power and perception in the soundinduced flash illusion. NeuroImage 125:724–30.
- Balz J, Roa Romero Y, Keil J, Krebber M, Niedeggen M, Gallinat J, and others. 2016b. Beta/gamma oscillations and event-related potentials indicate aberrant multisensory processing in schizophrenia. Front Psychol 7:106.
- Barutchu A, Freestone DR, Innes-Brown H, Crewther DP, Crewther SG. 2013. Evidence for enhanced multisensory facilitation with stimulus relevance: an electrophysiological investigation. PLoS One 8:e52978.
- Bauer M, Kennett S, Driver J. 2012. Attentional selection of location and modality in vision and touch modulates low-frequency activity in associated sensory cortices. J Neurophysiol 107:2342–51.
- Beauchamp MS, Sun P, Baum SH, Tolias AS, Yoshor D. 2012. Electrocorticography links human temporoparietal junction to visual perception. Nat Neurosci 15:957–9.
- Bhattacharya J, Shams L, Shimojo S. 2002. Sound-induced illusory flash perception: role of gamma band responses. Neuroreport 13:1727–30.
- 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.
- Cavanagh JF, Frank MJ. 2014. Frontal theta as a mechanism for cognitive control. Trends Cogn Sci 18:414–21.
- Cecere R, Rees G, Romei V. 2015. Individual differences in alpha frequency drive crossmodal illusory perception. Curr Biol 25:231–5.
- Covic A, Keitel C, Porcu E, Schröger E, Müller MM. 2017. Audio-visual synchrony and spatial attention enhance processing of dynamic visual stimulation independently and in parallel: a frequency-tagging study. NeuroImage 161:32–42.

- Craddock M, Poliakoff E, El-Deredy W, Klepousniotou E, Lloyd DM. 2016. Pre-stimulus alpha oscillations over somatosensory cortex predict tactile misperceptions. Neuropsychologia 96:9–18.
- Delorme A, Makeig S. 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 134:9–21.
- Engel AK, Fries P. 2010. Beta-band oscillations—signalling the status quo? Curr Opin Neurobiol 20:156–65.
- Evans N, Blanke O. 2013. Shared electrophysiology mechanisms of body ownership and motor imagery. NeuroImage 64:216–28.
- Fontolan L, Morillon B, Liegeois-Chauvel C, Giraud AL. 2014. The contribution of frequency-specific activity to hierarchical information processing in the human auditory cortex. Nature Commun 5:1–10.
- Fries P. 2015. Rhythms for cognition: communication through coherence. Neuron 88:220–35.
- Giordano BL, Ince R, Gross J, Schyns PG, Panzeri S, Kayser C. 2017. Contributions of local speech encoding and functional connectivity to audio-visual speech perception. Elife 6:e24763.
- Gips B, van der Eerden JPJM, Jensen O. 2016. A biologically plausible mechanism for neuronal coding organized by the phase of alpha oscillations. Eur J Neurosci 44:2147–61.
- Giraud AL, Poeppel D. 2012. Cortical oscillations and speech processing: emerging computational principles and operations. Nat Neurosci 15:511–7.
- Gleiss S, Kayser C. 2014a. Acoustic noise improves visual perception and modulates occipital oscillatory states. J Cogn Neurosci 26:699–711.
- Gleiss S, Kayser C. 2014b. Oscillatory mechanisms underlying the enhancement of visual motion perception by multisensory congruency. Neuropsychologia 53:84–93.
- Göschl F, Engel AK, Friese U. 2014. Attention modulates visual-tactile interaction in spatial pattern matching. PLoS One 9:e106896.
- Göschl F, Friese U, Daume J, König P, Engel AK. 2015. Oscillatory signatures of crossmodal congruence effects: an EEG investigation employing a visuotactile pattern matching paradigm. NeuroImage 116:177–86.
- Grabot L, Kösem A, Azizi L, van Wassenhove V. 2017. Prestimulus alpha oscillations and the temporal sequencing of audio-visual events. J Cogn Neurosci 83:1–17.
- Hamilton RH, Wiener M, Drebing DE, Coslett HB. 2013. Gone in a flash: manipulation of audiovisual temporal integration using transcranial magnetic stimulation. Front Psychol 4:571.
- He Y, Gebhardt H, Steines M, Sammer G, Kircher T, Nagels A, and others. 2015. The EEG and fMRI signatures of neural integration: An investigation of meaningful gestures and corresponding speech. Neuropsychologia 72:27–42.
- Henry MJ, Obleser J. 2012. Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. Proc Natl Acad Sci U S A 109:20095–100.
- Hipp JF, Engel AK, Siegel M. 2011. Oscillatory synchronization in large-scale cortical networks predicts perception. Neuron 69:387–96.
- Huang Y, Rao RPN. 2011. Predictive coding. WIREs Cogn Sci 2:580–93.

