Center for Mind/Brain Sciences – University of Trento
Doctoral School in Cognitive and Brain Sciences

Investigating multisensory integration in human early visual and auditory areas with intracranial electrophysiological recordings: insights and perspectives

Phd student: Stefania Ferraro

Advisor: Prof. Dr. Olivier Collignon

XXVIII Cycle-December 2016

Index

ABSTRACT

Cross-modal processing and multisensory integration (MSI) can be observed at early stages of sensory processing in the cortex. However, the neurophysiological mechanisms underlying these processes and how they vary across sensory systems remain elusive. The aim of this study was to investigate how cross-modal processing and MSI are reflected in power and phase of oscillatory neuronal activity at different temporal scales in different sensory cortices. To this goal, we recorded stereo-electroencephalographic (SEEG) responses from early visual (calcarine and pericalcarine) and auditory (Heschl's gyrus and planum temporale) regions in patients with drug-resistant epilepsy while performing an audio-visual oddball task.

To Investigate crossmodal processing and MSI in the power domain of oscillatory activity, we explored a wide range of frequency bands (theta/alpha band: 5-13Hz; beta band: 13-30 Hz; gamma band: 30-80 Hz; high-gamma band: 80-200 Hz) during the first 150 ms post-stimulus onset. Differently, to investigate crossmodal processing and MSI in the phase domain of oscillatory activity, we explored a narrow range of frequency bands (theta/alpha band: 5-13Hz; beta band: 13-30 Hz; gamma band: 30-80 Hz) during the first 300 ms post-stimulus onset.

In the power domain, we showed that cross-modal processing occurs mainly in the high-gamma band (80-200Hz) in both cortices. However, we evidenced that the way MSI is expressed across modalities differs considerably: in the visual cortex, MSI relies mainly on the beta band, however it is also evident, to a lesser extent, in the gamma and high-gamma band, while the auditory cortex reveals widespread MSI in the high-gamma band and, to a lesser extent, across the gamma band and the other investigated frequency bands.

In the phase domain, we showed that cross-modal processing is differently expressed across modalities: in the auditory cortex it induces an increased phase concentration index (PCI) in ongoing oscillatory activity across all the investigated frequency bands, while, in the visual cortex,

it induces an increased PCI particularly evident in the theta/alpha band with few or no effect respectively in the gamma and beta band. Importantly in both cortices, the most part of the COIs showing increased PCI, were not accompanied by a concomitant increase in power. These results indicate that in both auditory and visual cortex, cross-modal processing induces a pure phase resetting of the oscillatory activity. During MSI processing we observed, in both cortices, a stronger increase in PCI, in comparison to the intramodal processing, in the theta/alpha band and in the gamma band.

Our results confirm the presence of cross-modal information representations at neuronal populations level and conform to a model where the cross-modal input induces phase-locked modulation of the ongoing oscillations. Importantly, our data showed that the way MSI is expressed in power modulations differs between the investigated sensory cortices suggesting the presence of different types of neurophysiological interactions during this process. These results are discussed in the framework of the current literature.

1. Introduction

Multiple sensory systems process the flow of information present in the environment: to create coherent perceptual representations, the brain must bind the related information and segregate the unrelated ones (Stein and Meredith, 1993). This ability to rapidly and seamlessly integrate information coded by apparently separate sensory systems, defined multisensory integration (MSI), is a fundamental perceptual function at the basis of adaptive behaviours. Where and how the brain integrates multisensory signals is matter of intense research.

The classical basic tenet of neocortical organization is that different regions of the cortex separately receive visual, auditory, tactile, olfactory, and gustatory sensations. These sensory regions are thought to be independent from each other and are believed to send projections that converge on 'association areas' (e.g. parietal), which then enable the binding between the different senses. However, starting from the end of the 90's, new studies produced provocative findings that overturned this hierarchical model. Functional imaging (G.A. Calvert 2001; Macaluso 2000; Martuzzi et al. 2007; Kayser et al. 2005; Foxe et al. 2002) and electrophysiological studies in animals and humans (Giard & Peronnet 1999; Foxe et al. 2000; Gomez-ramirez et al. 2011; Molholm et al. 2002; Raij et al. 2010; Kayser et al. 2008; Lakatos, Chen, et al. 2007; Schroeder & Foxe 2002; Brosch et al. 2005; Ghazanfar 2005; Murray et al. 2004) provided evidence for multisensory interactions already at the level of early sensory cortices.

These studies represent a change of scientific paradigm in cognitive neuroscience: they unhinged the idea that primary sensory areas are exclusively sensitive to sensory input from one modality only and led to the hypothesis that multisensory integration (MSI) is present in almost all the neocortex, including the earliest cortical stages of sensory processing (Ghazanfar & Schroeder, 2006; Murray et al., 2016).

Even if this new conceptualization of how the senses interact is a pervasive breakthrough in the literature, the debate about the presence of MSI in early sensory regions (Kayser 2010) as well as the potential underlying computations and mechanisms is still far from being resolved (Schroeder & Lakatos 2009; Mercier et al. 2013; Kayser et al. 2009). Recent findings have just begun to reveal the complexity and the heterogeneity of early MSI mechanisms/effects (Iurilli et al. 2012; Lakatos et al. 2007; Mercier et al. 2015; Mercier et al. 2013); moreover some studies have failed to observe MSI in early sensory regions (Lemus et al. 2010; Quinn et al. 2014), further challenging this proposal.

In the first part of this work, I will give a broad overview of the studies investigating neural MSI and I will contextualize our research. I will start from the origins, presenting the earliest works of the 60's showing evidence of crossmodal convergence in putative primary sensory cortices, I will introduce the hierarchical model (Felleman & Van Essen 1991) and then I will present the studies that allowed the change of scientific paradigm showing the presence of MSI in early sensory regions. A section will be dedicated to the seminal works of Stein & Meredith (1993) who empirically defined the principles and integrative rules of MSI in neurons of the superior colliculus of anesthetized cat that heavily influenced the subsequent literature. Then, I will provide an overview of animal studies supporting the presence of anatomical connectivity between early sensory areas and I will describe the most recent works that have identified possible neural mechanisms of cross-modal processing and MSI. To contextualize our work, I will introduce the framework of the spectral fingerprints hypothesis and I will discuss the study of oscillations in the time-frequency domain as a means to reveal crossmodal processing and MSI.

In the second part of the work, I will present our investigation conducted with stereo-electroencephalogram (SEEG) in early human visual and auditory areas in 8 chronic epileptic patients. In the framework of the spectral fingerprints hypothesis, I will provide evidence that that

crossmodal processing and MSI are present in early visual and auditory areas across different temporal scales in both power and phase, suggesting the presence of multiple mechanisms at the basis of these processes.

I want to highlight here that the main strength of this work is the used technique (SEEG): indeed in order to unequivocally demonstrate early MSI, one has to show that it occurs in primary sensory regions at an early timing after the presentation of the event. However, the main techniques used in cognitive neurosciences in humans cannot separately disambiguate this question. For example, the hemodynamic nature of the functional magnetic resonance imaging (fMRI) signal is inadequate to capture fast dynamic neural processes and methods measuring human brain activity with temporal resolution in the range of neural dynamics, such as electroencephalography (EEG) do not present adequate spatial resolution. SEEG is characterized by high spatial and temporal resolution and, notably, it allows to record high-frequency oscillations, index of local cortical processing, not detectable with EEG/MEG investigations (Lachaux et al. 2012).

1.1. A brief history of multisensory research

1.1.1. Origins of multisensory research

Multisensory research has a long tradition with the first studies dating back to the last years of the 19th century. One of the earliest study that reported crossmodal effects was published in 1896: Stratton, investigating if the inverted retinal image was a prerequisite of seeing our perceptual world in the upright position, reported the effect of vision-distorting prism glasses on somatosensory perception. Several years later the neuroscientific community developed a great interest for sensory integration of inputs from different modalities: Adrian (1949), Nobel Prize with Sir Sherrington in 1932 for his studies on neurons, highlighted that MSI is a fundamental aspect of cognition. Along these lines, Gonzalo (1952) proposed a coherent theoretical framework of the sensory organization based on the observations of polysensorial syndromes in patients with parieto-occipital cortical lesions. Based on the concept of '*sistema de gradientes'*, he proposed that the more central was the lesion, the more the risk to present a polysensorial syndrome.

Subsequently, between the '50s and the '70s, several animal studies evidenced cross-modal responses in sensory cortices. Bremar (1952), for the first time, revealed acoustic evoked potentials in the occipital cortex. Then, other investigations showed that acoustical, vestibular and tactile stimuli modulated the activity of visual neurons (Hirsch et al. 1961; Kornhuber & Da Fonseca 1964). Interestingly, Murata et al. (1965) showed the presence, in the primary visual cortex, not only of bisensory neurons but also of trisensory neurons, with the convergence of visual, auditory and tactile stimuli. The important debate at that time was if the observed activity in the visual cortex was due to the effect of a generalized arousal, or was a pure cross-sensory effect (Morrell 1972). In the following years it was clear that the cross-modal activity in the visual cortex was not a generalized effect. Spinelli (1968) showed that some neurons of the visual cortex

were specifically activated by selected frequencies of the sound input, providing therefore the first evidence of specific activity of the visual cortex during cross-modal stimulation processes. Morrell (1972) provided additional evidence of visual neurons showing activity during processing of auditory inputs in specific location, indicating the presence of a possible auditory-spatial map in the visual cortex. The possibility that these audio-visual neurons in the visual cortex played some role in the spatial localization of the auditory and visual stimuli was reinforced by the study of Fishman & Michael, (1973), who also showed that bimodal cells were present in distinct anatomical cluster segregated from visual cells. More recently, specific and non-specific modulatory effects were observed for vestibular stimulation mainly in regions of the primary visual cortex representing the peripheral visual fields (Vanni-Mercier & Magnin 1982).

After these twenty and more years of blooming period of studies investigating crossmodal processing in sensory areas, the interest for this important feature of the neocortex was strongly reduced and in the 90's it disappeared from the scientific debate, opening the avenue for the 'hierarchical model' (Felleman & Van Essen, 1991). According to Ghazanfar et al. (2005), the 'hierarchical model' found its roots in early animal studies showing few or no evidence of anatomical connections among different sensory cortices (Kuypers et al. 1965), and in experimental studies showing that lesions of specific sensory regions produced unimodal deficits (Massopust & Barnes 1965). However, it would be reductive not to appreciate in the success of the 'hierarchical model' the part played by the theoretical proposal of the 'modularity of mind' proposed by Fodor (1983), some years before. Indeed domain specificity and informational encapsulation, two cardinal aspects of this new conceptualization of cognition, surely were the scaffolding of the sensory modalities intended as pure unisensory areas characterized by a specific sensory domain and a strong segregation of information. Interestingly, in the same years in which the hierarchical model was defined in its main aspects, Stein & Meredith (1993) provided seminal

studies revealing the computations and principles of MSI at neuronal level in the superior colliculus of anesthetized cats. These studies heavily influenced the subsequent studies in MSI field: the scientific community, assuming that these principles and integrative models have important heuristic values also for data obtained with population-based techniques (such as PET, fMRI, EEG, MEG), extensively used these empirical observations to investigate neural MSI in the neocortex. It is now important to give a brief overview of the 'hierarchical model' (Felleman & Van Essen 1991) and of the principles and integrative operations of MSI proposed by Stein & Meredith (Stein & Meredith 1993) to contextualize the MSI studies in early sensory areas.

1.1.2. The hierarchical model

Hubel and Wiesel (Hubel & Wiesel 1962) hypothesized the first serial feedforward hierarchical model of the brain: they explained the increasing complexity of the characteristics of the visual receptive fields with the proposal that in the visual cortex information processing occurred in a hierarchical manner (1962) with a serial and feedforward flow of information. Some years later, the observations of reciprocal and parallel anatomical pathways in the visual cortex (Lennie, 1980; Rockland & Pandya, 1979; Shapley, 1990) enriched this simple model. Indeed, since the first anatomical tract-tracing studies (Rockland & Pandya, 1979), a coherent pattern of anatomical connections was identified in animal visual cortex. These coherent results showed that parent neurons of rostral directed connections were mainly located in the supragranular layer and projected with their axon terminals in layer 4 of the target region; while, parent neurons of caudal directed connections were mainly located in the infragranular layers and projected outside the layer 4 of their target regions (Kuypers et al. 1965; Felleman & Burkhalter 1997; Colby & Duhamel 1991; Weller & Kaas 1985; Maunsell & Essen 1983; Spatz et al. 1970; Cragg 1969). Based on the previous knowledge of the thalamic-cortical connections, the rostral directed

pathways were assumed to channel information from lower to higher-order cortical areas and were coined feedforward connections, while the caudal directed connections were assumed to channel information from higher-order to lower-order cortical areas and were labeled feedback connections (Rockland & Pandya, 1979). Importantly, the different laminar patterns of feedback and feedforward connections were showed to determine different types of interareal interactions: from the functional point of view, it was showed that feedforward signals generate receptive field properties, while feedback signals have a modulatory effect on the target regions (Hupe et al. 1998). It is important to note, however, that recent works have evidenced several inconsistencies in this simple categorization (for a review see Markov & Kennedy 2013).

In their seminal work, Felleman & Van Essen (1991) showed that the primate visual areas presented a strong anatomical hierarchical organization, therefore refining the initial proposal of Hubel & Wiesel (1962). Based on the laminar patterns of parent neurons and of axon terminals, they classified the directionality of cortico-cortical connections and ranked each one of 32 visual regions on one level (out of 11) using a pair-wise comparison of feedback and feedforward connections between those areas. These authors provided evidence that the visual system presented extensive parallel pathways supporting a strong segregation of information streams and cross-talk among different stages of visual computation. Interestingly, the authors, highlighting that their model was grounded in anatomy, stated:

'…the hierarchical scheme for visual cortex that we have presented is grounded explicitly on anatomical criteria. Whether each level of the hierarchy represents a distinct stage of information processing is a separate issue that must be addressed mainly by physiological and behavioral approaches'.

Notably, although the general implant of their work was not rejected, a posterior quantitative analysis of their data showed the indeterminateness of the proposed model (up to 150.000 solutions), mainly because Felleman & Van Essen (1991) did not consider the anatomical distances separating the different stages of processing (Hilgetag et al. 1996; Hilgetag et al. 2000). An energetic debate raised from this dispute: I quote here the 'Summary' published on *Science* relative the work of Hilgetag et al. (1996).

'The classic view of how the brain areas that control vision are connected is a complicated wiring diagram devised by manual sorting on the basis of existing anatomical data (D. J. Felleman and D. C. Van Essen, (1991). Now, in this issue's 'Enhanced Perspective', Hilgetag and co-workers have used a computer algorithm to test whether there is a better way to organize the connections. They find that the brain is surprisingly indeterminate, and that no single hierarchy can satisfactorily represent the order implied by the anatomical data'.

Based on the observations of Hilgetag et al. (1996), Barone et al. (2000) proposed to investigate the hierarchical organization of the visual cortex using as an index of the hierarchical distance such as the proportion of supragranular layer neurons involved in feedback and feedforward connections (SLN model). The rational for this was related to the fact that long-distance feedforward connections were showed to originate only in supragranular layers, while approaching the target region, were showed to originate also from infragranular layers. Similarly, the long-range feedback connections were showed to originate in infragranular layers, while approaching the target region, were showed to originate also from supragranular layers (Barone et al. 2000). Importantly, the comparison between the hierarchical model of Felleman & van Essen (1991) and the SLN model produced almost overlapping results (Markov & Kennedy 2013; Barone

et al. 2000) except in the frontal eye fields, where the SLN model showed that this area was feedforward-connected with V4.

Despite the important knowledge gained in these years, the debate about the anatomical hierarchical organization is clearly still unresolved and the functional role of feedforward and feedback pathways is still to be elucidated in its fundamental aspects (Markov & Kennedy 2013). Consequently, we still lack of an operational definition of higher-order and lower-order areas (Markov et al. 2014; Markov & Kennedy 2013). It is clear that in the above context, the phenomenon of early cross-modal processing and MSI in putative 'unisensory areas' has profound implications: they challenge a strict interpretation of the hierarchical model of Felleman & Van Essen (1991) and offer a new perspective over the inconsistencies of this field.

1.1.3. Principles and integrative operations of multisensory integration in neurons of superior colliculus.

Since the pioneering work of Horn & Hill (1966), a long tradition of studies (Newman & Hartline 1981; McIlwain & Buser 1968; Feldon et al. 1970; King & Palmer 1985) used the animal superior colliculus as a model to investigate MSI at level of single neuron, the most basic functional unit of MSI. Indeed, the intermediate layers of superior colliculus integrate different sensory modalities and motor systems and present about 50% of multisensory neurons (Meredith & Stein 1986b). In this tradition, Stein and Meredith (1993) investigated the neurophysiology of the neurons of the superior colliculus of the anesthetized cats, and empirically identified integrative operations and principles of MSI occurring at neuronal level.

According to the their work (Stein & Meredith 1993), the presence of MSI at neuronal level can be tested by means of two models: the maximum and the additive model. The maximum model assumes that MSI is present at neuronal level if the number of impulses evoked by the

bimodal (B) condition (i.e. simultaneous presence of different modalities inputs) is statistically different from the number of impulses evoked by the most effective of the unimodal (U) conditions (i.e. a single modality input presented alone)(B vs. U_{max}). According to this model MSI can therefore result in enhanced or depressed neuron's activity. In the case of multisensory enhancement, differences between the bimodal and unimodal inputs might reflect different underlying non-linear computations: the largest enhancements are due to superadditive combinations of different modality inputs, while the smallest are due to sub-additive combinations (Stein and Stanford, 2008). To indentify these two different forms of multisensory enhancement, they proposed the additive model: the numbers of impulses evoked by each U condition are summed and then compared with the number of impulses evoked during the bimodal condition (B vs. $U_1 + U_2$). Importantly, Stein & Meredith evidenced that the most commonly bimodal responses exceeded the most effective unimodal responses showing therefore multisensory enhancement, in particular when the inputs were in spatial and temporal register, and emphasized that the integrative operations at single-neuron level were mostly non-linear combinations of the information from different modalities (Meredith et al. 1987).

However, in a recent review (Stein et al. 2011), they evidenced that the additive model presents several theoretical and empirical challenges. First of all, not only non-linear but also linear computation of cross-modal inputs might be at the basis of MSI at neuronal level; second, MSI responses are not always different from the most effective unimodal input (U_{max}), suggesting a 'maximizing' computation (Stanford & Stein 2007). Following these lines, they recommended the use of the traditional criterion (B vs. U_{max}) when investigating MSI in single-neurons responses. Notably, they suggested that this recommendation does not hold for the population-based techniques (EEG, MEG, fMRI): we will enter this topic in the next chapter.

Importantly, the MSI responses do not reflect only the computational properties of the specific neuron, but also the specific stimuli to which they respond to. Based on the neuronal responses to the properties of the stimuli, Stein & Meredith in their works (Meredith & Stein 1986a; Meredith & Stein 1986b; M. A. Meredith et al. 1987; Meredith & Stein 1996; M. Meredith et al. 1987) identified three fundamental principles that rule MSI in neurons of superior colliculus: the spatial, the temporal and the inverse effectiveness principles.

The spatial principle

Multisensory neurons in superior colliculus present specific receptive fields for each single modality to which they respond to. According to the spatial principle, different information from different modalities are integrated by the multisensory neuron if these inputs are localized in the overlapping receptive fields. Generally, the different receptive fields of each multisensory neuron are in spatial register with one another: therefore if the inputs are in spatial register, the multisensory neuron will present enhanced or depressed activity when compared to the unimodal condition, while if the inputs are in spatially different locations, the activity of multimodal neurons will not change. The spatial principle is of particular importance for superior colliculus in consideration of the capital role of this structure in determining an orienting behaviour (Stein and Stanford, 2008).

The temporal principle

Multisensory neurons in superior colliculus integrate different inputs from different modalities when they are presented within a relatively large temporal window (lasting several hundred milliseconds). Importantly, this temporal binding window allows the brain to take into account the different speeds at which signals arrive to the respective sense (i.e. retina, cochlea) and the different response latencies among senses (i.e. visual system, auditory system). The

magnitude of the spiking rates of the integrated response is sensitive to the temporal overlap of the responses and is usually maximal when the peak periods of activity coincide.

The inverse effectiveness principle

Multisensory neurons in superior colliculus integrate different inputs from different modalities in an inverse way to the effectiveness of the individual inputs that are being combined. Therefore, individual inputs eliciting robust responses will have a small effect when combined together, while individual inputs eliciting small responses, will have a great effect when combined together. In these cases the multisensory response can exceed the arithmetic sum of their individual responses (superadditive effect)(Stein et al. 2011).

1.1.4. Metrics to identify neural MSI: challenges in the use of the additive model

As we have seen in the above paragraph, Stein et al. (2011) suggested to rely on the maximum model when investigating MSI at single-neuron level. However, they recognized that the population-based techniques (EEG, MEG, fMRI) face different challenges that need to be addressed relying on the additive model. Indeed, population-based techniques have extensively used, to assess neural MSI, the additive criterion (B vs. U1+U2) (Kayser et al. 2007; Kayser et al. 2005; Beauchamp 2005; M H Giard & Peronnet 1999; G.A. Calvert 2001; G A Calvert 2001) in order to avoid to erroneously identify as MSI responses, pure unimodal responses provided from different unisensory neuronal populations during the presentation of the bimodal input (Stein et al. 2011).

However, several authors raised concerns (Populin & Yin 2002; Laurienti et al. 2005) about the use of this model derived from single neurons of the superior colliculus of anesthetized cats and argued for a more cautious approach when investing MSI with population-based techniques. Here I will discuss only the relevant problems for the electrophysiological data.

According to Besle et al. (2009), it is mandatory to use the additive criterion when investigating MSI in neurophysiological data. Indeed, due to volume conduction, an electrode records activities from distant generators. For this reason, it is not possible to compare the bimodal response with the single unimodal response (maximum criterion: B vs. U_{max}), since the bimodal response includes volume conduction effects from the two modalities. Therefore, the use of the additive criterion (B vs. U_1+U_2) allows to remove these confounding effects (Besle et al., 2009). However, using this model, possible unknown cognitive factors present in each single condition can introduce potential biases: unknown cognitive factors will be present one time in the left part of the equation, but two times in the right part of equation (B vs U_1+U_2). According to some authors, this is not a problem when studying early stages of MSI: EEG literature has shown that these cognitive not controlled activities usually begin about 200 ms post-stimulus, while earlier latencies are typical of sensory-specific responses (Hillyard et al., 1998). However, other authors suggest to insert in the paradigm to be performed a control condition (such as Null condition), which signal will be summed to the bimodal condition. This will allow to have the unknown cognitive process present two times in both terms of the equation (B+N vs U_1+U_2) (Besle et al., 2009).

