



UNIVERSITY
OF TRENTO

CIMeC - CENTER FOR MIND/BRAIN SCIENCES

CiMeC
Center for Mind/Brain Sciences

Doctoral School in Cognitive and Brain Sciences
XXVII cycle

**Audiotactile interactions: psychophysical and
neuroimaging approaches**

Elisa Leonardelli

advisor
prof. Massimiliano Zampini

co-advisor
prof. Christoph Braun

27/04/2015

There are many people I'd like to acknowledge for their vicinity during these years.

Thanks to my family first of all. My parents, my siblings, my nephews and my brother in law.

I'd like to express my gratitude to my advisors Massimiliano Zampini, Christoph Braun and Nathan Weisz for their precious support all along the way.

Special thanks also to the MEG-group, because the completion of this thesis would not have been possible without their help, especially the two GPs.

Thank you very much to my friends, especially Chrysa and the Amedei. To Michela, Martino O., Silvia, Eleonora, Maria Paola, Lorenzo, Patrizia, Francesca, Raffaele, Tommaso, Mirco, Martina, Guido, Giuliano, Camilla, Marcello, Shuzzy, Charlie, Emiliano, Martino S., Andrea, Valentina.

Also thanks to my wider family: to the 5 aunts, 7 uncles and 7 crazy cousins.

Contents

ABSTRACT	5
CHAPTER 1 General introduction	6
1.1 A multisensory world	7
1.2 Interactions between the senses of audition and touch	9
1.3 Behavioral evidence of AT multisensory interactions	10
1.4 Multisensory looming signals	13
1.4.1 Looming (and receding) signals	13
1.4.2 Multisensory looming signals	14
1.5 Neural correlates of multisensory interactions: the hierarchical model and new findings	16
1.6 AT multisensory association areas	16
1.7 Multisensory AT interactions in animals in primary or close to primary cortical areas	17
1.8 Neural correlates of AT interactions in humans	19
1.9 Methodological considerations	22
1.9.1 Definition of the principles of multisensory integration from cat's superior culliculus (SC)	22
1.9.2 The linear additive model	24
CHAPTER 2 Short description of the project	27
2.1 Global overview of the project	28
2.2 Study I	29
2.2.1 Introduction	29
2.2.2 Results	32
2.2.3 Discussion	35
2.3 Study II	37
2.3.1 Introduction	37

Table of content

2.3.2 Results	39
2.3.3 Discussion	42
2.4 General conclusions	46
2.5 Final considerations and future directions	49
CHAPTER 3 References	51
CHAPTER 4 Papers and contributions	62
4.1. Contributions	62
4.2 Study I	64
Auditory stimuli characterized by different saliencies modulate differently the neural processing of tactile stimuli: evidence from MEG recordings	64
4.3 Study II	89
Prestimulus oscillatory alpha power and connectivity patterns predispose perceptual integration of an audio and a tactile stimulus	89
APPENDIX	112
Introduction	113
Methods	115
Results	120
Discussion	129
Additional references used in the appendix	Error! Bookmark not defined.

ABSTRACT

In daily life, we are immersed in a continuous flow of stimuli targeting each of our different senses. Far from being independently processed, accumulating evidence has been widely documented by studies showing that stimuli from different modalities largely interact. However, despite the increasing interest, the interpretations of the results of experiments studying multisensory interaction are still controversial and the underlying mechanisms remain broadly unknown.

The aim of this thesis is to investigate the interactions that occur between the senses of audition and touch. Audiotactile interactions have been far less studied than the ones existing between other modality pairings. Maybe because they go often unnoticed though being well present in many everyday life situations. This thesis focuses mainly on two aspects that concern interactions: understanding the impact of the relative saliency between the stimuli and investigating the mechanism behind perceptual integration. These questions are addressed respectively in two studies conducted by means of magnetoencephalography.

The thesis is structured as following: in chapter 1, I provide the theoretical background to my scientific questions. A brief synthesis of the two main studies is presented in chapter 2. The two studies are entirely reported under the form of manuscripts in chapter 4.

Finally, in appendix a behavioral study that investigates spatial aspects of AT interactions is reported. Although the results of this study are of pertinence of the project, given the preparatory character and the preliminary state of the study we decided to show them in the appendix rather than include them in the main body of the thesis.

CHAPTER 1

General introduction

The scope of this chapter is to give an overview of the state of art of audiotactile (AT) multisensory research. It has been the basis of my scientific questions and motivated the experimental choices of the two magnetoencephalography (MEG) studies presented in chapter 2 and of the behavioral study reported in the appendix.

In particular, this introduction is about the field of multisensory research (s. 1.1), the reasons for studying AT interactions (s. 1.2), the existent behavioral results in humans of AT interactions (s. 1.3), and evidences suggesting a particular role of peripersonal space and of looming signals (s. 1.4) for multisensory interactions. Another part of this chapter is dedicated to the presentation of studies that have contributed in delving into the neural substrates subserving audiotactile interactions in both animals and humans (s. 1.5, 1.6, 1.7, 1.8). Finally, I discuss some concerns related to methodological issues that need to be taken into consideration when dealing with the studying of multiple modalities at once (s. 1.9).

1.1 A multisensory world

In daily life, we receive a continuous and simultaneous flow of information from each of our different senses. This information is merged in the brain, resulting in a seamless perception of the world. For many years research tended to focus on each of our senses in isolation. However, it has now been amply documented how stimuli from different modalities largely interact. Despite the increasing interest in interactions between senses, there is a long standing debate concerning the understanding of the underlying mechanism of multisensory interaction.

In the past the limiting factors that caused crossmodal interactions to be poorly, or almost not, investigated have been the assumption that such a complex structure as the brain and the functional architecture of the mind can be described in term of modules. In the 1980's, Jerry Fodor provided an extensive discussion on the idea of the modularity of the mind (Fodor, 1983). According to Fodor, modules are cognitive domains that only respond to certain kind of stimuli (i.e., they are specialized), are neurally specific, innately specified, fast and process information in a mandatory manner. In addition, they need not refer to other cognitive systems in order to operate (i.e., 'informationally-encapsulated' in Fodor's terminology) and therefore function in an essentially autonomous fashion. From this perspective, early sensory areas have typically been conceived as the quintessential modules. They were thought to deal with information from a single sensory modality at a time, within neurally specific and innately specified structures, and in a highly autonomous and 'informationally encapsulated' fashion. Not only, it may also partially reflect the fact that some of our perceptual experiences might appear unisensory. For example, when people try to understand what another person is saying they typically think that all the information they get comes from what they hear. However, when we listen to someone speaking in a noise environment (such

as at one in cocktail party), our perception can be facilitated if we not only hear the person's voice together with those of many others, but can also see their face- and lip-movements (e.g. Sumbly and Pollack, 1954). Although this influence is outside the observer's awareness, it becomes apparent when auditory and visual information are somewhat incongruent. For instance, whenever we go to the cinema we perceive the voices of the actors on the screen as emanating from their lips, despite the fact that the sounds are actually presented from loudspeakers situated elsewhere (Spence and Driver, 2000). This illusion known as the 'ventriloquism effect', demonstrates that people often mislocalize sounds toward their apparent visual source (e.g., Bertelson and Schersleben, 1998). Moreover, another paradigmatic case of crossmodal interaction is the well-known McGurk's effect (McGurk and MacDonald, 1976), where the presentation of an auditory syllable, paired with presentation of a visual different syllable, leads to the modified auditory perception of a third syllable.

In all of the examples of sensory conflicts reported so far, vision dominate when the information where somewhat incongruent. However, this is not always the case. Sometimes what we hear can also change what we see and what we feel. For instance, Shams and colleagues (2000) have shown that a beeping sound can lead to the illusory flashing of a light. People in their study fixate on a spot on a computer monitor and were asked to judge how frequently a white disk was flashed on the periphery of the screen. When presented in silence participants were extremely accurate in reporting the number of flashes, however, when a single flash of light was accompanied with multiple auditory beeps, participants were fooled into believing that the light had flashed several times. This is a clear demonstration that we hear can lead to a qualitative change in what we see.

Auditory cues can also influence (or dominate) tactile perception in the absence of visual cues as shown by the 'parchment-skin' illusion (Jousmäki and Hari, 1998). In this experiment, people were asked to rate the perceived roughness

/dryness of their own hands while rubbing them together. The hand rubbing-sounds were picked up by a microphone placed close to the hands, and fed back over headphones. This audio feedback could either be identical to the original sound or else manipulated to boost or cut the high-frequencies (i.e., above 2 kHz). People reported that their hands felt smother/dryer when the high-frequency sounds were amplified, but cutting (or reducing) sounds in this range led to the hands feeling rougher/smoothed. These examples clearly demonstrate how limiting it is to focus on one sensory modality in isolation in order to deeply understand how we perceive the external world and the need to discover the general rules of how senses interact.

1.2 Interactions between the senses of audition and touch

A specific focus of this thesis is the interaction between the auditory and tactile senses. AT-interactions have been far less studied than interactions between any other pairs of modality. While a rough search in the database PubMed using the keyword 'audiotactile' yields 49 results, a search with the keyword 'audiovisual' produces over 8000 results (visuotactile ranks 108). Nevertheless, AT interactions deserve their interest because they can link a remote sense (audition) to a bodily sense (touch). Moreover, though they go often unnoticed, they are present in everyday life. For example, when touching an object, we perceive its consistency through the touch but also the sound produced by the touch is informative about the object's properties, when leafing through a newspaper, or when typing on a keyboard or the buzz of a mosquito and its bite, knocking at a door, playing instruments, walking, shaving etc. Furthermore, a number of similarities exist between the senses of hearing and touch that may suggest the existence of a favored link between both modalities. In the fifties of the past century, von Békésy highlighted a number of similarities between the senses of audition and touch, suggesting to adopt the sense of touch as supporting model for studying audition

(e.g., von Békésy, 1955, 1957, 1959). For instance, he noticed that mechanisms for transducing sensory signals are analogous for sounds and vibrotactile stimuli. In particular, the auditory and tactile receptors are activated by the same type of physical energy, namely the mechanical pressure of the travelling waves (von Békésy, 1959; cf. Nicolson, 2005). Furthermore, there are analogies in the physiological mechanisms (see Corey, 2003; Gillespie & Müller, 2009) supporting the idea that the two sensory modalities have common origins (see Soto-Faraco & Deco, 2009; von Békésy, 1959, for reviews). From this point of view, prenatal studies of ontogenetic brain development in humans show that the structures of the inner ear gradually evolved from skin tissue (Driver and Kelley, 2009; Vickaryous and Hall, 2006). It is also interesting to note that the developments of neuros in the the anterior ectosylvian follows a precise order, first, the tactile neurons, second the auditory neurons and finally the visual neurons (Wallace et al., 2006). Therefore, the closer development of the auditory and tactile systems could affect the subsequent multisensory interactions having some effect upon the successive strength, direction, and amount of reciprocal connections between them (e.g., Gregory, 1967; Katuski, 1965). Moreover, the fact that auditory and tactile signal transduction times from the ear and skin are faster and comparable in time, as opposed to visual signals with much longer transduction times (Barnett-Cowan and Harris, 2009). Finally, studying AT interactions is interesting not only in and of itself, but can provide insights into the broad understanding of how different senses interact, regardless of their modalities.

1.3 Behavioral evidence of AT multisensory interactions

Perhaps the most striking evidence of interaction between the senses of audition and touch is revealed in the Parchment-Skin Illusion (Jousmäki and Hari, 1998) already mentioned in this chapter (s. 1.1). Beside this illusion, several

behavioral studies reported AT interactions and how they are sensitive to different factors like the space around the body where the stimuli are presented, the distance of the audio stimuli from the body, the complexity of the sound, and the spatial arrangement of the stimuli.

In Farnè and Làdavas (2002), patients suffering from tactile extinction failed to report most of the tactile stimuli presented on the contralesional (peri-head) side when an auditory stimulus was presented simultaneously on the ipsilesional side, even though they could report contralesional tactile stimuli nearly perfectly, when they were presented in isolation. The magnitude of this extinction effect was greater when the auditory stimuli were presented from close to the participant's head (20 cm) than when they were presented far from the head (70 cm away) for complex stimuli (white noise bursts) rather than pure tones (see also Làdavas et al., 2001, for similar results)

Despite these findings, research on healthy participants suggested a weaker spatial link for interactions between signals presented in the auditory and somatosensory modalities than those that exist between other pairs of modalities. For instance, people's performance in crossmodal Temporal Order Judgment (TOJ) tasks is typically worse when the two stimuli in different modalities are presented from the same location than when they are presented from different locations, perhaps because crossmodal integration (i.e., temporal ventriloquism; Morein-Zamir, Soto-Faraco, & Kingstone, 2003) is enhanced for stimuli presented from the same direction. This spatial modulatory effect on crossmodal TOJs has been reported between auditory and visual stimuli (Zampini et al., 2003a, 2003b), and between visual and tactile stimuli (Spence et al., 2003). However, somewhat surprisingly, Zampini et al. (2005) recently found no such spatial modulatory effects on audiotactile TOJs across three experiments. Moreover, in an EEG study, Murray et al. (2005) observed no difference in RTs facilitation for AT presented in the peri-hand space, when the stimuli were presented left/right aligned or

misaligned (interaction's neural correlates were as well indistinguishable between spatially aligned and misaligned and originated in auditory area, leading the authors to suggest the existence of large bilateral auditory RF in this region). Similar behavioral results were reported by Zampini et al. (2007) who presented AT stimuli from separated spatial location (front/back) and observed a facilitation of RTs irrespective of whether the stimuli were in spatial register or not (see also Lloyd et al., 2003; Zampini et al., 2005 similar results).

But in all the AT studies reported above, the stimuli were presented to participants' hands. Successive studies suggests that AT interactions in humans depends to a large extent on the particular region of space in which the auditory and tactile stimuli happen to occur and that, among the various regions of space surrounding the different parts of the body, stimuli delivered to the space around the head (peri-head space) have a higher degree of salience. This fits well with the consideration of the high importance of head for our survivor and the fact that vision cannot monitor its surface. Kitagawa et al. (2005) investigated AT temporal order judgments for the first time in healthy humans in the region behind the head. They showed how, accordingly to the results found between other modality pairings, in this portion of space performance was worse when auditory distractors were presented on the same side to the tactile stimuli than when they were presented on the opposite side. This crossmodal distractor interference effect was more pronounced when white noise distractors (vs pure tones) were presented from close to the head (20 cm) than when they were presented far from the head (70 cm). Tajadura-Jimenez and colleagues (2009) applied electrocutaneous stimuli to either earlobe while auditory stimuli were presented from the congruent or incongruent side and at two distances, close (20 cm) or far (70 cm) from the participant's head. They showed a spatial modulation of AT interaction for stimuli presented close to the head, with faster RTs for congruent with respect to incongruent stimuli. Interestingly, when the same stimuli were delivered to the hands, which were

placed either close to or far from the head, they could not prove any spatial modulation of AT interactions, showing a specificity of the particular body part stimulated rather than to the region of space (i.e. around the head) where the stimuli were presented.

In literature, it has been distinguished between different partitions of the space as a function of their distance to the body. The space directly surrounding the body is called peripersonal space (PPS) and possesses a particular relevance, since it is the space where physical interactions of a subject with the external world takes place (Rizzolatti, 1997). Evidence from studies in humans mentioned in above in this section (Kitagawa et al., 2005; Tajadura-Jiménez et al., 2009), together with similar results from other modality pairings (e.g. Làdavias, 2002) and from in neurophysiological studies in monkeys (see section 1.6 or Graziano et al., 2004 for a review) suggested the existence of a specialized brain system that specifically represents the PPS in a multisensory way (e.g. Holmes and Spence, 2004).

1.4 Multisensory looming signals

Among the multiple worries a living being has to daily deal with is the protection from external physical threats (e.g., Dosey and Murray 1969; Cavallin and Houston 1980 see Cooper and Vitt 2002; Eilam 2005 for evidence in animals). One of the possible functions suggested for the special relevance of multisensory in PPS representation is thought to be the better monitoring of the space around the body, and the better encoding of approaching stimuli that could be potentially harmful in order to prepare defensive behaviors.

1.4.1 Looming (and receding) signals

A considerable amount of evidence shows that the increase in sound intensity or in the size of a visual shape can reliably be interpreted as an

approaching object and are therefore increasingly used to simulate looming objects in experimental settings (e.g. Ghazanfar et al. 2002; Seifritz et al. 2002; Riskind et al. 2014). Approaching or looming signals are often related to extremely relevant environmental events, such as collisions or threats. Thus, the activation of an appropriate defensive behavior, such as attention orienting or motor acts (Bach et al., 2009; Grassi, 2010), in response to a looming object potentially threatening the body, is of vital importance for the maintenance of the integrity of a living being. Along this line, studies have shown that humans tend to overestimate the change in loudness of sounds with increasing level compared to those decreasing by an equivalent amount (Neuhoff, 1998) and, conversely, to underestimate the distance or time-to-arrival of looming sounds (Neuhoff, 2001). Similarly, the detection of looming sounds is prioritized as compared to static or receding sounds, as reflected by shorter reaction times in the former than in the latter conditions (Bach et al., 2009; Cappe et al., 2009). Looming sounds are also rated as more unpleasant, arousing, intense, salient and threatening and elicit increased skin conductance responses than receding sounds (Bach et al., 2009; Tajadura-jiménez et al., 2010).

1.4.2 Multisensory looming signals

Given the special relevance of prompt detection of looming signal (see previous section) given the fact that we live in a multisensory environment rich of information that comes from different sensory modalities and the presence of stimuli from different modalities has been largely shown to provide behavioral benefits, one could reasonably expect to obtain larger multisensory effects from the presence of looming than static signals. Indeed it has been shown in monkeys (e.g., Graziano et al., 1997) how multisensory neurons respond better to approaching visual stimuli and that, when electrically stimulated, these neurons evoke defensive movements (Cooke et al., 2003). Accordingly, it has been found that the

multisensory neurons in PZ and VIP are highly responsive to objects rapidly approaching the body (Graziano et al., 2002). When these sites are electrically stimulated, a specific pattern of defensive behaviors is evoked, such as facial grimace, head turn, hand lifting in the proximal space. The same pattern of behaviorally observable reactions is nevertheless triggered by presenting airpuffs close to the animal's face (Cooke et al., 2003). By virtue of their functional feature, these neurons have been considered as the neural substrate of the capability to detect objects potentially threatening and noxious for the body and to implement appropriate motor responses (Graziano and Cooke, 2006). In the audiovisual domain, looming (vs. receding or static) signals have been shown to induce preferential orientation responses in 3-month-old infants (Lewkowicz, 2008), human adults (Leo et al., 2011) as well as in monkeys (Maier et al., 2004) and shorter reaction times (Cappe et al., 2009). Although still tentative, studies on the neural underpinnings of the audiovisual looming effects show that audiovisual looming (vs. receding) signals determine an enhanced BOLD response in primary visual and auditory areas as well as in multisensory areas (superior temporal sulcus STS; parietal and frontal structures; Tyll et al., 2013; see also Hall and Moore 2003 for a review). Moreover, the processing of looming sounds is associated with an increased connectivity of bilateral superior temporal sulcus with visual areas (Tyll et al. 2013), possibly subserving a modulatory effect of looming sounds on the excitability of visual cortex (Romei et al., 2009).

With regard to the interactions between auditory looming signals and tactile stimuli, in a study conducted by Serino's group (Canzoneri et al., 2012) a tactile stimulus was administered at the hand at different delays from the onset of a looming or of a receding sound. Results showed that the moving auditory stimulus speeded up the processing of the tactile stimulus as long as the sound was perceived at a limited distance from the hand for both sounds. Moreover, when comparing looming vs receding sounds, the approaching one seems to have a

stronger effect in speeding up the processing of a tactile stimulus.

1.5 Neural correlates of multisensory interactions: the hierarchical model and new findings

At a cortical level, traditionally the processing of different senses has been believed to occur in a first step separately for each sensory modality and to converge successively, through a feedforward pathway, in a strictly hierarchical manner from primary to secondary sensory specific cortices and to regions of ‘association’ or ‘heteromodal’ cortex (hierarchical model). The classical multisensory association regions for AT were identified by a number of studies in the frontal lobe, as parts of the premotor cortex, in the parietal lobe, particularly in the intraparietal sulcus (IPS), and in the temporal lobes, particularly the STS.

In the last 15 years a multitude of studies have shown multisensory interactions already at a very early latency of cortical process and in primary or close to primary cortical areas. These results challenge the assumption that multisensory integration takes place only in high-order association cortices. Though several attempts have been made (see Ghazanfar and Schroeder, 2006 for a review), a clear understanding of the mechanisms of interactions is still lacking.

1.6 AT multisensory association areas

In monkeys, the premotor cortex contains multisensory neurons responding to visual, auditory, and somatosensory inputs presented in proximity of the face, arm and upper torso (Graziano and Gandhi, 2000; Graziano and Gross, 1998; Graziano et al., 1999). For the most part, multisensory neurons are located in a ‘polysensory area’ in the dorsal part of premotor cortex. The function of these neurons appear to be involved with the production of defensive behavior: they are particularly sensitive to looming sounds or looming visual stimuli (Ghazanfar et al.,

2002), i.e. to stimuli that indicate an approaching potential danger. This view is supported by defensive movements that are elicited by microstimulation of the 'polysensory zone' (Graziano et al., 2002).

The ventral intraparietal region (VIP), located in the fundus of the intraparietal sulcus, has been shown to respond to visual, somatosensory (Duhamel et al., 1998), and auditory (Schlack et al., 2005) stimuli, especially moving stimuli, and also to vestibular stimuli (Bremmer et al., 2002).

In macaque monkeys, along the fundus of the posterior superior temporal sulcus (STS), in a region, also labeled 'superior temporal polysensory' (STP) area, neurons have been detected that are responding to auditory, visual and somatosensory stimuli (Bruce et al., 1981). A multitude of neuroimaging studies in humans reported responses to auditory and visual stimulation in STS (Beauchamp et al., 2004; Calvert et al., 2001; Noesselt et al., 2007)(Beauchamp et al., 2004; Calvert et al., 2001; Noesselt et al., 2007) and also to somatosensory stimulation (Beauchamp et al., 2008).

1.7 Multisensory AT interactions in animals in primary or close to primary cortical areas

The caudomedial auditory area (CM) is an area adjacent to primary auditory cortex (AI), which is thought to participate in early processing of sounds, especially complex noises, as a first stage (or "belt") auditory association cortex. This area has been shown to be a site of multisensory convergence, since it shows robust somatosensory and auditory responses, at early stages of cortical processing. The first report of non-auditory inputs in CM is a study in 2001 by Schroeder and colleagues (Schroeder et al., 2001). By using multi-contact electrodes implanted in monkeys they reported somatosensory input in CM. Awake monkeys were presented with binaural clicks, pure tones and band-passed noise via headphones,

and somatosensory stimulation was delivered by means of contralateral median nerve stimulation. The timing and the laminar profile of activity in the caudomedial (CM) belt auditory area were similar for somatosensory and auditory inputs, and in both cases responses expressed a feedforward profile. (i.e., the initial responses began in and near lamina 4 and spread to extragranular layers). On the contrary, no somatosensory activation was registered in AI.

Fu and colleagues (Fu et al., 2003), by means of multi-neuron cluster recordings, confirmed that the majority of recording sites that responded to auditory stimulation in CM, but not in AI, responded also to somatosensory stimulation. They further investigated which body parts and which somatosensory modalities best activate the CM belt area and found an activation preference in response to hand and head surfaces and for cutaneous stimulation.

Kayser and colleagues (Kayser et al., 2005) showed by means of functional Magnetic Resonance Imaging (fMRI) technique in anaesthetized monkeys a supra-additive integration effect in response to a touch of the hand or foot, when presented simultaneously to a sound. This effect was localized in the CM, but also in the caudolateral (CL) belt areas of the auditory cortex. Integration was stronger for temporally coincident stimuli and obeyed the principle of inverse effectiveness, i.e. greater enhancement for less effective stimuli.

Cappe and Barone (2005) injected retrograde tracers in unimodal auditory, somatosensory and visual cortical areas of marmoset and found several direct connections between cortical areas involved in processing information of different modalities. Among others, they found projections from secondary somatosensory cortex (SII) to the auditory cortex and the anterior bank of the lateral sulcus. Furthermore, neurons were detected that respond to auditory stimuli at a short latency in a somatosensory region including SII. These findings suggest that this area, together with the posterior auditory areas, is involved in the integration of auditory and somatosensory stimuli. However, there is a lack in physiological

studies investigating activation or convergence in somatosensory cortex for visual and/or auditory responses (but see Zhou and Fuster, 2004).