- Jensen O, Mazaheri A. 2010. Shaping functional architecture by oscillatory alpha activity: gating by inhibition. Front Hum Neurosci 4:186.
- Jessen S, Obleser J, Kotz SA. 2012. How bodies and voices interact in early emotion perception. PLoS One 7:e36070.
- Jones SR. 2016. When brain rhythms aren't "rhythmic": implication for their mechanisms and meaning. Curr Opin Neurobiol 40:72–80.
- Kambe J, Kakimoto Y, Araki O. 2015. Phase reset affects auditory-visual simultaneity judgment. Cogn Neurodyn 9:487–93.
- Kamke MR, Vieth HE, Cottrell D, Mattingley JB. 2012. Parietal disruption alters audiovisual binding in the sound-induced flash illusion. NeuroImage 62:1334–41.
- Kanayama N, Kimura K, Hiraki K. 2015. Cortical EEG components that reflect inverse effectiveness during visuotactile integration processing. Brain Res 1598:18–30.
- Kanayama N, Sato A, Ohira H. 2007. Crossmodal effect with rubber hand illusion and gamma-band activity. Psychophysiology 44:392–402.
- Kanayama N, Sato A, Ohira H. 2009. The role of gamma band oscillations and synchrony on rubber hand illusion and crossmodal integration. Brain Cogn 69:19–29.
- Kanayama N, Tamè L, Ohira H, Pavani F. 2012. Top down influence on visuo-tactile interaction modulates neural oscillatory responses. NeuroImage 59:3406–17.
- Kayser C, Petkov CI, Logothetis NK. 2008. Visual modulation of neurons in auditory cortex. Cerebral Cortex 18:1560–74.
- Kayser SJ, Philiastides MG, Kayser C. 2017. Sounds facilitate visual motion discrimination via the enhancement of late occipital visual representations. NeuroImage 148:31–41.
- Keil J, Müller N, Hartmann T, Weisz N. 2014. Prestimulus beta power and phase synchrony influence the sound-induced flash illusion. Cerebral Cortex 24:1278–88.
- 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 22:221–31.
- Keil J, Pomper U, Senkowski D. 2016. Distinct patterns of local oscillatory activity and functional connectivity underlie intersensory attention and temporal prediction. Cortex 74:277–88.
- Keil J, Senkowski D. 2017. Individual alpha frequency relates to the sound-induced flash illusion. Multisensory Res 30:565–78.
- Keller AS, Payne L, Sekuler R. 2017. Characterizing the roles of alpha and theta oscillations in multisensory attention. Neuropsychologia 99:48–63.
- Kösem A, Gramfort A, van Wassenhove V. 2014. Encoding of event timing in the phase of neural oscillations. NeuroImage 92:274–84.
- Krebber M, Harwood J, Spitzer B, Keil J, Senkowski D. 2015. Visuotactile motion congruence enhances gamma-band activity in visual and somatosensory cortices. NeuroImage 117:160–9.
- Kumagai T, Mizuhara H. 2016. Top-down and bottom-up attention cause the ventriloquism effect with distinct electroencephalography modulations. Neuroreport 27:647–51.
- Kumar GV, Halder T, Jaiswal AK, Mukherjee A, Roy D, Banerjee A. 2016. Large scale functional brain networks

underlying temporal integration of audio-visual speech perception: an EEG study. Front Psychol 7:1321.