Although the additive model with the correction proposed by Besle et al. (2009) is widely used in population based techniques (Quinn et al. 2014), it is important to note that, quite recently, both linear and non-linear responses were showed to be present in multisensory neurons (Perrault & Vaughan 2005). It is clear that the fact that there is no default computational mode in MSI neurons exposes the studies using the additive model to the problems of false negative. This issue is further exacerbated by the discovery of unimodal multisensory neurons (subthreshold neurons) (Stein and Stanford, 2008). These particular neurons are strongly activated by intramodal inputs, but not by cross-modal inputs; however they show modulation of their responses

(enhanced or suppressed) with multisensory inputs. Subthreshold neurons have been observed in visual (Allman and Meredith, 2007; Allman et al., 2009), somatosensory (Dehner et al., 2004; Clemo et al., 2007) and auditory (Allman et al., 2009) cortices in different species. Notably, subthreshold multisensory effects occur in up to 66% of the neuronal population in their respective areas (Dehner et al., 2004). It must be emphasized that the presence of subthreshold effects reduces or increases the response to inputs from the single modality and does not render bimodal the response distribution of the neuron. Very important, because such neurons fail to show a response to the 'subthreshold' modality when presented alone, super-additive criterion cannot verify these effects as multisensory. According to Allman and collaborators (Allman et al., 2009), these unimodal multisensory neurons are an intermediate form in a continuum from totally unisensory neurons to bimodal neurons.

These observations clearly discourage the use of the additive criterion to detect neural MSI when possible, such as with Intracranial recordings that are affected very modestly by volume conduction problems.

1.1.5. The change of paradigm

In this chapter I will present an overview of the main studies that, from the last years of the '90, have challenged the idea of purely unisensory cortex (Ghazanfar & Schroeder, 2006) and the model, at least in its strong meaning, of the hierarchical organization of the neocortex as proposed by Felleman and van Essen (1991). Importantly, I will identify the possible neurophysiological mechanisms that were proposed to explain cross-modal processing and MSI.

After the proposal of the hierarchical model of the neocortex (Felleman & Van Essen 1991), the clear first suggestions that crossmodal processing and MSI were prerogative also of early sensory areas, were mainly due to human EEG studies (Foxe 2005). Interestingly, at least two seminal fMRI investigations possibly found evidence for crossmodal processing and MSI in early sensory areas (Calvert, 1997; Macaluso, 2000), however they were interpreted or in agreement with psycholinguistic theories (Calvert 1997) or as the activity of feedback projections from higher-level cortical areas (Macaluso 2000). In particular the work of Macaluso et al. (2000) was really emblematic in understanding the very strong influence of the hierarchical model (Felleman & Van Essen 1991). The authors, investigating crossmodal spatial attention with fMRI, showed early MSI in their results: tactile stimulation indeed enhanced the activity in the visual cortex when the visual target was spatially congruent with the tactile target. However, these authors concluded, using effective connectivity based on apriori assumption, that their findings were due to feedback projections from associative (parietal) areas (Macaluso 2000).

Giard et al. (1999) provided the first EEG study suggesting that crossmodal processing and MSI occurred in early human visual cortex: they showed the presence of audio-visual interactions at very early stages of the sensory analyses (40 msec post-stimulus onset) in the occipito-parietal site. The authors observed that this activity was very similar in its topography and latency to the event related potential (ERP) components generated by the primary visual cortex when stimulated

with flashed checkerboard, therefore they argued that the origin of this effect might have been the striate cortex.

These unexpected findings (Giard & Peronnet 1999), were then strongly supported by several investigations in humans. A very similar EEG study (Molholm et al., 2002) confimed early audio-visual MSI in the parieto-occipital scalp electrodes about 40-50 ms post-stimulus onset. Foxe et al. (2000, 2002) in two different studies tested early audio-somatosensory interactions. In the first study (Foxe et al. 2000), conducted with EEG, they showed that MSI processing occurred at very early onset (from 50msec to 80msec) in central/post-central region, in areas compatible with generators located in somatosensory and auditory cortex. These authors argued that the early onset of MSI was incompatible with possible feedback processing from higher-order processing areas. In the second study, Foxe et al. (2002), possibly to solve the issue relative the localization of the sources of these early effects, took advantage of the high spatial resolution of fMRI and showed that the simultaneous auditory and somatosensory stimulation led to an increased activity in an auditory area identified as the human homologue to caudal-medial belt area of macaque. This auditory area had been just identified as the site of convergent auditory and somatosensory inputs in monkeys (Schroeder et al., 2001). Notably, Foxe et al. (2002) concluded that early feedforward MSI occurs in human auditory cortex.

Despite the important above results, the most striking evidence that MSI occurred in early sensory cortices, were provided by two important investigations on awake monkeys by the same group (Schroeder et al., 2001; Schroeder & Foxe, 2002). In the first study (Schroeder et al., 2001) the authors demonstrated that somatosensory and auditory inputs converged within the posterior auditory cortex (mainly in the caudal-medial belt), a region considered to perform early cortical processing, where the somatosensory input was unexpected (Schroeder et al., 2001); moreover, the authors showed that both auditory and somatosensory inputs in this area presented a laminar

profile (as showed by CSD and MUA) typical of feedforward processing: the activity indeed began in layer 4 and then spread to the extragranular layers at about 12ms post-stimulus onset. In the subsequent work the authors (Schroeder & Foxe, 2002), examined the laminar profile of the identified area (i.e. posterior auditory cortex) during auditory, somatosensory and visual stimuli. Interestingly, beyond replicating their previous findings (i.e. feedfoward pattern for auditory and somatosensory inputs), they showed that the visual inputs presented a feedback activity profile, with an initial response above and below layer 4.

After these first studies, the subsequent investigations were mainly focused in identifying the principles governing early cross-modal processing and MSI and the underlying neurophysiological mechanisms.

Murray et al. (2004) presented the first EEG study that explicitly tested if the spatial principle, observed at the level of the superior colliculus (Stein & Meredith 1993), hold also for neural MSI in early auditory cortex. During an audio-somatosensory task with aligned and misaligned stimuli, the authors showed that MSI occurred early in time (50-95 ms post-stimulus onset), with a topography compatible with sources in the posterior auditory cortex controlateral to the somatosensory stimulation. More importantly, they found no difference between the aligned and misaligned stimuli. It is clear that the these findings challenged the notion that interactions between different modalities are restricted to spatially congruent stimuli in early sensory cortices. On the other side, Kayser et al. (2008) and Lakatos et al. (2007) provided strong evidence that the principle of inverse effectiveness hold also in the early auditory cortex at the level of local fiel potentials (LFPs) and neuronal spiking activity.

1.2. Possible neurophysiological mechanisms at the basis of crossmodal processing and multisensory integration

One of the most basic question emerged from the neural MSI literature of the recent years is relative the neural level at which cross-modal processing and MSI occurr in early sensory cortices: do these phenomena occurr at the level of single neurons with the convergence of different sensory inputs on the same neuron as observed in the superior colliculus (Stein & Meredith 1993)? or are these subthreshold phenomenona which effects can be observed only at neuronal population level (Kayser et al. 2007)?

Animal single neuron investigations have showed that neuronal activity can be modulated by nonauditory input in belt and core areas of the auditory cortex (Bizley et al. 2007; Brosch et al. 2005; Kayser et al. 2007; Kayser et al. 2010). Interestingly, in Bizley et al. (2007) investigation conducted on anesthetized ferrets, the visual input induced response suppression in the firing rates of the auditory cortex that consistently showed a visual spatial localization coding. Importantly, Kayser et al. (2008), recording both single units and local field potentials (LFP), showed that single unit firing rates were both suppressed and enhanced, while the local population activity, expressed by the LFP, was found to be increased. Importantly, the results of these authors reconciled the above results with fMRI findings that generally showed increased activity during MSI in the auditory cortex (Calvert et al., 1997; Kayser et al., 2007; Van Atteveldt, Formisano, Goebel, & Blomert, 2004). Importantly, it was recently showed that firing rates encode the information during multisensory stimulation processing (Kayser et al., 2010).

However, other studies have failed to reveal a direct effect at the level of single neurons. Chandrasekaran et al. (2013), investigating face/voice interactions in the belt regions of the auditory cortex of monkeys, found no changes in firing rates when the co-specific face was accompanied by voice. However, they observed a speed up of the onset of the spiking responses.

Similarly, the work of Lemus et al. (2010): during a flutter discrimination task, two monkeys had to discriminate the frequency of two subsequent stimuli. The authors found that, as expected, the somatosensory and the auditory cortex encoded the intramodal flutter frequency (respectively somatosensory and auditory), however, very few units responded to the cross-modal flutter frequency (respectively auditory and somatosensory) and, more importantly, without the specific encoding of the frequency itself. The authors argued that detailed information of the cross-modal sensory inputs is not present in primary auditory and somatosensory cortices. Notably, they did not find any evidence of MSI during this task. A commentary to this paper by Kayser (2010), stated that different mechanisms can influence the activity of local cortical neurons without necessarly involving a direct representation of the cross-sensory information at the single-neurons level. In agreement, one of these putative mechanisms seems to be the modulation of low-frequency oscillations during cross-modal inputs, as showed by the seminal work of Lakatos et al. (2007).

In their work, Lakatos et al. (2007), hypothesized that somatosensory input induces phase reset of local cortical activity of the auditory cortex, optimizing the efficacy of the auditory input processing during MSI. To test this hypothesis, they investigated the auditory and somatosensory activity in the primary auditory cortex of awake monkeys. Using current source density (CSD), an index of synaptic activity, and multiunit activity (MUA), an index of neuronal spiking rates, they were able to dissociate the presence of phase resetting from the presence of amplitude modulations. The most relevant finding was that the somatosensory input alone did not provoke any changes in neuronal spiking rates with no increase in the amplitude of LFPs, but induced phase-locked oscillations in gamma, theta and low-delta band. This effect (increase in phase consistency but not in amplitude) is called ‘pure phase resetting’ (Makeig et al. 2004) and indicates that the cross-modal input aligns consistently the phase of each single trial signal around a certain value, inducing that greatest reponses during MSI. Based on these results, the authors argued:

'..it does mean that neuronal acivity in auditory cortex is related to either somatosensory or visual perceptual experiences…On the contrary we think that it is likely that appropriately timed somatosensory and visual inputs to auditory cortex help us to hear better.'

Interestingly, the authors also observed that this phase resetting occurred in supragranular layers and before the occurrence of the auditory input processing; importantly these laminar and functional profiles were not compatible with feedforward interactions (activity in granular layer followed by extragranular activity).

Subsequent studies confirmed in both auditory and visual cortices that during MSI there is the reset of the phase of ongoing oscillations in the low-frequency band, although the 'strong' version of pure phase resetting was inconsistently showed (Kayser et al., 2008; Manuel Mercier et al., 2013). For example, Kayser et al. (2008), showed that the audio-visual input aligns consistently the phase of ongoing oscillations in primary auditory cortex: differently from Lakatos et al. (2007), their results strongly supported that the cross-modal input modulates the spiking responses in the primary auditory cortex. Moreover, they showed that: 1) the sites with enhanced spiking activity during audio-visual stimulation showed higher phase consistencies at 10 Hz in comparison to the auditory input; 2) the most part of these sites, always during the audio-visual stimulation, was characterized by an 'optimal' phase at 10 Hz. Importantly, they also evidenced that the amplitude of spiking responses to the intramodal input (i.e. auditory) was dependant on the phase of the oscillatory activity in the low-frequency bands (5-10 Hz) confirming that the activity of the primary auditory cortex is modulated by the low-frequency bands phase (Lakatos et al. 2005). Notably, differently from Lakatos et al. (2007), they concluded that:

'*Our data conform to a model where the visual stimulus enhances auditory responses by systematically setting the phase of ongoing slow-wave oscillations to an optimal phase….the activity communicated from these auditory areas to subsequent sensory processing stages not only contains information about acoustical events but also reflects aspects of the visual enviroment.'*

Interestingly, an intracranial study (Mercier et al., 2013) in the human visual cortex showed evidence of auditory-driven phase reset in the visual cortex. They first investigated the presence of non-random increase in phase consistency across trials during the different condition (intramodal, cross-modal and bimodal) in the alpha, beta and gamma band: as expected they found an important increase in phase concentration during the intramodal and bimodal condition, but interestingly they also detected some contacts showing an increase in phase concentration during the cross-modal condition. Then, they evaluated if this increase in phase concentration was accompanied or not by an increase in power in the same frequency bands. They observed that only a very small proportion of contacts presenting increase in phase concentration, showed also increase in power during the cross-modal condition, differently from the intramodal condition and the bimodal condition. It is important to note that another study in humans conducted with intracranial EEG did not find evidence of MSI in primary visual cortex (Quinn et al. 2014) investigating the high-gamma power band.

Notably, very recently, two important studies in mouse evidenced inhibitory mechanisms at the basis of MSI in the primary visual cortex and confirmed the presence of modulation of neuronal spiking activity during cross-modal processing. Iurilli et al. (2012) showed hyperpolarization of the neurons of layer 2/3 of the primary visual cortex during auditory processing of noise burts and suggested that this hyperpolarization is due to an inhibitory network involving neurons of layer 5. Ibrahim et al., (2016) confirmed the effect of sound processing in the

primary visual cortex: they showed that the orientation selectivity of the excitatory neurons of layer 2/3 is modulated by the presence of an auditory input or during the optogenetic activation of axonal projections from the primary auditory cortex to the primary visual cortex. This modulation, differently from the one observed in Iurilli et al. (2012), seems to be under the control of inhibitory and disinhibitory circuits possibly mediated by the neuronal activity in layer 1, strongly innervated by axon terminals of layer 5 neurons of the primary auditory cortex.

Returning to the main basic question I proposed at the beginning of this chapter, namely whether early MSI occurs at the level of single neurons with the convergence of different sensory inputs on the same neuron (Stein & Meredith 1993) or is a subthreshold phenomenon which effects can be observed only at neuronal population level (Kayser et al. 2007), it is now possible to give an answer, although with all the cautions that are necessary when comparing studies conducted with very different techniques: cross-modal processing and MSI mechanisms seem to be at work at both neuronal and population levels.

Moreover, based on the above studies, it is possible to conclude that: 1. In contrast with Lakatos et al. (2007), Lemus et al. (2010) and Chandrasekaran et al. (2013), neuronal activity during cross-modal stimulation plays an important role in MSI (C. Kayser et al. 2008; Bizley et al. 2007; Brosch et al. 2005; Ibrahim et al. 2016; Iurilli et al. 2012; Kayser et al. 2010); if this is true, possibly these cross-modal information are passed to higher-level sensory areas; 2. the phase resetting of the ongoing oscillatory activity in the low-frequency bands, seems to be an important mechanism at the basis of MSI, however, based on the above results, it is difficult to accept in its 'strong' meaning, namely the 'pure phase resetting' (Lakatos et al., 2007; Mercier et al., 2013); 3. complex mechanisms seem to be at work in the primary sensory cortices: Ibrahim et al., (2016) and Iurilli et al. (2012) studies provided evidence of strong neuronal activity during the cross-

modal stimulation in neurons of the layer 2/3 under the control of the inhibitory effect of layer 1 or 5.

In the next paragraph I will introduce the relevant anatomical studies that support the presence of cross-modal processing and MSI and suggest the possible functional interactions among early sensory areas with a particular attention to the early visual and auditory area.

1.2. Anatomical pathways of crossmodal processing and MSI

Although early anatomical studies did not find evidence of strong anatomical links between primary sensory areas of different modalities (Jones and Powell, 1970), in these last decades, the availability of new techniques (i.e. new anterograde and retrograde tracers) provided important novel insights. We will now revise the current knowledge relative the anatomical connections between visual and auditory cortices (i.e. interareal): these pathways constitute the anatomical bases of cross-modal processing and MSI.

From auditory area to early visual areas

In the last 15 years, several animal studies have provided coherent evidence of direct projections (monosynaptic) from several regions of the auditory cortex to early visual areas. However, until 2002, there was no evidence of direct pathways from A1 to early visual areas in monkeys. The apparent absence of this anatomical connection in monkeys was quite controversial because of the evidence of this pathway in Mongolian gerbil (Budinger et al. 2000) and, more importantly, because early visual areas responded to non-visual inputs in congenital blind individuals (Büchel, Price, Frackowiak, & Friston, 1998; Cohen, Weeks, Sadato, Celnik, & Ishii, 1999; Sadato, Pascual-Leone, Grafman, & Ibañez, 1996; Weeks, Horwitz, Aziz-Sultan, & Tian, 2000). The seminal work of Falchier et al. (2002), provided the first evidence of this pathway in macaque monkeys: using retrograde tracers in different portion of V1 they observed the parental neurons of these pathways in the layer 6 of the core, belt and parabelt areas of the auditory cortex. Importantly, they demonstrated an important feature of these projections: the peripheral areas (representing the peripheral visual field) of V1 received projections of moderate density, while the central areas received very few connections. Notably, the evidence that peripheral areas of V1 received projection from auditory areas was strongly supported in subsequent studies,

while, the evidence that central area of V1 received sparse connections from auditory areas, did not receive support.

The important results of Falchier et al. (2002) work were partially confirmed one year later by another group: Rockland et al. (2003), showed that belt areas of the auditory cortex projected to the peripheral areas of V1 and V2. These projections were sparse in V1 and denser in V2**.** However, the authors did not find any projections from the belt and parabelt areas of the auditory cortex to the central areas of V1 and V2.

Several studies supported the above findings also in other mammals: in Mongolian gerbil Henschke et al. (Henschke et al. 2015), confirmed that A1 sends moderate projections to V1; in marmoset, monkeys characterized by a lissencephalic brain. Cappe et al. (2005), did not find evidence of anatomical pathways projecting from the auditory regions to the central regions of V2, therefore indirectly supporting the notion that heteromodal connections of area V2 might be restricted to the peripheral representations of the visual field. In cats, Hall et al. (Hall & Lomber 2008) using retrograde tracer in both central and peripheral retinotopic representations of V1 and V2, identified the parent neurons of these pathways in the posterior auditory field. In a second group of cats, the authors refined their previous observations: using small deposit in precise locations of the visual cortex (central, paracentral and peripheral visual field) they found that parent neurons in the posterior auditory cortex target neurons of the visual periphery field. Hall and colleagues clearly showed that the organization of the heteromodal pathways of the cats was consistent with the one observed in monkeys. Very importantly, they revealed that neurons in the peripheral auditory fields presented a weak retinotopic organization. This resembles the neurophysiological work of Morrell (Morrell 1972a) who showed that neurons of the primary visual cortex of the cats presented a spatial map of the auditory inputs.

Based on the above it is now widely accepted that in primates core, belt and parabelt areas of the auditory cortex project moderate/dense projections to the peripheral field of V2 and sparse projections to the peripheral field of V1. Importantly, all the above studies reported that the observed projections presented a laminar profile compatible with feedback projections.

Ibrahim et al. (2016) confirmed, with viral injections, that the primary auditory cortex of mouse projects axon terminals to the primary visual cortex, in particular the layer 1. This is in agreement with the results of the connectome map of the mouse brain (Oh et al. 2014).

From visual area to early auditory areas

The evidence of reciprocal connections between the auditory and visual cortex is still far from being solved. However in recent years, different studies have tried to investigate pathways from the visual areas targeting the auditory cortex.

The first evidence that the core region of the auditory cortex is the target of projections from the visual cortex was produced by Cappe et al. (2005). In the same study cited above, the authors showed convincing evidence that the core region of the auditory cortex was the target of visual neurons localized in the superior temporal sulcus. Budinger et al. (2006) in Mongolian gerbil, confirmed the existence of visual pathways targeting A1 from secondary visual areas. Importantly, they stressed that this pathway presented the features of feedback connections.

The important work of Bizley et al. (2007) in ferrets, showed the existence of pathways targeting the auditory areas from the early visual cortex. These authors, finding evidence of the visual activity in neurons of the primary auditory cortex of ferret, explored its possible anatomical substrate. They showed that both primary and non -primary auditory cortices are target of neurons situated in early visual areas (V1 and V2). In particular, they showed that the primary

visual cortex project sparsely to A1 and that higher-order visual areas projected to auditory areas in a field specific manner.

Falchier et al. (2010) confirmed the existence of these anatomical connections in macaque monkeys. After their work on the pathways targeting the early visual cortex from the auditory cortex, they were interested in the possible reciprocal connections (i.e. from early visual to the caudal auditory areas) and, if any, in the type of connections. They showed that V2 and prostriata, but not V1, projected to the caudal auditory cortex by means of feedback pathways. The absence of projections from V1 was not unexpected: V1 indeed receives very few non reciprocal input from different association areas. Importantly the authors, based on their previous study, concluded that both auditory and low-level visual cortices are connected by feedback pathways in a bidirectional loop. A very recent study in Mongolian gerbil, cited also in the above paragraph, (Henschke et al. 2015) questioned the absence of neurons in V1 targeting the auditory areas. These authors showed evidence that V1 provides faint connections to A1 and that these connections are mainly of feedforward type. Interestingly, Ibrahim et al. (2016) failed to reveal connections from V1 to A1 in mouse, suggesting that in some species the connectivity between the primary visual and auditory cortex is unidirectional.

The above studies strongly support the presence of heteromodal direct connections in both early auditory and visual areas and their reciprocal connections, despite some doubts persist about the presence of connection from V1 to A1. At this point, the next challenge of the scientific community involved in this field, seems to be the discovery of possible organizational principles of these pathways: indeed Hall et al. (2008) showed that neurons in the peripheral auditory fields presented a weak retinotopic organization, while the old, although very actual, work of Morrell ((1972), showed that the auditory input in the visual cortex presented a precise spatial organization.

In the next paragraph I will illustrate the relevant key-concepts, not explicitly grounded in the field of MSI, that are critical for the research I have conducted during these years of study: neural oscillations, the spectral fingerprints framework and the pure phase resetting model.

1.3. Neural oscillations: a key to understand the underpinnings of MSI?

Until a decade ago, the most part of human studies conducted with EEG investigated the topic of interest by means of event related potentials (ERPs) and MSI studies, at this regard, were not an exception (M.H. Giard & Peronnet 1999; Molholm et al. 2002; Besle, Fort, Delpuech, et al. 2004; Bonath et al. 2007). The high temporal precision and the accuracy of ERPs and the extensive literature of ERPs findings to which to compare the results, constituted an enormous advantage in EEG investigations. However, more recently, we have assisted to a great amelioration of EEG acquisitions, and to the advent of new sophisticated analyses of the electrophysiological signal in the time-frequency domain (Tallon-Baudry & Bertrand 1999): all this progress has allowed to deeply investigate brain oscillations.

In the following paragraph I will delineate the most important differences between the time domain analyses and the time-frequency domain analyses and then I will discuss the importance to investigate oscillations in particular in the MSI field.

1.4.1. Time domain and time-frequency domain analyses

Please, consider that the following observations can be done also for magnetoencephalography (MEG) and intracranial electrophysiological recordings (iEEG).

The EEG signal provides an extraordinary powerful window onto the dynamic of the brain function. It is characterized by a multidimensional space comprising space, time, frequency and power and phase of the specific frequency band (Cohen, 2014).

EEG data are analyzed by means of time and time-frequency domain approaches. In the time domain approach, EEG epochs, time-locked to the event of interest, are averaged for each single time point. This analysis produces a 1–D temporal series (event-related potentials – ERPs) of mean potential deviations from baseline expressed in μV (Makeig et al. 2004).