Lakatos and colleagues (Lakatos et al., 2007) aimed to further investigate neuronal mechanisms and functional significance of low-level multisensory interaction in monkeys in primary auditory cortex (AI). They penetrated primary auditory cortical area of macaque monkeys with multielectrodes and recorded current source densities and multi-unit activity. They stimulated the animals with sound clicks and via median nerve electrical pulses first separately and found that somatosensory stimuli have a modulatory effect in AI, rather than increasing the neural firing rate. By comparing the sum of singular responses to those resulting from simultaneous AT stimulation using different intensities of the sound, they found AT interaction effects in AI. These effects were super-additive for weak auditory stimuli (below 50 dB) and additive for stronger auditory stimuli (from 50 dB to 80 dB), nicely in accordance with the principle of inverse effectiveness. Furthermore, by temporally shifting the occurrence of the auditory stimulus (with sound of fixed intensity) with respect to the onset of the tactile stimulus, they showed how the auditory-somatosensory interaction effect does not fall linearly or exponentially with the stimulus onset asynchronies (SOAs), but rather it follows a non-linearity cycle that coincides with the frequency-periods of the classical spontaneous oscillatory activity in AI (for this see Lakatos et al., 2005) detected in brain (i.e. gamma, theta and delta band). These results suggest that a somatosensory stimulus, rather than inducing an increase in neuronal firing rates, induces a reset of local oscillatory activity to an optimal excitability phase. This mechanism would ensure that the crucial input arrives at the moment of maximum excitability in the oscillatory cycle present in AI and enhances the response to an auditory stimulus.

1.8 Neural correlates of AT interactions in humans

Since some decades, the question of whether interactions between auditory and somatosensory can be detected at neuronal level in humans using non-invasive imaging techniques such as electroencephalography (EEG), MEG or fMRI has been raised.

Lam and colleagues (1998) reported that the activity in a region that was considered to reflect ipsilateral SII was suppressed by continuous auditory (music) stimulation.

In a high-density EEG study, Foxe and colleagues (Foxe et al., 2000) provided for the first time topographical maps of the AT interaction potential. They showed AT interactions already in early stages of cortical processing. The earliest significant AT interaction peaked at ~65 ms, in the hand representation area of the postcentral gyrus. A later interaction at ~80 ms was consistent with a contribution from activity within the posterior auditory cortices. The responses to AT stimuli were stronger than the sum of responses to auditory and tactile stimuli presented alone.

In a complementary fMRI study, Foxe and colleagues (2002) showed convergence of somatosensory and auditory inputs to a sub-region of the human auditory cortex along the superior temporal gyrus, a possible human homologue of the macaque monkey CM belt area. Moreover, they showed how simultaneous stimulation in both, auditory and somatosensory modalities, resulted in activity exceeding that predicted by summing the responses to unisensory stimuli, interpreted as facilitatory effect of AT interaction.

In a MEG study, (Lütkenhöner et al., 2002), subjects were stimulated with binaural tones and unilateral tactile pressure pulses to the right thumb, either separately or simultaneously. MEG recordings indicated AT interaction in the hemisphere contralateral to the tactile stimuli, in six of eight subjects, with the major effects at ~ 140 and ~220 ms. The responses to AT stimuli were weaker than the sum of responses to auditory and tactile stimuli presented alone and were

identified as reflecting partial inhibition of SII cortex.

Gobbelé and colleagues (2003) attempted to extend the study of Lütkenhöner by applying spatially and temporally coincident stimuli, so that the stimuli were also perceptually bound together. They recorded MEG responses to unilateral auditory or tactile stimuli, or spatiotemporally coincident bimodal AT stimuli, which could be presented either from the right or left side of the body midline. They identified suppressive AT interaction at about 75–85 ms, in the contralateral posterior parietal cortex, and at about 105–130 ms in the contralateral temporoparietal areas, between SII and auditory cortices.

In an EEG experiment, Murray and colleagues (2005) investigated the spatial alignment as the critical parameter for the temporally earliest multisensory interactions observed in lower-level sensory cortices. They recorded psychophysical and electrophysiological data, where participants performed a reaction time task in response to unisensory stimulus or to multisensory AT stimuli arranged in different spatial configurations: aligned (e.g. Left hand touch/ left sided sound), or spatially misaligned stimulation (e.g. left hand touch/ right-sided sound). The psychophysical data showed that participants responded more rapidly to multisensory than unisensory stimuli, irrespectively of whether the stimuli were presented spatially aligned or misaligned. Similarly, the EEG data showed (facilitatory) multisensory interactions at just 50-95 ms for both spatially aligned and misaligned stimuli. The auditory-somatosensory interaction was localized in auditory association areas contralateral to the side of somatosensory stimulation.

In Sperdin et al. (2009) in order to link early, low-level multisensory interactions and behavioral indices of multisensory processing, a different analysis was performed on the eight conditions of the dataset collected in the study of Murray et al. (2005) cited right above. In order to evaluate whether early AT interactions impact subsequent behavior, trials were sorted according to slow or fast reaction times for each condition. Results provide evidence that AT interactions

relate to the later response speed. However, no evidence was found of distinct networks engaged for fast and slow reaction times, but rather a modulation in their strength.

1.9 Methodological considerations

It is important to note how the different approaches mentioned in the previous sections of this chapter, utilized different methods to investigate and to define multisensory interactions. For example, in the case of anatomical tracer studies, typically connections between sensory-specific areas are investigated. In the case of single-cell studies, it is typically assessed how single neurons respond to more than one sensory modality when stimuli are presented in isolation or combined. Since it is quite widely debated, how multisensory integration is explored by different approaches the following paragraph aims to highlight the principles and methods (and the consequent strengths, precautions and limits) inherent in the study of multisensory interaction in humans and, thus of relevance for two of the three studies presented in this thesis.

1.9.1 Definition of the principles of multisensory integration from cat's superior colliculus (SC)

The milestone in multisensory research field is the early work conducted by Stein and Meredith (1993) on cat's superior colliculus (SC), later complemented by studies in the SC of macaque monkey (Wallace and Stein, 2001). The SC is a midbrain structure that controls orientation behavior (for instance gaze shifts), and coordination of movements. SC is characterized by the presence of many multisensory neurons, i.e. neurons responding to stimuli of more than a single modality. Operationally, multisensory integration has been defined as the presence

of a number of neuronal impulses significantly higher (superadditive) when evoked by a crossmodal combination of stimuli than when evoked by the most effective of these stimuli individually. Thus, according to this definition and based on single cell recordings in SC, the three basic principles of multisensory were formulated. The first principle is the ‘spatial rule’ that states that multisensory integration is more likely (or largest) when the unisensory stimuli to be integrated occur at the same location in space. Indeed, in SC, multisensory neurons possess multiple receptive fields (RFs), one for each modality. These RFs overlap in space, and, if two stimuli are in the same spatial register, the brain activity is enhanced. On the contrary, if two stimuli are not spatially aligned no enhancement is present, or even a depression. The second principle, the ‘temporal rule’, states that multisensory integration is more likely (or largest) when the unisensory stimuli to be integrated are presented at approximately the same time. In general, it has been hypothesized that there exists a possible window of time (in the order of few up to several hundreds of milliseconds) (Spence and Squire, 2003), during which integration can take place. This enables also to take into account different modalities speeds (e.g. light speed and sound speed), different response latencies and different mechanical transduction speeds. The third principle, the ‘rule of inverse effectiveness’, states that multisensory integration is more likely (or largest) if the uni-sensory stimuli are rather weak. This is also intuitively logic: while highly salient cues are easily detected and their combination has a modest effect on their detection and localization, combination of weak cues evokes a substantially enhanced response when compared to the individual weak stimuli.

This operational definition of multisensory integration and the three principles empirically derived from SC recordings, has provided a simple conceptual framework for interpreting multisensory findings for neuroimaging studies in humans.

1.9.2 The linear additive model

AT integration in human neuroimaging studies is usually defined by means of the additive linear model. This model is based and inspired on the studies on SC reviewed in previous section. According to the linear model, the data recorded during the simultaneous bimodal stimulation (AT) is contrasted with the sum of the data obtained from its unisensory components presented singularly (Audio A, tactile T). The rationale is that if the two dimensions of the bimodal stimulus are processed independently, the neural activities obtained by the algebraic sum of the responses to the unimodal constituent of the bimodal stimulus to the response should be equal to the activity produced by the bimodal stimulus itself. Any activity exceeding (superadditive effect) or falling below (depression) the sum is regarded as interaction. In other words, AT interactions = response to (AT) – (response to (A) + response to (T)).

This model has been widely used in human studies of crossmodal interactions but unluckily it has two main limitations. The first limitation is that this estimation is sensitive to several potential biases, and experiments have to be designed carefully in order to avoid or minimize them (see Besle et al., 2004). It assumes that the analyzed brain responses do not include any activity common to all three stimulation (AT, A, T). If this assumption is not true, this common activity would be added once and subtracted twice, leading to a wrong estimation of the multisensory integration effect. These common activities could be task-related neural activities associated with late cognitive processes, response selection, or motor processes. However, these activities arise usually after about 200 ms post-stimulus, while previous latencies are characterized by sense-specific answers (see Hillyard et al., 1998). Thus, with EEG and MEG techniques usually the analysis is limited to a time range below 200 ms, while with fMRI this is not possible, rendering this technique unsuitable for an ‘additive model’ based approach.

However, also in EEG/MEG preparatory states, i.e. slow potentials/fields that begin prior to stimulus onset, can introduce a further potential artifact (Teder-Sälejärvi et al., 2002). Another assumption of this approach is that, during unimodal stimulation, the non-stimulated cortex is not affected by the stimulation in the other modality. However, this is not entirely the case since several studies have reported deactivations in sensory-specific cortex when subjects were presented with continuous stimulation in another modality (e.g, Laurienti et al., 2002). Furthermore, it assumes that the attentional load required in the two unimodal and in the unimodal conditions is the same. This is obtained usually by requiring the same task in all three conditions to the subjects. However, it may be the case that the task may require less attentional effort in one condition than in the other.

Moreover, this model is blind to linear interactions between senses. In fact, as an implicit consequence of its design it can detect only super- or sub-additive interactions. On the contrary, senses can interact linearly. For example, if they converge (and are integrated) on the same pool of neurons, but they responses still approximate the sum of the responses to its modality-specific components. Due to poorness of spatial localization of imaging techniques, this interaction cannot be distinguished from activation of separating pools of unisensory neurons. Since studies using noninvasive methods like fMRI and ERP have become increasingly important tools for the study of brain function and are constrained to focus specifically on response non-linearities, non-linearities (in particular superadditivity) have become the main signature for the presence of multisensory integration and multisensory integration has become synonymous with response nonlinearity. However, in this regard it is important to mention more recent findings on SC in cats (Stanford et al., 2005; see also Perrault et al., 2005; Stanford and Stein, 2007) that emphasize the incidence of linear interactions with respect to the non-linear ones. In these studies in fact, the most multisensory interactions were approximated by simple summation of modality-specific input, while

superadditive interactions were common only in responses to very ineffective modality-specific stimuli (while subadditive interactions were relatively rare).

CHAPTER 2

Short description of the project

In this chapter a brief description of the entire research project is given. At first an overview of the project is provided (s. 2.1). Afterwards, the two MEG studies are synthetically introduced, described and discussed (Study I, s. 2.2; Study II, s. 2.3). Finally, I outline some general conclusions and future perspective (s. 2.4).

For a more detailed description of the studies see chapter 4 where the manuscripts of the two studies are reported. To note that a third behavioral study has not been included in the main body of the thesis but in the appendix. In fact, although results of this study are of pertinence of the project, it is at a preliminary stage.

2.1 Global overview of the project

The overall aim of this project was to expand the current knowledge in multisensory interactions, in particular between the senses of audition and touch. Although AT interactions have been poorly investigated (see section 1.2) so far, they may significantly increase our knowledge of the multisensory field. As a conclusion of the literature review previously presented, three main questions were issued.

In the first MEG experiment, the role of the saliency of the stimuli over audiotactile interactions was investigated. This was achieved using three sounds that, although identical in their frequency content, were modulated in amplitude obtaining sounds characterized by three different perceived saliences. The goal was to study how the different saliencies of the stimuli can modulate the neural processing of a tactile stimulus that was always the same throughout the entire experiment.

The second MEG study investigated how perceptual integration occurs. Perceptual integration refers to the process by which stimuli from different modalities are perceived as emanating from the same event (or vice versa as two separate events; i.e., perceptual segregation). In particular, in this study the temporal relationship between an auditory and a tactile stimulus was manipulated in order to elicit (or not) perceptual integration.

Finally, spatial aspects of AT interactions were investigated in a behavioral experiment that is reported in the Appendix. The results of this last study are preliminary and may be used for a comprehensive MEG experiment.

Several methodological features of the project are worth to be highlighted. First of all, two studies are conducted using MEG, a neuroimaging technique that allows high spatial resolution, especially when combined with the latest methods of source localization such as beamforming algorithms. Moreover, MEG data possess

a temporal resolution of millisecond allowing to follow the cortical stages of processing and to observe changes in the network patterns within this temporal resolution. Furthermore, given the several controversial issues that concern the design of multisensory experiments involving neuroimaging techniques (for a review Besle et al. 2004 or see section 1.9), a particular effort was put in this aspect by designing two experiments that allow us to overcome the usage of the linear additive model. Moreover, in all of our experiments, stimuli were presented in the peri-head space where possible effect of AT interactions are supposed to be stronger than in other portion of the space/body (Kitagawa et al. 2005; Tajadura-Jiménez et al., 2009; or see section 1.3).

Finally, regarding the design of the stimuli: all three studies utilized the same sound sample, i.e. a one second long sound of a mosquito. The reason for choosing the mosquito sound is that this sound represents an everyday situation and it is usually accompanied by a tactile event, i.e. the mosquito bite. Besides the ecological validity of the stimuli used, this sound consists of different sound frequencies, the typology of sound which has been proved to be integrated with tactile stimulation at a higher extent than pure tones. Thus, this sound sample was manipulated in amplitude and virtually localized so as to obtain the desired auditory stimuli according to the paradigm required by each of the studies.

2.2 Study I

2.2.1 Introduction

Traditionally AT interactions have been defined as the difference between the sum of the activity elicited from unimodal stimuli and bimodal stimuli, i.e. in the case of audio-tactile AT interaction = $AT - (A + T)$ ('linear additive model', as discussed in detail in the section 1.9). In particular, interactions can show enhanced profiles (defined also as 'facilitative') if the bimodal activity exceeds the sum of

the activity elicited by its unisensory components when presented individually ($AT > A + T$), while they are defined 'suppressive' in the opposite case ($AT < A + T$). In the literature interactions have been found not only under the form of the more intuitive facilitative profiles, but also of suppressive nature. To briefly exemplify some of the results, Lütkenhöner et al. (2002) found suppressive AT interactions at 140 ms that they interpreted as partial inhibition of SII, whereas Gobbelé and colleagues (2003) showed an AT-interaction profile at around 120 ms that was more similar to T than to A, and the authors suggested a suppression of auditory responses. Conversely, Foxe and colleagues (2000) and Murray and colleagues (2005) found facilitative effects (Foxe and colleagues at 65 and 80 ms, Murray and colleagues at 50ms) that they localized as originating from somatosensory and auditory cortices. As it clearly emerges from the few studies mentioned here and from similar studies on AT interactions reviewed in the previous chapter, all in all the pattern of results that emerge is extremely mixed.

A well-established finding, though often under-considered in the design and discussion of multisensory experiments, is that stimuli interact differently according to their respective effectiveness or saliency (Stein and Meredith 1993). For example, electrical stimulation paired to a soft "bip" will generate an AT interaction that differs from the interaction generated by the same tactile stimulus but paired to a loud binaural white noise and these interactions will differ from the ones generated by the same sounds but associated with electrical stimulus and so on. Thus, one major candidate responsible for such strikingly discrepant patterns of results reported in literature is the wide range of different types of stimuli used among these studies. For instance, the participants in Lütkenhöner et al.'s study (2002), who claimed partial inhibition of SII, reported that auditory stimuli were perceived as "more intense" than tactile stimuli (p. 518). Conversely those in Gobbelé and colleagues's study (2003), who reported similar AT-interaction profiles, claimed that the tactile inputs (i.e., the electrical stimulation of the median

nerve) "dominated during AT stimulation" (p. 510).

Indeed, when considering the equation of the linear model used for calculating the interaction, the single component of each stimulus is identical in the bimodal as in the unimodal stimulation, i.e. from both sides of the equation $AT = A + T$. However, the effect of relative saliency between stimuli is present only in the case of bimodal stimulation and thus it is not eliminated by applying the model, on the contrary is well preserved in the found interaction (correctly, since it constitutes an intrinsic part of it). However, since relative saliency varies with the stimuli chosen from study to study, this might represent one of the main causes behind the mixed pattern of results found in literature.

In the present study, in order to investigate the possible role of relative perceived saliencies of the stimuli, we compared MEG recordings of the tactile activity elicited by an identical somatosensory stimulus (i.e. an air puff) that was preceded by various different sounds. In fact the auditory stimuli we used, though identical in their content (a one second long buzz of a mosquito), were manipulated in amplitude so to appear as static and localized near the participants or as static but far from the participants or as approaching the participants (looming). These three sounds, according to literature and as confirmed by a behavioral questionnaire, possess a different saliency. Thus, in the present study we used three sounds characterized by different saliencies that were followed by an always identical somatosensory input. As the tactile event is the event of interest, and as it was always presented right at the end of the sounds and not simultaneously, we compared the activity elicited by three identical tactile events that were however preceded by sounds characterized by different saliencies, thus we can safely attribute all differences we found to our manipulation of sounds' features. Noteworthy, this paradigm avoided to adopt the linear additive model approach, which can introduce multiple confounds if not carefully applied (see s. 1.9).

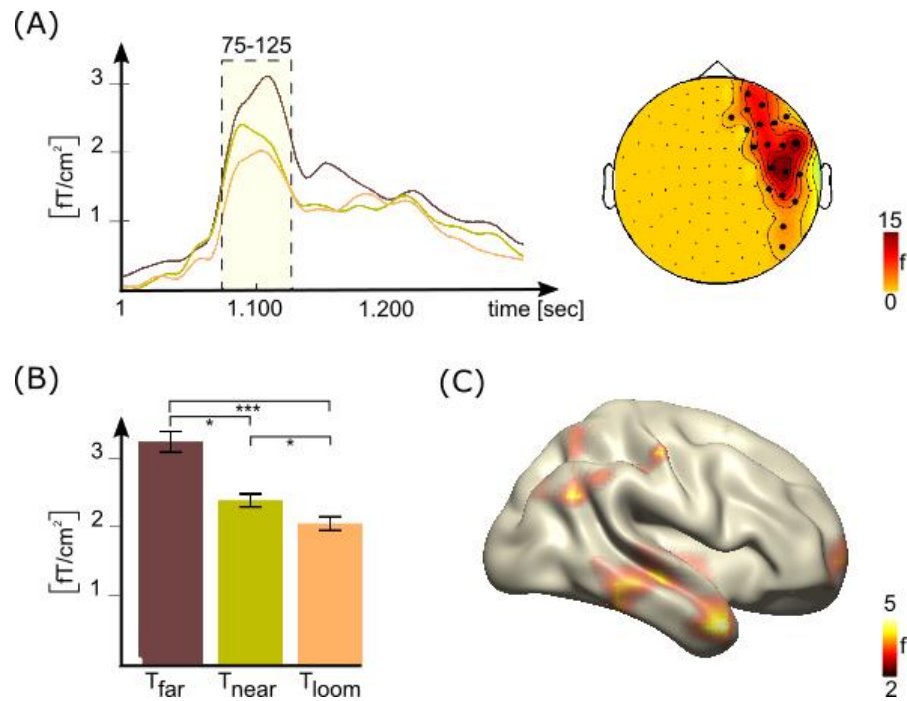
2.2.2 Results

The computation of the differential magnitude of tactile responses across the three conditions allowed us to define the spread of the influence of the different auditory stimuli across time and space. The main findings we report are:

- a change in the magnitude of the elicited somatosensory responses, in the time-locked results: the tactile stimulus responses following the looming sound elicited a smaller activity than the ones following the two static conditions and, among the latter, the smallest response was observed in the static-near condition. This difference in the event related fields (ERF) is at early latency (from 1075 ms to 1125 ms, 1000 correspond to auditory offset and tactile onset) and encompasses the right hemisphere (contralateral with respect to tactile stimulus), peaking over superior temporal sulcus (STS) and temporal pole (TP). (see Figure 2.1)
- a change in the magnitude of the elicited somatosensory responses in the time-frequency results that followed the same pattern as in the time-lock results, i.e. the tactile stimulus responses following the looming sound elicited a smaller activity than the ones following the two static conditions and, among the latter, the smallest response was observed in the static-near condition. Two clusters were identified: a first difference is localized over the right SII (peaking at 1100 ms; 6-8 Hz) and, on a later stage, over the contralateral frontal areas (peaking at 1250 ms; 4-6 Hz). (see Figure 2.2)

Figure 2.1. Summarized results of the time-course analysis.

Time point 1 sec represents the offset of the auditory stimulus and the onset of the tactile stimulus.



(A) Right: event-related field trace of the positive sensor cluster for the tactile event preceded by a static but far sound (T_{far} ; brown), preceded by a static-near sound (T_{near} ; light brown) or by a looming sound (T_{loom} ; orange). The highlighted rectangle shows the statistically significant time-window (1075-1125 ms).

Left: Topography (f values) of the positive sensor cluster (1087-1125 ms), masked for statistical significance. The dots indicate the channels of the significant cluster (1075-1130 ms).

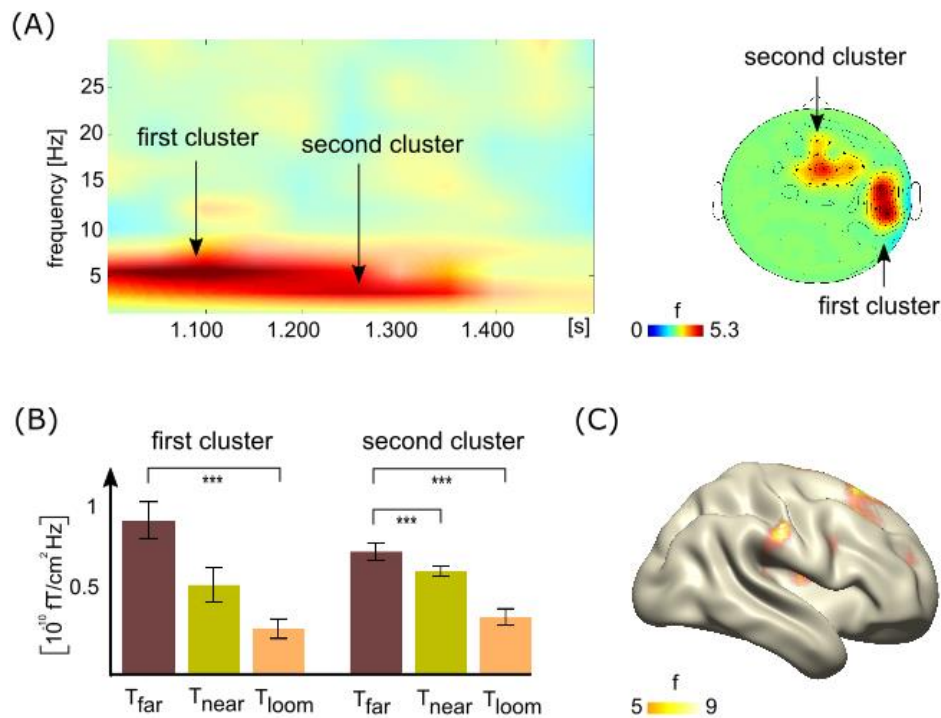
(B) the barplots represent the mean average for the sensors and time-window of the

significant cluster. Error bars show within-subjects 95% confidence intervals. In order to reveal the origin of the statistical differences revealed with the f test, post-hoc analysis was conducted (paired, Bonferroni corrected, t -test). $*p \leq 0.05$, $**p \leq 0.01$, $***p \leq 0.001$.

(C) Figure 2.1D shows the projected f -values obtained from statistical comparison in sensor space between conditions T_{near} , T_{loom} , T_{far} . Like in the topography shown in Figure 2.1A-right, differences are spread in the right hemisphere encompassing frontal, parietal and temporal lobes. The main generators are located in the temporal lobe (TP and STS).

Figure 2.2. Summarized results of the time-frequency analysis.

Time point 1 sec represents the offset of the auditory stimulus and the onset of the tactile stimulus.



(A) Left: time-frequency representation for the time window 1000-1500 ms. Shown are the f - values for comparison of conditions the $T_{/far}$, $T_{/near}$, $T_{/loom}$ T of the combined gradiometers. The arrows indicate roughly the two significant clusters identified. First cluster: $p < .03$, 1000–1270 ms, 6-8 Hz, peaking at 1100 ms. Second cluster: $p < .02$, 1030–1370 ms, 4-6 Hz, peaking at 1250 ms. Non-significant values have been reduced to 40% opacity.

Right: Topography of the significant clusters. The first cluster is localized centrally towards the front, whereas the second is right-lateralized.

(B) The bar plots show the means and the standard errors of the mean of the power for the sensors, time-window and frequency of the two significant clusters. Error bars show within-subjects 95% confidence intervals. $*p \leq 0.05$, $**p \leq 0.01$, $***p \leq 0.001$.

