

Published in final edited form as:

Neuroimage. 2013 October ; 79: 19–29. doi:10.1016/j.neuroimage.2013.04.060.

Auditory-driven phase reset in visual cortex: Human electrocorticography reveals mechanisms of early multisensory integration

Manuel R. Mercier^{1,3}, John J. Foxe^{1,2}, Ian C. Fiebelkorn¹, John S. Butler¹, Theodore H. Schwartz³, and Sophie Molholm^{1,2}

¹The Sheryl and Daniel R. Tishman Cognitive Neurophysiology Laboratory, Children's Evaluation and Rehabilitation Center, Departments of Pediatrics and Neuroscience, Albert Einstein College of Medicine, Bronx, New York 10461, USA

²The Cognitive Neurophysiology Laboratory, Nathan S. Kline Institute for Psychiatric Research, 140 Old Orangeburg Road, Orangeburg, New York 10962, USA

³Department of Neurological Surgery, Weill Cornell Medical College, New York Presbyterian Hospital, New York, New York 10021, USA

Abstract

Findings in animal models demonstrate that activity within hierarchically early sensory cortical regions can be modulated by cross-sensory inputs through resetting of the phase of ongoing intrinsic neural oscillations. Here, subdural recordings evaluated whether phase resetting by auditory inputs would impact multisensory integration processes in human visual cortex. Results clearly showed auditory-driven phase reset in visual cortices and, in some cases, frank auditory event-related potentials (ERP) were also observed over these regions. Further, when audiovisual bisensory stimuli were presented, this led to robust multisensory integration effects which were observed in both the ERP and in measures of phase concentration. These results extend findings from animal models to human visual cortices, and highlight the impact of cross-sensory phase resetting by a non-primary stimulus on multisensory integration in ostensibly unisensory cortices.

Keywords

multimodal sensory integration; oscillations; human visual cortex; electro-corticography (ECoG); intracranial; multisensory interaction

1. Introduction

Investigations into the mechanisms of sensory processing have traditionally focused on activity in the sensory cortices as a function of their respective primary sensory inputs (e.g.,

© 2013 Elsevier Inc. All rights reserved.

Corresponding author: Sophie Molholm, The Sheryl and Daniel R. Tishman Cognitive Neurophysiology Laboratory, Children's Evaluation and Rehabilitation Center (CERC), Departments of Pediatrics and Neuroscience, Albert Einstein College of Medicine, 1225 Morris Park Avenue, Bronx, New York, NY 10461, phone: +1 (718) 862 1823 fax: +1 (718) 862 1807.

Part of the data analysis was performed using the Fieldtrip toolbox for EEG/MEG-analysis, developed at the Donders Institute for Brain, Cognition and Behavior (Oostenveld et al., 2011; <http://www.ru.nl/neuroimaging/fieldtrip>).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

modulation of activity in visual cortex in response to visual stimuli). More recently, there has been concerted interest in whether and how ancillary cross-sensory inputs influence early sensory processing in so-called unisensory cortex. This has resulted in a major re-conceptualization of how the sensory systems interact to influence perception and behavior, with converging evidence that neural activity in a given cortical sensory region is modulated not only by its primary sensory inputs, but also by stimulation of the other sensory systems (Foxy and Schroeder, 2005; Meredith et al., 2009; Schroeder and Foxy, 2005; Schroeder and Foxy, 2002).

Non-invasive electrophysiological recordings in humans have revealed that multisensory inputs interact within the timeframe of early sensory processing (Fort et al., 2002; Foxy et al., 2000; Giard and Peronnet, 1999; Mishra et al., 2007; Molholm et al., 2002; Murray et al., 2005; Naue et al., 2011; Rajj et al., 2010; Thorne et al., 2011). Further, based on topographical mapping and dipole modeling of the underlying neuronal generators of these early effects, this likely occurs within sensory cortices. More precise anatomical localizations obtained from functional magnetic resonance imaging (fMRI) also support the occurrence of multisensory processing in hierarchically early sensory cortices (Foxy et al., 2002; Kayser et al., 2007).

A body of literature from animal studies has been building that provides important guidance on the nature of early multisensory modulations in sensory cortices. One particularly remarkable finding is that within sensory cortex, cross-sensory inputs serve to modulate neural responsiveness to a given region's principle sensory input. Thus, whereas an auditory input to visual cortex might not necessarily elicit a detectable response when presented in isolation, it is seen to modulate the evoked response when presented in conjunction with a visual stimulus (Allman et al., 2008; Allman and Meredith, 2007).

Providing a potential mechanism through which this is realized, electrophysiological recordings in non-human primates have revealed that the phase of ongoing oscillatory activity in primary and secondary auditory cortices can be "reset" by somatosensory or visual inputs (Kayser et al., 2008; Lakatos et al., 2007). The thesis has been forwarded that the phase of these intrinsic oscillations serves to alternate local cortical excitability between high and low states, driving neurons toward or away from their firing threshold (Lakatos et al., 2005). Under this account, the responsiveness of local neuronal populations to their primary sensory inputs is modulated by cross-sensory inputs through phase reset. Recent psychophysical and electrophysiological studies support the notion that these cross-sensory phase reset mechanisms have tangible implications for perceptual outcomes and behavior (Diederich et al., 2012; Fiebelkorn et al., 2011; Fiebelkorn et al., 2013; Thorne et al., 2011).

Here we took advantage of access to patients with subdural electrodes placed over posterior cortex to directly test for auditory responsiveness in traditionally visual regions, and to assess whether and how auditory stimulation would influence visual responses under conditions of multisensory stimulation. Specific analyses were directed at testing whether auditory stimulation results in phase reset of ongoing oscillations, and if in turn such phase resetting plays a role in audiovisual multisensory effects observed in human visual cortex.

2. Material and Methods

2.1. Patients

Data were collected from 5 patients implanted with subdural electrodes prior to undergoing pre-surgical evaluation for intractable epilepsy. Participants provided written informed consent, and the procedures were approved by the Institutional Review Boards of both the

Nathan Kline Institute and Weill Cornell Presbyterian Hospital. The conduct of this study strictly adhered to the principles outlined in the Declaration of Helsinki.

2.2. Electrode placement and localization

Subdural electrodes (stainless steel electrodes from AD-Tech Medical Instrument Corporation, Racine, WI) are highly sensitive to local field potentials and much less sensitive to distant activity, which allows for improved localization of the underlying current sources relative to scalp-recorded EEG. The number of electrodes per patient ranged from 100 to 126, and their use, placement and density were dictated solely by medical purpose.

The precise location of each electrode was determined through co-registration of pre-operative structural Magnetic Resonance Imaging (sMRI), post-operative sMRI, and CT scans. The preoperative sMRI provided accurate anatomic information, the postoperative CT scan provided an undistorted view of electrode placements, and the postoperative sMRI (i.e., a sMRI conducted while the electrodes were still implanted) allowed for an assessment of the entire co-registration process and the correction of brain deformation due to the presence of the electrodes. Co-registration procedures, normalization into Montreal Neurological Institute (MNI) space, electrode localization and image reconstruction were done through the BioImage suite software package and results projected on the MNI-colin27 brain (<http://www.bioimagesuite.org>; X. Papademetris, M. Jackowski, N. Rajeevan, H. Okuda, R.T. Constable, and L.H Staib. BioImage Suite: An integrated medical image analysis suite, Section of Bioimaging Sciences, Dept. of Diagnostic Radiology, Yale School of Medicine. <http://www.bioimagesuite.org>). Localizations were confirmed using the Statistical Parameter Mapping (SPM8) toolbox developed by Wellcome Department of Imaging Neuroscience at UCL (<http://www.fil.ion.ucl.ac.uk/spm/>) in conjunction with MRIcro (Rorden and Brett, 2000)

2.3. Selection of contacts of interest

The focus of this study was on auditory and multisensory activity in visual cortical regions. As such data analyses were directed at contacts placed over visual cortices. These cortices were defined anatomically and included the occipital, cuneus, lingual, fusiform and angular gyri. In addition to the use of these landmarks on an individual brain basis, we also confirmed that the MNI coordinates of the contacts of interests corresponded to one of the following Brodmann areas: 17, 18, 19, 37, 39, and posterior portions of both 7 and 20. This was done using the Yale Brodmann Areas Atlas Tool and confirmed by Talairach Daemon (<http://www.talairach.org>; Lancaster (Lancaster et al., 2000).

The contacts of interest (COIs) included 126 of the 576 contacts recorded from across 5 patients. Among the 126 electrodes that were identified as falling within our anatomical boundaries, 18 were excluded because of abnormal signals (either due to epileptic activity or noise artifacts), resulting in 108 COIs used for the analysis.

2.4. Stimuli and task

Auditory-alone, visual-alone, and audiovisual stimuli were presented equiprobably and in random order using Presentation software (Neurobehavioral systems). The inter-stimulus interval was randomly distributed between 750 and 3000 ms. The auditory stimulus, a 1000-Hz tone with a duration of 60 ms (5 ms rise/fall times), was presented at a comfortable listening level that ranged between 60 to 70dB, through Sennheiser HD600 headphones; the visual stimulus, a centered red disc subtending 3° on the horizontal meridian, was presented on a CRT (Dell Trinitron, 17") monitor for 60 ms, at a viewing distance of 75 cm. Patients maintained central fixation and responded as quickly as possible whenever a stimulus was detected, regardless of stimulus type (auditory-alone, visual-alone, or audiovisual). All

participants responded with a button press, using their right index finger (for previous application of this paradigm to probe multisensory processing, see (Molholm et al., 2002; Molholm et al., 2006), (Brandwein et al., 2011), and (Brandwein et al., 2012)). Each block included 100 stimuli, and the patients completed between 12 and 15 blocks. To maintain focus and prevent fatigue, patients were encouraged to take frequent breaks. Eye position was monitored by the experimenters.

