Fachbereich Erziehungswissenschaft und Psychologie
der Freien Universität Berlin

Neural oscillatory signatures of auditory and audiovisual illusions

Dissertation
zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften (Dr. rer. nat.)

vorgelegt von
Dipl. Psych. Mathis Kaiser

Berlin, 2020

Erstgutachter:	Prof. Dr. Felix Blankenburg
Zweitgutachter:	Prof. Dr. Julian Keil

Tag der Disputation: 26.11.2020

Abstract

Questions of the relationship between human perception and brain activity can be approached from different perspectives: in the first, the brain is mainly regarded as a recipient and processor of sensory data. The corresponding research objective is to establish mappings of neural activity patterns and external stimuli. Alternatively, the brain can be regarded as a self-organized dynamical system, whose constantly changing state affects how incoming sensory signals are processed and perceived.

The research reported in this thesis can chiefly be located in the second framework, and investigates the relationship between oscillatory brain activity and the perception of ambiguous stimuli. Oscillations are here considered as a mechanism for the formation of transient neural assemblies, which allows efficient information transfer. While the relevance of activity in distinct frequency bands for auditory and audiovisual perception is well established, different functional architectures of sensory integration can be derived from the literature. This dissertation therefore aims to further clarify the role of oscillatory activity in the integration of sensory signals towards unified perceptual objects, using illusion paradigms as tools of study.

In study 1, we investigate the role of low frequency power modulations and phase alignment in auditory object formation. We provide evidence that auditory restoration is associated with a power reduction, while the registration of an additional object is reflected by an increase in phase locking. In study 2, we analyze oscillatory power as a predictor of auditory influence on visual perception in the sound-induced flash illusion. We find that increased beta-/gamma-band power over occipitotemporal electrodes shortly before stimulus onset predicts the illusion, suggesting a facilitation of processing in polymodal circuits. In study 3, we address the question of whether visual influence on auditory perception in the ventriloquist illusion is reflected in primary sensory or higher-order areas. We establish an association between reduced theta-band power in mediofrontal areas and the occurrence of illusion, which indicates a top-down influence on sensory decision-making. These findings broaden our understanding of the functional relevance of neural oscillations by showing that different processing modes, which are reflected in specific spatiotemporal activity patterns, operate in different instances of sensory integration.

Zusammenfassung

Fragen nach dem Zusammenhang zwischen menschlicher Wahrnehmung und Hirnaktivität können aus verschiedenen Perspektiven adressiert werden: in der einen wird das Gehirn hauptsächlich als Empfänger und Verarbeiter von sensorischen Daten angesehen. Das entsprechende Forschungsziel wäre eine Zuordnung von neuronalen Aktivitätsmustern zu externen Reizen. Dieser Sichtweise gegenüber steht ein Ansatz, der das Gehirn als selbstorganisiertes dynamisches System begreift, dessen sich ständig verändernder Zustand die Verarbeitung und Wahrnehmung von sensorischen Signalen beeinflusst.

Die Arbeiten, die in dieser Dissertation zusammengefasst sind, können vor allem in der zweitgenannten Forschungsrichtung verortet werden, und untersuchen den Zusammenhang zwischen oszillatorischer Hirnaktivität und der Wahrnehmung von mehrdeutigen Stimuli. Oszillationen werden hier als ein Mechanismus für die Formation von transienten neuronalen Zusammenschlüssen angesehen, der effizienten Informationstransfer ermöglicht. Obwohl die Relevanz von Aktivität in verschiedenen Frequenzbändern für auditorische und audiovisuelle Wahrnehmung gut belegt ist, können verschiedene funktionelle Architekturen der sensorischen Integration aus der Literatur abgeleitet werden. Das Ziel dieser Dissertation ist deshalb eine Präzisierung der Rolle oszillatorischer Aktivität bei der Integration von sensorischen Signalen zu einheitlichen Wahrnehmungsobjekten mittels der Nutzung von Illusionsparadigmen.

In der ersten Studie untersuchen wir die Rolle von Leistung und Phasenanpassung in niedrigen Frequenzbändern bei der Formation von auditorischen Objekten. Wir zeigen, dass die Wiederherstellung von Tönen mit einer Reduktion der Leistung zusammenhängt, während die Registrierung eines zusätzlichen Objekts durch einen erhöhten Phasenangleich widergespiegelt wird. In der zweiten Studie analysieren wir oszillatorische Leistung als Prädiktor von auditorischem Einfluss auf visuelle Wahrnehmung in der *sound-induced flash illusion*. Wir stellen fest, dass erhöhte Beta-/Gamma-Band Leistung über occipitotemporalen Elektroden kurz vor der Reizdarbietung das Auftreten der Illusion vorhersagt, was auf eine Begünstigung der Verarbeitung in polymodalen Arealen hinweist. In der dritten Studie widmen wir uns der Frage, ob ein visueller Einfluss auf auditorische Wahrnehmung in der *ventriloquist illusion* sich in primären sensorischen oder übergeordneten Arealen widerspiegelt. Wir weisen einen Zusammenhang von reduzierter Theta-Band Leistung in mediofrontalen Arealen und dem Auftreten der Illusion nach, was einen *top-down* Einfluss auf sensorische Entscheidungsprozesse anzeigt. Diese Befunde erweitern unser Verständnis der funktionellen Bedeutung neuronaler Oszillationen,

indem sie aufzeigen, dass verschiedene Verarbeitungsmodi, die sich in spezifischen räumlich-zeitlichen Aktivitätsmustern spiegeln, in verschiedenen Phänomenen von sensorischer Integration wirksam sind.

Danksagung

Diese Arbeit wäre nicht möglich gewesen ohne die Begleitung und Unterstützung durch viele verschiedene Menschen, denen ich an dieser Stelle meinen Dank aussprechen will:

Julian: Danke, dass du die Grundlage für meinen akademischen Werdegang durch Einführung in die MEG-Forschung in Konstanz sowie die Gelegenheit zu diesem Projekt geschaffen hast, und ihm durch stetige Rückendeckung und Mentoring zum Abschluss verholfen hast. Danke für dein Vertrauen! Ich bin froh, dass unsere Wege sich mehrmals so produktiv gekreuzt haben, und hoffe dass das auch in Zukunft wieder passiert.

Daniel: Danke für die freundliche Aufnahme in deine Arbeitsgruppe, und damit die praktischen Voraussetzungen für die empirische Forschung, sowie hilfreiche Wegweiser beim wissenschaftlichen Schreiben und Publizieren, und den einen oder anderen lustigen Glühweinabend.

Meinen Betreuern an der Berlin School of Mind and Brain, Felix und Niko für konstruktives Feedback zu meinen Arbeiten und ihre kontinuierliche Bereitschaft zur Unterstützung.

Meinen Kolleg*innen in der Arbeitsgruppe Johanna, Yadira, Martin, Georgios und James: es war mir eine große Freude mit euch zu arbeiten (und zu pausieren). Der Austausch mit euch hat viel dazu beigetragen, die kleinen Frustrationen und Hindernisse bei der wissenschaftlichen Arbeit zu bewältigen.

Joseph für tatkräftige Unterstützung beim Erheben der EEG-Daten, Christian für die Durchführung eines Großteils der MRT-Aufnahmen, und allen anderen Hilfskräften, die einen Beitrag geleistet haben.

Meinen alten wissenschaftlichen Wegbereiter*innen und -gefährt*innen in Konstanz und Trento, Nathan, Nadia, Thomas, Sarang, und Britta.

Meinen Freund*innen aus den Zeiten in Freiburg, Konstanz, Toronto und Berlin: Danke euch für die fortwährende Verbundenheit, die so schöne Teile meines Lebens ausmacht und die mich die härteren Zeiten durchstehen ließ. Ein persönlicher Dank geht an meine ehemaligen Mitbewohner*innen in Berlin und lieben Freund*innen Julia, Daniel, Christoph, und Robin, deren Gesellschaft ich nie missen möchte.

Brian nicht nur für das Korrekturlesen dieser Dissertation, sondern vor allem für unsere lange, enge Freundschaft, die auch mein Ankommen in Berlin so beschwingt hat.

Meinem Bruder Niklas, ohne dessen liebevolle Sorge für unsere Großeltern ich niemals den nötigen Bewegungsspielraum für dieses Promotionsprojekt gehabt hätte. Ich bin dankbar, dass wir uns so nahe sind.

Meinem absoluten Lieblingsmenschen Larissa: ohne dich hätte ich das alles nicht geschafft. Danke, dass du an meiner Seite bist! Unsere beständig wachsende Liebe erfüllt mich mit Glück, Neugier und Zuversicht auf alles was kommen mag.

Besonders auch denen, die wirklich immer für mich da waren und mir alles mitgegeben haben, bei diesem Abschluss aber nicht mehr selbst dabei sein können: Brigitte, Renate und Leo. Diese Dissertation ist ihrem Andenken gewidmet.

List of original articles

This dissertation is based on the following articles:

Kaiser, M., Senkowski, D., Roa Romero, Y., Riecke, L., & Keil, J. (2018). Reduced low-frequency power and phase locking reflect restoration in the auditory continuity illusion. *European Journal of Neuroscience, 48,* 2849-2856.

Kaiser, M., Senkowski, D., Busch, N. A., Balz, J., & Keil, J. (2019). Single trial prestimulus oscillations predict perception of the sound-induced flash illusion. *Scientific Reports*, *9*, 5983.

Kaiser, M., Senkowski, D., Keil J. (2020). Mediofrontal Theta-band oscillations reflect top-down influence in the ventriloquist illusion. *Manuscript submitted for publication.*

Contents

Abbreviations

BOLD	blood-oxygen-level-dependent
EEG	electroencephalography
ERP	event-related potential
fMRI	functional magnetic resonance imaging
GABA	gamma-Aminobutyric acid
ITC	inter-trial coherence
MEG	magnetoencephalography
SIFI	sound-induced flash illusion

1 Introduction

Research on the neural foundations of perception has long operated in a framework where the brain reacts to external stimuli and then further processes these inputs, constructing increasingly complex representations that underlie generalized recognition (DeYoe & Van Essen, 1988; Riesenhuber & Poggio, 1999). However, there is now extensive evidence that neural activity is not a mere consequence of sensory input, but can have an influence on how that input is processed. An example for this notion of the brain always being in dynamic exchange with its environment is presented by predictive coding theories (Clark, 2013; Rao & Ballard, 1999), where higher-order areas constantly generate predictions about lower-order neural activity elicited by perceptual objects, and the predictions are calibrated using the available sensory evidence. In much simpler terms, the activation pattern of neural ensembles at a given point in time may have an influence on how incoming signals are processed (Arieli et al., 1996), an idea already advanced by Hebb (1949, p. 7). The research in this thesis is inspired by the idea that ongoing, rhythmic activity in the brain reflects network states that are closely linked to the integration and segregation of sensory processing streams, and ultimately to the coherent representation and conscious awareness of one's environment. Oscillations are here regarded as a functional organization principle of brain activity that supports effective information transfer between neural populations, and thereby the routing, filtering, and reverberation of sensory data, as well as the implementation of predictions and mental imagery. Hence, it can be hypothesized that oscillations reflect or even determine the subjective experience in a given sensory situation. The studies reported here investigate this hypothesis by exploring the relationship between specific spatiotemporal patterns of neural oscillatory activity and the perception of ambiguous stimuli.

In the course of this introduction, I will first provide a brief overview of the generation mechanisms, measurement, and analysis of cortical oscillations. I will then discuss their relevance for efficient neuronal communication and the resulting implications for sensory processing. Afterwards, I will review the functional roles of oscillations in distinct frequency bands, especially in the context of auditory, visual and audiovisual perception. Subsequently, I will summarize the literature on auditory perception and multisensory integration on the backdrop of the conducted experiments. I will thereby develop a methodological and theoretical framework where the empirical studies of this dissertation are embedded. Finally, I will derive the aims of this dissertation from the discussed literature.

1.1 The generation and measurement of neural oscillations

In this section, I will briefly summarize the generation of the signals measured in electroencephalography (EEG), following a recent review by Buzsáki, Anastassiou and Koch (2012). Ionic processes at neuronal membranes result in current flows in the surrounding brain volume. At any given location, a voltage potential can then be measured with respect to a reference. Due to the conductance of the intracranial, skull and scalp tissues, these potentials can also be measured on the level of the head surface via EEG. The measured signal depends on the magnitude, sign, orientation, spatial density and temporal synchrony of the sources, and on the conductivities of the intermediate tissues. It decreases with distance between measured source and recorded site and is subject to spatial averaging. Therefore, in contrast to intracranial methods, the EEG has no identifiable relationship with neuronal firing rates in the underlying tissues, but mainly reflects current flows along spatially aligned apical dendrites of cortical neurons. These current flows are due to synchronous postsynaptic potentials. Thus, there are spatial and temporal constraints on the activity that is measurable with EEG. As I will outline in more detail below, the synchrony of neuronal activity has a close relationship with network oscillations, and these oscillations constitute the focus of the research performed in this thesis. But what are the building blocks of temporally coordinated neural activity?

On the level of connected single neurons, the generation of oscillations can be ascribed to certain properties of ion channels (Llinas, 1988). These properties allow networks or single neurons to act as pacemakers (i.e. they generate oscillations) or resonators (i.e. they preferentially respond to and maintain activity at a given frequency). Resonance can arise from a combination of low- and highpass filtering at the levels of membrane and voltage-gated ion channels, respectively (Hutcheon & Yarom, 2000). For instance, many interneurons have resonator properties in the gamma-band (Bartos et al., 2007), likely supporting binding and segregation of neuronal assemblies. Oscillations at high frequencies are also supported by direct electrical coupling via gap junctions (Draguhn et al., 1998).

The field oscillations measured in EEG do not directly depend on individual firing rates, but emerge on the population level. Moving to the level of minimal network configurations capable of generating oscillations, there are excitatory-excitatory, inhibitory-inhibitory, and excitatory-inhibitory networks. The terms of excitation and inhibition refer to the depolarization and hyperpolarization of the postsynaptic neuronal membrane, resulting from the release of the neurotransmitters glutamate and gamma-Aminobutyric acid (GABA) from the presynapse and their binding to postsynaptic receptors, respectively. Although intuitively capable of generating oscillatory dynamics, a mechanism of mutual excitation cannot generate robust fast oscillations due to the slow time constant of glutamatergic decay (Wang, 2010). Reciprocal inhibition in a network of GABAergic interneurons, on the other hand, has been shown to produce synchronized responses in a wide frequency range, depending on the external input strength and synaptic weights (Wang & Buzsáki, 1996). This is due to a coupling by slow synapses leading to simultaneous inhibition and firing cycles. Another network mechanism for the generation of synchronous oscillations consists in feedback loops of excitatory populations with strong and delayed inhibition. Here, excitatory neurons slowly activate inhibitory ones, which eventually decrease population firing, leading to decreased self-inhibition, which then drives population activity up again (Wang, 2010).

The signals measured in EEG reflect the activity of much more complex networks than the minimal configurations described above. They represent changes in the relative level of membrane depolarization in large numbers of neurons, with artifactual muscle activity and external noise superimposed. To increase signal-to-noise ratio, EEG data has traditionally been analyzed by averaging the time courses, time locked to events of interest, across trials of a particular experimental condition. This results in condition-specific waveforms called event-related potentials (ERPs) that can be statistically compared (Luck, 2014). However, some researchers consider ERPs to arise from a phase reset of ongoing oscillations, rather than from the activation of previously idling neurons by external input (Makeig et al., 2002; Sauseng et al., 2007; but see Mazaheri & Jensen, 2006). This idea is in line with the notion that the brain state at the time of stimulus onset has an effect on further processing. More importantly, there is an implicit assumption in ERP analyses that external stimuli evoke activity in a phase-locked manner. Any activity that is non-phase-locked or occurs in the prestimulus baseline is considered random variation to be averaged out. However, the perspective taken here is that seemingly random variation should be regarded as an effect of self-organized activity, which reflects current or past network states. Therefore, non-phase-locked responses and the prestimulus period may contain meaningful information that is not captured by ERP analyses, and investigating the brain states that predispose or ensue variability in perception may provide insights into necessary conditions for the phenomena of interest. An analysis method more suitable to this perspective is the transformation of time-series data to the frequency domain, thereby preserving sensitivity to non-phase-locked oscillatory components (Makeig et al., 2004). Different properties of the oscillation such as the amplitude in specific bands, or the phase relationships across trials or locations can then be related to cognitive processes.

1.2 The functional relevance of neural oscillations

On a fundamental level, oscillations can be employed to integrate and segregate neural information flows, as well as provide a framework for the temporal ordering of perception and action. In terms of purely physiological limitations, without taking into account functional questions, oscillations play an important role in effective signaling along neural pathways. In complex networks, the timing of events plays a crucial role for further processing. A small number or irregular timing of incoming excitatory postsynaptic potentials does usually not suffice to discharge a target neuron, but the temporal coordination of input results in increased efficiency (König et al., 1996). Oscillations offer a mechanism to modulate both the firing rates of neurons in the network and the excitability of the postsynaptic side via the membrane potential. On the level of single neurons, this idea is supported by in vitro results which showed that the phase of subthreshold membrane oscillations affect the further processing of incoming synaptic potentials and ultimately the output, thereby creating temporal windows of information transfer (Lampl & Yarom, 1993). Generally, the integration window is determined by the time within which a postsynaptic potential returns to baseline, which is in turn affected by previous input and therefore by the network activity. On the level of populations, the temporal windows of integration are determined by the excitatory states of ongoing oscillations. Slow oscillations have longer integration windows, which makes them suitable for the recruitment of larger populations because they are less limited by conductance speed, while faster oscillations are thought to integrate information more locally (Buzsáki, 2006). Harris et al. (2003) proposed that the optimal time window for neural synchrony and communication falls together with the period of the gamma oscillation and matches the time window for synaptic plasticity. As a general framework about the consequences of oscillations for neuronal communication, Fries (2005, 2015) advanced the communication-through-coherence hypothesis, which states that oscillations constitute excitability changes that affect the output and the sensitivity to input of neuronal assemblies. Consequently, Womelsdorf et al. (2007) experimentally demonstrated a dependence of mutual influence between neuronal groups on the ongoing phase in the gamma-band.

These considerations lead to the question of the functional relevance of efficient information routing in neural networks. A fundamental problem where the importance becomes obvious is the binding problem, which can be stated as follows: at a given point in time, a large number of stimuli are processed by the sensory systems. A single visual object elicits activity in different brain areas specialized e.g. for shape, color and movement processing. These activity patterns have to be bound together to coherently represent the object, and this process has to take place for many objects simultaneously and track changes of the visual scene. The problem can be extended to include the binding of information from different sensory modalities (see section 1.5.3), and the matching of sensory and motor activity in decision-making (Roelfsema et al., 1997). Traditionally, the binding of different stimulus features has been assumed to be accomplished by a feed-forward hierarchical processing scheme, where features are bound by upstream neurons. While this model is well supported for basic visual features (Hubel & Wiesel, 1968), it is difficult to reconcile with the abundance of feedback pathways in the brain and faces other problems such as single points of failure. An elegant solution to the binding problem is provided by synchronization: the necessary neural assemblies can be formed by temporally coherent activity. The temporal coincidence of signals provides a tag for perceptual binding or integration, while signals that diverge in time are segregated. This "temporal binding", or "binding by synchronization" hypothesis was first fully formulated by von der Malsburg (1981), experimentally corroborated by the research group of Singer (Gray et al., 1989; for review see Singer, 1999; Singer & Gray, 1995), and informed later theories such as communication-through-coherence (Fries, 2005, 2015). The simultaneous, parallel and distributed processing mode postulated by these authors also allows rapid engagement and is therefore well-suited for dynamic environments. One corollary of the binding by synchronization hypothesis is that the temporal organization of neural activity is crucial for perception, and therefore, electrophysiological methods with sufficient temporal resolution are necessary to shed light on the relevant mechanisms.

Related to the binding problem is the question how high-level sensory objects are represented in the brain, since there is a near infinite combination of basic sensory properties in the external world. This again poses a problem for hierarchical accounts. Admittedly, Quiroga et al. (2005) demonstrated the selective activation of single neurons by different pictures and the written name of individuals, which exemplifies a sparse coding scheme. Notwithstanding, the representation of high-level perceptual objects can likely not be accomplished by a finite number of single neurons in general. The problem can be more readily solved by population coding, where properties are represented by patterns of network activation, which are instantiated by oscillations. Oscillations therefore provide a flexible neural architecture for the representation of perceptual objects. However, although a discussion of the evidence contradicting the temporal binding hypothesis is beyond the scope of this introduction, I want to point out that it is not without critics (see e.g. Shadlen & Movshon, 1999).

In an influential review, Engel and colleagues (2001) discuss the role of oscillations in top-down processing. Building on the temporal binding hypothesis, they consider the evidence for a dynamicist account of top-down processing, where larger-scale dynamics reflecting attention, expectations or predictions about the sensory environment can influence local neuronal activity. Here, internally generated activity is thought to carry templates of expected patterns against which afferent sensory information is matched, thus relating to the central ideas of predictive coding (Rao & Ballard, 1999). The sensitivity to specific input patterns could then be increased by coordinated excitability changes across the network. This idea is supported by evidence for attentional modulation of synchronized firing (Steinmetz et al., 2000) and expectancy-related increase of blood-oxygen-level-dependent (BOLD) signals in extrastriate areas (Kastner et al., 1999).

Based on these considerations, it becomes clear that a general function of oscillations consists in the dynamic formation of neural assemblies. The resulting functional network topology determines how information is transferred along existing anatomical pathways, thereby affecting if processing occurs locally or is spread to other brain regions. According to Varela et al. (2001), the large-scale integration of distributed neural processing is accomplished by transient phase synchronization between involved assemblies, which results in metastable activation states. In recent years, the temporal correlation patterns in resting state electrophysiological data have been used to characterize functional networks that show a good spatial agreement with those derived from functional magnetic resonance imaging (fMRI) data (Brookes et al., 2011). Furthermore, the fluctuation patterns of oscillatory power across the brain can be leveraged to identify frequency-specific network hubs (Hipp et al., 2012).

1.3 A functional taxonomy of EEG oscillations

As described above, oscillations are assumed to fulfill important functions in terms of effective neural signaling, integration and segregation of sensory input, top-down modulation of processing, and the formation of neural networks. These functions are supported by oscillations in distinct frequency bands, which are discernible in electrophysiological measurements. The long-term spectrum of neural activity is characterized by a 1/f power density, which can be observed at small and large spatial scales and across species, and is likely composed by a spatiotemporal integration of various local oscillators (Buzsáki, 2006). On shorter time scales, transient deviations in spectral power can occur, from which an inference on neural processes relevant for perception is possible. Although the exact spectral range of individual bands is a matter of debate, the EEG spectrum is usually separated into Delta (1-3 Hz), Theta (4-7 Hz), Alpha (8-12 Hz), Beta (12-30 Hz), and Gamma (30-100 Hz) frequency bands. A similar spectral classification can be formed by linear progression on a natural logarithmic scale (Buzsáki & Draguhn, 2004). I now turn to the question which modulations in specific frequency bands of cortical oscillations have been shown to be associated with functions that are relevant in the context of this thesis, progressing from lower to higher frequencies. My account of a functional taxonomy will necessarily be condensed and exclude some areas such as long-term memory (Axmacher et al., 2006), sleep (Siapas & Wilson, 1998), motor function (Pfurtscheller & Lopes da Silva, 1999) and psychiatric disorders (Uhlhaas & Singer, 2006).

1.3.1 Delta and Theta

Oscillations in the delta- and theta-band in auditory cortex play a critical role in aligning time windows of optimal neural excitability with the envelope of incoming speech signals, which have a similar temporal structure. They thereby support parsing of the auditory input stream (Doelling et al., 2014; Giraud & Poeppel, 2012). These findings relate to the general idea that the entrainment of slow oscillations can serve as a selection mechanism for rhythmical input in different sensory modalities, affecting local processing (as reflected by increased gamma-band power) and behavioral response speed (Lakatos et al., 2008). Moreover, slow oscillations in sensory areas could implement a temporal prediction mechanism for regular stimuli (Arnal & Giraud, 2012). Going beyond immediate sensory processing, Theta oscillations that originate in the frontal cortex are assumed to reflect canonical computations relevant for executive functions such as cognitive control and action monitoring (Cavanagh & Frank, 2014). An increase in frontal theta power can consistently be observed for novel stimuli, response and stimulus conflicts, and behavioral errors (Cavanagh et al., 2012). Theta modulations can be observed during working memory retention (Raghavachari et al., 2001; Sarnthein et al., 1998), and they index task demands such as memory load (Jensen & Tesche, 2002) and interference during the Stroop task (Hanslmayr et al., 2008).

1.3.2 Alpha

An increase of alpha-band activity over occipital cortex when the eyes are closed has been observed in pioneering EEG studies (Adrian & Matthews, 1934; Berger, 1929) and has first been associated with a state of neural idling (Pfurtscheller et al., 1996). More recently, this passive notion of the role of alpha has been amended by a more active one, where alpha reflects the functional inhibition of neural populations. This view is supported by studies reporting an increase of alpha power when the retrieval of information is suppressed, when responses are withheld, or when sensory processing is inhibited by top-down control (Klimesch et al., 2007). Similarly, alpha activity is lower over the occipital cortex contralateral to visually attended locations (Worden et al., 2000), and the interhemispheric alpha balance in attention tasks predicts the detection of lateral targets (Thut et al., 2006). This attention-indexing role of alpha can also be observed in intersensory paradigms, where alpha is higher over occipital cortex when attention is directed to the auditory compared to the visual modality (Foxe et al., 1998). Based on these findings, Jensen and Mazaheri (2010) proposed that alpha activity reflects the gating of information by an inhibition of task-irrelevant populations. This gating mechanism is assumed to operate in a phase-dependent manner, with cycles of decreased excitability providing transient inhibition windows for sensory processing. Accordingly, the phase of ongoing alpha oscillations in occipital areas around stimulus onset predicts whether participants can detect upcoming visual stimuli that are masked or near the perceptual threshold (Busch et al., 2009; Mathewson et al., 2009). Furthermore, modulations of alpha activity have been associated with visual, auditory and multisensory illusions, where perception differs from physical stimulus properties (Lange et al., 2014). A common theme in the literature is that alpha power is interpreted to be inversely related with excitability, which accordingly affects stimulus processing.

1.3.3 Beta

Moving up the spectrum, activity in the beta-band is thought to signal the maintenance of the current sensorimotor or cognitive set and the dominance of top-down influence over the processing of novel sensory input (Engel & Fries, 2010). This view was recently extended by suggesting that beta-band mediated population activity supports the formation and reactivation of content-specific representations of sensory input that is relevant to the current task (Spitzer & Haegens, 2017). One example for this are correlates of working memory that are independent of the sensory modality (Spitzer et al., 2014). Oscillations in the beta-band likely underlie the dynamic formation of connections between different cortical areas that are necessary for sensorimotor integration, supramodal representations and top-down modulations (Brovelli et al., 2004; Gross et al., 2004; von Stein et al., 1999). Therefore, beta-band activity has been proposed to be a signature of canonical computations in these contexts, which require large-scale neuronal interactions (Siegel et al., 2012).

1.3.4 Gamma

As mentioned previously, oscillations in higher frequency bands such as gamma have been demonstrated to reflect local cortical processing and subserve the binding of different stimulus features. In the visual cortex, neurons within a cortical column respond to optimally oriented stimuli with synchronized oscillatory activity around 40 Hz (Gray & Singer, 1989). Moreover, responses in the gamma range between spatially separated cortical columns are synchronized when the populations have similar orientation preferences, leading to the conclusion that higher-order stimulus properties such as contour are encoded by the temporal coherence of gamma-band activity (Gray et al., 1989). More generally, gamma-band synchronization has been proposed to be a mechanism for the construction of object representations from sensory input and the activation of representations through top-down modulation (Tallon-Baudry et al., 1997; Tallon-Baudry & Bertrand, 1999). A modulation of gamma-band responses by attention has been demonstrated in the auditory (Tiitinen et al., 1993) and visual modalities (Fries et al., 2001), likely subserving an amplification of processing for behaviorally relevant stimuli. Furthermore, gamma-band synchronization correlates with perceptual awareness in paradigms employing binocular rivalry (Fries et al., 1997), ambiguous pictures (Rodriguez et al., 1999), or near-threshold stimuli (Melloni et al., 2007). These findings have led to the hypothesis that gamma synchrony might be a necessary condition for conscious perception, although some studies have demonstrated consciousness in the absence of gamma and vice versa (Koch et al., 2016).

1.3.5 Interactions across the spectrum

In the literature discussed so far, the modulation of oscillatory activity in single frequency bands is often interpreted to mediate specific functions relevant for perception and cognition. However, there is extensive evidence that complex relations of nested activity and phase-power interactions across the spectrum are prevalent in the measured signals.

Different ranges of oscillatory activity can simultaneously be observed in the brain, reflecting the spatial scale of integration, from local gamma-mediated processing to long-range theta-mediated top-down influences (von Stein & Sarnthein, 2000). Similarly, Donner & Siegel (2011) propose a framework of gamma-band activity as a signature of local encoding and beta activity as a marker of integrative function. For top-down modulations to exert their influence on local processing, some mechanism of interaction must exist. In the macaque auditory cortex, a hierarchy of nested interactions has been demonstrated: delta phase modulates theta amplitude and theta phase modulates gamma amplitude (Lakatos, 2005). Taking into account that sound onsets can reset delta oscillations, this hierarchy offers an effective way to align windows of excitability with naturally occurring sounds. The spatial pattern of coupling between theta phase and gamma amplitude, such that gamma oscillations occur at the trough of the theta wave, depends on the behavioral task, thus indicating that transient cross-frequency coupling supports the flexible control of cognitive operations (Canolty et al., 2006). Another proposed function for a theta-gamma code is the ordered representation of multi-item information (Lisman & Jensen, 2013). A phase-power relationship across the layers of striate cortex has also been reported, with alpha phase in deeper layers modulating gamma power in granular and superficial layers (Spaak et al., 2012).

With regards to relationships between concurrent oscillatory activity in distinct bands and the direction of inter-areal information flow, Bastos et al (2015) showed that feedforward influences in macaques' visual areas occur in the theta- and gamma-bands, while feedback communication occurs in the beta-band. Interestingly, top-down influence has been shown to enhance subsequent bottom-up responses in the same bands (Richter et al., 2017). Homologue areas in the human brain exhibit a similar pattern of directed interactions, with dominant feedforward information flow in the gamma-band and feedback in the alpha- and beta-bands (Michalareas et al., 2016). For auditory areas, again a similar interplay has been demonstrated, with low-frequency phase modulating gamma power (Fontolan et al., 2014). This organization may also be related to layer-specific spectral profiles of oscillatory activity reflecting differences in laminar connectivity (Buffalo et al., 2011; van Kerkoerle et al., 2014), and constitute an architecture for predictive coding (Arnal & Giraud, 2012).

1.4 Auditory object perception

The first study in this dissertation employed the auditory continuity illusion, where the perception of an interruption in an amplitude-modulated sound depends on the spectral characteristics of a superimposed mask (Warren et al., 1972). In this paradigm, the auditory system has to perform an analysis of object boundaries to determine the continuity of sounds. The third study employed the ventriloquist illusion, where the location of visual stimuli affects location judgments of concurrently presented sounds (Bruns, 2019; Choe et al., 1975). Therefore, I will briefly summarize relevant literature on auditory perception with a focus on object processing in primary auditory cortex and higher areas.