(C) Projection in source space of the statistical comparison (f values) across the three conditions for time window 1000–1370 ms in the frequency 4 to 8 Hz. The areas corresponding to the first cluster (1250 ms, 4-6Hz) can be identified with the right secondary somatosensory cortex (BA40 spreading into BA2). The second cluster originate in correspondence with the right dorsolateral prefrontal cortex (BA 8).

2.2.3 Discussion

In Study I, by changing the saliency of a (preceding) auditory stimulus and keeping a (subsequent) tactile stimulus always identical, we investigated how the relative saliency between the stimuli impact the magnitude of the activity elicited by the tactile stimulus. We further confirmed thus the relevance for multisensory interactions of the relative saliency between stimuli. Possibly, this factor plays a

crucial role in explaining the discrepant pattern of results found in the literature of AT interactions.

Additionally, we show an inverse relationship between the perceived saliency of the acoustic stimulus and the strength of this modulation, with more salient auditory stimuli being associated to less tactile activity. The fact that these three auditory stimuli possess a different saliency is supported by evidence in the literature (e.g. Ahveninen et al., 2014; Bach et al., 2009; Hall and Moore, 2003; Kopco et al., 2012; Tonetto et al., 2014) and has been further confirmed by a questionnaire we performed during in the experiment. We believe that the perceived saliency of the auditory stimuli, rather than acoustic intensity cues per se, impact the following tactile activity because the rapidly approaching sound is followed by the smallest tactile activity, whereas the near stimulus – that overall possesses the highest intensity but relative less saliency, given that is static – is followed by an intermediate response and the far sound – static and low-intensity – by the biggest response.

This pattern of results is well in accordance to the law of inverse effectiveness (Stein and Meredith, 1993) which has long been considered one of the basic principles of the multisensory field. According to the inverse effectiveness rule, the strength of multisensory interactions increases when the stimuli efficacy (usually measured in terms of saliency) of the isolated components are relatively weak by themselves.

However, an alternative explanation to our findings is possible when recurring to the crossmodal shared attention mechanism, i.e. by assuming attention to be a limited resource and stimulus representations and different modalities are in competition with one another based on their representational strength (saliency): the more salient a stimulus representation, the more it will dominate the competition (e.g. Spence and Driver, 2004). The different types of sounds used in the present study are endowed with different levels of saliency, so differences in

attentional bias cannot be excluded. However it is worth noting that in our paradigm the sounds were already off in time window on which the analysis was centered and that potential sound offset effects were ruled out.

Another very interesting aspect of the results emerging from Study I is that this differential magnitude of responses across the three conditions allowed us to track the spread of the exerted crossmodal influence across time and space. Our results show a first modulation of amplitude at early latency (75 ms after tactile stimulus onset) over the right temporal lobe (STS and TP). Time-frequency results further revealed differences across the three conditions: the earlier cluster is localized over secondary somatosensory cortex (BA40/2; 6-8Hz, peak of activity at 100 ms after tactile stimulus onset) and finally, differences spread towards more frontal areas (BA 8; 4-6Hz, peak of activity at 250 ms after tactile stimulus onset). Noteworthy, the differences start in area close to primary auditory, extend to secondary somatosensory and finally move towards more frontal. This pattern further support the idea that multisensory interactions can be present at different level and stages of cortical processing (see s. 1.5).

2.3 Study II

2.3.1 Introduction

A fundamental ability of our brain is to generate a coherent and seamless representation of the external multisensory world. To this aim, the brain has to perceptually integrate (or bind or unify) information coming from different sensory modalities that belong to the same object/event, while unrelated incoming signals are perceptually segregated.

To note that a confound arises from the fact that in the multisensory literature, the term “integration” has been used for referring to every multisensory phenomenon, also to phenomenon generated by combination of relatively

meaningless multisensory stimuli, as for example a “bip” paired to and electrocutaneous stimulus (Foxy et al. 2000). A clarification in this sense, come from Driver and Noesselt (2008) where they differentiate between multisensory “interplay” (or interaction), when the processing of a sensory modality is affected by the presence of one (or more) stimuli from other modalities. While the term multisensory “integration” properly refers to the presence of multisensory interplay and of an integrated perception of the stimuli, as emanating from one event. In this study we aimed to investigate multisensory integration in the proper sense and in the present thesis we refer to this nomenclature.

It is generally thought that the solution of this cross-modal integration problem, depends on stimulus-driven factors such as spatial, temporal properties of the stimuli (Calvert et al., 2004) as well as on cognitive factors (the “unity assumption” (Vatakis and Spence, 2007; Vatakis et al., 2008; Welch and Warren, 1980). Paradigmatic examples are the cases where multisensory discordant cues are perceptually unified and generate illusory percepts. As in the “ventriloquism effect”, where the presence of an apparent speaker alters the perceived location of speech sounds as coming from the direction of the speaker (Bertelson, 1998). Interestingly, when the visual stimulus becomes blurred the opposite occurs with visual stimulus captured by audition (Alais and Burr, 2004). Another paradigmatic example, is the McGurk illusion in which conflicting audio (the syllable “ba”) and visual (the labial of the syllable “fa”) information is erroneously bind together generating a third new illusory auditory percept (the syllable “ga”). However, little is known about the neural mechanism underlying the binding operation (though see Bushara et al., 2003 for and fMRI study on this topic).

In this MEG study we investigated perceptually integrated (INT) or segregated (SEG) multisensory AT perceptions. In particular, we created a bistable paradigm, i.e. identical AT stimuli that could be alternatively perceived either as emanating from one event (INT) or from two distinct events (SEG). To this aim we

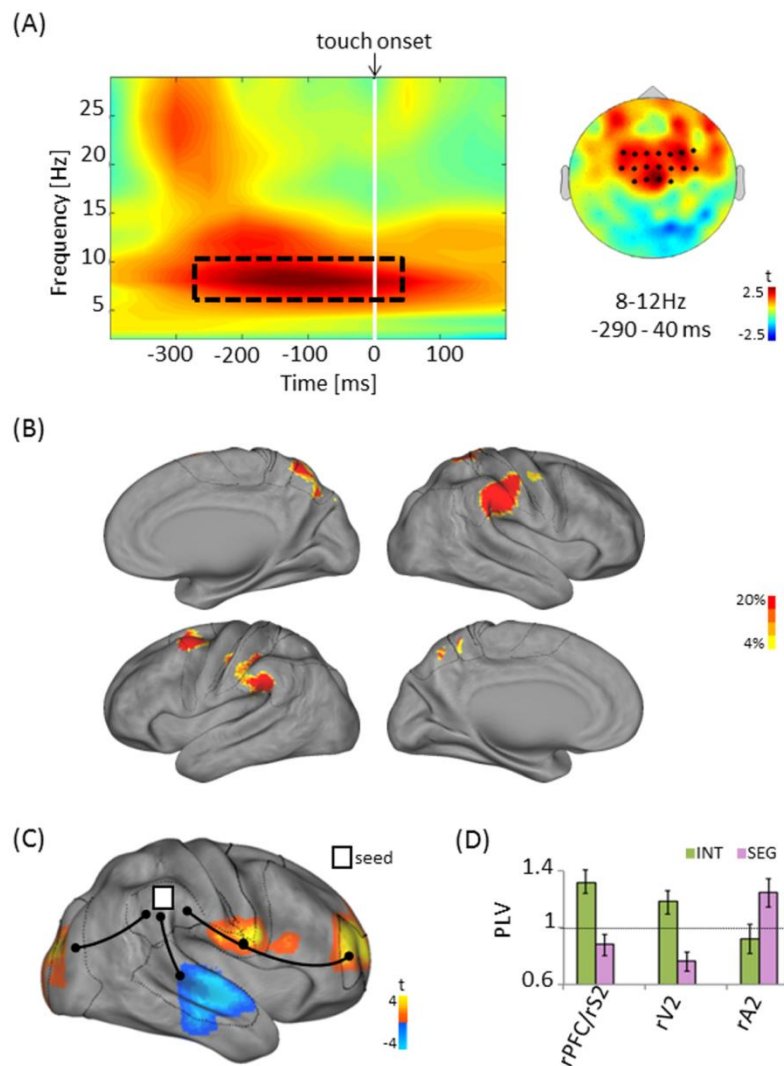
used an approaching one-second long auditory stimulus (the buzz of a mosquito) and a short tactile stimulus (delivered at the peri-oral area). The stimuli were spatially related (both presented in the left peri-head space) and semantically correlated (the sound of a mosquito and a touch that resembled its bite), while we manipulated the relative timing (stimulus onset asynchrony, SOA). In fact, perceptual integration rely critically on the temporal congruence of the stimuli (Calvert et al., 2004; Welch, 1999), more than the spatial one (e.g., the ventriloquist effect). So, in the experiment stimulation was delivered at a particular SOA (individually determined) that rendered the temporal relationship between the audio and the tactile stimuli ambiguous: the tactile stimulus in 50% of the cases was perceived as the bite of the approaching mosquito (integrated percept) and 50% of the cases as two different events (segregated percept). In this manner, we could dissociate the simple co-occurrence of two stimuli from cross-modal binding and investigate the underlying mechanism. This paradigm allows direct comparison of physically identical multisensory stimuli that however elicit different multisensory percepts. In this way, we overcome the limitations introduced by the linear additive model such as sensitivity only to responses nonlinearities or introduction of possible confounds (for a deeper discussion of the topic see section 1.9). Since with our paradigm we exclude the fact that integrated and segregated percepts are due to differences between stimuli because they are physically identical, we hypothesize that different brain-states prior to the upcoming tactile stimulus predispose different multisensory percepts. Indeed, an increasing body of evidence is showing an influence of prestimulus activity on subsequent perception, especially in the case of near-threshold stimuli (see Ruhnau et al., 2014 for a review) but also in multisensory perception (e.g. Hipp et al., 2011; Keil et al., 2012; see Lange et al., 2014 for a revision).

2.3.2 Results

By comparing physically identical trials that however elicited different percepts (segregated vs integrated) we found:

- time-frequency analysis revealed that the perception of integrated AT stimulus with respect to a segregated percept is associated with a prestimulus increase in alpha power (9 Hz) . This difference is localized bilaterally in inferior and superior parietal lobules (IPLs and SPLs) among a time window of 300 ms prior to tactile stimulus. (see Figure 2.3 A-B)
- prior to tactile stimulus, connectivity analysis, seeded in the rIPL (point of maximum difference between conditions) revealed that prior to a unified percept with respect to segregated percept, this region for the frequency of 9 Hz is characterized by more synchronized oscillations with right prefrontal cortex (rPFC), right secondary somatosensory and visual cortices (rSII, rVII) , while less synchronized oscillations with auditory regions. (see Figure 2.3 C-D)

Figure 2.3



(A) Results of the statistical comparison at sensor level for condition INT vs condition SEG. Red color indicates relatively more power during condition INT. Time 0 ms indicates the onset of touch.

Left: time-frequency representation (TFR) of the statistically significant sensors (cluster- $p < 0.05$) for the interval pre- and post- tactile stimulus. The dashed rectangle highlight the significant time-frequency window

Right: Topographical representation of the statistical results for the time-window -290 to 40 ms in the frequency band 8-12 Hz. Black dots represent the sensors that gave statistically significant results at sensor level and whose TFR are plotted on the left side of the figure.

(B) Projection in source space of the grand average of INT versus SEG for the time-frequency window and frequency of the significant results at sensors level.

(C) Phase synchrony between the seed region, rIPL ((60 – 40 40), MNI coordinates), and the whole-brain volume. T-values of INT versus SEG percept masked for statistical significance are displayed. Dashed lines represent the borders of Brodmann areas of interest. rAII exhibited significant decoupling with the rIPL, while rPFC/rSII and rVII showed increased coupling with rIPL during integrated compared to segregated percept.

(D) We calculated PLV of each significant cluster showed in (C) normalized to the baseline prior to the sound onset in order to evaluate if the differences reported in (C), originate by, respectively, a coupling and a decoupling in both conditions with respect to baseline or, on the contrary, the effect is driven only by a coupling (or decoupling) in one condition. PLV greater than 1, indicated by the dashed line, indicates higher coupling with respect to baseline and vice versa. Barplots show differences of PLV in the rPFC/rSII and rVII originate from an increase of coupling for INT and a decrease of coupling for SEG with respect to the baseline, while for the rAII the difference is primarily driven by an increase of coupling for the SEG percept

2.3.3 Discussion

In Study II we investigated the neural correlates of perceptual integration between an audio and a tactile stimulus. Experiments in multisensory field usually compared neuronal responses between well-defined bimodal and unimodal stimuli and focused on the non-linear differences in activation, according to the additive linear model (see s. 1.9). Here, we created a bi-stable paradigm, i.e. upon invariant bimodal AT stimulation, participants reported different percepts. The advantage of our paradigm is that it allows direct comparison of instances of identical bimodal AT stimuli that are however differently perceived (integrated or segregated), thus rendering the experiment sensitive also to linear AT interactions and moreover limiting the interpretation of results only to our manipulation, given that a part different percepts the two conditions (segregated and integrated) were physically identical. We found differences localized bilaterally over IPLs (and SPLs), in a prestimulus time-window of about 300 ms in the frequency band of alpha.

Several studies in macaque monkeys showed auditory and somatosensory convergence in area adjacent to the primary auditory cortex (AI), the caudomedial (CM) belt area (Brosch et al., 2005; Fu et al., 2003; Kayser et al., 2005; Schroeder and Foxe, 2002; Schroeder et al., 2001). In monkeys, CM is positioned between AI, the temporal parietotemporal areas (Tpt), and retroinsular area (Ri). Studies of auditory cortex on humans individuate their homologues in the parietal operculum and IPL (Galaburda and Sanides, 1980; Hackett et al., 2001; Sweet et al., 2005). Previous studies in humans revealed correlates of audiotactile interaction in approximately this region (Caetano and Jousmäki, 2006; Foxe et al., 2002, 2000; Gobbelé et al., 2003; Lütkenhöner et al., 2002; Schürmann et al., 2006). Our findings further support that IPL is a site of where AT interactions occur and the human homologue of the multisensory area CM in monkeys. Interestingly, in previous literature, IPLs and SPLs have been described as being activated during a localization task for both auditory and tactile stimuli (Renier et al., 2009), and relevant in localization tasks for stimuli of all modalities (e.g. Bushara et al., 1999;

Reed et al., 2005). Furthermore, rIPL appears to be involved in perception of time (e.g. Harrington et al., 1998; Rao et al., 2001) and the detection of temporal synchrony between auditory and visual stimuli (Adhikari et al., 2013; Bushara et al., 1999; Dhamala et al., 2007), and in tasks where multisensory temporal and spatial information need to be evaluated together (Assmus et al., 2005, 2003). In our paradigm, we manipulated the temporal relationship between the two stimuli. Given that the sound is looming, time is translated into the location of the sound source so that the integration/segregation task also had a spatial component, i.e. the localization of the flying mosquito, based on the loudness of its buzzing sound, at the time the tactile stimulus is delivered. Although the role of IPLs and SPLs described above has been determined based on post-stimulus activities, in our experiment the difference in IPLs and SPLs prior to the application of the tactile stimulus between conditions INT and SEG suggests that the ongoing brain state affects upcoming computations of spatio-temporal relations between stimuli of different modalities (here, audio and tactile).

Regarding the role of alpha, activity in the alpha-band has long being associated with a state of idling (Pfurtscheller et al., 1996). However, according to current literature, activity in this frequency band plays an active role in the brain, in particular in the inhibitory-excitatory balance: higher alpha in a certain brain region is related to local inhibition and vice versa (Jensen and Mazaheri, 2010). In particular, differences in the level of alpha power in the prestimulus window can modulate the perception of near-threshold (NT) stimuli (e.g. Hanslmayr et al., 2007; Jones et al., 2010; Romei et al., 2010; Thut et al., 2006; van Dijk et al., 2008; Weisz et al., 2014). In the sense that low levels of alpha power preceded “hits” (the NT stimulus is perceived), while high levels of alpha power preceded misses (the NT stimulus is not perceived). Interestingly, recent studies using super-threshold stimuli extended this binary role of alpha in either inhibiting or processing of a weak stimulus, to modulation of the quantity of supra-threshold stimuli perceived

(one or two) within visual illusions such as fusion effect (FE) and double flash illusion (DFI) (Lange et al., 2013) or within two tactile stimuli when presented in close temporal proximity (Baumgarten et al., 2014). In these studies, low level of alpha power preceded detection of two stimuli (independent from the fact that they were real, as in the FE or illusory as in the DFI) while, following higher alpha-power levels, only one stimulus was perceived. All in all our results are well in line with previous findings (Baumgarten et al. 2014; Lange et al. 2013) in the alpha band with increased excitability preceding the perception of two stimuli (segregated percept) and vice versa. For the first time we show a similar mechanism in binding across modalities.

Given our finding on IPLs (and SPLs) and the proven responsiveness of the areas to multisensory stimuli, we can reasonably suppose that these areas are critical for perceptual integration suggesting a specific role of these in spatio-temporal separation of incoming AT stimuli and expanding our knowledge of the functional role of alpha band to crossmodal perception.

A second very interesting aspect of our study are the results obtained by calculating brain connectivity with a seed placed in the point of maximum difference between conditions (rIPL) which exhibited differential patterns in a prestimulus time window. This support the hypothesis that not only local power differences over relevant regions, but also that the inclusion or exclusion of these regions into a globally distributed functional network critically influences subsequent perception (Hanslmayr et al., 2007; Keil et al., 2012; Ploner et al., 2010; Weisz et al., 2014). Our results suggest rather than a specific multisensory area, integrated/segregated perception is predisposed by a network comprising different areas. Integration of AT stimuli was preceded by a large-scale functional network involving distant region of the brain. It goes behind the purpose and the possibilities of this experiment to infer in detail which the role and the ratio of the particular functional architecture we found, however some remarks are noteworthy

to be mentioned. For condition INT, with respect to conditions SEG, rIPL is oscillating more in synchrony with rPFC, and PFC has been shown to be critical in situations when top-down processing is required (e.g. Miller, 2000). rIPL is disconnected from the superior temporal gyrus-sulcus (STG-STS). This region is a polysensory region (audio, visual, tactile) in monkeys as well in humans (see s.1.6). The fact that the differentiated network comprised somatomotor and even visual areas, despite the presentation of AT stimuli, suggests a non-modality specific integration/segregation network, possibly for exploiting any upcoming information. We can speculate that the precise network configuration is obtained through the local inhibitory/excitatory balance, as inhibition is essential for the establishment of long-range networks (Jonas and Buzsaki, 2007) so that the differential alpha levels observed in rIPLs in the two conditions is related to the formation of the network relevant for multisensory integration.

2.4 General conclusions

During the past decade, the field of multisensory research has expanded and has received increasing interest. However, as the mixed and contrasting pattern of results present in the literature shows, little is known about how sensory information interacts. The overall aim of this PhD study was to investigate further into AT interactions.

Interactions between these two sensory modalities have been under-investigated, maybe because they go often unnoticed. However, they are well present in many everyday life situations. Furthermore, a number of similarities exist between the senses of hearing and touch that may suggest the existence of a favored link between both modalities. For instance, mechanisms for transducing sensory signals are analogous for sounds and vibrotactile stimuli, analogies in the physiological mechanisms supporting the idea that the two sensory modalities have

common origins or the fact that auditory and tactile signal transduction times from the ear and skin are faster and comparable in time, as opposed to visual signals with much longer transduction times (see s. 1.2).

Studying AT interactions is interesting not only in and of itself, but can provide insights into the broad understanding of how different senses interact, regardless of their modalities. So we attempted to exploit these two modalities to study several interesting aspects strictly related to AT interactions but that however are also of broad interest of the multisensory research field.

In order to pursue our goals, we stressed methodological features. First of all we conducted the research in the portion of the space surrounding the head, a portion of space where AT interactions possess a particular relevance, though studies on AT usually has been conducted in the space of the hands. Possibly because in the space of the head is more difficult to provide stimulation (piezoelectrical or electrical stimulation for example cannot be provided). Moreover, we broadly exploited a particularly relevant class of signals for multisensory interactions, i.e. the looming ones. Furthermore, multisensory research is particularly demanding since, by increasing the number of modalities involved, their interaction raises the possible interpretation of results and the number of confounds and in order to control for them an exponential number of control conditions are usually required. We put a lot of effort in designing the experimental paradigms so to limit interpretation of our results to purely our desired manipulation (e.g. relative saliency or perceptual binding). Finally, by using MEG we could rely on a temporal resolution in the order of millisecond and on a spatial resolution of centimeter.

All in all our results add new interesting findings to crucial aspects of multisensory interactions at cortical level. A first aspect is that our study confirms how relative saliency between crossmodal stimuli can dramatically impact the activity elicited. Although this fact is known since decades (see “inverse

effectiveness rule”, Stein and Meredith 1993), it has been scarcely considered in the design of paradigms of “classical” neuroimaging studies on AT interactions. However, our results highlight the main role this factor plays in the interaction elicited and how more effort should be put into taking it into account. Thus, our results suggest how relative saliency of the stimuli has to be considered as a major factor able to dramatically influence the results of any multisensory experiment. Another interesting highlight emerged in Study I is the time and space across which the auditory stimulus impact the tactile processing. Noteworthy, it started in area close to primary auditory, extended to secondary somatosensory and finally moved towards more frontal. This pattern further support the idea that multisensory interactions can be present at different level and stages of cortical processing (e.g. Ghazanfar and Schroeder 2006). Though, here even crossmodal modulations are present at several levels of cortical processing, the “traditional” hierarchy of cortical processing is preserved.

In Study II we found that different level of alpha power, which according to current literature indicated different level of cortical excitability, over IPLs (and SPLs) in a prestimulus (tactile stimulus) predispose a different AT percept (segregated vs integrated). This finding is line with a growing body of evidence which underlines the importance of prestimulus brain states, in particular in the alpha band, in emerging of a certain percept (see Ruhnau et al. 2014). In the multisensory research field, the importance of prestimulus brain-state has already been shown in the creation of illusory percept such as the McGurk illusion, the double-flash and the fusion effect (see Lange et al., 2014 for a review). Interestingly, our nicely complement these findings extending its influence also to creation of an integrated/segregated crossmodal percepts.

Moreover, in this prestimulus time-window, rIPL (point where difference in alpha power is maximum) is oscillating with different synchrony within a network that encompasses several traditionally unisensory areas and prefrontal cortex.

Large-scale interactions modify the communication structure between brain areas that are functionally and anatomically specialized, leaving anatomical connections unchanged (Varela et al., 2001) and create distinct functional network architectures that constitute privileged pathways along which neural information will flow. Thus, large-scale interactions provide flexibility and plasticity and allow complex coordination between different brain regions. These fundamental properties are required to integrate incoming inputs from a dynamic and multifaceted world and endogenous activity and could constitute a mechanism of interactions between senses that complements the traditional hierarchical model. This further finding is in line with the most recent framework introduced by Weisz and colleagues (2014) according to which not only local prestimulus oscillatory power, but also global prestimulus connectivity architectures predispose different percepts, possibly by predetermining information flow.

2.5 Final considerations and future directions

All in all our results support the significance of AT interactions, especially for experiments conducted in the space of head. Although it is not obvious that mechanism holding for interactions between A and T can be extended straightforward to other modality pairings, still studying of AT interactions can add relevant insights into the complex interactions between senses that can occur in the brain.

Study I, suggest the importance of relative stimuli saliencies. It would be of broad interest to further deep our understanding the effects of this factor on interactions, for giving a more comprehensive interpretation of multisensory findings. For example, one might think about the reverse design of Study I, i.e. keeping auditory stimulus constant while changing the tactile stimulus saliency (e.g. air-puff vs electrical stimulation vs piezoelectrical stimulation) or experiments

with a similar paradigm but involving other modality pairings. In general, further experiments should be conducted where interactions between all senses are evaluated in function of parametric changes of saliencies. Ideally, by conducting such experiments one could think of creating a scale for evaluating relative stimuli saliencies that could be then used as a metric to account for relative saliencies used across multisensory experiments.

Study II suggests the importance of prestimulus brain states, alpha power over multisensory regions and different connectivity states, in creation of different multisensory percepts. Similar paradigms could be used to study perceptual integration between also different modality pairings, especially in light of our findings that involve also visual cortices. Moreover, it could be interesting to further investigate the role of rIPL in building of crossmodal percepts by using TMS. Finally, it would be interesting to investigate the role of expectation by creating a paradigm where one or other percepts are cued with a different probability.