2.5. Intracranial EEG recording and preprocessing

Continuous intracranial EEG (iEEG) was recorded using BrainAmp amplifiers (Brain Products, Munich, Germany) and sampled at 1000 Hz (Low/High cut off = 0.1/250 Hz). A frontally placed electrode was used as the reference during the recordings.

Offline, the iEEG was epoched from –1250 to 1250 ms, time-locked to stimulus onset. These epochs (+/- 500ms padding) then underwent artifact rejection. The threshold for rejecting a given trial was set at four times the standard score, with standardized z-values calculated across time, independently for each channel. Detrended epochs were further preprocessed to remove line noise (60/120/180 Hz) using a discrete Fourier transform, and high-pass (2 Hz) and low-pass (50 Hz) filtered using a two-pass 6th order Butterworth filter. Baseline correction was conducted over the entire epoch.

Local field potentials (LFP) were used to estimate the spatial derivative of the voltage axis (Butler et al., 2011; Gomez-Ramirez et al., 2011b; Perrin et al., 1987). A composite local reference scheme was applied in which the composite was defined by the number of immediate electrode neighbors on the horizontal and/or vertical plane (see equations 1 and 2). This number varied from 1 to 4 on the basis of the montage (grid or strip), and the reliability of the electrical signal (i.e., electrodes contaminated by electrical noise were not included). For instance, a five-point formula was applied when there were 4 immediate neighbors (grids), whereas a three-point formula was used when there were 2 immediate neighbors (strips). This approach was used to ensure maximum representation of the local signal, independent of the reference, and minimum contamination through diffusion of currents from more distant generators (i.e., volume conduction).

$$\mathbf{ERFP}_{i,j} = V_{i,j} - 1/4 \times (V_{i+1,j} + V_{i-1,j} + V_{i,j+1} + V_{i,j-1}) \quad (1)$$

$$\mathbf{ERFP}_k = V_k - 1/2 \times (V_{k-1} + V_{k+1}) \quad (2)$$

where $V_{i,j}$ (or V_k) denotes the recorded field potential at the i^{th} row and j^{th} column (or k^{th} position) in the electrode grid (or strip).

2.6. Analyses and Statistics

2.6.1. Testing the Race Model—A commonly used behavioral index of multisensory interaction (Brandwein et al., 2012; Brandwein et al., 2011; Harrison et al., 2010; Molholm et al., 2002; Molholm et al., 2006; Senkowski et al., 2006), Miller's race model (Miller, 1982), was computed for each participant. The race model places an upper limit on the cumulative probability (CP) of a reaction time at a given latency for stimulus pairs with redundant targets (i.e., targets indicating the same response). For any latency, t , the race model holds when this CP value is less than or equal to the sum of the CP from each of the single target stimuli (the unisensory stimuli). For each participant, the reaction time range within the valid reaction times (in this case, 100–800 ms) was calculated over the 3 stimulus types (auditory-alone, visual-alone, and audiovisual) and divided into quantiles from the 5th to 100th percentile in 5% increments (5, 10, ..., 95, 100%). At the individual level, a

participant was said to have shown race model violation if the CP of his/her RT to the audiovisual stimulus was larger than that predicted by the race model. A “Miller inequality” value is calculated by subtracting the value predicted by the race model from this CP value, and positive values represent the presence and amount of race-model violation.

2.6.2. ERP analysis—To compute ERPs, all re-referenced non-rejected trials were averaged for each stimulus type (auditory-alone, visual-alone, and audiovisual). To determine whether the ERPs represented a statistically significant modulation from baseline, post-stimulus amplitudes (from 0ms to 300ms) were compared to baseline amplitude values (from -100ms to 0ms). This was done using a paired-random permutation test for each post-stimulus time-point: For each trial, a time point was randomly selected from within the baseline period. These two paired values (pre- and post- stimulus time points) were then permuted or not (determined randomly), and the difference calculated. Finally, a mean value was calculated across all trials. This process was repeated 1000 times to generate a distribution that was used to determine whether or not the observed difference was statistically significant. P-value was computed by counting the proportion of permuted value than were larger or smaller (two tailed test: $p < 2.5\%$ and $p < 97.5\%$) than the observed mean of the differences. A unique distribution was generated for each condition for each time point at each COI.

Our choice to utilize the above procedure was motivated by what we consider to be a bias in the common approach to testing for significance against a baseline period. In the standard approach, baseline values are either randomly picked at different latencies across trials, or averaged across the baseline period, to build the surrogate distribution against which the observed values are compared. Therefore, if any prestimulus time-locked event exists, it will be washed out in the trial average. However, several studies have demonstrated that the prestimulus signal can carry information that is time-locked to stimulation. Our group and others have shown that this is especially the case in the context of multisensory experimental designs and can reflect anticipation, entrainment, or fluctuations in sustained attention (Besle et al., 2011; Fiebelkorn et al., 2013; Giard and Peronnet, 1999; Gomez-Ramirez et al., 2011a; Lakatos et al., 2009; Teder-Salejarvi et al., 2002). To confidently attribute post-stimulus activity to the stimulus input (rather than, for example, anticipatory activity), it is therefore necessary to demonstrate that the observed effect is different from time-locked activity in the prestimulus baseline period. We compared the two methods empirically, and observed that the present paired method was much more conservative than the commonly used approach.

2.6.3. Frequency analysis—To perform time-frequency decomposition, individual trials were convolved with complex Morlet wavelets, which had a width equal to five cycles (f_0/σ_f). The frequency range of these wavelets was 6 to 50 Hz (this cutoff circumvents issues with 60 Hz line-noise due to making recordings at the bedside in a hospital environment), every 10 ms, increasing in 1-Hz steps (i.e., 6, 7... 50 Hz). As such, the frequency and time resolutions ranged from 2.4 to 20 Hz and 265.3 to 31.8 ms, respectively (i.e.: spectral bandwidth, $2\sigma_f$; and wavelet duration, $2\sigma_t$). Power and phase concentrations were computed based on the complex output of the wavelet transform (Oostenveld et al., 2011; Roach and Mathalon, 2008; Tallon-Baudry et al., 1996). To avoid any back-leaking from post-stimulus activity into the pre-stimulus period, the baseline used for the time-frequency analysis was from -840 to -420 ms. This period was chosen such that, at the lowest frequency of interest (i.e., $833.3/2$ for 6 Hz), the temporal extent of the wavelet, which was convolved with the last time point of the baseline period, did not overlap with the post-stimulus period.

2.6.4. Analysis of phase concentration—To evaluate the presence or absence of systematic increases in phase concentration across trials, the phase concentration index (PCI; introduced as Phase Locking Factor [PLF] in (Tallon-Baudry et al., 1996), and also referred to as Inter Trial Coherence [ITC] in (Delorme and Makeig, 2004; Makeig et al., 2002) was computed as follows: The complex result of the wavelet convolution for each time point and frequency within a given trial was normalized by its amplitude such that each trial contributed equally to the subsequent average (in terms of amplitude). This provided an indirect representation of the phase concentration across trials, with possible values ranging from 0 (no phase locking) to 1 (perfect phase locking). To test for significant PCIs relative to stimulus onset, we used a Monte-Carlo Bootstrap procedure (Delorme, 2006; Delorme and Makeig, 2004) that was applied for each frequency and each time point (at each COI). A surrogate data distribution was computed by randomly selecting a complex value (normalized by its amplitude; see above) from within the defined baseline (−840 ms to −420 ms) from each trial, and then averaging across trials. After 1000 repetitions, the resulting baseline-PCI distribution was used to determine statistical significance. A one-tailed approach was used to determine statistical significance (p was considered as significant if $p < 0.05$) because we were specifically interested in detecting phase resetting, which is indicated by PCIs closer to 1.

2.6.5. Analysis of power—To assess evidence for phase resetting, we determined whether or not significant changes in phase concentration occurred in the absence of changes in power (see also section 3.2.2.). Event-related spectral perturbations were visualized by computing spectral power relative to baseline (i.e., the power value at each post-stimulus time point was divided by the mean of the baseline values). The significance of increases or decreases in power from baseline was calculated using the same Monte-Carlo Bootstrap procedure described above.

For both the ERP and power analyses, post-stimulus activity could be either positive or negative relative to baseline. Therefore a two-tailed threshold was used to determine statistical significance (p was considered as significant if $p < 0.025$ or $p < 0.975$). The same two-tailed threshold was used to assess the multisensory effects (both maximum criterion model and additive model, see below).

2.6.6. Multisensory statistics—The *maximum criterion model* measures whether the multisensory response differs from the maximum unisensory response. In the context of the current analysis, because the contacts of interest are restricted to so-called unisensory visual regions, the null hypothesis is that any response to an auditory stimulus, whether presented alone or paired with a visual stimulus, should be limited to noise. Any response to an audiovisual stimulus should therefore match the response to a visual-alone stimulus. On the other hand, if the null hypothesis is false, meaning that some information about the auditory stimulus is processed within so-called unisensory visual regions, we should observe significant changes in activity either in response to auditory-alone stimulation (analyzed as above), or through comparison of the responses to audiovisual and visual-alone stimulation. Therefore, to test whether auditory stimulation modulated the visual response, we compared the audio-visual and visual-alone responses. The application of the *maximum criterion* model allowed us to assess cross-sensory enhancement ($AV > V$) or suppression ($AV < V$) of the visual response.

When applying the maximum criterion model, statistical significance was determined using an unpaired randomization-permutation procedure. Single trials from the two stimulus conditions were randomly assigned to two separate pools of trials, with the only constraint being that the two randomized pools maintained their original size (between 129 and 377 trials). The average of the first pool was then subtracted from the average of the second

pool. By repeating this procedure 1000 times, a distribution was built from which significance thresholds could be determined.