In the time-frequency domain approach, EEG spectral power and phase are investigated. Pre to post-stimulus changes of EEG spectral power are averaged across epochs, time-locked to the event of interest, in each single time-frequency point, to obtain event-related spectral perturbations (ERSPs). EEG spectral phase allows to compute the phase concentration index concentration across trials, time-locked to the event of interest, at each single time-frequency point (Tallon-Baudry et al. 1996; Makeig et al. 2004): the phase of EEG activity can collapse around specific values indicating the presence of phase resetting.

Generally speaking, the EEG activity recorded during the execution of a task is classified as phase-locked or non phase-locked. The phase-locked activity (or 'evoked' activity) is time-locked and phase-locked to the onset of the event: this activity is observed in both time domain and in time-frequency domain. The non-phase-locked activity (or 'induced' activity) is time-locked but not-phase locked to the onset of event: it can be observed in the EEG spectral power (ERSPs) but not in phase concentration index and in the ERPs. It is clear that non-phase-locked oscillations, are inaccessible to the common ERPs analyses, while they are prone to be investigated in the EEG spectral power.

Despite the above methodological distinctions, it is still unclear what are the neurophysiological basis of phase-locked and non-phase-locked activity, however non-phase-locked activity is considered a strong evidence of the presence of oscillations (Donner & Siegel 2011; Tallon-Baudry & Bertrand 1999).

Neural oscillations are rhythmic fluctuations in the excitability of neuronal populations and are observed across different temporal and spatial scales (Cohen 2014). Specific oscillations have been linked to many cognitive events (Buzsáki 2006; Engel et al. 2001; Klimesch et al. 2007). From the neurophysiological point of view, oscillations are mainly produced by interactions between synchronized inhibitory interneurons (GABAergic) and excitatory neurons (pyramidal cells). These

interactions give rise to a shifting balance between excitation and inhibition of the neuronal populations (Cohen, 2014). Neural oscillations can be described by their frequency (the number of cycles per second), power (the squared amplitude of the oscillations in a time-frequency point) and phase (the position along the sine wave at a specific time point). It is important to note that power and phase, the complex output of the time-frequency analyses, are independent from the mathematical point of view: however, when the power is small, the phase is difficult to detect (Cohen, 2014).

1.4.2. Spectral fingerprints of cognitive processes

Several neurophysiological studies have suggested interesting relationships between oscillations at different frequencies bands and neurophysiological processes (Donner & Siegel, 2011; Buzsaki & Wang, 2012; Engel & Fries, 2010; Gregoriou et al., 2009; Kucewicz et al., 2014). Cortical oscillations indeed might present specific local spectral profiles that are process-dependent and are considered useful indicator of the underlying neural canonical computations; for this reason spectral profiles were coined spectral fingerprints in the works by Donner et al. (2011) and Siegel et al. (2012). This theoretical framework (i.e. spectral fingerprints) postulates that local changes in oscillations, in different frequency bands, encode different sensory/cognitive processes in a dynamic spectral fingerprint (Siegel et al., 2012). The spectral fingerprints are generally referred to the coherence in phase between different cortical areas, however, also local power oscillations are considered spectral fingerprints potentially able to reveal the underlying canonical computations (Siegel et al., 2012).

The hypothesis of the link between the local spectral profile of the oscillations and specific processes raised from the consistent observation that local gamma band oscillations are specific index of excitatory–inhibitory interactions among different neuronal populations (Donner & Siegel

2011). For example, it was consistently showed that the presence of a visual input increases the activity in the gamma band oscillations and reduces the activity in alpha band oscillations. Differently, during binocular rivalry, the oscillatory activity of the input awareness is characterized by increased activity in the low-frequency band (Maier et al. 2008; Wilke & Logothetis 2006).

Despite the enormous importance of this empirical and theoretical framework, the link between specific oscillations, specific cognitive processes and specific canonical computations is still debated and far from being solved (Wang 2010). Several theoretical proposals have been provided to make a coherent framework of the interpretation of the spectral fingerprints.

A broad categorization was provided by Donner & Siegel (2011), who divided the interactions among neuronal populations as encoding or integrative: encoding mechanisms would be mediated by local cortical network interactions, while integrative mechanisms would be mediated by long-range interactions among distant regions. From the oscillatory perspective, these mechanisms would induce very different local cortical oscillation patterns: local cortical network interactions would induce gamma-band oscillations, while long-range interactions low-frequency band oscillations (mainly in the beta band). In agreement, the frequencies of network interactions were showed to be determined by the biophysical properties of the neuronal networks; it was suggested that the frequency of synchronization of interacting neuronal population may depend from conduction delays: beta frequency is able to synchronize over long conduction delay, while this is not possible for the gamma oscillations (Kopell et al. 2000). However, this physical effect is not able to account for the gamma modulation in top-down visual attention (Siegel et al. 2008; Fries et al. 2008) and gamma synchronization in long-range interactions (Gregoriou et al. 2009; Hipp et al. 2011) and therefore other mechanisms, currently unknown, may play a role in the specific frequency activity.

More recently, coherent findings have provided evidence that different frequencies of coherent oscillations might subtend the directions of information stream (Wang 2010a; Buffalo et al. 2011). Since 2000 (von Stein et al. 2000) laminar recordings have provided evidence that gamma band coherence was linked to bottom-up interaction, while beta band coherence to top-down interaction. Buschman et al. (2007) compared the synchronization between the prefrontal and parietal areas during different effects of attention in monkeys. The authors showed that gamma band synchronization was related to bottom-up attention, while beta band synchronization to top-down attention. These findings suggested that information in different directions travel using different frequency bands. Along the same lines, Buffalo et al. (2011) evidenced that the superficial layers of the visual cortex in monkeys presented a strong synchronization in the gamma band, while the deep layers in the alpha/beta band. The authors argued that the synchronization in these two frequency bands is transferred to different targets due to the long standing notion that the superficial layers (mainly layers 2/3) provide feedforward projections, while the deep layers feedback projections.

These important findings elegantly explained the results of Bastos et al. (2015), who investigated interareal communication in visual areas of macaques showing that information travelled on distinct frequency channels. Bastos et al. (2015) showed that feedforward influences were carried by the gamma band, while feedback influences by the theta band. Importantly, the superficial layers are the primary source of feedforward projections, while the deep layers of the feedback projections. Fontolan et al. (2014) found evidence of this mechanism in human auditory cortex: feedforward influences were reported in the gamma band, while feedback influences in the low-frequency bands. Similarly, van Kerkorle et al. (2014) tested directly the hypothesis that low-frequency band and gamma band channel the direction of the information in the visual cortex. They showed that gamma oscillations started in layer 4 and then spread to the superficial and

deep layers of the cortex, while alpha oscillations followed the opposite direction. Moreover, recording simultaneously from V1 and V4, they confirmed that gamma oscillations propagated in feedforward direction, while alpha oscillations in the feedback direction. Importantly, when V1 was stimulated, activity in the gamma band was evidenced in V4, consistent with feedforward mechanisms, while when V4 was stimulated, activity in the low-frequency band was evidenced in V1, consistent with feedback mechanisms.

Taken together, these data refined and enriched the model proposed by Donner & Siegel (2011) who divided the possible interactions among neuronal populations as encoding (gamma-band oscillations) or integrative (low-frequency oscillations), suggesting that the brain strongly segregates information, flowing in different directions, in different frequency channels.

Moreover, these last empirical findings complement the hypothesis that synchronization of neural signals in specific frequency channels among different brain areas might build specific and dynamic neuronal networks, allowing a context-dependent binding of distributed processes, the selection of relevant information, and the efficient routing of information (Womelsdorf et al., 2007; Fries, 2009). The synchronization of neural activity in specific frequency bands may solve the binding problems that occur in distributed architectures (Engel et al., 2012a), such as during crossmodal processing and MSI.

1.4.3. Phase resetting of ongoing oscillations in MSI

In recent years, several studies have investigated the role of phase modulations of oscillatory responses during MSI, suggesting that the phase of the neural signals may be a relevant mechanism for the processing of cross-modal and multisensory inputs (Lakatos, Chen, O'Connell, A. Mills, et al. 2007; Mercier et al. 2013; Christoph Kayser et al. 2008).

Neuronal excitability was showed to be modulated by the ongoing neural oscillations (Romei, Brodbeck, et al. 2008; Romei, Rihs, et al. 2008; Romei et al. 2010; Lakatos, Chen, O'Connell, Ai. Mills, et al. 2007; Schroeder & Lakatos 2009). Along the same lines, it was also showed that ongoing oscillations are shaped by sensory inputs: the sensory inputs would reset the phase of the ongoing oscillations (Lakatos, Chen, O'Connell, Ai. Mills, et al. 2007; Schroeder & Lakatos 2009). This mechanism would allow an efficient and specific process of the sensory inputs.

In their seminal study, Lakatos et al. (2007), analyzing phase-locked oscillations in monkey primary auditory cortex (A1), showed that the cross-modal condition induces phase-locked oscillations with specific phase angles with very low amplitudes. According to previous studies, phase-locked oscillations might be originated by a stimulus-evoked response and/or by a stimulus-induced phase resetting (Makeig et al. 2004). In particular, stimulus-evoked responses are accompanied by increased power, while, stimulus-induced phase resetting are originated by an increase in phase synchrony across trials, but not by an increase in power (Makeig et al. 2004). Based on these notions, Lakatos et al. (2007) provided the first evidence that cross-modal processing induces a pure phase-resetting of ongoing oscillations (i.e. presence of phase locked oscillations characterized by very low amplitude) arguing that this modulation could functionally prepare the system (e.g., visual) for the processing of the intramodal stimulus (Schroeder & Lakatos 2009).

Kayser et al., (2008), provided further support to the phase-resetting as neural mechanism at the basis of cross-modal processing and MSI: they showed that the presentation of a simple auditory stimuli gives rise in monkey A1 to an evoked response, which amplitude was related to the phase of the low-frequency band activity (5-10 Hz). More importantly, they showed that the signal amplitude of audio-visual processes depended from the degree of phase concentration across trials: the enhanced sites were characterized by a strong phase consistency at 10 Hz. The authors stated that the low-frequency oscillations modulate the response amplitude of audio-visual interactions supporting the notion that the reset of ongoing oscillations in the low frequency band might be an important mechanism for sensory integration.

Important results in this direction were also obtained by two other studies investigating the coherence of phases between unisensory region (primary auditory cortex) and associative region (superior temporal sulcus). Maier et al. (2008) showed that audio-visual looming signals enhanced the coherence of the phases between the primary auditory cortex and the superior temporal sulcus, while Kayser et al. (2009) showed that the auditory cortex is provided of visual influences by the superior temporal sulcus by means of the same mechanism (i.e. enhanced coherence in the phase in these two regions).

Recently, evidences of this mechanism were observed also in humans. Thorne et al. (2011) presented evidence that the visual input resets the phase of oscillatory activity in auditory cortex in theta and alpha frequencies, importantly they showed that degree of phase resetting correlated with the RTs. Naue et al. (Naue et al. 2011), showed that the stimulus onset asynchrony (SOAs) of audiovisual stimuli, constituted by white noise burst and led flashes, modulate the beta responses suggesting the presence of phase-reset. A very recent iEEG study of patients undergoing epilepsy treatment has showed, using a simple detection paradigm, that the crossmodal input modulates the large part of the visual cortex activity by the phase resetting mechanism (Mercier et al., 2013).

Remarkably, Diederich et al. (Diederich et al. 2012) reported phase reset using a purely behavioural paradigm recording saccadic response times to audio-visual stimuli with variable SOAs. These data showed specific periodicities suggestive of phase resetting at 20-40 Hz.

Despite the above findings, some concerns were raised relative the possibility to detect the pure phase resetting phenomenon (increased in phase concentration not accompanied by an increase in the amplitude of ongoing oscillations). Ding et al. (2013), in a stochastic model, showed that, even without phase resetting, the phase emerges as synchronized over trials: more importantly, statistical properties renders more easily to detect the phase resetting, in comparison to the increase in power. The authors therefore questioned the validity of the methods used to investigate the phase resetting based on the evidence of increased phase concentration with no (or limited) increased power.

In the next paragraph I will introduce the major advantages and pitfalls of the technique (stereo-electroencephalography) I have used to investigate the spectral fingerprints of cross-modal processing and MSI.

1.4. Human intracranial recordings: advantages and pitfalls

Human intracranial recording (iEEG) is an invasive technique developed to clinically investigate subsamples of patients that meet particular pathological conditions. The use of iEEG is not an achievement of these last decades, but conversely, it can be dated back to the early electrophysiological studies: Berger performed the first electrocorticogram in 1924 during a neurosurgical operation on a young boy and in 1929 reported his seminal observations of the alpha and beta waves (Tudor et al. 2004). In these last years, an increasing number of groups have used the intracranial recordings to investigate the human cognition, providing important insights, just to cite few, relatively the functional organization of the visual system (Allison et al. 1999) and the human place cells in hippocampus (O'Keefe 1999) and for the development of brain-computer interfaces (Leuthardt et al. 2004).

Two main types of intracranial recordings techniques were developed: electrocorticography (ECoG) and stereo-electroencephalography (SEEG). ECoG uses subdural grids composed by arrays of contacts (spaced 1 cm from each other) that are positioned, with a large craniotomy, on the surface of the brain. SEEG uses depth needle-like electrodes, comprising several contacts (up to 18 and spaced 2 mm from each other), positioned with a minimal surgery (i.e. no need of craniotomy). It is clear that ECoG provides an extensive sampling of the gyral activity of the lateral surface of the brain, while SEEG, although not providing the extensive coverage of ECoG, allows to record from both gyri and sulci and from deep cortical structures (Jerbi et al. 2009). Based on the above observations, It is clear that, in the context of the investigation of MSI in early visual (calcarine scissure) and auditory areas (Heschl's gyrus), SEEG seems to be more suitable in comparison to ECoG.

Among clinical populations implanted with iEEG, selected pharmacologically resistant epileptic patients provide an incredible opportunity to investigate human cognition mainly

because iEEG tests the functional organization of the epileptogenic network. This implies that the recording sites are distributed not only in the epileptic focus, but also in sparse and normal (i.e. healthy) brain regions (Jerbi et al. 2009). A second pragmatic aspect is that these patients are chronically implanted for weeks for the pre-surgical evaluations of the epileptic focus (Jerbi et al. 2009): this creates an extended temporal window in which the patient recover almost completely from the surgery before undergoing the experimental settings and allows him to perform the cognitive tasks in the most convenient setting.

Obviously, the fact that the data are collected from clinical populations constitute one possible pitfall of iEEG. In principle, this could induce a physiological interpretation of pathological processes. To avoid this problem, it is used to inspect the activity and MRI location of each single contact and to discard any contact showing any type of abnormal discharge referred to the epileptic activity or located in dysplasic regions of the brain. Another well known pitfall of iEEG recording is that the anatomical origin of the signal cannot be directly compared among patients: indeed, due to the fact that the implantation of the electrodes is guided only by clinical questions, there are not standard regions under investigations in all the patients. This clearly renders the iEEG human studies very similar to electrophysiological studies in monkeys (Jerbi et al. 2009).

Despite the above drawbacks, intrinsically solvable, iEEG investigation provides a unique window over the dynamics of cognitive processes of the human brain for three extremely important reasons.

First, it presents both high temporal and spatial resolution (Engel 2005), high resistance to possible artefacts (muscle contractions, eye blinks), and an exceptional signal to noise ratio when compared to classic surface electrophysiological methods used in humans (Lachaux et al. 2012). The temporal resolution, depending from the sampling rate (up to 2000 Hz with the modern

system), is equal to the resolution provided by EEG/MEG techniques; the spatial resolution is similar, or even better, than the resolution provided by fMRI (Jerbi et al. 2009).

Second, it allows to record the local filed potentials (LFP), namely the coherent activity of neuronal populations and, very importantly, high-gamma frequency band activity, linked to high-frequency synaptic and spiking activity (Manning et al. 2009; Buzsáki et al. 2012) and almost inaccessible to the traditional EEG/MEG recordings. This aspect was particular relevant for the explosion of iEEG studies at the end of '90 in consideration of the hypothesized role of the gamma-band in communication among different neuronal populations and in neural representations, expressed in the 'binding by synchrony' hypothesis (Singer 1999), and of the evidence of the ubiquity of the high-frequency activity in the neocortex (Crone et al. 1998; Aoki et al. 1999). In the subsequent years, it became clear that high-frequency activity was an index of local cortical processing (Lachaux et al. 2012). Although the origin of high-frequency neural activity is currently under debate, it's explained by an increase in firing rates and by local synchronization mechanisms producing high frequency oscillations at vary latencies and frequencies. According to some authors, studying high-frequency neural activity with iEEG recordings can provide insights into the neural bases of human cognition that cannot be obtained from fMRI/EEG/MEG investigations (Lachaux et al. 2012).

The third great advantage of iEEG is related to its less sensitivity, in comparison to EEG/MEG acquisitions, to volume conduction problems. Volume conduction constitutes a great deal for the most part of cognitive investigations conducted with electrophysiological techniques.

Moreover, we will see in the next paragraph that this problem is particularly relevant for cross-modal processing and MSI studies, in particular when testing the maximum model on data obtained with population-based techniques (EEG/MEG).

2. Experimental part

2.1. Aims of the work

The main aim of this work is to contribute to the current ongoing debate about the nature of cross-modal processing and MSI in early human visual and auditory cortex relying on the spectral fingerprints framework. To this goal, we investigated early visual (V1 and V2) and auditory (Heschl's gyrus and planum temporale) areas in 8 epileptic patients by means of SEEG, an invasive technique with high temporal and spatial resolution and an exceptional signal to noise ratio (Lachaux et al. 2012).

First of all, due to the lack of a coherent framework in the interpretation of neural oscillations, we investigated in a comprehensive manner intracranial EEG spectral power (theta/alpha band: 5-13Hz; beta band: 13-30 Hz; gamma band: 30-80 Hz; high-gamma band: 80-200 Hz). This approach allowed us to observe whether, during cross-modal processing and MSI, power modulations are present in the high-gamma band, and in the low-frequency (theta/alpha and beta) and gamma band.

Modulations in the high-gamma band are indicators of the high-frequency synaptic and spiking activity (Manning et al. 2009; Buzsáki et al. 2012). The possible presence of this effect in the investigated regions should therefore indicate that cross-modal and multisensory information are coded in neuronal populations. This is particularly important in the light of a human study in the early visual cortex (Quinn et al. 2014) that failed to reveal the presence of high-gamma band modulations and of animal investigations in early sensory regions that have inconsistently showed neuronal responses during cross-modal processing (Christoph Kayser et al. 2008; Lakatos, Chen, O'Connell, Ai. Mills, et al. 2007; Bizley et al. 2007; Lemus et al. 2010; Chandrasekaran et al. 2013).

Modulations in the low-frequency (theta/alpha and beta) and gamma band are likewise interesting: recent investigations showed that feedback and feedforward information travel

segregated in different frequency channels during intramodal interactions, respectively in the oscillatory activity of the low-frequency bands and in the oscillatory activity of the gamma band (Fontolan et al. 2014; van Kerkoerle et al. 2014; A. M. A. M. Bastos et al. 2015). Based on the possibility that also inter-areal interactions might occur through specific frequency channels, the presence of low-frequency and gamma power modulations would allow to gain insight on the possible type of interactions during MSI.

Secondly, due to the recent interest of animal and human electrophysiological studies showing that phase resetting of oscillatory activity might play a role in early MSI (Lakatos et al. 2007; Kayser et al. 2008; Mercier et al. 2013; Naue et al. 2011; Thorne et al. 2011), we investigated intracranial EEG spectral phase (theta/alpha band: 5-13Hz; beta band: 13-30 Hz; gamma band: 30-80 Hz). This investigation is particularly important, because we still lack of evidence of this mechanism in human early auditory and visual areas.

Despite of human EEG studies showing that the auditory cortex presents phase resetting mechanism during visual and audio-visual processing (Naue et al. 2011; Thorne et al. 2011), these researches lack the essential spatial resolution to test whether this mechanism occurs in early auditory areas. Importantly, Mercier et al (2013), by means of ECoG, provided the first and only evidence that phase resetting is present in the human visual cortex. Although the interesting results, they did not report their findings relative to early visual areas (V1 and V2). Another important aim of our study, is therefore to confirm the presence of the phase resetting phenomenon in human early visual and auditory areas during cross-modal processing and MSI.

2.2. Participants

Data were collected from 8 participants (mean age: 34 years ± 11; 4 females), suffering from drug-resistant epilepsy, stereotactically implanted with intracerebral electrodes (Cardinale et al. 2013). Each electrode (diameter of 0.8 mm) comprised from 10 to 18 contacts (number of contacts per patient: from 146 to 181) spaced 1.5 mm apart (DIXI, Besançon, France). All patients had cognitive abilities in the normal range as assessed by a neuropsychological exam and did not have specific deficits in visual and auditory functions. Moreover, the data were collected not before than three days after electrode implantation, 24 hours before and after spontaneous seizures, except for one patient experiencing one short seizure 2 hours before the acquisition and lasting less than 30 seconds, occurring with the recovery of the usual interictal intracranial EEG activity in less than 30 minute. All the electrodes were implanted only according to clinical criteria, and the conduction of this study did not influence the clinical procedures. The research project was approved by the Institutional Review Boards of the Hospital Niguarda Ca' Granda of Milan and adhered to the Declaration of Helsinki. The participants provided written informed consent.

2.3. **Paradigm**

The paradigm (fig.1) was implemented and administered using Presentation software (https://www.neurobs.com/). Each participant, seated in front of the screen at a distance of 60 cm, was presented with several blocks (from 2 blocks to 12 blocks; median for all participants: 4 blocks) containing non-target and target auditory (A), visual (V) and audio-visual (AV) condition. A "nothing" condition (N), without any stimulation, was used as control condition to record and compensate for any anticipatory brain responses at the times a stimulation is predicted to typically occur (Teder-Sälejärvi et al. 2002; Talsma & Woldorff 2005; Mishra et al. 2007; Gondan et al. 2005; Bonath et al. 2007). In every single block, each non-target and N conditions were presented 40 times, while each target condition was presented 4 times. Non-target, target and N conditions were randomly interleaved. The V stimuli (subtending 14.8 deg. of visual angle) were presented for 100ms in the centre of the screen: the non-target V stimulus was a black and white checkerboard, while the target V stimulus was a coloured checkerboard. The A stimuli were presented binaurally through inserted earphones at a comfortable auditory level for each individual participant. The non-target A stimulus was a 100ms segment of white noise (5 ms fade in/out), while the target a 100ms pure tone (2000Hz; 5ms fade in/out). The non-target AV condition was a combination of the non-target V and non-target A condition. During the non-target AV condition, the V stimulus was presented 30 ms before the onset of the A stimulus, since behavioural studies have shown that V stimuli have to be presented before A stimuli (between 20 and 90 ms) to be perceived as simultaneous (Zampini et al. 2003) and to obtain the strongest behavioural gain during the audio-visual condition in comparison to the intramodal condition (J. D. Thorne et al. 2011); in agreement, neurophysiological studies have indicated that neural MSI in auditory areas occurs when the V stimulus precedes the A stimulus (depending from the studies,

between 30 and 75 ms; (Christoph Kayser et al. 2008; J. D. Thorne et al. 2011; Musacchia & Schroeder 2009).