CHAPTER 3

References

- Adhikari, B. M., Goshorn, E. S., Lamichhane, B., & Dhamala, M. (2013). Temporal-order judgment of audiovisual events involves network activity between parietal and prefrontal cortices. *Brain Connectivity*, *3*(5), 536–45. doi:10.1089/brain.2013.0163
- Ahveninen, J., Kopco, N., & Jääskeläinen, I. (2014). Psychophysics and neuronal bases of sound localization in humans. *Hearing Research*, *307*, 86–97. doi:10.1016/j.heares.2013.07.008.Psychophysics
- Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology : CB*, *14*(3), 257–62. doi:10.1016/j.cub.2004.01.029
- Assmus, a, Marshall, J. C., Noth, J., Zilles, K., & Fink, G. R. (2005). Difficulty of perceptual spatiotemporal integration modulates the neural activity of left inferior parietal cortex. *Neuroscience*, *132*(4), 923–7. doi:10.1016/j.neuroscience.2005.01.047
- Assmus, A., Marshall, J. C., Ritzl, A., Noth, J., Zilles, K., & Fink, G. R. (2003). Left inferior parietal cortex integrates time and space during collision judgments. *NeuroImage*, *20*, S82–S88. doi:10.1016/j.neuroimage.2003.09.025
- Bach, D. R., Neuhoff, J. G., Perrig, W., & Seifritz, E. (2009). Looming sounds as warning signals: the function of motion cues. *International Journal of Psychophysiology*, *74*(1), 28–33. doi:10.1016/j.ijpsycho.2009.06.004
- Barnett-Cowan, M., & Harris, L. R. (2009). Perceived timing of vestibular stimulation relative to touch, light and sound. *Experimental Brain Research*, *198*(2-3), 221–31. doi:10.1007/s00221-009-1779-4
- Baumgarten, T. J., Schnitzler, A., & Lange, J. (2014). Prestimulus Alpha Power Influences Tactile Temporal Perceptual Discrimination and Confidence in Decisions. *Cerebral Cortex* . doi:10.1093/cercor/bhu247
- Beauchamp, M. S., Argall, B. D., Bodurka, J., Duyn, J. H., & Martin, A. (2004). Unraveling multisensory integration: patchy organization within human STS multisensory cortex. *Nature Neuroscience*, *7*(11), 1190–2. doi:10.1038/nn1333
- Beauchamp, M. S., Yasar, N. E., Frye, R. E., & Ro, T. (2008). Touch, sound and vision in human superior temporal sulcus. *NeuroImage*, *41*(3), 1011–1020.

- Bertelson, P. (1998). Starting from the ventriloquist: The perception of multimodal events. In M. Sabourin, F.I.M. Craik, & M. Robert. In P. Press (Ed.), *Advances in psychological science: Vol. 2. Biological and cognitive aspects* (Vol. 2, pp. 419–439). Hove, England.
- Bertelson, P., & Schersleben, G. (1998). Automatic visual bias of perceived auditory location. *Psychonomic Bulletin & Review*, 5(3), 482–489.
- Besle, J., Fort, A., & Giard, M.-H. (2004). Interest and validity of the additive model in electrophysiological studies of multisensory interactions. *Cognitive Processing*, 5, 189–192. doi:10.1007/s10339-004-0026-y
- Bremmer, F., Klam, F., Duhamel, J.-R., Ben Hamed, S., & Graf, W. (2002). Visual-vestibular interactive responses in the macaque ventral intraparietal area (VIP). *European Journal of Neuroscience*, 16(8), 1569–1586. doi:10.1046/j.1460-9568.2002.02206.x
- Brosch, M., Selezneva, E., & Scheich, H. (2005). Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 25(29), 6797–806. doi:10.1523/JNEUROSCI.1571-05.2005
- Bruce, C., Desimone, R., & Gross, C. G. (1981). Visual properties of neurons in a polysensory area in the superior temporal sulcus of the macaque. *Journal of Neurophysiology*, 46(2), 369–384.
- Bushara, K. O., Hanakawa, T., Immisch, I., Toma, K., Kansaku, K., & Hallett, M. (2003). Neural correlates of cross-modal binding. *Nature Neuroscience*, 6(2), 190–5. doi:10.1038/nn993
- Bushara, K. O., Weeks, R. a, Ishii, K., Catalan, M. J., Tian, B., Rauschecker, J. P., & Hallett, M. (1999). Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. *Nature Neuroscience*, 2(8), 759–66. doi:10.1038/11239
- Caetano, G., & Jousmäki, V. (2006). Evidence of vibrotactile input to human auditory cortex. *NeuroImage*, 29(1), 15–28. doi:10.1016/j.neuroimage.2005.07.023
- Calvert, G. a, Hansen, P. C., Iversen, S. D., & Brammer, M. J. (2001). Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. *NeuroImage*, 14(2), 427–38. doi:10.1006/nimg.2001.0812
- Calvert, G. A., Spence, C., & Stein, B. E. (2004). *The handbook of multisensory processes*.
- Canzoneri, E., Magosso, E., & Serino, A. (2012). Dynamic sounds capture the boundaries of peripersonal space representation in humans. *PloS One*, 7(9), e44306. doi:10.1371/journal.pone.0044306
- Cappe, C., & Barone, P. (2005). Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. *The European Journal of Neuroscience*, 22(11), 2886–902. doi:10.1111/j.1460-9568.2005.04462.x

- Cappe, C., Thut, G., Romei, V., & Murray, M. M. (2009). Selective integration of auditory-visual looming cues by humans. *Neuropsychologia*, *47*(4), 1045–52. doi:10.1016/j.neuropsychologia.2008.11.003
- Cavallin, B. A., & Houston, B. K. (1980). Aggressiveness, maladjustment, body experience and the protective function of personal space. *J. Clin. Psychol.*, *36*, 170–176. doi:10.1002/1097-4679(198001)36
- Cooke, D. F., Taylor, C. S. R., Moore, T., & Graziano, M. S. a. (2003). Complex movements evoked by microstimulation of the ventral intraparietal area. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(10), 6163–8. doi:10.1073/pnas.1031751100
- Cooper, W., & Vitt, L. (2002). Increased predation risk while mate guarding as a cost of reproduction for male broad-headed skinks (*Eumeces laticeps*). *Acta Ethologica*, *5*(1), 19–23. doi:10.1007/s10211-002-0058-1
- Dhamala, M., Assisi, C. G., Jirsa, V. K., Steinberg, F. L., & Kelso, J. a S. (2007). Multisensory integration for timing engages different brain networks. *NeuroImage*, *34*(2), 764–73. doi:10.1016/j.neuroimage.2006.07.044
- Dosey, M., & Murray, M. (1969). Personal space and self-protection. *Journal of Personality and Social Psychology*, *11*(2), 93–97.
- Driver, E. C., & Kelley, M. W. (2009). Specification of cell fate in the mammalian cochlea. *Birth Defects Research Part C: Embryo Today: Reviews*, 212–221. doi:10.1002/bdrc.20154.Specification
- Driver, J., & Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on “sensory-specific” brain regions, neural responses, and judgments. *Neuron*, *57*(1), 11–23. doi:10.1016/j.neuron.2007.12.013
- Duhamel, J., Colby, C. L., & Goldberg, M. E. (1998). Ventral Intraparietal Area of the Macaque: Congruent Visual and Somatic Response Properties. *Journal of Neurophysiology*, *79*, 126–136.
- Eilam, D. (2005). Die hard: a blend of freezing and fleeing as a dynamic defense-- implications for the control of defensive behavior. *Neuroscience and Biobehavioral Reviews*, *29*(8), 1181–91. doi:10.1016/j.neubiorev.2005.03.027
- Farnè, A., & Làdavas, E. (2002). Auditory peripersonal space in humans. *Journal of Cognitive Neuroscience*, *14*(7), 1030–43. doi:10.1162/089892902320474481
- Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT press.
- Foxe, J. J., Morocz, I. A., Murray, M. M., Higgins, B. A., Javitt, D. C., & Schroeder, C. E. (2000). Multisensory auditory-somatosensory interactions in early cortical processing revealed by high-density electrical mapping. *Brain Research. Cognitive Brain Research*, *10*, 77–83. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10978694>
- Foxe, J. J., Wylie, G. R., Martinez, A., Schroeder, C. E., Daniel, C., Guilfoyle, D., ... Schwartz, T. H. (2002). Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. *Journal of Neurophysiology*, *88*, 540–543.

- Fu, K.-M. G., Johnston, T. a, Shah, A. S., Arnold, L., Smiley, J., Hackett, T. a, ... Schroeder, C. E. (2003). Auditory cortical neurons respond to somatosensory stimulation. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 23(20), 7510–5. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12930789>
- Galaburda, a, & Sanides, F. (1980). Cytoarchitectonic organization of the human auditory cortex. *The Journal of Comparative Neurology*, 190(3), 597–610. doi:10.1002/cne.901900312
- Ghazanfar, A. a, Neuuhoff, J. G., & Logothetis, N. K. (2002). Auditory looming perception in rhesus monkeys. *Proceedings of the National Academy of Sciences of the United States of America*, 99(24), 15755–7. doi:10.1073/pnas.242469699
- Ghazanfar, A. a, & Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences*, 10(6), 278–85. doi:10.1016/j.tics.2006.04.008
- Gobbelé, R., Schürmann, M., Forss, N., Juottonen, K., Buchner, H., & Hari, R. (2003). Activation of the human posterior parietal and temporoparietal cortices during audiotactile interaction. *NeuroImage*, 20(1), 503–511. doi:10.1016/S1053-8119(03)00312-4
- Grassi, M. (2010). Sex difference in subjective duration of looming and receding sounds. *Perception*, 39(10), 1424–1426. doi:10.1068/p6810
- Graziano, M. S. ., Gross, C. G., Taylor, C. S. R., & Moore, T. (2004). A system of multimodal areas in the primate brain. In Spence C., Driver J. In O. U. Press (Ed.), *Crossmodal space and crossmodal attention* (pp. 51–67). Oxford, England.
- Graziano, M. S. a, & Cooke, D. F. (2006). Parieto-frontal interactions, personal space, and defensive behavior. *Neuropsychologia*, 44(6), 845–59. doi:10.1016/j.neuropsychologia.2005.09.009
- Graziano, M. S. a., & Gandhi, S. (2000). Location of the polysensory zone in the precentral gyrus of anesthetized monkeys. *Experimental Brain Research*, 135, 259–266. doi:10.1007/s002210000518
- Graziano, M. S. A., Taylor, C. S. R., & Moore, T. (2002). Complex movements evoked by microstimulation of precentral cortex. *Neuron*, 34, 841–851.
- Graziano, M. S., & Gross, C. G. (1998). Spatial maps for the control of movement. *Current Opinion in Neurobiology*, 8(2), 195–201. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9635202>
- Graziano, M. S., Hu, X. T., & Gross, C. G. (1997). Visuospatial properties of ventral premotor cortex. *Journal of Neurophysiology*, 77(5), 2268–92. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9163357>
- Graziano, M. S., Reiss, L. a, & Gross, C. G. (1999). A neuronal representation of the location of nearby sounds. *Nature*, 397(6718), 428–30. doi:10.1038/17115
- Gregory, R. L. (1967). Origin of Eyes and Brains. *Nature*, 213, 369–372.

- Hackett, T. A., Preuss, T. M., & Kaas, J. O. N. H. (2001). Architectonic Identification of the Core Region in Auditory Cortex of Macaques, Chimpanzees, and Humans, *222*(September), 197–222. doi:10.1002/cne.1407
- Hall, D. a., & Moore, D. R. (2003). Auditory neuroscience: the salience of looming sounds. *Current Biology : CB*, *13*(3), R91–3. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12573234>
- Hanslmayr, S., Aslan, A., Staudigl, T., Klimesch, W., Herrmann, C. S., & Bäuml, K.-H. (2007). Prestimulus oscillations predict visual perception performance between and within subjects. *NeuroImage*, *37*(4), 1465–73. doi:10.1016/j.neuroimage.2007.07.011
- Harrington, D. L., Haaland, K. Y., & Knight, R. T. (1998). Cortical networks underlying mechanisms of time perception. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *18*(3), 1085–95. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9437028>
- Hipp, J. F., Engel, A. K., & Siegel, M. (2011). Oscillatory synchronization in large-scale cortical networks predicts perception. *Neuron*, *69*(2), 387–96. doi:10.1016/j.neuron.2010.12.027
- Holmes, N. P., & Spence, C. (2004). The body schema and the multisensory representation(s) of peripersonal space. *Cognitive Processing*, *5*(2), 94–105. doi:10.1007/s10339-004-0013-3
- Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Frontiers in Human Neuroscience*, *4*(November), 186. doi:10.3389/fnhum.2010.00186
- Jonas, P., & Buzsaki, G. (2007). Neural inhibition.
- Jones, S. R., Kerr, C. E., Wan, Q., Pritchett, D. L., Hämäläinen, M., & Moore, C. I. (2010). Cued spatial attention drives functionally relevant modulation of the mu rhythm in primary somatosensory cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(41), 13760–5. doi:10.1523/JNEUROSCI.2969-10.2010
- Jousmäki, V., & Hari, R. (1998). Parchment-skin illusion: sound-biased touch. *Current Biology*, *8*(6), 190.
- Katuski, Y. (1965). Comparative neurophysiology of hearing. *Physiological Reviews*, *45*, 380–423.
- Kayser, C., Petkov, C. I., Augath, M., & Logothetis, N. K. (2005). Integration of touch and sound in auditory cortex. *Neuron*, *48*(2), 373–84. doi:10.1016/j.neuron.2005.09.018
- Keil, J., Müller, N., Ihssen, N., & Weisz, N. (2012). On the variability of the McGurk effect: audiovisual integration depends on prestimulus brain states. *Cerebral Cortex (New York, N.Y. : 1991)*, *22*(1), 221–31. doi:10.1093/cercor/bhr125
- Kitagawa, N., Zampini, M., & Spence, C. (2005). Audiotactile interactions in near and far space. *Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation Cérébrale*, *166*(3-4), 528–37. doi:10.1007/s00221-005-2393-8

- Kopco, N., Huang, S., Belliveau, J. W., Raij, T., Tengshe, C., & Ahveninen, J. (2012). Neuronal representations of distance in human auditory cortex. *Proceedings of the National Academy of Sciences*, *109*(27), 11019–11024. doi:10.1073/pnas.1119496109
- Làdavas, E. (2002). Functional and dynamic properties of visual peripersonal space. *Trends in Cognitive Sciences*, *6*(1), 17–22.
- Làdavas, E., Pavani, F., & Farne, A. (2001). Auditory Peripersonal Space in Humans : a Case of Auditory – Tactile Extinction. *Neurocase*, *7*, 97–103.
- Lakatos, P., Chen, C.-M., O’Connell, M. N., Mills, A., & Schroeder, C. E. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron*, *53*(2), 279–92. doi:10.1016/j.neuron.2006.12.011
- Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., & Schroeder, C. E. (2005). An Oscillatory Hierarchy Controlling Neuronal Excitability and Stimulus Processing in the Auditory Cortex. *Journal of Neurophysiology*, *94*, 1904–1911. doi:10.1152/jn.00263.2005.
- Lam, K., Kakigi, R., Kaneoke, Y., Naka, D., Maeda, K., & Suzuki, H. (1998). Effects of visual and auditory stimulation on somatosensory evoked magnetic fields, *7*.
- Lange, J., Keil, J., Schnitzler, A., van Dijk, H., & Weisz, N. (2014). The role of alpha oscillations for illusory perception. *Behavioural Brain Research*, *271*, 294–301. doi:10.1016/j.bbr.2014.06.015
- Lange, J., Oostenveld, R., & Fries, P. (2013). Reduced occipital alpha power indexes enhanced excitability rather than improved visual perception. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *33*(7), 3212–20. doi:10.1523/JNEUROSCI.3755-12.2013
- Laurienti, P. J., Burdette, J. H., Wallace, M. T., Yen, Y.-F., Field, A. S., & Stein, B. E. (2002). Deactivation of sensory-specific cortex by cross-modal stimuli. *Journal of Cognitive Neuroscience*, *14*(3), 420–9. doi:10.1162/089892902317361930
- Leo, F., Romei, V., Freeman, E., Ladavas, E., & Driver, J. (2011). Looming sounds enhance orientation sensitivity for visual stimuli on the same side as such sounds. *Experimental Brain Research*, *213*(2-3), 193–201. doi:10.1007/s00221-011-2742-8
- Lewkowicz, D. J. (2008). Perception of Dynamic and Static Audiovisual Sequences in 3- and 4-Month-Old Infants. *Child Development*, *79*(5), 1538–1554.
- Lloyd, D. M. ., Merat, N. ., McGlone, F. ., & Spence, C. (2003). Crossmodal links between audition and touch in covert endogenous spatial attention. *Perception & Psychophysics*, *65*(6), 901–924.
- Lütkenhöner, B., Lammertmann, C., Simões, C., & Hari, R. (2002). Magnetoencephalographic correlates of audiotactile interaction. *NeuroImage*, *15*(3), 509–22. doi:10.1006/nimg.2001.0991
- Maier, J. X., Neuhoff, J. G., Logothetis, N. K., & Ghazanfar, A. A. (2004). Multisensory Integration of Looming Signals by Rhesus Monkeys Max Planck Institute for Biological Cybernetics. *Neuron*, *43*, 177–181.

- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264(5588), 746–748.
- Miller, E. K. (2000). The prefrontal cortex and cognitive control. *Nature Reviews Neuroscience*, 1, 59–65.
- Murray, M. M., Molholm, S., Michel, C. M., Heslenfeld, D. J., Ritter, W., Javitt, D. C., ... Foxe, J. J. (2005). Grabbing your ear: rapid auditory-somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cerebral Cortex (New York, N.Y. : 1991)*, 15(7), 963–74. doi:10.1093/cercor/bhh197
- Neuhoff, J. G. (1998). Perceptual bias for rising tones. *Nature*, 395(6698), 123–4. doi:10.1038/25862
- Neuhoff, J. G. (2001). An Adaptive Bias in the Perception of Looming Auditory Motion. *Ecological Psychology*, 13(2), 87–110. doi:10.1207/S15326969ECO1302_2
- Nicolson, T. (2005). Fishing for key players in mechanotransduction. *Trends in Neurosciences*, 28(3), 140–4. doi:10.1016/j.tins.2004.12.008
- Noesselt, T., Rieger, J. W., Schoenfeld, M. A., Kanowski, M., Hinrichs, H., Heinze, H.-J., & Driver, J. (2007). Audiovisual temporal correspondence modulates human multisensory superior temporal sulcus plus primary sensory cortices. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 27(42), 11431–41. doi:10.1523/JNEUROSCI.2252-07.2007
- Perrault, T. J., Vaughan, J. W., Stein, B. E., & Wallace, M. T. (2005). Superior colliculus neurons use distinct operational modes in the integration of multisensory stimuli. *Journal of Neurophysiology*, 93(5), 2575–86. doi:10.1152/jn.00926.2004
- Pfurtscheller, G., Stancák, a, & Neuper, C. (1996). Event-related synchronization (ERS) in the alpha band--an electrophysiological correlate of cortical idling: a review. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, 24(1-2), 39–46. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8978434>
- Ploner, M., Lee, M. C., Wiech, K., Bingel, U., & Tracey, I. (2010). Prestimulus functional connectivity determines pain perception in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 107(1), 355–60. doi:10.1073/pnas.0906186106
- Rao, S. M., Mayer, A. R., & Harrington, D. L. (2001). The evolution of brain activation. *Nature Neuroscience*, 4(3), 317–323.
- Reed, C. L., Klatzky, R. L., & Halgren, E. (2005). What vs. where in touch: an fMRI study. *NeuroImage*, 25(3), 718–26. doi:10.1016/j.neuroimage.2004.11.044
- Renier, L. a, Anurova, I., De Volder, A. G., Carlson, S., VanMeter, J., & Rauschecker, J. P. (2009). Multisensory integration of sounds and vibrotactile stimuli in processing streams for “what” and “where”. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 29(35), 10950–60. doi:10.1523/JNEUROSCI.0910-09.2009

- Riskind, J. H., Kleiman, E. M., Seifritz, E., & Neuhoff, J. (2014). Influence of anxiety, depression and looming cognitive style on auditory looming perception. *Journal of Anxiety Disorders*, 28(1), 45–50. doi:10.1016/j.janxdis.2013.11.005
- Rizzolatti, G. (1997). The Space Around Us. *Science*, 277(5323), 190–191. doi:10.1126/science.277.5323.190
- 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), 8692–7. doi:10.1523/JNEUROSCI.0160-10.2010
- Romei, V., Murray, M. M., Cappe, C., & Thut, G. (2009). Preperceptual and stimulus-selective enhancement of low-level human visual cortex excitability by sounds. *Current Biology : CB*, 19(21), 1799–805. doi:10.1016/j.cub.2009.09.027
- Ruhnau, P., Hauswald, A., & Weisz, N. (2014). Investigating ongoing brain oscillations and their influence on conscious perception – network states and the window to consciousness. *Frontiers in Psychology*, 5:1230. doi:10.3389/fpsyg.2014.01230
- Schlack, A., Sterbing-D'Angelo, S. J., Hartung, K., Hoffmann, K.-P., & Bremmer, F. (2005). Multisensory space representations in the macaque ventral intraparietal area. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 25(18), 4616–25. doi:10.1523/JNEUROSCI.0455-05.2005
- 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), 187–98. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12063142>
- Schroeder, C. E., Lindsley, R. W., Specht, C., Marcovici, A., Smiley, J. F., & Javitt, D. C. (2001). Somatosensory input to auditory association cortex in the macaque monkey. *Journal of Neurophysiology*, 85, 1322–1327.
- Schürmann, M., Caetano, G., Hlushchuk, Y., Jousmäki, V., & Hari, R. (2006). Touch activates human auditory cortex. *NeuroImage*, 30(4), 1325–31. doi:10.1016/j.neuroimage.2005.11.020
- Seifritz, E., Neuhoff, J. G., Bilecen, D., Scheffler, K., Mustovic, H., Schächinger, H., ... Di Salle, F. (2002). Neural processing of auditory looming in the human brain. *Current Biology : CB*, 12, 2147–51. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12498691>
- Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature*, 408, 788.
- Spence, C., & Driver, J. (2000). Attracting attention to the illusory location of a sound : re-examine crossmodal orienting and ventriloquism. *Neuroreport*, 11(9), 2057–2061.
- Spence, C., & Driver, J. (2004). *Crossmodal space and crossmodal attention*. Oxford: Oxford Univ. Press.