Sounds can generally be described in terms of pitch, intensity, duration, location and timbre. These features are encoded and analyzed by the auditory system to define and segregate perceptual objects, likely on the basis of constructing sound images with time and frequency dimensions (Griffiths & Warren, 2004). The transformation of auditory signals, characterized by a temporal modulation of air pressure, to a frequency-specific code happens on the level of the cochlea, which is sensitive to high frequencies near the base and low frequencies at the apex. This tonotopic organization is preserved along the peripheral auditory system and can be revealed at the level of primary auditory cortex using magnetoencephalography (MEG) and fMRI (Romani et al., 1982; Talavage et al., 2004). An important sound property that assists in parsing auditory input into separate objects is amplitude modulation, which is also preserved along the auditory pathway and can entrain synchronized activity in primary auditory cortex (Joris et al., 2004). Similarly, short sounds presented at different rates, and longer sounds modulated by amplitude or frequency, elicit a steady state response at different stages of the auditory neuraxis, with the largest responses at lower frequencies and around 40 Hz (Picton et al., 2003). Based on electrophysiological evidence, Näätänen & Winkler (1999) argued that primary auditory cortex is the first analysis stage where complete stimulus representations, which are accessible to conscious perception and top-down modulation, emerge. The planum temporale, located posterior to Heschl's gyrus, is assumed to serve as a computational hub for analyzing complex sounds, segregating the complex acoustic world into spectrotemporal patterns and matching them with learned representations (Griffiths & Warren, 2002). Although differences between auditory and visual object perception exist with regards to the accessibility of overlaid objects, there seems to be an analogous organization of downstream processing into "what" and "where" pathways (Kubovy & Van Valkenburg, 2001). In the auditory system, the spatial stream originates in caudal auditory areas and projects to parietal cortex (Rauschecker & Tian, 2000).

1.5 Principles of audiovisual interactions

While perception has long been investigated under the assumption of modularity (if only for methodological limitations), recent research in the field of multisensory integration has emphasized the occurrence of diverse interactions between different sensory modalities (Shimojo & Shams, 2001). These can take the form of a crossmodal influence from visual information on auditory perception, such as in the ventriloquist illusion. An inverse influence can be observed in the sound-induced flash illusion (SIFI), the paradigm employed in the second study of this thesis, where the number of perceived flashes is affected by concurrently presented sounds (Shams et al., 2000). Another prominent example is the McGurk illusion, where discrepant audiovisual speech can even result in a percept that is qualitatively different from both unisensory signals (McGurk & MacDonald, 1976). I will now outline the literature on multisensory interactions, divided into sections on principles and models, anatomical foundations, and oscillatory mechanisms, although these different aspects are often intertwined in practice.

1.5.1 Principles and models

Different theoretical principles have been proposed to explain the occurrence and characteristics of audiovisual interactions (Welch & Warren, 1980). The "modality precision" hypothesis states that the modality with higher acuity dominates in multisensory perception, explaining visual bias in localization tasks such as the ventriloquist illusion. According to the somewhat similar "modality appropriateness" hypothesis, not the precision, but the suitability of a given sensory modality for perceptual judgment in a given context determines the direction of influence. For instance, the dominance of audition in the temporal rate perception of discrepant audiovisual streams is attributed to its better suitability for such tasks (Welch et al., 1986). Welch & Warren (1980) lastly propose a model of intersensory bias that emphasizes the tendency to perceive in a way that is congruent with a single external cause, mediated by attentional factors, stimulus and task properties. Linked to this idea are approaches that formulate probabilistic models of perception based on cue reliability under the assumption of a single cause. In this framework, multisensory illusions are not considered to be a flawed perceptual judgment, but the result of an optimal joint estimate of noisy multimodal information. When the variance associated with unisensory visual and tactile object size estimates is used to construct a model of maximum-likelihood integration, this model arrives at similar multimodal perceptual decisions as human subjects, including a dominance of the more reliable modality (Ernst & Banks, 2002). Similarly, the ventriloquist effect can be modeled in terms of an optimal weighing of auditory and visual signals according to their localization variance (Alais & Burr, 2004); and perception of the SIFI is also consistent with optimal cue combination, albeit not necessarily assuming a single source (Shams, Ma, et al., 2005). This framework hints at the importance of multisensory processing in integrating complementary, but noisy information about external events to minimize uncertainty. Taking this approach one step further, recent studies have started to elucidate how humans infer the causal structure of sensory events using Bayesian models that take into account different possible stimulus constellations underlying the sensory data (Körding et al., 2007). These models infer whether heteromodal signals have a common cause and accordingly predict if and how signals are combined, with a better fit to human performance than previous models. By combining Bayesian model estimates and fMRI data using multivariate decoding techniques, the representations of different assumptions about the causal structure of sensory signals can be localized in the brain, allowing the identification of processing hierarchies (Rohe & Noppeney, 2015). Similar approaches have been applied to M/EEG data to additionally delineate the temporal dynamics of causal inference (Aller & Noppeney, 2019; Cao et al., 2019; Rohe et al., 2019).

1.5.2 Anatomical foundations

The sketched computations underlying audiovisual interactions pose the question of anatomical substrates. Furthermore, basic principles of integration have also been inferred from the characteristics of associated neural activity patterns. I will now summarize the evidence related to these aspects, which has been gained from anatomical tracing, single cell and neuroimaging studies. Classically, multisensory integration has been conceptualized in terms of hierarchical stages, with a feedforward convergence of unisensory pathways in multimodal areas after initial sensory-specific processing. One prominent example is the superior temporal sulcus, where pathways from primary sensory areas converge (Jones & Powell, 1970), and which is often implicated in the processing of audiovisual stimuli, especially speech (Calvert & Thesen, 2004). This area shows a patchy organization of subregions responding to different modalities, consistent with the organization of homologue areas in nonhuman primates, and possibly allowing joint processing in intermittent patches (Beauchamp, 2005; Beauchamp et al., 2004). Another example are intraparietal regions, which are associated with the integration of multimodal spatial signals (Andersen et al., 1997). Some of these areas have initially been delineated based on conjunction analyses of unimodal activation maps (Lewis et al., 2000), although this approach does not reveal areas exhibiting superadditive responses, which are often regarded as a marker of a true integration process. This property, also termed "multisensory enhancement", can be identified by computing differences in the magnitude of multimodal responses to the sum of the respective unisensory responses (Stanford & Stein, 2007).

Regarding subcortical structures, single cell studies have established responses to heteromodal stimuli already on the level of superior colliculus (Meredith & Stein, 1983). The response properties of neurons observed in these studies were instrumental in establishing principles of multisensory integration like the spatial and temporal rules, which postulate that the congruence of audiovisual stimuli determines the neural response magnitude (King & Palmer, 1985). The importance of temporal coincidence has been underlined in a recent study, where many aspects of human multisensory perception could be replicated by a simple correlation detector model (Parise & Ernst, 2016), and also applies to crossmodal influence, which is constrained by spatial and temporal factors (Sekuler et al., 1997; Shams et al., 2002; Slutsky & Recanzone, 2001). Another important principle first established in superior colliculus neurons is that of inverse effectiveness, which states that bimodal response enhancement is largest when weak unimodal stimuli are used that elicit small responses in isolation (Meredith & Stein, 1986). The validity of this principle has also been corroborated using other methods such as ERPs (Senkowski et al., 2011).

Recent reviews have compiled evidence for integrative processing in primary or early sensory areas as opposed to later convergence (Driver & Noesselt, 2008; Kayser & Logothetis, 2007; Schroeder & Foxe, 2005). Based on the wealth of studies that show sensitivity of ostensibly sensory-specific areas to heteromodal input, Ghazanfar & Schroeder (2006) have called for a reconsideration of the notion that sensory processing ever occurs independently, and formulated the provocative hypothesis that neocortical processing is fundamentally multisensory in nature. They argue that this is the most efficient organization for a unified representation of the external world, given that cognitive development and everyday perception are multisensory. Murray et al. (2016) reviewed the evidence limited to the question if primary visual cortex can be considered multisensory and arrived at a positive conclusion. Supporting this idea on the anatomical level, Falchier et al. (2002) found direct connections from auditory to striate visual areas. Using fMRI, Calvert et al. (1997) demonstrated that watching lip movements is sufficient to activate auditory cortex, and Kayser et al. (2007) showed a modulation of caudal auditory areas by visual input. ERP studies have adduced the early temporal scale of effects to substantiate claims that modality-specific processing is crossmodally influenced (Giard & Peronnet, 1999; Molholm et al., 2002). However, as Cappe et al. (2010) noted, early ERP modulations in response to audiovisual stimuli may more closely reflect distinct source configurations that include higher-order areas, than changes of unisensory response amplitudes. This indicates that some heteromodal responses in primary areas can also be attributed to feedback from multisensory areas.

Notwithstanding the compelling evidence that processing in primary sensory areas can be influenced by input from other modalities, Driver & Noesselt (2008) suggest that functional specialization is still the prevalent cortical organization principle because most areas exhibit a preference for certain modalities or combinations thereof. Furthermore, subregions of primary sensory areas may differ in their connectivity and sensitivity to heteromodal input; and the mechanisms of ostensible multisensory influence are often unspecified. For instance, nonspecific changes in arousal, orienting responses, modulations and driving of heteromodal responses have to be distinguished. In their view, the organization of multisensory processing could plausibly be conceived as one of parallel processing streams interacting already at early stages.

1.5.3 Oscillatory mechanisms

Substantiating the idea that crossmodal influence can affect primary sensory areas, a number of studies have suggested this can be achieved by an oscillatory phase-reset: Lakatos et al. (2007) showed that somatosensory input phase-resets ongoing oscillations in the auditory cortex of macaques, such that concurrent auditory inputs arrive at high-excitability phase and responses are enhanced with inverse effectiveness. The authors ascribed this modulatory influence to thalamic projections. Similarly, a phase reset in auditory cortex by visual stimuli was demonstrated by Kayser et al. (2008). In the inverse direction, Romei et al. (2012) showed a phase modulation of occipital EEG alpha oscillations by auditory inputs, along with periodic excitability changes in visual areas phase-locked to sounds. This finding was subsequently corroborated in intracranial recordings (Mercier et al., 2013).

Recently, the possible roles of neural oscillations in crossmodal interactions more generally have been highlighted. Based on earlier oscillatory binding theories and the changing view of multisensory integration from the primacy of feedforward integration towards a recognition of lateral and feedback interaction, Senkowski et al. (2008) proposed that synchronized activity provides a mechanism for the dynamic formation of crossmodal assemblies. According to van Atteveldt et al. (2014), the diverse findings concerning the principles of integration suggest flexible and dynamic mechanisms, which nevertheless can be explained by a combination of canonical computations. Multisensory processing is here considered to be representative of more general neural mechanisms for information integration. The importance of binding principles and the involvement of networks is task and context-dependent, indicating that multisensory integration is not uniform but adaptive. The general operations proposed by the authors include divisive normalization (Fetsch et al., 2013), which is a model for the interaction of excitatory input in convergence regions. Because this operation is mediated by interneurons, it is likely reflected by gamma-range oscillations. The other operation is phase-reset, as discussed above. Both operations are suggested to be complementary with regards to affected regions, oscillatory frequencies and the temporal predictability of sensory signals. Keil & Senkowski (2018) proposed an integrative framework for the role of oscillations in multisensory perception, underlining that complementary mechanisms are reflected in distinct spectral signatures. For instance, stimulus-driven feedforward integration is characterized by increased gamma-band power, while feedback and top-down modulations are reflected by modulations in lower frequencies.

1.6 Aims of thesis

Based on the notion of perception as an active and selective process (Engel et al., 2001), the research in this thesis aims to further clarify the role of neural oscillations in the integration of sensory signals towards unified perceptual objects, using illusion paradigms as tools of study. In the first study, I used the auditory continuity illusion, a purely auditory paradigm, and focused on low frequency-power and phase modulations in response to stimulus constellations that elicit the sensation of a continuous or interrupted tone. In the second and third studies I used well-established crossmodal illusion paradigms, where auditory signals influence the perception of visual stimuli or vice versa, and focused on prestimulus power modulations. In both audiovisual paradigms, perception varies on a trial-by-trial basis while physical input is constant. Perceptual variability can then be attributed to intrinsic differences of neural activity, allowing inferences on mechanisms of sensory processing. A common theme of the three studies is the binding of auditory or audiovisual objects (Bizley et al., 2016; Griffiths & Warren, 2004), with the related research question how oscillatory dynamics reflect the integration and segregation of sensory information within and across modalities. The research therefore represents an inquiry into distinct, but related problems for sensory systems: how can the separation of information related to an auditory object from that related to the rest of the auditory scene be achieved; or in the audiovisual case: what determines whether auditory and visual information should be integrated and attributed to a single source? On a more abstract level, this question can also be linked to studies operating in the framework of Bayesian inference, which investigated how the brain represents and estimates the underlying causal structure of sensory signals. Following Driver & Noesselt (2008), I focused on multisensory interplay as an example of integration, which can more strictly be defined in terms of a unified percept.

The research addressed different aspects of oscillatory dynamics, from stimulus-induced power and phase modulations, over ongoing prestimulus oscillations that bias the perception of bistable stimuli, to top-down influences. As I have discussed above, different architectures of audiovisual interactions can be derived from the available literature, including lateral projections between presumably unisensory areas, feedforward convergence in higher-order multisensory areas, or feedback from those to unisensory areas. A thorough understanding of their roles in specific multisensory phenomena is currently lacking (Driver & Noesselt, 2008). One aim that can accordingly be derived for this thesis is to investigate how crossmodal influence in different illusions is related to neural activity patterns associated with these possible architectures: do the effects reflect activity in primary sensory areas, multisensory hubs or top-down influences? The main general goal of the empirical work in this dissertation was therefore to investigate which specific spectrotemporal activity patterns are associated with different processing modes assumed to mediate auditory and audiovisual illusions.

2 Summary of experiments

In this chapter, I will summarize the three empirical studies that together constitute this dissertation. In all experiments, EEG was measured using a cap with 128 electrodes. This allows a uniform and dense scalp coverage, which is beneficial for topological inferences and reconstruction of cortical sources. To transform data to the frequency domain, often a convolution of the timeseries with Morlet wavelets is used (Tallon-Baudry et al., 1997). Here, we used a related method based on the application of a single hanning taper or multiple tapers (Mitra & Pesaran, 1999), depending on the frequency range of interest. In the third publication included in this thesis, we also reconstructed the timecourses of neural activity on the level of brain sources before further analysis, using realistic boundary element method headmodels and beamforming (Gramfort et al., 2010; Hillebrand & Barnes, 2005; Van Veen et al., 1997).

2.1 Study 1: The roles of slow oscillatory power and phase alignment in auditory restoration

In this study, we investigated the role of induced oscillatory power and phase alignment in response to composite auditory stimuli that could be perceived as continuous or interrupted. Auditory perception is remarkably resistant to external disturbances, likely to maintain stable representations of auditory objects under difficult acoustic conditions. This can lead to the perceptual restoration of physically interrupted sounds when the interruption is masked by a broadband noise burst (Miller & Licklider, 1950). When this noise mask contains a spectral gap around the frequency of the tracked auditory object, interruptions are readily perceived (Warren et al., 1972).

Previous research on neural correlates of the continuity illusion has identified modulations in primary auditory cortex when simple sounds are used (Petkov et al., 2007; Riecke et al., 2007, 2009), and modulations in higher-order areas when speech signals are used (Heinrich et al., 2008; Riecke et al., 2011; Shahin et al., 2009). For simple amplitude modulated sounds, BOLD responses were weaker for full-spectrum compared to partial masks, and weaker for continuity illusions compared to veridical gap perception given physically identical stimuli (Riecke et al., 2007). Using EEG, an important role of slow oscillations in the continuity illusion could be demonstrated, roughly matching the BOLD pattern: the total power of Theta oscillations shortly after noise onset is increased for interrupted compared to continuous stimuli, for partial compared to full masks, and for continuity illusions compared to veridically perceived interruptions (Riecke et al., 2009). In line with an emerging framework of auditory perception highlighting the role of oscillatory entrainment (Schroeder & Lakatos, 2009), the authors suggested that differences in phase alignment could have contributed to the observed effects, but no corresponding analyses were performed. In this study, we therefore aimed to replicate the previous finding of theta power modulations in the auditory continuity illusion and clarify whether these can be attributed to an alignment of oscillatory phases.

The study consisted of a 2 x 2 factorial design with factors Tone Continuity and Masking Level. We used pure tones, amplitude-modulated at 3 Hz, which could either be continuous or interrupted in the central portion. The noise mask encompassed 2 octaves around the tone frequency, with the 0.6 octave range around the tone frequency notch-filtered for the partial mask, and was overlaid with the central portion of the tone. Participants were asked to rate the continuity of the tone.

The EEG data analysis was focused on induced oscillations, i.e. data spectrally transformed after removing time-locked (or evoked) activity, and inter-trial coherence (ITC, Tallon-Baudry et al., 1996), in the 3-7 Hz range. EEG data were averaged over a mediocentral electrode cluster and the time range from 0.1 to 0.4 s after noise onset in accordance with a previous study (Riecke et al., 2009). We performed a two-factorial statistical analysis on the behavioral ratings, oscillatory power and ITC, respectively.

The behavioral data analysis showed that participants rated the continuity of interrupted tones lower if a partial mask was used, compared to when a full mask was used. In other words, partially unmasking the gap resulted in gap detection, while a full mask resulted in the continuity illusion (Fig. 1a). The analysis of induced power revealed a corresponding effect around 3 Hz: Power was increased for partially compared to fully masked interrupted tones, i.e. for stimuli where a gap was detected compared to continuity illusions. Furthermore, power was increased for continuous compared to interrupted fully masked tones (Fig. 1b). For phase alignment, we found a different pattern: ITC was increased for fully compared to partially masked continuous tones, and for continuous compared to interrupted fully masked tones (Fig. 2c). Thus, 3 Hz power and phase alignment was reduced for the stimulus eliciting a continuity illusion.

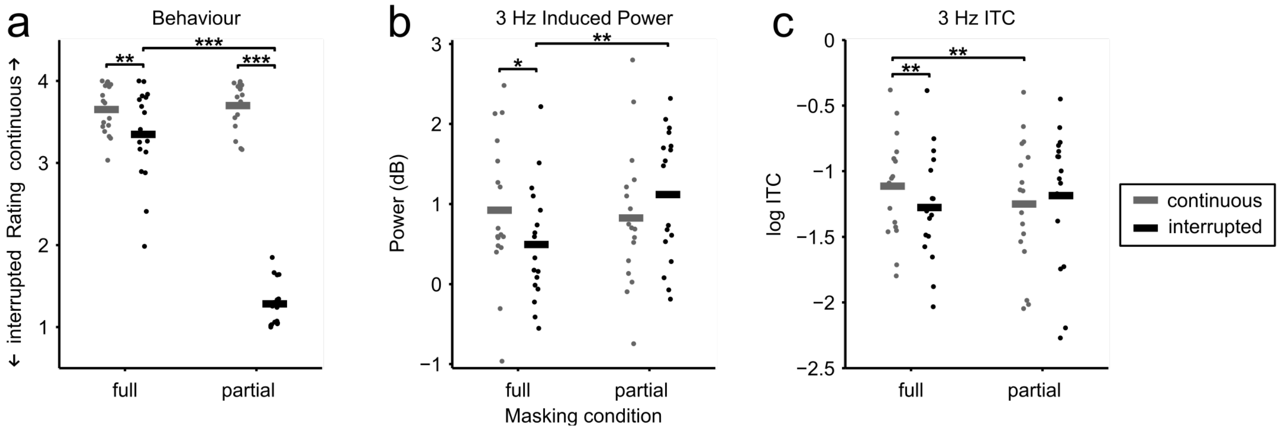

*Figure 1: (a) Mean behavioral ratings, indicated by bars, and individual ratings, indicated by dots. Significant results from follow-up tests are indicated by asterisks (*p < 0.05; **p < 0.01; ***p < 0.001). (b) same as a, but for induced 3 Hz power changes. (c) same as a, but for log-transformed inter-trial coherence at 3 Hz.*

Adapted from "Reduced low-frequency power and phase locking reflect restoration in the auditory continuity illusion", by M. Kaiser et al., 2018, European Journal of Neuroscience, 48, p. 4. Copyright 2018 by Federation of European Neuroscience Societies and John Wiley & Sons Ltd. Adapted with permission.

The pattern of the power effect was broadly consistent with a previous study, which found increased theta power for partially masked, interrupted tones, and decreased theta power for continuity illusions compared to veridical gap detection (Riecke et al., 2009). However, compared to this study, we found relatively higher power in the continuous conditions, which may be explained by differences in analysis strategies and could reflect the encoding of a change in the auditory scene, induced by the onset of the mask. Related to this idea, we found that ITC was selectively increased for fully masked, continuous tones, which we suggested to reflect the registration of the mask as a newly appearing object covering the amplitude-modulated tone. In contrast, when interrupted tones were fully masked, a spectral portion of the mask was attributed to the interrupted tone, which reduces the perceived volume of the mask (McAdams et al., 1998), and the phase locking elicited by it. Our results therefore support the hypothesis that "the restoration of a sound depends on on the suppression of neuronal phase-locking to that sound's acoustic structure" (Riecke et al., 2009, p. 556). However, the hypothesis that the theta increase for gap detection is due to phase locking to the tone modulation frequency, enhancing the salience of the gap, could not be corroborated. Taken together, this study replicated and extended previous work by clarifying the relevance of power and phase modulations in the continuity illusion. We demonstrated that auditory restoration is associated with a reduction of low-frequency power, while gap detection is associated with an increase of power, and that an increase in phase locking may reflect registering an additional object in the auditory scene.

2.2 Study 2: Single-trial modeling of perception in the sound-induced flash illusion

Whereas the first study investigated a purely auditory illusion with a focus on auditory objects, the second study examined crossmodal influence from the auditory to the visual modality. We investigated the trial-by-trial relationship of oscillatory brain activity prior to stimulus presentation and the perception of the sound-induced flash illusion. In this paradigm, participants are presented with one or more bright visual stimuli, concurrently with one or more brief sounds (Shams et al., 2000), and asked to report the number of perceived flashes. The occurrence of the illusion is defined by responses that are biased by auditory information. Importantly, this illusion occurs with a large variability between and within participants, which allows to investigate the neural activity patterns linked with its occurrence. Previous studies have demonstrated that BOLD activity in primary visual cortex reflects the number of subjectively perceived flashes, while activity in the right superior temporal sulcus reflects crossmodal influence regardless of the number of perceived flashes (Watkins et al., 2006, 2007). M/EEG studies have found early modulations of ERPs (Shams et al., 2001; Shams, Iwaki, et al., 2005) and gamma-band power (Bhattacharya et al., 2002; Mishra et al., 2007) related to the illusion. Some studies have also shown a link between oscillatory brain activity prior to stimulus onset and subsequent perception of the illusion: Keil et al. (2014) found increased beta-band power in temporal areas, and Lange et al. (2013) found decreased alpha-band power in occipital areas and increased gamma-band power in occipitotemporal areas. However, no study thus far had incorporated trial-by-trial information in the analysis, which is potentially informative regarding variable perceptual outcomes (Samaha et al., 2017). We therefore aimed to extend previous findings using a new method based on a logistic regression model that predicts the perceptual outcome based on oscillatory power in each trial.

The experimental design comprised different combinations of brief auditory and visual stimuli, only one of which was analyzed here: the combination of two sounds and one flash. On average, participants perceived this combination either as one or two flashes with roughly equal frequencies.

To quantify the relationship between single-trial EEG power in the prestimulus time window and behavioral outcome, we calculated logistic regression weights between the spectral activity in a 500 ms time window prior to stimulus presentation and the binary perceptual rating (i.e. illusion vs. no illusion) for each electrode, frequency and timepoint. The thereby calculated regression weights were used to construct individual dummy distributions by sampling from a normal distribution with the same mean and standard deviation, and subsequently compared against this dummy data using a cluster-based permutation test. Within the identified cluster, a conventional t-test of power values, averaged over trials of different perceptual conditions (illusion vs. no illusion), was used to check the robustness of the modeling result.

We found a significant cluster of regression weights over occipital electrodes between 25 and 41 Hz, from 0.17 s to 0.05 s before stimulus onset, indicating a positive relationship between upper beta-/ lower gamma-band power and subsequent perception of the illusion (Fig. 2). The most robust effects were concentrated in a right occipitotemporal area. The conventional t-test confirmed that power was higher in illusion compared to no-illusion trials.

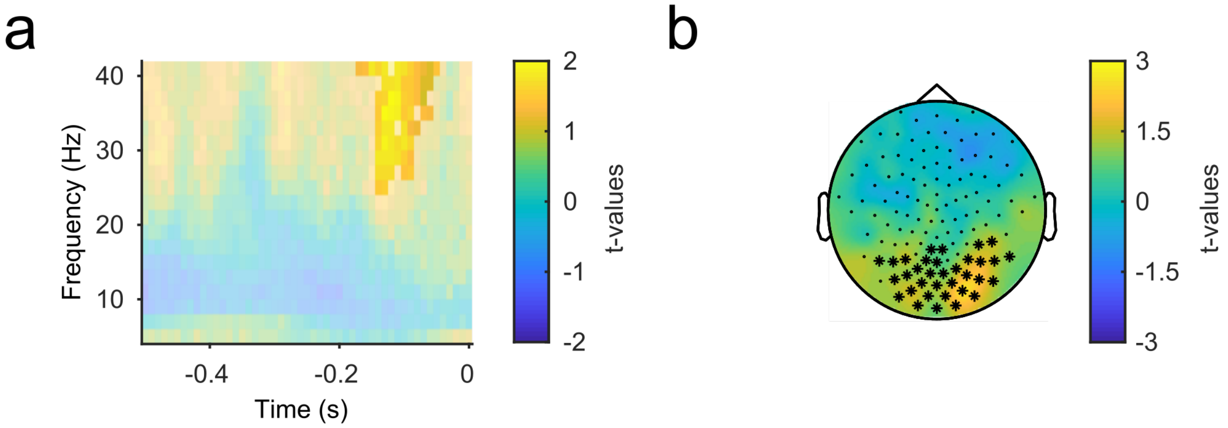

Figure 2: (a) Time frequency spectrum of t-values for the comparison of observed regression weights against dummy data, averaged over significant channels. Significant regions are indicated by saturation. (b) Topography of t-values, averaged over the significant time-frequency region. Significant channels are marked by asterisks.
Adapted from "Single trial prestimulus oscillations predict perception of the sound-induced flash illusion", by M. Kaiser et al., 2019, Scientific Reports, 9, p. 3. Copyright 2019 by the author.

Following Lange et al. (2013), who reported a prestimulus modulation of excitability in occipital and temporal areas entailing illusory flash perception, we suggested that the observed effect has a similar functional role in facilitating crossmodal influence. Another recent EEG study using the SIFI paradigm has found a correlation between oscillatory gamma activity and the prior belief that signals originate from a common source, with a spatial and spectrotemporal extent consistent with our results (Rohe et al., 2019). A study that used brain stimulation to regulate activity in different cortical regions found that the activation level in visual areas has an inverse relationship with illusion rates in the SIFI, while the activation level in temporal areas shows a positive relationship (Bolognini et al., 2011). This indicates that changes in excitability in primary sensory areas induced by stimulation promote the processing of unisensory stimuli, while the effect in our study may reflect enhanced excitability of polysensory pathways. Later modulations of cortical activity that were observed in the SIFI may be primed by the modulation of ongoing oscillations already before stimulus onset.

2.3 Study 3: The ventriloquist illusion: auditory asymmetry, multisensory areas, or frontal influence?

In the third study, we again investigated crossmodal influence, this time from the visual to the auditory modality. We analyzed source-reconstructed neural oscillations to address the question whether perceived sound shifts in the ventriloquist illusion affect the balance of interhemispheric auditory responses, and whether the occurrence of the illusion is related to prestimulus modulations in multisensory or other higher-order areas.

In the ventriloquist illusion, the location of visual stimuli affects the perceived location of concurrently presented sounds (Bruns, 2019; Choe et al., 1975). A previous fMRI study suggested a hierarchical processing account of inferences regarding the sources of spatially disparate audiovisual stimuli, where the posterior intraparietal sulcus represents location under the assumption of a common source and forced fusion of the signals (Rohe & Noppeney, 2015). This brain area has also been implicated in other studies using similar paradigms (Park & Kayser, 2019; Zierul et al., 2017). An EEG study of the ventriloquist illusion showed that multimodal ERP difference waves are larger over the hemisphere contralateral to a perceived peripheral sound shift and localized this effect in the Sylvian fissure (Bonath et al., 2007). The authors suggested that this effect is mediated by connections from visual areas over multimodal areas to auditory cortex. Other EEG studies have provided additional evidence for an early auditory processing account of the ventriloquist illusion (Colin et al., 2002; Stekelenburg et al., 2004). Based on these findings, we aimed to investigate oscillatory prerequisites of crossmodal influence in the ventriloquist illusion with a focus on higher-order areas, as well as the relationship of asymmetrical neural oscillations in auditory areas and sound shifts towards the center or periphery.

To localize space-sensitive auditory areas, we first had participants listen to unilateral 40 Hz amplitude-modulated sounds and reconstructed the sources of cortical responses to the modulation frequency. In the main experiment, we presented combinations of sounds and bright visual stimuli from the left, center, and right, and asked participants to indicate the perceived sound origin. The response patterns indicated that visual stimuli affected the perceived sound location, and that the illusion occurred in a bistable manner in roughly two thirds of the participants. We selected those participants for further analysis. To investigate the relationship of prestimulus oscillations and crossmodal influence, we pooled illusion and no-illusion trials over the different stimulus conditions and compared the resulting time-frequency spectra on the scalp and source level. This analysis showed that activity over mediocentral electrodes around 4 Hz, from 0.5 to 0.12 s before stimulus onset, was decreased in the illusion trials (Fig. 3). On the source level, we found a comparable effect in mediofrontal regions, which was temporally spread to the whole analysis window (Fig. 4). An analysis of a parietal region of interest resembling the area shown to be associated with the fusion of audiovisual signals (Rohe & Noppeney, 2015) resulted in no significant effects.

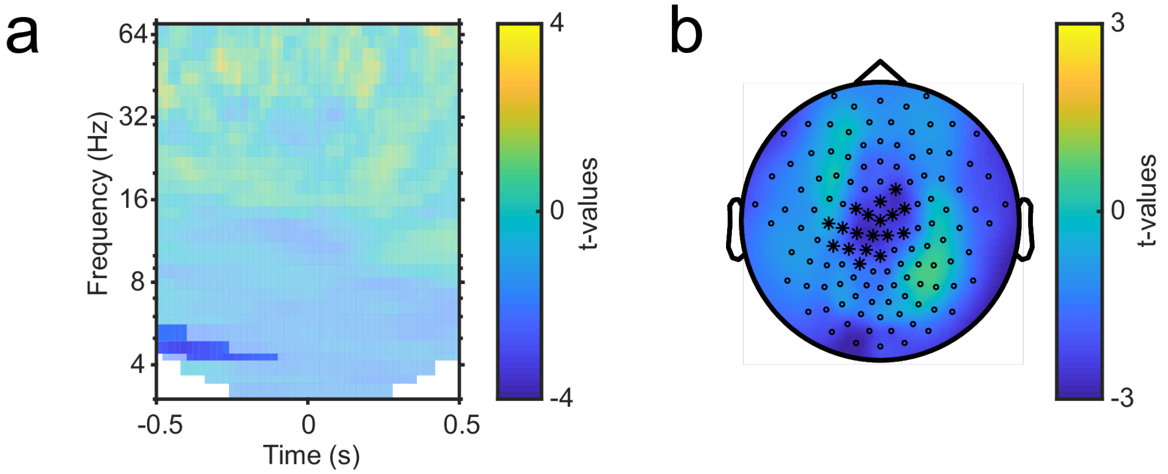

Figure 3: (a) Time-frequency spectrum of t-values for the illusion / no illusion comparison, averaged over significant channels. Significant regions are indicated by saturation. (b)Topography of t-values, averaged over the significant time-frequency region. Significant channels are marked by asterisks.