In a second level of analysis, the COIs that demonstrated cross-sensory responses (either response to auditory-alone stimulation or modulation of the visual response by auditory input as assessed by the max model) were assessed for whether auditory-driven modulations interacted non-linearly with the visual response. Such non-linear multisensory effects were identified as either supra- or sub- additive by application of the *additive criterion model* (AV vs. [A+V]) (see e.g. (Avillac et al., 2007; Kayser et al., 2008; Stanford et al., 2005; Stein, 1998; Stein and Meredith, 1993)). For this, a randomization method was used in which the average audiovisual response was compared to a representative distribution of the summed 'unisensory' signals (see e.g., this approach as applied in (Senkowski et al., 2007)). This distribution was built from a random subset of all possible summed combinations of the unisensory trials (baseline corrected), with the number of summed trials corresponding to the number of audiovisual trials. All randomization procedures were performed independently for each time point, COI, and for the frequency analyses, for each frequency-band. For tests of MSI effects in power and phase concentration, the unisensory responses were summed before being transformed through a wavelet convolution due to the non-linearity of the wavelet transform (Senkowski et al., 2007; Senkowski et al., 2006).

2.6.7. Control for multiple comparisons—All p-values were corrected (in the time dimension for ERP, and both the time and frequency dimensions for PCI and power) using the commonly applied False Discovery Rate procedure from (Benjamini and Hochberg, 1995). This correction, a sequential Bonferroni-type procedure, is highly conservative and thus favors certainty (Type II errors) over statistical power (Type I errors - see (Groppe et al., 2011) for consideration of different approaches to controlling for multiple comparisons). This approach is widely used to control for multiple comparisons in neuroimaging studies (Genovese et al., 2002). The False Discovery Rate was set at 5%.

2.6.8. Quantitative summary—To descriptively quantify the results of the statistical analyses across COIs, we used the maximum absolute value over the window of analysis (0 to 300 ms post-stimulus onset) that retained significance following the stringent correction procedures, at a given COI, and for a given analytic approach. For analyses conducted in the frequency domain, this quantification was done separately for Theta/low- Alpha (6–10Hz), Beta (13–30Hz) and Gamma (30–50Hz), to distinguish effects among the different frequency bands.

All data analyses were performed in Matlab (the MathWorks, Natick, MA, USA) using personal scripts and the Fieldtrip Toolbox (Oostenveld et al., 2011).

3. Results

3.1. Behavioral data

Participants easily performed the task, with hit rates close to ceiling (95%±4%). Mean reaction-time data demonstrated the commonly observed multisensory redundant target effect (Gingras et al., 2009; Molholm et al., 2002; Molholm et al., 2006; Moran et al., 2008; Senkowski et al., 2006), with RTs to audiovisual stimuli (AV average across participant: 325±68ms) faster than RTs to either of the unisensory conditions (A: 365±49ms and V: 370±95ms). This pattern was observed in all participants. Additionally we tested the race model (Miller, 1982) to determine if multisensory response facilitation could be accounted for by simple probability summation of the fastest unisensory responses (Brandwein et al., 2012; Brandwein et al., 2011; Molholm et al., 2002; Molholm et al., 2006; Senkowski et al., 2006). There was race model violation in all five subjects in the fastest quantiles, indicating

that the joint probability of responding to the unisensory inputs cannot account for the fastest reaction times observed in the audiovisual condition.

3.2. Electrophysiological data: Primary and Cross-sensory responses

3.2.1. Event related potentials—In total, 69% of the COIs showed a visual evoked response (VEP) (74 of the 108 COIs) for the visual-alone and the audiovisual conditions. This was assessed visually by considering the transient and phasic aspect of the response that characterizes a typical ERP, and confirmed by a significant post-stimulus response. Rather remarkably, 20% of the COIs also showed significant responses to the auditory-alone condition (22 of 108, see waveform exemplars in Figure 1; see also Figure 2). Auditory evoked activity in visual cortical regions was notably different in character from the classic auditory evoked potentials (AEPs) seen over auditory cortex (see Figure S1, S2 and (Molholm et al., 2006), for more classic looking intracranial AEPs). Instead, the response was characterized by a slower oscillatory pattern that only reached significance for short periods of time corresponding to its peaks and/or troughs. Exemplars for each participant are depicted in Figure 1. This illustrates that visual-alone and audiovisual responses vary in form, amplitude and latency, as a function of COI. This is not surprising given that the COIs were placed over heterogeneous regions of visual cortex for which the preferred stimulation characteristics and timing of responses would be expected to differ. The figure also reveals the presence of responses to the auditory-alone condition in some of the COIs, although as would be expected these are of much smaller amplitude than the responses to stimuli containing a visual element. While the direction of the auditory evoked response usually followed that of the VEP, this was not always the case. Differences in polarity of response for different types of stimulation likely reflect different underlying neural populations with differing dipolar configurations. Localizations of COIs presenting significant ERP responses to auditory-alone stimulation are depicted on an MNI brain reconstruction in Figure 3A. Except for the extreme posterior region of the occipital pole, where representation was relatively sparse, the spatial distribution of auditory responses appeared relatively homogeneous with regard to COI coverage.

3.2.2. Time-frequency analysis—In a first stage of analysis, phase modulation of ongoing oscillatory activity in response to the different stimulus conditions was assessed with measurements of the PCI, as described in the methods. For both visual-alone and audiovisual stimulation conditions the PCI was qualitatively similar at a given COI, as one would expect given that they both include the primary visual response. For both conditions, a significant increase in phase concentration compared to baseline was seen in Theta/low-Alpha (6–10Hz) bands in 69% of the COIs, in the Beta band (13–30Hz) for 65% of the COIs, and in the Gamma band (30–50Hz) for 51% of the COIs (proportions were similar across participants; see Figure 2). In contrast, the auditory-alone condition resulted in a smaller PCI increase that was often restricted to the lower frequencies (i.e. 6–10Hz; see in Figure 1, row D, and Figure 2).

For the auditory-alone condition, increased PCI was observed in the Theta/low-Alpha bands in 27% of the COIs, in the Beta band in 20% of the COIs, and in the Gamma band in 11% of the COIs. This finding of non-random increases in phase concentration following stimulus presentation indicates that auditory stimuli modulate processing in visual cortex. Further, across the full frequency spectrum analyzed here, increases in PCI as a function of auditory stimulation were observed in 42% of the COIs (45 of 108), suggesting that this was a widespread phenomenon over visual cortical regions. Localizations of COIs showing significant changes in phase concentration after auditory-alone stimulus presentation (significant corrected PCI) are depicted in Figure 3B (indicated by square symbols). Their

spatial distribution, with regard to COI coverage, didn't show any distinctive pattern, suggesting a generalized phenomenon over visual cortices.

In a second stage of analysis of the PCI data, we examined whether maximum significant PCI increase was accompanied by a significant increase in power (for the same frequency band and at the same latency, see materials and methods section for a detailed description). That is, while primary sensory inputs tend to lead to both increases in PCI and increases in signal power, it is possible that non-primary cross-sensory inputs instead tend to lead to changes in phase concentration in the absence of significant increases in power, suggestive of a phase-resetting mechanism. This latter situation represents a potential facilitatory influence for simultaneously presented primary inputs, providing a mechanism of cross-sensory facilitation. This analysis revealed that, for audiovisual and visual-alone conditions respectively, 44% and 31% of the COIs with significant PCI were accompanied by statistically significant increases in power in the same frequency band and at the same latency. For the auditory-alone condition, a considerably smaller proportion of the COIs showing significant PCI also showed significant increases in power (6%). Thus, despite the fact that one cannot infer the absence of an effect from the lack of statistical significance, here it is clear that more COIs had significant increases in power for the visual-alone and audiovisual conditions than for the auditory-alone condition. To summarize, auditory stimuli tended to modulate the phase of ongoing oscillatory activity in visual cortical regions, and in contrast to visual-alone and audiovisual stimuli, for the majority of COIs this modulation was not accompanied by significant increases in power.

3.3. Electrophysiological data: Multisensory Effects

3.3.1. Modulation of the visual response by simultaneous auditory input:

Testing the maximum model—To assess whether auditory-alone stimulation modulated the visual ERP, we compared the audiovisual and visual-alone ERP responses (the max model (AV vs. V)). This revealed significant differences for 28 of the 79 contacts that showed a visual ERP (35%), with the audiovisual response tending to be larger than the visual-alone response (in 61% of the cases: 17 of 28 COIs).

53.6% of the COIs with significant MSI by the max model also had significant ERPs in response to auditory-alone stimulation (15 of 28; small red square with white surround in Figure 3A). Within the remaining COIs, for which there was no indication of an auditory ERP (small red circle with white surround in Figure 3A), fully 61.5% (8 of 13) showed significant increases in PCI in response to the auditory-alone condition (squares in Figure 3B). Thus, overall, 82% (23 of 28) of the COIs with MSI effects as assessed by the max model also showed statistically significant responses to the auditory-alone condition.

The max model was next applied to assess multisensory effects on phase concentration. Comparison of the PCI for visual-alone and audiovisual conditions revealed significant differences in phase concentration for 74% of the COIs (80 of 108; see Fig 2). The effect was more often found in the Beta and Gamma frequency bands (47% and 49% of the COIs, respectively) than Theta/low-Alpha bands (19% of the COIs). Furthermore, the direction of the effect was most often seen as an enhancement of phase concentration for the audiovisual condition (Theta/low-Alpha: 13% ; Beta: 34% ; Gamma band 27%) than a reduction compared to the visual-alone condition (Theta/low-Alpha: 6%; Beta: 13%; Gamma band 22%). Note that a contact might show both enhancement and suppression of PCI, since the data were divided into three frequency bands for the quantitative analyses (see exemplar in Figure 4). While it could be argued that by the max model, auditory-driven increases in PCI will inevitably lead to enhanced PCI for audiovisual versus visual-alone stimulation, thus diminishing the significance of multisensory effects, our data showed both enhanced and

decreased PCI multisensory effects. Thus it is clear that the auditory and visual inputs interacted to influence phase concentration.

