RESEARCH ARTICLE

WILEY

# Mediofrontal theta-band oscillations reflect top-down influence in the ventriloquist illusion

Mathis Kaiser[1,2] 🅾 | Daniel Senkowski[1] | Julian Keil[1,3]

[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

**Correspondence**
Mathis Kaiser, Department of Psychiatry and Psychotherapy, Charité Universitätsmedizin—Berlin, Berlin, Germany.
Email: mathis.kaiser@mailbox.org

## 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 readily discriminable 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**
EEG, multisensory perception, oscillations, theta

## 1 | 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 the cortical level, the location of unisensory auditory stimuli is processed along a dorsal stream, from caudal primary auditory cortex toward 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 260 ms was larger over the hemisphere contralateral versus 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) versus 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 (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. Such a finding would constitute evidence for the laterality of auditory responses as a more general mechanism for subjective sound localization. Furthermore, no study has investigated the relationship between perception in the VI and neural oscillations. Synchronization of neural oscillations has been proposed to orchestrate the integration of information across sensory modalities and involved brain areas (Keil & Senkowski, 2018; Senkowski et al., 2008). Therefore, if lateralized responses in the VI are mediated by connectivity between auditory, visual and multimodal areas, one might expect this to be reflected in frequency-specific modulations of 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 toward 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.

## 2 | METHODS

### 2.1 | 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 when no meaningful 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).

## 2.2 | Experimental design

### 2.2.1 | 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 min.

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, IL) at 72 dB SPL.

### 2.2.2 | 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.

### 2.2.3 | 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 Figure 1). In audiovisual trials, the auditory and visual stimulus onsets were simultaneous. 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.

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 ($A_CV_R$, $A_CV_L$, $A_RV_C$, $A_LV_C$), *congruent* trials, where the locations coincided ($A_LV_L$, $A_CV_C$, $A_RV_R$), and *divergent* trials, were the locations were on different sides relative to the central fixation ($A_LV_R$, $A_RV_L$). Two-hundred trials were presented for each of the four ventriloquist conditions, 100 trials for each of the three congruent conditions, and 50 trials for each of the two divergent conditions. Furthermore, 120 *unisensory auditory* trials ($A_LV_0$, $A_CV_0$, $A_RV_0$) 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 min.
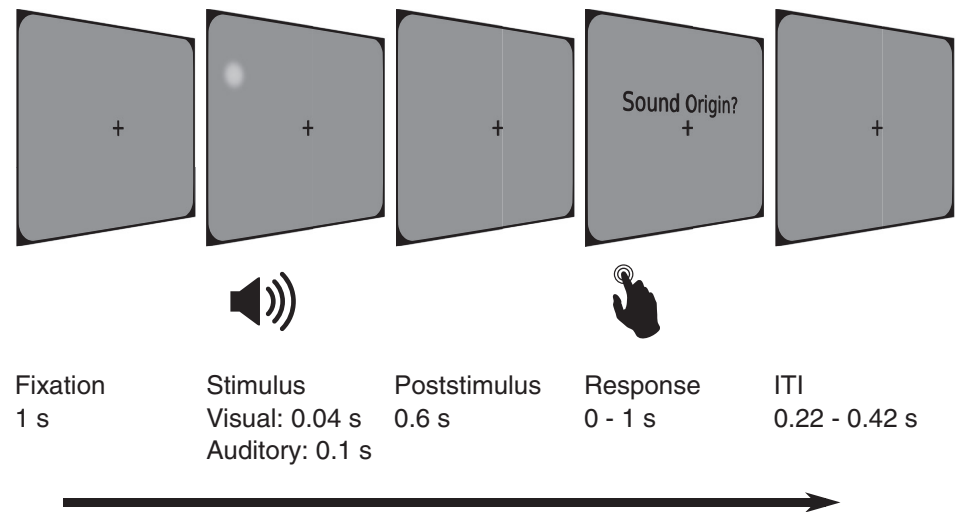
### 2.2.4 | Unisensory auditory experiment

This behavioral experiment was conducted to assess discrimination thresholds by determining the angles where the responses 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.

## 2.3 | Acquisition and preprocessing of EEG data

Prior to the experiment, individual head fiducials, electrode positions, and headshape were digitized using a Polhemus Patriot (Polhemus, VT). EEG was recorded using a 128-channel passive electrode cap (EasyCap,

**FIGURE 1** Timeline of one trial in the ventriloquist experiment

Fixation
1 s

Stimulus
Visual: 0.04 s
Auditory: 0.1 s

Poststimulus
0.6 s

Response
0 - 1 s

ITI
0.22 - 0.42 s

Sound Origin?

Herrsching, Germany), including one horizontal and one vertical electro-oculography (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 1,000 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 −6 dB cutoff frequency and an order of 3,300 for the high-pass, and −6 dB cutoff frequencies of 49.5 and 50.5 Hz and an order of 3,300 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).