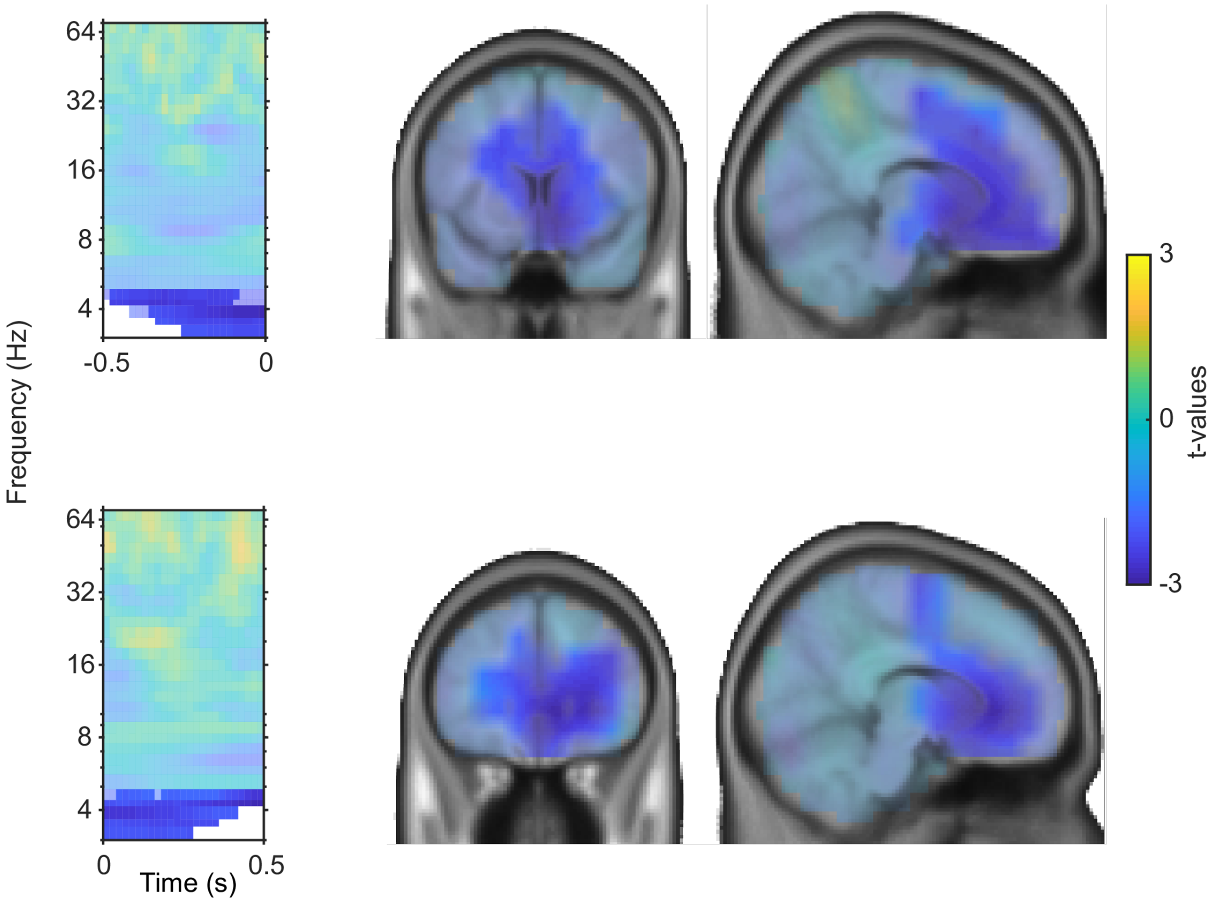

Figure 4: Time-frequency spectrum and sourceplot of t-values for the illusion vs. no illusion comparison on the source level. Significant regions / virtual channels are indicated by saturation. The prestimulus period is shown at the top, the poststimulus period at the bottom.

To investigate the symmetry of auditory responses, we computed the hemispheric laterality index of activity in the previously localized region. We expected this index to show asymmetrical activity when participants correctly perceived a peripheral sound, or when they illusorily perceived a central sound shifted to the periphery, whereas symmetrical activity was expected when central sounds were localized correctly or when a peripheral sound was shifted to the center. However, no corresponding effects were revealed in the analysis, and a simplified analysis roughly equivalent to the previous study by Bonath et al. (2007) could not corroborate the finding that peripheral sound shifts result in ERP asymmetries.

We suggested that the decreased mediofrontal theta-band power before stimulus onset reflects a state of diminished cognitive control over the appraisal of conflicting multisensory information (Cavanagh & Frank, 2014), which leads to a reliance on more salient visual information. Other studies have also shown that the ventriloquist illusion is susceptible to top-down influence (Bruns et al., 2014; Maiworm et al., 2012). Therefore, our results are consistent with the notion that crossmodal influence in the ventriloquist illusion depends primarily on frontal control, rather than on modulations of activity in primary sensory or multisensory areas.

3 General discussion

In the three studies performed as parts of this dissertation, I investigated oscillatory signatures of auditory object processing and crossmodal influence in different directions, using illusion paradigms that elicit variable perceptual judgments. The empirical research was motivated by an aspiration to better understand the neural activity patterns related to sensory binding within and across the senses. The results obtained in the individual studies highlight the link between activity in different spectral bands and specific sensory processing modes, collectively tying into a functional framework regarding the spectrotemporal characteristics of the human EEG. The main contributions of the individual research articles are: a clarification of the role of low frequency power modulations and phase alignment for object formation in the auditory continuity illusion (study 1), a prediction of perception in the sound-induced flash illusion based on increased upper beta-/ lower gamma-band activity, presumably reflecting excitability in visual and multisensory areas (study 2), and a demonstration of frontal influence in the theta-band on perception in the ventriloquist illusion (study 3).

3.1 Different oscillatory signatures in different phenomena of sensory binding

In study 1, the rationale of analysis was to reveal modulations of slow oscillatory power and phase, which were driven by differences of composite auditory stimuli. This represents an investigation into externally driven brain responses, rather than differences in perception resulting from internally generated, variable neural activity. The constructive aspect of the employed paradigm lies in the perceptual restoration of an interrupted auditory object. However, Riecke et al. (2009) showed that the theta power decrease in response to the combination of an interrupted tone and a full mask, i.e. the composite sound eliciting a continuity illusion, can also be observed when comparing trials that were perceived as continuous to those perceived as interrupted, given the same physical stimuli. This behavior-based analysis indicates that the suppression of induced oscillations observed in our study reflects an interpolation of the sound that is variable over trials, determining perceived continuity. The behavioral pattern and the power modulation we observed broadly replicate the results from the earlier study, which can be considered important given the recently declared replication crisis in psychology and neuroscience (Button et al., 2013; Open Science Collaboration, 2015). As the primary contribution of our study, we extended the previous finding by analyzing phase alignment, which had been suggested to underlie the total power modulation by Riecke et al. (2009). Our results corroborated the idea that restoration depends on reduced phase-locking to acoustic structure, but not the suggestion that increased phase-locking supports the detection of gaps. Instead, ITC was largest when a full mask overlaid the continuous sound, whose modulation frequency was likely tracked by auditory cortex (Liégeois-Chauvel et al., 2004; Picton et al., 1987). This may reflect a phase reset that amplifies afferent activity caused by the mask, and thereby supports its registration as a separate object. Conversely, the reduction of slow frequency power and phase locking in the illusion condition may reflect a diminished processing of mask-evoked activity, leading to blurred boundaries and the perception of a constant auditory object. These results underline the importance of slow oscillations in auditory cortex as a general mechanism for tracking amplitude modulations and thereby segmenting the auditory scene (Giraud & Poeppel, 2012; Luo & Poeppel, 2007). Interpreting these results in the context of sensory binding, it appears that a diminished encoding of constituent parts of the auditory scene results in their coalescence. An alternative account of the role of low-frequency modulations in the framework of predictive auditory coding is presented by a recent MEG study (Recasens et al., 2018). Here, mismatching stimuli were associated with increased theta-band power and phase locking, as well as feedforward influences from auditory to frontal areas in the theta-band, reflecting the generation of prediction errors. Similarly in our study, the perception of a gap, which occurred roughly in a quarter of trials only, could elicit a prediction error, while perceptual restoration induced by spectral filling reduces it. Another MEG study has investigated the interaction of low-frequency entrainment in auditory areas with oscillatory activity that could support predictive top-down modulations, and found that entrainment in the left superior temporal gyrus was modulated by beta power in frontal areas (Keitel et al., 2017). Taken together, the results of this study clarified the contributions of low-frequency power and phase modulations in auditory object formation.

In study 2, we turned our attention to the modulation of single-trial oscillatory power as a predictor of subjective perception in the sound-induced flash illusion. We found that increased upper beta-/ lower gamma-band power shortly before stimulus onset facilitated crossmodal influence, as exemplified by the perception of an additional flash induced by a concurrently presented supernumerous sound. Here, the binding aspect consists in a perceptual combination of auditory inducer signals and visual stimuli in illusion trials, such that both are perceived as originating from the same source (Shams, Ma, et al., 2005). Interestingly, a recent EEG study

(Rohe et al., 2019) has found high correlation values with the prior assumption that visual and auditory signals in the SIFI are due to a single cause in a spectrotemporal region consistent with our modeling result. The prestimulus power modulation we observed might therefore reflect a tendency to bind audiovisual information. This tendency is possibly mediated by changes of excitability or functional connectivity, leading to a multimodal processing advantage between primary sensory and multisensory areas such as the superior temporal sulcus, which have previously been shown to be associated with the illusion (Keil et al., 2012; Lange et al., 2013; Watkins et al., 2006, 2007). The topography and spectral extent of the observed effect in the beta- and gamma-bands implies that sensory pathways are affected directly, as opposed to the more distal influence from frontal areas observed in study 3. An important contribution of this study was that we successfully applied a new method for the analysis of single-trial relationships between neural activity and behavior, consisting of a logistic regression of a binary perceptual outcome on oscillatory power, which allowed us to explicitly analyze how modulations of time- and frequency-resolved activity affect variability of crossmodal binding across trials. This analysis revealed that perception of the illusion can be predicted by power modulations in the upper beta-/ lower gamma-band.

In study 3, we again investigated the role of oscillations in a classical crossmodal illusion paradigm, where auditory localization is affected by concurrently presented visual stimuli at divergent positions. Similarly to study 2, sensory binding here refers to the attribution of auditory signals to a visual stimulus in illusion trials. This view of the ventriloquist illusion is supported by a correlation between crossmodal localization bias and reports of perceptual unity (Wallace et al., 2004). We aimed to corroborate a previous report of a link between subjectively perceived sound location and the symmetry of responses in auditory cortex (Bonath et al., 2007), as well as an association between processing in higher-order multisensory areas and the representation of a forced-fusion estimate of audiovisual signals (Rohe & Noppeney, 2015). However, consistent with behavioral studies showing top-down modulation in the ventriloquist illusion (Bruns, 2019), we found evidence for a susceptibility of the illusion to frontal influence in the theta-band. We suggested that the reduced power we found for illusion trials reflects diminished cognitive control, which facilitates crossmodal influence from salient visual signals. A recent MEG study combining modeling of behavior with multivariate decoding found that frontal areas arbitrate between different representations of multimodal information, and are therefore a likely candidate region for the regulation of relevant sensory circuits (Cao et al., 2019). Likewise, frontal regions have been implicated in the reliability-dependent processing of incongruent audiovisual signals (Noppeney et al., 2010), and the modulation of audiovisual integration via modality-specific attention (Rohe & Noppeney, 2018). The results of study 3 underline that frontal areas influence sensory processing and perception via oscillations in the theta band.

To summarize, the three studies provided evidence for frequency-specific modulations of neural oscillations in different instances of sensory binding: in study 1, we found a reduction of 3 Hz power and phase alignment in response to the perceptual restoration of an interrupted tone, underlining the importance of slow oscillations in auditory cortex for segmenting the auditory scene. In study 2, we found that upper beta-/ low-gamma band oscillation over occipitotemporal regions predicted the subsequent perception of the sound-induced flash illusion, in line with a facilitation of multisensory processing in visual and multimodal circuits. In study 3, we established an association between reduced theta-band power in mediofrontal areas and the occurrence of the ventriloquist illusion, which points to a more effortful mode of operation possibly involving a resolution of response conflict due to incongruent, but salient visual information, or a monitoring of the demanding auditory localization task. The integrated account of the three experiments represents a progression of processing modes, which are reflected in different spatiotemporal activity patterns, from stimulus-related modulations of activity in primary sensory areas (study 1), over prestimulus differences of excitability or connectivity in multisensory circuits (study 2), to frontal influences on sensory decision-making (study 3). A graphical synopsis of these results is presented in Figure 5.

The combined evidence from studies 2 and 3 allowed inferences on the architectures of multisensory integration differentially involved in these phenomena, which was outlined as an important issue in the field by Driver & Noesselt (2008): while the sound-induced flash illusion depends more on activity along sensory pathways, the ventriloquist illusion is affected by top-down influence from frontal areas. In turn, this leads to the question why no common signature of multisensory integration was found in these studies. Likely reasons for the lack of comparable effects include differences in task demands and sensory processing in the involved modalities. For instance, the temporal numerosity judgment task in the SIFI might preferentially recruit circuits that are sensitive to the rate of visual stimuli (Fox & Raichle, 1984), while the spatial localization task in the ventriloquist illusion depends more on higher order areas (Bushara et al., 1999). Similarly, visual but not auditory processing depends on cycles of increased excitability (VanRullen et al., 2014). Consequently, a correlation of perception in the

SIFI with the individual alpha frequency has been described and replicated (Cecere et al., 2014; Keil & Senkowski, 2017), but no similar relationship has yet been discovered in the ventriloquist illusion.

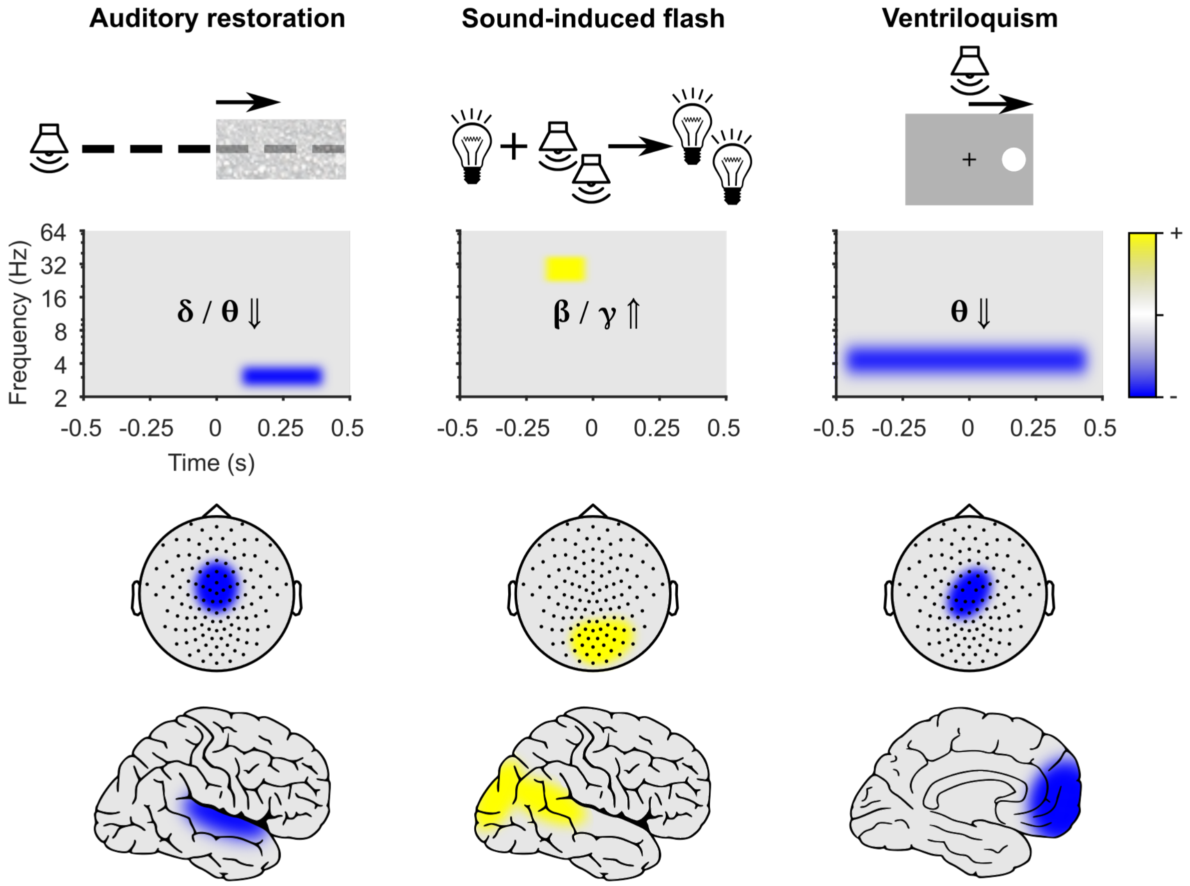

Figure 5: An integrated summary of the three studies. Top row: schematic depiction of the paradigm. Second and third row: illustration of the spectrotemporal and spatial extent of the modulation. Time point zero refers to the onset of the gap/mask in auditory restoration, and to stimulus onset in the other studies. Bottom row: assumed neural sources. Please note that we only performed source analysis in the study on ventriloquism. Sources for the other studies are derived from the literature (Keil et al., 2014; Lange et al., 2013; Riecke et al., 2007, 2009; Watkins et al., 2006, 2007).

To the extent that we focused on prestimulus modulations in studies 2 and 3, the experimental approach stands orthogonal to studies that investigate the sampling and processing of multisensory information. Ernst & Bulthoff (2004) distinguish subsequent stages of sensory combination, where available information is maximized, and integration, where sensory uncertainty is reduced. This idea of hierarchical processing is supported by studies showing a succession of estimation stages in primary sensory and higher-order areas (Aller & Noppeney, 2019; Cao et al., 2019; Rohe & Noppeney, 2015). The research in this thesis focused on oscillatory signatures that reflect the tendency of sensory systems to bind information, starting already before stimulus presentation. Accordingly, it aligns with several studies demonstrating an influence of prestimulus activity on perception (Hanslmayr et al., 2007; Iemi et al., 2017; Linkenkaer-Hansen et al., 2004; Weisz et al., 2014), especially in multisensory contexts (Keil et al., 2012, 2014; Lange et al., 2013; Rohe et al., 2019). The role of prestimulus oscillations for subsequent binding can be likened to that of prerequisite processes carrying no sensory content for conscious experience (Aru et al., 2012; Ruhnau et al., 2014), in that both may facilitate specific consequences regarding actual sensory processing.

3.2 Limitations

In addition to specific limitations outlined in the research articles, there are some caveats on the level of integrating the findings in a more general framework regarding the relation between oscillations and perception. Recently, Keil & Senkowski (2018) developed a framework for the role of oscillations in multisensory perception, which underlined the multifaceted and complementary mechanisms at play. The present thesis adds to this body of research by providing evidence for specific modulations in different binding phenomena, falling into place in different sections of an existing reference frame of oscillatory activity. However, an integration of these findings into a unified framework seems difficult, not least because mechanisms might differ according to sensory modality and task context, as indicated above. While the field has not yet reached a consensus on relevant mechanisms, van Atteveldt et al. 2017 have outlined two general operations possibly underlying many phenomena. The first is crossmodal phase reset, which we did not investigate in studies 2 and 3. Especially in the context of the ventriloquist illusion, an analysis of phase or phase-amplitude-coupling may prove fruitful, because the theta-band is often implicated in modulations of these measures (Schroeder & Lakatos, 2009). Similarly, although theta-band oscillations are often associated with long-range connectivity (von Stein & Sarnthein, 2000), and despite implicit assumptions about the relationship between neural oscillations and connectivity especially in the context of multisensory integration (Senkowski et al., 2008), we did not perform analyses of functional network measures. Such an analysis would allow stronger inferences on the mode of influence that e.g. the observed mediofrontal theta-band modulation exerts on perception. Assuming this modulation reflects an adaptation to conflict or diminished control, does the mechanism then consist in an enhanced processing of auditory information or the inhibition of visual information? These accounts make different predictions about the targets of top-down modulation, that could be tested explicitly using measures of functional or effective connectivity. Furthermore, the lack of source-level analyses in studies 1 and 2 hampers strong conclusions on affected regions. For study 1, a localization in auditory cortex can reasonably be derived from previous research (Riecke et al., 2009). Despite the similarity of the modulated frequency band, the different sources and temporal extent of effects in studies 1 and 3 illusion suggest different functional roles. On the methodological level, EEG is limited in its ability to resolve cortical interactions at a spatial scale that is relevant for multisensory processing, as exemplified by the patchy organization of superior temporal sulcus (Beauchamp et al., 2004) or differences in multimodal connectivity within primary sensory areas (Driver & Noesselt, 2008). This limitation seems especially relevant in the interpretation of the SIFI results, where an exact localization of the effect to visual or superior temporal areas was not realized. However, the results of Lange et al. (2013) indicate that a network consisting of occipital as well as temporal areas is implicated in prestimulus gamma modulations predicting crossmodally-induced flash perception. More generally, the correlative nature of our findings impedes specifying the links between the measured signal, neural mechanisms and subjective perception. Working towards that goal, an explicit account of the relationship between properties of neural circuits, timescales of rhythmic activity and the computations implemented by the circuits was recently developed by Womelsdorf et al. (2014).

On the conceptual level, the coherence of the studies included in this thesis derives from a common focus on the relation of neural oscillations and the perceptual integration of signals in terms of a common underlying object, or their segregation into separate objects. In the case of the continuity illusion, this refers to the constituent parts of the tones and the mask, in the case of the SIFI, this refers to the number of flashes and beeps, and in the case of ventriloquism, this refers to localized auditory and visual signals. However, no explicit analyses concerning the way the brain estimates and represents the causal structure of the environment were conducted in this thesis. Furthermore, it is arguable whether illusions should be used as a framework for successful integration. Following van Atteveldt (2017), crossmodal interactions can be taken as representative of general integrative operations in the brain, providing a unique window into these processes. This amounts to an extended binding concept compared to the one employed in studies concerned with the binding of stimulus properties that are processed in different functionally specialized brain areas. Nevertheless, conceived from the result of the binding process, the perceptual restoration induced by a fusion of mask and interrupted sounds in the continuity illusion, or crossmodal influence that leads to unified percepts, can be regarded as examples that are related to the same problem, requiring specific explanations.

3.3 Outlook

In addition to the explicit analyses of connectivity outlined above, aiming at a mechanistic account of the relation between theta-band activity and the illusion, some venues for future research can be derived from the present thesis. Building on the ideas of Engel and colleagues (2001), it holds promise to explicitly relate spectrotemporal activity patterns with specific predictions about sensory input, to arrive at a better idea how active and selective processing affects perception. For instance, one could manipulate the rhythmicity of constituent signals as a proxy for temporal predictability, and analyze corresponding changes in behavior and oscillatory signatures. Similarly, the context dependency and automaticity of the phenomena could be investigated by manipulating cognitive factors such as attentional resources. Such work was recently initiated by Michail & Keil (2018). Finally, moving towards a causal account of oscillatory signatures in perception, rhythmic brain stimulation methods should be used to investigate the impact of elevated oscillatory activity on perceptual binding.

3.4 Conclusion

The research in this dissertation investigated different aspects of oscillatory activity and their relation to specific phenomena of sensory binding within and across modalities. Auditory restoration is reflected by reduced power and phase locking of slow entrained oscillations, indicating a diminished segregation of constituent parts of the auditory scene. Crossmodal influence from auditory information on visual numerosity judgments can be predicted by occipitotemporal upper beta-/lower gamma-band power, indicating changes of excitability between visual and multisensory areas. Finally, crossmodal influence from visual information on auditory localization is associated with reduced frontal theta-band power, indicating a top-down influence on sensory processing. Taken together, the empirical findings provide further evidence for frequency-specific modulations of neural activity, depending on the processing mode required in auditory and audiovisual perceptual illusions.

References

Adrian, E. D., & Matthews, B. H. C. (1934). The Berger rhythm: potential changes from the occipital lobes in man. *Brain*, *57*, 355–385. https://doi.org/10.1093/brain/57.4.355

Alais, D., & Burr, D. (2004). The Ventriloquist Effect Results from Near-Optimal Bimodal Integration. *Current Biology*, *14*(3), 257–262. https://doi.org/10.1016/j.cub.2004.01.029

Aller, M., & Noppeney, U. (2019). To integrate or not to integrate: Temporal dynamics of hierarchical Bayesian causal inference. *PLOS Biology*, *17*(4), e3000210. https://doi.org/10.1371/journal.pbio.3000210

Andersen, R. A., Snyder, L. H., Bradley, D. C., & Xing, J. (1997). Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annual Review of Neuroscience*, *20*, 303–330. https://doi.org/10.1146/annurev.neuro.20.1.303

Arieli, A., Sterkin, A., Grinvald, A., & Aertsen, A. (1996). Dynamics of Ongoing Activity: Explanation of the Large Variability in Evoked Cortical Responses. *Science*, *273*(5283), 1868–1871. https://doi.org/10.1126/science.273.5283.1868

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

Aru, J., Bachmann, T., Singer, W., & Melloni, L. (2012). Distilling the neural correlates of consciousness. *Neuroscience & Biobehavioral Reviews*, *36*(2), 737–746. https://doi.org/10.1016/j.neubiorev.2011.12.003

Axmacher, N., Mormann, F., Fernández, G., Elger, C. E., & Fell, J. (2006). Memory formation by neuronal synchronization. *Brain Research Reviews*, *52*(1), 170–182. https://doi.org/10.1016/j.brainresrev.2006.01.007

Bartos, M., Vida, I., & Jonas, P. (2007). Synaptic mechanisms of synchronized gamma oscillations in inhibitory interneuron networks. *Nature Reviews Neuroscience*, *8*(1), 45–56. https://doi.org/10.1038/nrn2044

Bastos, A. M., Vezoli, J., Bosman, C. A., Schoffelen, J.-M., Oostenveld, R., Dowdall, J. R., De Weerd, P., Kennedy, H., & Fries, P. (2015). Visual Areas Exert Feedforward and Feedback Influences through Distinct Frequency Channels. *Neuron*, *85*(2), 390–401. https://doi.org/10.1016/j.neuron.2014.12.018

Beauchamp, M. S. (2005). See me, hear me, touch me: multisensory integration in lateral occipital-temporal cortex. *Current Opinion in Neurobiology*, *15*(2), 145–153. https://doi.org/10.1016/j.conb.2005.03.011

Beauchamp, M. S., Argall, B. D., Bodurka, J., Duyn, J. H., & Martin, A. (2004). Unraveling multisensory integration: patchy organization within human STS multisensory cortex. *Nature Neuroscience*, *7*(11), 1190–1192. https://doi.org/10.1038/nn1333

Berger, H. (1929). Über das Elektrenkephalogramm des Menschen. *Archiv für Psychiatrie und Nervenkrankheiten*, *87*(1), 527–570. https://doi.org/10.1007/BF01797193

Bhattacharya, J., Shams, L., & Shimojo, S. (2002). Sound-induced illusory flash perception: role of gamma band responses. *Neuroreport*, *13*(14), 1727–1730.