The inter-stimulus interval was jittered between 1000 and 1500 ms. The target AV condition was a combination of the target V and target A condition with the same temporal characteristic of the non-target AV stimulus. Timing accuracy of the presented stimuli was controlled offline with the Black Box toolkit (http://www.blackboxtoolkit.com). Participants were asked to maintain central fixation and to respond as fast as possible when the target conditions were presented. All patients were able to detect the target conditions with high accuracy (accuracy: 100% for all participants except for one (63%); false hits: from 0% to 7%, median 2%). We investigated the electrophysiological activity recorded during the non-target conditions in order to avoid any confounds in the signal that were linked to the motor response of the participant. Hereby, when mentioning A, V and AV conditions we will refer to the non-target conditions. Moreover, *intramodal* input describes the stimulus matching the sensory representations (Stein & Stanford 2008) of the investigated cortex (auditory input in the auditory cortex, visual input in the visual cortex). In contrast, *cross-modal* input refers to the stimulus not matching the main sensory representations of the investigated cortex (auditory input in the visual cortex, visual input in the auditory cortex). Finally, *MSI input* is the combined intramodal and cross-modal inputs.

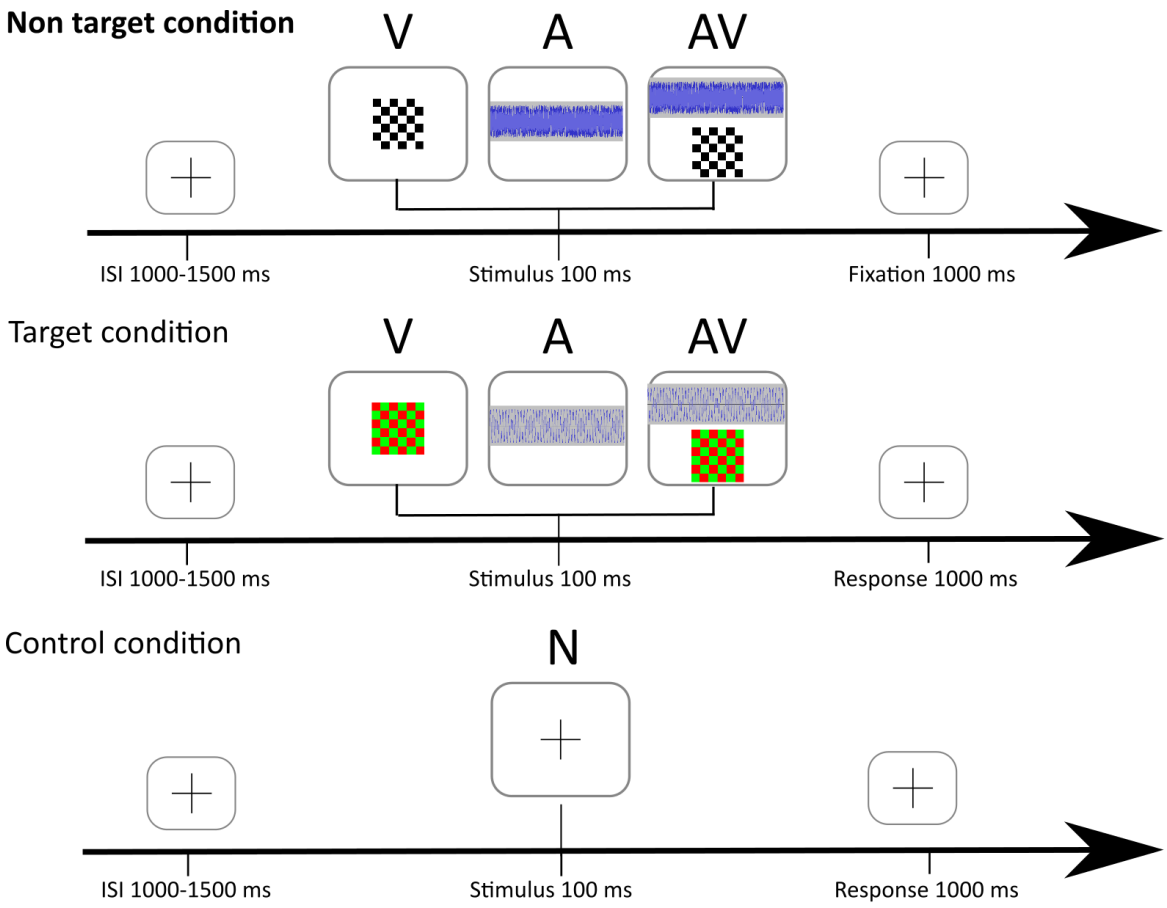

Figure 1. The paradigm presented non-target and target auditory (A), visual (V) and audiovisual (AV) conditions. A "Nothing" (no stimulus presented) condition (N) was used as control condition. Non-target V stimulus was a black and white checkerboard while the target a colored checkerboard. The non-target A stimulus was a white noise, while the target a pure tone. A and V stimuli were presented for 100 ms. The non-target AV condition was a combination of the non-target V and non-target A condition. Our analyses were focalized on the non-target conditions to avoid motor confound in our calculation of MSI (see method).

2.4. Localization of contacts of interest

We performed the analyses for contacts of interest (COIs). These contacts were anatomically localized in calcarine and pericalcarine regions (n=44) for the occipital-visual cortex and in Heschl's gyrus and planum temporale (n=47) for the temporal-auditory cortex. To identify COIs in each participant, post-implantation intraoperative Cone-Beam-CT (CBCT) scan (192 axial slices, 512 x 512 matrix, 0.415 x 0.415 x 0.833 mm anisotropic voxels) was registered to pre-implantation MR (3D fast field echo T1-weighted sequence, contiguous axial slices with 560 x 560 reconstruction matrix, 0.46 x 0.46 x 0.9 mm voxel, no inter-slice gap) by means of FLIRT 6.0 (Jenkinson & Smith 2001). Such CBCT scans provide undistorted images of the electrodes (fig. 2). Co-registered images were then normalized to MNI space. The anatomical location of COIs was assessed with two steps-procedures: 1) an expert medical epileptologist (R.M.) performed a visual inspection of the co-registered images in the native MRI space with 3D Slicer 4.3.1 software (Fedorov et al. 2012), and identified possible COIs; 2) by means of Freesurfer 5.3.0 (Fischl 2012), normalized brain tissue was segmented and cerebral surfaces were reconstructed and parcellated. Electrodes inserted in visual (i.e. calcarine and pericalcarine) and auditory regions of interest (i.e. Heschl's gyrus and planum temporale) were identified as COIs.

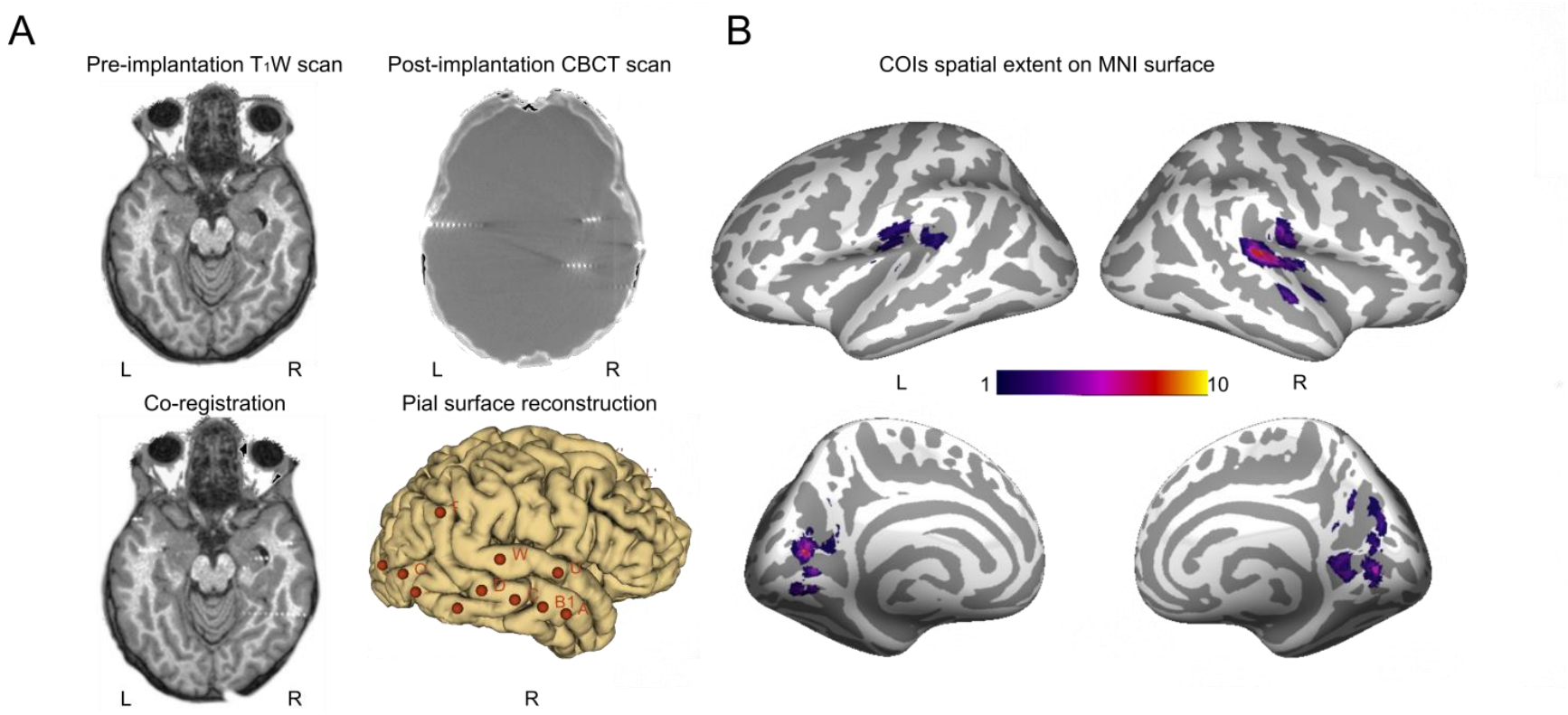

Figure 2. (A) Exemplar of pre-implantation T1W scan, post-implantation CBCT scan and simultaneous visualization of both co-registered datasets from one of the studied subjects. From the same participant, pial surface reconstruction with fiducial markups at electrode entry points is also depicted; (B) Spatial extension of COIs from all participants superimposed on MNI surface (auditory cortex: n = 47; visual cortex n = 44). COIs = contacts of interest; L = left; R = right.

2.5. Data acquisitions and preprocessing

Continuous SEEG was recorded by means of EEG-1200 Neurofax (Nihon Kohden), comprising 192 channels (1000 Hz sampling rate). A medical epileptologist (R.M.) visually inspected the raw signal and did not detect any pathological activity in any investigated COI. We performed all the analyses with Fieldtrip software package (http://www.fieldtriptoolbox.org/), using a unipolar montage (Quinn et al. 2014; Mercier et al. 2015): a contact in the white matter was used as a neutral reference. The raw signal was resampled at 500 Hz. Each trial was detrended, epoched in a time-window of 900 ms pre to 800 ms post-stimulus onset and high-pass filtered (1Hz). Trials showing artefacts were removed. The pre-processed and cleaned data were used for subsequent time-frequency analyses and statistics.

2.6. Time-frequency analyses: power domain

2.6.1. Power domain: time-frequency analyses and statistics

Time-frequency analysis was performed by convolving the pre-processed signal of each individual trial with complex Morlet wavelets in a time-window in steps of 10 ms. Signal decomposition was performed in two different frequency-windows: 5-30 Hz and 30-200 Hz in order to optimize the trade-off between temporal and frequency precision (Cohen 2014).

For the 5-30 Hz frequency window, the power values of each condition of interest (i.e. A, V, AV and N) were estimated in steps of 2 Hz with wavelet widths ranging from 4 to 5 cycles. For the 30-200 Hz frequency window, the power values were estimated in steps of 5 Hz with wavelet widths ranging from 5 to 10 cycles. For both frequency-windows, the wavelet widths changed linearly as a function of frequency. In these analyses we used as output function the power-spectra (http://www.fieldtriptoolbox.org/).

A baseline correction was performed: Wilcoxon signed-rank tests were used to compute z-scores of the power difference between the post-stimulus activity and the baseline activity (500-300ms pre-stimulus onset) of each condition for each single COI-frequency-time point.

Statistical analyses were performed on the first 150 ms after stimulus onset and separately on the alpha (5-13 Hz), beta (13-30 Hz), gamma (30-80 Hz), and high-gamma band (80-200 Hz) by means of cluster-based permutation tests (Maris & Oostenveld 2007), as implemented in Fieldtrip toolbox. Briefly, the conditions of interest were compared by two-tailed paired-samples *t*-tests. Statistically significant samples ($p<0.05$) were clustered based on temporal, spatial and spectral proximity. For each identified cluster the *t*-values were summed to compute cluster-level statistics. The maximum of the cluster-level statistics was taken as the test statistic. Significance probability was computed by Monte Carlo simulations (1000 permutations). Conditions were considered significantly different, if the probability of the maximum cluster mass was $p<0.05$. By

means of the statistical analyses (table 1), we assessed the presence of: 1) intramodal input processing, comparing the intramodal input with the control input (A vs. N in the auditory cortex; V vs. N in the visual cortex); 2) cross-modal input processing, comparing the cross-modal input with the control input (A vs. N in the visual cortex; V vs. N in the auditory cortex); 3) MSI input processing, comparing the MSI input with intramodal input eliciting the maximum response in that cortex (AV vs. A in the auditory cortex; AV vs. V in the visual cortex), in agreement with the maximum model (Stein & Meredith, 1993). This model assessed whether the neural activity was enhanced or depressed during the cross-modal condition in comparison to the intramodal condition, without any assumption about the type of integration of the two stimuli; 4) non-additive MSI processing, comparing the sum of the MSI input and control input with the sum of the intramodal input and cross-modal input (AV+N vs. A+V), in agreement with the additive model (Stein & Meredith, 1993). To apply this model, we used fictitious data obtained summing the power values of each single COI-trial-time-frequency point of AV condition with each single COI-trial-time-frequency point of N condition (AV+N). We did the same for A and V conditions (A+V). Importantly, we did not use the approach proposed by Senkowski et al. (Senkowski et al. 2007): in this method the authors randomly summed the pre-processed signals of the trials from the two unimodal conditions, and only then performed the time-frequency analyses. However, we believe that with this method both the non-phase locked activity and the phase-locked activity differing in phases between conditions would be lost during the sum of the two pre-processed signals.

The additive model assessed whether the neural activity during the cross-modal condition presented sub-additive or supra-additive effects, making the assumption of non-linear integration of the two stimuli when combined together. The equation includes the control input in order to control possible unknown cognitive factors present across all the stimuli, which would be present

once in the left part of the equation, but twice in the right part of the equation (Teder-Sälejärvi et al. 2002; Talsma & Woldorff 2005; Mishra et al. 2007; Gondan et al. 2005; Bonath et al. 2007).

Along the lines of previous electrophysiological studies in the field (Christoph Kayser et al. 2008), any COI showing the presence of the intramodal input response was classified as a functional COI (fCOI). Moreover, fCOIs were labelled as: 1) '*unimodal fCOIs* ', when they responded only to the presence of the intramodal input but not to the cross-modal and MSI input (as assessed with both maximum and additive models); 2) '*MSI unimodal fCOIs* ', when they responded to the intramodal stimulation but not to the cross-modal stimulation, and presented significant differences when tested for the maximum and/or additive model; furthermore, these fCOIs were defined as '*MSI unimodal additive fCOIs*' when showing statistical differences when tested for the maximum model but not for the additive model, while '*MSI unimodal non-additive fCOIs*' when showing statistical differences when tested for the additive model; 3) '*MSI bimodal fCOIs* ', when they responded to both the intramodal and cross-modal stimulation presented separately (Kayser et al., 2008). These fCOIs might or might not present a significant modulation of the intramodal stimulation due the presence of the cross-modal stimulation (i.e. during MSI input) as assessed by the maximum model and the additive model; therefore they were further defined '*MSI only bimodal fCOIs*', when they did not show differences when tested for both the maximum model and the additive model, '*MSI bimodal additive fCOIs*', when presenting differences when tested for the maximum model but not when tested for the additive model, '*MSI bimodal non-additive fCOIs*' when they showed statistical differences when tested for the additive model.

To characterize the results in the alpha, beta, gamma and high-gamma band, we assessed the numbers of the classified fCOIs separately for the two cortices in each single frequency band across all the participants and we expressed them as percentages of COIs or of fCOIs.

To test for differences between the two cortices in the numbers of fCOIs and of MSI fCOIs, we collapsed the number of these COIs across the different frequency bands, and we applied two-tailed Fisher's exact tests *(p*<0.05, Bonferroni corrected). To describe whether fCOIs showed enhanced or depressed responses and sub-additive or supra-additive responses during the MSI input, we computed the enhancement index ([(AV-A)/(AV+A)]*100[1] and the additivity index [(AV+N)-(A+V)]/ [(AV+N)+(A+V)]*100) (Kayser et al., 2008). Specifically for each fCOI, z-scores, obtained from the baseline correction, were averaged between 0 and 150ms after stimulus onset in the alpha, beta, gamma and high-gamma band. The averaged z-scores of each fCOIs were used to compute the enhancement and additive index. fCOIs were classified as enhanced (enhancement index>0) or depressed (enhancement index<0); and sub-additive (additivity index<0) or supra-additive (additivity index>0). To summarize these results, we assessed, within each cortex in the different frequency bands of all the participants, the number of fCOIs showing enhancement or depression and of fCOIs showing sub-additive or supra-additive responses.

To evaluate differences between the auditory and visual cortices in the numbers of enhanced and depressed fCOIs and in the numbers of sub-additive and supra-additive fCOIs, we used two-tailed Fisher's exact tests (*p*<0.05, Bonferroni corrected).

[1] This formula accounts for calculating multisensory enhancement in auditory cortex. In the visual cortex, the formula is then ([(AV-V)/(AV+V)]*100

	CONTRASTS used for statistical comparison	
COIs labels	**AUDITORY CORTEX**	**VISUAL CORTEX**
fCOIs	A vs. N	V vs. N
unimodal fCOIs	only A vs. N	only V vs. N
MSI unimodal fCOIs		
MSI unimodal add. fCOIs	A vs. N & AV vs. A	V vs. N & AV vs. V
MSI unimodal non-add. fCOIs	A vs. N & AV vs. A & AV+N vs. A+V	V vs. N & AV vs. V & AV+N vs. A+V
MSI bimodal fCOIs		
MSI only bimodal fCOIs	A vs. N & V vs. N	V vs. N & A vs. N
MSI bimodal add. fCOIs	A vs. N & V vs. N & AV vs. A	V vs. N & A vs. N & AV vs. V
MSI bimodal non-add. fCOIs	A vs. N & V vs. N & AV vs. A & AV+N vs. A+V	V vs. N & A vs. N & AV vs. V & AV+N vs. A+V

Table 1.

Contacts of interest (COIs) definition and relative contrasts used for statistical comparisons. fCOIs = functional contacts of interest; MSI = multisensory integration; add.= additive; non-add. = non additive; A = auditory input; V = visual input; AV = audio-visual input.

2.6.2. Power domain: time-frequency results

A summary of these data is presented in fig. 3A and B and table 2; response exemplars of two fCOIs are provided in fig. 4.

fCOIs (i.e. all the COIs responding to the intramodal stimulation) were present across all the investigated frequency bands in both cortices (from 83% to 100% of all the auditory COIs; from 20% to 61% of all the visual COIs), with smaller percentages in the visual cortex (Fisher's exact test, $p<0.001$). Despite this functional asymmetry, both cortices showed very high percentages of fCOIs in the high-gamma band (auditory cortex: 100%; visual cortex: 61%) and the smallest in the alpha band (auditory cortex: 83%; visual cortex: 20%). Importantly, the two cortices showed *MSI fCOIs* (both MSI *unimodal fCOIs* and *MSI bimodal fCOIs*) across all the frequency bands (from 30% to 57% of the auditory fCOIs, from 11% to 41% of the visual fCOIs), with no differences between them in the total numbers of these fCOIs (Fisher's exact test, not significant (n.s.)). However, the two cortices showed the largest percentages of *MSI fCOIs* in different frequency bands: MSI expressed mostly in the high-gamma band (57% of fCOIs) for the auditory cortex and in the beta band for the visual cortex (41% of fCOIs). Further reinforcing this observation, the auditory cortex always showed greater proportions of MSI fCOIs in comparison to the visual cortex in all the frequency bands, except for the beta band where this effect was inverted. Importantly, this preferential MSI activity was mainly sustained by additive mechanisms (85% of MSI auditory fCOIs in the high-gamma band; 100% of MSI visual fCOIs in the beta band). Notably, in each sensory cortex, *MSI bimodal fCOIs* revealed that intramodal and cross-modal inputs processing were present simultaneously in each investigated frequency band (from 9% to 26% of the auditory fCOIs; from 11% to 26% of visual fCOIs) in particular in the high-gamma band (26% of fCOIs in auditory as well as in visual cortex). However, an important exception was constituted by the gamma band in the visual cortex, where MSI occurred only by the means of *MSI unimodal fCOIs*.

Interestingly, these fCOIs were characterized by non-additive responses. Very importantly*, MSI bimodal fCOI* and *MSI unimodal fCOIs* showed a different spectral profile: the *MSI bimodal fCOIs* expressed their activity mainly in the high-gamma band (26% of fCOIs in both cortices), while the *MSI unimodal fCOIs* mainly in the gamma band (33% of fCOIs in the visual cortex and 34% of fCOIs in the auditory cortex). Moreover, *MSI only bimodal fCOIs* (i.e. fCOIs responding to both intramodal and cross-modal input but without showing any difference when tested for the maximum model and the additive model), were present in the auditory cortex in small percentages across all the frequency bands (from 2% to 10% of fCOIs), while in the visual cortex they were present in a more consistent percentage but only in the high-gamma band (15% of fCOIs).

The enhancement index (fig.3C), showed that the two cortices differed significantly in the total numbers (across all frequency bands) of enhanced and depressed fCOIs (Fisher's exact test =0.009, $p<0.05$): indeed, in the auditory cortex the number of enhanced fCOIs was slightly superior (53% of fCOIs) in comparison to depressed fCOIs, while in the visual cortex the number of depressed fCOIs (69% of fCOIs in visual cortex) was slightly superior in comparison to enhanced fCOIs. Despite these differences, it was possible to observe that only the visual cortex showed a very strong tendency towards depressed fCOIs in the alpha band (89% of fCOIs) during the MSI input processing when compared to the intramodal input. The two cortices did not differ significantly for the total number of sub- and supra-additive fCOIs (Fisher's exact test, n.s., 72% of supra-additive fCOIs in the auditory cortex and 50% in the visual cortex).