Finally, of the COIs showing MSI effects under the max model when measured by changes in phase concentration (PCI), 51% (41) also showed significant responses to the auditory-alone condition (an ERP and/or phase modulation for the auditory-alone condition).

Thus, overall, auditory-driven modulation was observed in 90 of the 108 COIs (83%) and was characterized either by a response to the auditory-alone condition (ERP: 22 COIs, PCI: 45 COIs) and/or by a modulation of the visual response, as measured with the max. model (ERP: 28 COIs, PCI: 80).

3.3.2. MSI Characterization: The additive model—We next sought to characterize the nature of the auditory cross-sensory effects on multisensory processing. That is, for contacts in which an auditory response was observed or in which the max model was violated, we asked whether the responses to the multisensory inputs responded in a linear or a non-linear manner. Application of the *additive criterion model* to the ERP data (AV vs. A+V) revealed significant non-linear responses in 38% of these contacts (34 of the selected 90 COIs; see Fig 2), with the audiovisual response tending to be smaller than the summed response (sub-additive in 20 of the 34 COIs showing non-linear effects). A summary of these results is shown in Figure 3A as a 3D projection onto the MNI brain.

47% of these contacts (16 of the 34) also showed an ERP to auditory-alone stimulation (small white square with red surround in Figure 3A). Within the other 53% of contacts (i.e. for which there was no ERP response to the auditory-alone condition), 44% (8 of 18) showed significant increases in PCI in response to the auditory alone condition (squares in Figure 3B). Thus, overall, 71% of the COIs showing non-additive MSI also showed statistically significant responses to the auditory-alone condition.

Comparison of the PCI for the audiovisual and summed auditory-alone and visual-alone conditions (additive model; see methods section for details) revealed significant differences in 89 of the 90 selected COIs (61% in Theta/low-Alpha band, 81% in Beta and 86% in Gamma). These non-linear effects were more often supra-additive in the Theta/low-Alpha and Beta band (respectively 48% and 44% of the 90 COIs), although sub-additive effects were also observed (respectively 12% and 36% of the 90 COIs). In contrast, in the Gamma band the effects were more often sub-additive; 66% of COIs were sub-additive, and 19% were supra-additive. These results are summarized in Figure 2 and Figure 3B, the latter of which captures the widespread presence of MSI effects on PCI over visual cortices.

Finally, considering MSI effects measured by changes in phase concentration (PCI), under the additive model 54% (48 of 89) of the COIs also showed a significant ERP/or increase in phase concentration in response to the auditory-alone condition.

3.4. Consideration of the latency of the MSI effects as a function of COI location

Observation of Figure 1 reveals that the latency of the MSI effects could vary considerably over contacts. This is consistent with the recording of data from different neuronal populations with different functional properties. Under the assumption of a temporal processing hierarchy from posterior to anterior visual cortices, we considered the possibility that MSI effects would follow a similar temporal progression. To test for a simple relationship between anatomical coordinates and latency, correlation analyses were performed between each of the three MNI coordinates (x, y, and z; with our focus on the y coordinate, which represents the posterior to anterior axis) and the latency of the maximum

amplitude among the significant MSI effects. This analysis did not reveal significant relationships between latency and COI location ($p > 0.15$ in all three cases, uncorrected).

4. Discussion

Intracranial recordings in humans were used to test for auditory-driven modulation of neural activity in posterior visual cortices. The data revealed both frank auditory evoked potentials and widespread auditory-driven modulation of the phase of ongoing oscillatory activity in visual cortices (i.e. Brodmann Areas 17, 18, 19, 37 and 39). Furthermore, presentation of audiovisual stimuli led to clear MSI effects over these same regions, which were observed in both the ERP and in increased phase concentration of ongoing oscillations. In what follows, we describe these results in greater detail, parsing them in the context of related findings in the literature.

4.1. Auditory responses in visual cortex

A number of investigations suggest that auditory stimulation can modulate visual responses in visual cortex to influence early sensory-perceptual processing (Fort et al., 2002; Giard and Peronnet, 1999; Mishra et al., 2007; Molholm et al., 2002; Rajj et al., 2010). To date, however, there has been scant direct evidence for responses to auditory-alone stimulation in *human* visual cortex. Here, using intracranial recordings, we found that auditory-alone stimuli modulate neuronal activity in human visual cortices. These responses were generally consistent with an oscillatory pattern that emerged from baseline and was characterized by small amplitude variation, contrasting with the higher-amplitude classical ERPs evoked by visual-alone and audiovisual stimulation (see Figure 1, Figure S1 and S2).

This type of auditory response in visual cortex can be considered within the context of underlying mechanisms of ERP generation. The prevailing view stipulates that, following an incoming sensory input, an ERP ‘emerges’ from baseline due to synchronization of neuronal assemblies. Such synchronization results from stimulus-driven resetting of the phase of ongoing oscillations that may or may not be accompanied by an increase in neural activity (Becker et al., 2008; Makeig et al., 2004; Shah et al., 2004). The auditory responses recorded here in visual cortices suggest that cross-sensory stimulation influences sensory cortical activity by resetting the phase of ongoing activity, in the absence of increases in power. That is, the vast majority of the COIs demonstrated a profile typical of phase resetting. In contrast, presentation of a visual stimulus, the primary sensory input to these regions, was much more likely to lead to both increases in phase concentration and increases in power.

This pattern of activation in sensory cortex by cross-sensory inputs is reminiscent of findings from non-human primate studies (Kayser et al., 2008; Lakatos et al., 2007). For example, recording local field potentials and multiunit activity in macaques, Lakatos and colleagues (2007) have meticulously investigated the influence of somatosensory inputs on neural activity in primary auditory cortex. They showed that somatosensory inputs reset the phase of ongoing oscillations in the supragranular layers of this region. They also revealed that this doesn’t lead to increased firing, but rather reorganizes subthreshold membrane potential fluctuations (ongoing oscillations), rendering the neuronal ensemble more (or less) susceptible to discharge in response to stimulation from the primary sensory modality (Lakatos et al., 2007; Lakatos et al., 2005).

Although the current method does not allow us to determine the pathway by which auditory-driven influences occurred, there are several highly plausible possibilities to consider. These include a direct cortico-cortical auditory to visual pathway (Cappe and Barone, 2005; Clavagnier et al., 2004; Falchier et al., 2002), subcortical thalamic influences (Jones, 2001;

Schroeder and Lakatos, 2009; Sherman, 2007; Sherman and Guillery, 2002), or a mediating higher-order multisensory region such as posterior superior temporal gyrus (Tyll et al., 2012) (Werner and Noppeney, 2010) or the intra-parietal sulcus (Leitao et al., 2012).

4.2. Visual responses in visual cortices

While the visual evoked responses revealed more classic evoked componentry, with large amplitude peaks and troughs, the morphology of the response varied considerably from contact to contact (see exemplars in Figure 1). This is likely because, unlike with the scalp recorded ERP, there is not a strong spatial filtering of the intracranial ERP, and thus the different COIs represent the functional specificity of subsets of neurons in visual cortices. The quantitative analysis of the visual responses showed that, across the COIs, a large fraction showed an increase in PCI (81%, across frequencies) and the presence of an ERP (69%), whereas proportionally fewer had a significant increase in power (respectively 44% and 31% for AV and V; across frequencies) (see Figure 2). One possibility is that regions that are functionally specialized for a given type of input (e.g., based on topographic location, or stimulus feature) will tend to show an increase in power whereas regions that are not are less likely to. In a human intracranial study, Vidal and colleagues (Vidal et al., 2010) investigated category-specific visual responses recorded from electrodes distributed over broad regions of cortex and found that, depending on the electrode location in the brain and visual stimulus category, the response varied considerably both in terms of the presence of an ERP, power within the frequency band considered, and the relationship between the two. The role of phase resetting of ongoing oscillatory activity and increases in power and their respective contribution to the ERP remain an area of debate (Penny et al., 2002) that clearly cannot be resolved here. Nevertheless, the present data suggest that there are regions of visual cortices that respond to a given visual stimulus by phase modulation in the absence changes of power, and that such a response can lead to a VEP.

4.3. Multisensory integration effects: Auditory modulation of the visual response

To assess the impact of auditory stimulation on the processing of a co-occurring visual stimulus, we applied the maximum criterion model. The maximum criterion model revealed that ERPs to audiovisual stimulation were more often enhanced compared to visual-alone stimulation (in 60% of cases). Notably, such MSI effects did not systematically imply that auditory-alone stimulation evoked a significant ERP at the same COI. This latter observation fits well with findings from a number of animal investigations (Kayser et al., 2008; Lakatos et al., 2007; Meredith et al., 2009). Meredith and colleagues, for example, reported analogous findings from a series of single-cell studies in cat visual areas AEV (Anterior Ectosylvian Visual area) and PLLS (PosteroLateral Lateral Suprasylvian visual area) using audiovisual stimulation, and in somatosensory area RSS (Rostral SupraSylvian sulcus) using audio-tactile stimulation (Allman and Meredith, 2007; Clemo et al., 2007; Meredith and Allman, 2009; Meredith et al., 2009). They found increased firing rates for bisensory compared to unisensory, primary stimulation, and this was observed for two classes of neurons. The first were termed “subthreshold” neurons, which only responded to the primary sensory modality, but whose response was modulated by simultaneous stimulation in a non-primary modality. The second were bisensory neurons, which responded to both types of unisensory inputs but showed a clear preference for the primary sensory input.