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

Bolognini, N., Rossetti, A., Casati, C., Mancini, F., & Vallar, G. (2011). Neuromodulation of multisensory perception: A tDCS study of the sound-induced flash illusion. *Neuropsychologia*, *49*(2), 231–237. https://doi.org/10.1016/j.neuropsychologia.2010.11.015

Bonath, B., Noesselt, T., Martinez, A., Mishra, J., Schwiecker, K., Heinze, H.-J., & Hillyard, S. A. (2007). Neural Basis of the Ventriloquist Illusion. *Current Biology*, *17*(19), 1697–1703. https://doi.org/10.1016/j.cub.2007.08.050

Brookes, M. J., Woolrich, M., Luckhoo, H., Price, D., Hale, J. R., Stephenson, M. C., Barnes, G. R., Smith, S. M., & Morris, P. G. (2011). Investigating the electrophysiological basis of resting state networks using magnetoencephalography. *Proceedings of the National Academy of Sciences*, *108*(40), 16783–16788. https://doi.org/10.1073/pnas.1112685108

Brovelli, A., Ding, M., Ledberg, A., Chen, Y., Nakamura, R., & Bressler, S. L. (2004). Beta oscillations in a large-scale sensorimotor cortical network: directional influences revealed by Granger causality. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(26), 9849–9854. https://doi.org/10.1073/pnas.0308538101

Bruns, P. (2019). The Ventriloquist Illusion as a Tool to Study Multisensory Processing: An Update. *Frontiers in Integrative Neuroscience*, *13*. https://doi.org/10.3389/fnint.2019.00051

Bruns, P., Maiworm, M., & Röder, B. (2014). Reward expectation influences audiovisual spatial integration. *Attention, Perception, & Psychophysics*, *76*(6), 1815–1827. https://doi.org/10.3758/s13414-014-0699-y

Buffalo, E. A., Fries, P., Landman, R., Buschman, T. J., & Desimone, R. (2011). Laminar differences in gamma and alpha coherence in the ventral stream. *Proceedings of the National Academy of Sciences*, *108*(27), 11262–11267. https://doi.org/10.1073/pnas.1011284108

Busch, N. A., Dubois, J., & VanRullen, R. (2009). The Phase of Ongoing EEG Oscillations Predicts Visual Perception. *Journal of Neuroscience*, *29*(24), 7869–7876. https://doi.org/10.1523/JNEUROSCI.0113-09.2009

Bushara, K. O., Weeks, R. A., Ishii, K., Catalan, M.-J., Tian, B., Rauschecker, J. P., & Hallett, M. (1999). Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. *Nature Neuroscience*, *2*(8), 759–766. https://doi.org/10.1038/11239

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), 365–376. https://doi.org/10.1038/nrn3475

Buzsáki, G. (2006). *Rhythms of the Brain.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195301069.001.0001

Buzsáki, G., Anastassiou, C. A., & Koch, C. (2012). The origin of extracellular fields and currents — EEG, ECoG, LFP and spikes. *Nature Reviews Neuroscience*, *13*(6), 407–420. https://doi.org/10.1038/nrn3241

Buzsáki, G., & Draguhn, A. (2004). Neuronal Oscillations in Cortical Networks. *Science*, *304*(5679), 1926–1929. https://doi.org/10.1126/science.1099745

Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C. R., McGuire, P. K., Woodruff, P. W. R., Iversen, S. D., & David, A. S. (1997). Activation of Auditory Cortex During Silent Lipreading. *Science*, *276*(5312), 593–596. https://doi.org/10.1126/science.276.5312.593

Calvert, G. A., & Thesen, T. (2004). Multisensory integration: methodological approaches and emerging principles in the human brain. *Journal of Physiology-Paris*, *98*(1–3), 191–205. https://doi.org/10.1016/j.jphysparis.2004.03.018

Canolty, R. T., Edwards, E., Dalal, S. S., Soltani, M., Nagarajan, S. S., Kirsch, H. E., Berger, M. S., Barbaro, N. M., & Knight, R. T. (2006). High Gamma Power Is Phase-Locked to Theta Oscillations in Human Neocortex. *Science*, *313*(5793), 1626–1628. https://doi.org/10.1126/science.1128115

Cao, Y., Summerfield, C., Park, H., Giordano, B. L., & Kayser, C. (2019). Causal Inference in the Multisensory Brain. *Neuron*, *102*(5), 1076-1087.e8. https://doi.org/10.1016/j.neuron.2019.03.043

Cappe, C., Thut, G., Romei, V., & Murray, M. M. (2010). Auditory–Visual Multisensory Interactions in Humans: Timing, Topography, Directionality, and Sources. *Journal of Neuroscience*, *30*(38), 12572–12580. https://doi.org/10.1523/JNEUROSCI.1099-10.2010

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

Cavanagh, J. F., Zambrano-Vazquez, L., & Allen, J. J. B. (2012). Theta lingua franca: A common mid-frontal substrate for action monitoring processes. *Psychophysiology*, *49*(2), 220–238. https://doi.org/10.1111/j.1469-8986.2011.01293.x

Cecere, R., Rees, G., & Romei, V. (2014). Individual Differences in Alpha Frequency Drive Crossmodal Illusory Perception. *Current Biology*. https://doi.org/10.1016/j.cub.2014.11.034

Choe, C. S., Welch, R. B., Gilford, R. M., & Juola, J. F. (1975). The “ventriloquist effect”: Visual dominance or response bias? *Perception & Psychophysics*, *18*(1), 55–60.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(03), 181–204. https://doi.org/10.1017/S0140525X12000477

Colin, C., Radeau, M., Soquet, A., Dachy, B., & Deltenre, P. (2002). Electrophysiology of spatial scene analysis: the mismatch negativity (MMN) is sensitive to the ventriloquism illusion. *Clinical Neurophysiology*, *113*(4), 507–518. https://doi.org/10.1016/S1388-2457(02)00028-7

DeYoe, E. A., & Van Essen, D. C. (1988). Concurrent processing streams in monkey visual cortex. *Trends in Neurosciences*, *11*(5), 219–226. https://doi.org/10.1016/0166-2236(88)90130-0

Doelling, K. B., Arnal, L. H., Ghitza, O., & Poeppel, D. (2014). Acoustic landmarks drive delta–theta oscillations to enable speech comprehension by facilitating perceptual parsing. *NeuroImage*, *85*, 761–768. https://doi.org/10.1016/j.neuroimage.2013.06.035

Donner, T. H., & Siegel, M. (2011). A framework for local cortical oscillation patterns. *Trends in Cognitive Sciences*, *15*(5), 191–199. https://doi.org/10.1016/j.tics.2011.03.007

Draguhn, A., Traub, R. D., Schmitz, D., & Jefferys, J. G. R. (1998). Electrical coupling underlies high-frequency oscillations in the hippocampus in vitro. *Nature*, *394*(6689), 189–192. https://doi.org/10.1038/28184

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., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–433. https://doi.org/10.1038/415429a

Ernst, M. O., & Bülthoff, 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

Falchier, A., Clavagnier, S., Barone, P., & Kennedy, H. (2002). Anatomical Evidence of Multimodal Integration in Primate Striate Cortex. *Journal of Neuroscience*, *22*(13), 5749–5759. https://doi.org/10.1523/JNEUROSCI.22-13-05749.2002

Fetsch, C. R., DeAngelis, G. C., & Angelaki, D. E. (2013). Bridging the gap between theories of sensory cue integration and the physiology of multisensory neurons. *Nature Reviews Neuroscience*, *14*(6), 429–442. https://doi.org/10.1038/nrn3503

Fontolan, L., Morillon, B., Liegeois-Chauvel, C., & Giraud, A.-L. (2014). The contribution of frequency-specific activity to hierarchical information processing in the human auditory cortex. *Nature Communications*, *5*(1), 1–10. https://doi.org/10.1038/ncomms5694

Fox, P. T., & Raichle, M. E. (1984). Stimulus rate dependence of regional cerebral blood flow in human striate cortex, demonstrated by positron emission tomography. *Journal of Neurophysiology*, *51*(5), 1109–1120. https://doi.org/10.1152/jn.1984.51.5.1109

Foxe, J., Simpson, G., & Ahlfors, S. (1998). Parieto-occipital ~10 Hz activity reflects anticipatory state of visual attention mechanisms. *Neuroreport*, *9*(17), 3929–3933.

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

Fries, P., Reynolds, J. H., Rorie, A. E., & Desimone, R. (2001). Modulation of Oscillatory Neuronal Synchronization by Selective Visual Attention. *Science*, *291*(5508), 1560–1563. https://doi.org/10.1126/science.1055465

Fries, P., Roelfsema, P. R., Engel, A. K., König, P., & Singer, W. (1997). Synchronization of oscillatory responses in visual cortex correlates with perception in interocular rivalry. *Proceedings of the National Academy of Sciences*, *94*(23), 12699–12704. https://doi.org/10.1073/pnas.94.23.12699

Ghazanfar, A. A., & Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences*, *10*(6), 278–285. https://doi.org/10.1016/j.tics.2006.04.008

Giard, M. H., & Peronnet, F. (1999). Auditory-Visual Integration during Multimodal Object Recognition in Humans: A Behavioral and Electrophysiological Study. *Journal of Cognitive Neuroscience*, *11*(5), 473–490. https://doi.org/10.1162/089892999563544

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience*, *15*(4), 511–517. https://doi.org/10.1038/nn.3063

Gramfort, A., Papadopoulo, T., Olivi, E., & Clerc, M. (2010). OpenMEEG: opensource software for quasistatic bioelectromagnetics. *BioMedical Engineering OnLine*, *9*(1), 45. https://doi.org/10.1186/1475-925X-9-45

Gray, C. M., König, P., Engel, A. K., & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, *338*(6213), 334. https://doi.org/10.1038/338334a0

Gray, C. M., & Singer, W. (1989). Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proceedings of the National Academy of Sciences*, *86*(5), 1698–1702. https://doi.org/10.1073/pnas.86.5.1698

Griffiths, T. D., & Warren, J. D. (2002). The planum temporale as a computational hub. *Trends in Neurosciences*, *25*(7), 348–353. https://doi.org/10.1016/S0166-2236(02)02191-4

Griffiths, T. D., & Warren, J. D. (2004). What is an auditory object? *Nature Reviews Neuroscience*, *5*(11), 887–892. https://doi.org/10.1038/nrn1538

Gross, J., Schmitz, F., Schnitzler, I., Kessler, K., Shapiro, K., Hommel, B., & Schnitzler, A. (2004). Modulation of long-range neural synchrony reflects temporal limitations of visual attention in humans. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(35), 13050–13055. https://doi.org/10.1073/pnas.0404944101

Hanslmayr, S., Aslan, A., Staudigl, T., Klimesch, W., Herrmann, C. S., & Bäuml, K.-H. (2007). Prestimulus oscillations predict visual perception performance between and within subjects. *NeuroImage*, *37*(4), 1465–1473. https://doi.org/10.1016/j.neuroimage.2007.07.011

Hanslmayr, S., Pastötter, B., Bäuml, K.-H., Gruber, S., Wimber, M., & Klimesch, W. (2008). The Electrophysiological Dynamics of Interference during the Stroop Task. *Journal of Cognitive Neuroscience*, *20*(2), 215–225. https://doi.org/10.1162/jocn.2008.20020

Harris, K. D., Csicsvari, J., Hirase, H., Dragoi, G., & Buzsáki, G. (2003). Organization of cell assemblies in the hippocampus. *Nature*, *424*(6948), 552–556. https://doi.org/10.1038/nature01834

Hebb, D. O. (1949). *The organization of behavior: a neuropsychological theory*. John Wiley and Sons.

Heinrich, A., Carlyon, R. P., Davis, M. H., & Johnsrude, I. S. (2008). Illusory Vowels Resulting from Perceptual Continuity: A Functional Magnetic Resonance Imaging Study. *Journal of Cognitive Neuroscience*, *20*(10), 1737–1752. https://doi.org/10.1162/jocn.2008.20069

Hillebrand, A., & Barnes, G. R. (2005). Beamformer analysis of MEG data. *International Review of Neurobiology*, *68*, 149–171. https://doi.org/10.1016/S0074-7742(05)68006-3

Hipp, J. F., Hawellek, D. J., Corbetta, M., Siegel, M., & Engel, A. K. (2012). Large-scale cortical correlation structure of spontaneous oscillatory activity. *Nature Neuroscience*, *15*(6), 884–890. https://doi.org/10.1038/nn.3101

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, *195*(1), 215–243. https://doi.org/10.1113/jphysiol.1968.sp008455

Hutcheon, B., & Yarom, Y. (2000). Resonance, oscillation and the intrinsic frequency preferences of neurons. *Trends in Neurosciences*, *23*(5), 216–222. https://doi.org/10.1016/S0166-2236(00)01547-2

Iemi, L., Chaumon, M., Crouzet, S. M., & Busch, N. A. (2017). Spontaneous Neural Oscillations Bias Perception by Modulating Baseline Excitability. *Journal of Neuroscience*, *37*(4), 807–819. https://doi.org/10.1523/JNEUROSCI.1432-16.2016

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

Jensen, O., & Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task: Frontal theta increases with memory load. *European Journal of Neuroscience*, *15*(8), 1395–1399. https://doi.org/10.1046/j.1460-9568.2002.01975.x

Jones, E. G., & Powell, T. P. S. (1970). An anatomical study of converging sensory pathways within the cerebral cortex of the monkey. *Brain*, *93*(4), 793–820. https://doi.org/10.1093/brain/93.4.793

Joris, P. X., Schreiner, C. E., & Rees, A. (2004). Neural Processing of Amplitude-Modulated Sounds. *Physiological Reviews*, *84*(2), 541–577. https://doi.org/10.1152/physrev.00029.2003

Kaiser, M., Senkowski, D., Busch, N. A., Balz, J., & Keil, J. (2019). Single trial prestimulus oscillations predict perception of the sound-induced flash illusion. *Scientific Reports*, *9*, 5983. https://doi.org/10.1038/s41598-019-42380-x

Kaiser, M., Senkowski, D., Roa Romero, Y., Riecke, L., & Keil, J. (2018). Reduced low-frequency power and phase locking reflect restoration in the auditory continuity illusion. *European Journal of Neuroscience*, *48*, 2849–2856. https://doi.org/10.1111/ejn.13861

Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1999). Increased Activity in Human Visual Cortex during Directed Attention in the Absence of Visual Stimulation. *Neuron*, *22*(4), 751–761. https://doi.org/10.1016/S0896-6273(00)80734-5

Kayser, C., & Logothetis, N. K. (2007). Do early sensory cortices integrate cross-modal information? *Brain Structure and Function*, *212*(2), 121–132. https://doi.org/10.1007/s00429-007-0154-0

Kayser, C., Petkov, C. I., Augath, M., & Logothetis, N. K. (2007). Functional Imaging Reveals Visual Modulation of Specific Fields in Auditory Cortex. *Journal of Neuroscience*, *27*(8), 1824–1835. https://doi.org/10.1523/JNEUROSCI.4737-06.2007

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., Müller, N., Hartmann, T., & Weisz, N. (2014). 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., 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., & 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. *The Neuroscientist*, *24*(6), 609–626. https://doi.org/10.1177/1073858418755352

Keitel, A., Ince, R. A. A., Gross, J., & Kayser, C. (2017). Auditory cortical delta-entrainment interacts with oscillatory power in multiple fronto-parietal networks. *NeuroImage*, *147*, 32–42. https://doi.org/10.1016/j.neuroimage.2016.11.062

King, A. J., & Palmer, A. R. (1985). Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Experimental Brain Research*, *60*(3), 492–500. https://doi.org/10.1007/bf00236934

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition–timing hypothesis. *Brain Research Reviews*, *53*(1), 63–88. https://doi.org/10.1016/j.brainresrev.2006.06.003

Koch, C., Massimini, M., Boly, M., & Tononi, G. (2016). Neural correlates of consciousness: progress and problems. *Nature Reviews Neuroscience*, *17*(5), 307–321. https://doi.org/10.1038/nrn.2016.22

König, P., Engel, A. K., & Singer, W. (1996). Integrator or coincidence detector? The role of the cortical neuron revisited. *Trends in Neurosciences*, *19*(4), 130–137. https://doi.org/10.1016/S0166-2236(96)80019-1

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., & Shams, L. (2007). Causal Inference in Multisensory Perception. *PLoS ONE*, *2*(9), e943. https://doi.org/10.1371/journal.pone.0000943

Kubovy, M., & Van Valkenburg, D. (2001). Auditory and visual objects. *Cognition*, *80*(1–2), 97–126. https://doi.org/10.1016/S0010-0277(00)00155-4

Lakatos, P. (2005). An Oscillatory Hierarchy Controlling Neuronal Excitability and Stimulus Processing in the Auditory Cortex. *Journal of Neurophysiology*, *94*(3), 1904–1911. https://doi.org/10.1152/jn.00263.2005

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

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of Neuronal Oscillations as a Mechanism of Attentional Selection. *Science*, *320*(5872), 110–113. https://doi.org/10.1126/science.1154735

Lampl, I., & Yarom, Y. (1993). Subthreshold oscillations of the membrane potential: a functional synchronizing and timing device. *Journal of Neurophysiology*, *70*(5), 2181–2186. https://doi.org/10.1152/jn.1993.70.5.2181

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

Lange, J., Oostenveld, R., & Fries, P. (2013). Reduced Occipital Alpha Power Indexes Enhanced Excitability Rather than Improved Visual Perception. *The Journal of Neuroscience*, *33*(7), 3212–3220. https://doi.org/10.1523/JNEUROSCI.3755-12.2013

Lewis, J. W., Beauchamp, M. S., & DeYoe, E. A. (2000). A Comparison of Visual and Auditory Motion Processing in Human Cerebral Cortex. *Cerebral Cortex*, *10*(9), 873–888. https://doi.org/10.1093/cercor/10.9.873

Liégeois-Chauvel, C., Lorenzi, C., Trébuchon, A., Régis, J., & Chauvel, P. (2004). Temporal Envelope Processing in the Human Left and Right Auditory Cortices. *Cerebral Cortex*, *14*(7), 731–740. https://doi.org/10.1093/cercor/bhh033

Linkenkaer-Hansen, K., Nikulin, V. V., Palva, S., Ilmoniemi, R. J., & Palva, J. M. (2004). Prestimulus Oscillations Enhance Psychophysical Performance in Humans. *Journal of Neuroscience*, *24*(45), 10186–10190. https://doi.org/10.1523/JNEUROSCI.2584-04.2004

Lisman, J. E., & Jensen, O. (2013). The Theta-Gamma Neural Code. *Neuron*, *77*(6), 1002–1016. https://doi.org/10.1016/j.neuron.2013.03.007

Llinas, R. R. (1988). The intrinsic electrophysiological properties of mammalian neurons: insights into central nervous system function. *Science*, *242*(4886), 1654–1664. https://doi.org/10.1126/science.3059497

Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique* (2nd edition). MIT Press. https://mitpress.mit.edu/books/introduction-event-related-potential-technique-second-edition

Luo, H., & Poeppel, D. (2007). Phase Patterns of Neuronal Responses Reliably Discriminate Speech in Human Auditory Cortex. *Neuron*, *54*(6), 1001–1010. https://doi.org/10.1016/j.neuron.2007.06.004

Maiworm, M., Bellantoni, M., Spence, C., & Röder, B. (2012). When emotional valence modulates audiovisual integration. *Attention, Perception, & Psychophysics*, *74*(6), 1302–1311. https://doi.org/10.3758/s13414-012-0310-3

Makeig, S., Debener, S., Onton, J., & Delorme, A. (2004). Mining event-related brain dynamics. *Trends in Cognitive Sciences*, *8*(5), 204–210. https://doi.org/10.1016/j.tics.2004.03.008

Makeig, S., Westerfield, M., Jung, T.-P., Enghoff, S., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2002). Dynamic Brain Sources of Visual Evoked Responses. *Science*, *295*(5555), 690–694. https://doi.org/10.1126/science.1066168

Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., & Ro, T. (2009). To See or Not to See: Prestimulus α Phase Predicts Visual Awareness. *Journal of Neuroscience*, *29*(9), 2725–2732. https://doi.org/10.1523/JNEUROSCI.3963-08.2009

Mazaheri, A., & Jensen, O. (2006). Posterior α activity is not phase-reset by visual stimuli. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(8), 2948–2952. https://doi.org/10.1073/pnas.0505785103

McAdams, S., Botte, M.-C., & Drake, C. (1998). Auditory continuity and loudness computation. *The Journal of the Acoustical Society of America*, *103*(3), 1580–1591.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*(5588), 746–748. https://doi.org/10.1038/264746a0

Melloni, L., Molina, C., Pena, M., Torres, D., Singer, W., & Rodriguez, E. (2007). Synchronization of Neural Activity across Cortical Areas Correlates with Conscious Perception. *Journal of Neuroscience*, *27*(11), 2858–2865. https://doi.org/10.1523/JNEUROSCI.4623-06.2007

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

Meredith, M. A., & Stein, B. E. (1983). Interactions among converging sensory inputs in the superior colliculus. *Science*, *221*(4608), 389–391. https://doi.org/10.1126/science.6867718

Meredith, M. A., & Stein, B. E. (1986). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology*, *56*(3), 640–662. https://doi.org/10.1152/jn.1986.56.3.640

Michail, G., & Keil, J. (2018). High cognitive load enhances the susceptibility to non-speech audiovisual illusions. *Scientific Reports*, *8*(1), 1–11. https://doi.org/10.1038/s41598-018-30007-6

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

Miller, G. A., & Licklider, J. C. (1950). The intelligibility of interrupted speech. *The Journal of the Acoustical Society of America*, *22*(2), 167–173.

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

Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). Multisensory auditory–visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Cognitive Brain Research*, *14*(1), 115–128. https://doi.org/10.1016/S0926-6410(02)00066-6

Murray, M. M., Thelen, A., Thut, G., Romei, V., Martuzzi, R., & Matusz, P. J. (2016). The multisensory function of the human primary visual cortex. *Neuropsychologia*, *83*, 161–169. https://doi.org/10.1016/j.neuropsychologia.2015.08.011

Näätänen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin*, *125*(6), 826–859. https://doi.org/10.1037/0033-2909.125.6.826

Noppeney, U., Ostwald, D., & Werner, S. (2010). Perceptual Decisions Formed by Accumulation of Audiovisual Evidence in Prefrontal Cortex. *Journal of Neuroscience*, *30*(21), 7434–7446. https://doi.org/10.1523/JNEUROSCI.0455-10.2010

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716–aac4716. https://doi.org/10.1126/science.aac4716

Parise, C. V., & Ernst, M. O. (2016). Correlation detection as a general mechanism for multisensory integration. *Nature Communications*, *7*, 11543. https://doi.org/10.1038/ncomms11543

Park, H., & Kayser, C. (2019). Shared neural underpinnings of multisensory integration and trial-by-trial perceptual recalibration in humans. *ELife*, *8*, e47001. https://doi.org/10.7554/eLife.47001

Petkov, C. I., O'Connor, K. N., & Sutter, M. L. (2007). Encoding of Illusory Continuity in Primary Auditory Cortex. *Neuron*, *54*(1), 153–165. https://doi.org/10.1016/j.neuron.2007.02.031

Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology*, *110*(11), 1842–1857. https://doi.org/10.1016/S1388-2457(99)00141-8

Pfurtscheller, G., Stancák, A., & Neuper, C. (1996). Event-related synchronization (ERS) in the alpha band — an electrophysiological correlate of cortical idling: A review. *International Journal of Psychophysiology*, *24*(1), 39–46. https://doi.org/10.1016/S0167-8760(96)00066-9

Picton, T. W., John, M. S., Dimitrijevic, A., & Purcell, D. (2003). Human auditory steady-state responses. *International Journal of Audiology*, *42*(4), 177–219.

Picton, T. W., Skinner, C. R., Champagne, S. C., Kellett, A. J. C., & Maiste, A. C. (1987). Potentials evoked by the sinusoidal modulation of the amplitude or frequency of a tone. *The Journal of the Acoustical Society of America*, *82*(1), 165–178. https://doi.org/10.1121/1.395560

Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C., & Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature*, *435*(7045), 1102–1107. https://doi.org/10.1038/nature03687

Raghavachari, S., Kahana, M. J., Rizzuto, D. S., Caplan, J. B., Kirschen, M. P., Bourgeois, B., Madsen, J. R., & Lisman, J. E. (2001). Gating of Human Theta Oscillations by a Working Memory Task. *The Journal of Neuroscience*, *21*(9), 3175–3183. https://doi.org/10.1523/JNEUROSCI.21-09-03175.2001

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.

Rauschecker, J. P., & Tian, B. (2000). Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proceedings of the National Academy of Sciences*, *97*(22), 11800–11806. https://doi.org/10.1073/pnas.97.22.11800

Recasens, M., Gross, J., & Uhlhaas, P. J. (2018). Low-Frequency Oscillatory Correlates of Auditory Predictive Processing in Cortical-Subcortical Networks: A MEG-Study. *Scientific Reports*, *8*(1), 1–12. https://doi.org/10.1038/s41598-018-32385-3

Richter, C. G., Thompson, W. H., Bosman, C. A., & Fries, P. (2017). Top-Down Beta Enhances Bottom-Up Gamma. *The Journal of Neuroscience*, *37*(28), 6698–6711. https://doi.org/10.1523/JNEUROSCI.3771-16.2017

Riecke, L., Esposito, F., Bonte, M., & Formisano, E. (2009). Hearing Illusory Sounds in Noise: The Timing of Sensory-Perceptual Transformations in Auditory Cortex. *Neuron*, *64*(4), 550–561. https://doi.org/10.1016/j.neuron.2009.10.016

Riecke, L., van Opstal, A. J., Goebel, R., & Formisano, E. (2007). Hearing Illusory Sounds in Noise: Sensory-Perceptual Transformations in Primary Auditory Cortex. *Journal of Neuroscience*, *27*(46), 12684–12689. https://doi.org/10.1523/JNEUROSCI.2713-07.2007

Riecke, L., Walter, A., Sorger, B., & Formisano, E. (2011). Tracking Vocal Pitch through Noise: Neural Correlates in Nonprimary Auditory Cortex. *Journal of Neuroscience*, *31*(4), 1479–1488. https://doi.org/10.1523/JNEUROSCI.3450-10.2011

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, *2*(11), 1019–1025. https://doi.org/10.1038/14819

Rodriguez, E., George, N., Lachaux, J.-P., Martinerie, J., Renault, B., & Varela, F. J. (1999). Perception's shadow: long-distance synchronization of human brain activity. *Nature*, *397*(6718), 430–433. https://doi.org/10.1038/17120

Roelfsema, P. R., Engel, A. K., König, P., & Singer, W. (1997). Visuomotor integration is associated with zero time-lag synchronization among cortical areas. *Nature*, *385*(6612), 157–161. https://doi.org/10.1038/385157a0

Rohe, T., Ehlis, A.-C., & Noppeney, U. (2019). The neural dynamics of hierarchical Bayesian causal inference in multisensory perception. *Nature Communications*, *10*(1), 1–17. https://doi.org/10.1038/s41467-019-09664-2

Rohe, T., & Noppeney, U. (2015). Cortical Hierarchies Perform Bayesian Causal Inference in Multisensory Perception. *PLoS Biol*, *13*(2), e1002073. https://doi.org/10.1371/journal.pbio.1002073

Rohe, T., & Noppeney, U. (2018). Reliability-Weighted Integration of Audiovisual Signals Can Be Modulated by Top-down Attention. *Eneuro*, *5*(1), ENEURO.0315-17.2018. https://doi.org/10.1523/ENEURO.0315-17.2018

Romani, G. L., Williamson, S. J., & Kaufman, L. (1982). Tonotopic organization of the human auditory cortex. *Science*, *216*(4552), 1339–1340. https://doi.org/10.1126/science.7079770

Romei, V., Gross, J., & Thut, G. (2012). Sounds Reset Rhythms of Visual Cortex and Corresponding Human Visual Perception. *Current Biology*, *22*(9), 807–813. https://doi.org/10.1016/j.cub.2012.03.025

Ruhnau, P., Hauswald, A., & Weisz, N. (2014). Investigating ongoing brain oscillations and their influence on conscious perception – network states and the window to consciousness. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.01230

Samaha, J., Iemi, L., & Postle, B. R. (2017). Prestimulus alpha-band power biases visual discrimination confidence, but not accuracy. *Consciousness and Cognition*, *54*, 47–55. https://doi.org/10.1016/j.concog.2017.02.005

Sarnthein, J., Petsche, H., Rappelsberger, P., Shaw, G. L., & von Stein, A. (1998). Synchronization between prefrontal and posterior association cortex during human working memory. *Proceedings of the National Academy of Sciences*, *95*(12), 7092–7096. https://doi.org/10.1073/pnas.95.12.7092

Sauseng, P., Klimesch, W., Gruber, W. R., Hanslmayr, S., Freunberger, R., & Doppelmayr, M. (2007). Are event-related potential components generated by phase resetting of brain oscillations? A critical discussion. *Neuroscience*, *146*(4), 1435–1444. https://doi.org/10.1016/j.neuroscience.2007.03.014

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. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, *32*(1), 9–18. https://doi.org/10.1016/j.tins.2008.09.012

Sekuler, R., Sekuler, A. B., & Lau, R. (1997). Sound alters visual motion perception. *Nature*, *385*(6614), 308–308. https://doi.org/10.1038/385308a0

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

Senkowski, D., Schneider, T. R., Foxe, J. J., & Engel, A. K. (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

Shadlen, M. N., & Movshon, J. A. (1999). Synchrony Unbound: A Critical Evaluation of the Temporal Binding Hypothesis. *Neuron*, *24*(1), 67–77. https://doi.org/10.1016/S0896-6273(00)80822-3

Shahin, A. J., Bishop, C. W., & Miller, L. M. (2009). Neural mechanisms for illusory filling-in of degraded speech. *NeuroImage*, *44*(3), 1133–1143. https://doi.org/10.1016/j.neuroimage.2008.09.045

Shams, L., Iwaki, S., Chawla, A., & Bhattacharya, J. (2005). Early modulation of visual cortex by sound: an MEG study. *Neuroscience Letters*, *378*(2), 76–81. https://doi.org/10.1016/j.neulet.2004.12.035

Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature*, *408*(6814), 788. https://doi.org/10.1038/35048669

Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research*, *14*(1), 147–152.

Shams, L., Kamitani, Y., Thompson, S., & Shimojo, S. (2001). Sound alters visual evoked potentials in humans. *Neuroreport*, *12*(17), 3849–3852.

Shams, L., Ma, W. J., & Beierholm, U. (2005). Sound-induced flash illusion as an optimal percept. *Neuroreport*, *16*(17), 1923–1927.

Shimojo, S., & Shams, L. (2001). Sensory modalities are not separate modalities: plasticity and interactions. *Current Opinion in Neurobiology*, *11*(4), 505–509. https://doi.org/10.1016/S0959-4388(00)00241-5

Siapas, A. G., & Wilson, M. A. (1998). Coordinated Interactions between Hippocampal Ripples and Cortical Spindles during Slow-Wave Sleep. *Neuron*, *21*(5), 1123–1128. https://doi.org/10.1016/S0896-6273(00)80629-7

Siegel, M., Donner, T. H., & Engel, A. K. (2012). Spectral fingerprints of large-scale neuronal interactions. *Nature Reviews Neuroscience*, *13*(2), 121–134. https://doi.org/10.1038/nrn3137

Singer, W. (1999). Neuronal Synchrony: A Versatile Code for the Definition of Relations? *Neuron*, *24*(1), 49–65. https://doi.org/10.1016/S0896-6273(00)80821-1

Singer, W., & Gray, C. M. (1995). Visual Feature Integration and the Temporal Correlation Hypothesis. *Annual Review of Neuroscience*, *18*(1), 555–586. https://doi.org/10.1146/annurev.ne.18.030195.003011

Slutsky, D. A., & Recanzone, G. H. (2001). Temporal and spatial dependency of the ventriloquism effect. *NeuroReport*, *12*(1), 7–10.