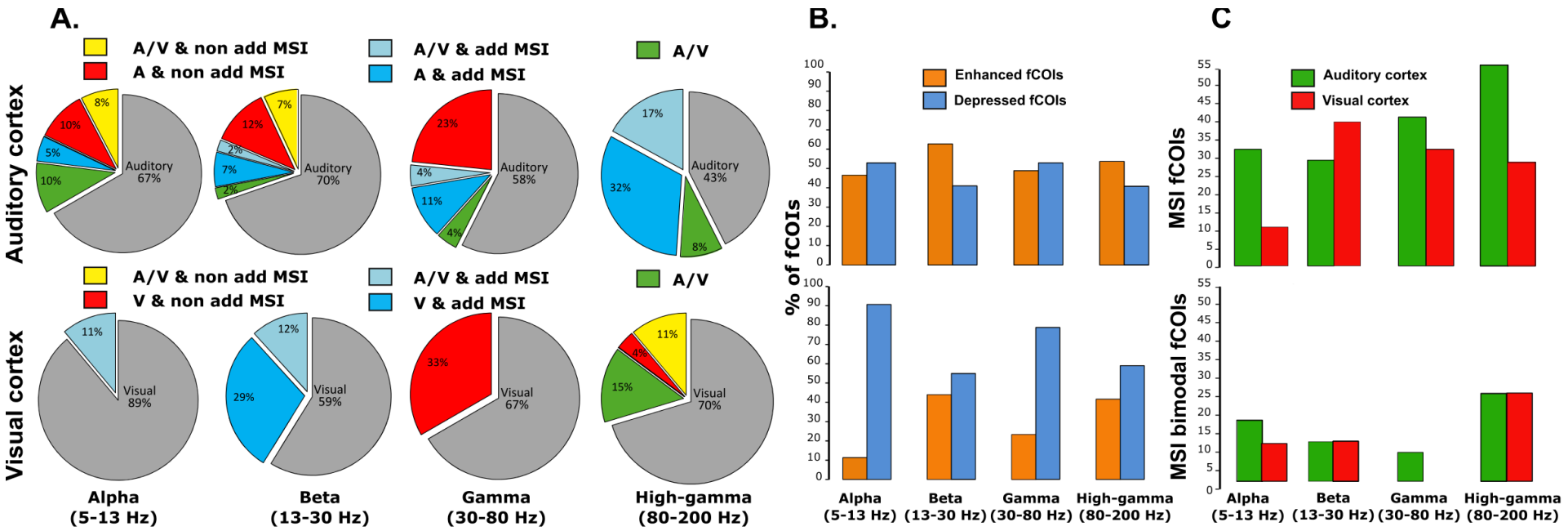

Figure 3. (A) Distribution of the classified functional contacts of interest (fCOIs) across the investigated frequency bands (alpha: 5-13 Hz; beta: 13-30 Hz; gamma: 30-80 Hz; high gamma: 80-200 Hz) in early visual (pericalcarine) and auditory areas (Heschl's gyrus, planum temporale). Gray: fCOIs responding only to intramodal stimuli; remaining colors: fCOIs showing some multisensory modulations. (B) Distribution of fCOIs responding to both intramodal and cross-modal stimuli (MSI bimodal fCOIs: fCOIs responding to the intramodal and cross-modal stimulation presented separately) and distribution of fCOIs presenting multisensory integration effects (MSI unimodal fCOIs: fCOIs responding to the intramodal stimulation but not to the cross-modal stimulation and presenting a significant modulation of the intramodal stimulation due the presence of the crossmodal input; MSI bimodal fCOIs: fCOIs responding to the intramodal and cross-modal stimulation presented separately presenting or not a significant modulation of the intramodal stimulation due the presence of the crossmodal input). (C) Percentages of enhanced (enhancement index>0) and depressed (enhancement index<0) fCOIs during audio-visual stimulation as assessed by means of the enhancement index ([(AV-A)/(AV+A)]*100. Visual = fCOIs responding to the intramodal visual stimuli; Auditory = fCOIs responding to the intramodal auditory stimuli; A/V =

fCOIs responding to the intramodal and crossmodal stimuli; V & add MSI = visual fCOIs responding to the intramodal visual stimuli only and presenting additive MSI effects; A & add MSI = auditory fCOIs responding to the intramodal stimuli only and presenting additive MSI effects; V & non-add MSI = visual fCOIs responding to the intramodal stimuli only and presenting non-additive MSI effects; A & non-add MSI = auditory fCOIs responding to the intramodal stimuli and presenting non-additive MSI effects; A/V & add MSI = fCOIs responding to the intramodal and crossmodal stimuli and presenting additive MSI effects; A/V & non-add MSI = fCOIs responding to the intramodal and crossmodal stimuli and presenting non-additive MSI effects.

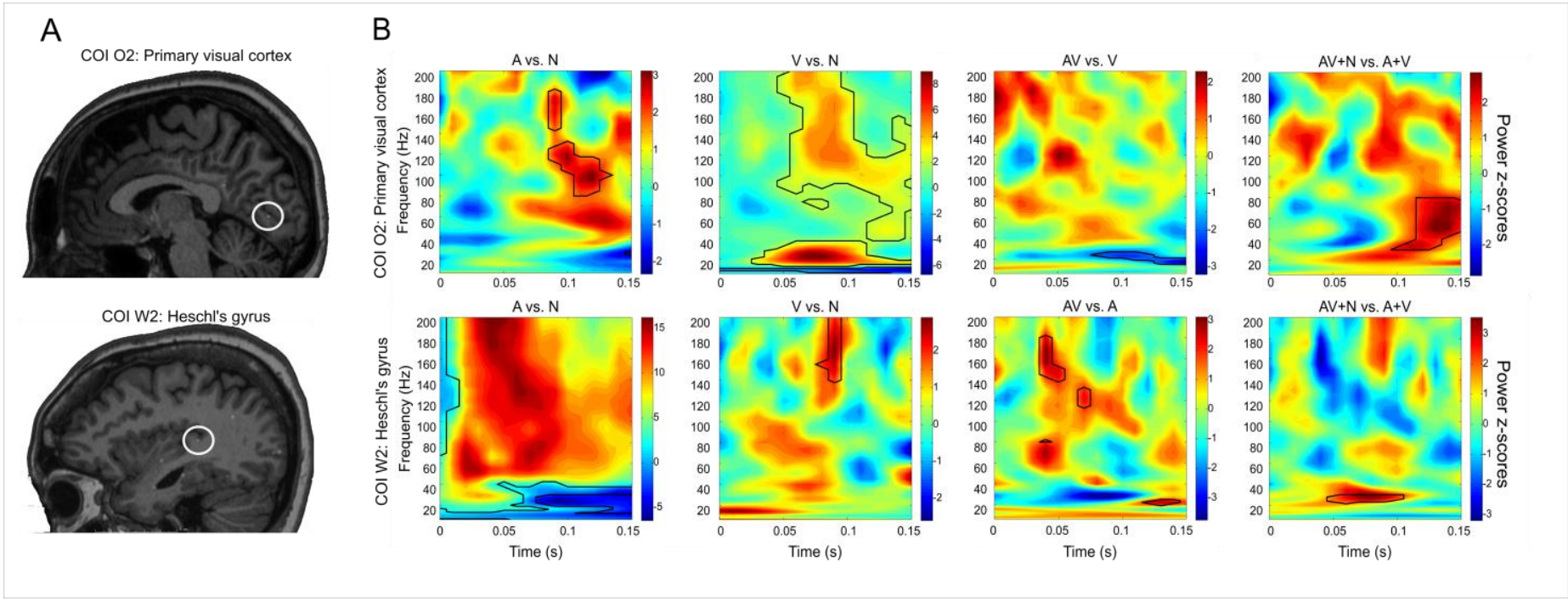

Figure 4. Response exemplars of two functional contacts of interest (fCOIs): (A). MRI images with superimposed CBCT scan showing the localization of the fCOIs (top figure: Heschl's gyrus; bottom figure: calcarine scissure). (B) Time-frequency representations of the auditory condition, visual condition, audio-visual condition and audio-visual condition + control condition expressed as baseline corrected z-scores in the relative fCOI. Solid black lines represent significant responses for the relative comparisons.

Both fCOIs in V1 and Heschl's gyrus, responded to both visual and auditory conditions in the high-gamma band. The visual fCOI showed MSI effect when tested for the maximum model in the beta band, while the auditory fCOI in the high gamma band. Both fCOIs presented additive MSI effect in these two frequency bands. The visual fCOIs showed that the gamma band presented effects only for the matching stimulation accompanied by MSI non-additive effects.

	5-13 Hz		13-30Hz		30-80 Hz		80-200Hz	
	AUD cort	VIS cort	AUD cort	VIS cort	AUD cort	VIS cort	AUD cort	VIS cort
n. total COIs	47	44	47	44	47	44	47	44
fCOIs*	39(83%)	9(20%)	43(91%)	17(39%)	47(100%)	15(34%)	47(100%)	27(61%)
unimodal fCOIs*	26(55%)	8(18%)	30(63%)	10(23%)	27(57%)	10(23%)	20(43%)	19(43%)
unimodal fCOIs**	26(67%)	8(89%)	30(70%)	10(59%)	27(57%)	10(67%)	20(43%)	19(70%)
MSI fCOIs**	13(33%)	1(11%)	13(30%)	7(41%)	20(43%)	5(33%)	27(57%)	8(30%)
tot. MSI unimodal fCOIs**	6(15%)	0(0%)	8(19%)	5(29%)	16(34%)	5(33%)	15(32%)	1(4%)
MSI unimodal add. fCOIs**	2(5%)	0(0%)	3(7%)	5(29%)	5(11%)	0(0%)	15(32%)	0(0%)
MSI unimodal non-add. fCOIs**	4 (10%)	0(0%)	5(12%)	0(0%)	11(23%)	5(33%)	0(0%)	1(4%)
tot. MSI bimodal fCOIs**	7(18%)	1(11%)	5(12%)	2(12%)	4(9%)	0(0%)	12(26%)	7(26%)
MSI only bimodal fCOIs**	4(10%)	0(0%)	1(2%)	0(0%)	2(4%)	0(0%)	4(8%)	4(15%)
MSI bimodal add. fCOIs**	0(0%)	1 (11%)	1(2%)	2(12%)	2(4%)	0(0%)	8(17%)	0(0%)
MSI bimodal non-add. fCOIs**	3(8%)	0(0%)	3(7%)	0(0%)	0(0%)	0(0%)	0(0%)	3(11%)
tot. MSI additive fCOIs***	2(15%)	1(100%)	4(31%)	7(100%)	7(35%)	0(0%)	23(85%)	0(%)
tot. MSI non-additive fCOIs ***	7(54%)	0(0%)	8(62%)	0(0%)	11(55%)	5(100%)	0(0%)	4(50%)
enhanced fCOIs**	22(56%)	1(11%)	26(60%)	9(56%)	23(49%)	6(40%)	31(66%)	13(48%)
depressed fCOIs**	17(44%)	8(89%)	17(40%)	7(44%)	24(51%)	9(60%)	16(34%)	14(51%)

Table 2. Numbers and percentages of the relevant contacts of interest (COIs) expressed relatively to the total number of identified COIs (*), to the total numbers of fCOIs(**), to the total number of MSI fCOIs (***) in each sensory cortex for each frequency band. fCOIs = functional contacts of interest; add. = additive; non-add = non-additive; MSI = multisensory integration.

2.7. Time-frequency analyses: phase concentration index

2.7.1. Phase concentration index: time-frequency analyses and statistics

Phase concentration index (PCI) (Tallon-Baudry et al. 1996; Makeig et al. 2004) is a measure of phase angles distribution across trials of the EEG spectral phase in the polar space at each single time-frequency point (Makeig et al. 2002). Mathematically, it is the length of the average of the vectors of the phase angles: it is comprised between zero (phase angles uniformly distributed across the polar space) and one (phase angles perfectly aligned) (Cohen 2014). From the neurophysiological point of view, the phase seems to reflect the timing of the level of activity of the underlying population of neurons (Cohen 2014; Kayser et al. 2008).

Across the relevant conditions, we were interested in assessing the presence of increased PCI and of the 'phase resetting' phenomenon (Lakatos, Chen, O'Connell, Ai. Mills, et al. 2007), defined as an increase in phase concentration (as showed by the phase concentration index - PCI) not accompanied by a difference in power in the same frequency band (Shah et al. 2004).

First, we performed the time frequency analyses: signal decompositions were performed as for the time-frequency analyses in the power domain, but in this case we used as output function the complex Fourier-spectra. As described for the analyses in power domain, we convolved the pre-processed signal of each individual trial with complex Morlet wavelets in steps of 10 ms. Then, signal decompositions were performed in two different frequency-windows: 2-30 Hz and 30-80 Hz, in order to optimize the trade-off between temporal and frequency precision (Cohen 2014). For the 2-30 Hz frequency window, the complex Fourier-spectra values of each condition of interest (i.e. A, V, AV and N) were estimated in steps of 2 Hz with wavelet widths ranging from 4 to 5 cycles. For the 30-80 Hz frequency window, the complex Fourier-spectra values were estimated in steps of 5 Hz with wavelet widths ranging from 5 to 10 cycles. For both frequency-windows, the wavelet widths changed linearly as a function of frequency.

Second, we computed the PCI in each single time-frequency point in a time-window from 0 to 300ms post-stimulus onset for each condition (A, V, AV and N) and the statistical differences between the conditions of interest (A vs. N, V vs. N, AV vs. N., AV vs. A and AV vs. V), separately for alpha ([5:13] Hz), beta ([13:30] Hz), and gamma band ([30:80] Hz). We did not extend our analyses to the high gamma band because of the unlikely possibility that phase concentration is maintained for several cycles in the high-frequency bands (Cohen 2014).

The complex output was used to compute PCI in each single time-frequency point (Mercier et al. 2013) in a time-window from 0 to 300ms post-stimulus onset. To this end, for each single trial the complex result of the wavelet convolution was normalized by its amplitude. The mean of the normalized values across trials (comprised from 0 and 1) is a representation of the phase concentration across trials (i.e. PCI) in each single time-frequency point.

We computed between-conditions differences in PCI in the time-window from 0 to 300ms post-stimulus onset, using the function 'ft_statfun_diff_itc' available in Fieldtrip. This function allows to compute the significance of the difference in PCI between two conditions. This is obtained using the Monte-Carlo procedure (1000 permutations): by randomly shuffling the trials between the two conditions and computing repeatedly the PCI difference between the two conditions, a PCI distribution of this difference is obtained for each single time-frequency point. The obtained PCI distribution then was used to compute the statistical significance of the difference in PCI between the two conditions ($p < 0.05$) by means of one-tailed paired-samples t-tests. Also in this case, statistically significant samples ($p<0.05$) were clustered-corrected based on temporal, spatial and spectral proximity to control for the problem of multiple comparisons.

Third, to test for the presence of the pure phase resetting phenomenon in each relevant condition (A, V and AV), we identified the time-frequency point, in each frequency band, corresponding to the maximum PCI value, emerged as significant in the relevant comparisons (A vs. N, V vs. N, AV

vs. N., AV vs. A, AV vs. A). Then, we assessed whether the same time-frequency point was significant also power (Mercier et al. 2013).

2.7.2. Phase concentration index: time-frequency results

A resume of these data is present in table 3 and fig. 5 and 6.

Our results showed that very high proportions of COIs in both cortices presented increased PCI (with or without detectable differences in power at the same time-frequency point) across all the investigated frequencies bands during intramodal (i.e. A input in the auditory cortex and V input in the visual cortex) and audio-visual conditions (intramodal input: from 87% to 100% of the auditory COIs; from 61% to 80% of the visual COIs; audio-visual input: 100% of the auditory COIs, from 50% to 82% of the visual COIs). It is possible to observe that there are no substantial differences between the intramodal condition and the audio-visual condition in the percentages of COIs showing increase in PCI, as previously observed by Mercier et al. (2013) in the visual cortex.

The percentages of COIs showing increased PCI during cross-modal condition, as expected, were smaller in comparison to the intramodal condition. For the visual alone condition, we observed an increased PCI in 19% of the auditory COIs in alpha band, 32% in beta band and 40% in gamma band. For the auditory alone condition, we observed an increased PCI in 25% of visual COIs in alpha band, and 11% in gamma band. No increase in PCI was observed in any visual COIs in the beta band.

As a second step, we tested the presence of the pure resetting phenomenon: therefore we observed, if significant increased PCI was accompanied by an increased power (Mercier et al. 2013). During the intramodal condition, we observed that, in each frequency band of interest (i.e. alpha, beta and gamma), from 64% to 88% of the auditory COIs and from 29% to 33% of the visual COIs showing an increase in PCI were accompanied by an increase in power. Similarly, during the AV condition from 60% to 87% of the auditory COIs and from 32% to 36% of the visual COIs showing an increase in PCI also showed an increase in power. Also in this case, the intramodal and audio-visual condition showed similar percentages of visual and auditory COIs.

During the crossmodal condition, a small proportion of auditory COIs showing increased PCI also showed increased power: this effect involved the 11% of the auditory COIs in both alpha (1 out 9) and gamma band (2 out 19). Importantly no auditory COIs showed this effect in the beta band despite the presence of 15 COIs showing increased PCI. Interestingly, no visual COIs showing an increase in PCI was accompanied by an increase in power in the alpha and gamma band (the beta band, showed no visual COIs presenting any increase in PCI, as we observed above).

To assess the presence of the effect of MSI processing on PCI, we tested the maximum model (AV vs. A in the auditory cortex; AV vs. V in the visual cortex). In both cortices, we observed increased PCI restricted to the alpha band and the gamma band (13% of auditory COIs in the alpha and gamma band; 30% of visual COIs in the alpha band, 7% in the gamma band). No differences in PCI were detected in the beta band in both cortices. Importantly, in visual and auditory COIs showing significant differences in PCI when tested for the maximum model, the audio-visual processing was always characterized by increased PCI, in comparison to the intramodal processing. This clearly indicates that the presence of the audio-visual condition organizes the phase of the ongoing oscillations in a stronger way in comparison to the intramodal condition in the alpha and gamma band in both cortices.

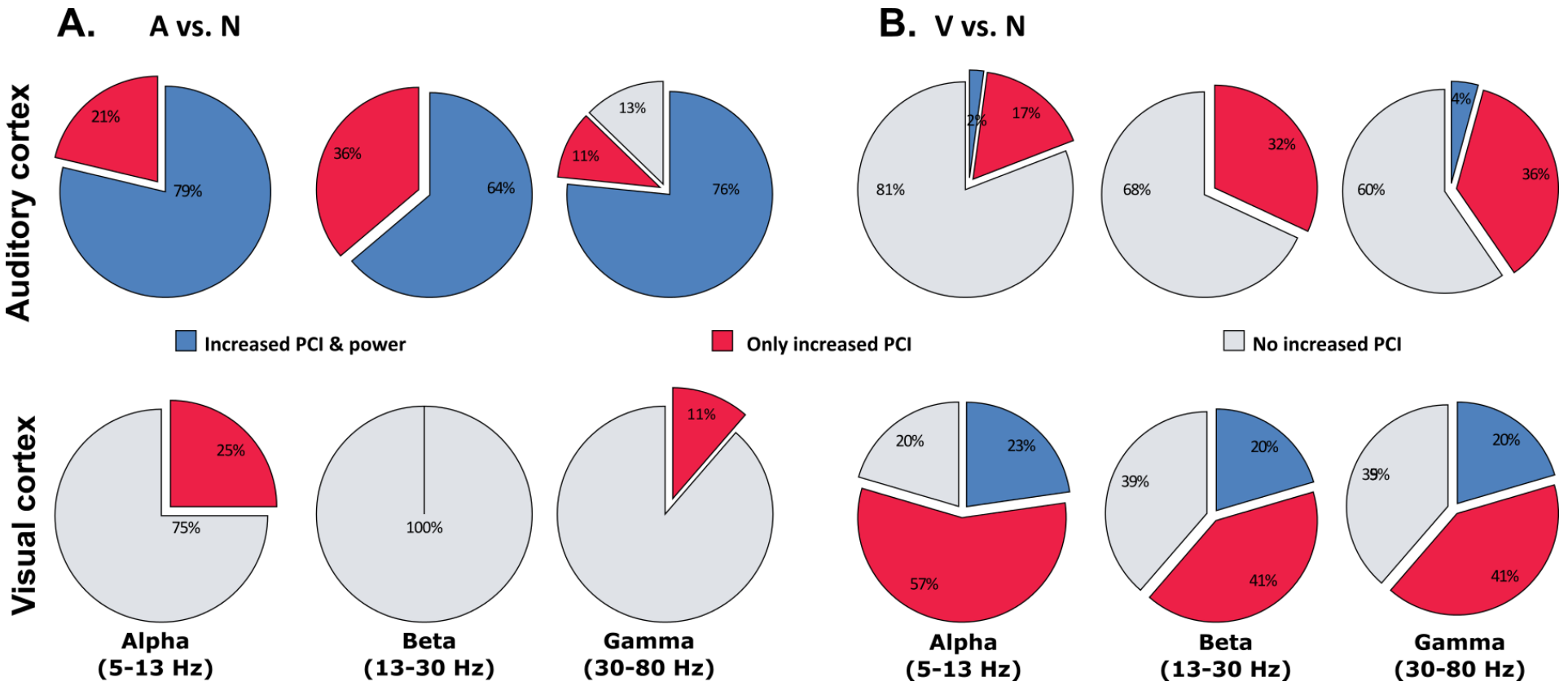

Figure 5. (A) Distribution of the contacts of interest (fCOIs) across the investigated frequency bands showing evidence of only increased phase concentration index (PCI), increased in PCI and power, no increased PCI during the comparison of the auditory condition vs the null condition (alpha: 5-13 Hz; beta: 13-30 Hz; gamma: 30-80 Hz) in early visual (pericalcarine) and auditory areas (Heschl's gyrus, planum temporale).

(B) Distribution of the contacts of interest (fCOIs) across the investigated frequency bands showing evidence of only increased PCI, increased in PCI and power, no increased PCI during the comparison of the visual condition vs the null condition (alpha: 5-13 Hz; beta: 13-30 Hz; gamma: 30-80 Hz) in early visual (pericalcarine) and auditory areas (Heschl's gyrus, planum temporale).