## 2.4 | 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.

## 2.5 | Analysis of behavioral data

### 2.5.1 | Ventriloquist experiment

The illusion rate was computed for each participant as the percentage of ventriloquist trials where the response was at the visual location. To investigate the influence of visual information on auditory localization in the ventriloquist experiment, we compared the proportions of correct responses (defined by true auditory location) in the *ventriloquist*, *congruent*, and *unisensory auditory* trials using a $3 \times 2$ 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 posthoc tests, the estimated marginal means were contrasted where applicable using the emmeans package in R (Lenth, 2019), with Holm-correction for multiple comparisons. Finally, we computed an audiovisual weight index ($w_{AV}$, Rohe & Noppeney, 2016) for each subject to quantify the relative

contribution of auditory and visual information on responses (coded as −1/0/1 for left/center/right, respectively) in the ventriloquist trials. The $w_{AV}$ index is obtained by linear regression of auditory and visual locations on perceived locations, then computing the four-quadrant inverse tangent of the visual and auditory parameter estimates. It ranges from 0° to 90°, for pure auditory to pure visual influence, respectively. We also computed a circular-linear correlation between the $w_{AV}$ indices and illusion rates across participants using the CircStat toolbox (Berens, 2009; https://github.com/circstat/circstat-matlab; RRID:SCR_016651).

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.

### 2.5.2 | Unisensory auditory experiment

For each participant, we estimated the proportion of left/center/right responses along auditory azimuths using local linear fitting (Zchaluk & Foster, 2009). The resulting psychometric functions were averaged across participants. Individual and averaged discrimination thresholds were determined by calculating the angles were adjacent responses were equally likely to occur.

### 2.6 | 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.

### 2.6.1 | 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 specific stimulus combinations. 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* = .01, 1,000 permutations). On the source level, the pre-stimulus 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. Three-dimensional clusters were defined along neighboring electrodes (scalp level) or grid points (source level), time points and frequencies. 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 in favor of the null or alternative hypothesis.

Finally, we 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).

### 2.6.2 | 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 1 s 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, Oostenveld, and Schoffelen (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* = .01, 1,000 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* < .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 Figure 2). Since we used a symmetric dipole pair for the source reconstruction, the virtual channel showing the maximum and the homolog position in the left hemisphere were selected for the next analysis steps.

After defining ROIs for the further analysis, we selected trials from the ventriloquist experiment according to the location of the visual stimulus relative to the auditory stimulus, and according to perception. 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
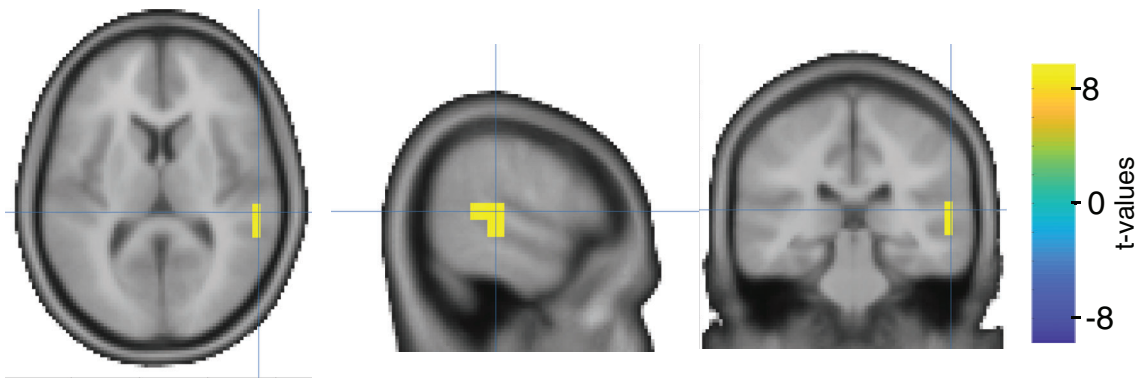
**FIGURE 2** Source reconstruction of 40 Hz-coherence (stimulation vs. baseline) in the auditory localizer experiment. *t*-values are masked using 95% of the maximal value, which is indicated by the crosshairs. *N* = 15

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, Handel, and Jensen (2011)), and ERP difference waves (*contralateral − ipsilateral*) were computed. Trials were combined across different stimulus locations because our analysis approach was based on the relative ipsi-/contralaterality of stimulus location and brain activity. Importantly, this approach captures asymmetries defined by responses in auditory cortex ipsi-/contralateral to stimulus location, irrespective of absolute hemisphere/ stimulus locations and absolute values of responses. 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* = .05, 1,000 permutations). Two-dimensional

clusters were defined along neighboring time points and frequencies. 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 ±20 ms average around 250 ms in the analysis. Averaged peak amplitudes were subjected to a repeated-measures ANOVA with Holm correction for four latencies. For posthoc 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.