The preponderance of enhanced MSI responses under the max model observed here may reflect the so-called *principle of inverse effectiveness* (Stein and Meredith, 1993). That is, recordings of single cell activity, LFPs, and surface EEG have repeatedly demonstrated that multisensory enhancement tends to be greatest when the unisensory inputs are minimally effective when presented in isolation (Avillac et al., 2007; Kayser et al., 2008; Meredith and Stein, 1986; Senkowski et al., 2011). For example, using single and multiunit recordings in

nonhuman primates, Kayser et al. (2008) reported that for a given neuron or set of neurons in primary and secondary auditory cortex, a visual stimulus that accompanied an “optimal” auditory stimulus was found to suppress the activity of this neuron. Conversely, a visual stimulus presented with a relatively ineffective auditory stimulus lead to an enhanced response. In the present study, the stimuli can be considered globally as suboptimal with regard to the large number of neurons’ activity necessarily represented in the response. That is, since we recorded at the population level and over different functional visual areas, fewer neurons or groups of neurons are likely to have been specifically tuned to the particular stimulus that we utilized. Thus a prevalence of enhanced over diminished MSI effects might be expected.

In summary, across the COIs, the maximum criterion model showed net enhanced MSI effects, with the ERP response to the audiovisual stimulus being, on average, larger in amplitude than the ERP response to the visual-alone stimulus.

We next sought to address how phase modulation by cross-sensory auditory inputs interacted with phase modulation evoked by the primary visual input. Assessing MSI at the phase level, the maximum criterion model revealed greater stimulus-driven phase concentration for the audiovisual condition, with the audiovisual stimulus leading to greater phase concentration compared to the visual-alone stimulus.

A previous study similarly sought to investigate this question, with the use of low-density scalp EEG recordings (Naue et al., 2011). Here they also manipulated the SOA between the auditory and visual components, to consider how this might influence multisensory processing. Though there were potentially interesting data in this study, in testing for MSI, the additive model was not applied. Reliance on the max model is problematic for scalp recorded data, where there is significant volume conduction of the electrical signal (Besle et al., 2004). Thus in comparing the responses between the audiovisual and visual conditions, even when focusing on signal from electrodes over posterior scalp regions, differences may represent processing of the auditory signal in auditory cortices (that is, due to the conduction properties of the scalp, signal from auditory cortex can be conducted to electrodes over more posterior regions of the scalp). In contrast, the use of intracranial recordings, as in the present study, rules out the possibility that volume conduction of signal from auditory cortex accounts for findings recorded over visual cortices.

4.4. Characterization of multisensory integration effects

To characterize the nature of the MSI effects, we applied the additive model. This allowed us to determine if the responses to the multisensory inputs were linear or a non-linear, and when non-linear, if they were supra- or sub- additive. For the ERP data, for the majority of contacts the results were consistent with a linear response (for 62%, the AV response was not statistically different from the A+V response). Of the 38% that showed a statistically significant non-linear response, the response to AV stimulation tended to be significantly smaller than the sum on the unisensory responses (A plus V), resulting in a preponderance of sub-additive MSI responses (59%). When Kayser et al. (2008) applied the additive model, they similarly found mainly sub-additive interactions (60% of the recording sites at the LFP level). In contrast to the ERP data, analysis of the PCI data revealed a preponderance of non-linear responses (99% of the COIs considered). Thus, for most of the visual regions assayed, there was stronger phase concentration after audiovisual stimulation relative to what would be predicted by the simple summation of the auditory-alone and visual-alone responses. Such a finding – increased PCI following bisensory stimulation- has been previously hinted at (Lakatos et al., 2007), and shown to be selectively present at 10Hz for sites where there were enhanced MSI effects in the local field potential response (Kayser et al., 2008). In the present findings, non-linear PCI MSI effects were present for the COIs that showed ERP

MSI regardless of the direction of the effect as well as of event-related brain potentials/fields in critical frequency bands.

4.5. MSI in the absence of cross-sensory inputs?

It is notable that, whether assessed using ERPs or PCI, there were COIs that showed the presence of MSI in the absence of evidence for any direct auditory input to the implicated region during the auditory-alone condition (in the ERP analysis of MSI: 18% of cases where significant multisensory responses were determined under the max model and in 29% of cases where the same was seen using the additive model; in the PCI analysis of MSI, in 49% and in 46%, respectively). This raises the question of how a multisensory effect can be observed in the absence of input from one of the constituent modalities. A number of possible explanations pertain.

First, to quantify all the results, they have to be described in binary terms: the presence or the absence of an effect based objectively on a predefined threshold and stringent multiple comparison corrections. Therefore it could be, in some cases, that an effect was extremely close to the corrected threshold, clearly visible in the mean, but did not reach statistical criteria. Such an instance is seen in Figure 1. The COI exemplar from participant S1 shows an ERP-like modulation in response to the auditory-alone condition that does not survive correction (statistical results with and without correction are depicted as transparent and solid respectively).

Second, it is reasonable to suppose that some of these multisensory effects reflect feedback from up-stream multisensory processing. Indeed, prior to the discovery of direct cortico-cortical connections between the sensory cortices (Falchier et al., 2002), it was assumed that the sensory cortices were not directly involved in multisensory integration, with sensory inputs only being integrated in higher-order multisensory regions (such as IPS or STS) following extensive unisensory processing. Thus neuroimaging data showing multisensory effects in sensory cortices were thought to reflect feedback from upstream higher-order associative regions (Calvert, 2001). While current understanding has led to a revision of this view, the early sensory responses that have been observed in these higher-order multisensory regions (Molholm et al., 2006) and the presence of extensive bidirectional anatomical connectivity with early sensory cortices (Falchier et al., 2002; Neal et al., 1990; Rockland and Ojima, 2003; Seltzer and Pandya, 1994), is clearly consistent with this possibility. Effective connectivity analysis of fMRI data in a multisensory paradigm also supports the notion that multisensory effects in unisensory cortices are mediated by both direct (cross-sensory) and indirect (via higher-order multisensory regions) connectivity (Werner and Noppeney, 2010).

Finally, yet another plausible explanation is limitations inherent to our analysis. Because of the interstimulus interval that we used, it was not possible in the present study to investigate frequencies below 6 Hz. However, work from our group and others has shown the importance of the Delta band (1 to 4 Hz) in carrying cross-sensory information (Fiebelkorn et al., 2011; Fiebelkorn et al., 2013; Lakatos et al., 2007). Therefore it is possible that auditory-driven activity in the slower frequencies impacted some of the MSI effects reported here.

4.6. MSI and the different frequency bands

Studies on the role of the different oscillatory frequency bands in sensory information processing suggest different functional roles for the higher and lower frequency bands. Because of their cycle length, slower oscillations support better functional coupling of networks over much larger distances due to conduction delay properties (Kopell et al.,

2000). On the other hand, the (faster) Gamma band seems to be more linked to bottom-up processing (Engel and Fries, 2010; Gray et al., 1989). Therefore, while faster oscillations reflect more localized and direct stimulus-driven inputs, the lower frequencies seem to reflect spatially dispersed modulatory inputs (Belitski et al., 2008; Rasch et al., 2008). In what follows we see that the lower frequency bands (Theta/low-Alpha) appear to be most involved in auditory modulation of visual cortical activity, which requires longer distance communication, whereas the higher frequencies (Gamma band) are where the MSI effects dominate, seemingly reflecting the influence of auditory inputs on local visual processing. Following a summary of the auditory and MSI effects, we conclude with consideration of the broader Beta band effects, which we link to general sensory-motor processes.

4.6.1. Theta/low-Alpha Band—Auditory-driven increases in phase concentration over visual cortices was most often observed in the Theta/low-Alpha frequency band, strongly suggesting a central role for this frequency band in mediating communication between auditory and visual cortices. Non-human primate studies also point to the significance of these lower frequency bands in phase reset of ongoing oscillations by a non-primary modality in sensory cortices. For instance, Lakatos et al. (2007) found that contra- versus ipsi- lateral somatosensory stimulation differentially influenced the phase of Theta oscillations in primary auditory cortex, with contra- and ipsi- lateral stimulations leading to phase concentrations at counter phase values. They suggested that this differential modulation of phase might play a role in the *spatial principal* of multisensory integration, whereby optimal and non optimal phase setting, depending on spatial alignment, could explain respective enhancement and suppression observed in multisensory effects. A different example comes from Kayser et al. (2008), where it was shown that for multisensory sites characterized by enhanced multisensory responses, the pre-stimulus phase of ongoing Theta/low-Alpha (in their study, specifically 10Hz) oscillations in auditory cortex influenced the direction of audio-visual multisensory effects.

4.6.2. Gamma band—Across the visual regions assayed, multisensory influences on phase alignment (measured with the PCI) were observed more often in the Gamma band than in the Theta/low-Alpha band. Furthermore, whereas those effects in lower frequencies showed enhanced and/or supra-additive MSI, the direction of the effects in the Gamma band were more variable. One possibility is that this difference reflects the modulatory character of Theta/low-Alpha and the stimulus-related processing aspect of the Gamma band. Here we reported observations from contacts distributed over a wide range of visual areas. Therefore it is likely that we recorded from distinct “functional units”, which were differentially sensitive to the stimuli that we used. This difference in susceptibility/specialization toward different features of the stimulation could explain the greater variability in the direction of MSI effects in the Gamma band. Thus visual cortices would be “informed” about cross-sensory auditory inputs through phase modulations in the Theta-low Alpha band, with differential impact on gamma activity within a given visual functional unit depending on its functional tuning characteristics.