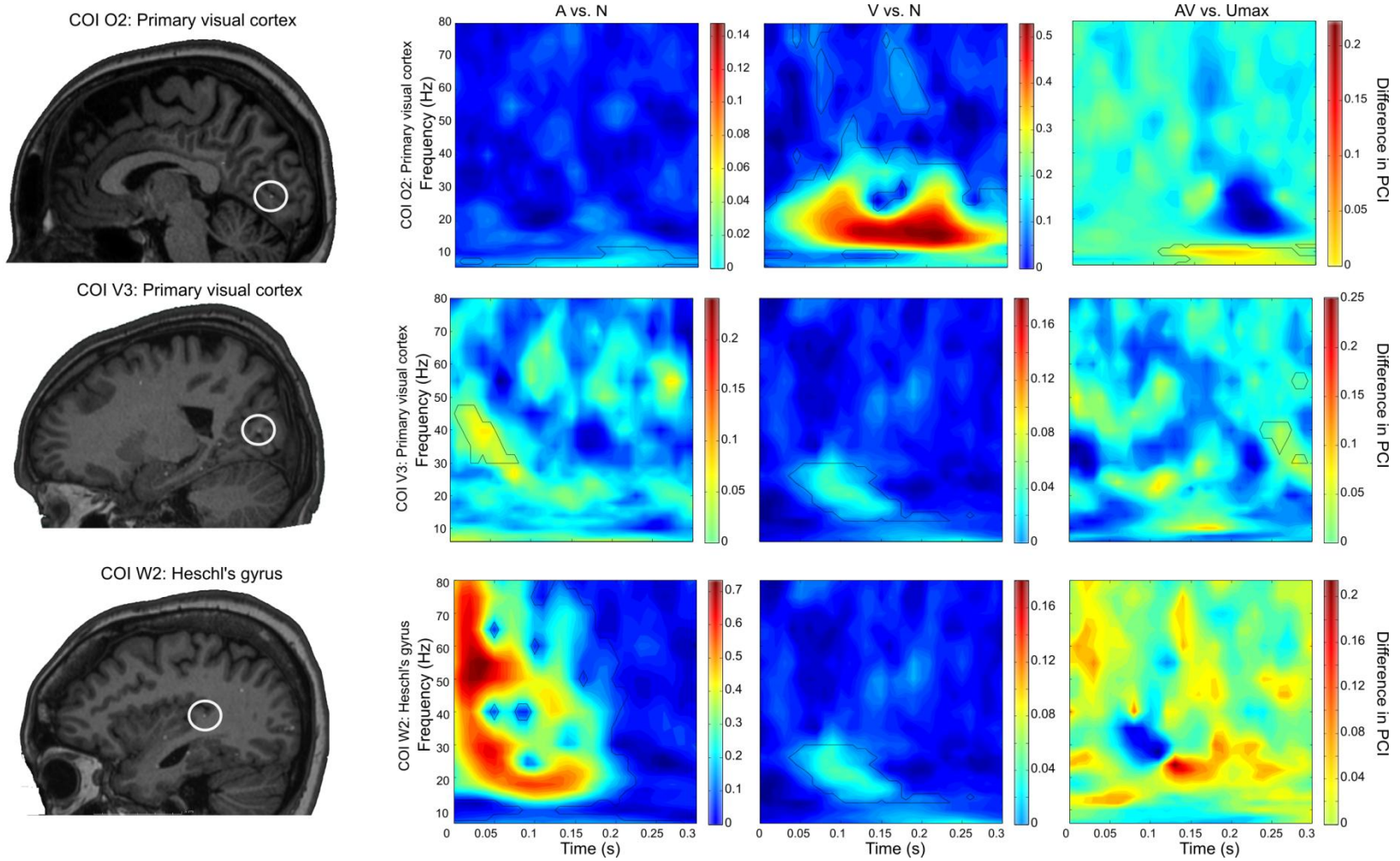

Figure 6. On the left: MRI images with superimposed CBCT scan showing the localization of the fCOIs (top and middle figure: primary visual cortex; bottom figure: Heschl's gyrus). On the right: Phase concentration index (PCI) differences between relevant conditions (A vs. N, V vs. N, AV vs. Umax). Solid black lines represent significant differences for the relative comparisons. A = auditory; N = null; V = visual; AV = audio-visual; Umax = unimodal condition eliciting the maximum response in that cortex.

	5-13 Hz					13-30Hz					30-80 Hz				
	A vs. N	V vs. N	AV vs. A	AV vs. V	AV vs. N	A vs. N	V vs. N	AV vs. A	AV vs. V	AV vs. N	A vs. N	V vs. N	AV vs. A	AV vs. V	AV vs. N
	INCREASE IN PHASE CONCENTRATION INDEX*														
AUD cort	46(98%)	5(11%)	4(9%)	44(94%)	46(98%)	45(96%)	11(23%)	0(0%)	46(98%)	46(98%)	39(83%)	14(30%)	5(11%)	40(85%)	47(100%)
VIS cort	10(23%)	35(80%)	28(64%)	8(18%)	33(75%)	0(0%)	25(57%)	18(41%)	0(0%)	21(48%)	4(9%)	23(52%)	12(27%)	2(5%)	19(43%)
	INCREASE IN PHASE CONCENTRATION INDEX AND POWER**														
AUD cort	37(80%)	0(0%)	2(50%)	28(64%)	32(70%)	30(67%)	0(0%)	0(0%)	30(65%)	33(72%)	35(90%)	2(14%)	0(0%)	32(80%)	41(87%)
VIS cort	0(0%)	10(29%)	13(46%)	2(25%)	12(36%)	0(0%)	9(36%)	7(39%)	0(0%)	6(29%)	0(0%)	9(39%)	9(75%)	0(0%)	8(42%)

Table 3. Numbers and percentages of the relevant contacts of interest (COIs) showing increased concentration index (PCI) expressed relatively to the total number of identified COIs (*), and increased PCI and power expressed relatively to the total number of COIs(**), showing increased PCI in each sensory cortex for each frequency band for the relevant contrasts. COIs = functional contacts of interest.

2.8. iERP (intracranial event related potentials)

2.8.1. iERPs: analyses and statistics

To complement our analyses, we computed iERPs. The pre-processed LFP signals were baseline corrected (0.200-0.050s pre-stimulus onset), low-pass filtered (30 Hz) and averaged across trials for each condition of interest. Between-conditions differences were assessed comparing the post-stimulus amplitude (0- 0.150s post-stimulus onset) between the relevant conditions as showed in Table 1. To test for significance, as for the time-frequency analyses, unpaired two-samples *t*-tests were implemented: statistically significant samples($p < 0.05$), were clustered based on temporal proximity. Then, each cluster-level statistic was determined summing the t-values within each identified cluster and taking the maximum cluster-level statistics. The significance probability was computed by Monte Carlo method (1000 re-sampling of the original data). Clusters were considered significant for $p < 0.05$.

2.9.1. **iERPs: results**

A resume of these data is present in fig. 7 and 8. The results showed that 38 COIs (81% of anatomical auditory COIs) in the auditory cortex and 29 COIs in the visual cortex (66% of anatomical visual COIs) responded to the intramodal stimulation (fCOIs).

The two cortices presented different MSI profiles. In the early auditory cortex MSI sites were constituted only by fCOIs responding to both intramodal and cross-modal stimulation (14% of fCOIs); notably, these fCOIs showed no differences when tested for the maximum model (AV vs. A). In the early visual cortex, MSI was observed in fCOIs showing different activity only when tested for the maximum model (AV vs. V) (21% of fCOIs), in this case no fCOIs responded to both intramodal and cross-modal stimulation. All these MSI visual fCOIs presented an enhanced

activity. Interestingly, when we tested the additive model (AV+N vs A+V), we found no evidence of non-linear process during the audio-visual stimulation in both cortices.

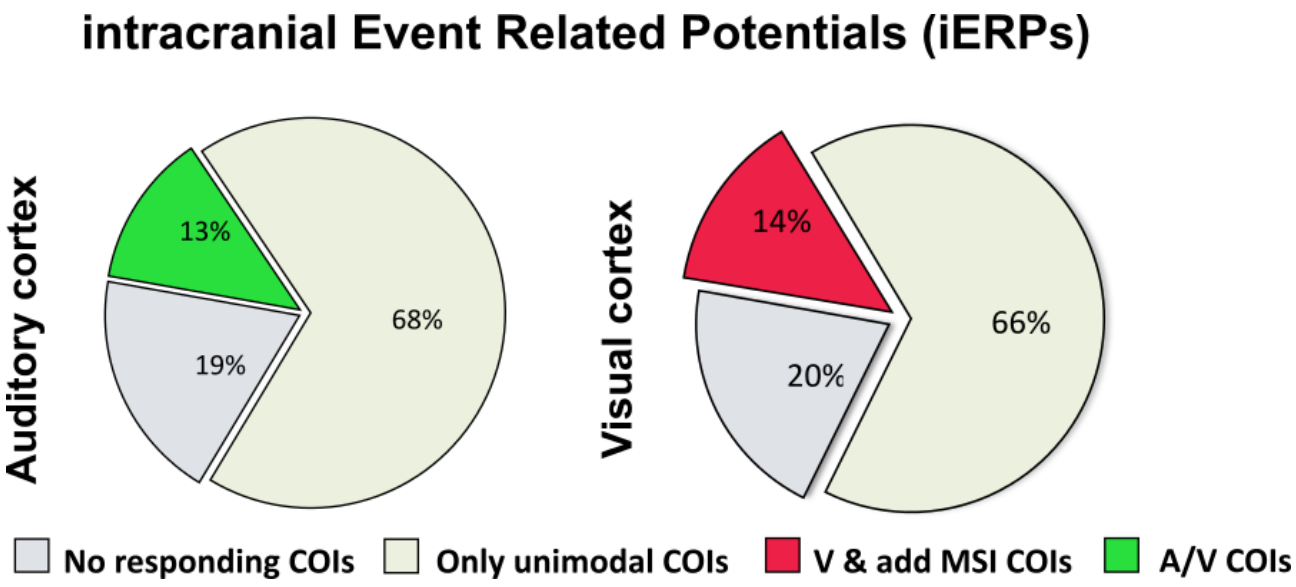

Figure 7. Distribution of the contacts of interest showing evidence of significant intracranial event related potentials (iERPs in early visual (pericalcarine) and auditory areas (Heschl's gyrus, planum temporale).No responding COIs: COIs showing no evidence of intramodal responses; Only unimodal COIs: COIs showing evidence of only intramodal responses; A/V COIs: COIs showing responses for both intramodal and cross-modal stimulation; V & add MSI: COIs showing responses for intramodal stimulation (V) and when tested for the maximum model (AV vs V).

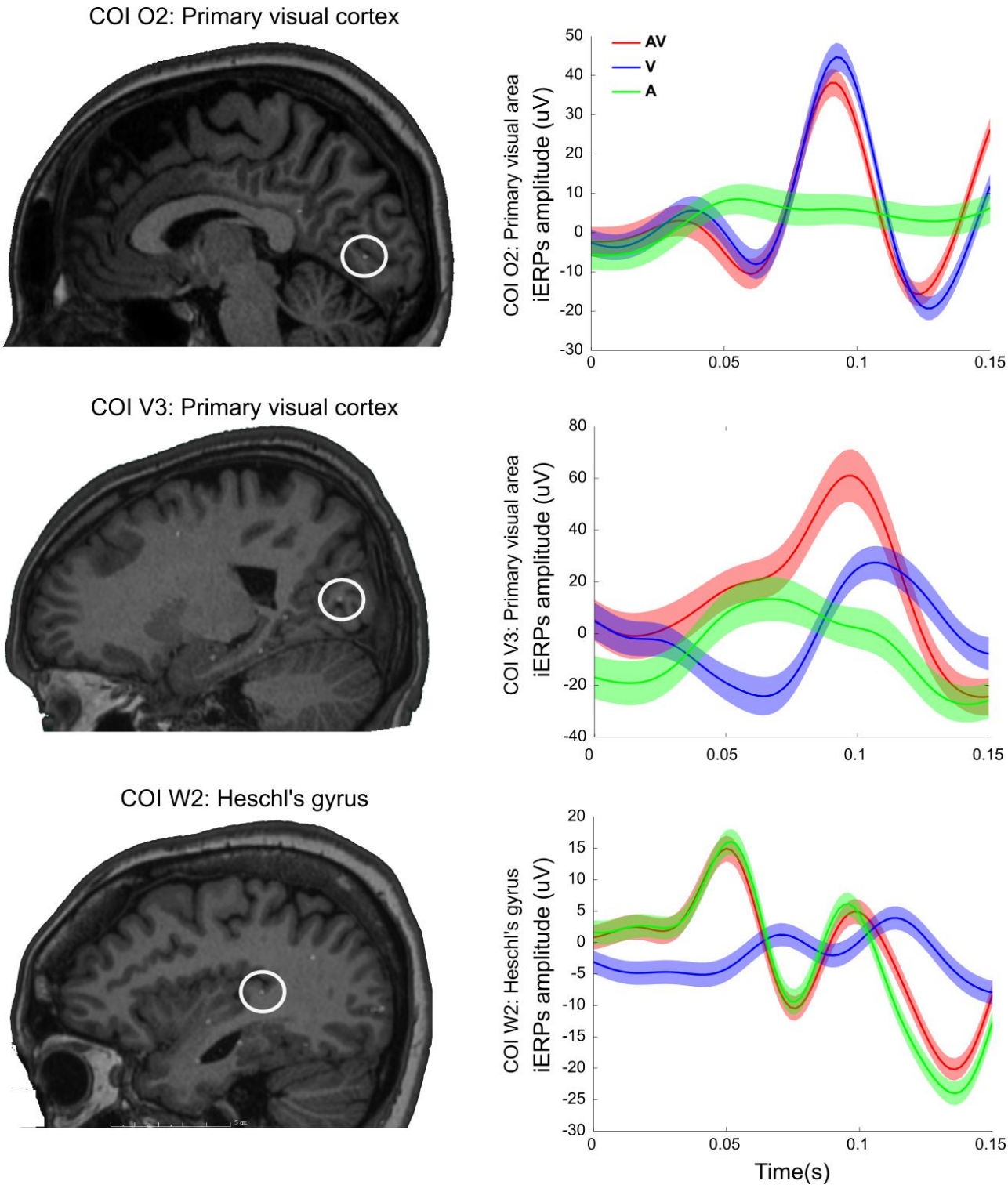

Figure 8. Response exemplars of three contacts of interest (fCOIs). On the left: MRI images with superimposed CBCT scan showing the localization of the fCOIs (top figure: Heschl's gyrus; bottom figure: calcarine scissure). On the right: intracranial event related potentials (iERPs) of the auditory condition, visual condition, audio-visual condition.

3. Discussion

We investigated the spectral fingerprints of cross-modal and MSI processing in early human visual (calcarine and pericalcarine regions) and auditory (Heschl's gyrus and planum temporale) brain areas. To this end, we conducted time-frequency analyses to investigate both power and phase of the intracranial EEG signal. We conducted the power analyses within the first 150ms after stimulus onset across different frequency bands (theta/alpha: 5-13 Hz; beta: 13-30 Hz; gamma: 30-80 Hz; high-gamma; 80-200 Hz) and the phase analyses within the first 300ms after stimulus onset in a narrow range of frequency bands (theta/alpha: 5-13 Hz; beta: 13-30 Hz; gamma: 30-80 Hz). Moreover, we complemented our study computing time-locked analyses within the first 150ms after stimulus onset. To this end, we relied on the unique spatio-temporal resolution of SEEG..

We provide compelling evidence that the activity of both early visual and auditory cortices is modulated by cross-modal input and shows MSI in both power and phase domain.

Crucially, we identified markers of cross-modal and MSI processing in the spectral fingerprints of local oscillations, possibly indicating different specific neurophysiological mechanisms at work during these neural computations.

In particular, we showed that cross-modal processing induces important modulations in the power of the high-gamma band and an organization of the oscillatory activity across all the investigated frequency bands in both cortices (except in the beta band of the visual cortex). The most part of the COIs showing evidence of post-stimulus organized oscillatory activity, showed a profile compatible with the pure phase resetting phenomenon (increase in PCI with no increase in power in the same time-frequency point).

Moreover, we showed that MSI induces power modulations in the early visual cortex mainly in the beta band (13-30 Hz) (41% of fCOIs), and in the early auditory cortex mainly in the high-gamma band (80-200Hz) (57% of auditory fCOIs), although it was also observed, to a lesser

extent, across all the other investigated frequency bands (auditory cortex: gamma: 43% of auditory fCOIs; alpha: 33%; beta: 30%; visual cortex: gamma: 33% of visual fCOIs; high-gamma: 30%; alpha: 11%). We also showed that MSI organizes the oscillatory activity in theta/alpha and gamma band in both cortices, but not in the beta band. Despite the organization of oscillatory activity in specific COIs in the theta/alpha band, a spread depression in the theta/alpha power was present in the visual cortex (86% of fCOIs).

Interestingly, iERPs analyses showed evidence of cross-modal responses only in the early auditory cortex but not in the early visual cortex. However, when COIs were tested for the maximum model (AV vs. U_{max}), the visual cortex, but not the auditory cortex, showed significant COIs (14% of the visual COIs).

Power domain: spectral fingerprints of cross-modal processing

Several neurophysiological studies in humans have showed that cross-modal input modulates the activity of early sensory regions (M.H. Giard & Peronnet 1999; Molholm et al. 2002; Raij et al. 2010; Mercier et al. 2013; Brang et al. 2015). Mercier et al. (2013) evidenced, for the first time in humans, that cross-sensory stimulation (auditory) resets the phase of ongoing activity without an increase in power (i.e. pure phase resetting) in visual areas. This mechanism, evidenced initially in primate primary auditory cortex (Lakatos, Chen, O'Connell, A. Mills, et al. 2007) with tactile inputs, is thought to play a key-role in cross-modal and MSI processing (Kayser 2009; Ghazanfar & Chandrasekaran 2007).

Our findings show another mechanism by which cross-modal processing occurs in early sensory areas. Indeed, we showed that cross-modal inputs modulate mainly the power of the high-gamma band in fCOIs (i.e. COIs responding to intramodal stimulation in that frequency band) of both early sensory cortices. This frequency band, prone to be extensively investigated by means

of intracranial recordings (Lachaux et al. 2012), is a direct and robust index of neuronal spiking synchrony reflecting local cortical processing (Buzsáki et al. 2012). Hence, in our work, the fact that cross-modal input mainly modulated the high-gamma band in fCOIs, is an evidence of the presence and of the activity of multisensory neuronal populations in early sensory regions. Although this is the first report of high-gamma band modulation during cross-modal processing in human early sensory cortices, these findings are not unexpected being in line with the long tradition of animal studies demonstrating the presence of cross-modal neurons in putative unisensory cortices (Bizley et al. 2007; Brosch et al. 2005; Kayser et al. 2008; Morrell 1972).

Power domain: spectral fingerprints of MSI processing

Despite both sensory cortices showed MSI across all the investigated frequency spectrum, there is a striking difference between the visual and auditory cortex: the visual cortex showed the greatest percentages of MSI fCOIs in the beta band, and to a lesser extent in the gamma and high-gamma band; while the auditory cortex in the high-gamma band and, to a lesser extent, across all the other investigated frequency bands.

Although we still lack of a comprehensive and strong theoretical framework to link patterns of different frequency oscillations with specific neural mechanisms, recent animal studies (van Kerkoerle et al. 2014; A. M. Bastos et al. 2015) have provided empirical evidence to the hypothesis that feedback and feedforward processing travel on different frequency channels (Wang 2010).
Van Kerkoerle et al. (2014), investigating the laminar pattern of different frequency oscillations in the primate visual cortex, revealed that low-order processing areas (V1) drove the activity of high-order processing areas (V4) by means of the gamma band, while high-order processing areas (V4) influenced the activity of low-order processing areas (V1) by means of the alpha rhythm. Similarly, Bastos et al. (2015), showed that, in primate visual cortex, the gamma rhythms of V1 influenced

the gamma rhythms of high-order processing area (7A). Interestingly, differently from Van Kerkoerle et al. (2014), Bastos et al. (2015) reported that feedback processing travels on the beta rhythm. Importantly, all the above findings seem to hold also in the human auditory cortex: a SEEG study revealed the predominance of delta-beta band activity for feedback processing and of gamma band activity for feedforward processing (Fontolan et al. 2014). All together, these very important results revealed precise inter-areal influences (in visual and auditory cortex) by means of specific frequency channels: the low frequency bands (in particular beta rhythms) seem to preside over feedback interactions, while the gamma frequency band over feedforward interactions (Fries 2015; Zheng & Colgin 2015). This model is confirmed by electrophysiological laminar studies that showed that supragranular layers, origin of feedforward connections (Felleman & Van Essen 1991) present mainly gamma-band oscillations, while infragranular layers, origin of the feedback pathways (Felleman & Van Essen 1991), beta band oscillations (Buffalo et al. 2011).

Despite the puzzling pattern of different frequency bands modulations emerged from our findings during MSI, it is possible to make very important inferences based on the above studies. These inferences are possible assuming that the observed mechanisms for inter-areal interactions might be at work also for cross-modal interactions. This assumption is based on animal anatomical studies showing strong evidence of direct (monosynaptic) heteromodal connections: the primary visual cortex, in particular regions of the peripheral visual field, was showed to receive direct projections from the auditory cortex (Rockland & Ojima 2003), mainly from the parabelt auditory areas and, in a small fraction, from A1 (Falchier et al. 2002; Hall & Lomber 2008; Henschke et al. 2015), while primary auditory cortex was showed to receive direct connections from the visual regions (Bizley et al. 2007; Cappe & Barone 2005; Falchier et al. 2010) comprising V1 (Henschke et al. 2015; Bizley et al. 2007).

Our results showed that, during MSI, the early visual cortex presented the most important power modulations in the beta band. In line with the above studies, we speculate that feedback interactions, expressed as modulations of the beta band power, might be the predominant phenomenona in the early visual cortex during MSI. These feedback interactions might originate during the dynamic interplay between the investigated early sensory cortices. This possibility is in good agreement with laminar studies (Falchier et al. 2010; Rockland & Ojima 2003; Falchier et al. 2002; Markov et al. 2014) showing the preponderant presence of feedback connections in the visual cortex and globally suggesting that early visual and auditory areas interact via a feedback loop (Falchier et al. 2010). Moreover, the evidence of modulations of the gamma band power in the early visual cortex, although to a lesser extent than beta modulations, might indicate feedforward interactions, possibly originate during interactions with high-processing visual areas. Notably, the gamma band in the visual cortex emerged as a prominent frequency band for MSI processing: this frequency band, differently from the other bands, responded only to input from the dominant modality (i.e. intramodal input), but showed power modulation during MSI. This might corroborate the hypothesis that these feedforward interactions are entrained with high-order visual processing areas during MSI. Notably, the possible feedback and feedforward interactions in the visual cortex are accompanied by a widespread, although subthreshold, depression in the alpha band in the early visual cortex, as showed by the enhancement index, and by modulations of neuronal spiking activity, as indexed by the relative strong activity in the high-gamma band. The activity in the alpha and in the high-gamma band seems to be strictly related: decreased alpha oscillations, biasing local cortical excitability, are strongly linked to local information processing (de Pesters et al. 2016; Klimesch et al. 2007; Romei et al. 2010), as showed by the observations that alpha oscillations modulate the firing rate of neuronal populations (Haegens et al. 2011). Importantly, increased alpha band activity in the visual cortex reflects a

functional cortical inhibition preventing the communication among connected neuronal populations (Zumer et al. 2014): our results therefore clearly showed that MSI is supported by active involvement of connected neuronal networks.