## 3 | RESULTS

### 3.1 | 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) %. Mean discrimination thresholds in the unisensory auditory experiment were − 4.6° (center vs. left) and 8.2° (center vs. right), which indicates that subjects could reliably discriminate sound locations well below the angle used in the main experiment (±17.5°). Figure 3 illustrates mean response rates in all conditions of the main experiment and Figure 4 illustrates mean response rates and averaged psychometric functions 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, $p$ < .001) and an interaction between Visual Stimulus and Auditory Location ($F_{(2,50)}$ = 3.41, $p$ = .0408, see Figure 5). The main effect of Auditory Location was not significant ($F_{(1,25)}$ = 3.17, $p$ = .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, $p$ < .0001), higher for congruent compared to adjacent Visual Stimulus ($t$ = 13.821, $p$ < .0001), and higher for no compared to adjacent Visual Stimulus ($t$ = 9.17, $p$ < .0001). This indicates that visual stimuli influenced auditory perception and induced the VI. Posthoc 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, $p$ = .0089), but not when the Visual Stimulus was congruent or adjacent to the auditory location (both $p$ > .2). Simple contrasts of Visual Stimulus within the levels of Auditory Location mirrored those of the main effect (all $p$ < .01). The mean $w_{AV}$ index was 40.8 ± 28.3° (SD), indicating an influence of visual stimulation on auditory perception. Across participants, $w_{AV}$ was significantly correlated with illusion rate ($r$ [24] = .97, $p$ < .001).

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, $p$ = .0025). Posthoc tests revealed that illusion rates were lower in the second, third and fourth quartile compared to the first quartile ($t$ = 2.84, $p$ = .0292; $t$ = 3.04, $p$ = .0169; and $t$ = 3.62, $p$ = .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.

## 3.2 | 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 ($p$ = .038, illusion < no-illusion; see Figure 6). The cluster ranged from −0.5 to −0.12 s, between 4.2 and 4.9 Hz, that is, the theta band. The across-subject correlation between illusion rates and illusion−no-illusion power differences in the cluster was not significant ($R$ = −.35, $p$ = .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 ($p$ = .05,
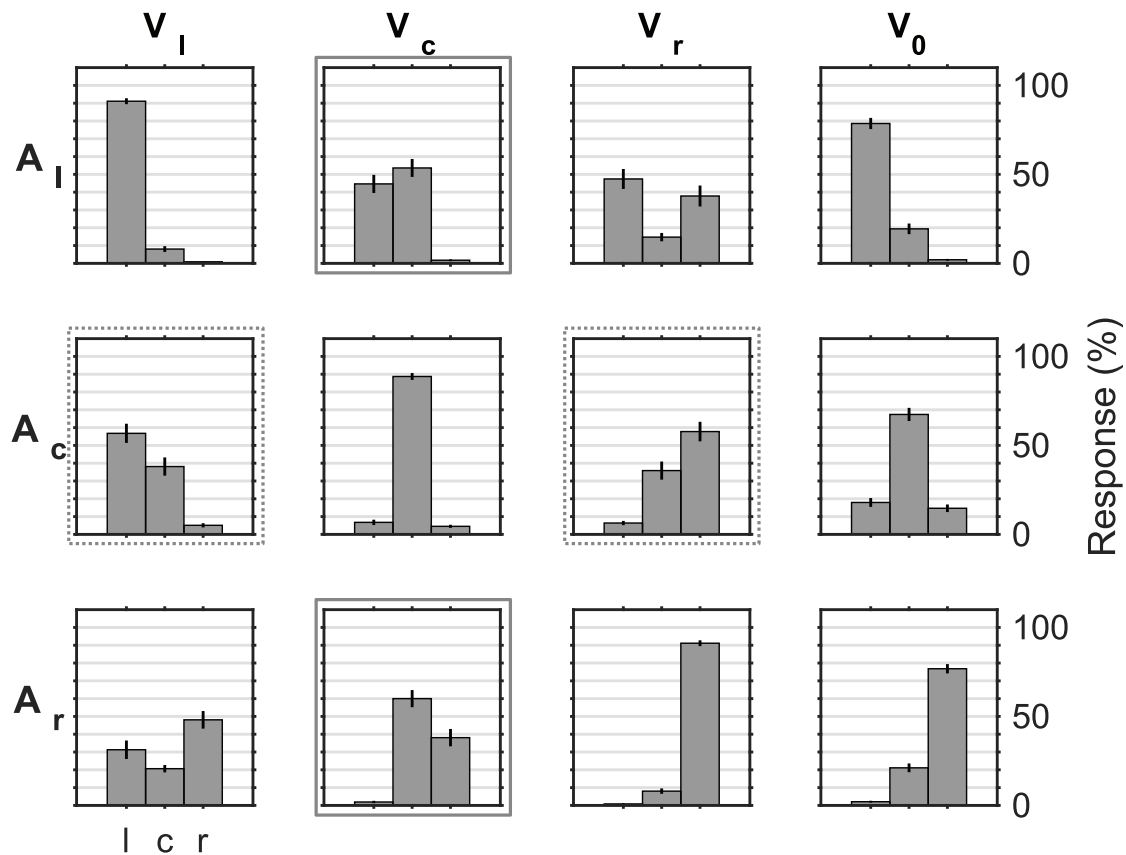


**FIGURE 3** mean response rates (l = left, c = center, r = right) for all conditions. Error bars indicate SEM. Solid and dashed gray lines around the plots indicate the central and peripheral ventriloquist conditions, respectively