- Spence, C., & Squire, S. (2003). Multisensory Integration: Maintaining the Perception of Synchrony. *Current Biology*, *13*(13), R519–R521. doi:10.1016/S0960-9822(03)00445-7
- Spence, Ch., Baddeley, R., Zampini, M., James, R., & Shore, D. (2003). Multisensory temporal order judgments : When two locations are better than one. *Perception & Psychophysics*, *65*(2), 318–328.
- Sperdin, H. F., Cappe, C., Foxe, J. J., & Murray, M. M. (2009). interactions impact reaction time speed. *Frontiers in Integrative Neuroscience*, *3*(March), 1–10. doi:10.3389/neuro.07
- Stanford, T. R., Quessy, S., & Stein, B. E. (2005). Evaluating the operations underlying multisensory integration in the cat superior colliculus. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *25*(28), 6499–508. doi:10.1523/JNEUROSCI.5095-04.2005
- Stanford, T. R., & Stein, B. E. (2007). Superadditivity in multisensory integration : putting the computation in context. *Neuroreport*, *18*(8), 787–792.
- Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. Cambridge, MA: MIT press.
- Sumbly, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of Acoustical Society of America*, *26*, 212–215.
- Sweet, R. a, Dorph-Petersen, K.-A., & Lewis, D. a. (2005). Mapping auditory core, lateral belt, and parabelt cortices in the human superior temporal gyrus. *The Journal of Comparative Neurology*, *491*(3), 270–89. doi:10.1002/cne.20702
- Tajadura-Jiménez, A., Kitagawa, N., Väljamäe, A., Zampini, M., Murray, M. M., & Spence, C. (2009). Auditory-somatosensory multisensory interactions are spatially modulated by stimulated body surface and acoustic spectra. *Neuropsychologia*, *47*, 195–203. doi:10.1016/j.neuropsychologia.2008.07.025
- Tajadura-jiménez, A., Väljamäe, A. A., & Västfjäll, D. (2010). Embodied Auditory Perception : The Emotional Impact of Approaching and Receding Sound Sources. *Emotion*, *10*(2), 216.
- Teder-Sälejärvi, W. A., McDonald, J. J., Russo, F. Di, & Hillyard, S. A. (2002). An analysis of audio-visual crossmodal integration by means of event-related potential (ERP) recordings. *Cognitive Brain Research*, *14*, 106–114.
- Thut, G., Nietzel, A., Brandt, S. a, & Pascual-Leone, A. (2006). Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *26*(37), 9494–502. doi:10.1523/JNEUROSCI.0875-06.2006
- Tonetto, L. M., Klanovicz Porto, C., & Spence, C. (2014). Modifying action sounds influences people ' s emotional responses and bodily sensations. *I-Perception*, *5*, 153–163.
- Tyll, S., Bonath, B., Schoenfeld, M. A., Heinze, H.-J., Ohl, F. W., & Noesselt, T. (2013). Neural basis of multisensory looming signals. *NeuroImage*, *65*, 13–22. doi:10.1016/j.neuroimage.2012.09.056

- Van Dijk, H., Schoffelen, J.-M., Oostenveld, R., & Jensen, O. (2008). Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 28(8), 1816–23. doi:10.1523/JNEUROSCI.1853-07.2008
- Varela, F., Lachaux, J., Rodriguez, E., & Martinerie, J. (2001). The brainweb: phase synchronization and large-scale integration. *Nature Reviews. Neuroscience*, 2(April), 229–239.
- Vatakis, A., Ghazanfar, A. A., & Spence, C. (2008). Facilitation of multisensory integration by the “unity effect” reveals that speech is special. *Journal of Vision*, 8(14), 1–11. doi:10.1167/8.9.14.Introduction
- Vatakis, A., & Spence, C. (2007). Crossmodal binding: evaluating the “unity assumption” using audiovisual speech stimuli. *Perception & Psychophysics*, 69(5), 744–56. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/17929697>
- Vickaryous, M. K., & Hall, B. K. (2006). Human cell type diversity, evolution, development, and classification with special reference to cells derived from the neural crest. *Biological Reviews of the Cambridge Philosophical Society*, 81(3), 425–55. doi:10.1017/S1464793106007068
- Wallace, M. T., Carriere, B. N., Perrault, T. J., Vaughan, J. W., & Stein, B. E. (2006). The development of cortical multisensory integration. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 26(46), 11844–9. doi:10.1523/JNEUROSCI.3295-06.2006
- Wallace, M. T., & Stein, B. E. (2001). Sensory and Multisensory Responses in the Newborn Monkey Superior Colliculus. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 21(22), 8886–8894.
- Weisz, N., Wühle, A., Monittola, G., Demarchi, G., Frey, J., Popov, T., & Braun, C. (2014). Prestimulus oscillatory power and connectivity patterns predispose conscious somatosensory perception. *Proceedings of the National Academy of Sciences of the United States of America*, 111(4), E417–25. doi:10.1073/pnas.1317267111
- Welch, R. B. (1999). *Meaning, attention, and the “unity assumption” in the intersensory bias of spatial and temporal perceptions.* (Elsevier, Ed.) (Cognitive.).
- Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, 88(3), 638–667.
- Zampini, M., Brown, T., Shore, D. I., Maravita, A., Roeder, B., & Spence, C. (2005). Audiotactile temporal order judgments. *Acta Psychologica*, 118, 277–291. doi:10.1016/j.actpsy.2004.10.017
- Zampini, M., Shore, D. I., & Spence, C. (2003a). Audiovisual temporal order judgments. *Experimental Brain Research*, 152(2), 198–210. doi:10.1007/s00221-003-1536-z
- Zampini, M., Shore, D. I., & Spence, C. (2003b). Multisensory temporal order judgments: the role of hemispheric redundancy. *International Journal of Psychophysiology*, 50(1-2), 165–180. doi:10.1016/S0167-8760(03)00132-6

- Zampini, M., Torresan, D., Spence, C., & Murray, M. M. (2007). Auditory-somatosensory multisensory interactions in front and rear space. *Neuropsychologia*, *45*(8), 1869–77.
doi:10.1016/j.neuropsychologia.2006.12.004
- Zhou, Y.-D., & Fuster, J. M. (2004). Somatosensory cell response to an auditory cue in a haptic memory task. *Behavioural Brain Research*, *153*(2), 573–8.
doi:10.1016/j.bbr.2003.12.024

CHAPTER 4

Papers and contributions

4.1. Contributions

Study I

Leonardelli E., Occelli V., Demarchi G., Grassi M., Braun C., Zampini M.

Auditory stimuli characterized by different saliencies modulate differently the neural processing of tactile stimuli: evidence from MEG recordings

Submitted to Neuroimage

§these authors contributed equally to this work

Leonardelli E. study design, data analysis, manuscript preparation, submission

Occelli V. study design, perform research, support in data analysis, manuscript preparation

Demarchi G. support in perform research

Grassi M. stimuli design

Braun C. study design, support in data analysis, support in manuscript preparation

Zampini M. study design, support in manuscript preparation

Study II

Leonardelli E., Braun C., Weisz N., Lithari C., Occelli V., Zampini M.

Prestimulus oscillatory alpha power and connectivity patterns predispose perceptual integration of an audio and a tactile stimulus.

Under re-review in Human Brain mapping

Leonardelli E. study and stimuli design, perform research, data analysis, manuscript preparation, submission

Braun C. study design, support in data analysis, support in manuscript preparation

Weisz N. study design, support in data analysis, support in manuscript preparation

Lithari C. support in manuscript preparation

Occelli V. support in manuscript preparation

Zampini M. study design, support in manuscript preparation

4.2 Study I

Auditory stimuli characterized by different saliencies modulate differently the neural processing of tactile stimuli: evidence from MEG recordings

Valeria Occelli^{1,2*§}, Elisa Leonardelli^{3*§}, Gianpaolo Demarchi³, Massimo Grassi⁴,
Christoph Braun^{3,5} and Massimiliano Zampini^{2,3}

¹Department of Neurology, Emory University, USA;

²Department of Cognitive and Education Sciences, University of Trento, Italy;

³CIMeC, Center for Mind/Brain Sciences, University of Trento, Italy;

⁴Department of General Psychology, University of Padua, Italy.

⁵Institute of Medical Psychology and Behavioral Neurobiology, University of
Tübingen, Germany;

*CORRESPONDENCE TO: Valeria Occelli, Department of Neurology, Emory
University School of Medicine, Woodruff Memorial Research Building, Suite 6209,
101 Woodruff Circle, Atlanta, GA 30322, USA.

Email:valeria.occelli@gmail.com

OR TO: Elisa Leonardelli, Center for Mind and Brain, University of Trento, Via delle
Regole, 101, Mattarello, 38123 TN, Italy.

Telephone: +39 0461 282753; Fax. +39 0464 808690; Email: elisa.leonardelli@unitn.it

§Equal contribution

SUBMITTED TO: *NeuroImage*

ABSTRACT

In literature, audiotactile (AT) interactions have been found at different locations, latencies and under the form of enhanced or suppressive profiles. The differences between the A and T stimuli involved, especially their relative saliency, are likely to contribute in determining the mixed pattern of found results.. In the present study, we investigate the modulation exerted by three auditory stimuli, characterized by different perceived salience (looming, static-near or static-far sounds, each lasting one second) on the activity evoked by an always identical tactile stimulus that was administered at the end of the sounds. We found that the most salient the preceding sound was, the less activity was elicited by the tactile stimulus. This differential amount of activity associated to an always identical tactile event, was found in the event related fields at early latencies (75-125ms after tactile stimulus onset) encompassing the right hemisphere, mainly the superior temporal sulcus (STS) and the temporal pole (TP). The same differential pattern was found in the time-frequency domain at 100ms after tactile stimulation (6-8 Hz) localized over the right secondary somatosensory cortex (S2) and at 250ms (4-6 Hz), over the right dorsolateral prefrontal cortex (BA 8). In sum, the activity triggered by somatosensory stimulation can be modulated by the saliency of a preceding sound at various stages of cortical processing, from primary cortices at early latencies to higher order areas.

KEYWORDS: auditory, tactile, multisensory, looming, peri-head, magnetoencephalography.

Introduction

A number of studies using various techniques have contributed in delving into the neural substrates subserving audiotactile interactions in both animals (e.g. Fu et al., 2003; Kayser et al., 2005; Lakatos et al., 2007) and humans (e.g. Foxe et al., 2002, 2000; Gobelé et al., 2003; Lütkenhöner et al., 2002; Murray et al., 2005). However, despite the large amount of evidence available at present, the overall pattern which emerges from these studies looks somehow mixed and tentative.

Traditionally, AT interactions have been defined as the difference between the neural activity elicited by the bimodal stimuli and the sum of the neural activities elicited by the unimodal stimuli, i.e. $AT_{interaction} = AT - (A + T)$ ('linear additive model', see Besle et al., 2004 for a review). AT interactions are identified with 'multisensory enhancement' when the bimodal neural activity exceeds the sum of the neural activity elicited by its unisensory components when presented individually ($AT > A + T$), while they are defined as suppressive in the opposite case ($AT < A + T$). In the literature, AT interactions have been found to result not only into enhanced profiles, but also, although more sparsely, into suppressive ones. For example, Lütkenhöner et al. (2002) found suppressive AT interactions, interpreted as partial inhibition of secondary somatosensory cortex (S2), whereas Gobelé and colleagues (2003) showed a profile of AT that was more similar to T than to A, suggesting a suppression of auditory responses. Conversely, Foxe and colleagues (2000) and Murray and colleagues (2005) found enhanced activity, that they localized as originating from somatosensory and auditory cortices. One major candidate responsible for such strikingly discrepant patterns is the wide range of different types of stimuli used among these studies (see Gobelé et al., 2003 for a similar discussion and Lakatos et al. 2007 for evidence on monkeys). The participants in Lütkenhöner et al.'s study (2002) (binaural tone burst, 100 ms, 60 dB SL, 1047 Hz delivered through plastic tubes fitted into the ear canals) were perceived as "more intense" than T stimuli (balloon diaphragm driven by bursts of compressed air on the right thumb; p. 518). Those in Gobelé and colleagues's study (2003) claimed that the tactile inputs (the electrical stimulation of the median nerve) "dominated during AT stimulation" (p. 510) (auditory stimulus is a monaural 50-ms 1-kHz square-wave sounds, presented at about 70 dB SPL and delivered to the subject by plastic tubes). Thus, when applying the linear additive model though the single component of each stimulus is identical in the bimodal as in the unimodal stimulation, i.e. from both sides of the equation $AT = A + T$, the relative saliency between stimuli changes.

In the present study, the modulation played by three auditory stimuli differing in salience on the activity evoked by subsequent identical somatosensory stimuli will be assessed. To create sounds with different salience, the intensity level of the sounds was manipulated in order to induce the impression of static sounds being delivered near, far from the head or a sound moving towards the head (looming). Although crude (Kopco et al., 2012), the amplitude manipulation to induce the impression of sounds

sources located at different distances has been demonstrated to effectively convey the perception of near vs. far sounds (see Ahveninen et al., 2014; Moore and King, 1999 for reviews). Along the same line, the increase in the intensity of a sound or in the size of a visual input can reliably be interpreted as an approaching object (Ghazanfar et al., 2002; Riskind et al., 2014; Seifritz et al., 2002). From the literature, it is known that being exposed to certain types of sounds reflect into specific behavioural patterns. For instance, louder sounds are more arousing than softer sounds (Tonetto et al., 2014) and have been demonstrated to act as salient exogenous cues, facilitating subsequent voluntary shift of the attentional focus (Alho et al., 2014). Looming sounds are rated as more unpleasant, arousing, intense, salient and threatening than receding sounds (Bach et al., 2009; Tajadura-Jiménez et al., 2010) and typically induces attention orienting or motor acts (Bach et al., 2009; Cappe and Barone, 2005; Kearsley, 1973; Lloyd et al., 2003). Moreover, they are perceptually more salient than receding sound. For example, looming sounds are longer (Grassi and Darwin, 2006; Grassi and Pavan, 2012; Grassi, 2010; Schlauch et al., 2001) louder (Stecker and Hafter, 2000), change more in loudness (Neuhoff, 2001, 1998; Olsen and Stevens, 2010; Olsen et al., 2010) than do receding sounds.

Finally, studies on humans have shown that effects are stronger for stimuli delivered to the space around the head (peri-head space) than when the same stimuli are delivered to the hands (Kitagawa et al., 2005; Farnè and Làdavas, 2002; Tajadura-Jimenez et al. 2009). Therefore, in order to increase the possible different impact of the various auditory stimuli on the tactile one, the tactile stimuli were presented on the head.

Thus, in the present study the interplay between looming vs. static far and near sounds and somatosensory inputs (i.e., air puffs) presented within the peri-head space are assessed. MEG will be used to characterize the different neuronal activity responses. If the manipulation of the sound is effective, it would be expected to observe a differential modulation of the oscillatory activity as a function of the type of sounds delivered. As additional measure of the efficacy of the manipulation we performed on sounds stimuli in order to obtain a different salience, behavioral ratings about various features of the sounds (i.e., pleasantness, remoteness, motion, arousal, power, distance, duration, ecological validity, threat) will be collected.

Materials and methods

Participants

Fourteen healthy volunteers (5 female, mean age 30 years, range 21-53 years; all right-handed) participated in the study. Participants provided written, informed consent to the experimental session. The study had prior approval by the Ethical Committee of the University of Trento, and an informed consent was obtained from

each subject prior the beginning of the experimental session. The participants received a reimbursement in return for taking part in the study.

Stimuli and experimental instrumentation

The sound used in the experiment was a mosquito sound (22050 Hz sample rate, 16 bits resolution). This kind of sound was chosen for different reasons. First, it has ecological validity especially when paired with a tactile stimulation in form of an air puff that can constitute a violation of the peripersonal zone. Moreover, the experience of being bitten by a mosquito is thought to constitute an unpleasant event, such to induce in the participants the sensation of a threat for the body. Lastly, the sound of a mosquito is composed of different sound frequencies, which is the typology of sound which has been proved to be integrated with tactile stimulation at a higher extent than pure tones (e.g., Farnè and Làdavas, 2002; Graziano et al., 1999). The original sample of the auditory stimulus can be found in the following webpage: <http://www.acoustics.org/press/132nd/3pab10.html>. The sample (1.4 sec) was edited in duration and amplitude. The duration was truncated to 1 sec. The amplitude was customly attenuated (or amplified) in various parts of the recording in order to obtain a sound as much as possible constant in root mean square amplitude over time. The resulting sound was presented either at a level of 40 dB SPL (far static sound), at a level of 75 dB SPL (near static sound) or further modulated in amplitude with an exponential envelope increasing of 35-dB (from 40 to 75 dB SPL) over the sound's duration (looming sound). This latter stimulus simulates the proximal stimulation pattern produced by the motion of an object emitting a steady-amplitude sound and moving at a constant speed toward the perceiver. An object as such generates a change in intensity at the listener's ear that is linear on a logarithmic scale. Note that a sound source moving as such does not produce any audible frequency shift at the perceiver's ear (Neuhoff & McBeath, 1997).

The auditory stimuli presented during the behavioural part were presented via headphones (Logitech USB Headset H330). During the MEG measurement, inside the shielded room, the stimuli were delivered to the participants earplugs via MEG-compatible pneumatic tubal insert foam plugs (Tip-300, Nicolet, Madison, WI, USA). The somatosensory stimulation was delivered via an air nozzle located ~5 cm from the face of the participant. The nozzle was fixed to a wooden MEG-compatible jib and set in such a way that it was pointing at the participant's left perioral area. The tactile stimulation was delivered via an inhouse built pneumatic stimulator, with computer controlled valves. The air stream had a pressure of 0.07 bar.

Procedure

At the beginning of the experimental session, the participants were asked to rate the three auditory stimuli (i.e., looming, static near and static far sounds) according to different dimensions by using 7-point Likert scales (modified from Bach et al., 2009).

The evaluated dimensions were: pleasantness (unpleasant – pleasant), remoteness (very close - very far), perception of motion (static – high motion), arousal (calming – activating), power (weak – strong), distance estimation (0-10 cm – 60-70 cm), perception of moving away (static – highly receding), perception of getting closer (static – highly approaching), duration (short – long). In addition, the participants were asked to evaluate how much the sounds heard resembled to an everyday life sounds (ecology) and how likely was that they could be followed by a threatening event (threat). The participants could listen to them as many times as they considered necessary in order to give an accurate rating. MEG recordings were performed in a magnetically shielded room. For the MEG measurement, they were asked to wear earplugs and sit upright with their heads comfortably resting within the mould of the helmet. The participants were instructed to look straight ahead, trying to fixate a fixation cross in the centre of the screen, to avoid any movements during the measurement, and to ignore the auditory and somatosensory stimuli. In order to keep the participants alert throughout the session, the colour of the fixation cross was programmed to randomly change from white to red during the time interval right before the start of the sound and the participants were asked to report at the end of each run the number of times the change occurred. Each session included fourteen runs. Eleven trials for each experimental condition were presented in each run in a random order. Participants were presented with the six following experimental conditions. Three experimental conditions were unimodal auditory: looming sound (condition *Aloom*), static near sound (condition *Anear*), static far sound (condition *Afar*). Three experimental conditions were bimodal auditory-somatosensory stimuli: each of the three types of sound was followed, after sound offset, by a tactile stimulus (conditions *Afar* → *T*, *Anear* → *T*, *Aloom* → *T*). The somatosensory stimulation was programmed such that the 23 ms delay due to the transmission of the air through the pipe was compensated and resulted as presented right at the end of the sound (see Figure 1).

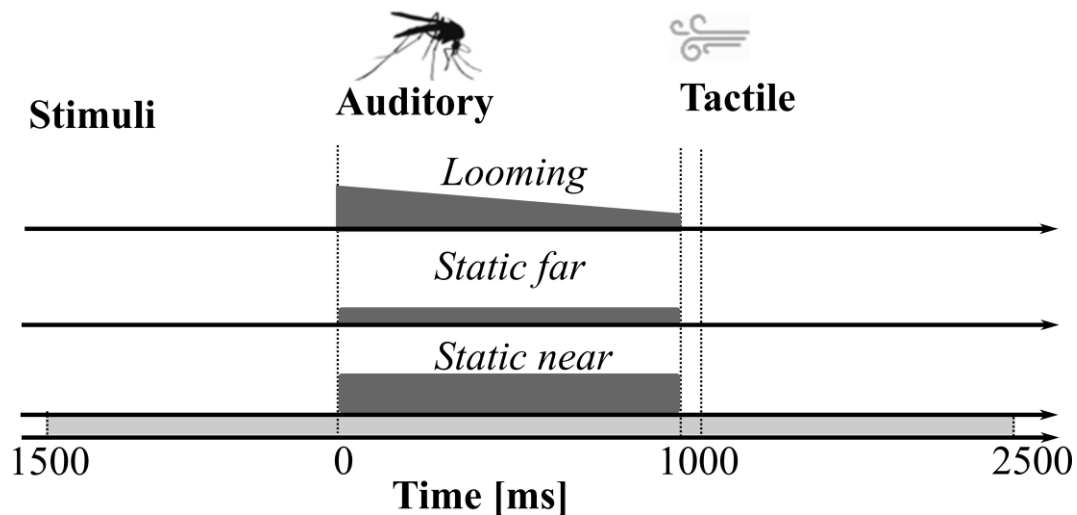


Figure 1

Schematic representation of the experimental design. In light gray, the time window considered in the data analysis.

Data Acquisition

The magnetic brain activity was recorded with a whole head Elekta Neuromag Vectorview® MEG system, with 306 channels (204 first order planar gradiometers, 102 magnetometers) in a magnetically shielded room (AK3B, Vakuum Schmelze, Hanau, Germany). Raw magnetic data were hardware band-pass filtered in the frequency range 0.1-330 Hz and subsequently sampled at a rate of 1kHz. Before the MEG measurements, five head position indicator (HPI) coils were placed on the head surface. The position of the coils and the participants' head shapes were digitized with a Fastrak 3D digitizer (Polhemus, Colchester, VT, USA, www.polhemus.com). In order to control for head movements, participants' head positions with respect to the MEG sensors were estimated at the beginning of each run. Stimulus presentation was controlled using E-prime software package (PST, Inc., Pittsburgh, PA), running on a Windows computer, and presented by a projector. MEG recordings were stored for off-line analysis.

Data analysis

The data from one participant were discarded due to excessive noise. The MEG data of the remaining participants (4 female, mean age 30 years, range 21-53 years) were analyzed offline using Fieldtrip (Oostenveld et al., 2011), an open source toolbox for Matlab (www.mathworks.com).

Data preprocessing

Epochs of four seconds were centered at the auditory stimulus onset, 1500 ms pre sound-onset and 2500 ms post-sound onset (i.e., 1500 seconds post-sound offset). Zero seconds therefore represents the onset of the auditory stimulus while one second represents the onset of the tactile stimulus. The recorded epochs were high-pass filtered at 1 Hz to remove very slow frequencies (DC offset) and then down-sampled to 400 Hz. Epochs were visually inspected for possible artifacts and contaminated trials were excluded from further processing.

Definition of the conditions of interest

Main focus of this study is to assess the modulation of the somatosensory activity as a function of the typology of the sound preceding the somatosensory stimulation, i.e. to compare the somatosensory activity elicited in three bimodal conditions: $A_{near} \rightarrow T$, $A_{loom} \rightarrow T$, $A_{far} \rightarrow T$. It is well known that the offset of an auditory sound produces an activity (e.g. Hari et al., 1987). In order to cancel out this auditory effect, we defined the three different types of touch as:

$$\begin{aligned} T/far &= (A_{far} \rightarrow T) - A_{far} \\ T/near &= (A_{near} \rightarrow T) - A_{near} \\ T/loom &= (A_{loom} \rightarrow T) - A_{loom} \end{aligned}$$

Event related fields (ERFs) calculation in sensors space

For the three bimodal conditions, $A_{near} \rightarrow T$, $A_{loom} \rightarrow T$, $A_{far} \rightarrow T$, and the three unimodal conditions, A_{near} , A_{far} , A_{loom} , single trials were low-pass filtered using a 30-Hz filter prior to averaging. We averaged the data over trials to obtain the ERF waveforms. In the analysis of the ERFs at sensor level, we considered only the combined gradiometers: horizontal and vertical planar gradients of the magnetic field were combined separately at each sensor and the root of the squared sum of both directions was computed to obtain a single positive-valued number representing the averaged ERF at each sensor. The conditions of interest $T/near$, $T/loom$, and T/far , were calculated for each participant by subtracting unimodal conditions A_{near} , A_{loom} , A_{far} , from the respective compound activities $A_{near} \rightarrow T$, $A_{loom} \rightarrow T$, $A_{far} \rightarrow T$.

Time-frequency (TF) calculation in sensors space

For each participant, time-frequency analysis was performed on single trials between 1-30 Hz for the three bimodal conditions, $A_{near} \rightarrow T$, $A_{loom} \rightarrow T$, $A_{far} \rightarrow T$, and the three unimodal conditions A_{near} , A_{loom} , A_{far} . We applied an adaptive sliding time window with a length (Δt) of 5 cycles of the respective frequency and shifted in steps of 50 ms between -0.5 to 1.5s. A Hanning taper was applied yielding a spectral smoothing of $1/\Delta t$. For time-frequency analysis at sensor level, we considered the combined gradiometers: orthogonal gradients of the magnetic field were calculated separately at each sensor. The root of the squared sum of both directions was computed

to obtain the power at each sensor irrespective of the orientation of the gradients (Medendorp et al., 2007). The TF representations of the T /near, T /loom, and T /far conditions were calculated for each participant by subtracting from the TFs of the three conditions $A_{near} \rightarrow T$, $A_{loom} \rightarrow T$, $A_{far} \rightarrow T$, their respective unimodal conditions A_{near} , A_{loom} , A_{far} (see above and section *Definition of the conditions of interest*).

Statistics for the event related fields and the time-frequency analysis in sensors space

For both time and frequency domain analysis, the T /near, T /loom, and T /far conditions were compared through a non-parametric cluster-based dependent-samples F-test with Monte-Carlo randomization (Maris and Oostenveld, 2007) performed on the time window encompassing the somatosensory stimulation, i.e. 1 s to 1.500 s. A post-hoc analysis was conducted for each couple of conditions by using a paired t test Bonferroni corrected. In the time domain the paired t test was performed on the individual mean amplitude of the T /near, T /loom, and T /far of the significant cluster, while in the time-frequency domain the individual mean power for each significant clusters.

Projection of the data in source space

For ten participants, the head shape obtained with Polhemus 3d-digitizer prior to the MEG measurement was co-registered with the individual structural MRI (4T Bruker MedSpec, Siemens). This procedure consists of a first course alignment between the three fiducials digitized with the Pholemus, and the corresponding points on the individual MRIs. Finally, the head shape points were fitted, through a rigid body transformation, to the individual head surface as extracted from the structural MRI. For three participants the individual MRIs were not available, thus their head shapes points obtained with Polhemus were fitted to a Montreal Neurological Institute (MNI) brain. Subsequently, a 3D grid covering the entire brain volume (1 cm spacing, 2982 points inside the brain) fitted to a template MNI brain was created. In order to compare the source activity across participants on a group level, for each participant the constructed grid was warped into individual headspace. In this manner, the grid points were not regularly spaced but they were located in the same area with respect to the MNI template across participants (Larson-Prior et al., 2013). Based on the implementation of Nolte (2003), an analytical single shell model was fitted to the individual segmented MRI and the lead field was calculated for each grid point. Both, magnetometers and gradiometers were taken into account after appropriate adjustment of the balancing matrix based on the distance of the gradiometers (17 mm). In order to project the data into source space, i.e. on the points of the grid, a linearly constrained minimum variance (LCMV; Van Veen et al., 1997) spatial filter was used: first a LCMV filter was individually estimated using for each condition for the data on the time interval -1 s to +2 s; then, the time-series of each trial was multiplied with this common filter.