4.6.3. Beta Band—Across all of the phase analyses conducted in the present paper, the Beta band was found to be consistently involved. Numerous COIs showed auditory-driven increases in phase concentration in this band, as well as multisensory effects on the phase of Beta oscillations. In the literature, the Beta band (13–30Hz) is most often associated with sensory-motor processing (Engel and Fries, 2010). With regard to multisensory processing, in a scalp EEG study, von Stein and colleagues (von Stein et al., 1999) observed an increase in Beta band coherence between temporal and parietal scalp regions during multisensory compared to unisensory stimulation. Similarly, Senkowski et al. (Senkowski et al., 2006) showed that Beta activity predicted response speed in the same audiovisual reaction time

task used in our study, suggesting that activity in the Beta band may play an important role in communication between sensory and motor cortical regions. In contrast to our results, no MSI effects on PCI in the beta band were reported in Lakatos et al. (2007, 2009) or in Kayser et al. (2008). Leaving aside the species differences and the scale of recording, one main difference is related to the performed task. In those studies the animals either observed the stimuli passively or responded to rare targets (which were not included in the main analyses), whereas in our study the participants had to respond as soon as they perceived a stimulus regardless of the modality. That is, all stimuli were relevant to task performance. Consequently, our study design may be more likely to lead to frequent communication between cortices dedicated to sensory processing/stimulus detection and motor cortices.

5. Conclusions

The present study establishes, within the context of a simple reaction time task, that auditory-driven phase reset of ongoing oscillations is a common phenomena over visual cortices. Characterization of audiovisual MSI within the same cortical region revealed effects both on the amplitude of the averaged ERP, as well as on the inter-trial consistency of phase. These results extend findings from animal models to human visual cortices, and indicate a role for cross-sensory phase resetting by a cross-sensory input on multisensory integration in visual cortices.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding:

This work was supported by a grant from the U.S. National Institute of Mental Health (MH 85322 to S.M. and J.J.F.) and a postdoctoral fellowship award from the Swiss National Science Foundation (PBELP3-123067 to M.R.M).

We would like to thank the five patients who donated their time and energy with enthusiasm at a challenging time for them.

References

- Allman BL, Bittencourt-Navarrete RE, Keniston LP, Medina AE, Wang MY, Meredith MA. Do cross-modal projections always result in multisensory integration? *Cereb Cortex*. 2008; 18:2066–2076. [PubMed: 18203695]
- Allman BL, Meredith MA. Multisensory processing in “unimodal” neurons: cross-modal subthreshold auditory effects in cat extrastriate visual cortex. *J Neurophysiol*. 2007; 98:545–549. [PubMed: 17475717]
- Avillac M, Ben Hamed S, Duhamel JR. Multisensory integration in the ventral intraparietal area of the macaque monkey. *J Neurosci*. 2007; 27:1922–1932. [PubMed: 17314288]
- Becker R, Ritter P, Villringer A. Influence of ongoing alpha rhythm on the visual evoked potential. *Neuroimage*. 2008; 39:707–716. [PubMed: 17977023]
- Belitski A, Gretton A, Magri C, Murayama Y, Montemurro MA, Logothetis NK, Panzeri S. Low-frequency local field potentials and spikes in primary visual cortex convey independent visual information. *J Neurosci*. 2008; 28:5696–5709. [PubMed: 18509031]
- Benjamini Y, Hochberg Y. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B-Methodological*. 1995; 57:289–300.

- Besle J, Fort A, Giard MH. Interest and validity of the additive model in electrophysiological studies of multisensory interactions. *Cognitive Processing*. 2004; 5
- Besle J, Schevon CA, Mehta AD, Lakatos P, Goodman RR, McKhann GM, Emerson RG, Schroeder CE. Tuning of the Human Neocortex to the Temporal Dynamics of Attended Events. *Journal of Neuroscience*. 2011; 31:3176–3185. [PubMed: 21368029]
- Brandwein AB, Foxe JJ, Butler JS, Russo NN, Altschuler TS, Gomes H, Molholm S. The Development of Multisensory Integration in High-Functioning Autism: High-Density Electrical Mapping and Psychophysical Measures Reveal Impairments in the Processing of Audiovisual Inputs. *Cereb Cortex*. 2012
- Brandwein AB, Foxe JJ, Russo NN, Altschuler TS, Gomes H, Molholm S. The development of audiovisual multisensory integration across childhood and early adolescence: a high-density electrical mapping study. *Cereb Cortex*. 2011; 21:1042–1055. [PubMed: 20847153]
- Butler JS, Molholm S, Fiebelkorn IC, Mercier MR, Schwartz TH, Foxe JJ. Common or redundant neural circuits for duration processing across audition and touch. *J Neurosci*. 2011; 31:3400–3406. [PubMed: 21368051]
- Calvert GA. Crossmodal processing in the human brain: insights from functional neuroimaging studies. *Cereb Cortex*. 2001; 11:1110–1123. [PubMed: 11709482]
- Canolty RT, Edwards E, Dalal SS, Soltani M, Nagarajan SS, Kirsch HE, Berger MS, Barbaro NM, Knight RT. High gamma power is phase-locked to theta oscillations in human neocortex. *Science*. 2006; 313:1626–1628. [PubMed: 16973878]
- Canolty RT, Knight RT. The functional role of cross-frequency coupling. *Trends Cogn Sci*. 2010; 14:506–515. [PubMed: 20932795]
- Cappe C, Barone P. Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. *Eur J Neurosci*. 2005; 22:2886–2902. [PubMed: 16324124]
- Clavagnier S, Falchier A, Kennedy H. Long-distance feedback projections to area V1: implications for multisensory integration, spatial awareness, and visual consciousness. *Cogn Affect Behav Neurosci*. 2004; 4:117–126. [PubMed: 15460918]
- Clemo HR, Allman BL, Donlan MA, Meredith MA. Sensory and multisensory representations within the cat rostral suprasylvian cortex. *J Comp Neurol*. 2007; 503:110–127. [PubMed: 17480013]
- Delorme A. Statistical methods. In: Webster, JG., editor. *Encyclopedia of Medical Device and Instrumentation*. Wiley interscience; 2006. p. 240-264.
- Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods*. 2004; 134:9–21. [PubMed: 15102499]
- Diederich A, Schomburg A, Colonius H. Saccadic reaction times to audiovisual stimuli show effects of oscillatory phase reset. *PLoS One*. 2012; 7:e44910. [PubMed: 23056186]
- Engel AK, Fries P. Beta-band oscillations--signalling the status quo? *Curr Opin Neurobiol*. 2010; 20:156–165. [PubMed: 20359884]
- Falchier A, Clavagnier S, Barone P, Kennedy H. Anatomical evidence of multimodal integration in primate striate cortex. *J Neurosci*. 2002; 22:5749–5759. [PubMed: 12097528]
- Fiebelkorn IC, Foxe JJ, Butler JS, Mercier MR, Snyder AC, Molholm S. Ready, set, reset: stimulus-locked periodicity in behavioral performance demonstrates the consequences of cross-sensory phase reset. *J Neurosci*. 2011; 31:9971–9981. [PubMed: 21734288]
- Fiebelkorn IC, Snyder AC, Mercier MR, Butler JS, Molholm S, Foxe JJ. Cortical cross-frequency coupling predicts perceptual outcomes. *Neuroimage*. 2013; 69:126–137. [PubMed: 23186917]
- Fort A, Delpuech C, Pernier J, Giard MH. Early auditory-visual interactions in human cortex during nonredundant target identification. *Brain Res Cogn Brain Res*. 2002; 14:20–30. [PubMed: 12063127]
- Foxe JJ, Morocz IA, Murray MM, Higgins BA, Javitt DC, Schroeder CE. Multisensory auditory-somatosensory interactions in early cortical processing revealed by high-density electrical mapping. *Brain Res Cogn Brain Res*. 2000; 10:77–83. [PubMed: 10978694]
- Foxe JJ, Schroeder CE. The case for feedforward multisensory convergence during early cortical processing. *Neuroreport*. 2005; 16:419–423. [PubMed: 15770144]

- Foxe JJ, Wylie GR, Martinez A, Schroeder CE, Javitt DC, Guilfoyle D, Ritter W, Murray MM. Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. *J Neurophysiol.* 2002; 88:540–543. [PubMed: 12091578]
- Genovese CR, Lazar NA, Nichols T. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage.* 2002; 15:870–878. [PubMed: 11906227]
- Giard MH, Peronnet F. Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J Cogn Neurosci.* 1999; 11:473–490. [PubMed: 10511637]
- Gingras G, Rowland BA, Stein BE. The differing impact of multisensory and unisensory integration on behavior. *J Neurosci.* 2009; 29:4897–4902. [PubMed: 19369558]
- Gomez-Ramirez M, Kelly SP, Molholm S, Sehatpour P, Schwartz TH, Foxe JJ. Oscillatory Sensory Selection Mechanisms during Intersensory Attention to Rhythmic Auditory and Visual Inputs: A Human Electroencephalographic Investigation. *Journal of Neuroscience.* 2011a; 31:18556–18567. [PubMed: 22171054]
- Gomez-Ramirez M, Kelly SP, Molholm S, Sehatpour P, Schwartz TH, Foxe JJ. Oscillatory sensory selection mechanisms during intersensory attention to rhythmic auditory and visual inputs: a human electroencephalographic investigation. *J Neurosci.* 2011b; 31:18556–18567. [PubMed: 22171054]
- Gray CM, König P, Engel AK, Singer W. Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature.* 1989; 338:334–337. [PubMed: 2922061]
- Groppe DM, Urbach TP, Kutas M. Mass univariate analysis of event-related brain potentials/fields I: a critical tutorial review. *Psychophysiology.* 2011; 48:1711–1725. [PubMed: 21895683]
- Harrison NR, Wuerger SM, Meyer GF. Reaction time facilitation for horizontally moving auditory-visual stimuli. *J Vis.* 2010; 10:16. [PubMed: 21163957]
- Jensen O, Colgin LL. Cross-frequency coupling between neuronal oscillations. *Trends Cogn Sci.* 2007; 11:267–269. [PubMed: 17548233]
- Jones EG. The thalamic matrix and thalamocortical synchrony. *Trends Neurosci.* 2001; 24:595–601. [PubMed: 11576674]
- Kayser C, Petkov CI, Augath M, Logothetis NK. Functional imaging reveals visual modulation of specific fields in auditory cortex. *J Neurosci.* 2007; 27:1824–1835. [PubMed: 17314280]
- Kayser C, Petkov CI, Logothetis NK. Visual modulation of neurons in auditory cortex. *Cereb Cortex.* 2008; 18:1560–1574. [PubMed: 18180245]
- Kopell N, Ermentrout GB, Whittington MA, Traub RD. Gamma rhythms and beta rhythms have different synchronization properties. *Proc Natl Acad Sci U S A.* 2000; 97:1867–1872. [PubMed: 10677548]
- Lakatos P, Chen CM, O'Connell MN, Mills A, Schroeder CE. Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron.* 2007; 53:279–292. [PubMed: 17224408]
- Lakatos P, O'Connell MN, Barczak A, Mills A, Javitt DC, Schroeder CE. The Leading Sense: Supramodal Control of Neurophysiological Context by Attention. *Neuron.* 2009; 64:419–430. [PubMed: 19914189]
- Lakatos P, Shah AS, Knuth KH, Ulbert I, Karmos G, Schroeder CE. An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *J Neurophysiol.* 2005; 94:1904–1911. [PubMed: 15901760]
- Lancaster JL, Woldorff MG, Parsons LM, Liotti M, Freitas CS, Rainey L, Kochunov PV, Nickerson D, Mikiten SA, Fox PT. Automated Talairach atlas labels for functional brain mapping. *Hum Brain Mapp.* 2000; 10:120–131. [PubMed: 10912591]
- Leitao J, Thielscher A, Werner S, Pohmann R, Noppeney U. Effects of Parietal TMS on Visual and Auditory Processing at the Primary Cortical Level - A Concurrent TMS-fMRI Study. *Cereb Cortex.* 2012
- Makeig S, Debener S, Onton J, Delorme A. Mining event-related brain dynamics. *Trends Cogn Sci.* 2004; 8:204–210. [PubMed: 15120678]
- Makeig S, Westerfield M, Jung TP, Enghoff S, Townsend J, Courchesne E, Sejnowski TJ. Dynamic brain sources of visual evoked responses. *Science.* 2002; 295:690–694. [PubMed: 11809976]