**FIGURE 4** Mean response rates and corresponding psychometric functions 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°)
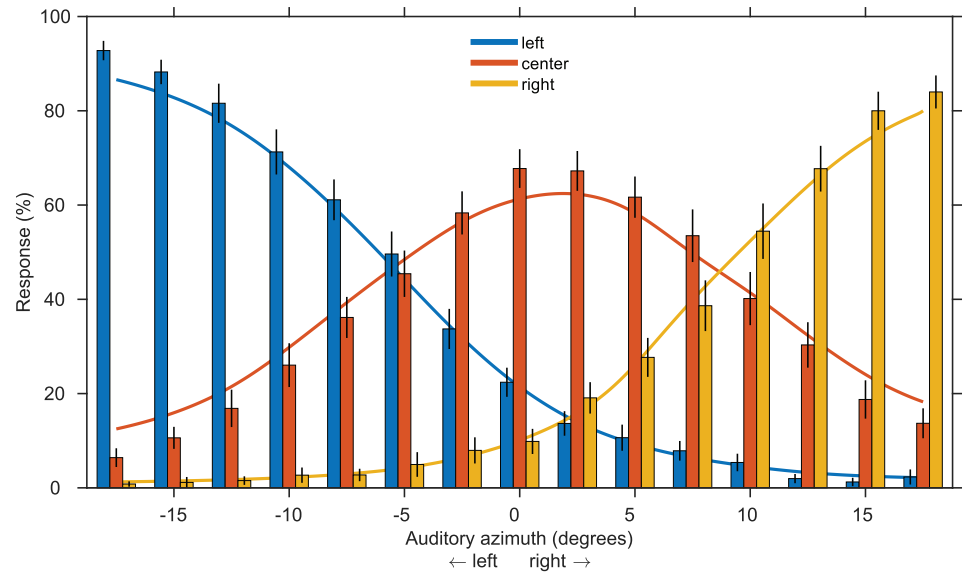


**FIGURE 5** Mean and individual response accuracies, indicated by bars and dots, respectively. Asterisks indicate significant posthoc tests for the 3 × 2 ANOVA (**$p$ < .01; ***$p$ < .001). Please note that the simple effects of visual stimulus, which mirrored the main effect, are not depicted
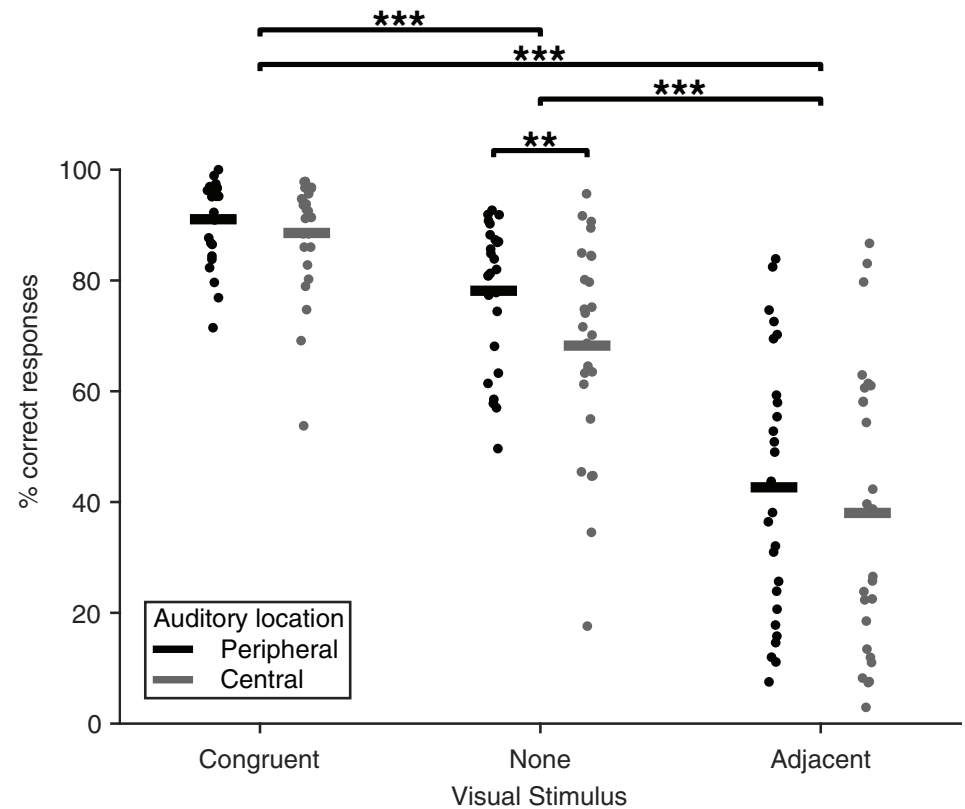