Localization of the task-dependent ERF modulation in source space

In order to identify the sources for the task dependent modulation of ERF on the sensor level the activity of bimodal conditions $A_{far} \rightarrow T$, $A_{near} \rightarrow T$, $A_{loom} \rightarrow T$ and unimodal conditions A_{far} , A_{near} , A_{loom} , were projected into source space using an LCMV Beamformer after preprocessing (see previous section). Then, the source activity for each voxel was low-pass filtered at 30 Hz and time-lock averaged with the same parameters as had been used for the ERF analysis. As before, by subtracting from the activity of the time-lock averaged bimodal conditions the activity of the time-lock averaged unimodal conditions, we obtained an estimate for the source activation T/far , $T/near$ and $T/loom$ (in source space). For identifying the origin of the significant difference found previously at sensor level, we searched for the same task-dependent parametric modulation of source data in the same time window as defined by the sensor level analysis. Using an F-test, the areas were identified by using the ERF-amplitudes T/far , $T/near$, and $T/loom$ as regressors: (i.e., resulting coefficients: 0.7, -.256, -0.444). F-values were plotted on an MNI brain. It is important to stress however, that statistical validation to ERF analysis is given by the statistics conducted at sensor level. At source level, the multiple comparisons problem renders statistic difficult and a rigorous pipeline of how conducting statistic is still lacking. For this reason no statistical comparison of source activation was performed, instead, only effects found on the sensor level were identified on the level of sources.

Localization of task-dependent TF modulation in source space

For bimodal conditions $A_{far} \rightarrow T$, $A_{near} \rightarrow T$, $A_{loom} \rightarrow T$ and unimodal conditions A_{far} , A_{near} , A_{loom} we applied the same parameters utilized in the TF at the sensor level to the preprocessed data projected to the source space (see previous paragraph). Source activity for each condition – T/far , $T/near$ and $T/loom$ – was then individually calculated and clusters of voxels adjacent in space and time revealing the parametric modulation observed in sensor space were identified. F-values were then plotted on an MNI brain. As for the analysis in time, the projection of TF to source space is meant to localize the origin of the statistical difference found at sensor level. Because of the multicomparison problem no additional hypothesis testing was done on the source level.

Results

The focus of the present study is the effect of different types of sound located or moving in space on a subsequent tactile stimulus. To this aim, we investigated a 500 ms time-window encompassing the activity generated by a tactile stimulus, always identical, that nevertheless could be preceded by a one second sound, either far static (T/far), near static ($T/near$) or looming ($T/loom$).

Behavioural data.

Each sound was rated by 12 participants before the session according to various dimensions (see Section of Methods). For each rated dimension, an ANOVA between the three typologies of sound was performed and, when significant, post-hoc paired comparisons were calculated. Whereas the three sounds were not evaluated as differing for valence and duration, they were perceived as significantly differing for other features: distance (i.e., looming sound ~20 cm distant, static far ~50 cm distant and the static near ~30 cm distant); arousal (looming significantly more arousing than static far); power (static far being significantly weaker than both static near and looming sounds); motion (looming sound being perceived as more endowed with motion than static far and, though not significantly, than static near); approaching (looming sound being perceived as more highly approaching than the two static sounds); ecological valence (looming sound being perceived as more ecological than the static near sound); thread valence (static near sound being perceived as more threatening than the static far) (see Table 1 for results). Taken together, these data show that looming sounds were perceived as being more salient than the static far sounds. Given that static near sounds were rated as significantly different from the looming sounds in some scales, and from static far sounds in other scales, it is reasonable to assume that they were perceived in the middle on a possible saliency continuum.

Table 1. Explicit subjective ratings.

Abbreviations: L = Looming; SN = Static Near; SF = Static Far. * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$.

<i>Dimension</i>	<i>Mean±SE subjective ratings</i>		<i>ANOVA</i>	<i>Post-hoc paired comparisons</i>
Valence (unpleasant – pleasant)	L	4.17±.64	p=.14	---
	SN	3.58±.58		
	SF	3.08±.29		
Remoteness (near – far)	L	3.36±.61	p=.003**	L vs. SN p=1.00
	SN	2.75±.43		L vs. SF p=.09
	SF	5.25±.49		SN vs. SF p=.002**
Distance estimation (0-10 cm – 60-70)	L	2.17±.21	p<.001***	L vs. SN p=.27
	SN	3.17±.52		L vs. SF p=.001***

cm)	SF	4.92±.57		SN vs. SF	p=.005**
Arousal (calming – activating)	L	6.00±.25	p=.004**	L vs. SN	p=1.00
	SN	5.75±.31		L vs. SF	p=.03*
	SF	4.75±.35		SN vs. SF	p=.06
Power (weak – strong)	L	4.83±.30	p<.001***	L vs. SN	p=.73
	SN	5.25±.37		L vs. SF	p<.001***
	SF	2.00±.30		SN vs. SF	p<.001***
Motion (static – high motion)	L	6.42±.19	p<.001***	L vs. SN	p=.07
	SN	4.92±.51		L vs. SF	p=.001***
	SF	3.75±.51		SN vs. SF	p=.16
Approaching (static – highly approaching)	L	6.67±.14	p<.001***	L vs. SN	p=.009**
	SN	3.83±.76		L vs. SF	p<.001
	SF	2.58±.47		SN vs. SF	p=.29
Duration (short – long)	L	3.42±.43	p=.03*	L vs. SN	p=.16
	SN	3.83±.42		L vs. SF	p=.63
	SF	3.00±.33		SN vs. SF	p=.08
Ecology (artificial – natural)	L	5.58±.45	p=.013*	L vs. SN	p=.03*
	SN	4.33±.56		L vs. SF	p=.72
	SF	5.17±.49		SN vs. SF	p=.22
Thread (low – high)	L	5.00±.39	p=.04*	L vs. SN	p=1.00
	SN	5.25±.35		L vs. SF	p=.29
	SF	4.17±.39		SN vs. SF	p=.04*

Event-Related Activity

In order to evaluate the differences in activation between *T/loom*, *T/near* and *T/far*, we compared their somatosensory event-related activities. We found a significant difference between the three conditions, $p = .001$, in the time window 1075–1125 ms, i.e. between 75 and 125 ms after tactile stimulus onset (Figure 2A left). The topography of the statistical values in the significant time window shows that the difference is localized in the hemisphere contralateral to the tactile stimulation and spreads along the entire hemisphere, peaking over fronto-temporal sensors (Figure 2A right).

Interestingly, the pattern of amplitude is that T_{loom} elicits the smallest amplitude ERF, while T_{near} intermediate amplitude, while T_{far} evokes the biggest activity. In order to better evaluate these differences, figure 2B shows the individual mean average of the significant sensors, within the statistically significant time-window. A post-hoc analysis (paired t-test, Bonferroni corrected) revealed a statistical difference between T_{loom} and T_{near} ($p = .02$), between T_{far} and T_{near} ($p = .003$) and between T_{far} and T_{loom} ($p = .0002$) (Figure 2B). Main generators are located in the right temporal cortex (Brodmann Area (BA) 22-STS and BA 21-TP) (Figure 2C). However, estimation of the sources shows activity on the right hemisphere that encompasses the frontal (BA 10), the parietal (BA 7, BA 9, BA 2) and the temporal lobe (BA 22, BA 21).

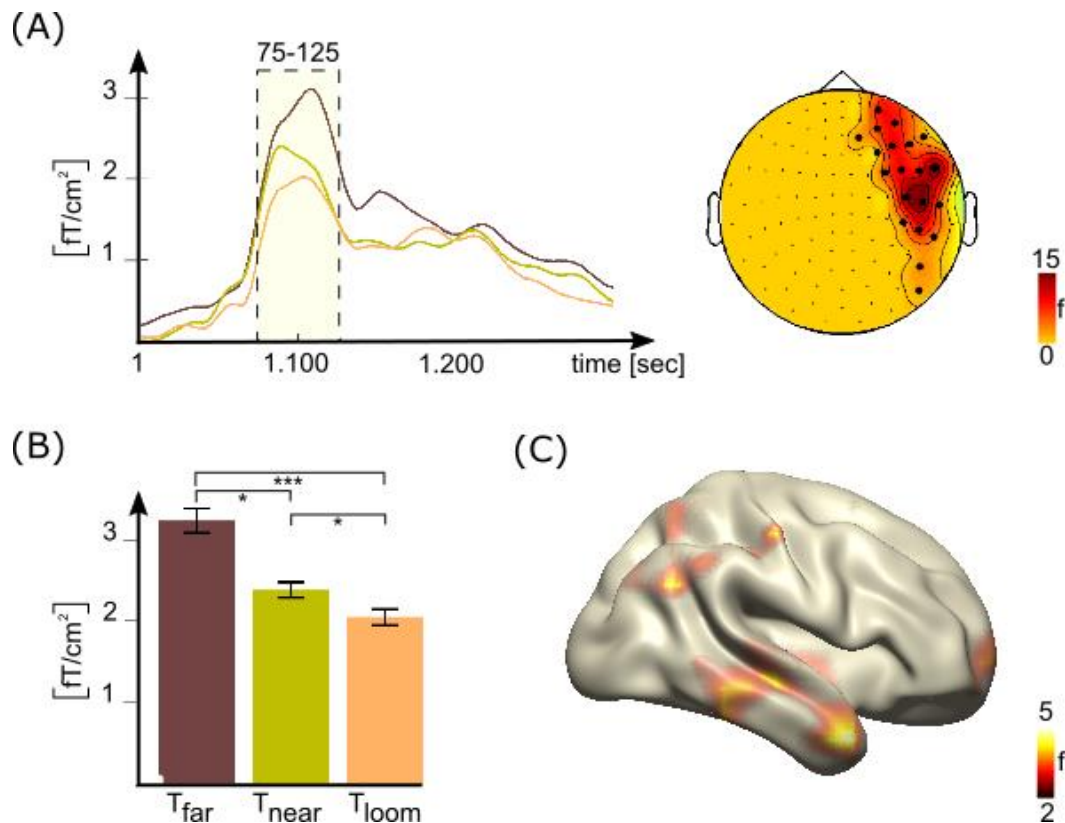


Figure 2. Summarized results of the time-course analysis.

(A) Right: event-related field trace of the positive sensor cluster for the T_{far} (brown), T_{near} (light brown), T_{loom} (orange). The highlighted rectangle shows the

statistically significant time-window (1.075-1.125 sec, 1sec is the auditory offset and tactile onset)

Left: Topography (f values) of the positive sensor cluster (1.087-1.125 ms), masked for statistical significance. The significant channels. The dots indicate the channels of the significant cluster (1075-1130 ms).

(B) the bar plots represent the mean average for the sensors and time-window of the significant cluster. Error bars show within-subjects 95% confidence intervals. In order to reveal the origin of the statistical differences revealed with the f test, post-hoc analysis was conducted (paired, Bonferroni corrected, t -test). $*p \leq 0.05$, $**p \leq 0.01$, $***p \leq 0.001$.

(C) Figure 2D shows the projected f -values obtained from statistical comparison in sensor space between conditions T_{near} , T_{loom} , T_{far} . Like in the topography shown in Figure 2A-right, differences are spread in the right hemisphere encompassing frontal, parietal and temporal lobes. The main generators are located in the temporal lobe (TP and STS).

Time-frequency results

We analyzed the oscillatory power pattern within a 500 ms interval encompassing the tactile event. Utilizing a cluster based permutation statistic, we identified two clusters of power that significantly differed between the three conditions. The first cluster ($p < .03$, 1000–1270 ms, 6-8 Hz, max at 1100 ms) was localized on the right part of the topography (contralateral to the tactile stimulus). A second cluster ($p < .02$, 1030–1370 ms, 4-6 Hz, peaking at 1250 ms) was localized centrally towards the frontal part of the topography (Figures 3A and 3B). Figure 3C shows the mean power of activity averaged across the sensors, and the time-frequency windows forming significant clusters. Mean power is depicted for all three conditions and for all clusters identified. For both clusters, the T_{loom} (in orange) elicits the smallest increase in power, followed by T_{near} (in green) and by T_{far} (in brown). Post-hoc analysis (paired t -test Bonferroni corrected) was performed. For the first cluster, a statistical difference was revealed between T_{far} and T_{loom} ($p = .0005$), while between T_{loom} and T_{near} and between T_{far} and T_{near} no statistical difference could be proven ($p = .10$ and $p = .14$, respectively). For the second cluster, a statistical difference was revealed between T_{loom} and T_{near} ($p = .0008$) and between T_{far} and T_{loom} ($p = .001$), while between T_{far} and T_{near} no statistical difference could be proven ($p = .17$). Figure 3D shows the projected f -values obtained from statistical comparison in source space between conditions T_{near} , T_{loom} , T_{far} . Each cluster was projected into source space centered at the point of its maximum value in sensor level (first cluster 1100 ms, 6-8 Hz; second cluster 1250 ms, 4-6Hz) and then two clusters were plotted on the same brain. By comparing the topography at sensor level and the projection of the cluster, the principal area originating in the first cluster (1250 ms, 4-6Hz) can be identified as

the right secondary somatosensory cortex (BA40 spreading towards BA2). The second cluster originates in the dorsolateral prefrontal cortex (BA 8).

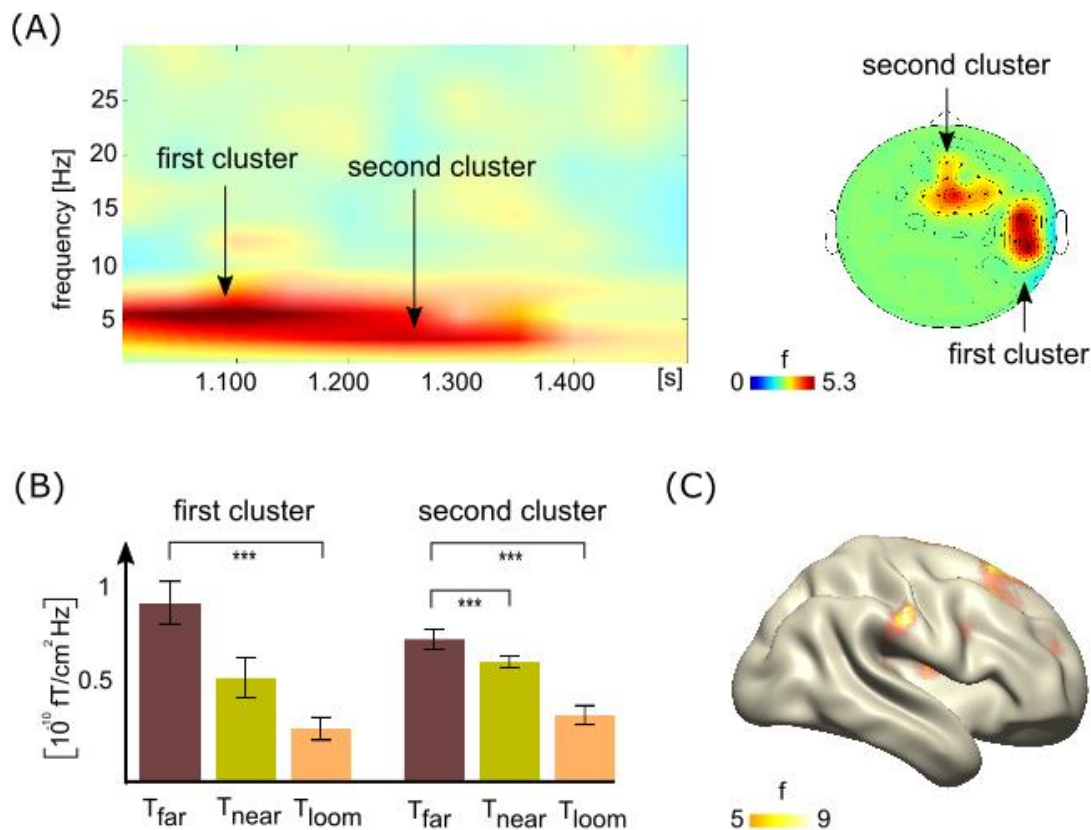


Figure 3. Summarized results of the time-frequency analysis

(A) Left: time-frequency representation for the time window 1000-1500 ms (1 is the auditory offset and tactile onset)). Shown are the f -values for comparison of conditions the T_{near} , T_{loom} , T_{far} of the combined gradiometers. The arrows indicates roughly the two significant clusters identified. First cluster: $p < .03$, 1000–1270 ms, 6-8 Hz, peaking at 1100 ms. Second cluster: $p < .02$, 1030–1370 ms, 4-6 Hz, peaking at 1250 ms. Non-significant values have been reduced to 40% opacity.

Right: Topography of the significant clusters. The first cluster is localized centrally towards the front, whereas the second is right-lateralized.

(B) The bar plots show the means and the standard errors of the mean of the power for the sensors, time-window and frequency of the two significant clusters. Error bars show within-subjects 95% confidence intervals. $*p \leq 0.05$, $**p \leq 0.01$, $***p \leq 0.001$.

0.001.

(C)Projection in source space of the statistical comparison (f-stat) across the three conditions for time window 1000–1370 ms in the frequency 4 to 8 Hz. The areas corresponding to the first cluster (1250 ms, 4-6Hz) can be identified with the right secondary somatosensory cortex (BA40 spreading into BA2). The second cluster originate in correspondence with the right dorsolateral prefrontal cortex (BA 8).

Discussion

In the present study, we have characterized how the neuronal activity elicited by a somatosensory stimulation is modulated by different types of preceding auditory stimuli. In particular, we created three sounds with different saliency. Our behavioral data confirm that the looming sounds were perceived as being more salient than the static far sounds, while the static near sounds were perceived somewhat in the middle. The three different sounds exert a different cross-modal influence on a subsequent tactile stimulus that, on the contrary, was kept constant among conditions. The observation we report is a change in the magnitude of the elicited somatosensory responses, in the time-locked as well as in the time-frequency results: the tactile stimulus responses following the looming sound elicited a smaller activity than the ones following the two static conditions and, among the latter, the smallest response was observed in the static near condition. The computation of the differential magnitude of responses across the three conditions allowed us to define the spread of the exerted cross-modal influence across time and space. The first difference, as revealed by the ERF analysis, is at early latency (from 75 ms to 125 ms after tactile stimulus onset) and encompassed the contralateral hemisphere (right), peaking over STS. Later, as highlighted by the time frequency domain, a difference is localized over the right S2 (peaking at 100 ms; 6-8 Hz) and, on a later stage, over the contralateral frontal areas (peaking at 250 ms; 4-6 Hz).

Decrease in the magnitude of tactile activity correlates with the salience of auditory stimulus.

Our results demonstrate the importance for multisensory interactions of the relative saliency between stimuli. In particular, by changing the saliency of a (preceding) auditory stimulus while keeping a (subsequent) tactile stimulus always identical, we show how the relative saliency between the stimuli impact the magnitude of the activity evoked by the tactile stimulus.

A first consideration is that this factor (the relative saliency) affects the interaction calculated with the linear additive model ($AT \text{ interaction} = AT - (A+T)$, see Besle et al. 2004 for a review of the model). In fact, when applying the model, the “absolute saliency” of each of the stimuli is taken into account by summing and subtracting the same stimuli when presented in combination (AT) and when presented singularly (A and T). However in the bimodal condition the factor “relative saliency”

between the stimuli is present while in the unimodal conditions is not, thus rendering the calculated AT interaction sensitive to the relationship between the particular type of stimuli used (for example, interaction between a strong electrical pulse and a weak “bip” will be different than interaction between a soft touch and a loud binaural noise). Differences in relative saliencies between stimuli is therefore likely to constitute one of the main reasons of the discrepant pattern of results found in the literature of AT interactions (see Introduction for a review), since these studies were conducting by using very different stimuli. Moreover, we believe that the perceived saliency, rather than acoustic intensity cues *per se*, impact the found interaction since the rapidly approaching sound is followed by the smallest tactile activity, whereas the near stimulus – that overall possesses the highest intensity but relative less saliency, given that is static – is followed by an intermediate response and the far sound – static and low-intensity – by the biggest response. Our participants' self-reports support the assumption that the three typologies of sounds were perceived differently in regard to some critical features: according to our predictions, the static far sound was perceived significantly weaker and less activating than the other two sounds (see Table 1). This finding is consistent with previous evidence, showing that looming sounds are associated with a higher warning significance (Bach et al., 2009; Klimesch et al., 1998; Neuhoff, 1998), and possibly with a stronger attentional orienting bias toward these stimuli compared to static ones (Balconi et al., 2009).

Additionally, we show an inverse relationship between the perceived saliency of the acoustic stimulus and the strength of this modulation, with more salient auditory stimuli being associated to less tactile activity. This pattern of results is well in accordance to the law of inverse effectiveness (Stein and Meredith, 1993) which has long been considered one of the basic principles of the multisensory field and which states that, the strength of multisensory interactions increases when the stimuli efficacy (usually measured in terms of saliency) of the isolated components are relatively weak by themselves (see also Kayser et al., 2005; Lakatos et al., 2007; Perrault et al., 2005; Stanford et al., 2005).

Noteworthy, in order to avoid differential manipulation of attention or expectancy between conditions, no tasks involving the auditory or the somatosensory stimulation was implied, and the auditory stimuli had an equal probability to be or not to be followed by a somatosensory stimulus. Thus, any potential discrepancy in the neural activation across the sounds can be genuinely attributed to the features of the auditory stimulation delivered. However, an alternative explanation to our findings is possible when recurring to the crossmodal shared attention mechanism. This mechanism assumes attention to be a limited resource and stimulus representations from different modalities are in competition with one another based on their representational strength (saliency): the more salient a stimulus representation, the more it will dominate the competition (e.g. Spence and Driver, 2004). The different types of sounds here used are endowed with different levels of saliency, so differences

in attentional bias cannot be excluded. However it is worth noting that in our paradigm the sounds were already off in time window on which the analysis was centered and that potential sound offset effects were ruled out (see Methods section).

Spatiotemporal profile of the modulation of tactile activity: from early STS and temporal pole (TP) to S2 to prefrontal areas

Besides the aspects discussed in previous section, this differential magnitude of responses across the three conditions allowed us to track the spread of the exerted cross-modal influence across time and space. Our results show modulations of amplitude in the activity of areas within or in close proximity to auditory and somatosensory cortices and a subsequent spread of this modulation towards more frontal regions.

The earliest significant difference between tactile activities appears at early latency (75 ms after tactile stimulus onset) and is in the activity evoked (stimulus-locked analysis). This first differential contribution to the evoked activity come from regions that encompasses a major part of contralateral hemisphere (with respect to tactile stimulation), peaking over the right temporal lobe (STS and TP). TPs are known to be involved in category-specific conceptual knowledge processing (Hanley, 2014; Noppeney and Price, 2002). Moreover, it has been found to be sensitive to complex auditory features, such as changes in the acoustic scale (von Kriegstein et al., 2007) or spectral composition (Kaiser et al., 2002) in a variety of sound categories, including animal sounds. Furthermore, a recent study investigating the large-scale functional connectivity pattern of the TP has shed further light on the functions of this area (Pascual et al., 2013). Interestingly, the TP shows connectivity with somatosensory primary cortex, in particular with its ventral portion, which subserves mouth and tongue movement, thus suggesting its role in the integration of auditory and somatosensory information corresponding to the mouth. The TP also shows connections with the orbitofrontal cortex and neighboring frontal regions as well as with subcortical structures related to higher order sensorimotor processing. The authors conclude that its connectivity pattern suggests that the TP might be involved in the integration of sensorimotor and auditory information and/or use this information to categorize sounds and build concepts (Pascual et al., 2013). Regarding STS region, this is located in proximity to the primary auditory cortices and known to be a multisensory area, responding to stimuli in vision, hearing and touch (Beauchamp, 2005; Beauchamp et al., 2008, 2004; Musacchia and Schroeder, 2009; Schürmann et al., 2006) (Schroeder and Foxe, 2005;). Beauchamp and coworkers demonstrated a significant enhancement of the BOLD response for simultaneous somatosensory and auditory stimulation as compared to unisensory stimulation (Beauchamp et al., 2008; see also Kayser et al., 2005; Schroeder et al., 2001; Foxe et al., 2002; Murray et al., 2005; for evidence on monkeys). Moreover, STS has been suggested to be the human homologue of the caudomedial (CM) belt area of primates' auditory association cortex (Schroeder and

Foxe, 2002) which consistently shows auditory and somatosensory stimuli convergence with a bias for the skin surfaces of the head and neck (Fu et al., 2003; Kayser et al., 2005; Schroeder and Foxe, 2002; Schroeder et al., 2001), similarly to the stimuli used here. The activity within STS has been found to be highly biased for looming versus receding auditory signals (Bach et al., 2009; Seifritz et al., 2002). Interestingly, when distance is conveyed, as in the present study, by amplitude changes, a right enhanced response over the right STS has been reported (Mathiak et al., 2003). STS also responds to audiovisual motion signals (Baumann and Greenlee, 2007; Lewis and Noppeney, 2010; Maier and Ghazanfar, 2007; Sadaghiani et al., 2009; Seifritz et al., 2002; Tyll et al., 2013; Werner and Noppeney, 2011), possibly suggesting its role in the selective enhancement of a class of salient stimuli – those rising in intensity – that could constitute a potentially threatening change of the environment (Tyll et al., 2013).