- Lakatos P, Shah AS, Knuth KH, Ulbert I, Karmos G, Schroeder CE. 2005. An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. J Neurophysiol 94:1904–11.
- Lakatos P, Chen CM, O'Connell MN, Mills A, Schroeder CE. 2007. Neuronal oscillations and multisensory interaction in primary auditory cortex. Neuron 53:279–92.
- Lange J, Christian N, Schnitzler A. 2013a. Audio-visual congruency alters power and coherence of oscillatory activity within and between cortical areas. NeuroImage 79:111–20.
- Lange J, Keil J, Schnitzler A, Van Dijk H, Weisz N. 2014. The role of alpha oscillations for illusory perception. Behav Brain Res 271:294–301.
- Lange J, Oostenveld R, Fries P. 2011. Perception of the touchinduced visual double-flash illusion correlates with changes of rhythmic neuronal activity in human visual and somatosensory areas. NeuroImage 54:1395–405.
- Lange J, Oostenveld R, Fries P. 2013b. Reduced occipital alpha power indexes enhanced excitability rather than improved visual perception. J Neurosci 33:3212–20.
- Lenggenhager B, Halje P, Blanke O. 2011. Alpha band oscillations correlate with illusory self-location induced by virtual reality. Eur J Neurosci 33:1935–43.
- 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. Hum Brain Mapp 36:3486–98.
- Lin Y, Liu B, Liu Z, Gao X. 2015. EEG gamma-band activity during audiovisual speech comprehension in different noise environments. Cogn Neurodyn 9:389–98.
- Macaluso E, Noppeney U, Talsma D, Vercillo T, Hartcher-O'Brien J, Adam R. 2016. The curious incident of attention in multisensory integration: bottom-up vs. top-down. Multisensory Res 29:557–83.
- Maris E, Fries P, van Ede F. 2016. Diverse phase relations among neuronal rhythms and their potential function. Trends Neurosci 39:86–99.
- McGurk H, MacDonald J. 1976. Hearing lips and seeing voices. Nature 264:746–8.
- Mercier MR, Foxe JJ, Fiebelkorn IC, Butler JS, Schwartz TH, Molholm S. 2013. Auditory-driven phase reset in visual cortex: Human electrocorticography reveals mechanisms of early multisensory integration. NeuroImage 79:19–29.
- Mercier MR, Molholm S, Fiebelkorn IC, Butler JS, Schwartz TH, Foxe JJ. 2015. Neuro-oscillatory phase alignment drives speeded multisensory response times: an electrocorticographic investigation. J Neurosci 35:8546–57.
- Mégevand P, Molholm S, Nayak A, Foxe JJ. 2013. Recalibration of the multisensory temporal window of integration results from changing task demands. PLoS One 8:e71608.
- Michalareas G, Vezoli J, van Pelt S, Schoffelen JM, Kennedy H, Fries P. 2016. Alpha-beta and gamma rhythms subserve feedback and feedforward influences among human visual cortical areas. Neuron 89:384–97.
- Miller J. 1986. Timecourse of coactivation in bimodal divided attention. Perception Psychophysics 40:331–43.

- Mishra J, Martinez A, Sejnowski TJ, Hillyard SA. 2007. Early cross-modal interactions in auditory and visual cortex underlie a sound-induced visual illusion. J Neurosci 27:4120–31.
- Mitra PP, Pesaran B. 1999. Analysis of dynamic brain imaging data. Biophys J 76:691–708.
- Morís Fernández L, Macaluso E, Soto-Faraco S. 2017a. Audiovisual integration as conflict resolution: the conflict of the McGurk illusion. Hum Brain Mapp 38:5691–705.
- Morís Fernández L, Torralba Cuello M, Soto-Faraco S. 2017b. Theta oscillations reflect conflict processing in the perception of the McGurk illusion. Eur J Neurosci. Epub Dec 18. doi:10.1111/ejn.13804.
- Ohki T, Gunji A, Takei Y, Takahashi H, Kaneko Y, Kita Y, and others. 2016. Neural oscillations in the temporal pole for a temporally congruent audio-visual speech detection task. Sci Rep 6:37973.
- Peelle JE, Sommers MS. 2015. Prediction and constraint in audiovisual speech perception. Cortex 68:169–81.
- 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:e103238.
- Pomper U, Keil J, Foxe JJ, Senkowski D. 2015. Intersensory selective attention and temporal orienting operate in parallel and are instantiated in spatially distinct sensory and motor cortices. Hum Brain Mapp 36:3246–59.
- Quinn BT, Carlson C, Doyle W, Cash SS, Devinsky O, Spence C, and others. 2013. Intracranial cortical responses during visual-tactile integration in humans. J Neurosci 34:171–81.
- Rao IS, Kayser C. 2017. Neurophysiological correlates of the rubber hand illusion in late evoked and alpha/beta band activity. Front Hum Neurosci 11:777–12.
- Rhone AE, Nourski KV, Oya H, Kawasaki H, Howard MA III, McMurray B. 2015. Can you hear me yet? An intracranial investigation of speech and non-speech audiovisual interactions in human cortex. Lang Cogn Neurosci 31:284–302.
- Roa Romero Y, Keil J, Balz J, Gallinat J, Senkowski D. 2016a. Reduced frontal theta oscillations indicate altered crossmodal prediction error processing in schizophrenia. J Neurophysiol 116:1396–407.
- Roa Romero Y, Keil J, Balz J, Niedeggen M, Gallinat J, Senkowski D. 2016b. Alpha-band oscillations reflect altered multisensory processing of the McGurk illusion in schizophrenia. Front Hum Neurosci 10:41.
- Roa Romero Y, Senkowski D, Keil J. 2015. Early and late betaband power reflect audiovisual perception in the McGurk illusion. J Neurophysiol 113:2342–50.
- Romei V, Gross J, Thut G. 2012. Sounds reset rhythms of visual cortex and corresponding human visual perception. Curr Biol 22:807–13.
- Ronconi L, Oosterhof NN, Bonmassar C, Melcher D. 2017. Multiple oscillatory rhythms determine the temporal organization of perception. Proc Natl Acad Sci U S A 103:13435–40.
- Schelenz PD, Klasen M, Reese B. 2013. Multisensory integration of dynamic emotional faces and voices: method