*illusion < no-illusion*). The cluster encompassed a −0.48 to 0 s time interval at a frequency of 2.7–4.4 Hz. Hence, the scalp level analysis, as well as the source level analysis revealed differential prestimulus low frequency power in the delta to theta range between illusion and no-illusion trials. 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 ($p$ = .014) ranging from 0 to 0.5 s, between 2.7 and 4.4 Hz was obtained (see Figure 7). The across-subject correlations between illusion rates and *illusion–no-illusion* power differences were not significant on the source level (prestimulus: $R$ = −.27, $p$ = .284; poststimulus: $R$ = −.44, $p$ = .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, $p$ = .0537), though there was a trend toward 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.
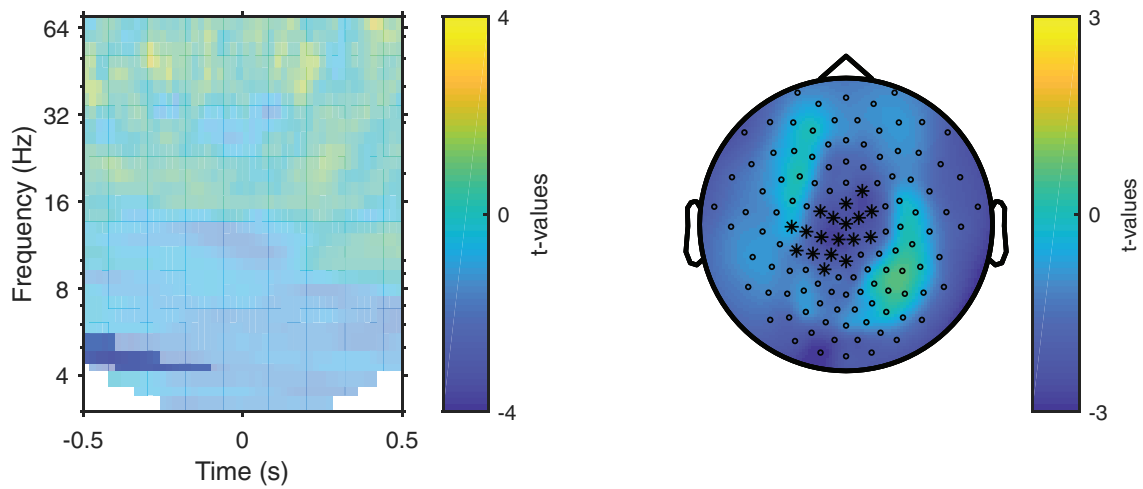
**FIGURE 6** Time-frequency spectrum and topography of *t*-values for the illusion/no illusion comparison on the sensor level. Significant regions/channels are indicated by saturation/asterisks, respectively
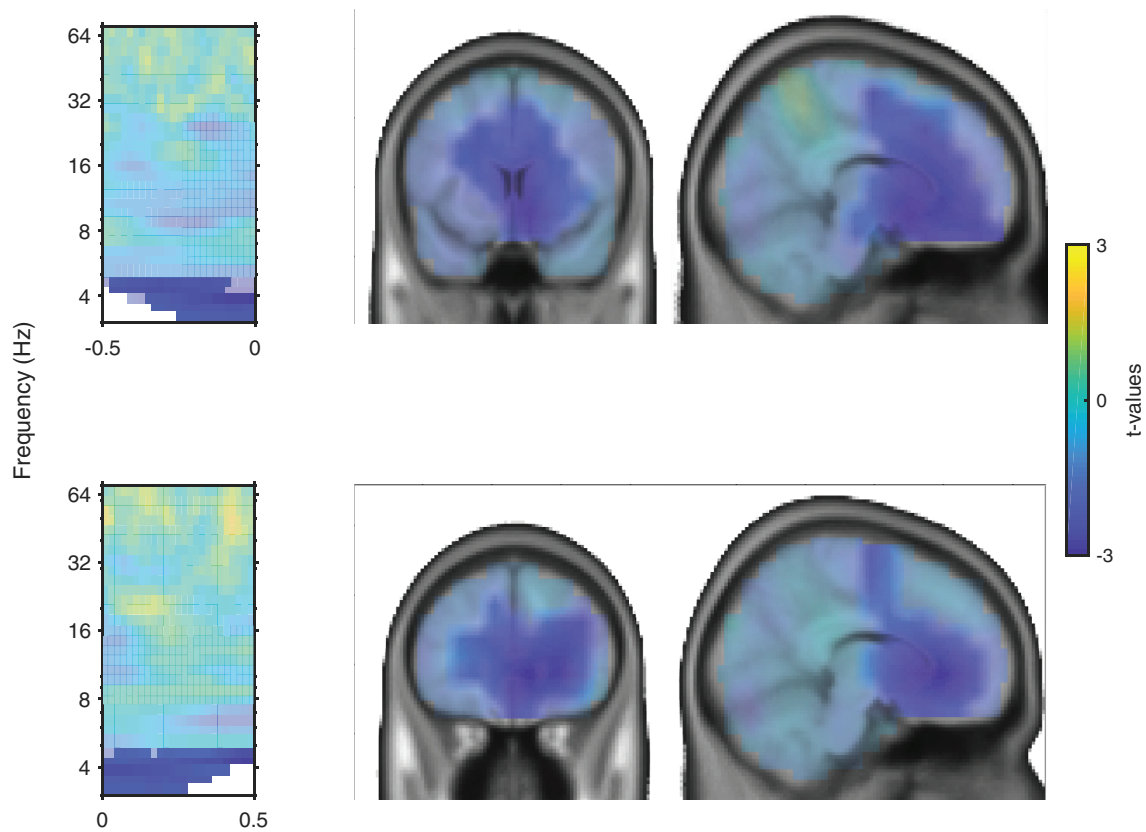