The assessment of the oscillatory activity induced by the tactile event indicated that the power frequency difference across the three conditions was right-lateralized and localizable in correspondence with two clusters. The earlier cluster (6-8Hz) is localized over secondary somatosensory cortex (S2). The peak of activity within this cluster observed in correspondence with S2 at 100 ms after stimulus onset is in agreement with the idea that S2 is the second stage of tactile processing – on which the analysis was centered – and with previous evidence on ERF latency timing of airpuff stimulation of peri-oral area (Nguyen et al., 2005). This evidence adds to the robust demonstration of S2 in selective somatosensory attention (Fujiwara et al., 2002; Johansen-berg and Lloyd, 2000; Meftah et al., 2009). However, S2 has also been proved to have multisensory properties, as shown by studies demonstrating its responsivity to sounds (Beauchamp and Ro, 2008). Along the same line is the evidence showing mutual connections between S2 and auditory cortices (Ro et al., 2013) and AT multisensory interactions within this area (Gobbelé et al., 2003; Lütkenhöner et al., 2002). Moreover, S2 is considered as providing a link between sensory inputs and motor areas (Lin and Forss, 2002).

It is worth noting, however, that the interpretation in terms of suppressive or enhanced effects of auditory and S2 cortices is somehow deceptive, given the spatial arrangement of these areas. Indeed, auditory and S2 cortices have opposite location in the Sylvian fissure and originate signals that exhibit opposite deflections and may cancel out, thus hampering a clear interpretation of whether the signal changes are due to suppression of one, facilitation of the other or a combination of both mechanisms (see also Gobbelé et al. 2003 on this point). The fact that, as demonstrated by the post-hoc analysis (see Results section), the only significant comparison across the three conditions is between the looming and the static far conditions is consistent with the conjecture that this cluster is more sensitive to distance cues than to sound motion properties. The self-reports of our participants support the assumption that these two typologies of sounds were perceived as significantly differing in distance: whereas the

looming sound was estimated as being ~20 cm distant, the static far sound was estimated as being ~30 cm distant (see Table 1).

Finally, differences spread towards more frontal areas, i.e. condition is the right dorsolateral prefrontal cortex (BA 8). Strongly connected with the motor system, this area has been traditionally mainly considered as a motor-related region involved in planning motor behaviors (Abe and Hanakawa, 2009; Chouinard and Paus, 2006; Schluter et al., 1999). Recently, the involvement of the dorsal prefrontal cortex has been suggested for cognitive tasks as well (Hanakawa et al., 2002), such as early filtering of irrelevant sensory information in order to accomplish specific goals in the most efficient way (Abe and Hanakawa, 2009; Bolton and Staines, 2011). The evidence that within this cluster a significant difference was observed between the looming sound and static near sound conditions contributes to rule out the hypothesis that the intensity and the perceived closeness of the auditory stimulation delivered is not sufficient *per se* to induce a differing pattern of results, but points to the functional specificity of motion. In support of this conjecture are the self-reports of the participants, who estimated the looming sound as being endowed with a higher degree of motion than the two static conditions (see Table 1). Moreover, the significant differences observed with the static far sound condition and the other two conditions indicates that the necessity of producing an effective motor response in conditions of threat or danger is of particular relevance in presence of aversive objects occurring within the peripersonal space.

Conclusions

In this study we have provided evidence that the activity triggered by somatosensory stimulation of the perioral area can be crossmodally modulated by a preceding sound. We manipulated the saliency of the preceding sound and showed how the magnitude of the tactile activity inversely correlated with this feature. Namely, following looming sounds (the most salient sound) less tactile activity was elicited as compared to the two static conditions, and among the two static conditions the near static sounds (more salient than the far static sound but less salient than the looming sounds) were followed by less activity than the far-static sound. This inverse relation between amplitude and saliency is well in accordance with the “inverse effectiveness rule”, though we cannot exclude an involvement of cross-modal attention mechanism. Importantly, our results show the effects of stimuli relative saliences on interactions between senses and underscore the need to appropriately considering this factor when designing and discussing experiments of multisensory nature. Interestingly, following this modulatory pattern we could show influences of the sounds on the activity elicited by a tactile event as early as 75 ms (e.g. Foxe et al., 2002, Gobelé et al., 2003) in regions encompassing the anterior temporal lobe. At later latencies this modulation spread towards secondary somatosensory and frontal regions. The computation of the spatiotemporal features of the interplay between sounds and somatosensory stimuli

allowed us to highlight influences at several levels of cortical processing (Ghazanfar and Schroeder, 2006) thus supporting the idea that multisensory interactions are present at various stages of cortical processing, from primary cortices at early latencies to higher order areas.

Future directions

Since audiotactile interactions have been shown to be more pronounced in the space behind the head (e.g., Farné and Làdavas, 2002), different portions of the peri-head space could be tested by presenting auditory stimuli from the front and behind the head. Another line of future research pertains the potential influence of visual experience on the development of the mechanism devoted to the coding of audiotactile interactions within the peri-head space. In particular, it would be interesting to assess whether the absence of visual experience induces the establishment of a higher sensitive system for the coding of audiotactile interplay in the peri-head space, and whether this is spatially-modulated.

ACKNOWLEDGEMENTS

We wish to thank Nathan Weisz, Christos Papadelis and Luigi Tamé for their invaluable suggestions and Massimo Vescovi, Alessandra Calliari and Gianpiero Monittola for their technical support.

REFERENCES

- Abe, M., Hanakawa, T., 2009. Functional coupling underlying motor and cognitive functions of the dorsal premotor cortex. *Behav. Brain Res.* 198, 13–23. doi:10.1016/j.bbr.2008.10.046
- Ahveninen, J., Kopco, N., Jääskeläinen, I., 2014. Psychophysics and neuronal bases of sound localization in humans. *Hear. Res.* 307, 86–97. doi:10.1016/j.heares.2013.07.008.
- Alho, K., Salmi, J., Koistinen, S., Salonen, O., Rinne, T., 2014. Top-down controlled and bottom-up triggered orienting of auditory attention to pitch activate overlapping brain networks. *Brain Res.* 1–10. doi:10.1016/j.brainres.2014.12.050
- Bach, D.R., Neuhoff, J.G., Perrig, W., Seifritz, E., 2009. Looming sounds as warning signals: the function of motion cues. *Int. J. Psychophysiol.* 74, 28–33. doi:10.1016/j.ijpsycho.2009.06.004
- Balconi, M., Brambilla, E., Falbo, L., 2009. Appetitive vs. defensive responses to emotional cues. Autonomic measures and brain oscillation modulation. *Brain Res.* 1296, 72–84. doi:10.1016/j.brainres.2009.08.056
- Baumann, O., Greenlee, M.W., 2007. Neural correlates of coherent audiovisual motion perception. *Cereb. Cortex* 17, 1433–43. doi:10.1093/cercor/bhl055
- Beauchamp, M.S., 2005. See me, hear me, touch me: multisensory integration in lateral occipital-temporal cortex. *Curr. Opin. Neurobiol.* 15, 145–53. doi:10.1016/j.conb.2005.03.011
- Beauchamp, M.S., Argall, B.D., Bodurka, J., Duyn, J.H., Martin, A., 2004. Unraveling multisensory integration: patchy organization within human STS multisensory cortex. *Nat. Neurosci.* 7, 1190–2. doi:10.1038/nn1333

- Beauchamp, M.S., Ro, T., 2008. Neural substrates of sound-touch synesthesia after a thalamic lesion. *J. Neurosci.* 28, 13696–702. doi:10.1523/JNEUROSCI.3872-08.2008
- Beauchamp, M.S., Yasar, N.E., Frye, R.E., Ro, T., 2008. Touch, sound and vision in human superior temporal sulcus. *Neuroimage* 41, 1011–1020.
- Besle, J., Fort, A., Delpuech, C., Giard, M.-H., 2004. Bimodal speech: early suppressive visual effects in human auditory cortex. *Eur. J. Neurosci.* 20, 2225–34. doi:10.1111/j.1460-9568.2004.03670.x
- Bolton, D. a E., Staines, W.R., 2011. Transient inhibition of the dorsolateral prefrontal cortex disrupts attention-based modulation of tactile stimuli at early stages of somatosensory processing. *Neuropsychologia* 49, 1928–37. doi:10.1016/j.neuropsychologia.2011.03.020
- Cappe, C., Barone, P., 2005. Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. *Eur. J. Neurosci.* 22, 2886–902. doi:10.1111/j.1460-9568.2005.04462.x
- Chouinard, P. a, Paus, T., 2006. The primary motor and premotor areas of the human cerebral cortex. *Neuroscientist* 12, 143–52. doi:10.1177/1073858405284255
- Cooke, D.F., Graziano, M.S. a, 2004. Sensorimotor integration in the precentral gyrus: polysensory neurons and defensive movements. *J. Neurophysiol.* 91, 1648–60. doi:10.1152/jn.00955.2003
- Duhamel, J., Colby, C.L., Goldberg, M.E., 1998. Ventral Intraparietal Area of the Macaque: Congruent Visual and Somatic Response Properties. *J. Neurophysiol.* 79, 126–136.
- Foxe, J.J., Morocz, I.A., Murray, M.M., Higgins, B.A., Javitt, D.C., Schroeder, C.E., 2000. Multisensory auditory-somatosensory interactions in early cortical processing revealed by high-density electrical mapping. *Brain Res. Cogn. Brain Res.* 10, 77–83.
- Foxe, J.J., Wylie, G.R., Martinez, A., Schroeder, C.E., Daniel, C., Guilfoyle, D., Ritter, W., Murray, M.M., Butler, J.S., Molholm, S., Fiebelkorn, I.C., Mercier, M.R., Schwartz, T.H., 2002. Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. *J. Neurophysiol.* 88, 540–543.
- Fu, K.-M.G., Johnston, T. a, Shah, A.S., Arnold, L., Smiley, J., Hackett, T. a, Garraghty, P.E., Schroeder, C.E., 2003. Auditory cortical neurons respond to somatosensory stimulation. *J. Neurosci.* 23, 7510–5.
- Fujiwara, N., Imai, M., Nagamine, T., Mima, T., Oga, T., Takeshita, K., Keiichiro, T., Shibasaki, H., 2002. Second somatosensory area (SII) plays a significant role in selective somatosensory attention. *Cogn. Brain Res.* 14, 389–397.
- Ghazanfar, A. a, Neuohoff, J.G., Logothetis, N.K., 2002. Auditory looming perception in rhesus monkeys. *Proc. Natl. Acad. Sci. U. S. A.* 99, 15755–7. doi:10.1073/pnas.242469699
- Ghazanfar, A. a, Schroeder, C.E., 2006. Is neocortex essentially multisensory? *Trends Cogn. Sci.* 10, 278–85. doi:10.1016/j.tics.2006.04.008
- Gobbelé, R., Schürmann, M., Forss, N., Juottonen, K., Buchner, H., Hari, R., 2003. Activation of the human posterior parietal and temporoparietal cortices during audiotactile interaction. *Neuroimage* 20, 503–511. doi:10.1016/S1053-8119(03)00312-4
- Grassi, M., 2010. Sex difference in subjective duration of looming and receding sounds. *Perception* 39, 1424–1426. doi:10.1068/p6810
- Grassi, M., Darwin, C.J., 2006. The subjective duration of ramped and damped sounds. *Percept. Psychophys.* 98, 1382–1392.
- Grassi, M., Pavan, A., 2012. The subjective duration of audiovisual looming and receding stimuli. *Atten. Percept. Psychophys.* 74, 1321–33. doi:10.3758/s13414-012-0324-x
- Graziano, M.S., 2001. A system of multimodal areas in the primate brain. *Neuron* 29, 4–6.
- Graziano, M.S. a, Cooke, D.F., 2006. Parieto-frontal interactions, personal space, and defensive behavior. *Neuropsychologia* 44, 845–59. doi:10.1016/j.neuropsychologia.2005.09.009
- Graziano, M.S., Reiss, L. a, Gross, C.G., 1999. A neuronal representation of the location of nearby sounds. *Nature* 397, 428–30. doi:10.1038/17115

- Hanakawa, T., Honda, M., Sawamoto, N., Okada, T., Yonekura, Y., Fukuyama, H., Shibasaki, H., 2002. The Role of Rostral Brodmann Area 6 in Mental-operation Tasks : an Integrative Neuroimaging Approach. *Cereb. cortex* 12, 1157–1170.
- Hari, R., Pelizzone, M., Mäkelä, J.P., Hällstroem, J., Leinonen, L., Lounasmaa, O. V., 1987. Neuromagnetic responses of the human auditory cortex to on- and offsets of noise bursts. *Audiology* 26, 31–43.
- Johansen-berg, H., Lloyd, D.M., 2000. The physiology and psychology of selective attention to touch. *Front. Biosci.* 5, 894–904.
- Kaiser, J., Lutzenberger, W., Ackermann, H., Birbaumer, N., 2002. Dynamics of Gamma-band Activity Induced by Auditory Pattern Changes in Humans 212–221.
- Kaysers, C., Petkov, C.I., Augath, M., Logothetis, N.K., 2005. Integration of touch and sound in auditory cortex. *Neuron* 48, 373–84. doi:10.1016/j.neuron.2005.09.018
- Kearsley, R.B., 1973. The Newborn 's response to Auditory Stimulation :A Demonstration of Orienting and Defensive Behavior. *Child Dev.* 44, 582–590.
- Klimesch, W., Doppelmayr, M., Russegger, H., Pachinger, T., Schwaiger, J., 1998. Induced alpha band power changes in the human EEG and attention. *Neurosci. Lett.* 244, 73–76.
- Kopco, N., Huang, S., Belliveau, J.W., Raj, T., Tengshe, C., Ahveninen, J., 2012. Neuronal representations of distance in human auditory cortex. *Proc. Natl. Acad. Sci.* 109, 11019–11024. doi:10.1073/pnas.1119496109
- Lakatos, P., Chen, C.-M., O'Connell, M.N., Mills, A., Schroeder, C.E., 2007. Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron* 53, 279–92. doi:10.1016/j.neuron.2006.12.011
- Larson-Prior, L.J., Oostenveld, R., Della Penna, S., Michalareas, G., Prior, F., Babajani-Feremi, a, Schoffelen, J.-M., Marzetti, L., de Pasquale, F., Di Pompeo, F., Stout, J., Woolrich, M., Luo, Q., Bucholz, R., Fries, P., Pizzella, V., Romani, G.L., Corbetta, M., Snyder, a Z., 2013. Adding dynamics to the Human Connectome Project with MEG. *Neuroimage* 80, 190–201. doi:10.1016/j.neuroimage.2013.05.056
- Lewis, R., Noppeney, U., 2010. Audiovisual synchrony improves motion discrimination via enhanced connectivity between early visual and auditory areas. *J. Neurosci.* 30, 12329–39. doi:10.1523/JNEUROSCI.5745-09.2010
- Lin, Y.Y., Forss, N., 2002. Functional characterization of human second somatosensory cortex by magnetoencephalography. *Behav. Brain Res.* 135, 141–145.
- Lloyd, D.M., Merat, N., McGlone, F., Spence, C., 2003. Crossmodal links between audition and touch in covert endogenous spatial attention. *Percept. Psychophys.* 65, 901–924.
- Lütkenhöner, B., Lammertmann, C., Simões, C., Hari, R., 2002. Magnetoencephalographic correlates of audiotactile interaction. *Neuroimage* 15, 509–22. doi:10.1006/ning.2001.0991
- Maier, J.X., Ghazanfar, A. a, 2007. Looming biases in monkey auditory cortex. *J. Neurosci.* 27, 4093–100. doi:10.1523/JNEUROSCI.0330-07.2007
- Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* 164, 177–90. doi:10.1016/j.jneumeth.2007.03.024
- Mathiak, K., Hertrich, I., Kincses, W., Riecker, A., Lutzenberger, W., Ackermann, H., 2003. The right supratemporal plane hears the distance of objects : neuromagnetic correlates of virtual reality. *Neuroreport* 14, 2003.
- Medendorp, W.P., Kramer, G.F.I., Jensen, O., Oostenveld, R., Schoffelen, J.-M., Fries, P., 2007. Oscillatory activity in human parietal and occipital cortex shows hemispheric lateralization and memory effects in a delayed double-step saccade task. *Cereb. cortex* 17, 2364–74. doi:10.1093/cercor/bhl145
- Meftah, E.-M., Bourgeon, S., Chapman, C.E., 2009. Instructed delay discharge in primary and secondary somatosensory cortex within the context of a selective attention task. *J. Neurophysiol.* 101, 2649–67. doi:10.1152/jn.91121.2008

- Moore, D.R., King, A.J., 1999. Auditory perception : The near and far of sound localization. *Curr. Biol.* 9, 361–363.
- Murray, M.M., Molholm, S., Michel, C.M., Heslenfeld, D.J., Ritter, W., Javitt, D.C., Schroeder, C.E., Foxe, J.J., 2005a. Grabbing your ear: rapid auditory- somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cereb. Cortex* 15, 963–974. doi:10.1093/cercor/bhh197
- Murray, M.M., Molholm, S., Michel, C.M., Heslenfeld, D.J., Ritter, W., Javitt, D.C., Schroeder, C.E., Foxe, J.J., 2005b. Grabbing your ear: rapid auditory-somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cereb. Cortex* 15, 963–74. doi:10.1093/cercor/bhh197
- Musacchia, G., Schroeder, C.E., 2009. Neuronal mechanisms, response dynamics and perceptual functions of multisensory interactions in auditory cortex. *Hear. Res.* 258, 72–9. doi:10.1016/j.heares.2009.06.018
- Neuhoff, J.G., 1998. Perceptual bias for rising tones. *Nature* 395, 123–4. doi:10.1038/25862
- Neuhoff, J.G., 2001. An Adaptive Bias in the Perception of Looming Auditory Motion. *Ecol. Psychol.* 13, 87–110. doi:10.1207/S15326969ECO1302_2
- Neuhoff, J.G., McBeath, M.K., 1997. Overcoming naive mental models in explaining the Doppler Shift: An illusion creates confusion. *Am. J. Phys.* 65, 618–621.
- Nguyen, B.T., Inui, K., Hoshiyama, M., Nakata, H., Kakigi, R., 2005. Face representation in the human secondary somatosensory cortex. *Clin. Neurophysiol.* 116, 1247–53. doi:10.1016/j.clinph.2005.01.018
- Nolte, G., 2003. The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors. *Phys. Med. Biol.* 48, 3637–52.
- Noppeney, U., Price, C.J., 2002. Retrieval of visual, auditory, and abstract semantics. *Neuroimage* 15, 917–26. doi:10.1006/nimg.2001.1016
- Olsen, K.N., Stevens, C.J., 2010. Perceptual overestimation of rising intensity: Is stimulus continuity necessary? *Perception* 39, 695–704. doi:10.1068/p6592
- Olsen, K.N., Stevens, C.J., Tardieu, J., 2010. Loudness change in response to dynamic acoustic intensity. *J. Exp. Psychol. Hum. Percept. Perform.* 36, 1631–44. doi:10.1037/a0018389
- Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.-M., 2011. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* 2011, 156869. doi:10.1155/2011/156869
- Pascual, B., Masdeu, J.C., Hollenbeck, M., Makris, N., Insausti, R., Ding, S.-L., Dickerson, B.C., 2013. Large-Scale Brain Networks of the Human Left Temporal Pole: A Functional Connectivity MRI Study. *Cereb. Cortex*. doi:10.1093/cercor/bht260
- Perrault, T.J., Vaughan, J.W., Stein, B.E., Wallace, M.T., 2005. Superior colliculus neurons use distinct operational modes in the integration of multisensory stimuli. *J. Neurophysiol.* 93, 2575–86. doi:10.1152/jn.00926.2004
- Riskind, J.H., Kleiman, E.M., Seifritz, E., Neuhoff, J., 2014. Influence of anxiety, depression and looming cognitive style on auditory looming perception. *J. Anxiety Disord.* 28, 45–50. doi:10.1016/j.janxdis.2013.11.005
- Sadaghiani, S., Maier, J.X., Noppeney, U., 2009. Natural, metaphoric, and linguistic auditory direction signals have distinct influences on visual motion processing. *J. Neurosci.* 29, 6490–9. doi:10.1523/JNEUROSCI.5437-08.2009
- Schlauch, R.S., Ries, D.T., DiGiovanni, J.J., 2001. Duration discrimination and subjective duration for ramped and damped sounds. *J. Acoust. Soc. Am.* 109, 2880. doi:10.1121/1.1372913
- Schluter, N.D., Rushworth, M.F.S., Mills, K.R., Passingham, R.E., 1999. Signal-, set-, and movement-related activity in the human premotor cortex. *Neuropsychologia* 37, 233–243.

- Schroeder, C.E., Foxe, J.J., 2002. The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. *Brain Res. Cogn. Brain Res.* 14, 187–98.
- Schroeder, C.E., Lindsley, R.W., Specht, C., Marcovici, A., Smiley, J.F., and Javitt, D.C., 2001. Somatosensory input to auditory association cortex in the macaque monkey. *J. Neurophysiol.* 85, 1322–1327.
- Schürmann, M., Caetano, G., Hlushchuk, Y., Jousmäki, V., Hari, R., 2006. Touch activates human auditory cortex. *Neuroimage* 30, 1325–31. doi:10.1016/j.neuroimage.2005.11.020
- Seifritz, E., Neuhoff, J.G., Bilecen, D., Scheffler, K., Mustovic, H., Schächinger, H., Elefante, R., Di Salle, F., 2002. Neural processing of auditory looming in the human brain. *Curr. Biol.* 12, 2147–51.
- Spence, C., Driver, J., 2004. *Crossmodal space and crossmodal attention.* Oxford Univ. Press, Oxford.
- Stanford, T.R., Quessy, S., Stein, B.E., 2005. Evaluating the operations underlying multisensory integration in the cat superior colliculus. *J. Neurosci.* 25, 6499–508. doi:10.1523/JNEUROSCI.5095-04.2005
- Stecker, G.C., Hafter, E.R., 2000. An effect of temporal asymmetry on loudness 107, 3358–3368.
- Stein, B.E., Meredith, M.A., 1993. *The merging of the senses.* MIT press, Cambridge, MA.
- Tajadura-jiménez, A., Väljamäe, A.A., Västfjäll, D., 2010. Embodied Auditory Perception : The Emotional Impact of Approaching and Receding Sound Sources. *Emotion* 10, 216.
- Tonetto, L.M., Klanovicz Porto, C., Spence, C., 2014. Modifying action sounds influences people ' s emotional responses and bodily sensations. *Iperception.* 5, 153–163.
- Tyll, S., Bonath, B., Schoenfeld, M.A., Heinze, H.-J., Ohl, F.W., Noesselt, T., 2013. Neural basis of multisensory looming signals. *Neuroimage* 65, 13–22. doi:10.1016/j.neuroimage.2012.09.056
- Van Veen, B.D., van Drongelen, W., Yuchtman, M., Suzuki, a, 1997. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Trans. Biomed. Eng.* 44, 867–80. doi:10.1109/10.623056
- Von Kriegstein, K., Smith, D.R.R., Patterson, R.D., Ives, D.T., Griffiths, T.D., 2007. Neural representation of auditory size in the human voice and in sounds from other resonant sources. *Curr. Biol.* 17, 1123–8. doi:10.1016/j.cub.2007.05.061

4.3 Study II

Prestimulus oscillatory alpha power and connectivity patterns predispose perceptual integration of an audio and a tactile stimulus

Elisa Leonardelli¹, Christoph Braun^{1,2}, Nathan Weisz¹, Chrysa Lithari¹, Valeria
Occelli³, Massimiliano Zampini¹

¹ Center for Mind/Brain Sciences, University of Trento, Trento, Italy,

² MEG Center, University of Tübingen, Tübingen, Germany

³ Department of Neurology, Emory University, Atlanta, Georgia, USA

CORRESPONDENCE TO:

Elisa Leonardelli
Center for Mind and Brain, University of Trento
Via delle Regole, 101
Mattarello, TN, Italy 38123
Tel. +39 0461 282753
Fax. +39 0464 808690
elisa.leonardelli@unitn.it

KEYWORDS: audiotactile, multisensory interactions, magnetoencephalography (MEG), functional connectivity, prestimulus

ABSTRACT

To efficiently perceive and respond to the external environment, our brain has to perceptually integrate or segregate stimuli of different modalities. The temporal relationship between the different sensory modalities is therefore essential for the formation of different multisensory percepts. In this magnetoencephalography (MEG) study, we created a paradigm where an audio and a tactile stimulus were presented by an ambiguous temporal relationship so that perception of physically identical audiotactile stimuli could vary between integrated (emanating from the same source) and segregated. This bistable paradigm allowed us to compare identical bimodal stimuli that elicited different percepts, providing a possibility to directly infer multisensory interaction effects.