The early auditory cortex, differently from the early visual cortex, during MSI processing showed the most important power modulation in the high-gamma band: this result clearly indicates robust neuronal activity and emphasizes the role of local processing mechanisms during the integration of intramodal and cross-modal information. Interestingly, we also observed a prevalence of feedforward interactions, expressed as modulations of the gamma band power, although, to a lesser degree, also feedback interactions were present. Based on the proposal that the most part of projections between visual and audiotry areas seem to be of feedback type, we speculate that feedback interactions are entrained with the early visual cortex, while the feedforward interactions with the high-order auditory processing areas.

Based on the above results, an asymmetry during MSI between the early visual cortex and the early auditory cortex clearly emerged: the early visual cortex presented more feedback interactions, compatible with modulatory effect, while the early auditory cortex more feedforward interactions, compatible with driving effect, accompanied by a strong local cortical processing.

Phase domain: spectral fingerprints of cross-modal processing and MSI

Our findings indicate that the cross-modal input in the early auditory and visual cortex induces phase-locked oscillatory responses resetting the phase of ongoing oscillations in several COIs. Notably, the two sensory cortices presented different profiles: in the auditory cortex increased PCI was distributed across all the different frequency bands, in particular in the gamma (40% of COIs) and beta band (32% of COIs); in the visual cortex it was more prevalent in the theta/alpha band (25% of COIs), but totally absent in the beta band.

Importantly, very few or no COIs were accompanied by power modulations in the same time-frequency point showing the highest PCI value in the investigated frequency band (theta/alpha, beta and gamma band). This suggests a pure resetting mechanism at work during cross-modal processing. In contrast, the intramodal input processing was accompanied, in the most part of the COIs, by increase in power in both cortices. Our findings are consistent with the results obtained by Mercier et al. (2013) in the human visual cortex by means of ECoG and with the results from animal studies in the primary auditory cortex (Kayser et al. 2008; Lakatos et al. 2007).

Mercier et al. (2013) reported evidence of phase-locked oscillatory responses to auditory stimulation in the human visual cortex. They observed that few sites with increased PCI across the different frequency bands (theta/alpha, beta, gamma) also showed increase in power, therefore suggesting that the great part of these COIs presented a typical profile of phase resetting (increase in PCI with no increase in power). Although we replicated the presence of this possible mechanism in the early visual cortex, our results showed that the oscillatory activity of the beta band is not modulated by the presence of the cross-modal input.

Lakatos et al. (2007), analyzing phase-locked oscillations in monkey primary auditory cortex (A1), showed that the cross-modal condition induces phase-locked oscillations with specific phase angles with very low amplitudes. According to previous studies, phase-locked oscillations might be originated by a stimulus-evoked response and/or by a stimulus-induced phase resetting (Makeig et al. 2004). In particular, stimulus-evoked responses are accompanied by increased power, while, stimulus-induced phase resetting are originated by an increase in phase synchrony across trials, but not by an increase in power (Makeig et al. 2004). Based on these notions, Lakatos et al. (2007) provided the first evidence that cross-modal processing induces a pure phase-resetting of ongoing oscillations (i.e. presence of phase locked oscillations characterized by very low amplitude). This

organization would lead to a certain degree of susceptibility of the neuronal population to subsequent inputs. Also Kayser et al. (2008) found results compatible with a phase resetting model in the primary auditory cortex of monkeys. These authors made two important observations: 1) the auditory response to the intramodal stimulation depends from the phase of the low-frequency oscillations (in particular from 5 to 10 Hz); 2) the enhanced cross-modal response correlates with the phase consistency across trials in the low-frequency bands and with the prevalence of trials showing an 'optimal phase' (i.e. phase that induces an optimal cortical excitability). The authors concluded stating that their results are compatible with the hypothesis that the intramodal activity is modulated by the phase of low-frequency oscillations, which in turn can be modulated by the presence of other sensory inputs, providing therefore a mechanism for MSI.

Testing the maximum model (AV vs. U_{max}), we showed that the audio-visual condition presented different PCI in comparison to the intramodal input processing in several auditory and visual COIs. Importantly, our results showed that the audio-visual condition was characterized by increased PCI in comparison to the intramodal input processing. Generally speaking, this finding again is in agreement with the results of Mercier et al. (2013). It is important to note however, that these authors found greater percentages of visual COIs showing significant PCI differences when testing the maximum model (beta band: 47% of COIs; gamma band: 49%, alpha band: 19%) in comparison to our results model (gamma band: 5%, alpha band: 18%). Moreover, our findings showed that the most interested frequency band by difference in PCI between the audio-visual condition and the visual condition in the visual cortex was the theta/alpha band, with a small involvement of the gamma band. In contrast with Mercier et al. (2013) results, no effect was present in the beta band.

To investigate phase resetting mechanisms, we used the same temporal window and the same frequency bands (theta/alpha, beta and gamma band) used in the work by Mercier et al. (2013).

This choice was done to replicate the results of these authors. Despite this, some inconsistencies, in particular relative the involvement of the beta band in phase resetting, emerged.

Several explanations might account for the observed differences: 1) Mercier et al. (2013) asked the participants to indicate the detection of any kind of inputs (auditory, visual or audio-visual) providing a motor response. Differently, in our task, we asked the participants a motor response only when they detected the rare targets. Notably, we analyzed only the non-target inputs to avoid possible response modulations due to the presence of the motor response. Importantly, our results are consistent with the neurophysiological studies in the primary auditory cortex, where the monkeys attended passively or responded only to rare target input, and where no evidence of increase in PCI was reported in the beta band (Lakatos, Chen, O'Connell, Ai. Mills, et al. 2007; Christoph Kayser et al. 2008); 2) Mercier et al. (2013) explored several regions of the visual cortex, comprising high-order visual processing areas, by means of ECoG, while we explored only early sensory areas (V1 and V2) with SEEG.

Despite we confirm the presence of pure-phase resetting during the cross-modal conditions, it is important to keep in mind two aspects: first, that the phase resetting of ongoing oscillatory activity are much more detectable than increase in power due to their statistical properties (Ding & Simon 2013), second that a causal relation between cross-modal interactions and oscillatory activity is still to be confirmed (Christoph Kayser et al. 2008).

Intracranial event related potentials (iERPs) of cross-modal processing and MSI

We complemented our analyses computing iERPs. Our results clearly showed that the two cortices present different MSI profiles: the auditory cortex showed that some COIs responding to the intramodal input also responded to the cross-modal input (14% of auditory fCOIs); this was not the case for the visual cortex. Interestingly, when tested for the maximum model (i.e. AV vs V), the early visual cortex showed a different activity between the audio-visual condition and the intramodal condition, while the early auditory cortex did not.

The finding that MSI responses can be present in sites where there is no evidence of cross-modal responses is consistent with previous animal (Lakatos, Chen, O'Connell, A. Mills, et al. 2007; Meredith & Allman 2009; Christoph Kayser et al. 2008) and human studies (Mercier et al. 2013). As observed by Mercier et al. (2013), these results fit well with the relatively recent observations of a new class of neurons (named subthreshold neurons) responding to intramodal but not to cross-modal input, and showing a vigorous response (different from the intramodal response) when stimulated with a multisensory input (Allman et al. 2009; Meredith & Allman 2009). Despite this speculation, it is important to note that we have explored the LFPs , and for this reason it is difficult to make a clear parallel with single-neuron study that provided evidence of subthreshold neurons (Allman et al. 2009; Meredith & Allman 2009).

The discrepancy between the visual and auditory areas can be interpreted based on the principle of inverse effectiveness. This principle, stating that the less effective is the intramodal input response the stronger is the response to multisensory input, was showed to hold at single-neuron, LFPs and scalp-EEG level (Stein & Meredith 1993; Christoph Kayser et al. 2008; Avillac et al. 2007; Senkowski et al. 2011): testing the maximum model, the finding of MSI effect in the visual cortex but not in the auditory cortex might be explained by the fact that the used visual input was possibly less effective in eliciting the activity in the visual cortex (66% of visual COIs responding to

the visual stimulation), in comparison to the white noise in the auditory cortex (81% of the auditory COIs responding to the auditory stimulation).

Importantly no effect was detected when we tested auditory and visual COIs for the additive model (AV+N vs. A+V), suggesting that the MSI processing are mainly linear, at least in early sensory regions and in the first 150 ms after stimulus onset. Our results therefore conform to the observations that super-additivity and sub- additivity responses are very rarely met at the population level (Stein et al., 2004) and that linear responses are quite common in multisensory neurons (Stein and Stanford, 2008).

These results are intriguing: the most part of the previous neurophysiological literature detected the presence of non-linear bimodal responses comparing the signal of the bimodal condition with the sum of the signals of the unimodal conditions alone (i.e. B vs U_1+U_2). In our case, following the observations of Besle et al. (2004), to control for possible unknown cognitive factor, we compared the sum of the signals during the bimodal condition and the control condition with the sum of the signals of the unimodal conditions (AV+N vs. A+V). This equation includes the control condition: in its original form (i.e. B vs U_1+U_2) possible unknown cognitive factors common to all these stimuli would be added twice in the right part of the equation, but would be present just one time on the left side of the equation (Teder-Sälejärvi et al. 2002; Talsma & Woldorff 2005; Mishra et al. 2007; Gondan et al. 2005; Bonath et al. 2007). It is clear that in its original form the additive model could lead to false positive. Notably EEG literature using the additive model to test for MSI in detection tasks, showed contradicting findings. Several studies (Fort et al. 2002; M.H. Giard & Peronnet 1999; Talsma et al. 2006) showed evidence of early audio-visual interactions in visual and auditory cortex by means of the additive model, but other studies did not (Molholm et al. 2002; Talsma & Woldorff 2005; Teder-Sälejärvi et al. 2002). Although

several speculations can be done, such as the nature of tasks (Besle et al. 2009), it is important to consider the possible effect of the problems raised by Besle et al (2004).

Thorne et al. (2011), presented evidence that the visual input resets the phase of oscillatory activity in theta and alpha frequencies., while Naue et al. (2011), showed that the stimulus onset of audio-visual stimuli modulated the beta responses compatibly with the presence of phase-resetting mechanism. Despite the importance of

Limitations

As all the study in humans taking advantage of intracranial recordings, we investigated a population of chronic epileptic patients. For this reason, we cannot exclude that the observed activity might be related to abnormal functional processing. However, all the patients did not show apparent deficits in visual and auditory functions. Moreover, we collected sEEG data at least three days after the surgery, 48 hours after or before the seizures and we focalized the analyses only in recordings sites far from the regions where epileptic activity was identified to occur and where there was no evidence of dysplasic tissues.

Another limitation could be the lack of matching between the audio and visual stimuli. However, we chose simple stimuli to avoid possible confounds due to the difference of salience across modalities. Indeed, neurophysiological activity induced by high-level stimuli, although matched for some aspect, would have been modulated by multiple cognitive aspects (such as recognition, semantic, emotional valence etc.) difficult to control across the different senses. Purposely relying on these simple, highly salient and meaningless stimuli (checkerboard and white-noise), known to trigger large response in the corresponding sensory modality, ensures that the salience of the auditory and visual stimuli are equal, avoiding the confound of the "matching" between the senses. Another possible limitation of our study is that the position of the COIs was not matched

between the two sensory areas. As explained above, for obvious ethical reasons, COIs positions were determined only by clinical purposes. This reduces the possibility to directly compare the two early sensory cortices.

Another limitation is constituted by the possible presence of volume conduction effect that would justify the important presence of MSI, when we tested for the maximum model (AV vs U_{max}), in the early visual and auditory cortices. Volume conduction can occur also in iEEG, however, as recently published, auditory LFPs generated in the auditory cortex might extends around 10 mm: therefore they might contaminate an auditory ERP recorded in the secondary somatosensory regions and in the visual and multisensory regions of the superior temporal sulcus (Kajikawa & Schroeder 2011). Importantly, we investigated two areas quite distant and we used very simple stimuli that should not involve higher order sensory areas (such as visual areas of the superior temporal sulcus). For this reason we think that volume conduction had no effect on our results.

4. Conclusion

Despite a great number of studies has begun to shed some lights on the complex MSI mechanisms, MSI remains still elusive in its fundamental as clearly showed by the negative findings reported by two recent studies (Quinn et al. 2014; Lemus et al. 2010) questioning the presence of MSI in primary auditory and visual cortices.

Exploring the spectral fingerprints of cross-modal interactions and early MSI processing we have showed the contribution of power and phase modulations of different frequency bands to these processes in human early visual and auditory cortices. We further demonstrate that, in addition to spectral manifestation common to both early auditory and visual cortices, MSI also possibly expresses via region-specific neurophysiological mechanisms. Our results, clearly emphasizing that crossmodal processing and MSI are distributed processes (Kayser et al. 2010), pave the way for future studies. Although we cannot fully exclude that the observed difference in MSI processing between the visual and auditory cortex might be related to the type of stimulation we relied on or to the localization of the investigated COIs, our findings might indicate that the mesoscopic MSI mechanisms may vary across sensory cortices, opening new avenue for future researches in the field.

Important aspects of our study constituted a very strong backbone for the interpretation of the results: the method used to collect the neurophysiological signal (sEEG) and the used paradigm.

Intracranial electroencephalographies record local field potentials with excellent temporal and spatial resolution (Lachaux et al. 2003). The high signal to noise ratio and the high resistance to possible artefacts (muscle contractions, eye blinks) present in the common used electrophysiological techniques in humans, allow a unique perspective over human brain processes. Notably, the high-gamma frequency band, strictly linked to neuronal spiking synchrony, constitutes the most important band for the application of these techniques, because of its

inaccessibility to MEG and EEG recordings (Lachaux et al. 2012). In our study we used sEEG: differently from electrocorticography, it allows to directly investigate the folded and not superficial regions of the brain, such as the calcarine scissure and Heschl's gyrus. These characteristics put us in the best possible condition to investigate early MSI processing in humans

As for as the used paradigm, several aspect can explain our compelling results. First of all, during the audio-visual condition, the visual stimuli lead the auditory stimuli of 30 ms. In the auditory cortex, the neurophysiological interactions of audio-visual stimuli are strongest when the visual stimulus lead by 20-80 ms (Kayser et al. 2008). In agreement, cross-modal phase resetting occurs when the visual input precedes the auditory input by 30-75 ms (Thorne et al. 2011). Second, the use of white noise as auditory stimulation. From a global perspective, a long tradition of studies (Harper 1979; Manjarrez et al. 2007; Lugo et al. 2008; Gleiss & Kayser 2014) 2014) has showed that auditory white noise can facilitate the sensitivity. From the specific perspective of our research, only clinical purposes defined the positions of COIs in our study: the use of white noise, allowed us to stimulate wide neuronal populations tuned to the range of different frequency bands present in the white noise. Therefore, we had greatly increased the possibility to detect strong matching and non-matching input processing, as well as MSI processing in both visual and auditory cortices, as we observed in our study. Third, in the analyzed conditions, (non-target conditions), we avoid the confounding factor of motor responses, whose presence would have increased the possibility of spurious or, at least less interpretable, results. The use of the white noise (highly unstructured and perceived binaurally from all the directions) as auditory stimulation and of a black and white checkerboard as visual stimulation (highly structured and foveally presented) well explain why the visual cortex in comparison to the auditory cortex, presented always smaller proportions of fCOIs in each single frequency band across all the conditions.

5. Concluding remarks and future avenues

To conclude, we can affirm that : 1) cross-modal processing and MSI modulate the power in the high-gamma band. These results allow to infer that in the early sensory cortices there is a neuronal representation of the two different information as suggested by previous animal studies (Christoph Kayser et al. 2008; Morrell 1972b); 2) MSI, in comparison to the intramodal processing, modulate mainly the power of the gamma band in the early auditory cortex and of the beta band in the early visual cortex. These results are suggestive of the possible prevalence of feedback interactions in the early visual cortex and of feedforward interactions in the early auditory cortex. Future studies should asses this possibility. Importantly, the two cortices presented specific spectral profile in power during MSI. This clearly speaks in favour of different mechanisms at the basis of MSI, although we cannot exclude that the position of the COIs and the type of stimulation used might have influences these results; 3) cross-modal processing organizes the oscillatory activity of the ongoing oscillations during cross-modal processing and MSI consistent with the notion of 'pure resetting mechanism' (Lakatos et al. 2007; Kayser et al. 2008; Mercier et al. 2013).

As a next future step, we will investigate the functional connectivity between early visual and auditory cortices during cross-modal processing and MSI using the collected SEEG data. Moreover, we will take advantage of the fact that these same patients, for clinical reasons, underwent to biphasic stimulation at 1 Hz for 30s in each single visual and auditory COIs; therefore we will observe whether perturbations in power and phase of the SEEG signal in the sensory cortex (e.g. early visual cortex) occur when the other sensory modality (e.g. early auditory cortex) is stimulated. In particular, based on the work of van Kerkoerle et al. (2014), we are particularly interested in observing whether and in which frequency band, power modulations might occur. As

a last step, we will verify the consistency of the results obtained from the connectivity analyses and from the analyses of stimulation data.

Bibliography

Adrian, E.D., 1949. *The Sherrington Lectures - Sensory integration*, Liverpool, University Press.

Allison, T. et al., 1999. Electrophysiological studies of human face perception. I: Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cerebral cortex*, 9(5), pp.415–30.

Allman, B.L. et al., 2009. Not Just for Bimodal Neurons Anymore: The Contribution of Unimodal Neurons to Cortical Multisensory processing. *Brain Topography*, 21(3–4), pp.157–167.

Aoki, F. et al., 1999. Increased gamma-range activity in human sensorimotor cortex during performance of visuomotor tasks. *Clinical Neurophysiology*.

Van Atteveldt, N. et al., 2004. Integration of letters and speech sounds in the human brain. *Neuron*, 43(2), pp.271–282.

Avillac, M., Ben Hamed, S. & Duhamel, J.-R., 2007. Multisensory integration in the ventral intraparietal area of the macaque monkey. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 27(8), pp.1922–32.

Barone, P. et al., 2000. Laminar distribution of neurons in extrastriate areas projecting to visual areas V1 and V4 correlates with the hierarchical rank and indicates the operation of a distance rule. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 20(9), pp.3263–3281.

Bastos, A.M. et al., 2015. Visual areas exert feedforward and feedback influences through distinct

Beauchamp, M.S., 2005. Statistical criteria in FMRI studies of multisensory integration. *Neuroinformatics*, 3(2), pp.93–113.

Besle, J., Fort, A., Delpuech, C., et al., 2004. Bimodal speech: Early suppressive visual effects in human auditory cortex. *European Journal of Neuroscience*, 20(8), pp.2225–2234.

Besle, J., Bertrand, O. & Giard, M.H.H., 2009. Electrophysiological (EEG, sEEG, MEG) evidence for multiple audiovisual interactions in the human auditory cortex. *Hearing Research*, 258(1–2), pp.143–151.

Besle, J., Fort, A. & Giard, M.-H.H., 2004. Interest and validity of the additive model in electrophysiological studies of multisensory interactions. *Cognitive Processing*, 5(3), pp.189–192.

Bizley, J.K. et al., 2007. Physiological and anatomical evidence for multisensory interactions in auditory cortex. *Cerebral Cortex*, 17(9), pp.2172–2189.

Bonath, B. et al., 2007. Neural Basis of the Ventriloquist Illusion. *Current Biology*, 17(19), pp.1697–1703.

Brang, D. et al., 2015. Peripheral sounds rapidly activate visual cortex: evidence from electrocorticography. *Journal of Neurophysiology*, (SEPTEMBER), p.jn.00728.2015.

Bremar, F., 1952. Les aires auditives de l'ecorce cerebrale. *Cours international d'audiologie clinique*.

Brosch, M., Selezneva, E. & Scheich, H., 2005. Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. *The Journal of neuroscience*, 25(29), pp.6797–806.

Büchel, C. et al., 1998. Different activation patterns in the visual cortex of late and congenitally blind subjects. *Brain*.

Budinger, E. et al., 2006. Multisensory processing via early cortical stages: Connections of the primary auditory cortical field with other sensory systems. *Neuroscience*, 143(4), pp.1065–1083.

Budinger, E., Heil, P. & Scheich, H., 2000. Functional organization of auditory cortex in the Mongolian gerbil (Meriones unguiculatus). III. Anatomical subdivisions and corticocortical connections. *European Journal of*.

Buffalo, E.A. et al., 2011. Laminar differences in gamma and alpha coherence in the ventral stream. *Proceedings of the National Academy of Sciences of the United States of America*, 108(27), pp.11262–7.

Buschman, T. & Miller, E., 2007. Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *science*.

Buzsáki, G., 2006. *Rhythms of the Brain*,

Buzsáki, G., Anastassiou, C. a & Koch, C., 2012. The origin of extracellular fields and currents--EEG, ECoG, LFP and spikes. *Nature reviews. Neuroscience*, 13(6), pp.407–20.

Buzsaki, G. & Wang, X., 2012. Mechanisms of Gamma Oscillations. *October*, 35(4), pp.203–225.

Calvert, G.A., 1997. Activation of Auditory Cortex During Silent Lipreading. *Science*, 276(5312), pp.593–596.

Calvert, G.A., 2001. Crossmodal processing in the human brain: insights from functional neuroimaging studies. *Cerebral cortex*, 11, pp.1110–23.

Cappe, C. & Barone, P., 2005. Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. *European Journal of Neuroscience*, 22(11), pp.2886–2902.

Cardinale, F. et al., 2013. Stereoelectroencephalography: Surgical methodology, safety, and stereotactic application accuracy in 500 procedures. *Neurosurgery*, 72(3), pp.353–366.

Chandrasekaran, C., Lemus, L. & Ghazanfar, A. a, 2013. Dynamic faces speed up the onset of auditory cortical spiking responses during vocal detection. *Proceedings of the National Academy of Sciences of the United States of America*, 110(48), pp.E4668-77.

Cohen, L. et al., 1999. Period of susceptibility for cross-modal plasticity in the blind. *Annals of*.

Cohen, M.X., 2014. *Analyzing Neural Time Series Data: Theory and Practice*,

Colby, C. & Duhamel, J., 1991. Heterogeneity of extrastriate visual areas and multiple parietal areas in the macaque monkey. *Neuropsychologia*.

Cragg, B., 1969. The topography of the afferent projections in the circumstriate visual cortex of the monkey studied by the Nauta method. *Vision research*.

Crone, N. et al., 1998. Functional mapping of human sensorimotor cortex with electrocorticographic spectral analysis. II. Event-related synchronization in the gamma band. *Brain*.

Diederich, A., Schomburg, A. & Colonius, H., 2012. Saccadic reaction times to audiovisual stimuli show effects of oscillatory phase reset. *PloS one*, 7(10), p.e44910.

Ding, N. & Simon, J., 2013. Power and phase properties of Oscillatory Neural Responses in the Presence of Background Activity. , 18(9), pp.1199–1216.

Donner, T.H. & Siegel, M., 2011. A framework for local cortical oscillation patterns. *Trends in Cognitive Sciences*, 15(5), pp.191–199.

Ekstrom, L., Roelfsema, P. & Arsenault, J., 2008. Bottom-up dependent gating of frontal signals in early visual cortex.

Engel, A.K. & Fries, P., 2010. Beta-band oscillations-signalling the status quo? *Current Opinion in Neurobiology*, 20(2), pp.156–165.