**FIGURE 7** Time-frequency spectrum and sourceplot of t-values for the illusion vs. no illusion comparison on the source level. Significant regions/virtual channels are indicated by saturation. The prestimulus period is shown at the top, the poststimulus period at the bottom

## 3.3 | 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 Figure 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* = .48). The analysis of the ERP

difference waves did also not reveal any significant effects (lowest negative cluster *p* = .67). ERPs and difference waves are illustrated in Figure 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* = .52; in the *peripheral* condition it was *p* = .42. In the ERPs, the lowest positive

cluster p-value for comparison in the *central* condition was *p* = .57. No clusters were found in the *peripheral* condition.

In the ANOVAs of averaged ERP amplitudes, no significant three-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* = .013). Post hoc tests revealed that
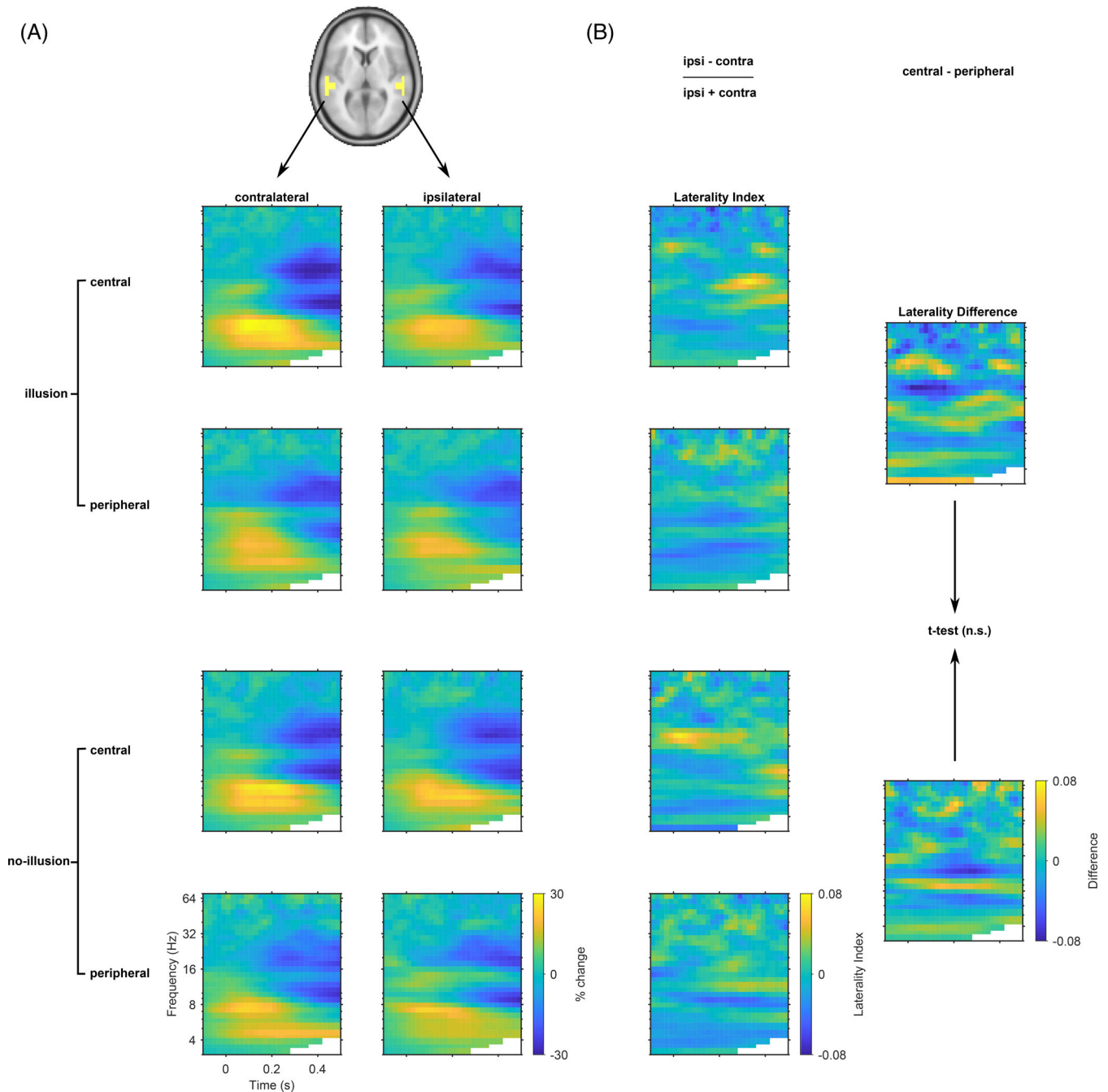