Spaak, E., Bonnefond, M., Maier, A., Leopold, D. A., & Jensen, O. (2012). Layer-Specific Entrainment of Gamma-Band Neural Activity by the Alpha Rhythm in Monkey Visual Cortex. *Current Biology*, *22*(24), 2313–2318. https://doi.org/10.1016/j.cub.2012.10.020

Spitzer, B., Fleck, S., & Blankenburg, F. (2014). Parametric Alpha- and Beta-Band Signatures of Supramodal Numerosity Information in Human Working Memory. *Journal of Neuroscience*, *34*(12), 4293–4302. https://doi.org/10.1523/JNEUROSCI.4580-13.2014

Spitzer, B., & Haegens, S. (2017). Beyond the Status Quo: A Role for Beta Oscillations in Endogenous Content (Re)Activation. *ENeuro*, *4*(4), ENEURO.0170-17.2017. https://doi.org/10.1523/ENEURO.0170-17.2017

Stanford, T., & Stein, B. (2007). Superadditivity in multisensory integration: putting the computation in context. *Neuroreport*, *18*(8), 787–792. https://doi.org/10.1097/WNR.0b013e3280c1e315

Steinmetz, P. N., Roy, A., Fitzgerald, P. J., Hsiao, S. S., Johnson, K. O., & Niebur, E. (2000). Attention modulates synchronized neuronal firing in primate somatosensory cortex. *Nature*, *404*(6774), 187–190. https://doi.org/10.1038/35004588

Stekelenburg, J. J., Vroomen, J., & de Gelder, B. (2004). Illusory sound shifts induced by the ventriloquist illusion evoke the mismatch negativity. *Neuroscience Letters*, *357*(3), 163–166. https://doi.org/10.1016/j.neulet.2003.12.085

Talavage, T. M., Sereno, M. I., Melcher, J. R., Ledden, P. J., Rosen, B. R., & Dale, A. M. (2004). Tonotopic Organization in Human Auditory Cortex Revealed by Progressions of Frequency Sensitivity. *Journal of Neurophysiology*, *91*(3), 1282–1296. https://doi.org/10.1152/jn.01125.2002

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. https://doi.org/10.1016/S1364-6613(99)01299-1

Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J. (1996). Stimulus Specificity of Phase-Locked and Non-Phase-Locked 40 Hz Visual Responses in Human. *Journal of Neuroscience*, *16*(13), 4240–4249. https://doi.org/10.1523/JNEUROSCI.16-13-04240.1996

Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J. (1997). Oscillatory γ-band (30–70 Hz) activity induced by a visual search task in humans. *Journal of Neuroscience*, *17*(2), 722–734. https://doi.org/10.1523/JNEUROSCI.17-02-00722.1997

Thut, G., Nietzel, A., Brandt, S. A., & Pascual-Leone, A. (2006). α-Band Electroencephalographic Activity over Occipital Cortex Indexes Visuospatial Attention Bias and Predicts Visual Target Detection. *Journal of Neuroscience*, *26*(37), 9494–9502. https://doi.org/10.1523/JNEUROSCI.0875-06.2006

Tiitinen, H. T., Sinkkonen, J., Reinikainen, K., Alho, K., Lavikainen, J., & Näätänen, R. (1993). Selective attention enhances the auditory 40-Hz transient response in humans. *Nature*, *364*(6432), 59–60. https://doi.org/10.1038/364059a0

Uhlhaas, P. J., & Singer, W. (2006). Neural Synchrony in Brain Disorders: Relevance for Cognitive Dysfunctions and Pathophysiology. *Neuron*, *52*(1), 155–168. https://doi.org/10.1016/j.neuron.2006.09.020

van Kerkoerle, T., Self, M. W., Dagnino, B., Gariel-Mathis, M.-A., Poort, J., van der Togt, C., & Roelfsema, P. R. (2014). Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proceedings of the National Academy of Sciences*, *111*(40), 14332–14341. https://doi.org/10.1073/pnas.1402773111

Van Veen, B. D., Van Drongelen, W., Yuchtman, M., & Suzuki, A. (1997). Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Transactions on Biomedical Engineering*, *44*(9), 867–880. https://doi.org/10.1109/10.623056

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

VanRullen, R., Zoefel, B., & Ilhan, B. (2014). On the cyclic nature of perception in vision versus audition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1641), 20130214. https://doi.org/10.1098/rstb.2013.0214

Varela, F., Lachaux, J.-P., Rodriguez, E., & Martinerie, J. (2001). The brainweb: Phase synchronization and large-scale integration. *Nature Reviews Neuroscience*, *2*(4), 229–239. https://doi.org/10.1038/35067550

von der Malsburg, C. (1981). *The Correlation Theory of Brain Function* [Internal Report 81-2]. Department of Neurobiology, Max Planck Institute for Biophysical Chemistry.

von Stein, A., Rappelsberger, P., Sarnthein, J., & Petsche, H. (1999). Synchronization Between Temporal and Parietal Cortex During Multimodal Object Processing in Man. *Cerebral Cortex*, *9*(2), 137–150. https://doi.org/10.1093/cercor/9.2.137

von Stein, A., & Sarnthein, J. (2000). Different frequencies for different scales of cortical integration: from local gamma to long range alpha/theta synchronization. *International Journal of Psychophysiology*, *38*(3), 301–313. https://doi.org/10.1016/S0167-8760(00)00172-0

Wallace, M. T., Roberson, G. E., Hairston, W. D., Stein, B. E., Vaughan, J. W., & Schirillo, J. A. (2004). Unifying multisensory signals across time and space. *Experimental Brain Research*, *158*(2). https://doi.org/10.1007/s00221-004-1899-9

Wang, X.-J. (2010). Neurophysiological and Computational Principles of Cortical Rhythms in Cognition. *Physiological Reviews*, *90*(3), 1195–1268. https://doi.org/10.1152/physrev.00035.2008

Wang, X.-J., & Buzsáki, G. (1996). Gamma Oscillation by Synaptic Inhibition in a Hippocampal Interneuronal Network Model. *Journal of Neuroscience*, *16*(20), 6402–6413. https://doi.org/10.1523/JNEUROSCI.16-20-06402.1996

Warren, R. M., Obusek, C. J., & Ackroff, J. M. (1972). Auditory Induction: Perceptual Synthesis of Absent Sounds. *Science*, *176*(4039), 1149–1151. https://doi.org/10.1126/science.176.4039.1149

Watkins, S., Shams, L., Josephs, O., & Rees, G. (2007). Activity in human V1 follows multisensory perception. *NeuroImage*, *37*(2), 572–578. https://doi.org/10.1016/j.neuroimage.2007.05.027

Watkins, S., Shams, L., Tanaka, S., Haynes, J.-D., & Rees, G. (2006). Sound alters activity in human V1 in association with illusory visual perception. *NeuroImage*, *31*(3), 1247–1256. https://doi.org/10.1016/j.neuroimage.2006.01.016

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*, *111*(4), E417–E425. https://doi.org/10.1073/pnas.1317267111

Welch, R. B., DutionHurt, L. D., & Warren, D. H. (1986). Contributions of audition and vision to temporal rate perception. *Perception & Psychophysics*, *39*(4), 294–300. https://doi.org/10.3758/BF03204939

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, *88*(3), 638–667. https://doi.org/10.1037/0033-2909.88.3.638

Womelsdorf, T., Schoffelen, J.-M., Oostenveld, R., Singer, W., Desimone, R., Engel, A. K., & Fries, P. (2007). Modulation of Neuronal Interactions Through Neuronal Synchronization. *Science*, *316*(5831), 1609–1612. https://doi.org/10.1126/science.1139597

Womelsdorf, T., Valiante, T. A., Sahin, N. T., Miller, K. J., & Tiesinga, P. (2014). Dynamic circuit motifs underlying rhythmic gain control, gating and integration. *Nature Neuroscience*, *17*(8), 1031–1039. https://doi.org/10.1038/nn.3764

Worden, M. S., Foxe, J. J., Wang, N., & Simpson, G. V. (2000). Anticipatory Biasing of Visuospatial Attention Indexed by Retinotopically Specific α-Band Electroencephalography Increases over Occipital Cortex. *The Journal of Neuroscience*, *20*(6), RC63–RC63. https://doi.org/10.1523/JNEUROSCI.20-06-j0002.2000

Zierul, B., Röder, B., Tempelmann, C., Bruns, P., & Noesselt, T. (2017). The role of auditory cortex in the spatial ventriloquism aftereffect. *NeuroImage*, *162*, 257–268. https://doi.org/10.1016/j.neuroimage.2017.09.002

Appendix A: Original publications

Study 1

Kaiser, M., Senkowski, D., Roa Romero, Y., Riecke, L., & Keil, J. (2018). Reduced low-frequency power and phase locking reflect restoration in the auditory continuity illusion. *European Journal of Neuroscience*, *48*, 2849–2856. https://doi.org/10.1111/ejn.13861

This article is not included in the online version of this dissertation for copyright reasons.

Study 2

Kaiser, M., Senkowski, D., Busch, N. A., Balz, J., & Keil, J. (2019). Single trial prestimulus oscillations predict perception of the sound-induced flash illusion. *Scientific Reports*, *9*, 5983. https://doi.org/10.1038/s41598-019-42380-x

www.nature.com/scientificreports

SCIENTIFIC REPORTS

OPEN

Single trial prestimulus oscillations predict perception of the sound-induced flash illusion

Received: 12 July 2018
Accepted: 1 April 2019

Mathis Kaiser[1,2], Daniel Senkowski[1], Niko A. Busch[3], Johanna Balz[1] & Julian Keil[4]

In the sound-induced flash illusion, auditory input affects the perception of visual stimuli with a large inter- and intraindividual variability. Crossmodal influence in this illusion has been shown to be associated with activity in visual and temporal areas. In this electroencephalography study, we investigated the relationship between oscillatory brain activity prior to stimulus presentation and subsequent perception of the illusion on the level of single trials. Using logistic regression, we modeled the perceptual outcome dependent on oscillatory power. We found that 25 Hz to 41 Hz activity over occipital electrodes from 0.17 s to 0.05 s prior to stimulus onset predicted the perception of the illusion. A t-test of power values, averaged over the significant cluster, between illusion and no-illusion trials showed higher power in illusion trials, corroborating the modeling result. We conclude that the observed power modulation predisposes the integration of audiovisual signals, providing further evidence for the governing role of prestimulus brain oscillations in multisensory perception.

In order to successfully navigate our environment, it is vitally important to integrate information from various sensory sources. Recent studies have highlighted the crossmodal influence between sensory modalities and the underlying neural mechanisms[1]. An established paradigm to study crossmodal influence is the sound-induced flash illusion (SIFI), where auditory input affects the processing and perception of visual input. In a typical SIFI paradigm, participants are presented with one or more brief sound stimuli concurrently with one or more bright visual stimuli. Participants are then asked to report the number of perceived flashes. The illusion manifests itself in responses that are influenced by auditory information; the number of flashes that participants subjectively perceive depends on the number of sounds they have heard[2]. We here regard occurrence of the SIFI as an example of multisensory integration resulting from crossmodal stimulation[3]. Importantly, the illusion does not occur in all trials, but with a large intra- and interindividual variability[4]. This allows for analysis approaches that differentiate the neural activity patterns that predict the perceptual events of interest.

With regard to the neuroanatomical substrates of crossmodal influence in the SIFI, activity in occipital and temporal areas has been shown to be associated with illusory flash perception: Studies using functional magnetic resonance imaging (fMRI) have established that the activity level in V1 reflects the number of flashes subjectively perceived, with increased activity for additionally perceived flashes and decreased activity for perceptual fusion, i.e. perception of one flash following presentation of two flashes and one sound[5,6]. Interestingly, activity in the right superior temporal gyrus is increased in both cases. This suggests that activity in lower and higher tier areas differentially reflects the processing and integration of multisensory stimuli. Consistent with these findings, event-related potentials (ERPs) over occipital areas, as obtained through electroencephalography (EEG), are modulated around 100 ms after stimulus onset in the SIFI[7,8]. Moreover, early gamma power over occipital areas is increased following the illusion[9,10], and the mid-latency gamma power increase in temporal areas is correlated with increased likelihood of perceiving the illusion[11]. While early modulations of ERPs and spectral activity in electrophysiological studies have traditionally been viewed as correlates of modality-specific processing in unisensory areas, influences from other unisensory, polymodal and frontal areas cannot be ruled out[12,13].

In recent years, the brain states and network configurations predisposing multisensory integration and – more generally – conscious perception have come into focus[14–16]. Several studies have established links between

[1]Department of Psychiatry and Psychotherapy, Charité–Universitätsmedizin Berlin, Große Hamburger Str. 5-11, 10115, Berlin, Germany. [2]Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, Luisenstraße 56, 10117, Berlin, Germany. [3]Institute of Psychology, University of Münster, Fliednerstr. 21, 48149, Münster, Germany. [4]Biological Psychology, Christian-Albrechts-University Kiel, Olshausenstraße 62, 24118, Kiel, Germany. Correspondence and requests for materials should be addressed to M.K. (email: mathis.kaiser@charite.de)

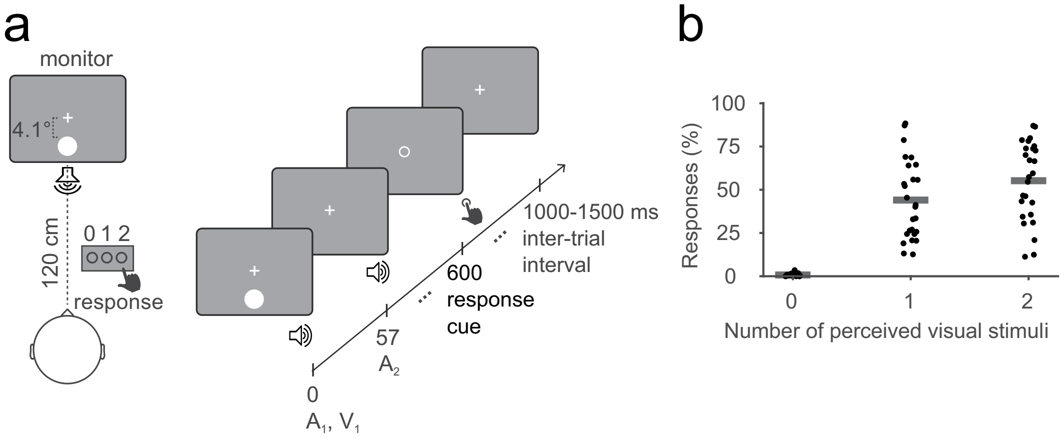

Figure 1. (**a**) Experimental design. (**b**) Response rates in the A_2V_1 condition. Individual response rates are indicated by dots, the average is indicated by horizontal bars.

oscillatory brain activity prior to stimulus onset and subjective perception of multisensory stimuli[17–19]. Using magnetoencephalography (MEG), Keil *et al.*[18] have found that beta band power is increased in the left temporal gyrus prior to perceiving the SIFI. Moreover, in a visuotactile flash illusion paradigm, Lange *et al.*[19] have found that decreased alpha band power in occipital areas and increased gamma band power in occipitoparietal and right temporal areas precede illusory flashes. These studies provide the primary research background for the current study. However, these studies analyzed averaged estimates of oscillatory power across trials, thereby neglecting the trial-by-trial fluctuations of cortical activity, which are potentially informative regarding variability in perception and other cognitive functions[20,21].

Here, we aimed to elucidate the relationship between oscillatory brain activity prior to stimulus presentation and perception of the SIFI on the level of single trials, initially focusing on the frequencies up to 41 Hz to follow previous analysis protocols[18] and then extending to higher frequencies. We employed a logistic regression model that predicts the binary perceptual outcome (*illusion*/*no illusion*) dependent on the level of oscillatory activity in each trial. By incorporating single trial information in the analysis, we aimed to detect subtle modulations of oscillatory activity that might go unnoticed when averaged power is statistically compared between different perceptual outcomes. When trial numbers are strongly unbalanced across conditions, trial numbers are often equalized in conventional analyses by random sampling, thereby losing information. Employing a regression model allowed us to circumvent this. Based on the previous work, we hypothesized that ongoing oscillatory activity influences upcoming perception on the trial-by-trial level. We also explored a possible relationship between prestimulus activity and early evoked potentials by correlating prestimulus oscillatory power differences with previously demonstrated ERP amplitude differences[10,22].

Results

In this experiment, six audiovisual stimulus combinations were presented: A_0V_1, A_0V_2, A_1V_1, A_2V_0, A_2V_1 and A_2V_2, where the indexed numbers represent the number of auditory (A) and visual (V) stimuli. A_2V_1 is the illusion condition, where either one or two flashes can be perceived. Illusion rates in the sample varied between 11 and 87% (mean: 55 ± 22% SD, also see Fig. 1b). Behavioral accuracies (referring to veridical visual perception) in the conditions were as follows (mean ± SD): A_0V_1: 83 ± 15%, A_0V_2: 83 ± 16%, A_1V_1: 94 ± 9%, A_2V_0: 92 ± 9%, A_2V_1: 44 ± 22%, A_2V_2: 93 ± 12%. Accuracies above 80% in all but the illusion condition (A_2V_1) indicate that subjects generally responded accurately to visual input.

When quantifying the relationship between single-trial EEG power in the prestimulus time window and behavioral outcome, we found a significant cluster of regression weights over occipital electrodes between 25 and 41 Hz, i.e. the high beta to low gamma band, from 0.17 s to 0.05 s before stimulus onset ($p = 0.006$, Fig. 2a,b). The individual regression coefficients, averaged across the cluster, were positive in 22 of 26 participants, indicating a positive relationship between increased beta/gamma band power and SIFI perception (Fig. 2d). We found high Bayes Factor values at a subset of right occipitotemporal electrodes, indicating strong support for the hypothesis that oscillatory power at these electrodes predicts perception on a trial-by-trial basis (Fig. 2c). The mean of Bayes factor values, averaged across the cluster, was 67.61. A dependent-samples t-test of normalized power values averaged over the significant channel/frequency/time window (*illusion* vs. *no illusion*) resulted in a significant effect ($t(25) = 3.13$, $p = 0.0045$). This further supports the notion that oscillatory power has predictive value for crossmodal influence. An unrestricted, cluster-corrected dependent-samples t-test for the same comparison did not reveal significant effects ($p = 0.51$ for the positive cluster with the largest effect and comparable extent to the modeling result). This indicates that the single-trial modeling approach can uncover relationships that are not picked up by conventional analyses comparing means between conditions.

Since the significant effect was located at the upper end of the analyzed frequency window, we additionally performed an analysis of gamma activity (40 to 100 Hz), which did not reveal any significant effects. The cluster with the largest effect in the gamma band extended from 40 to 45 Hz and from 0.13 to 0.05 s prior to stimulus onset ($p = 0.1018$).

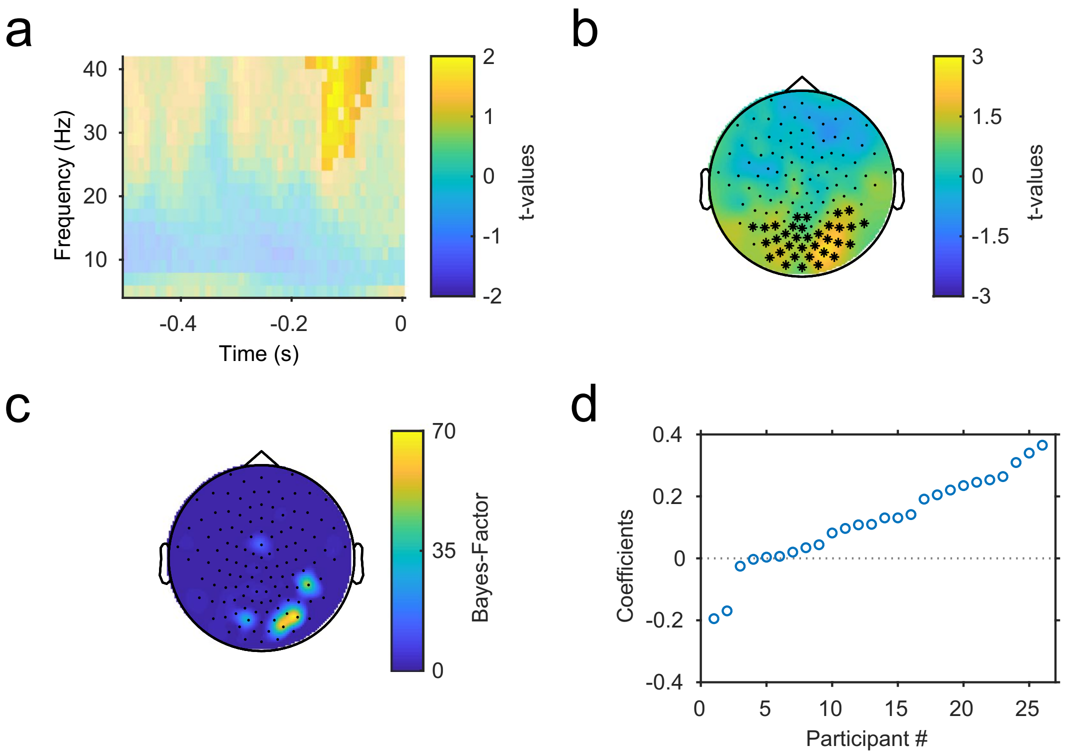

Figure 2. (**a**) Time frequency spectrum of t-values for the comparison of observed regression weights against dummy data, averaged over the significant channels. Significant regions are indicated by saturation. (**b**) Topography of t-values, averaged over the significant time-frequency region. Significant channels are marked by asterisks. (**c**) Topography of Bayes-Factor values, averaged over the significant time-frequency region. (**d**) Individual regression coefficients, averaged over the significant cluster. Participants are sorted in ascending order of the coefficient.

Using a cluster-corrected dependent-samples t-test, we replicated previous findings[10,22] of an ERP amplitude difference over central sensors in the time range of 0.1 to 0.17 s after stimulus onset ($p = 0.004$, see supplementary Fig. S1). Amplitudes of this early negative component were larger in *illusion* compared to *no illusion* trials. To link pre- and poststimulus activity, we calculated a Pearson's correlation between the identified differences in oscillatory power and ERP amplitudes across participants. However, the correlation did not yield a significant result ($r = 0.16$, $p = 0.43$, see supplementary Fig. S1).

Discussion

In this study, we investigated the relationship between single-trial prestimulus oscillatory brain activity and audiovisual crossmodal influence in the sound-induced flash illusion using a logistic regression model. Across participants, increased power over occipital electrodes between 25 and 41 Hz, starting 170 ms before stimulus onset, predicted subsequent perception of the illusion. Although the effect was located at the upper end of the analyzed frequency window, a subsequent analysis of high-frequency gamma band power resulted in no significant effects. There was a trend level cluster in the gamma analysis that could be part of the cluster obtained in the lower-frequency analysis. However, since that cluster extended only to 45 Hz, this would not substantially change the conclusion that the contribution of oscillatory power to illusory perception seems to be restricted to the upper beta/lower gamma frequency range. In a follow-up analysis, we calculated Bayes-Factor values for the t-values resulting from the comparison of regression weights against random data with the same mean and standard deviation, which quantify the empirical support for the alternative hypothesis (i.e. that the regression weights are significantly different from noise). We found the most robust effects over right occipitotemporal channels, suggesting a contribution of occipital and, possibly, right temporal areas to the observed effect, although this suggestion must remain speculative at this point since no source analysis was performed. The exploratory analysis of a relationship between prestimulus power and ERP amplitude differences did not reveal a significant effect. It is worth noting that this does not preclude a link between prestimulus oscillatory activity and other types of stimulus-induced responses.

Previous studies investigating prestimulus oscillatory activity in double flash illusions using MEG have found that crossmodal influence is associated with increased beta band power (13–21 Hz) in left temporal areas[18]; enhanced excitability in visual areas, as reflected by decreased alpha band power; and increased gamma band power in a more extensive cortical network including temporal areas[19]. Prestimulus alpha band power is also decreased in the triple flash illusion, where two rapidly presented flashes are occasionally perceived as three[23]. Interestingly, the optimal delay between flashes is correlated with the subject-specific impulse response frequency, pointing to a contribution of oscillatory reverberation to illusory visual perception. Similarly, the phase of ongoing alpha band oscillations influences crossmodal synchrony perception[24]. A phase reset of ongoing slow oscillations in unisensory areas is a likely mechanism for the modulation of activity following heteromodal input[13,25,26]. The ensuing phase alignment increases local, modality-specific excitability and promotes efficient communication between different cortical areas. However, prestimulus alpha band power did not have predictive value for perception of the SIFI in the present study, and analyses of the possible influence of oscillatory phase were beyond scope.

Previously, Lange *et al.*[19] reported that a prestimulus modulation of excitability in occipital and temporal areas, as reflected by decreased alpha and increased gamma band power, entails perception of the SIFI. This modulation of excitability might facilitate the crossmodal influence of auditory input on visual stimulus processing. Although the frequency range of the currently observed effect is lower than the gamma band effect previously reported, the functional roles might be similar. Interestingly, a recent study combining EEG and Bayesian modeling of behavioral responses in a SIFI paradigm has found a correlation between prestimulus oscillatory gamma activity over occipital electrodes and the perceptual prior that signals originate from a common source[27]. In other words, high prestimulus occipital gamma power increased the tendency to perceptually bind audiovisual signals in the SIFI. The spectrotemporal extent of the highest correlation values has a substantial overlap with the significant cluster in the present study.

A study using transcranial direct current stimulation (tDCS) has shown that up-regulation of occipital cortex results in decreased illusion rates in the SIFI, while up-regulation of temporal cortex results in increased illusion rates[28]. Furthermore, the detection of TMS-induced phosphenes, which are enhanced by concurrent auditory or tactile stimulation, selectively benefits from tDCS over temporal and parietal, but not occipital cortex[29]. These findings suggest that variations in crossmodal influence are a function of excitability in the cortical areas where influence originates, rather than in the target sites of influence. A possible reason for this is that excitability changes induced by tDCS promote unisensory processing in the target sites. Related to this notion, inferring the spatial location of incongruent audiovisual signals is performed by a hierarchy of cortical areas: primary visual and auditory areas estimate location under the assumption of independent sources, while more parietal areas represent location under the assumption of a common source and integrate the estimates, which results in crossmodal influence[30]. This means that, although functional interactions and direct projections between heteromodal unisensory areas have been demonstrated[25,31], the representation resulting from weighting of perceptual priors and sensory input seems to rely more on multisensory than unisensory areas in the case of crossmodal influence. Given the context of these findings, the predictive value of high-beta/low-gamma band oscillations for the SIFI that we found in the present study might capture a biasing influence of auditory information on multisensory integration areas (e.g. superior temporal gyrus). Related to this, a previous study showed that increased lower beta band functional connectivity between auditory and multisensory temporal areas but decreased connectivity between temporal and visual areas is related to the perception of the SIFI[18]. Studies investigating the mechanisms of interareal information transfer in the visual system have shown that the direction of influence is correlated with activity in distinct frequency bands: feedforward influence is carried by gamma band synchronization, and feedback influence is carried by beta band synchronization[32,33]. Furthermore, top-down influence in the beta band has been shown to enhance bottom-up gamma band responses around 100 ms later[34]. We therefore suggest that increased 25 Hz to 41 Hz oscillatory power that predicts perception of the SIFI in the present study could reflect the interplay of feedback and feedforward influence between auditory, visual and multisensory regions. Nevertheless, more conclusive evidence for this suggestion, as well as for the direction of influence and the importance of local excitability, should be obtained in future studies employing directed connectivity metrics on the source level.

Given reports that crossmodal interaction in the SIFI is associated with increased gamma band power after stimulus onset[9–11], one could argue that the current results are confounded with spectral leakage from the post-stimulus period. The time window we used for time frequency analysis is 3 cycles at each frequency, resulting in a time window of 0.12 s at 25 Hz (shorter for increasing frequencies). Hence, 25 Hz activity from up to 0.06 s later may have leaked into a given data point being analyzed (albeit less towards the edge of the time window due to tapering, and even less for higher frequencies). It is unlikely that the effect we found, which extended from 25 to 41 Hz and from 0.170 to 0.05 s prestimulus, was substantially affected by poststimulus activity, since it already arises long before substantial contamination may have occurred: only the last 0.01 seconds at the lower frequency edge of the cluster may have been affected. Another concern may be our non-canonical analysis approach, consisting of a logistic regression of a binary perceptual outcome on oscillatory activity and a subsequent test of the obtained regression weights against random noise. Since such an analysis has not been applied to EEG data in this manner before, we validated the finding by performing a classical t-test of power values averaged over trials of the two perceptual conditions, which confirmed that power was higher in *illusion* vs. *no illusion* trials. We therefore argue that our approach, although it requires further validation, merits application because it allows modeling perceptual variation on the single trial level, and resulted in findings that could be confirmed using more established methods.

Future research could address the role of crossmodal phase resetting in the SIFI and the direction of information transfer between relevant brain areas. Such studies would also allow insight into the relative importance of local excitability changes in sensory areas and feedback by higher-order areas[1]. Furthermore, the sources of modulations in oscillatory activity should be localized more accurately to characterize the pertinent network topology and corresponding activity patterns. Electrocorticography is one method that has been successfully employed to study the mechanisms of crossmodal influence with sufficiently high temporal and spatial resolution[26]. A limitation of the current study is that it is unclear whether the effect of prestimulus oscillations specifically affects the SIFI, i.e. the number of illusory flashes in the A_2V_1 condition. Instead, prestimulus oscillations might generally affect accuracy or bias in the flash-counting task irrespective of the veridical number of flashes. However, modeling the perceptual outcome in the other conditions was not feasible due to ceiling effects. In follow-up studies, it would therefore be interesting to modify stimuli such that conditions with two visual flashes and with no visual flashes (A_2V_2 and A_2V_0, respectively) yield greater variability in the number of reported flashes, and to include the veridical number of flashes as an additional predictor in the regression model. It would then be possible to differentiate the influence of visual input, prestimulus oscillations, and their interaction.

To conclude, the present study demonstrates that single-trial oscillatory activity predicts the integration of multisensory signals, as exemplified by an influence of auditory input on visual perception in the SIFI. A number of studies have found early modulations of cortical activity over visual areas related to the SIFI, which can be integrated in the context of the present findings: We propose that crossmodal influence in the SIFI is facilitated already before stimulus onset. This facilitation is reflected by reduced alpha band power in occipital areas and increased gamma band power in temporal areas[19], which might indicate enhanced excitability of polysensory pathways. Changes in excitability induced in visual cortex by tDCS, which result in changes of the illusion rate[28], are likely confined to unisensory visual pathways. Furthermore, multisensory areas in the superior temporal gyrus, as well as frontal areas, exert influence over audiovisual information flow, which is reflected by modulation of beta band oscillations[18]. These brain areas, which are relevant for a predisposition for crossmodal influence, are also involved in stimulus processing: activation levels in early visual areas indicate the number of subjectively perceived stimuli[5,8], while activity in superior temporal areas reflects crossmodal interaction more generally[6,11]. Importantly, these later modulations of cortical activity may already be primed by the ongoing prestimulus oscillations revealed in the present study.

Methods

Participants. Forty healthy volunteers participated in the study. All participants had normal hearing and normal or corrected-to-normal vision, and reported no history of neurological or psychiatric disorders. All participants gave written informed consent, the study was conducted in accordance with the Declaration of Helsinki and approved by the ethics committee of the Charité–Universitätsmedizin Berlin. Findings from this dataset focusing on independent aspects of neural processing have previously been reported[4,11,22]. In accordance with our previous analysis of alpha frequency from the same dataset, a subset of 26 participants who showed illusion rates between 10 and 90% were selected for analysis[4]. The illusion rate is computed as the proportion of trials containing two auditory stimuli and one visual stimulus (A_2V_1) where there is an illusory perception of two flashes. The reason for excluding participants with extreme illusion rates is that these participants either did not reliably perceive the illusion, or primarily relied on auditory instead of visual input in reporting the number of perceived flashes. The mean age of the 26 selected participants (8 female; 1 left-handed) was 33.7 years (range: 18–51 years).

Experimental design. The experiment was conducted in a sound-attenuated electrically shielded chamber. Visual stimuli were presented on a CRT monitor with a background luminance of 21 cd/m^2 and a refresh rate of 75 Hz. Auditory stimuli were presented from a central speaker below the screen. Six stimulus combinations were presented: A_0V_1, A_0V_2, A_1V_1, A_2V_0, A_2V_1 and A_2V_2, where the indexed numbers represent the number of auditory (A) and visual (V) stimuli. A_2V_1 is the illusion condition, while the other combinations served as behavioral control trials. We did not use all possible stimulus combinations due to time constraints. Visual stimuli were presented for 10 ms and consisted of a white disk subtending a visual angle of 1.6° with a luminance of 89 cd/m^2. Visual stimuli were presented at 4.1° centrally below the fixation cross. Auditory stimuli were presented for 7 ms and consisted of a 73 dB (SPL) 1000 Hz sine wave tone. Three hundred A_2V_1 trials and 150 trials per control condition were presented in random order in eight blocks. Subjects were asked to indicate how many visual stimuli they perceived (zero, one or two) with a button press of their right hand. For details of the experimental setup, see Fig. 1a.

Acquisition and Preprocessing of EEG data. EEG was recorded using a 128-channel active electrode cap (EasyCap, Herrsching, Germany), including one horizontal and one vertical electrooculography electrode to monitor eye movements, and Brainamp DC amplifiers (Brainproducts, Gilching, Germany). Data were recorded in reference to an electrode placed on the nose with a sampling frequency of 1000 Hz and a pass band from 0.016 to 250 Hz.

EEG data processing was performed in MATLAB (MathWorks, Natick, MA, USA) using the EEGlab[35] and FieldTrip[36] toolboxes and custom scripts. Data were filtered using a two-pass Hamming-windowed FIR filter, with an order of 2999 and a −6 dB cutoff frequency of 1 Hz for the high pass, and an order of 23 and a cutoff frequency of 125 Hz for the low pass. A 4th-order two-pass Butterworth filter with a stop band from 49 to 51 Hz was used to filter out line noise. Data were subsequently downsampled to 500 Hz and epoched from −1 to 3 s relative to the first auditory stimulus onset. Trials and channels that contained large artifacts were removed after visual inspection. Independent component analysis using an extended infomax algorithm[37] was performed on the truncated data and components that represented eye blinks or cardiac activity were removed. Electrooculography channels were removed from the data and rejected EEG channels were interpolated using spherical interpolation. Next, data were re-referenced to common average. Trials that still exceeded a threshold of ±100 µV after these procedures were rejected automatically. On average, 98.65 (±50.9 SD) trials, 1.35 (±1.39 SD) channels, and 15.65 (±6.7 SD) components were removed from each dataset. Finally, epochs were sorted according to combination of audiovisual stimuli and response and A_2V_1-trials, where 1 (*no illusion*) or 2 flashes (*illusion*) were perceived, were selected. After preprocessing, there were on average 144.65 (±59.11 SD) A_2V_1 -trials, in which an illusion was perceived, and 116.35 (±61.82) A_2V_1-trials, in which no illusion was perceived in the individual datasets.