Engel, A.K., Fries, P. & Singer, W., 2001. Dynamic predictions: oscillations and synchrony in top-down processing. *Nature reviews. Neuroscience*, 2(10), pp.704–16.

Falchier, A. et al., 2002. Anatomical evidence of multimodal integration in primate striate cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 22(13), pp.5749–59.

Falchier, A. et al., 2010. Projection from visual areas V2 and prostriata to caudal auditory cortex in the monkey. *Cerebral Cortex*, 20(7), pp.1529–1538.

Fedorov, A. et al., 2012. 3D slicers as an image computing platform for thw quantitative imaging network. *Magnetic resonance imaging*, 30(9), pp.1323–1341.

Feldon, S., Feldon, P. & Kruger, L., 1970. Topography of the retinal projection upon the superior colliculus of the cat. *Vision research*.

Felleman, D. & Burkhalter, A., 1997. Cortical connections of areas V3 and VP of macaque monkey extrastriate visual cortex. *The Journal of*.

Felleman, D.J. & Van Essen, D.C., 1991. Distributed hierachical processing in the primate cerebral cortex. *Cerebral Cortex*, 1, pp.1–47.

Fischl, B., 2012. FreeSurfer. *NeuroImage*, 62(2), pp.774–81.

Fishman, M.C. & Michael, C.R., 1973. Integration of auditory information in the cat's visual cortex. *Vision Research*, (8), pp.1415–1419.

Fodor, J., 1983. The modularity of mind: An essay on faculty psychology.

Fontolan, L. et al., 2014. The contribution of frequency-specific activity to hierarchical information processing in the human auditory cortex. *Nature communications*, 5(May), p.4694.

Fort, A., Delpuech, C. & Pernier, J., 2002. Early auditory – visual interactions in human cortex during nonredundant target identification. , 14, pp.20–30.

Foxe, J.J. et al., 2002. Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. *Journal of neurophysiology*, 88(1), pp.540–543.

Foxe, J.J. et al., 2000. Multisensory auditory-somatosensory interactions in early cortical processing revealed by high-density electrical mapping. *Brain research. Cognitive brain research*, 10(1–2), pp.77–83.

Fries, P., 2015. Rhythms for Cognition: Communication through Coherence. *Neuron*, 88(1), pp.220–235.

Fries, P., Womelsdorf, T. & Oostenveld, R., 2008. The effects of visual stimulation and selective visual attention on rhythmic neuronal synchronization in macaque area V4. *The Journal of*.

Ghazanfar, A.A. & Chandrasekaran, C.F., 2007. Paving the Way Forward: Integrating the Senses through Phase-Resetting of Cortical Oscillations. *Neuron*, 53(2), pp.162–164.

Ghazanfar, A.A. & Schroeder, C.E., 2006. Is neocortex essentially multisensory? *Trends in Cognitive Sciences*, 10(6), pp.278–285.

Ghazanfar, A. a. et al., 2005. Multisensory Integration of Dynamic Faces and Voices in Rhesus Monkey Auditory Cortex. *Journal of Neuroscience*, 25(20), pp.5004–5012.

Giard, M.H. & Peronnet, F., 1999. Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *Journal of cognitive neuroscience*, 11(5), pp.473–90.

Gleiss, S. & Kayser, C., 2014. Acustic noise improves visual perception and modulates occipital oscillatory states. *Journal of Cognitive Neuroscience*, 26(4), pp.699–711.

Gomez-ramirez, M. et al., 2011. Oscillatory Sensory Selection Mechanisms during Intersensory Attention to Rhythmic Auditory and Visual Inputs: A Human Electro-Corticographic Investigation. *J Neurosci*, 31(50), pp.18556–18567.

Gondan, M. et al., 2005. Multisensory processing in the redundant-target effect: a behavioral and event-related potential study. *Perception & psychophysics*, 67(4), pp.713–26.

Gonzalo, J., 1952. Las funciones cerebrales humanas según nuevos datos y bases fisiológicas: Una introducción a los estudios de Dinámica Cerebral. *Trabajos del Instituto Cajal de Investigaciones Biológicas*, 55, pp.157–195.

Gregoriou, G.G. et al., 2009. High-Frequency, Long-Range Coupling Between Prefrontal and Visual Cortex During Attention. *Science*, 324(5931), pp.1207–1210.

Haegens, S. et al., 2011. α-Oscillations in the monkey sensorimotor network influence discrimination performance by rhythmical inhibition of neuronal spiking. *Proceedings of the National Academy of Sciences of the United States of America*, 108(48), pp.19377–82.

Hall, A.J. & Lomber, S.G., 2008. Auditory cortex projections target the peripheral field representation of primary visual cortex. *Experimental Brain Research*, 190(4), pp.413–430.

Harper, D., 1979. Signal detection analysis of effect of white noise intensity on sensitivity to visual flicker. *Perceptual and motor skills*.

Henschke, J.U. et al., 2015. Possible anatomical pathways for short-latency multisensory integration processes in primary sensory cortices. *Brain structure & function*, 220(2), pp.955–977.

Hilgetag, C.C., O'Neill, M.A. & Young, M.P., 1996. Indeterminate organization of the visual system. *Science*, (271), pp.776–777.

Hilgetag, C.C., O'Neill, M. a & Young, M.P., 2000. Hierarchical organization of macaque and cat cortical sensory systems explored with a novel network processor. *Phil Trans R Soc B*, 355(1393), pp.71–89.

Hipp, J., Engel, A. & Siegel, M., 2011. Oscillatory synchronization in large-scale cortical networks predicts perception. *Neuron*.

Hirsch, J.F. et al., 1961. Short and long latencies responses to somesthetic stimulation in the cat. *Experimental Neurology*, 4, pp.562–583.

Horn, G. & Hill, R., 1966. Responsiveness to sensory stimulation of units in the superior colliculus and subjacent tectotegmental regions of the rabbit. *Experimental neurology*.

Hubel, D. & Wiesel, T., 1962. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology*.

Hupe, J. et al., 1998. Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature*.

Ibrahim, L.A. et al., 2016. Cross-Modality Sharpening of Visual Cortical Processing through Layer-1-Mediated Inhibition and Disinhibition. *Neuron*, 89(5), pp.1031–1045.

Iurilli, G. et al., 2012. Sound-Driven Synaptic Inhibition in Primary Visual Cortex. *Neuron*, 73(4), pp.814–828.

Jenkinson, M. & Smith, S., 2001. A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, 5(2), pp.143–156.

Jerbi, K. et al., 2009. Task-related gamma-band dynamics from an intracerebral perspective: Review and implications for surface EEG and MEG. *Human Brain Mapping*, 30(6), pp.1758–1771.

Kajikawa, Y. & Schroeder, C.E., 2011. How local is the local field potential? *Neuron*, 72(5), pp.847–858.

Kayser, C. et al., 2007. Functional Imaging Reveals Visual Modulation of Specific Fields in Auditory Cortex. *Journal of Neuroscience*, 27(8), pp.1824–1835.

Kayser, C. et al., 2005. Integration of touch and sound in auditory cortex. *Neuron*, 48(2), pp.373–84.

Kayser, C., 2009. Phase Resetting as a Mechanism for Supramodal Attentional Control. *Neuron*, 64(3), pp.300–302.

Kayser, C. et al., 2009. Spike-Phase Coding Boosts and Stabilizes Information Carried by Spatial and Temporal Spike Patterns. *Neuron*, 61(4), pp.597–608.

Kayser, C., 2010. The multisensory nature of unisensory cortices: A puzzle continued. *Neuron*, 67(2), pp.178–180.

Kayser, C. & Logothetis, N.K., 2009. Directed Interactions Between Auditory and Superior Temporal Cortices and their Role in Sensory Integration. *Frontiers in integrative neuroscience*, 3, p.7.

Kayser, C., Logothetis, N.K. & Panzeri, S., 2010. Visual Enhancement of the Information Representation in Auditory Cortex. *Current Biology*, 20(1), pp.19–24.

Kayser, C., Petkov, C.I. & Logothetis, N., 2008. Visual modulation of neurons in auditory cortex. *Cerebral Cortex*, 18(7), pp.1560–1574.

van Kerkoerle, T. et al., 2014. Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proceedings of the National Academy of Sciences*, 111(40), pp.14332–41.

King, A. & Palmer, A., 1985. Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Experimental Brain Research*.

Klimesch, W., Sauseng, P. & Hanslmayr, S., 2007. EEG alpha oscillations: The inhibition-timing hypothesis. *Brain Research Reviews*, 53(1), pp.63–88.

Kopell, N. et al., 2000. Gamma rhythms and beta rhythms have different synchronization properties. *Proceedings of the National Academy of Sciences of the United States of America*, 97(4), pp.1867–1872.

Kornhuber, H.H. & Da Fonseca, J.S., 1964. Optovestibula integration in the cat's cortex: a study of sensory convergence on cortical neurons. , pp.239–279.

Kucewicz, M.T. et al., 2014. High frequency oscillations are associated with cognitive processing in human recognition memory. *Brain*, 137(8), pp.2231–2244.

Kuypers, H., Szwarcbart, M. & Mishkin, M., 1965. Occipitotemporal corticocortical connections in the rhesus monkey. *Experimental*.

Lachaux, J.P., Rudrauf, D. & Kahane, P., 2003. Intracranial EEG and human brain mapping. *Journal of Physiology-Paris*, 97(4), pp.613–628.

Lachaux, J.-P.P. et al., 2012. High-frequency neural activity and human cognition: Past, present and possible future of intracranial EEG research. *Progress in Neurobiology*, 98(3), pp.279–301.

Lakatos, P. et al., 2005. An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *Journal of Neurophysiology*

Lakatos, P., Chen, C.M., O'Connell, M.N., Mills, A., et al., 2007. Neuronal Oscillations and Multisensory Interaction in Primary Auditory Cortex. *Neuron*, 53(2), pp.279–292.

Laurienti, P.J. et al., 2005. On the use of superadditivity as a metric for characterizing multisensory integration in functional neuroimaging studies. *Experimental Brain Research*, 166(3–4),

pp.289–297.

Lemus, L. et al., 2010. Do sensory cortices process more than one sensory modality during perceptual judgments? *Neuron*, 67(2), pp.335–348.

Lennie, P., 1980. Parallel visual pathways: a review. *Vision research*.

Leuthardt, E.C. et al., 2004. A brain–computer interface using electrocorticographic signals in humans. *Journal of Neural Engineering*, 1(2), pp.63–71.

Lugo, E., Doti, R. & Faubert, J., 2008. Ubiquitous crossmodal Stochastic Resonance in humans: auditory noise facilitates tactile, visual and proprioceptive sensations. *PloS one*, 3(8), p.e2860.

Macaluso, E., 2000. Modulation of Human Visual Cortex by Crossmodal Spatial Attention. *Science*, 289(5482), pp.1206–1208.

Maier, A. et al., 2008. Divergence of fMRI and neural signals in V1 during perceptual suppression in the awake monkey. *Nature*.

Makeig, S. et al., 2002. Dynamic brain sources of visual evoked responses. *Science*, 295(5555), pp.690–694.

Makeig, S. et al., 2004. Mining event-related brain dynamics. *Trends in Cognitive Sciences*, 8(5), pp.204–210.

Manjarrez, E. et al., 2007. Effects of auditory noise on the psychophysical detection of visual signals: cross-modal stochastic resonance. *Neuroscience letters*, 415(3), pp.231–6.

Manning, J.R. et al., 2009. Broadband shifts in local field potential power spectra are correlated with single-neuron spiking in humans. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 29(43), pp.13613–20.

Maris, E. & Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. *Journal of neuroscience methods*, 164(1), pp.177–90.

Markov, N.T. et al., 2014. Anatomy of hierarchy: Feedforward and feedback pathways in macaque visual cortex. *Journal of Comparative Neurology*, 522(1), pp.225–259.

Markov, N.T. & Kennedy, H., 2013. The importance of being hierarchical. *Current Opinion in Neurobiology*, 23(2), pp.187–194.

Martuzzi, R. et al., 2007. Multisensory interactions within human primary cortices revealed by BOLD dynamics. *Cerebral cortex*, 17(7), pp.1672–9.

Massopust, L. & Barnes, H., 1965. Auditory frequency discrimination in cortically ablated monkeys. *Journal of Auditory*.

Maunsell, J. & Essen, D. van, 1983. The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *The Journal of neuroscience*.

McIlwain, J. & Buser, P., 1968. Receptive fields of single cells in the cat's superior colliculus. *Experimental brain research*.

Mercier, M.R. et al., 2013. Auditory-driven phase reset in visual cortex: Human

electrocorticography reveals mechanisms of early multisensory integration. *NeuroImage*, 79, pp.19–29.

Mercier, M.R. et al., 2015. Neuro-Oscillatory Phase Alignment Drives Speeded Multisensory Response Times: An Electro-Corticographic Investigation. *Journal of Neuroscience*, 35(22), pp.8546–8557.

Meredith, M.A. & Allman, B.L., 2009. Subthreshold multisensory processing in cat auditory cortex. *Neuroreport*, 28(20(2)), pp.126–131.

Meredith, M.A., Nemitz, J.W. & Stein, B.E., 1987. Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *The Journal of neuroscience*, 7(10), pp.3215–29.

Meredith, M., Nemitz, J. & Stein, B., 1987. Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *The Journal of neuroscience*.

Meredith, M. & Stein, B., 1996. Spatial determinants of multisensory integration in cat superior colliculus neurons. *Journal of Neurophysiology*.

Meredith, M. & Stein, B., 1986a. Spatial factors determine the activity of multisensory neurons in cat superior colliculus. *Brain research*.

Meredith, M. & Stein, B., 1986b. Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of neurophysiology*.

Mishra, J. et al., 2007. Early cross-modal interactions in auditory and visual cortex underlie a sound-induced visual illusion. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 27(15), pp.4120–31.

Molholm, S. et al., 2002. Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Brain research. Cognitive brain research*, 14(1), pp.115–28.

Morrell, F., 1972. Visual system's view of acoustic space. *Nature*, 238(5358), pp.44–6.

Murata, K., Cramer, H. & Bach-y-Rita, P., 1965. Neuonal convergence of noxious, acoustic and visual stimuli in the visual cortex of the cat. *J. Neurophysiol.*, 231(5978), p.736.

Murray, M.M. et al., 2004. Grabbing Your Ear: Rapid Auditory-Somatosensory Multisensory Interactions in Low-level Sensory Cortices Are Not Constrained by Stimulus Alignment. *Cerebral Cortex*, 15(7), pp.963–974.

Murray, M.M. et al., 2016. The multisensory function of the human primary visual cortex. *Neuropsychologia*, 83, pp.161–9.

Musacchia, G. & Schroeder, C.E., 2009. Neuronal mechanisms, response dynamics and perceptual functions of multisensory interactions in auditory cortex. *Hearing Research*, 258(1–2), pp.72–79.

Naue, N. et al., 2011. Auditory Event-Related Response in Visual Cortex Modulates Subsequent Visual Responses in Humans. *Journal of Neuroscience*, 31(21), pp.7729–7736.

Newman, E. & Hartline, P., 1981. Integration of visual and infrared information in bimodal neurons of the rattlesnake optic tectum. *Science (New York, NY)*.

O'Keefe, J., 1999. Do hippocampal pyramidal cells signal non-spatial as well as spatial information? *Hippocampus*.

Oh, S.W. et al., 2014. A mesoscale connectome of the mouse brain. *Nature*, 508(7495), pp.207–214.

Perrault, T. & Vaughan, J., 2005. Superior colliculus neurons use distinct operational modes in the integration of multisensory stimuli. *Journal of Neurophysiology*, 93, pp.2575–2586.

de Pesters, A. et al., 2016. Alpha power indexes task-related networks on large and small scales: A multimodal ECoG study in humans and a non-human primate. *NeuroImage*, 134, pp.122–131.

Populin, L.C. & Yin, T.C.T., 2002. Bimodal interactions in the superior colliculus of the behaving cat. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 22(7), pp.2826–2834.

Quinn, B.T. et al., 2014. Intracranial Cortical Responses during Visual-Tactile Integration in Humans. *Journal of Neuroscience*, 34(1), pp.171–181.

Raij, T. et al., 2010. Onset timing of cross-sensory activations and multisensory interactions in auditory and visual sensory cortices. *The European journal of neuroscience*, 31(10), pp.1772–82.

Rockland, K. & Pandya, D., 1979. Laminar origins and terminations of cortical connections of the occipital lobe in the rhesus monkey. *Brain research*.

Rockland, K.S. & Ojima, H., 2003. Multisensory convergence in calcarine visual areas in macaque monkey. *International Journal of Psychophysiology*, 50(1–2), pp.19–26.

Romei, V., Rihs, T., et al., 2008. Resting electroencephalogram alpha-power over posterior sites indexes baseline visual cortex excitability. *Neuroreport*, 19(2), pp.203–8.

Romei, V., Brodbeck, V., et al., 2008. Spontaneous fluctuations in posterior alpha-band EEG activity reflect variability in excitability of human visual areas. *Cerebral cortex (New York, N.Y. : 1991)*, 18(9), pp.2010–8.

Romei, V., Gross, J. & Thut, G., 2010. On the role of prestimulus alpha rhythms over occipito-parietal areas in visual input regulation: correlation or causation? *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 30(25), pp.8692–7.

Sadato, N. et al., 1996. Activation of the primary visual cortex by Braille reading in blind subjects. *Nature*.

Schroeder, C.E. et al., 2001. Somatosensory input to auditory association cortex in the macaque monkey. *Journal of neurophysiology*, 85(3), pp.1322–1327.

Schroeder, C.E. & Foxe, J.J., 2002. The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. *Cognitive Brain Research*, 14(1), pp.187–198.

Schroeder, C.E. & Foxe, J.J., 2002. The timing and laminar profile of converging inputs to

multisensory areas of the macaque neocortex. *Brain research. Cognitive brain research*, 14(1), pp.187–198.

Schroeder, C.E. & Lakatos, P., 2009. Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, 32(1), pp.9–18.

Senkowski, D. et al., 2011. Multisensory interactions in early evoked brain activity follow the principle of inverse effectiveness. *NeuroImage*, 56(4), pp.2200–8.

Senkowski, D. et al., 2007. Multisensory processing and oscillatory activity: Analyzing non-linear electrophysiological measures in humans and simians. *Experimental Brain Research*, 177(2), pp.184–195.

Shah, A.S. et al., 2004. Neural Dynamics and the Fundamental Mechanisms of Event-related Brain Potentials. *Cerebral Cortex*, 14(5), pp.476–483.

Shapley, R., 1990. Visual sensitivity and parallel retinocortical channels. *Annual review of psychology*.

Siegel, M. et al., 2008. Neuronal synchronization along the dorsal visual pathway reflects the focus of spatial attention. *Neuron*.

Siegel, M., Donner, T.H. & Engel, A.K., 2012. Spectral fingerprints of large-scale neuronal interactions. *Nature Reviews Neuroscience*, 13(February), pp.20–25.

Singer, W., 1999. Neuronal synchrony: a versatile code for the definition of relations? *Neuron*.

Spatz, W., Tigges, J. & Tigges, M., 1970. Subcortical projections, cortical associations, and some intrinsic interlaminar connections of the striate cortex in the squirrel monkey (Saimiri). *Journal of Comparative*.

Spinelli, D., 1968. Auditory specificity in uniy recordings from cat's visual cortex. *ExpNeurology*, 22, pp.75–84.

Stanford, T.R. & Stein, B.E., 2007. Superadditivity in multisensory integration: putting the computation in context. *Neuroreport*, 18(8), pp.787–92.

von Stein, A., Chiang, C. & König, P., 2000. Top-down processing mediated by interareal synchronization. *Proceedings of the National Academy of Sciences of the United States of America*, 97(26), pp.14748–53.

Stein, B. & Meredith, M., 1993. *The merging of the senses.* MIT Press, ed.,

Stein, B.E. et al., 2011. Challenges in Quantifying Multisensory Integration: Alternative Criteria, Models, and Inverse Effectiveness. *Experimental Brain Research*, 198(2–3), pp.113–126.

Stein, B.E. & Stanford, T.R., 2008. Multisensory integration: current issues from the perspective of the single neuron. *Nature Reviews Neuroscience*, 9(4), pp.255–266.

Stratton, G., 1896. Some preliminary experiments on vision without inversion of the retinal image. *Psychological Review*.

Tallon-Baudry, C. et al., 1996. Stimulus specificity of phase-locked and non-phase-locked 40 Hz

visual responses in human. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 16(13), pp.4240–9.

Tallon-Baudry & Bertrand, 1999. Oscillatory gamma activity in humans and its role in object representation. *Trends in cognitive sciences*, 3(4), pp.151–162.

Talsma, D., Doty, T.J. & Woldorff, M.G., 2006. Selective Attention and Audiovisual Integration: Is Attending to Both Modalities a Prerequisite for Early Integration? *Cerebral Cortex*, 17(3), pp.679–690.

Talsma, D. & Woldorff, M.G., 2005. Selective attention and multisensory integration: multiple phases of effects on the evoked brain activity. *Journal of cognitive neuroscience*, 17(7), pp.1098–114.

Teder-Sälejärvi, W.A. et al., 2002. An analysis of audio-visual crossmodal integration by means of event-related potential (ERP) recordings. *Brain research. Cognitive brain research*, 14(1), pp.106–14.

Thorne, J.D. et al., 2011. Cross-Modal Phase Reset Predicts Auditory Task Performance in Humans. *Journal of Neuroscience*, 31(10), pp.3853–3861.

Tudor, M., Tudor, L. & Tudor, K., 2004. Hans Berger (1873-1941)- the history of electroencephalography. *Acta medica Croatica: časopis*.

Vanni-Mercier, G. & Magnin, M., 1982. Single neuron activity related to natural vestibular stimulation in the cat's visual cortex. *Experimental Brain Research*, 45, pp.451–455.

Wang, X., 2010a. Neurophysiological and Computational Principle of Cortical Rhythms in Cognition. *Physiol Rev*, 90(3), pp.1195–1268.

Wang, X., 2010b. Rhythms in Cognition. *Physiological Reviews*, 90(3), pp.1195–1268.

Weeks, R. et al., 2000. A positron emission tomographic study of auditory localization in the congenitally blind. *The Journal of Neuroscinece*.

Weller, R. & Kaas, J., 1985. Cortical projections of the dorsolateral visual area in owl monkeys: the prestriate relay to inferior temporal cortex. *Journal of Comparative Neurology*.

Wilke, M. & Logothetis, N., 2006. Local field potential reflects perceptual suppression in monkey visual cortex. *Proceedings of the*.

Zampini, M., Shore, D.I. & Spence, C., 2003. Audiovisual temporal order judgments. *Experimental Brain Research*, 152(2), pp.198–210.

Zheng, C. & Colgin, L.L., 2015. Beta and gamma rhythms go with the flow. *Neuron*, 85(2), pp.236–237.

Zumer, J.M. et al., 2014. Occipital Alpha Activity during Stimulus Processing Gates the Information Flow to Object-Selective Cortex. *PLoS Biology*, 12(10).