- Meredith MA, Allman BL. Subthreshold multisensory processing in cat auditory cortex. *Neuroreport*. 2009; 20:126–131. [PubMed: 19057421]
- Meredith MA, Allman BL, Keniston LP, Clemo HR. Auditory influences on non-auditory cortices. *Hear Res*. 2009; 258:64–71. [PubMed: 19303926]
- Meredith MA, Stein BE. Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *J Neurophysiol*. 1986; 56:640–662. [PubMed: 3537225]
- Miller J. Divided attention: evidence for coactivation with redundant signals. *Cogn Psychol*. 1982; 14:247–279. [PubMed: 7083803]
- Mishra J, Martinez A, Sejnowski TJ, Hillyard SA. Early cross-modal interactions in auditory and visual cortex underlie a sound-induced visual illusion. *J Neurosci*. 2007; 27:4120–4131. [PubMed: 17428990]
- Molholm S, Ritter W, Murray MM, Javitt DC, Schroeder CE, Foxe JJ. Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Brain Res Cogn Brain Res*. 2002; 14:115–128. [PubMed: 12063135]
- Molholm S, Sehatpour P, Mehta AD, Shpaner M, Gomez-Ramirez M, Ortigue S, Dyke JP, Schwartz TH, Foxe JJ. Audio-Visual Multisensory Integration in Superior Parietal Lobule Revealed by Human Intracranial Recordings. *J Neurophysiol*. 2006; 96:721–729. [PubMed: 16687619]
- Moran RJ, Molholm S, Reilly RB, Foxe JJ. Changes in effective connectivity of human superior parietal lobule under multisensory and unisensory stimulation. *Eur J Neurosci*. 2008; 27:2303–2312. [PubMed: 18445221]
- Murray MM, Molholm S, Michel CM, Heslenfeld DJ, Ritter W, Javitt DC, Schroeder CE, Foxe JJ. Grabbing your ear: rapid auditory-somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cereb Cortex*. 2005; 15:963–974. [PubMed: 15537674]
- Naue N, Rach S, Struber D, Huster RJ, Zaehle T, Korner U, Herrmann CS. Auditory event-related response in visual cortex modulates subsequent visual responses in humans. *J Neurosci*. 2011; 31:7729–7736. [PubMed: 21613485]
- Neal JW, Pearson RC, Powell TP. The connections of area PG, 7a, with cortex in the parietal, occipital and temporal lobes of the monkey. *Brain Res*. 1990; 532:249–264. [PubMed: 2282518]
- Oostenveld R, Fries P, Maris E, Schoffelen JM. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci*. 2011; 2011:156869. [PubMed: 21253357]
- Penny WD, Kiebel SJ, Kilner JM, Rugg MD. Event-related brain dynamics. *Trends Neurosci*. 2002; 25:387–389. [PubMed: 12127748]
- Perrin F, Pernier J, Bertrand O, Giard MH, Echallier JF. Mapping of scalp potentials by surface spline interpolation. *Electroencephalogr Clin Neurophysiol*. 1987; 66:75–81. [PubMed: 2431869]
- Raij T, Ahveninen J, Lin FH, Witzel T, Jaaskelainen IP, Letham B, Israeli E, Sahyoun C, Vasios C, Stufflebeam S, Hamalainen M, Belliveau JW. Onset timing of cross-sensory activations and multisensory interactions in auditory and visual sensory cortices. *Eur J Neurosci*. 2010; 31:1772–1782. [PubMed: 20584181]
- Rasch MJ, Gretton A, Murayama Y, Maass W, Logothetis NK. Inferring spike trains from local field potentials. *J Neurophysiol*. 2008; 99:1461–1476. [PubMed: 18160425]
- Roach BJ, Mathalon DH. Event-related EEG time-frequency analysis: an overview of measures and an analysis of early gamma band phase locking in schizophrenia. *Schizophr Bull*. 2008; 34:907–926. [PubMed: 18684772]
- Rockland KS, Ojima H. Multisensory convergence in calcarine visual areas in macaque monkey. *Int J Psychophysiol*. 2003; 50:19–26. [PubMed: 14511833]
- Rorden C, Brett M. Stereotaxic display of brain lesions. *Behavioural Neurology*. 2000; 12:191–200. [PubMed: 11568431]
- Schroeder CE, Foxe J. Multisensory contributions to low-level, ‘unisensory’ processing. *Curr Opin Neurobiol*. 2005; 15:454–458. [PubMed: 16019202]
- Schroeder CE, Foxe JJ. The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. *Brain Res Cogn Brain Res*. 2002; 14:187–198. [PubMed: 12063142]

- Schroeder CE, Lakatos P. Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci.* 2009; 32:9–18. [PubMed: 19012975]
- Seltzer B, Pandya DN. Parietal, temporal, and occipital projections to cortex of the superior temporal sulcus in the rhesus monkey: a retrograde tracer study. *J Comp Neurol.* 1994; 343:445–463. [PubMed: 8027452]
- Senkowski D, Gomez-Ramirez M, Lakatos P, Wylie GR, Molholm S, Schroeder CE, Foxe JJ. Multisensory processing and oscillatory activity: analyzing non-linear electrophysiological measures in humans and simians. *Exp Brain Res.* 2007; 177:184–195. [PubMed: 17051387]
- Senkowski D, Molholm S, Gomez-Ramirez M, Foxe JJ. Oscillatory beta activity predicts response speed during a multisensory audiovisual reaction time task: a high-density electrical mapping study. *Cereb Cortex.* 2006; 16:1556–1565. [PubMed: 16357336]
- Senkowski D, Saint-Amour D, Hofle M, Foxe JJ. Multisensory interactions in early evoked brain activity follow the principle of inverse effectiveness. *Neuroimage.* 2011; 56:2200–2208. [PubMed: 21497200]
- Shah AS, Bressler SL, Knuth KH, Ding M, Mehta AD, Ulbert I, Schroeder CE. Neural dynamics and the fundamental mechanisms of event-related brain potentials. *Cereb Cortex.* 2004; 14:476–483. [PubMed: 15054063]
- Sherman SM. The thalamus is more than just a relay. *Curr Opin Neurobiol.* 2007; 17:417–422. [PubMed: 17707635]
- Sherman SM, Guillery RW. The role of the thalamus in the flow of information to the cortex. *Philos Trans R Soc Lond B Biol Sci.* 2002; 357:1695–1708. [PubMed: 12626004]
- Stanford TR, Quessy S, Stein BE. Evaluating the operations underlying multisensory integration in the cat superior colliculus. *J Neurosci.* 2005; 25:6499–6508. [PubMed: 16014711]
- Stein BE. Neural mechanisms for synthesizing sensory information and producing adaptive behaviors. *Exp Brain Res.* 1998; 123:124–135. [PubMed: 9835401]
- Stein, BE.; Meredith, MA. *The merging of the senses.* MIT Press; Cambridge, Mass: 1993.
- Tallon-Baudry C, Bertrand O, Delpuech C, Pernier J. Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *J Neurosci.* 1996; 16:4240–4249. [PubMed: 8753885]
- Teder-Salejarvi WA, McDonald JJ, Di Russo F, Hillyard SA. An analysis of audio-visual crossmodal integration by means of event-related potential (ERP) recordings. *Brain Res Cogn Brain Res.* 2002; 14:106–114. [PubMed: 12063134]
- Thorne JD, De Vos M, Viola FC, Debener S. Cross-modal phase reset predicts auditory task performance in humans. *J Neurosci.* 2011; 31:3853–3861. [PubMed: 21389240]
- Tyll S, Bonath B, Schoenfeld MA, Heinze HJ, Ohl FW, Noesselt T. Neural basis of multisensory looming signals. *Neuroimage.* 2012
- Vidal JR, Ossandon T, Jerbi K, Dalal SS, Minotti L, Ryvlin P, Kahane P, Lachaux JP. Category-Specific Visual Responses: An Intracranial Study Comparing Gamma, Beta, Alpha, and ERP Response Selectivity. *Front Hum Neurosci.* 2010; 4:195. [PubMed: 21267419]
- von Stein A, Rappelsberger P, Sarnthein J, Petsche H. Synchronization between temporal and parietal cortex during multimodal object processing in man. *Cereb Cortex.* 1999; 9:137–150. [PubMed: 10220226]
- Werner S, Noppeney U. Distinct functional contributions of primary sensory and association areas to audiovisual integration in object categorization. *J Neurosci.* 2010; 30:2662–2675. [PubMed: 20164350]
- Whittingstall K, Logothetis NK. Frequency-band coupling in surface EEG reflects spiking activity in monkey visual cortex. *Neuron.* 2009; 64:281–289. [PubMed: 19874794]

Highlights

Auditory driven phase reset occurs over visual cortices and can lead to auditory evoked potential.