Analysis of EEG data. Single trial EEG data were time-frequency transformed using a single Hanning taper with a window length of 3 cycles at each frequency. Time-frequency analysis was performed for the time window from −1 to 1 s around the onset of the first auditory stimulus, with a step size of 10 ms, for the frequencies from 5 to 41 Hz, with a frequency resolution of 2 Hz. We additionally performed an analysis of activity in the gamma frequency band from 40 to 100 Hz in steps of 5 Hz, using multiple tapers with a window length of 200 ms and frequency smoothing of ±10 Hz. The only difference in the analysis protocols for the lower and higher frequencies was the use of single and multiple tapers, respectively. We have used multiple tapers for higher frequencies because they offer better control over temporal and spectral resolution, and thus better signal-to-noise ratio for higher frequencies. Statistical modeling was otherwise identical. Thus, the analysis for higher frequencies can be considered an extension of the primary analysis. Taken together, they should reveal the entire frequency range of the observed effect.

To quantify the relationship between single-trial EEG power in the prestimulus time window and behavioral outcome, we calculated logistic regression weights between the spectral activity in a 500 ms time window prior to stimulus presentation and the binary perceptual rating (i.e. *illusion* vs. *no illusion*) separately for each electrode, frequency, and time-point. Before calculation of the regression weights, time-frequency data were scaled between zero and one, and then the inverse of the normal cumulative distribution function was taken. This was done in order to approximate the data to a normal distribution due to concerns regarding the use of regression approaches with non-normal distributed data[38]. The regression model can be stated as follows:

$$\log\left(\frac{p_{illusion}}{1-p_{illusion}}\right)=\beta_0+\beta_1 X$$

where $p_{illusion}$ is the probability of perceiving the illusion, β_0 is the intercept, β_1 is the regression weight, and X is the normalized single-trial power.

For statistical evaluation, a cluster-based permutation test was used[39]. Within each subject, we calculated the mean and standard deviation of the observed regression weights across channels, frequencies and time-points. We then generated dummy regression weights for each subject that were randomly selected from a normal distribution with the mean and standard deviation calculated in the first step. For the group-level analysis, we compared the observed regression weights against these dummy data using a dependent-samples t-test with cluster correction. Samples at a given channel/frequency/time-point entered a cluster when the significance level exceeded 0.05 in at least three neighboring electrodes. The test statistic was computed as the sum of t-values within a cluster. The comparison was repeated with permuted condition labels 1000 times, yielding a distribution of test statistics under the null hypothesis. The significance level of the cluster was computed as the proportion of permutations that resulted in a test statistic exceeding the observed one. The aim of this procedure was to identify clusters of regression weights that stand out from the subject-specific distribution. Instead of testing the null hypothesis that the population level regression weights are zero, this procedure makes no assumptions about the real distribution of coefficients, which might be negatively or positively biased for single subjects. While this approach does not allow the identification of clusters of coefficients with high or low absolute values, it does detects those clusters of coefficients that differ from the subject-specific distribution. Based on the absolute regression weights within the cluster, we assessed the direction of the relationship between power and perception.

Additionally, within the channel/frequency/time cluster identified above, we performed a conventional t-test of normalized power values averaged over different perceptual conditions (*illusion* vs. *no illusion*) to check the robustness of the model results. We restricted this analysis to the time/frequencies/channels derived from the regression analysis to corroborate the result within the observed cluster. We performed an unrestricted cluster-based dependent samples t-test (*illusion* vs. *no illusion*) with the same parameters as in the statistical test for the regression weights to examine whether the modeling approach is more sensitive than conventional methods. We also calculated the Bayes factor (BF) based on the outcome of a dependent-samples t-test between observed regression weights and dummy data. The BF summarizes the ratio of evidence for a true effect vs. the evidence for the null hypothesis of no effect and was computed from the t-values using the *ttest.tstat* function from the BayesFactor R package[40].

To explore whether differences in prestimulus oscillatory power affect differences of ERP amplitudes related to different perceptual outcomes, we calculated a Pearson's correlation between both measures across participants. We first calculated ERPs for *illusion* and *no illusion* trials separately, applying baseline-correction with the 200 ms interval before stimulus onset and a 30 Hz low-pass filter (hamming-windowed sinc FIR, order 220). Then, we compared the ERPs from both conditions using a cluster-based dependent samples t-test with similar parameters as before, only the cluster alpha criterion decreased to 0.01 for a more focal cluster. Next, we calculated absolute amplitude differences, averaged over the resulting negative cluster for each participant. Similarly, we calculated absolute differences of normalized oscillatory power between *illusion* and *no-illusion* trials (averaged over the significant cluster identified from modeling). Finally, we correlated these differences of amplitude and power across participants.

Data Availability

Raw data cannot be made available because participants did not consent to public dissemination. Processed data can be made available upon reasonable request.

References

1. Keil, J. & Senkowski, D. Neural Oscillations Orchestrate Multisensory Processing. *The Neuroscientist* **24**, 609–626 (2018).
2. Shams, L., Kamitani, Y. & Shimojo, S. What you see is what you hear. *Nature* **408**, 788 (2000).
3. Stein, B. E. *et al.* Semantic confusion regarding the development of multisensory integration: a practical solution. *Eur. J. Neurosci.* **31**, 1713–1720 (2010).
4. Keil, J. & Senkowski, D. Individual Alpha Frequency Relates to the Sound-Induced Flash Illusion. *Multisensory Res.* **30**, 565–578 (2017).

5. Watkins, S., Shams, L., Josephs, O. & Rees, G. Activity in human V1 follows multisensory perception. *NeuroImage* **37**, 572–578 (2007).
6. Watkins, S., Shams, L., Tanaka, S., Haynes, J.-D. & Rees, G. Sound alters activity in human V1 in association with illusory visual perception. *NeuroImage* **31**, 1247–1256 (2006).
7. Shams, L., Iwaki, S., Chawla, A. & Bhattacharya, J. Early modulation of visual cortex by sound: an MEG study. *Neurosci. Lett.* **378**, 76–81 (2005).
8. Shams, L., Kamitani, Y., Thompson, S. & Shimojo, S. Sound alters visual evoked potentials in humans. *Neuroreport* **12**, 3849–3852 (2001).
9. Bhattacharya, J., Shams, L. & Shimojo, S. Sound-induced illusory flash perception: role of gamma band responses. *Neuroreport* **13**, 1727–1730 (2002).
10. Mishra, J., Martinez, A., Sejnowski, T. J. & Hillyard, S. A. Early Cross-Modal Interactions in Auditory and Visual Cortex Underlie a Sound-Induced Visual Illusion. *J. Neurosci.* **27**, 4120–4131 (2007).
11. Balz, J. *et al*. GABA concentration in superior temporal sulcus predicts gamma power and perception in the sound-induced flash illusion. *NeuroImage* **125**, 724–730 (2016).
12. Molholm, S. *et al*. Multisensory auditory–visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Cogn. Brain Res.* **14**, 115–128 (2002).
13. Kayser, C., Petkov, C. I. & Logothetis, N. K. Visual Modulation of Neurons in Auditory Cortex. *Cereb. Cortex* **18**, 1560–1574 (2008).
14. Lange, J., Keil, J., Schnitzler, A., van Dijk, H. & Weisz, N. The role of alpha oscillations for illusory perception. *Behav. Brain Res.* **271**, 294–301 (2014).
15. Senkowski, D., Schneider, T. R., Foxe, J. J. & Engel, A. K. Crossmodal binding through neural coherence: implications for multisensory processing. *Trends Neurosci.* **31**, 401–409 (2008).
16. Weisz, N. *et al*. Prestimulus oscillatory power and connectivity patterns predispose conscious somatosensory perception. *Proc. Natl. Acad. Sci.* **111**, E417–E425 (2014).
17. Keil, J., Müller, N., Ihssen, N. & Weisz, N. On the variability of the McGurk effect: audiovisual integration depends on prestimulus brain states. *Cereb. Cortex* **22**, 221–231 (2012).
18. Keil, J., Müller, N., Hartmann, T. & Weisz, N. Prestimulus beta power and phase synchrony influence the sound-induced flash illusion. *Cereb. Cortex* **24**, 1278–1288 (2014).
19. Lange, J., Oostenveld, R. & Fries, P. Reduced Occipital Alpha Power Indexes Enhanced Excitability Rather than Improved Visual Perception. *J. Neurosci.* **33**, 3212–3220 (2013).
20. Cohen, M. X. & Cavanagh, J. F. Single-Trial Regression Elucidates the Role of Prefrontal Theta Oscillations in Response Conflict. *Front. Psychol.* **2** (2011).
21. Samaha, J., Iemi, L. & Postle, B. R. Prestimulus alpha-band power biases visual discrimination confidence, but not accuracy. *Conscious. Cogn.* **54**, 47–55 (2017).
22. Balz, J. *et al*. Beta/Gamma Oscillations and Event-Related Potentials Indicate Aberrant Multisensory Processing in Schizophrenia. *Front. Psychol.* **7** (2016).
23. Gulbinaite, R., İlhan, B. & VanRullen, R. The Triple-Flash Illusion Reveals a Driving Role of Alpha-Band Reverberations in Visual Perception. *J. Neurosci.* **37**, 7219–7230 (2017).
24. Ikumi, N., Torralba, M., Ruzzoli, M. & Soto-Faraco, S. The phase of pre-stimulus brain oscillations correlates with cross-modal synchrony perception. *Eur. J. Neurosci.* **49**, 150–164 (2019).
25. Lakatos, P., Chen, C.-M., O'Connell, M. N., Mills, A. & Schroeder, C. E. Neuronal Oscillations and Multisensory Interaction in Primary Auditory Cortex. *Neuron* **53**, 279–292 (2007).
26. Mercier, M. R. *et al*. Auditory-driven phase reset in visual cortex: Human electrocorticography reveals mechanisms of early multisensory integration. *NeuroImage* **79**, 19–29 (2013).
27. Rohe, T., Ehlis, A.-C. & Noppeney, U. The neural dynamics of hierarchical Bayesian causal inference in multisensory perception. *Nat. Commun*. (2019).
28. Bolognini, N., Rossetti, A., Casati, C., Mancini, F. & Vallar, G. Neuromodulation of multisensory perception: A tDCS study of the sound-induced flash illusion. *Neuropsychologia* **49**, 231–237 (2011).
29. Convento, S., Vallar, G., Galantini, C. & Bolognini, N. Neuromodulation of Early Multisensory Interactions in the Visual Cortex. *J. Cogn. Neurosci.* **25**, 685–696 (2012).
30. Rohe, T. & Noppeney, U. Cortical Hierarchies Perform Bayesian Causal Inference in Multisensory Perception. *Plos Biol* **13**, e1002073 (2015).
31. Falchier, A., Clavagnier, S., Barone, P. & Kennedy, H. Anatomical Evidence of Multimodal Integration in Primate Striate Cortex. *J. Neurosci.* **22**, 5749–5759 (2002).
32. Bastos, A. M. *et al*. Visual Areas Exert Feedforward and Feedback Influences through Distinct Frequency Channels. *Neuron* **85**, 390–401 (2015).
33. Michalareas, G. *et al*. Alpha-Beta and Gamma Rhythms Subserve Feedback and Feedforward Influences among Human Visual Cortical Areas. *Neuron* **89**, 384–397 (2016).
34. Richter, C. G., Thompson, W. H., Bosman, C. A. & Fries, P. Top-Down Beta Enhances Bottom-Up Gamma. *J. Neurosci.* **37**, 6698–6711 (2017).
35. Delorme, A. & Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* **134**, 9–21 (2004).
36. Oostenveld, R., Fries, P., Maris, E. & Schoffelen, J.-M. FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Comput. Intell. Neurosci*. **2011** (2010).
37. Lee, T.-W., Girolami, M. & Sejnowski, T. J. Independent Component Analysis Using an Extended Infomax Algorithm for Mixed Subgaussian and Supergaussian Sources. *Neural Comput.* **11**, 417–441 (1999).
38. Kiebel, S. J., Tallon-Baudry, C. & Friston, K. J. Parametric analysis of oscillatory activity as measured with EEG/MEG. *Hum. Brain Mapp.* **26**, 170–177 (2005).
39. Maris, E. & Oostenveld, R. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* **164**, 177–190 (2007).
40. Morey, R. D. & Rouder, J. N. *BayesFactor: Computation of Bayes Factors for Common Designs*. (2015).

Acknowledgements

We thank Yadira Roa Romero for help with data acquisition and James Moran for proofreading the manuscript. This work was supported by grants from the German Research Foundation (KE1828/2-1 and KE1828/4-1 to JK, SE1859/4-1 to DS, and BU2400/9-1 to NAB) and the European Union (ERC-2010-StG-20091209 to DS).

Author Contributions

M.K., analyzed data, drafted and revised manuscript; N.A.B., discussed methods and results, revised manuscript; J.B., acquired and analyzed data; D.S., designed research, revised manuscript; J.K., designed research, analyzed data, revised manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-019-42380-x.

Competing Interests: The authors declare no competing interests.

Study 3

Kaiser, M., Senkowski, D., Keil J. (2020). Mediofrontal Theta-band oscillations reflect top-down influence in the ventriloquist illusion. *Manuscript submitted for publication.*

Title: Mediofrontal Theta-band oscillations reflect top-down influence in the ventriloquist illusion.

Running title: Reduced Theta in ventriloquist illusion

Authors: Mathis Kaiser[1,2,*], Daniel Senkowski[1], Julian Keil[1,3]

Affiliations

1. Department of Psychiatry and Psychotherapy, Charité Universitätsmedizin – Berlin, Berlin, Germany
2. Berlin School of Mind and Brain, Humboldt Universität zu Berlin, Berlin, Germany
3. Biological Psychology, Christian-Albrechts-Universität zu Kiel, Kiel, Germany

* Present address: Science of Intelligence, Technische Universität Berlin, Marchstraße 23, 10587 Berlin, Germany

Acknowledgments

This work was supported by grants from the German Research Foundation (KE1828/2-1 and KE1828/4-1 to JK, SE1859/3-1 to DS). We thank Joseph Wooldridge for assistance with data collection and preprocessing, Prof. Felix Blankenburg for helpful comments on an earlier version of the manuscript, and Dr. James Moran for proofreading.

Conflict of interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Data availability

The data that support the findings of this study cannot be made available because participants did not consent to public dissemination. Processed data are available from the corresponding author upon reasonable request.

Abstract

In the ventriloquist illusion, spatially disparate visual signals can influence the perceived location of simultaneous sounds. Previous studies have shown asymmetrical responses in auditory cortical regions following perceived peripheral sound shifts. Moreover, higher-order cortical areas perform inferences on the sources of disparate audiovisual signals. Recent studies have also highlighted top-down influence in the ventriloquist illusion and postulated a governing function of neural oscillations for crossmodal processing. In this EEG study, we analyzed source-reconstructed neural oscillations to address the question of whether perceived sound shifts affect the laterality of auditory responses. Moreover, we investigated the modulation of neural oscillations related to the occurrence of the illusion more generally. With respect to the first question, we did not find evidence for significant changes in the laterality of auditory responses due to perceived sound shifts. However, we found a sustained reduction of mediofrontal theta-band power starting prior to stimulus onset when participants perceived the illusion compared to when they did not perceive the illusion. We suggest that this effect reflects a state of diminished cognitive control, leading to reliance on more salient visual information and increased crossmodal influence. We conclude that mediofrontal theta-band oscillations serve as a neural mechanism underlying top-down modulation of crossmodal processing in the ventriloquist illusion.

Keywords: oscillations, EEG, multisensory perception, theta

Introduction

The ability to integrate and segregate information reaching us via our different senses is a fundamental requirement for forming a coherent mental representation of our environment. Since these processes must operate dynamically, the neural architecture subserving them should also be flexible. Consequently, the brain activity patterns preceding and accompanying multisensory integration have come into focus in recent years, with a specific emphasis on the role of neural oscillations (Keil & Senkowski, 2018; van Atteveldt, Murray, Thut, & Schroeder, 2014). Of special interest in this context are experimental paradigms where crossmodal influence varies across single trials, because they allow researchers to investigate which neural conditions are associated with differences in perception while sensory input is constant. This is the case in the audiovisual ventriloquist illusion (VI) paradigm, where the location of visual stimuli affects the perceived location of concurrently presented sounds (Bertelson & Radeau, 1981; Bruns, 2019; Chen & Vroomen, 2013; Choe, Welch, Gilford, & Juola, 1975).

Along the auditory pathway, the superior olivary complex in the brainstem is the first structure that receives input from both ears and can use interaural time and intensity differences to encode sound location (Goldberg & Brown, 1969). At this processing stage, the auditory location is assumed to be coded in a head-centered reference frame and not amenable to the integration of visual information, which is coded in an eye-centered reference frame at early processing stages. However, in the subsequent processing stages in the midbrain (Groh, Trause, Underhill, Clark, & Inati, 2001; Jay & Sparks, 1984), primary auditory cortex (Werner-Reiss, Kelly, Trause, Underhill, & Groh, 2003), and parietal cortex (Mullette-Gillman, Cohen, & Groh, 2005), auditory responses are affected by eye position. Hence, the interaction of eye position and sound location results in a hybrid reference frame in higher-level areas, which is in line with a model of multisensory integration depending on multidirectional sensory predictions (Avillac, Denève, Olivier, Pouget, & Duhamel, 2005).

At the cortical level, the location of unisensory auditory stimuli is processed along a dorsal pathway, from caudal primary auditory cortex towards parietal areas (Rauschecker & Tian, 2000). Auditory localization, compared to pitch judgment, is associated with increased BOLD activation in posterior temporal and parietal areas (Alain, Arnott, Hevenor, Graham, & Grady, 2001). In a task-free fMRI paradigm, location changes of auditory stimuli elicited activation in the posterior planum temporale (Warren & Griffiths, 2003). Similar auditory regions have also been shown to be modulated by visual stimuli. Using high-resolution fMRI of the macaque monkey, Kayser, Petkov, Augath, and Logothetis (2007) showed that convergent audiovisual information activates specific fields in the caudal auditory cortex, extending into the upper bank of the superior temporal sulcus.

Evidence for a modulation of activity in auditory areas by visual information in the VI comes from an EEG-fMRI study by Bonath et al. (2007). The authors analyzed multimodal difference waves between audiovisual stimuli comprising a central auditory and a peripheral visual stimulus, and unisensory auditory plus unisensory visual stimuli. The negative ERP difference wave after 260ms was larger over the hemisphere contralateral vs. ipsilateral to the perceived peripheral shift of the sound. Using dipole modeling, the authors localized this effect in the Sylvian fissure. In a separate fMRI experiment, a corresponding decrease of illusion-related BOLD activity in the ipsilateral planum temporale was observed. The authors suggested that these effects are mediated by connections from visual areas over multimodal areas to auditory cortex. Further EEG studies have provided evidence for an early auditory processing account of the VI: the mismatch negativity, an early ERP component in response to infrequent (deviant) vs. frequent (standard) sounds with sources in auditory areas, is suppressed when sounds are visually shifted to standard positions (Colin, Radeau, Soquet, Dachy, & Deltenre, 2002), but evoked when they are shifted to deviant positions (Stekelenburg, Vroomen, & de Gelder, 2004). In summary, the auditory cortex likely processes the crossmodal shift of perceived sound location in the VI.

Building on evidence that the ventriloquist effect is based on a statistically optimal weighting of sensory information (Alais & Burr, 2004), recent studies have focused on the question how the brain infers the causal structure of multisensory input. Rohe and Noppeney (2015) showed that a hierarchy of cortical areas performs inferences regarding the sources of disparate audiovisual stimuli. Primary sensory areas represent location under the assumption of separate sources, while the posterior intraparietal sulcus (IPS) represents a common source and the forced fusion of input signals. Finally, the anterior IPS performs Bayesian inference, weighing the signals according to their reliability. The IPS has also been shown to exhibit increased functional connectivity with auditory areas following adaptation to spatially disparate audiovisual stimuli (Zierul, Röder, Tempelmann, Bruns, & Noesselt, 2017). Furthermore, a recent MEG study by Park and Kayser (Park & Kayser, 2019) has found that parietal areas encode both past and current sensory evidence in a ventriloquist paradigm. Taken together, these studies indicate a crucial role of parietal cortex in inferring the location of audiovisual stimuli.

While the study by Bonath et al. (2007) has shown an ERP asymmetry associated with peripheral sound shifts in the VI, it is as yet unknown if central shifts also result in reduced asymmetry. Furthermore, no study has investigated the relationship between perception in the VI and neural oscillations. Hence, proceeding from and extending the findings of Bonath et al. (2007), one aim in this study was to investigate the relationship between neural oscillations and visually induced sound location shifts towards the center or periphery. We examined hemispheric asymmetries depending on the perceived sound location and hypothesized that the symmetry of ERPs and oscillatory activity in auditory areas depends on the occurrence of illusory perception and the direction of the sound shift. Specifically, we expected an interaction effect of perception and direction on indices of laterality: responses in auditory areas should be lateralized for peripheral illusions and accurately perceived peripheral sounds, but not for central illusions and accurately perceived central sounds.

Furthermore, we investigated the modulation of neural oscillations related to crossmodal influence, irrespective of the direction of shift. Since perceptual priors (Rohe & Noppeney, 2015) in the VI may already develop before stimulus onset, and fluctuations of ongoing oscillations presumably contribute to variability in perception (Iemi et al., 2019; Keil, Müller, Hartmann, & Weisz, 2014; Keil, Müller, Ihssen, & Weisz, 2012; Weisz et al., 2014), we included the prestimulus period in this analysis. In agreement with the findings of Rohe and Noppeney (2015), we expected a modulation of oscillatory prestimulus activity or induced responses, especially in the IPS.

Methods

Participants

Thirty-five participants were recruited from the general population (mean age 30.3 ± 7.8 (SD) years, 17 male, 3 left-handed). All participants gave written informed consent and the study was conducted in accordance with the 2008 Declaration of Helsinki and approved by the ethics committee of the Charité–Universitätsmedizin Berlin. Participants reported no history of neurological or psychiatric disorders and were screened for hearing impairments using 500 and 750 Hz tones with an exclusion threshold of 25 dB.

Nine participants had to be excluded from the further data analysis. One participant was excluded due to technical problems during EEG data acquisition. Two further participants were excluded during preprocessing due to excessive muscular artifacts. Three additional participants were excluded due to low auditory accuracy. Subjects were excluded when auditory accuracy was lower than 50% in at least one unisensory condition during the main experiment and no discrimination thresholds could be determined from the response patterns in the unisensory auditory experiment (see below for descriptions of the tasks). Two further participants were excluded due to low visual accuracy: one reported not seeing peripheral visual stimuli during the main experiment, and one repeatedly closed their eyes during the experiment. Finally, one participant with an illusion rate > 90% was excluded because they relied almost exclusively on visual information in the auditory localization task. Thus, 26 participants were included in the analysis (mean age 29.9 ± 8.2 (SD) years, 12 male, 1 left-handed). Subsets of 15 and 18 participants were selected for two different EEG data analysis strategies based on trial counts in relevant stimulus and response categories (see below).

Experimental design

General procedure

Participants were seated in an electrically and acoustically shielded chamber. The experiment consisted of an auditory and visual steady state localizer, the main ventriloquist experiment, and unisensory auditory and visual control experiments. Data from the visual steady state localizer and unisensory visual control experiment were not used in the current work. Therefore, they are not further reported. The total experimental runtime, excluding breaks, was about 90 minutes.

Visual stimuli were presented at 45 cm viewing distance on an LCD display with a gray background (mean luminance: 30 cd/m^2) and a refresh rate of 75 Hz. Auditory stimuli consisted of a

600 Hz pure tone, sampled at 44.1 kHz, and were presented via earphones (Etymotic Research, Illinois, USA) at 72 dB SPL.

Auditory steady state localizer

Participants passively listened to tones with 40 Hz amplitude modulation at 90% modulation depth, on the left or right ear. We used unilateral stimuli to avoid strongly correlated activity between hemispheres with bilateral stimulation, which is difficult to localize using beamforming techniques, and to specifically stimulate space-sensitive areas. A trial consisted of a prestimulus period of 1 s, the auditory stimulus of 1.25 s and an inter-trial interval between 0.54 and 0.64 s. Throughout the trial, a central fixation cross was presented on the screen. Thirty-five trials were presented to each ear, in random order.

Ventriloquist experiment

In the main experiment, unisensory auditory and combined audiovisual stimuli were presented. Each trial consisted of a central fixation cross for 1 s, the auditory or audiovisual stimulus, a 0.6 poststimulus period, the response window, and an intertrial interval (ITI) randomly sampled between 0.22 and 0.42 s (for details, see Fig. 1). Participants were asked to indicate the perceived sound origin (left / center / right) on each trial with a button press using the index, middle, or ring finger of their right hand within a 1 s response interval. Before the start of data collection, participants completed a self-chosen number of training runs, where feedback about response timing was provided.

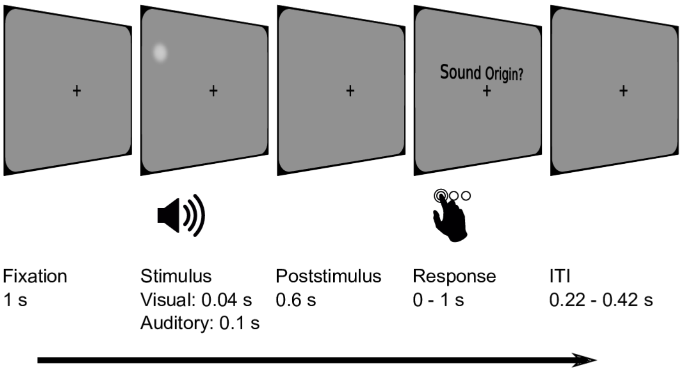

Figure 1: Timeline of one trial in the ventriloquist experiment

Auditory stimuli were presented for 0.1 s. The apparent origin of the sound was manipulated via the interaural time difference on three levels: -17.5° (A_L), 0° (A_C), 17.5° (A_R). Visual stimuli consisted of a light gray (75 % luminance) circular gaussian blob subtending 0.33° (at full width half maximum), presented for 0.04 s on a gray (50% luminance) background, 4° above the fixation cross and laterally displaced at either -17.5° (V_L), 0° (V_C), 17.5° (V_R) relative to fixation. Visual stimuli were presented above fixation to avoid proximity to the blind spot.

Auditory and visual stimuli were combined according to three categories: *ventriloquist* trials, where the visual location was adjacent to the auditory location ($A_C V_R$, $A_C V_L$, $A_R V_C$, $A_L V_C$), *congruent* trials, where the locations coincided ($A_L V_L$, $A_C V_C$, $A_R V_R$), and *divergent* trials, were the locations were on different sides relative to the central fixation ($A_L V_R$, $A_R V_L$). Two-hundred trials were presented per ventriloquist condition, 100 trials per congruent condition, and 50 trials per divergent condition. Furthermore, 120 *unisensory auditory* trials ($A_L V_0$, $A_C V_0$, $A_R V_0$) were presented per location, for a total of 1560 trials. These trial numbers were chosen to allow perception-based comparisons in the *ventriloquist* conditions, while avoiding inferences from visual on auditory location by the participants. Ventriloquist trials were categorized as *no-illusion* when auditory stimuli were localized correctly, or as *illusion* when auditory stimuli were perceived at the visual location. The order of the various stimulus conditions was pseudo-randomized across the length of the experiment. The experiment was split into 12 blocks of 130 trials each, with a self-paced break after each block. The total experimental runtime excluding breaks was approximately 75 minutes.

Unisensory auditory experiment

This behavioral experiment was conducted to assess the individual discrimination threshold by visually determining the approximate angles where the highest response rate transitioned from one direction to the next (e.g. from "center" to "right"). Auditory location was manipulated in 2.5° steps ranging between -17.5° and 17.5°. Ten trials per location were presented in random order. Trial timing, task and response mode were identical to the ventriloquist experiment.

Acquisition and Preprocessing of EEG data

Prior to the experiment, individual head fiducials, electrode positions, and headshape were digitized using a Polhemus Patriot (Polhemus, Vermont, USA). EEG was recorded using a 128-channel passive electrode cap (EasyCap, Herrsching, Germany), including one horizontal and one vertical electrooculography (EOG) electrode to monitor eye movements, and Brainamp DC amplifiers (Brainproducts, Gilching, Germany). Data were recorded in reference to an electrode placed on the nose with a sampling frequency of 1000 Hz and a pass band from 0.016 to 250 Hz. EEG data were processed and analyzed using the EEGlab (Delorme & Makeig, 2004; http://sccn.ucsd.edu/eeglab, RRID:SCR_007292) and Fieldtrip (Oostenveld, Fries, Maris, & Schoffelen, 2010; http://www.fieldtriptoolbox.org, RRID:SCR_004849) toolboxes for MATLAB (http://www.mathworks.com/products/matlab/, RRID: SCR_001622), and custom scripts. Parts of the statistical analyses were performed in R (R Core Team, 2013; http://www.r-project.org/, RRID:SCR_001905).

Raw EEG data were filtered using the default FIR filter settings in EEGlab, with a 0.5 Hz -6dB cutoff frequency and an order of 3300 for the high-pass, and -6dB cutoff frequencies of 49.5 and 50.5 Hz and an order of 3300 for the bandstop to filter out line noise. Data were resampled to 500 Hz and epoched into trials from -1.1 to 1.1 s around stimulus onset. Trials containing large artifacts and noisy channels were removed following visual inspection. After re-referencing to the common average, data were subjected to independent component analysis using an extended infomax algorithm (Lee, Girolami, & Sejnowski, 1999). Components representing blinks, lateral eye movement or cardiac artifacts were removed following visual inspection. Removed channels were interpolated using spherical spline interpolation and EOG channels were removed from the data. Trials still exceeding an absolute threshold of 100 mV after these procedures were removed automatically. On average, 6 ± 3 channels, 130 ± 85 trials and 4 ± 2.2 ICA components were removed from the individual datasets (mean ± SD).

Construction of forward models and source reconstruction

Realistic boundary element method (BEM) headmodels and lead fields were created from individual T1-weighted MRI scans, acquired on a 3T scanner (Siemens, Germany), and digitized electrode positions using OpenMEEG (Gramfort, Papadopoulo, Olivi, & Clerc, 2010; http://openmeeg.gforge.inria.fr/, RRID:SCR_001905). A template source grid with a resolution of 1 cm in MNI space was constructed, and individual grids were inverse-warped to the template positions for comparability across subjects. For one subject where no MRI scan was available, a template headmodel and standard electrode positions were used.

Virtual channel time courses in source space were reconstructed using an LCMV beamformer (Van Veen, Van Drongelen, Yuchtman, & Suzuki, 1997) with noise regularization of 5%. For each subject, *ventriloquist* trials were selected from the data and these data were analyzed using the LCMV beamformer. To this end, a spatial filter was constructed from the covariance matrix across the whole epochs, and the virtual channel time courses at each grid position were computed by multiplying EEG data with the spatial filter.

Analysis of behavioral data

To quantify the influence of visual information on auditory localization in the ventriloquist experiment, we compared the proportion of correct responses in the *ventriloquist*, *congruent,* and *unisensory auditory* trials using a 3x2 factorial repeated measures ANOVA with the factors Visual Stimulus (congruent; adjacent; none) and Auditory Location (peripheral; central). Response accuracies were averaged across visual and auditory locations according to these levels. For post-hoc tests, the estimated marginal means were contrasted where applicable using the emmeans package in R (Lenth, 2019), with Holm-correction for multiple comparisons.

To examine the possible influence of fatigue on the VI, we divided each participant's behavioral data into four quartiles along the length of the experiment, and calculated a repeated-measures ANOVA with the factor Quartile for the dependent variable illusion rate. This was done to dissociate potential perception-related effects in the EEG data from experimental runtime.