for simultaneous EEG-fMRI measurements. Front Hum Neurosci 7:729.

- Schepers IM, Schneider TR, Hipp JF, Engel AK, Senkowski D. 2013. Noise alters beta-band activity in superior temporal cortex during audiovisual speech processing. NeuroImage 70:101–12.
- Senkowski D, Schneider T, Foxe J, Engel A. 2008. Crossmodal binding through neural coherence: implications for multisensory processing. Trends Neurosci 31:401–9.
- Senkowski D, Talsma D, Grigutsch M, Herrmann CS, Woldorff MG. 2007. Good times for multisensory integration: effects of the precision of temporal synchrony as revealed by gamma-band oscillations. Neuropsychologia 45:561–71.
- Senkowski D, Talsma D, Herrmann CS, Woldorff MG. 2005. Multisensory processing and oscillatory gamma responses: effects of spatial selective attention. Exp Brain Res 166:411–26.
- Shams L, Kamitani Y, Shimojo S. 2000. Illusions. What you see is what you hear. Nature 408:788.
- Siegel M, Donner TH, Engel AK. 2012. Spectral fingerprints of large-scale neuronal interactions. Nat Rev Neurosci 13:121–34.
- Simon DM, Wallace MT. 2017. Rhythmic modulation of entrained auditory oscillations by visual inputs. Brain Topogr 30:565–78.
- Stein BE, Burr D, Constantinidis C, Laurienti PJ, Alex Meredith M, Perrault TJ, Jr, and others. 2010. Semantic confusion regarding the development of multisensory integration: a practical solution. Eur J Neurosci 31:1713–20.
- Stein BE, Meredith MA. 1993. The merging of the senses. Cambridge: MIT Press.
- Tallon-Baudry C, Bertrand O. 1999. Oscillatory gamma activity in humans and its role in object representation. Trends Cogn Sci 3:151–62.
- van Atteveldt N, Murray MM, Thut G, Schroeder CE. 2014. Multisensory integration: flexible use of general operations. Neuron 81:1240–53.
- van Kerkoerle T, Self MW, Dagnino B, Gariel-Mathis MA, Poort J, van der Togt C, and others. 2014. Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. Proc Natl Acad Sci U S A 111:14332–41.
- Wang XJ. 2010. Neurophysiological and computational principles of cortical rhythms in cognition. Physiol Rev 90: 1195–268.
- Yuan X, Li H, Liu P, Yuan H, Huang X. 2016. Pre-stimulus beta and gamma oscillatory power predicts perceived audiovisual simultaneity. Int J Psychophysiol 107: 29–36.
- Yuval-Greenberg S, Deouell LY. 2007. What you see is not (always) what you hear: induced gamma band responses reflect cross-modal interactions in familiar object recognition. J Neurosci 27:1090–6.
- Yuval-Greenberg S, Tomer O, Keren AS, Nelken I, Deouell LY. 2008. Transient induced gamma-band response in EEG as a manifestation of miniature saccades. Neuron 58:429–41.