**FIGURE 8** (a) Time-frequency spectra in the contra- and ipsilateral source regions of interest for all conditions. Data are baseline-corrected using the 500 ms prestimulus period. Note that this correction was only applied for illustrative purposes, but not in the analyzed data, where a similar correction is implicit in the calculation of the laterality index. Also note that the left/right brain hemispheres do not represent physical brain locations, but contra–/ipsilaterality with regard to the stimulus. (b) Time-frequency spectra of the laterality index and tested differences
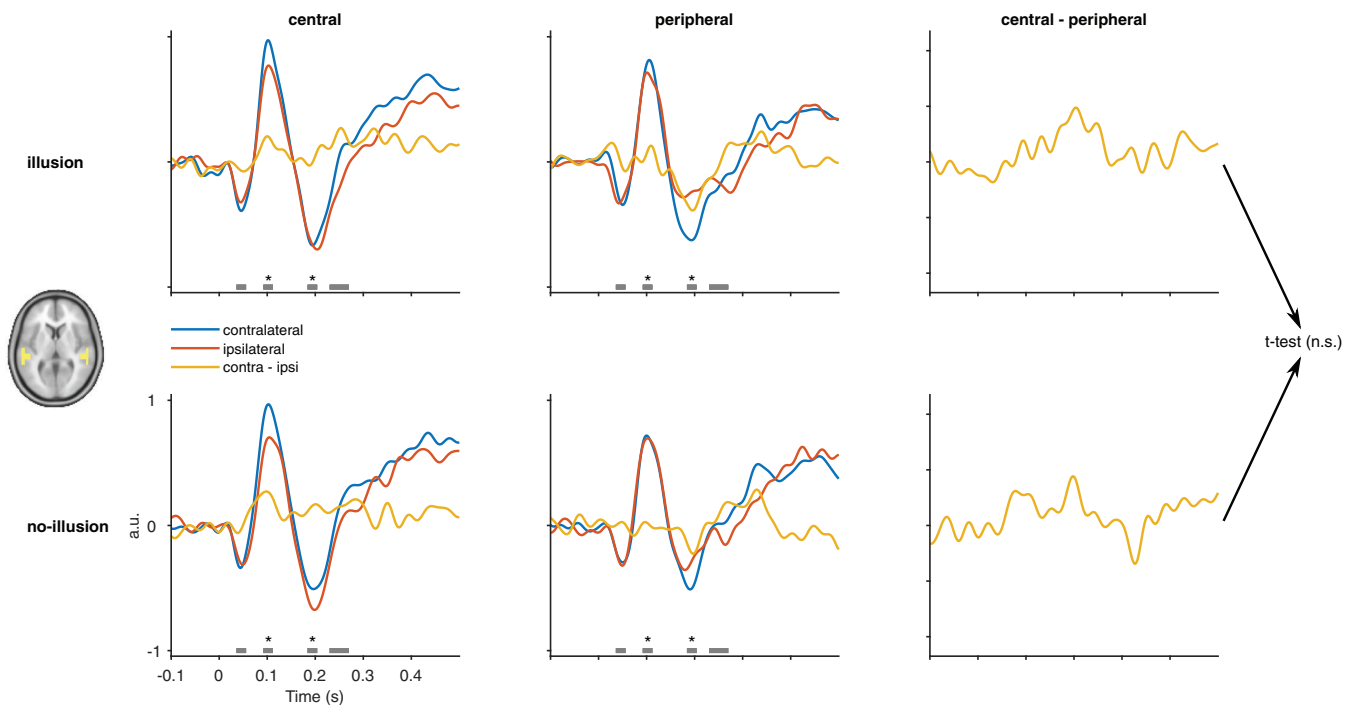
**FIGURE 9** Event-related potentials and difference waves in the source regions of interest. Time periods used for the ANOVA of averaged amplitudes are indicated by gray lines above the abscissa, and periods where main or interaction effects were found are indicated by asterisks. The figure illustrates 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

**TABLE 1** F- and p-values for the 3-factor ANOVA of ERP amplitudes.

| 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 | .3727 | 6.17 | .0788 | 7.24 | .0703 | 1.93 | .3727 |
| Perception | 1,14 | 1.41 | .5108 | 0.96 | .5108 | 2.14 | .4965 | 6.5 | .0924 |
| Hemisphere | 1,14 | 0.29 | .6 | 12.54 | .013* | 5.46 | .1046 | 3.72 | .1488 |
| Direction × perception | 1,14 | 0.05 | 1 | 0.03 | 1 | 0.12 | 1 | 0 | 1 |
| Direction × hemisphere | 1,14 | 0.78 | .394 | 6.14 | .0796 | 15.16 | .0065* | 1.84 | .394 |
| Perception × hemisphere | 1,14 | 0.16 | 1 | 0 | 1 | 1.8 | .8046 | 0.08 | 1 |
| Direction × perception × hemisphere | 1,14 | 0.07 | 1 | 0.45 | 1 | 0.02 | 1 | 0.52 | 1 |

*Note:* The reported p-values are Holm-corrected for comparisons at four latencies.
* Significant p-values.