Analysis of EEG data

We pursued two complementary analysis strategies. The first focused on oscillatory correlates of crossmodal influence, irrespective of specific sound locations. The second focused on lateralization of activity in auditory cortex dependent on perceived sound location.

Crossmodal influence

For this line of analysis, we compared *illusion* and *no-illusion* trials pooled across the four *ventriloquist* conditions. Participants with at least 13 *illusion* and *no-illusion* trials in each *ventriloquist* condition were selected, resulting in a subset of 18 participants. Trial counts were equalized across conditions by random selection to ensure similar signal-to-noise ratios, and that potential effects were not confounded with real or perceived stimulus location. The number of 13 trials was chosen to reach a minimum combined count of 52 trials after pooling across the four conditions, separately for *illusion* and *no-illusion* trials. Analyses of neural oscillations were performed on both the scalp level and source level. Data were time-frequency transformed using multiple tapers with a window length of 5 cycles and spectral smoothing of 20% of the analyzed frequency, from -0.5 to 0.5 s in steps of 20 ms. The analyzed frequencies were logarithmically scaled between 2 and 70 Hz.

For the scalp level EEG data a cluster-based permutation dependent-samples t-test (Maris & Oostenveld, 2007) in the time range of -0.5 to 0.5 s peristimulus was used (*illusion* vs. *no-illusion,* cluster threshold $p = 0.01$, 1000 permutations). On the source level, the prestimulus period was initially analyzed separately, due to computational (RAM) limitations and because a scalp level effect was found in this period. The analysis was then extended to the poststimulus period. Absolute power changes between *illusion* and *no-illusion* trials, averaged over significant clusters from the permutation test, were then correlated with the illusion rates across subjects, and the Bayes Factor (BF), considering the correlation coefficient and sample size, was computed according to Wetzels and Wagenmakers (2012) to assess statistical evidence.

Finally, we also tested for differences in oscillatory power between illusion and no-illusion trials averaged over a region of interest (ROI) consisting of virtual channels within the inferior parietal gyrus (defined from the AAL-atlas). This ROI resembled the posterior parietal sulcus region, which has previously been shown to be associated with fusion of spatially divergent audiovisual signals (Rohe & Noppeney, 2015).

Lateralization

For the second line of analysis, we first analyzed data from the auditory localizer task. This was done to define regions of interest for the further statistical analysis. For the combined left and right ear steady state stimulation trials, the Fourier spectrum was computed for the time period from -1 to 1s peristimulus using a single Hanning taper. Data from -0.775 to -0.125 and 0.25 to 0.9 s peristimulus were selected as baseline and stimulation periods, respectively. In line with recent recommendations for the source analysis of auditory steady state responses by Popov et al. (2018), the sources of cortical responses phase-locked to a synthetic 40 Hz-signal were reconstructed, using a DICS-beamformer with noise regularization of 5% and a symmetric dipole pair as the source model. A cluster-based permutation dependent-samples t-test (cluster threshold $p = 0.01$, 1000 permutations) was used to compare stimulation and baseline periods. The location of maximal activation was identified based on the maximal t-value within the resulting significant cluster ($p < 0.05$). The MNI coordinates of the maximum were [60 -30 10], located in the right superior temporal gyrus (according to the AAL-atlas), adjacent to Heschl's gyrus (see Fig. 2). Since we used a symmetric dipole pair for the source reconstruction, the virtual channel showing the maximum and the homologue position in the left hemisphere were selected for the next analysis steps.

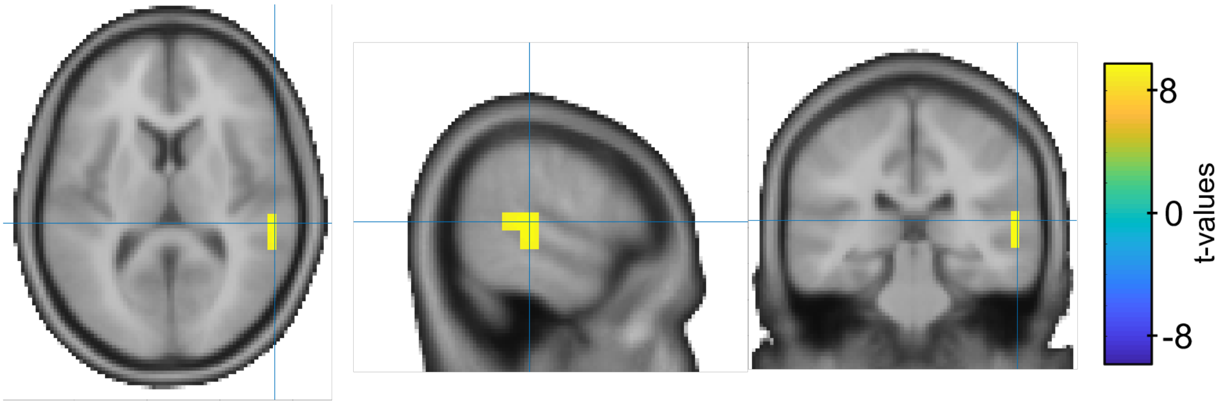

Figure 2: Source reconstruction of 40 Hz-coherence (stimulation vs. baseline) in the auditory localizer experiment. t-values are masked using 95% of the maximal value, which is indicated by the crosshairs. N = 15.

After defining ROIs for the further analysis, we combined trials from the ventriloquist experiment according to the location of the visual stimulus relative to the auditory stimulus. We will refer to A_CV_R and A_CV_L trials as *peripheral* ventriloquist trials, because a central sound is perceived peripherally in case of the illusion. Accordingly, we will refer to A_RV_C and A_LV_C trials as *central* ventriloquist trials. Participants with at least 50 *illusion* and *no-illusion* trials in both the central and peripheral conditions were selected for further analysis. This resulted in a subset of 15 participants (14 of which had also been selected for the crossmodal influence analysis). To avoid differences in signal-to-noise ratio, trial counts were equalized between the four conditions by random selection. The analysis focused on auditory regions of interest consisting of the symmetric virtual channels at the positions identified in the auditory localizer experiment, plus their respective five immediate grid neighbors. One additional neighbor was located outside the brain volume and was not included in the region of interest. Data were time-frequency transformed using the same parameters as in the crossmodal influence analysis. Additionally, ERPs with baseline correction from -0.2 to 0 s peristimulus and a 30 Hz low-pass FIR filter were computed for the same time window as in the time-frequency analysis (-0.5 to 0.5 s). ERPs were included in this line of analysis because previous research had demonstrated illusion-related ERP asymmetries (Bonath et al., 2007). Time-frequency (TFR) data and ERPs were averaged separately over the virtual channels ipsi- and contralateral relative to auditory location for the *central* condition, and ipsi- and contralateral relative to visual location for the *peripheral* condition. Next, the TFR laterality index (*ipsilateral* - *contralateral* / *ipsilateral* + *contralateral*, see Haegens et al. (2011)), and ERP difference waves (*contralateral* – *ispilateral*) were computed. A double difference approach was used to test for interaction effects of the direction of perceptual shift (*central/*

peripheral) and perception (*illusion/no- illusion)*. Specifically, activity in the *peripheral* condition was first subtracted from activity in the *central* condition (for *illusion* and *no-illusion* trials separately). Then, these differences were submitted to a cluster-based permutation dependent-samples t-test (*illusion* vs. *no-illusion,* cluster threshold $p = 0.05$, 1000 permutations). To directly examine whether there were simple perception-related differences, especially within the *peripheral* condition as described by Bonath et al. (2007), *illusion* against *no-illusion* was also tested in the *central* and *peripheral* conditions separately.

For the ERP analysis we also computed a 3-factorial ANOVA (with an additional factor Hemisphere instead of forming a difference wave) to complement the difference wave analysis and to specifically test for lower-level interactions and main effects of Hemisphere, Direction and Perception. Peak latencies of the components in a +/-20 ms window around 50, 100, and 200 ms were first extracted from the average across all conditions. Then, amplitudes were averaged over a +/-10 ms window around the identified peak. To directly test for asymmetrical evoked responses in the time window identified by Bonath et al. (2007), we also included the +/-20ms average around 250 ms in the analysis. Averaged peak amplitudes were subjected to a repeated-measures ANOVA with Holm correction for four latencies. For post-hoc tests, the estimated marginal means were contrasted where applicable using the emmeans package in R (Lenth, 2019), with Holm-correction for multiple comparisons. We did not compute a corresponding ANOVA for the TFR data because this would require a selection of time-frequency regions of interest, which is not as straightforward as in the case of ERPs.

Results

Behavior

Mean accuracy in the unisensory auditory conditions of the main experiment was 74.7 ± 9.7 (SD) %. The illusion rate across *ventriloquist* conditions was 55.8 ± 23.3 (SD) %. Responses in the unisensory auditory control experiment indicated that subjects could reliably discriminate sound locations well below the angle used in the main experiment. Fig. 3 illustrates mean response rates in all conditions of the main experiment and Fig. 4 illustrates mean response rates for the unisensory auditory control experiment. A 3x2 repeated measures ANOVA of response accuracy with factors Visual Stimulus and Auditory Location revealed a main effect of Visual Stimulus ($F(2,50) = 98.93$, $p < 0.001$) and an interaction between Visual Stimulus and Auditory Location ($F(2,50) = 3.41$, $p = 0.0408$, see Fig. 5). The main effect of Auditory Location was not significant ($F(1,25) = 3.17$, $p = 0.087$). Post-hoc tests for the main effect of Visual Stimulus revealed that accuracy was higher for congruent compared to no Visual Stimulus ($t = 4.65$, $p < 0.0001$), higher for congruent compared to adjacent Visual Stimulus ($t = 13.821$, $p < 0.0001$), and higher for no compared to adjacent Visual Stimulus ($t = 9.17$, $p < 0.0001$). This indicates that visual stimuli influenced auditory perception and induced the VI. Post-hoc tests for the interaction between Visual Stimulus and Auditory Location revealed that accuracy was higher for peripheral compared to central Auditory Location when no Visual Stimulus was presented ($t = 2.75$, $p = 0.0089$), but not when the Visual Stimulus was congruent or adjacent to the auditory location (both $p > 0.2$). Simple contrasts of Visual Stimulus within the levels of Auditory Location mirrored those of the main effect (all $p < 0.01$).

Next, we analyzed the influence of experimental runtime on illusion rate. The factor Quartile had a significant influence on the illusion rate ($F(3,25) = 5.23$, $p = 0.0025$). Post-hoc tests revealed that illusion rates were lower in the second, third and fourth quartile compared to the first quartile ($t = 2.84$, $p = 0.0292$; $t = 3.04$, $p = 0.0169$; and $t = 3.62$, $p = 0.0029$, respectively). This decrease of illusion rates only from the first to the subsequent quartiles suggests that fatigue did not substantially influence perception of the VI. Otherwise, a continuous increase or decrease would have been expected. The reduction in illusion rates may rather reflect an initial training effect.

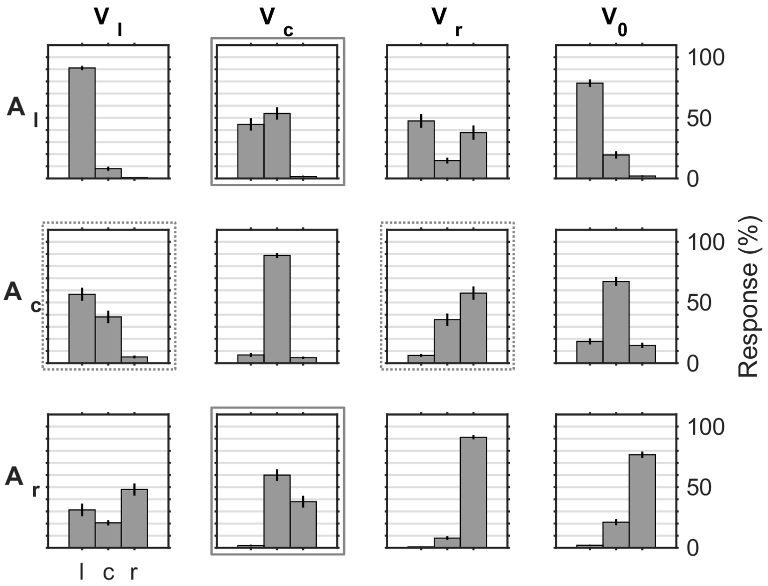

Figure 3: mean response rates (l = left, c = center, r = right) for all conditions. Error bars indicate SEM. Solid and dashed gray lines around the plots indicate the central and peripheral ventriloquist conditions, respectively.

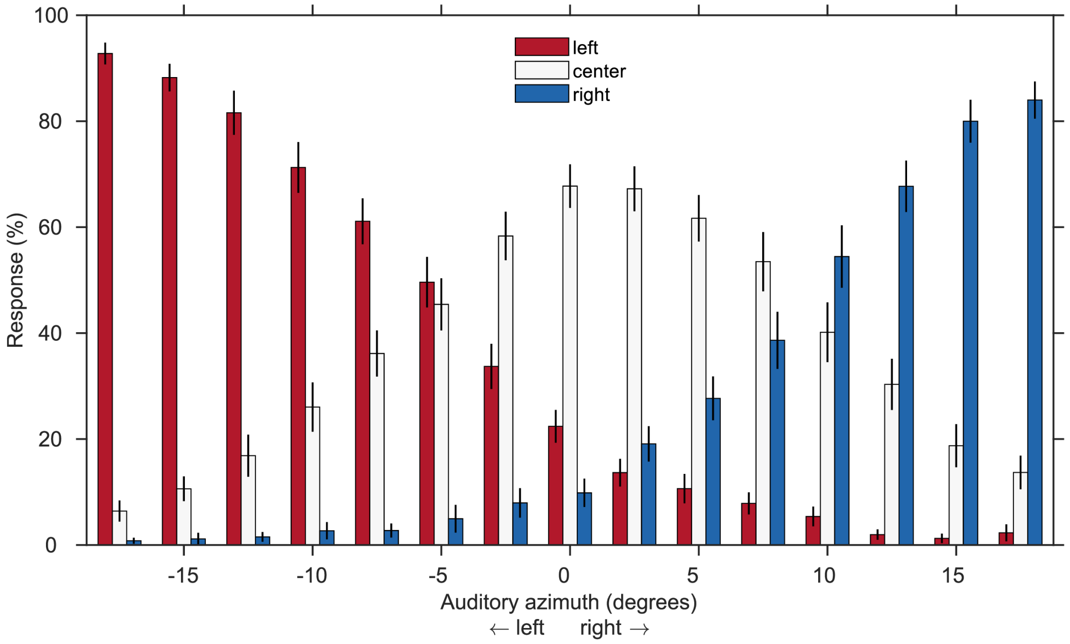

Figure 4: mean response rates in the unisensory auditory experiment. Error bars indicate SEM. This figure demonstrates that participants could discriminate central and peripheral sounds well below the angle used in the main experiment (17.5°).

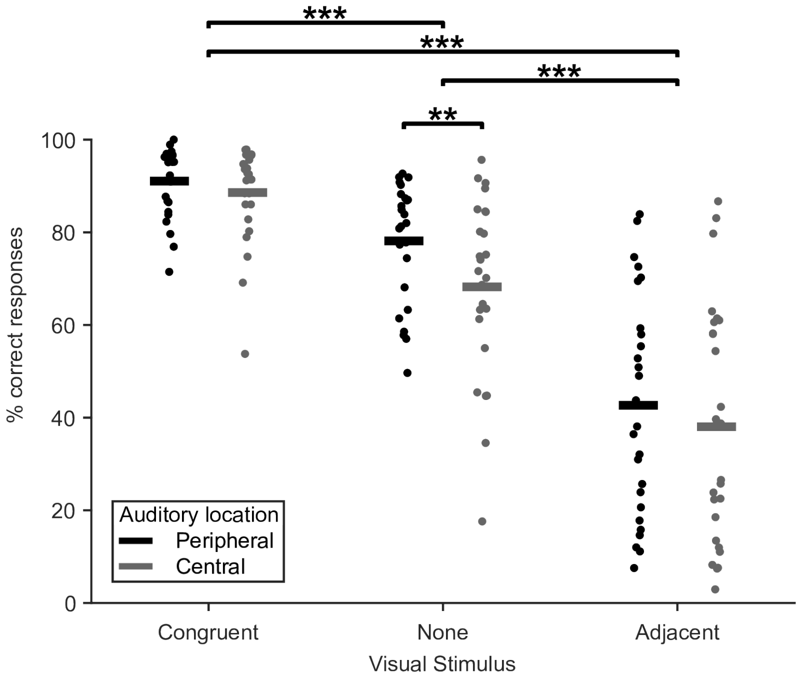

*Figure 5: Mean and individual response accuracies, indicated by bars and dots, respectively. Asterisks indicate significant post-hoc tests for the 3x2 ANOVA (** $p < 0.01$; *** $p < 0.001$). Please note that the simple effects of Visual Stimulus, which mirrored the main effect, are not depicted.*

Crossmodal influence

We compared illusion with no-illusion trials pooled across different stimulus locations. This was done to test for oscillatory power modulations related to crossmodal influence. For the scalp level analysis, a significant negative electrode cluster was found over mediocentral channels in the prestimulus period (p = 0.038, *illusion* < *no-illusion*; see Fig. 6). The cluster ranged from -0.5 to -0.12 s, between 4.2 and 4.9 Hz, i.e. the theta band. The across-subject correlation between illusion rates and *illusion – no-illusion* power differences in the cluster was not significant (R = -0.35, p = 0.1553). The corresponding BF was 0.49.

For the source space analysis, a significant negative cluster was found at medial frontal regions in the prestimulus period (p = 0.05, *illusion* < *no-illusion*). The cluster encompassed a -0.48 to 0 s time interval at a frequency of 2.7 to 4.4 Hz. Hence, the scalp level analysis, as well as the source level analysis revealed differential prestimulus low frequency power in the delta to theta range between illusion and no-illusion trials. Because the source level cluster extended to the time of stimulus onset, we also analyzed the poststimulus period subsequently. In the poststimulus period, a comparable cluster (p = 0.014) ranging from 0 to 0.5 s, between 2.7 and 4.4 Hz was obtained (see Fig. 7). The across-subject correlations between illusion rates and *illusion - no-illusion* power differences were not significant on the source level (prestimulus: R = -0.27, p = 0.284; poststimulus: R = -0.44, p = 0.0668). The BFs were 0.33 and 0.95, respectively.

For the region of interest analysis in the inferior parietal gyrus, no significant effect was found. We also repeated the control analysis on behavioral data for fatigue in this participant subsample and did not find a significant influence of quartile (F(3,17) = 2.72, p = 0.0537), though there was a trend towards significance. In summary, both the scalp and source level analyses revealed an influence of medial frontal theta band power prior to stimulus onset on the perception of the VI.

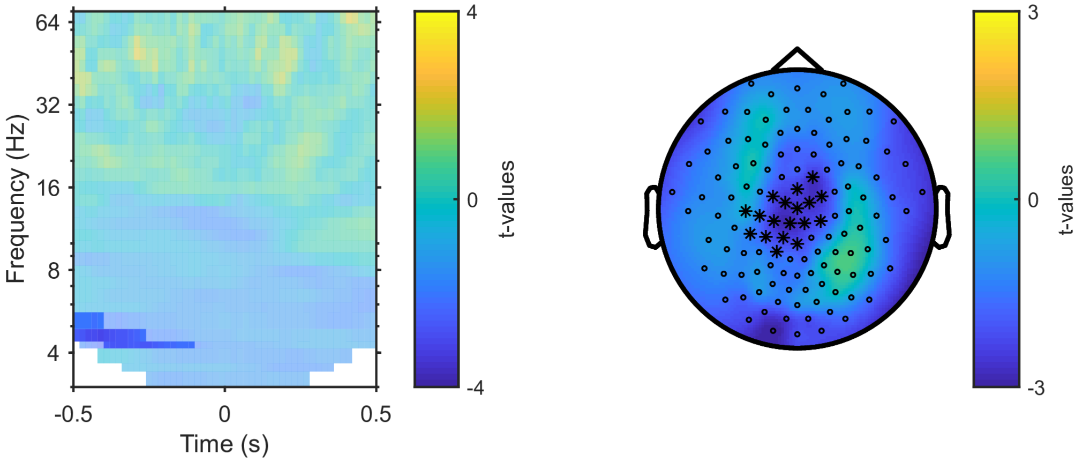

Figure 6: Time-frequency spectrum and topography of t-values for the illusion / no illusion comparison on the sensor level. Significant regions / channels are indicated by saturation / asterisks, respectively.

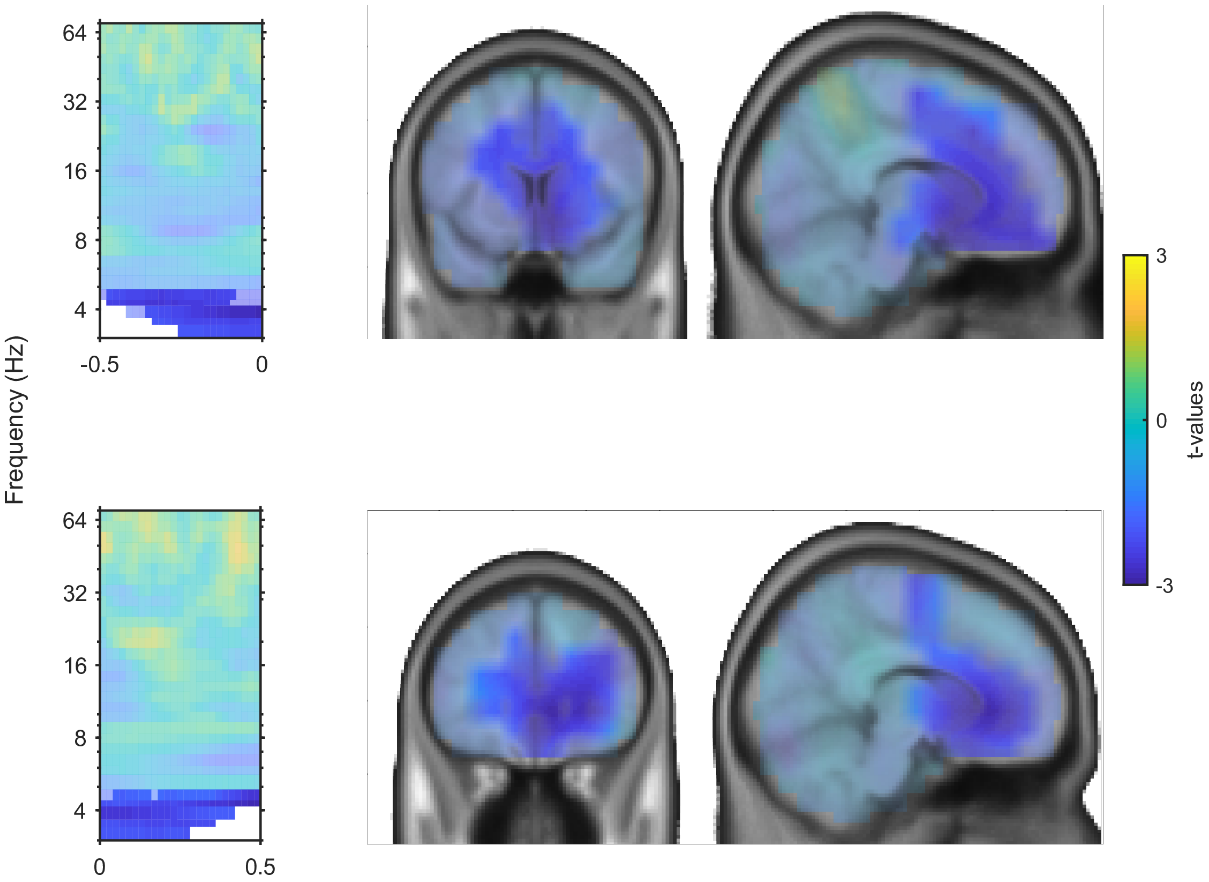

Figure 7: Time-frequency spectrum and sourceplot of t-values for the illusion vs. no illusion comparison on the source level. Significant regions / virtual channels are indicated by saturation. The prestimulus period is shown at the top, the poststimulus period at the bottom.

Lateralization

In this line of analysis, we analyzed the symmetry of auditory cortical responses related to the occurrence of sound shifts toward the center or the periphery. Spectra for baseline-corrected activity in the four conditions and the laterality index, as well as the tested differences are illustrated in Fig. 8. The analysis of the TFR laterality index revealed no significant effects for the comparison between the *central* minus *peripheral* condition differences within the *illusion* and *no-illusion* conditions (lowest negative cluster $p = 0.48$). The analysis of the ERP difference waves did also not reveal any significant effects (lowest negative cluster $p = 0.67$). ERPs and difference waves are illustrated in Fig. 9.

When testing simple effects of *illusion* vs. *no-illusion*, for the *central* and *peripheral* conditions separately, again no significant differences were obtained. In the TFR data, the lowest negative cluster p-value for the comparison in the *central* condition was $p = 0.52$; in the *peripheral* condition it was $p = 0.42$. In the ERPs, the lowest positive cluster p-value for comparison in the *central* condition was $p = 0.57$. No clusters were found in the *peripheral* condition.

In the ANOVAs of averaged ERP amplitudes, no significant 3-way interaction effects were found, contrary to the hypothesis that ERP amplitudes should depend on an interaction of hemisphere, the direction of perceptual shift, and the occurrence of the illusion. However, a significant main effect of Hemisphere was found around 102 ms (F(1,14) = 12.54, p= 0.013). Post hoc tests revealed that amplitudes around 102 ms were larger in the contralateral compared to the ipsilateral hemisphere (t = 3.54, p = 0.0033). Moreover, a Direction x Hemisphere interaction effect was observed around 204 ms (F(1,14) = 15.16, p = 0.0065). Post hoc tests revealed that amplitudes were larger in the contralateral compared to the ipsilateral hemisphere when visual stimuli were presented peripherally (t = 4.44, p = 0.0005). No such effect was found when visual stimuli were presented centrally (t = 1.25, p = 0.6015). Moreover, ERP amplitudes around 204 ms were larger for centrally presented visual stimuli, compared to peripherally presented visual stimuli within the ipsilateral hemisphere (t = 4.33, p = 0.007). Thus, ERP amplitudes around 204 ms differed between hemispheres for peripherally, but not for centrally presented visual stimuli. Taken together, the current results do not support the notion that lateralized cortical activity reflects subjective sound location in the VI.

Factor	**DF**	**36-56 ms**		**92-112 ms**		**184-204 ms**		**230-270 ms**	
		F-value	**p-value**	**F-value**	**p-value**	**F-value**	**p-value**	**F-value**	**p-value**
Direction	1,14	0.92	0.3727	6.17	0.0788	7.24	0.0703	1.93	0.3727
Perception	1,14	1.41	0.5108	0.96	0.5108	2.14	0.4965	6.5	0.0924
Hemisphere	1,14	0.29	0.6	12.54	0.013*	5.46	0.1046	3.72	0.1488
Direction x Perception	1,14	0.05	1	0.03	1	0.12	1	0	1
Direction x Hemisphere	1,14	0.78	0.394	6.14	0.0796	15.16	0.0065*	1.84	0.394
Perception x Hemisphere	1,14	0.16	1	0	1	1.8	0.8046	0.08	1
Direction x Perception x Hemisphere	1,14	0.07	1	0.45	1	0.02	1	0.52	1

Table 1: F- and p-values for the 3-factor ANOVA of ERP amplitudes. The reported p-values are Holm-corrected for comparisons at four latencies.

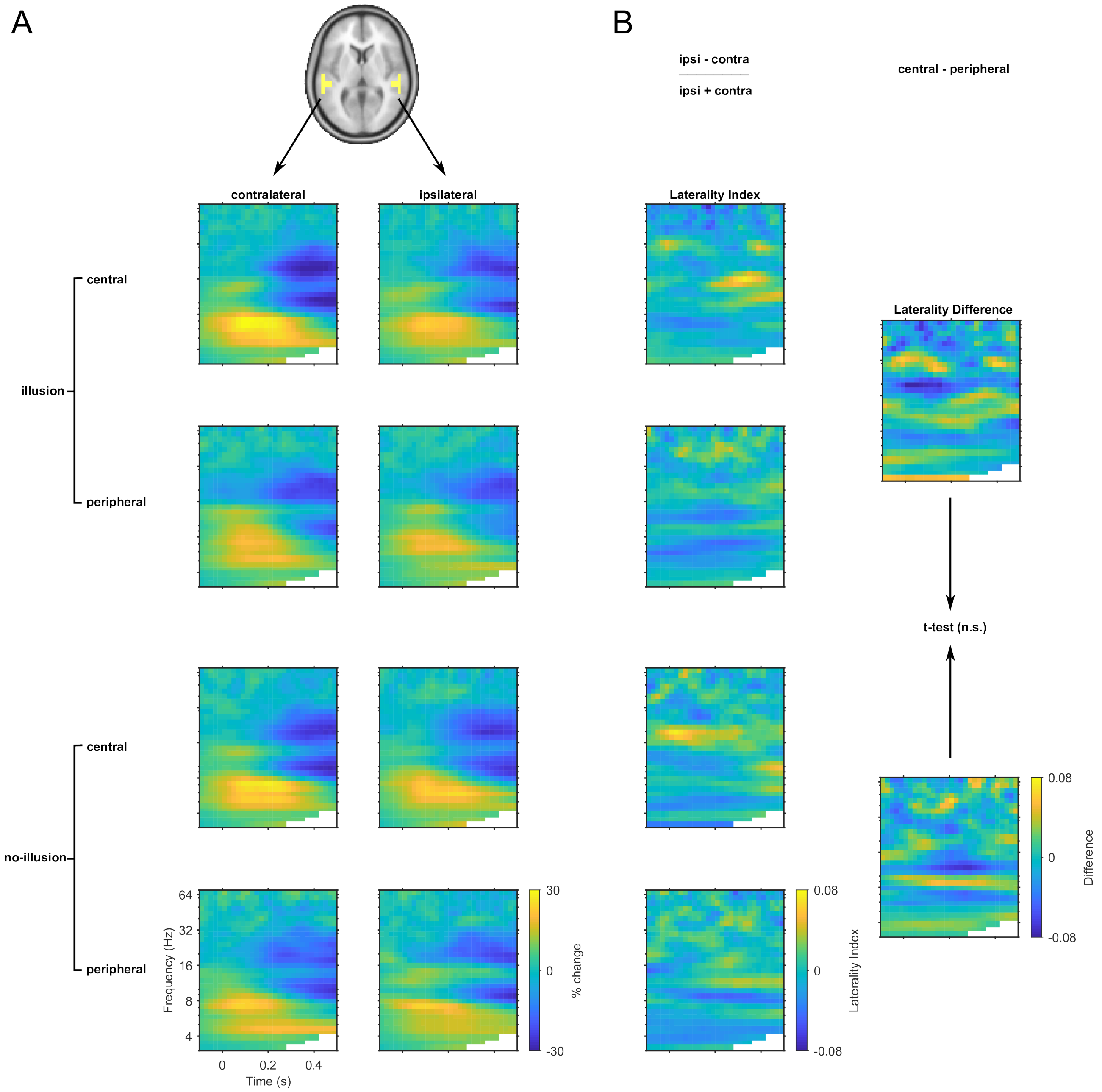

Figure 8: A: Time-frequency spectra in the contra- and ipsilateral source regions of interest for all conditions. Data are baseline-corrected using the 500 ms prestimulus period. Note that this correction was only applied for illustrative purposes, but not in the analyzed data, where a similar correction is implicit in the calculation of the laterality index.

B: Time-frequency spectra of the laterality index and tested differences.