Multisensory interactions occur extensively in visual cortices.

In visual regions, auditory phase resetting interacts with the evoked visual activity.

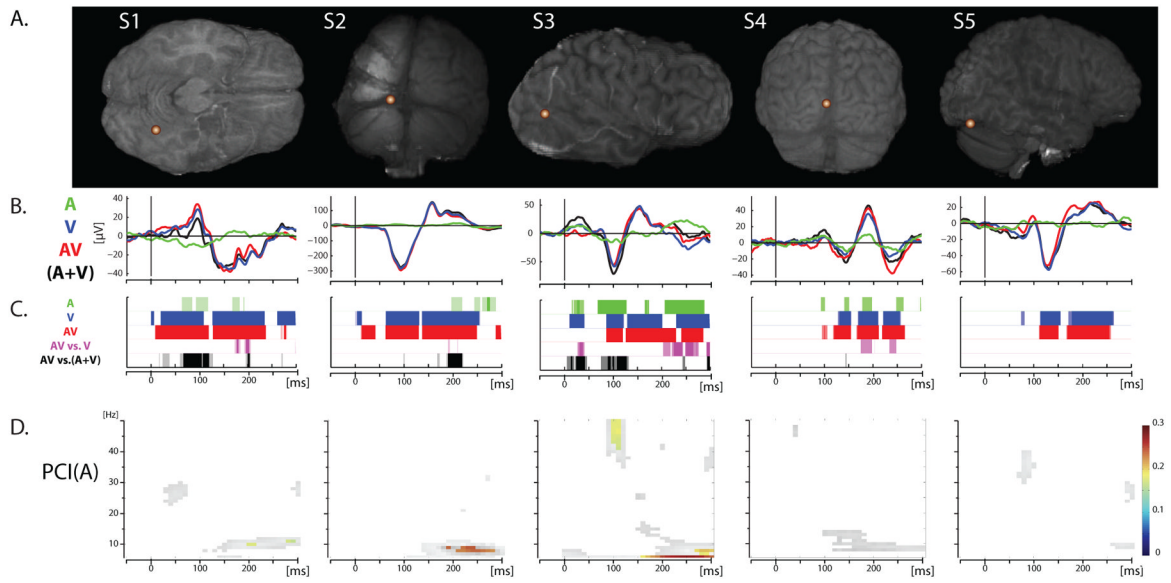


Figure 1.

ERP waveforms exemplars. **(A)** COI location for each of the five different participants brains, with COIs selected from different regions of visual cortex, representing a range of observed responses. **(B)** ERP responses for auditory-alone (A; Green), visual-alone (V; Blue) and audio-visual (AV; Red) conditions. The summed response (the sum of the unisensory responses (A+V)) is depicted with a black line. **(C)** Statistical results representing comparison of each condition against baseline and between conditions following the *maximum* model (AV vs. V; magenta) and the *additive* model (AV vs. (A+V); black). Corrected significant p-values are represented as solid color, uncorrected p-values appear transparent. **(D)** Statistics on Phase Consistency Index (PCI) are depicted for the auditory-alone condition. Only significant values are shown, in color after correction, and in black and white without the multiple comparison correction.

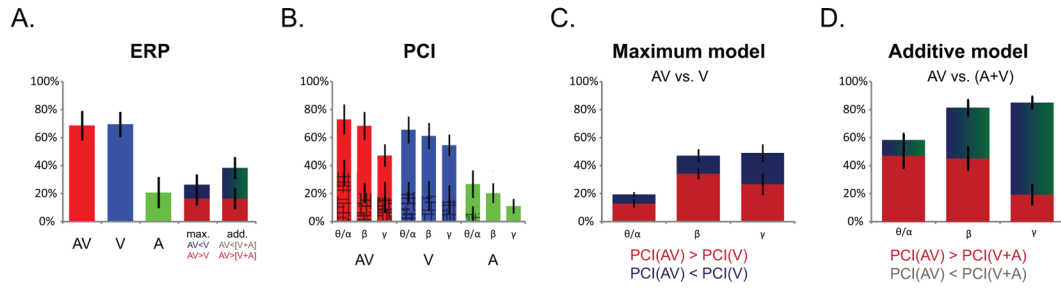
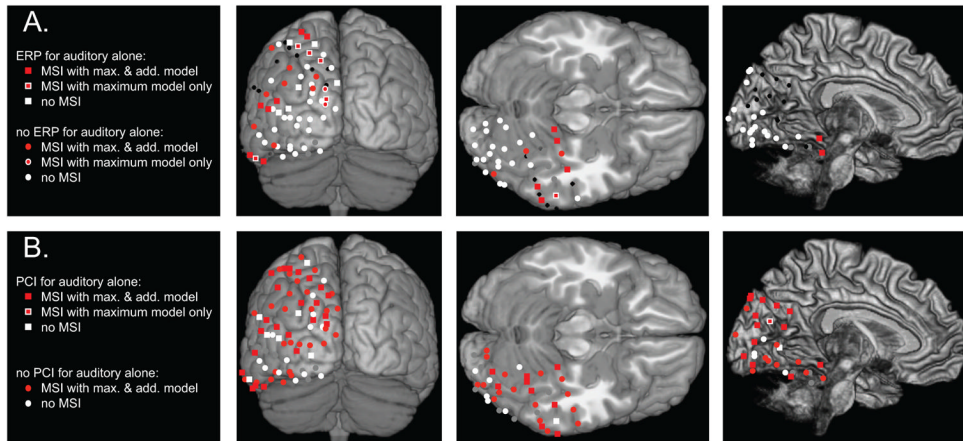


Figure 2.

Bar graphs depicting the average proportion (across participants) of COIs ($n = 108$) for which a given effect was observed. **(A)** ERPs: percentage of COIs showing a significant response (compared to baseline) for Audio-Visual (AV-red), Visual-alone (V-blue) and Auditory-alone (A-green) conditions, and showing a significant multisensory effect assessed using the maximum model criterion (AV vs. V ; with dark red representing enhanced effect and dark blue a suppression effect), and the additive model criterion (AV vs. A+V ; with dark red representing supra-additive effect and dark blue-green a sub-additive effect). **(B)** PCIs: percentage of COIs showing a significant difference from baseline for Audio-Visual (AV-red), Visual-alone (V-blue) and Auditory-alone (A-green) conditions. Results for Theta-Alpha (6–10Hz), Beta (13–30Hz) and Gamma (30–50Hz) bands are depicted separately. The patched area represents co-occurring significant power modulations compared to baseline. **(C)** PCIs: percentage of COIs showing a significant auditory modulation of the visual response assessed using the the maximum model (AV vs. V ; with dark red representing enhanced effect, and dark blue a suppression effect). **(D)** PCIs: percentage of selected COIs ($n = 90$; that presented cross-sensory auditory response and/or auditory modulation of the visual response as assessed by the max model) showing non-linear MSI effects as measured by the additive model (AV vs. A+V ; with dark red representing a supra-additive effect, and dark blue-green a sub-additive effect). For both (C) and (D), the results of Theta-Alpha (6–10Hz), Beta (13–30Hz) and Gamma (30–50Hz) bands are depicted separately.

**Figure 3.**

Summary of responses observed at the COIs for all participants, projected on the MNI brain, with right site contacts collapsed to the left hemisphere, from three views: posterior, ventral (cerebellum trimmed), and mesial (cortical surface squeezed). **(A)** ERP results: for each COI the shape indicates if an ERP was observed for auditory-alone stimuli (square) or not (circle) ; color code for presence (red) or absence (white) of multisensory integration, MSI being assessed using Maximum model (center) and Additive model (outline). COIs without any ERP response or contaminated by artifacts are depicted respectively in black and grey. **(B)** Phase resetting results: shape indicates if the phase of ongoing oscillations was reset by auditory-alone stimuli (square) or not (circle) ; color code for presence (red) or absence (white) of multisensory integration effects on phase, MSI being assessed using Maximum model (center) and Additive model (outline). COIs contaminated by artifacts are depicted in grey.

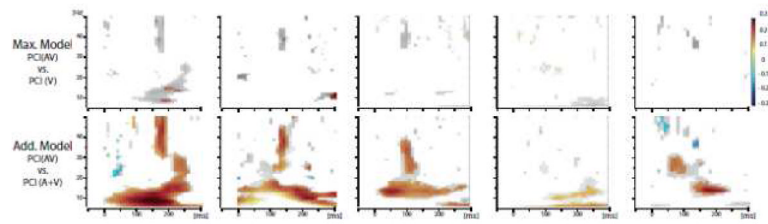


Figure 4.

Example statistic for multisensory effects observed on Phase Consistency Index (PCI), for the five same contacts depicted in Figure 1. Top panel: the maximum criterion model. Bottom panel: the additive criterion model. Only significant values are shown, in color after correction, in black and white without correction.