amplitudes around 102 ms were larger in the contralateral compared to the ipsilateral hemisphere ($t = 3.54$, $p = .0033$). Moreover, a Direction × Hemisphere interaction effect was observed around 204 ms ($F_{(1,14)} = 15.16$, $p = .0065$). Post hoc tests revealed that amplitudes were larger in the contralateral compared to the ipsilateral hemisphere when visual stimuli were presented peripherally ($t = 4.44$, $p = .0005$). No such effect was found when visual stimuli were presented centrally ($t = 1.25$, $p = .6015$). Moreover, ERP amplitudes around 204 ms were larger for centrally presented visual stimuli, compared to peripherally presented visual stimuli within the ipsilateral hemisphere ($t = 4.33$, $p = .007$). 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 (Table 1).

## 4 | 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 oscillations in mediofrontal areas, starting prior to stimulus onset, facilitated the illusion.

## 4.1 | 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). The audiovisual weight index was tightly correlated with illusion rate, suggesting that the illusion rate directly reflects the relative contribution of auditory and visual information to perceptual judgments in the VI. 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.

## 4.2 | 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 and 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.

## 4.3 | Lateralization

This analysis focused on hemispheric asymmetries in auditory areas, depending on perceptual shifts toward 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* versus *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). Furthermore, the tuning of responses in auditory cortex to the contralateral hemifield depends on the presence of interaural time difference cues (Ortiz-Rios et al., 2017), which would be absent for visually induced shifts. 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.

## 4.4 | Limitations and future directions

The current analysis has a number of limitations. First, 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 and 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. Secondly, the frequency range differed somewhat between scalp- and source-level effects. However, the significance of this difference is hard to evaluate given the spectral resolution of 1 Hz (resulting from a 1 s analysis time window) and the limited inferences on the extent of effects warranted by cluster-based statistics (Sassenhagen & Draschkow, 2019). We consider the overlap from 4.2 to 4.4 Hz between effects on the scalp- and source-level to indicate a common signal origin in the theta-band.

Finally, 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. Our sample size in the EEG analyses was reduced by excluding a relatively large number of participants to reach a minimum trial count of 50 for the statistical comparisons. This number was chosen to reach sufficient signal-to-noise ratio of the EEG parameters (Cohen, 2014, p.65).

## 5 | CONCLUSION

Our findings indicate that modulations of theta band power, starting prior to stimulus onset in mediofrontal 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 mediofrontal theta-band oscillations serve as a neural mechanism underlying top-down control of crossmodal influence in the ventriloquist illusion.

## CONFLICT OF INTEREST

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

## DATA AVAILABILITY STATEMENT

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.

## ORCID

*Mathis Kaiser* https://orcid.org/0000-0002-8004-4083

## 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

Berens, P. (2009). CircStat: A MATLAB toolbox for circular statistics. *Journal of Statistical Software*, *31*(1), 1–21. https://doi.org/10.18637/jss.v031.i10

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*, 51. 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. (2014). *Analyzing neural time series data: Theory and practice* (1st ed.). Cambridge, MA: The MIT Press.

Cohen, M. X., & Cavanagh, J. F. (2011). Single-trial regression elucidates the role of prefrontal theta oscillations in response conflict. *Frontiers in Psychology*, *2*, 30–. 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

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

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 prefrontal 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

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*(156869). https://doi.org/10.1155/2011/156869

Ortiz-Rios, M., Azevedo, F. A. C., Kuśmierek, P., Balla, D. Z., Munk, M. H., Keliris, G. A., ... Rauschecker, J. P. (2017). Widespread and opponent fMRI signals represent sound location in macaque auditory cortex. *Neuron*, *93*(4), 971–983. https://doi.org/10.1016/j.neuron.2017.01.013

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*, 711. 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 Biology*, *13*(2), e1002073. https://doi.org/10.1371/journal.pbio.1002073

Rohe, T., & Noppeney, U. (2016). Distinct computational principles govern multisensory integration in primary sensory and association cortices. *Current Biology*, *26*(4), 509–514. http://dx.doi.org/10.1016/j.cub.2015.12.056

Rohe, T., & Noppeney, U. (2018). Reliability-weighted integration of audiovisual signals can be modulated by top-down attention. *Eneuro*, *5*(1). https://doi.org/10.1523/ENEURO.0315-17.2018

Sassenhagen, J., & Draschkow, D. (2019). Cluster-based permutation tests of MEG/EEG data do not establish significance of effect latency or location. *Psychophysiology*, *56*(6), e13335. https://doi.org/10.1111/psyp.13335

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 Atteveldt, N., Murray, M. M., Thut, G., & Schroeder, C. E. (2014). Multisensory integration: Flexible use of general operations. *Neuron*, *81*(6), 1240–1253. https://doi.org/10.1016/j.neuron.2014.02.044

Van 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.

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

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

Zchaluk, K., & Foster, D. H. (2009). Model-free estimation of the psychometric function. *Attention, Perception, & Psychophysics*, *71*(6), 1414–1425. https://doi.org/10.3758/APP.71.6.1414

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