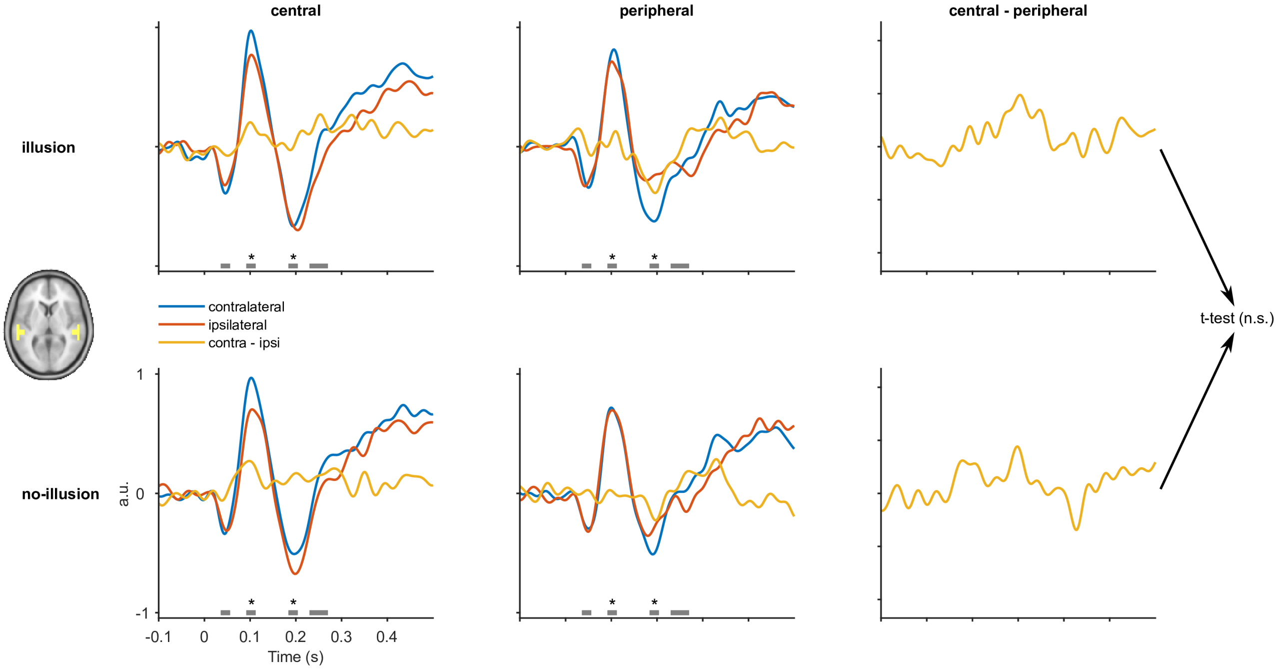

Figure 9: Event-related potentials and difference waves in the source regions of interest. Time periods used for the ANOVA of averaged amplitudes are indicated by gray lines above the abscissa, and periods where main or interaction effects were found are indicated by asterisks. The figure illustrates that amplitudes around 100 ms are generally larger in the contralateral vs. ipsilateral hemisphere, while amplitudes around 200 ms show this difference only for peripheral visual stimuli.

Discussion

In this study, we investigated neural correlates of the ventriloquist illusion. We focused on examining oscillatory activity related to crossmodal influence, and asymmetrical activity related to perceived sound location. Our study revealed that a decrease of slow wave prestimulus oscillations in mediofrontal areas facilitated the illusion.

Behavior

The response patterns indicated that participants could accurately localize unisensory auditory stimuli. Since the illusion rates in the ventriloquist conditions were markedly higher than the unisensory error rates, there is strong evidence that visual stimuli shifted the perceived sound location, in line with previous studies (Bruns, 2019). Our analysis of response accuracy showed that accuracy was higher for congruent audiovisual trials compared to unisensory auditory trials and ventriloquist trials, and higher for unisensory auditory trials compared to ventriloquist trials. This confirms that visual stimuli biased the perceived sound localization. In addition, we found that accuracy was higher for peripheral compared to central auditory location, but only when sounds were presented alone. This suggests that the advantage for peripheral sounds was superseded by visual influence.

An analysis of illusion rates across the duration of the experiment showed that illusion rates were higher in the first quartile compared to the following three. This suggests that participants got better at discriminating the sounds after initial practice, but contradicts the idea that fatigue had a strong impact on illusion rates. Taken together, our study replicated prior observations that visual stimuli can affect the perceived location of sounds.

Cross-modal influence

In this line of analysis, we investigated modulations of oscillatory activity related to the occurrence of the VI, irrespective of the direction of perceptual shift. The analysis was performed on the level of the scalp, the whole brain source level, and with a focus on inferior parietal sources.

In our analysis of scalp-level activity, we found a significant prestimulus modulation of frontal theta band power: illusory perception was associated with decreased theta power. This finding could reflect a state of diminished cognitive control (Cavanagh & Frank, 2014) that leads to a reliance on more salient visual information. The topography of the effect was consistent with frontal sources, a notion that was further supported by the source-level analysis. On the source level, we found a sustained decrease of theta power in the prestimulus and poststimulus periods in mediofrontal regions associated with the illusion. Theta band oscillations have been implicated in the monitoring of response conflict (Cohen & Cavanagh, 2011) and are well suited for long-range information transfer across cortical regions (von Stein & Sarnthein, 2000). Therefore, they may serve as a neural mechanism for perceptual adjustment and action selection in multisensory tasks, where information might be disparate and has to be integrated across different sensory regions. Thus, the frontal theta modulation could reflect activity of populations representing action goals, ultimately biasing sensory circuits involved in response selection (E. K. Miller, 2000). In line with this assumption, theta band functional connectivity has been suggested to signal changing task demands and dynamically modulates the integration of cortical areas into distributed networks (Keil, Pomper, & Senkowski, 2016). Similarly, Rohe and Noppeney (2018) argued for a task-dependent modulation of functional networks by frontal areas in audiovisual perception. Finally, our finding of reduced frontal theta oscillations in the VI is also consistent with recent evidence that the VI is susceptible to top-down influence and is not a purely perceptual phenomenon. For instance, it has been shown that reward expectations (Bruns, Maiworm, & Röder, 2014) and emotional valence (Maiworm, Bellantoni, Spence, & Röder, 2012) modulate the VI. The VI can also be induced by imagined visual stimuli (Berger & Ehrsson, 2013, 2014), further supporting the notion that it results not solely from bottom-up processing. In line with these recent findings, early evidence for a contribution of response bias as opposed to perceptual changes to the ventriloquist effect came from the study by Choe et al. (1975), where the authors argued for an influence of shifts of the decision criteria.

We did not find a substantial influence of experimental runtime on illusion rates in the participants from this analysis. This suggests that the effects in the low-frequency range are not due to fatigue increasing with experimental duration, but unfold on a shorter time scale. Taken together with the observation of temporally sustained decrease of theta power, this indicates that the variability of multisensory integration in the VI is due to modulations of cognitive control that span several trials, but not longer time periods. Furthermore, illusion rates were not significantly correlated with theta power changes across subjects. Therefore, the observed modulation appears not to relate to interindividual differences in the tendency to perceive the VI, but to perceptual variability within each individual.

Contrary to our hypotheses and the conclusions of Rohe & Noppeney (2015), we found no illusion-related modulation of prestimulus activity or induced responses in the inferior parietal region of interest. Possible reasons for the lack of an effect include the reduced spatial resolution of source-level EEG compared to fMRI or a location difference between the analyzed regions across studies. However, we also found no corresponding effect in nearby regions in the whole-brain analysis. Furthermore, it is unclear whether differences in oscillatory activity between perceptual conditions have enough sensitivity to the perceptual prior for a common source, which was computed from behavioral data and then correlated with the BOLD signal in the original study by Rohe and Noppeney (2015). In summary, our results did not corroborate an involvement of parietal cortex in the ventriloquist illusion.

Lateralization

This analysis focused on hemispheric asymmetries in auditory areas, depending on perceptual shifts towards the center or periphery. Contrary to our hypotheses and the results of Bonath et al. (2007), we found no significant effects on ERP amplitudes or neural oscillations in the laterality analyses. Interhemispheric balance in the auditory region of interest did not reflect the perceived auditory stimulus location. This conclusion is based on the lack of an interaction between Direction and Perception in the analysis of the TFR laterality index and the ERP difference waves. We also tested a simple contrast of *illusion* vs. *no-illusion* within the *peripheral* condition, thereby trying to replicate the finding of Bonath et al. (2007) more directly. However, this analysis also resulted in no significant effects.

One possible reason for the lack of perception-related effects is that, in contrast to Bonath et al. (2007), we did not analyze multimodal differences, which might be more sensitive to modulation by crossmodal integration. However, if the effect described by Bonath et al. is due to an integrative process, we consider it plausible that it should also be detected when analyzing responses to multimodal stimuli directly. An ERP study using an audiotactile ventriloquist paradigm (Bruns & Röder, 2010) found enlarged central ERPs in a similar time range as Bonath et al. (2007) for central compared to lateral sound perception, irrespective of the physical sound location. Importantly, the ERP asymmetry effects described by Bonath et al. (2007) could not be replicated in that study either. Another possible reason for the lack of laterality effects is that more realistic stimuli including spectral cues might be necessary to drive salient responses in auditory cortex (Callan, Callan, & Ando, 2013). In the same vein, hemispheric asymmetries are more consistently found for monoaural compared to lateralized binaural stimuli (Woldorff et al., 1999). However, we found prominent ERPs in our auditory region of interest, including some components that showed hemispheric dominance, using simple sounds with temporal location cues. On a more fundamental level, there is evidence from primate studies that acoustic space is represented by population codes which might not be easily resolvable using EEG (L. M. Miller &

Recanzone, 2009). Hence, it remains an open question whether the physical or subjective location of auditory stimuli is reflected in the EEG. Lastly, it is possible that our region of interest did not include relevant neural loci to capture the effect. However, the selected region of interest should be considered suitable because it showed the largest modulation in response to lateralized sounds in the localizer experiment.

Whereas we did not find perception-related changes in auditory areas, we found stimulus-related modulation of ERPs in the ANOVA of peak amplitudes. In the N1 range, evoked potentials were enhanced in the contralateral relative to the ipsilateral auditory cortex, demonstrating a well-known contralateral dominance of the auditory system (Pantev, Ross, Berg, Elbert, & Rockstroh, 1998; Picton et al., 1999). In the P2 range, this enhancement was only observed for peripheral visual stimuli. This indicates that the location of visual stimuli had an influence on auditory processing. However, the lack of interactions with perception indicates that the altered auditory processing had no direct impact on perception. In summary, whereas we could not replicate the ERP asymmetry effects described by Bonath et al. (2007), we found evidence for a modulation of auditory cortical activity by spatial visual information. This modulation, however, was not related to the VI.

Limitations and future directions

The current analysis has a number of limitations. Firstly, the exact temporal localization of the effect is difficult due to the low temporal precision in the low-frequency range. We found modulations of prestimulus theta-band activity over a time span of 400 ms in the sensor-level analysis, which spread across the whole trial in the source-level analysis. Due to the width of the sliding temporal window in the time-frequency analysis, activity from minus to plus 500 ms around a given time point is included in the spectral estimate at 5 Hz. This hampers strong conclusions on the temporal dynamics of the effect. Interesting evidence regarding the time course of causal inference in the VI comes from a recent study by Aller & Noppeney (2019), who found that the brain estimates auditory and visual signal location under the prior of forced fusion in the time range of 100 to 250 ms after stimulus onset.

Although there is evidence for left-lateralized processing of ITD cues (Tardif, Murray, Meylan, Spierer, & Clarke, 2006), low trial numbers did not allow us to analyze all conditions separately. Instead, we pooled contra- and ipsilateral electrodes across conditions for the laterality analyses. Therefore, we could not make inferences about hemispheric asymmetries or the processes underlying shifts in specific directions. However, averaging across several conditions was required to uncover more general correlates of the crossmodal influence of visual signals on auditory spatial perception, such as the reduction in theta-band power that we observed. We also did not find a significant correlation between illusion rates and power differences in the analyzed clusters, possibly due to the low statistical power with a sample size of 18 participants. The Bayes Factors between 0.33 and 0.95 for the correlation values and sample size indicated no substantial evidence for or against the null hypothesis, therefore the correlation between theta power and illusion rates should be reexamined in a larger sample.

Conclusion

Taken together, our findings indicate that modulations of theta band power in frontal cortical areas influence the perception of the VI. In contrast to a previous study (Rohe & Noppeney, 2015), we did not find a representation of a perceptual prior for forced fusion of audiovisual signals in parietal cortex. Overall, our analyses of laterality and cross-modal influences as mechanisms underlying the VI support earlier notions of top-down influences and shifts of decision criteria, rather than a modulation of cortical activity in primary sensory or parietal areas.

Our study shows that reduced pre- and poststimulus theta power in mediofrontal regions is associated with the perception of the VI. This suggests that diminished top-down control over the demanding auditory localization task leads to stronger crossmodal influence and hence, a stronger VI. We could not corroborate earlier results of a relationship between perceived auditory location and interhemispheric balance. Instead, our findings support the notion that frontal theta-band oscillations serve as a neural mechanism underlying top-down control of crossmodal influence in the ventriloquist illusion.

References

Alain, C., Arnott, S. R., Hevenor, S., Graham, S., & Grady, C. L. (2001). "What" and "where" in the human auditory system. *Proceedings of the National Academy of Sciences*, *98*(21), 12301–12306.

Alais, D., & Burr, D. (2004). The Ventriloquist Effect Results from Near-Optimal Bimodal Integration. *Current Biology*, *14*(3), 257–262. https://doi.org/10.1016/j.cub.2004.01.029

Aller, M., & Noppeney, U. (2019). To integrate or not to integrate: Temporal dynamics of hierarchical Bayesian causal inference. *PLOS Biology*, *17*(4), e3000210. https://doi.org/10.1371/journal.pbio.3000210

Avillac, M., Denève, S., Olivier, E., Pouget, A., & Duhamel, J.-R. (2005). Reference frames for representing visual and tactile locations in parietal cortex. *Nature Neuroscience*, *8*(7), 941–949. https://doi.org/10.1038/nn1480

Berger, C. C., & Ehrsson, H. H. (2013). Mental Imagery Changes Multisensory Perception. *Current Biology*, *23*(14), 1367–1372. https://doi.org/10.1016/j.cub.2013.06.012

Berger, C. C., & Ehrsson, H. H. (2014). The Fusion of Mental Imagery and Sensation in the Temporal Association Cortex. *Journal of Neuroscience*, *34*(41), 13684–13692. https://doi.org/10.1523/JNEUROSCI.0943-14.2014

Bertelson, P., & Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Perception & Psychophysics*, *29*(6), 578–584.

Bonath, B., Noesselt, T., Martinez, A., Mishra, J., Schwiecker, K., Heinze, H.-J., & Hillyard, S. A. (2007). Neural Basis of the Ventriloquist Illusion. *Current Biology*, *17*(19), 1697–1703. https://doi.org/10.1016/j.cub.2007.08.050

Bruns, P. (2019). The Ventriloquist Illusion as a Tool to Study Multisensory Processing: An Update. *Frontiers in Integrative Neuroscience*, *13*. https://doi.org/10.3389/fnint.2019.00051

Bruns, P., Maiworm, M., & Röder, B. (2014). Reward expectation influences audiovisual spatial integration. *Attention, Perception, & Psychophysics*, *76*(6), 1815–1827. https://doi.org/10.3758/s13414-014-0699-y

Bruns, P., & Röder, B. (2010). Tactile capture of auditory localization: an event-related potential study. *European Journal of Neuroscience*, *31*(10), 1844–1857. https://doi.org/10.1111/j.1460-9568.2010.07232.x

Callan, A., Callan, D. E., & Ando, H. (2013). Neural correlates of sound externalization. *NeuroImage*, *66*, 22–27. https://doi.org/10.1016/j.neuroimage.2012.10.057

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

Chen, L., & Vroomen, J. (2013). Intersensory binding across space and time: A tutorial review. *Attention, Perception, & Psychophysics*, *75*(5), 790–811. https://doi.org/10.3758/s13414-013-0475-4

Choe, C. S., Welch, R. B., Gilford, R. M., & Juola, J. F. (1975). The "ventriloquist effect": Visual dominance or response bias? *Perception & Psychophysics*, *18*(1), 55–60.

Cohen, M. X., & Cavanagh, J. F. (2011). Single-Trial Regression Elucidates the Role of Prefrontal Theta Oscillations in Response Conflict. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00030

Colin, C., Radeau, M., Soquet, A., Dachy, B., & Deltenre, P. (2002). Electrophysiology of spatial scene analysis: the mismatch negativity (MMN) is sensitive to the ventriloquism illusion. *Clinical Neurophysiology*, *113*(4), 507–518. https://doi.org/10.1016/S1388-2457(02)00028-7

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21.

Goldberg, J. M., & Brown, P. B. (1969). Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. *Journal of Neurophysiology*, *32*(4), 613–636. https://doi.org/10.1152/jn.1969.32.4.613

Gramfort, A., Papadopoulo, T., Olivi, E., & Clerc, M. (2010). OpenMEEG: opensource software for quasistatic bioelectromagnetics. *BioMedical Engineering OnLine*, *9*(1), 45. https://doi.org/10.1186/1475-925X-9-45

Groh, J. M., Trause, A. S., Underhill, A. M., Clark, K. R., & Inati, S. (2001). Eye Position Influences Auditory Responses in Primate Inferior Colliculus. *Neuron*, *29*(2), 509–518. https://doi.org/10.1016/S0896-6273(01)00222-7

Haegens, S., Handel, B. F., & Jensen, O. (2011). Top-Down Controlled Alpha Band Activity in Somatosensory Areas Determines Behavioral Performance in a Discrimination Task. *Journal of Neuroscience*, *31*(14), 5197–5204. https://doi.org/10.1523/JNEUROSCI.5199-10.2011

Iemi, L., Busch, N. A., Laudini, A., Haegens, S., Samaha, J., Villringer, A., & Nikulin, V. V. (2019). Multiple mechanisms link prestimulus neural oscillations to sensory responses. *ELife*, *8*, e43620. https://doi.org/10.7554/eLife.43620

Jay, M. F., & Sparks, D. L. (1984). Auditory receptive fields in primate superior colliculus shift with changes in eye position. *Nature*, *309*(5966), 345–347. https://doi.org/10.1038/309345a0

Kayser, C., Petkov, C. I., Augath, M., & Logothetis, N. K. (2007). Functional Imaging Reveals Visual Modulation of Specific Fields in Auditory Cortex. *Journal of Neuroscience*, *27*(8), 1824–1835. https://doi.org/10.1523/JNEUROSCI.4737-06.2007

Keil, J., Müller, N., Hartmann, T., & Weisz, N. (2014). 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., 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., 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

Keil, J., & Senkowski, D. (2018). Neural Oscillations Orchestrate Multisensory Processing. *The Neuroscientist*, *24*(6), 609–626. https://doi.org/10.1177/1073858418755352

Lee, T.-W., Girolami, M., & Sejnowski, T. J. (1999). Independent Component Analysis Using an Extended Infomax Algorithm for Mixed Subgaussian and Supergaussian Sources. *Neural Computation*, *11*(2), 417–441. https://doi.org/10.1162/089976699300016719

Lenth, R. (2019). *Emmeans package: Estimated Marginal Means, aka Least-Squares Means. R package version 1.3.5.1*. Retrieved from https://CRAN.R-project.org/package=emmeans

Maiworm, M., Bellantoni, M., Spence, C., & Röder, B. (2012). When emotional valence modulates audiovisual integration. *Attention, Perception, & Psychophysics*, *74*(6), 1302–1311. https://doi.org/10.3758/s13414-012-0310-3

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Miller, E. K. (2000). The prefontral cortex and cognitive control. *Nature Reviews Neuroscience*, *1*(1), 59–65. https://doi.org/10.1038/35036228

Miller, L. M., & Recanzone, G. H. (2009). Populations of auditory cortical neurons can accurately encode acoustic space across stimulus intensity. *Proceedings of the National Academy of Sciences*, *106*(14), 5931–5935. https://doi.org/10.1073/pnas.0901023106

Mullette-Gillman, O. A., Cohen, Y. E., & Groh, J. M. (2005). Eye-Centered, Head-Centered, and Complex Coding of Visual and Auditory Targets in the Intraparietal Sulcus. *Journal of Neurophysiology*, *94*(4), 2331–2352. https://doi.org/10.1152/jn.00021.2005

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2010). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*, *2011*. https://doi.org/10.1155/2011/156869

Pantev, C., Ross, B., Berg, P., Elbert, T., & Rockstroh, B. (1998). Study of the Human Auditory Cortices Using a Whole-Head Magnetometer: Left vs. Right Hemisphere and Ipsilateral vs. Contralateral Stimulation. *Audiology and Neurotology*, *3*(2–3), 183–190. https://doi.org/10.1159/000013789

Park, H., & Kayser, C. (2019). Shared neural underpinnings of multisensory integration and trial-by-trial perceptual recalibration in humans. *ELife*, *8*, e47001. https://doi.org/10.7554/eLife.47001

Picton, T. W., Alain, C., Woods, D. L., John, M. S., Scherg, M., Valdes-Sosa, P., … Trujillo, N. J. (1999). Intracerebral Sources of Human Auditory-Evoked Potentials. *Audiology and Neurotology*, *4*(2), 64–79. https://doi.org/10.1159/000013823

Popov, T., Oostenveld, R., & Schoffelen, J. M. (2018). FieldTrip Made Easy: An Analysis Protocol for Group Analysis of the Auditory Steady State Brain Response in Time, Frequency, and Space. *Frontiers in Neuroscience*, *12*. https://doi.org/10.3389/fnins.2018.00711

R Core Team. (2013). *R: A Language and Environment for Statistical Computing*. Retrieved from http://www.R-project.org

Rauschecker, J. P., & Tian, B. (2000). Mechanisms and streams for processing of “what” and “where” in auditory cortex. *Proceedings of the National Academy of Sciences*, *97*(22), 11800–11806. https://doi.org/10.1073/pnas.97.22.11800

Rohe, T., & Noppeney, U. (2015). Cortical Hierarchies Perform Bayesian Causal Inference in Multisensory Perception. *PLoS Biol*, *13*(2), e1002073. https://doi.org/10.1371/journal.pbio.1002073

Rohe, T., & Noppeney, U. (2018). Reliability-Weighted Integration of Audiovisual Signals Can Be Modulated by Top-down Attention. *Eneuro*, *5*(1), ENEURO.0315-17.2018. https://doi.org/10.1523/ENEURO.0315-17.2018

Stekelenburg, J. J., Vroomen, J., & de Gelder, B. (2004). Illusory sound shifts induced by the ventriloquist illusion evoke the mismatch negativity. *Neuroscience Letters*, *357*(3), 163–166. https://doi.org/10.1016/j.neulet.2003.12.085

Tardif, E., Murray, M. M., Meylan, R., Spierer, L., & Clarke, S. (2006). The spatio-temporal brain dynamics of processing and integrating sound localization cues in humans. *Brain Research*, *1092*(1), 161–176. https://doi.org/10.1016/j.brainres.2006.03.095

Van Veen, B. D., Van Drongelen, W., Yuchtman, M., & Suzuki, A. (1997). Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Transactions On Biomedical Engineering, 44*(9), 867–880.

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

von Stein, A., & Sarnthein, J. (2000). Different frequencies for different scales of cortical integration: from local gamma to long range alpha/theta synchronization. *International Journal of Psychophysiology*, *38*(3), 301–313. https://doi.org/10.1016/S0167-8760(00)00172-0

Warren, J. D., & Griffiths, T. D. (2003). Distinct mechanisms for processing spatial sequences and pitch sequences in the human auditory brain. *The Journal of Neuroscience*, *23*(13), 5799–5804.

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*, *111*(4), E417–E425. https://doi.org/10.1073/pnas.1317267111

Werner-Reiss, U., Kelly, K. A., Trause, A. S., Underhill, A. M., & Groh, J. M. (2003). Eye Position Affects Activity in Primary Auditory Cortex of Primates. *Current Biology*, *13*(7), 554–562. https://doi.org/10.1016/S0960-9822(03)00168-4

Wetzels, R., & Wagenmakers, E.-J. (2012). A default Bayesian hypothesis test for correlations and partial correlations. *Psychonomic Bulletin & Review*, *19*(6), 1057–1064. https://doi.org/10.3758/s13423-012-0295-x

Woldorff, M. G., Tempelmann, C., Fell, J., Tegeler, C., Gaschler-Markefski, B., Hinrichs, H., … Scheich, H. (1999). Lateralized auditory spatial perception and the contralaterality of cortical processing as studied with functional magnetic resonance imaging and magnetoencephalography. *Human Brain Mapping*, *7*(1), 49–66. https://doi.org/10.1002/(SICI)1097-0193(1999)7:1<49::AID-HBM5>3.0.CO;2-J

Zierul, B., Röder, B., Tempelmann, C., Bruns, P., & Noesselt, T. (2017). The role of auditory cortex in the spatial ventriloquism aftereffect. *NeuroImage*, *162*, 257–268. https://doi.org/10.1016/j.neuroimage.2017.09.002

Appendix B: Erklärung über den Eigenanteil

Erklärung gemäß § 7 Abs. 3 Satz 4 der Promotionsordnung über den Eigenanteil an den veröffentlichten oder zur Veröffentlichung vorgesehenen eingereichten wissenschaftlichen Schriften im Rahmen meiner publikationsbasierten Arbeit

I. Name, Vorname: Kaiser, Mathis
Institut: Arbeitsbereich Neurocomputation and Neuroimaging
Promotionsfach: Psychologie
Titel: Neural oscillatory signatures of auditory and audiovisual illusions

II. Nummerierte Aufstellung der eingereichten Schriften:

1. Kaiser, M., Senkowski, D., Roa Romero, Y., Riecke, L., & Keil, J. (2018). Reduced low-frequency power and phase locking reflect restoration in the auditory continuity illusion. *European Journal of Neuroscience, 48,* 2849-2856.

2. Kaiser, M., Senkowski, D., Busch, N. A., Balz, J., & Keil, J. (2019). Single trial prestimulus oscillations predict perception of the sound-induced flash illusion. *Scientific Reports*, *9*, 5983.

3. Kaiser, M., Senkowski, D., Keil J. (submitted). Mediofrontal Theta-band oscillations reflect top-down influence in the ventriloquist illusion. *Human Brain Mapping.*

III. Darlegung des eigenen Anteils der Schriften:

Zu II.1: Konzeption (in Teilen), Datenvorverarbeitung (überwiegend), Datenauswertung (überwiegend), Ergebnisdiskussion (überwiegend), Erstellen des Manuskripts (überwiegend)

Zu II.2: Konzeption (in Teilen), Datenauswertung (überwiegend), Methodenentwicklung (in Teilen), Ergebnisdiskussion (überwiegend), Erstellen des Manuskripts (überwiegend)

Zu II.3: Konzeption (überwiegend), Versuchsdesign (mehrheitlich), Rekrutierung der Teilnehmer*innen (mehrheitlich), Datenerhebung (überwiegend), Datenvorverarbeitung (überwiegend), Datenauswertung (vollständig), Ergebnisdiskussion (überwiegend), Erstellen des Manuskripts (überwiegend)

IV. Namen und Anschriften nebst E-Mail der jeweiligen Mitautor*innen

Die Anschriften und E-Mail-Adressen sind in der Online-Version aus Gründen des Datenschutzes nicht enthalten.

Zu II.1: Daniel Senkowski

Yadira Roa Romero

Lars Riecke

Julian Keil

Zu II.2: Daniel Senkowski, s.o.

Niko A. Busch

Johanna Balz

Julian Keil, s.o.

Zu II.3: Daniel Senkowski, s.o.

Julian Keil, s.o.

26.02.2021 Mathis Kaiser

Datum, Unterschrift des Doktoranden

Ich bestätige die von Mathis Kaiser unter III. abgegebene Erklärung

Name: Daniel Senkowski Unterschrift: ____________________________________

Name: Yadira Roa Romero Unterschrift: ____________________________________

Name: Lars Riecke Unterschrift: ____________________________________

Name: Julian Keil Unterschrift: ____________________________________

Name: Niko A. Busch Unterschrift: ____________________________________

Name: Johanna Balz Unterschrift: ____________________________________

Appendix C: Eigenständigkeitserklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit eigenständig und ohne unerlaubte Hilfe verfasst habe, und dass Ideen und Gedanken aus Arbeiten anderer entsprechend gekennzeichnet wurden.

Berlin, 26.02.2021
Ort, Datum

Mathis Kaiser
Unterschrift

Appendix D: Lebenslauf

Ausbildung

Seit 10/2014	Promotion in Psychologie an der Freien Universität Berlin und der Berlin School of Mind and Brain, Humboldt Universität zu Berlin
09/2010 – 04/2011	Auslandsstudium an der York University, Toronto (Kanada)
10/2007 – 02/2014	Diplom in Psychologie an der Universität Konstanz, Schwerpunkte Kognitive Neurowissenschaften und klinische Psychologie, Nebenfach Soziologie

Berufserfahrung

seit 11/2019	Laborleiter im Exzellenzcluster „Science of Intelligence“, Technische Universität Berlin
09/2014 – 09/2019	Wissenschaftlicher Mitarbeiter in der AG Multisensorische Integration, Klinik für Psychiatrie und Psychotherapie, Charité Universitätsmedizin – Berlin
01/2013 – 07/2014	Wissenschaftliche Hilfskraft / Wissenschaftlicher Mitarbeiter im Neuroelectromagnetic Oscillations Lab, Klinische Neuropsychologie, Universität Konstanz
08/2012 – 10/2012	Praktikum im MEG Lab, Center for Mind/Brain Sciences (CIMeC), Universität Trento (Italien)
05/2011 – 07/2011	Praktikum im Critical Hermeneutics Lab, History and Theory of Psychology, York University, Toronto (Kanada)
11/2008 – 06/2010	Wissenschaftliche Hilfskraft im Ongoing-Brain-Oscillations-and-Behavior Lab, Klinische Neuropsychologie, Universität of Konstanz


model
task
slow
activation
inhibition
stimuli
combination
cognitive
pattern
reflect
band
ventriloquist
rate
interaction
frequency
temporal
connectivity
properties
idea
context
prestimulus
provide
data
pathway
cortical
object
area
binding
induced
illusion
feedback
masked
literature
compared
reduced
assemblies
state
problem
analyses
measured
hypothesis
influence
single
top
range
occipital
inference
audiovisual
spectral
structure
organization
gamma
oscillation
integration
across
research
alpha
excitability
primary
higher
EEG
onset
beta
framework
increase
stimulus
sound
question
unisensory
input
occur
location
showed
scale
neuronal
account
ongoing
theta
phase
activity
general
signature
dissertation
relevant
different
signal
terms
affected
information
order
human
mechanism
paradigm
suggested
tone
potential
observed
spatial
excitatory
phenomena
multimodal
thesis
lower
method
associated
modulation
perception
found
oscillatory
modality
level
response
superior
alignment
represent
increased
restoration
representation
based
related
formation
mode
relationship
frontal
cortex
studies
auditory
interrupted
study
mask
investigate
prediction
within
behavioral
power
multisensory
effect
continuity
region
demonstrated
participants
previous
affect
modalities
brain
window
principle
low
locking
sensory
trial
depend
neural
proposed
reflected
instance
investigated
reset
perceived
used
evidence
population
presented
continuous
operation
network
similar
gap
result
processing
specific
ERP
neuron
support
scene
time
amplitude
SIFI
flash
finding
difference
analysis
important
system
localization
supported
perceptual
role
frequencies
condition
source
change
functional
visual
dynamic
common
modulated
external
crossmodal
function