Local differences in alpha power over bilateral inferior parietal lobules (IPLs) and superior parietal lobules (SPLs) preceded integrated versus segregated percepts of the two stimuli (audio and tactile). Furthermore, differences in long-range cortical functional connectivity seeded in rIPL (region of maximum difference) revealed differential patterns that predisposed integrated or segregated percepts encompassing secondary areas of all different modalities and prefrontal cortex. We showed that the prestimulus brain states predispose the perception of the audiotactile stimulus both in a global and a local manner. Our findings are in line with a recent consistent body of findings on the importance of prestimulus brain states for perception of an upcoming stimulus. This new perspective on how stimuli originating from different modalities are integrated suggests a non-modality specific network predisposing multisensory perception.

Introduction

Inputs through multiple senses are continuously processed in our brain, resulting in a unique and coherent perception of the world. Using data recorded from single neurons in the cat's superior colliculus, Stein and Meredith [1993] were the first to propose specific principles of multisensory interactions. They showed that, rather than being independent, processing of stimuli of different modalities depends on their relative spatiotemporal properties and salience. However, the mechanism of crossmodal interactions in the human brain remains largely unknown. Specifically, audio-tactile (AT) interactions have been far less investigated than those of other sensory modalities (e.g. audio-visual or visuo-tactile), despite their frequent occurrence in everyday life (knocking at a door, playing instruments, walking, shaving, perceiving the buzzing and the sting of a mosquito).

To date, neurons responsive to both audio and tactile stimuli have been found in primates in the ventral premotor cortex, the ventral intraparietal region and the superior temporal sulcus [e.g. Bruce et al. 1981; Graziano et al. 1999; Schlack et al. 2005]. Further evidence suggests that AT interactions may occur already at early latencies of the sensory processing in brain regions close to or within primary sensory areas. More specifically, somatosensory stimulation was found to evoke rapid activations in belt areas of the auditory cortex [e.g. Schroeder and Foxe 2002; Fu et al. 2003] and tactile stimuli modulated the activity of primary auditory cortex [Lakatos et al., 2007].

In human neuroimaging studies, AT multisensory interactions have been identified in secondary somatosensory cortex, the auditory belt area, and the posterior parietal cortex also before 100 ms [Beauchamp et al., 2008; Foxe et al., 2000; Foxe et al., 2002; Gobbelé et al., 2003; Hofer et al., 2013; Kassuba et al., 2013; Lütkenhöner et al., 2002; Murray et al., 2005; Schürmann et al., 2006].

Traditionally, crossmodal interactions have been estimated through the so-called "linear additive model". This model defines multisensory interaction as the difference between the neural activity generated by the bimodal stimuli and the linear sum of the unimodal neural responses (e.g. AT interaction = AT-[A+T]). This model has been inspired and supported by observations of non-linear superposition of single neuron measurements in the cat's superior colliculus [e.g. Stein and Meredith 1993]. The model is sensitive to super/sub-additivity of neural responses, i.e. neural responses elicited by the stimuli presented in combination that exceed/diminish the sum of the activity elicited by the stimuli when these are presented singularly. However, it fails when it comes to detection of linear interactions, which have been as well extensively reported [e.g. Stanford et al., 2005] and is sensitive to multiple confounds, when not applied appropriately [Besle et al., 2004].

In the present study, we aimed to elucidate the mechanism of perceptual integration/segregation of AT stimuli. It is well known that to create a coherent and correct representation of the external world, stimuli emanating from the same object or event are perceptually integrated, while stimuli independently generated are

perceptually segregated [Ghazanfar and Schroeder 2006]. To investigate this process, we created a bistable paradigm that allowed direct comparison of identical bimodal stimuli that however could elicit different (integrated or segregated) percepts. A sound of one second, simulating a looming mosquito approaching the head, was utilized as the auditory stimulus. A short touch, delivered on the face and resembling the mosquito landing, was the tactile stimulus. As the timing between the different modalities is critical for perceptually merging two stimuli [Calvert et al., 2004; Welch, 1999], we manipulated the stimulus onset asynchronies (SOAs) between the stimuli, rendering their temporal relation “ambiguous”: i.e. the perception of the physically identical AT stimuli could vary from integrated (condition INT) (audio and tactile stimuli perceived as the approaching mosquito and its landing) to segregated (condition SEG) (audio and tactile perceived as independent events) from trial to trial. These stimuli were selected because of their ecological relevance, and because approaching sounds [Canzoneri et al., 2012] and the space around the head [e.g. Kitagawa et al. 2005] have been shown to evoke stronger AT interactions.

Since with our paradigm we exclude the fact that integrated and segregated percepts are due to physical differences between stimuli, we hypothesize that different brain-states prior to the upcoming tactile stimulus predispose different multisensory percepts. Indeed, an increasing body of evidence is showing an influence of prestimulus activity on subsequent perception, especially in the case of near-threshold stimuli [Ruhnau et al., 2014] and in multisensory perception [Hipp et al., 2011; Keil et al., 2012]. We focused our analysis on the prestimulus time-window and characterized the multimodal AT interaction by local measures of power and global measures of inter-areal coupling [Weisz et al., 2014].

Materials and methods

Participants

Sixteen participants (5 females, mean age \pm s.d., 26 ± 7 years, all right-handed) were recruited from a participant database of the Center for Mind/Brain Sciences at the University of Trento. Visual acuity was normal or corrected-to-normal and all participants reported a normal sense of hearing and touch. The study was approved by the Ethical Committee of the University of Trento and was conducted in accordance with the Declaration of Helsinki as revised in October 2008. All participants gave written informed consent prior to participation and received monetary compensation for their time. The approximate duration of the study including preparation was 90 minutes.

Stimuli and experimental instrumentation

The auditory stimuli consisted of a sound simulating a flying mosquito (44100 Hz sample rate, 16 bits resolution). The original stimulus length (see <http://www.acoustics.org/press/132nd/3pab10.html>) was truncated to one second and

its intensity was modulated with an exponential envelope increasing of 35 dB (from 25 to 60 dB) over the duration of the sound to create the impression of a looming sound approaching the participant's head. Stimuli were presented binaurally via tubal insert headphones compatible with MEG. The sound was filtered with the head related transfer function (HRTF) of a KEMAR dummy head. The function was selected from the CIPIC HRTF Database [Algazi et al., 2001, dataset freely downloadable at <http://interface.cipic.ucdavis.edu/sound/hrtf.html>]. The simulated sound source appeared in the horizontal plane, at an azimuth angle of 30° left of the direction of the participant's gaze and was in accordance with the spatial position of the tactile stimulus. The tactile stimulation consisted of a light touch (diameter = 3 mm) at the participant's left perioral area. The tactile stimulation was delivered via a custom-built apparatus. A plastic filament was attached to a piston that could be moved back and forth by a double-acting pneumatic cylinder that was driven by compressed air with a pressure of 0.7 bars. The system was controlled electrically by a valve via a custom-built relay box. The final part of the stimulation equipment was mounted on a wooden MEG-compatible pedestal (that had no direct contact with the participant). The pneumatic system was placed outside the magnetically shielded room, so the noise produced by the device was not audible inside. The plastic filament was the terminal part of a fiber optic system (Keyence series FS-N, Neu-Isenburg, Germany). One fiber of the system was connected to an infra-red light and the other to a light sensor installed outside the magnetically shielded room. The fiber approaching the skin during stimulation resulted in an increase of reflected light until the fiber reached the skin. It was therefore possible to measure the precise timing of the delivery of the tactile stimulation. Participants' view of the tactile stimulation apparatus was obstructed to avoid them performing the task using visual cues provided by the movement of the apparatus.

Procedure

In each trial, the participants listened to the one second mosquito sound and felt the tactile stimulus described above (see Figure 1A). While the individual stimuli were always the same, their relative timing, i.e. stimulus onset asynchrony (SOA), could vary from trial to trial.

The SOAs were individually determined prior to recording in a pilot run. Aim of this pilot was to determine the individual psychophysical curve (Fig. 1B) that describes how, in function of their SOAs, perception of the auditory and the tactile stimuli varies between integrated (i.e. the touch being judged as caused by the approaching mosquito) and segregated (i.e. the touch being judged as unrelated to the approaching mosquito). In particular, we were interested in the two "transition" SOAs, one before and one after the end of the sound, at which the occurrence of the touch elicited 50% of time an integrated and 50% of the time a segregated percept. Each one

of these two SOAs was determined by staircase procedures conducted using a staircase toolbox for Matlab (<https://code.google.com/p/matlabstaircase/>). Four (i.e. two for each of the two SOAs) independent staircase procedures were randomly interleaved during the same pilot run. For each of the two SOAs one staircase started from a clearly integrated percept (touch occurring at the sound offset) and one from a clearly segregated percept (touch occurring at the sound onset or 1 second after sound offset) and they moved towards the ambiguity threshold with a one-up/one-down step procedure. In order to quickly approach the ambiguity region, the step-sizes at the beginning were bigger (80ms) and progressively diminished until (after 7 reversals) becoming stable with a step-size of 16 ms. Each staircase procedure stopped when a defined number of trials (40) or of reversals (30), or of boundaries-hits (5) was reached. At this point the SOA was determined by averaging the two thresholds obtained.

In the recording session, according to previously individually estimated SOAs, two types of trials were presented. In the first type of trials, the touch occurred before the end of the sound and the SOA matched the 50% rate of integration/segregation as previously determined. In the second type of trials, the tactile stimulation was delivered after the end of the sound and matched the previously determined 50% rate of integration/segregation (Fig. 1B). Additionally, a third type of trials was presented where SOAs between audio and tactile stimuli were randomly chosen from a uniform distribution of SOA ranging from sound onset (1000 ms prior to the end of the sound) to 1000 ms after the sound offset. The latter trials were presented to increase the difficulty level of the perception evaluation task and were discarded from further analyses.

The three types of trials were randomly presented within a block. Each block consisted of 78 trials with 26 trials of each trial type. Each participant underwent nine blocks. The participants' task was to report after each trial whether the two sensory inputs belonged to the same event or were distinct events. Henceforth, INT response refers to the perception of the auditory and the tactile stimuli being perceived as a single event (i.e. the touch being judged as caused by the approaching mosquito); SEG response refers to the perception of the auditory and the tactile stimuli as distinct events (i.e. the touch being judged as unrelated to the approaching mosquito).

The experiment design allowed direct comparison of physically identical stimuli that produced either the percept of integrated or segregated AT stimuli. During the trial, participants were instructed to fixate on a cross on the display screen. After a period of 2000 ± 400 ms (uniform distribution) following the sound onset, participants were asked to report their perception by answering the question "Do the auditory and the tactile stimuli belong together?" Responses were delivered via button presses within the given time limit of 2000 ms using the left and right index fingers to respond "yes" or "no" respectively. The buttons corresponding to positive and negative were changed randomly on a trial-by-trial basis, in order to avoid possible confounds due to

motor-response mapping preparation. Trials were eliminated if participants responded prior to the onset of the tactile stimulus or failed to make a response within the limit. Such responses accounted for less than 1% of trials overall and were not further analyzed.

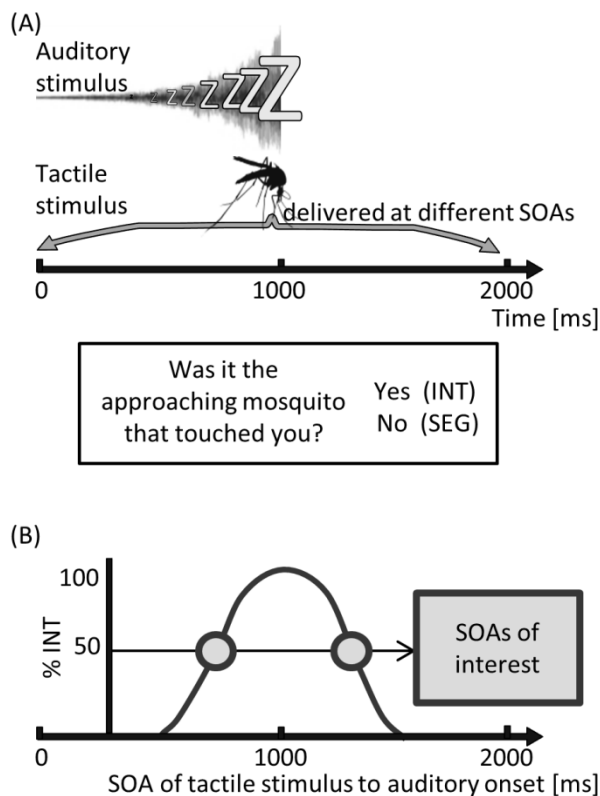


Figure 1

(A) Each trial consisted of an auditory stimulus, a one-second sound of a mosquito approaching the participant's head from the left side, and a spatially correlated tactile stimulus that was intended to simulate the bite of a mosquito. Different (SOAs) between audio and tactile stimuli yielded different percepts: integrated, i.e. the approaching mosquito and the touch belonged to the same event (condition INT, the impression that it was the approaching mosquito that touched them) or segregated (condition SEG, audio and tactile stimuli were unrelated events). SOAs could vary from trial to trial. Participants were asked to report their percept after each trial.

(B) Schematic psychophysical function describing the relationship between different SOAs and an integrated percept. SOAs of interest (prior and following sound

offset) were those in which integrated and segregated percepts were equally probable.

Data Acquisition

MEG data were recorded at a sampling rate of 5 kHz using a 306-channel (204 first order planar gradiometers, 102 magnetometers) VectorView MEG system (Elekta-Neuromag Ltd., Helsinki, Finland) in a magnetically shielded room (AK3B, Vakuum Schmelze, Hanau, Germany). Hardware filters were adjusted to band-pass the MEG signal in the frequency range of 0.01–1000 Hz. Prior to the MEG recording, 3 localization coils were placed on the forehead and 2 behind the ears. The position of the coils and the participants' head shapes were digitized with a Fastrak 3D digitizer (Polhemus, Colchester, VT, USA, <http://www.polhemus.com>). Prior to each experimental block, participants' head positions with respect to the MEG sensors were estimated. The neurophysiological MEG data were analyzed offline using Fieldtrip (Oostenveld et al., 2011), an open source toolbox for Matlab (www.mathworks.com).

Data preprocessing

As we were interested in the integration of a tactile stimulus with an ongoing auditory stimulus, epochs of four seconds (2.5 sec pre and 1.5 sec post) were centered on the tactile stimulus onset; onsets were obtained from the optical fiber device (see Stimuli and experimental instrumentation). Zero seconds therefore represents the onset of the tactile stimulus. Epochs were high-pass filtered at 1 Hz to remove very slow frequencies (DC offset) and filtered for line noise removal, then down-sampled to 400 Hz. Trials were visually inspected for possible artefacts and contaminated trials were excluded from further processing. As mentioned in the Procedure section, we are interested only in trials where the auditory and the tactile stimulus are related by a certain SOA that produced a rate of INT/SEG responses as close as possible to 50%, while the other trials (i.e. trials with random SOA) were discarded. Remaining trials were divided into two conditions based on participant responses: SEG and INT conditions. Since the rate of integrated/segregated percepts was not always 50%, we equalized the trial number by randomly discarding trials from the condition having a larger number of trials for each subject. This procedure was conducted separately for the case of touch occurring prior to the end of the sound and after the end of the sound. Equalization of trial number ensured comparable signal-to-noise ratios and ensured that any other common activation, including activity due to the auditory stimulus in the condition where touch arrives prior to the end of the sound, was cancelled when contrasting SEG and INT. Since we were interested in the neural mechanism underlying perceptual integration, trials in which touch occurred prior and after sound offset were collapsed.

Event related fields (ERFs) analysis in sensor space

For each participant, for condition INT and SEG, we averaged the preprocessed

data over trials to obtain the ERF waveforms. Since we were interested in the perception of a tactile stimulus with an ongoing auditory stimulus, statistical analysis was performed on the post-stimulus time after the tactile stimulus was delivered. A non-parametric cluster-based dependent-samples t-test with Monte-Carlo randomization [Maris and Oostenveld 2007] was conducted on the time window 0 ms to 200 ms (0 ms refers to the tactile stimulus onset). The time window was selected to encompass the first peak of activation following tactile stimulation.

Time-frequency analysis in sensors space

For each participant, time-frequency analysis was performed on single trials between 1-40 Hz. We applied an adaptive sliding time window with a length (Δt) of 5 cycles of the respective frequency and shifted in steps of 50 ms between -1.2 to .9s. A Hanning taper was applied yielding a spectral smoothing of $1/\Delta t$. For the gradiometers, orthogonal gradients of the magnetic field were calculated separately at each position. The sum of both directions was computed to obtain the power at each sensor irrespective of the orientation of the gradients (Medendorp et al., 2007). In order to statistically examine sensor level differences between the condition SEG and INT, we performed two non-parametric cluster-based dependent-samples t-tests with Monte-Carlo randomization for gradiometer recordings [Maris and Oostenveld, 2007]. A first statistical test of this type was run on a time-frequency window that encompassed the pre-(tactile) stimulus period (-450 to 50 ms, 1-40 Hz), while the second test was run for the post-(tactile) stimulus time-window (0 to 500 ms, 1-40 Hz). This type of statistical test was used because it controls for the multiple comparisons problem (type I error rate). In particular, a t-test for each sensor-time-frequency pair is performed. Subsequently, t-values exceeding a certain threshold (here 0.05) are considered and, based on spatial, temporal and frequency adjacencies, clusters of significant differences over space, time, and frequency are identified. Finally, cluster-level statistics, i.e. the sum of t-values within each identified cluster, are evaluated under the permutation distribution of the maximum cluster-level statistic. This permutation distribution is obtained from a random draw of the observed data (random partition) repeated a high number of times (here 1000). Thus, the proportion of random partitions that resulted in a larger value than the true observed one (maximum cluster-level statistics) is calculated and represents significance probability (p-value) under which cluster-level statistics are evaluated. Since a significant positive cluster of condition INT versus SEG was revealed in the time window from -300 ms to 50 ms in the alpha range (peak at 9Hz), further investigation in the prestimulus window was focused on this frequency band and time window.

Localization of the time-frequency findings in source space

The individual head shapes obtained with Polhemus were co-registered with the individual structural MRI (4T Bruker MedSpec, Siemens). First, a course alignment

was performed on the three fiducials, as collected with Polhemus prior to the MEG measurement, to the corresponding points on the individual MRI. Second, a rigid body transformation fitted the digitized head shape points of the individual to the head surface as extracted from the structural MRI. For two participants for whom individual MRIs were not available, a Montreal Neurological Institute (MNI) brain was morphed to fit the individual head shape. To compare the source activity across participants on a group level, a grid (1 cm spacing, 2982 points inside the brain) fitted to a template MNI brain was warped to individual headspace for each participant. Resulting grid points were not regularly spaced but they were located in the same area with respect to the MNI template across participants [Larson-Prior et al., 2013]. An analytical single shell model was fitted to the individual segmented MRI [Nolte 2003] and the leadfield was calculated for each grid point.

Both magnetometers and gradiometers were taken into account for source estimation after appropriate adjustment of the balancing matrix based on the distance of the gradiometers (17 mm), applied in order to take into account the different measurement units of magnetometer [T] and gradiometers [T/m] sensors. We projected the preprocessed data, i.e. each trial, into source space, i.e. on the points of the grid, using the linearly constrained minimum variance (LCMV) spatial filter [Van Veen et al. 1997]: first, for each participant, a LCMV filter was estimated on the combined conditions SEG and INT on the time interval -0.8 ms to +0.8 ms with respect to touch onset; then, the time-series of each trial was multiplied with this common filter. This allowed time-frequency analysis with the same parameters utilized at the sensor level for conditions INT and SEG. The grand average of the difference between the two conditions for the significant time interval (from -291 ms to +42 ms) at the significant frequency (9 Hz) was then plotted on an MNI brain.

Connectivity analysis in source space

The spectral analysis suggested strong pre-touch alpha power differences (see Results). To investigate potential connectivity patterns that predispose different percepts, we examined brain functional connectivity to and from the voxel with the maximum power effect (MNI coordinates: [60 -40 40]; corresponding to BA40). As a connectivity metric we used phase locking values (PLV) [Lachaux et al., 1999]. The frequency of interest was the peak frequency in the time-frequency results (9 Hz) in the time window chosen for the source analysis (-291ms +42ms). We used the fast Fourier transform (FFT) algorithm on the sensor data (multitaper analysis, Hanning window) and projected the complex Fourier coefficients into source space by multiplying them with the respective spatial filters. Spatial filters were constructed using Dynamic Imaging of Coherent Sources (DICS) beamformer [Gross et al., 2001] from the cross-spectral density matrix, calculated using a multitaper FFT transformation over the time window and frequency of interest resulting from the analysis on sensor level (-0.291 0.042 ms, 9 Hz, temporal interval centered at the maximum effect found in the time

frequency analysis, length of 3 cycles/9 Hz = 333 ms) and the respective leadfields. We then calculated PLV between the reference voxel and all other voxels.

To perform the statistical analysis, we used a non-parametric cluster-based dependent-samples t-test with Monte-Carlo randomization for condition INT versus SEG. Neighboring voxels above a given threshold with a minimum cluster size were considered as significant. The minimum size required for a cluster to be considered significant was defined according to a cluster-based approach common in fMRI research and implemented as 3DClustSim within the AFNI suite [Cox 1996; Cox and Hyde 1997]. The program simulates random t-values on a grid provided by the user, thresholds them according to a p-value and records the size of remaining clusters. This process is repeated 10,000 times and calculates the probability that this purely random activity will produce a significant cluster of a given size. To account for correlation between neighboring voxels, we applied a smoothing to the random activity. We used a smoothing factor of 1cm on the random data prior to cluster identification. The value of 1 cm was chosen as it coincides to the estimated spatial resolution of MEG [Hansen et al., 2010]. This leads to a minimum cluster size of 15 voxels to be significant, for an alpha threshold of $p \leq 0.05$.

Results

In each trial in the experimental paradigm, an audio and a tactile stimulus were presented in a way that, although physically identical, could be perceived either as integrated, i.e. as emanating from the same source, or as segregated, i.e. they were independently generated. The perception fluctuated on a trial-by-trial basis. As the stimuli were physically identical, it is expected that any differences between conditions INT and SEG will arise from intrinsic neural dynamics and reflect only the differences related to perception.

Behavior

In a first pilot run, we studied for each participant the individual psychophysical function that links perception of integration/segregation to SOAs, so as to determine the two “transition” SOAs at which stimuli are ambiguously related and the rate of INT/SEG is 50. For touch occurring before the end of the sound, mean “transition” SOA is at 665 ms (± 238 ms s.d., 0 ms refers to the onset of sound). Exploration of the “right” part of the psychophysical function, i.e. for touch occurring after the end of the sound, resulted in a mean “transition” SOAs of 1338 ms (± 188 ms s.d., 0 ms refers to the onset of sound). In the subsequent part of the experiment, stimulation at these “transition” SOAs resulted in a percentage of INT of $44 \pm 14\%$ trials (mean \pm s.d) for the touch occurring before the end of the sound, and $46 \pm 15\%$ trials for the other case. For three participants we were unable to elicit an ambiguous percept for touch before the end of the sound. For these participants only trials belonging to “transition” SOAs for touch after the end of the sound were used in the analysis.

Event related fields analysis

Touch elicited strong responses in both conditions, with the evoked response showing the most pronounced peak ~85 ms. Source analysis of the 0-200ms time window versus prestimulus baseline revealed pronounced generators in left and right primary somatosensory regions. Bilateral activation for lip/face tactile stimulation is in accordance with the literature [e.g. Schulz et al. 2004]. Interestingly, the ERF following the tactile stimulus did not differ between INT and SEG (Figure 2A). In the following analysis we focused on the period before the upcoming tactile stimulus.

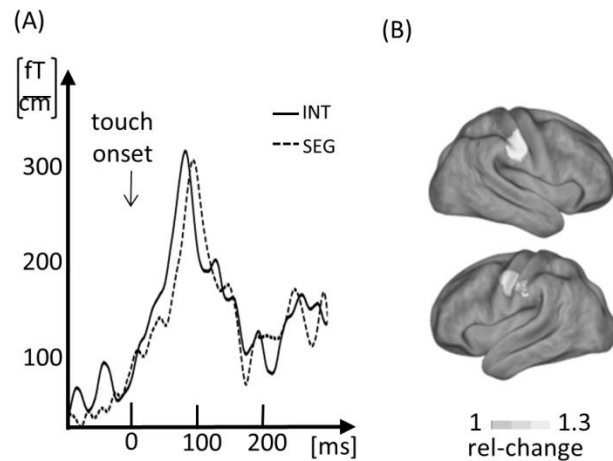


Figure 2

(A) The event-related field for the combined gradiometers for condition INT (black line) and SEG (dotted line). A non-parametric cluster-based dependent-samples *t*-test with Monte-Carlo randomization was conducted on the time window 0 ms to 200 ms and failed to reveal any statistically significant difference between the two conditions. (B) Source analysis shows the localization of touch-related evoked responses - independent of condition - in primary somatosensory cortex. These results are shown as a sanity check of our data

Time-frequency analysis

We statistically compared the time-frequency representations of INT trials versus SEG trials at sensor level. This comparison yielded a positive cluster ($p < 0.05$) starting from approximately 290 ms prior to touch onset and lasting until 40 ms after touch onset. The effect was in the alpha range (Figure 3A), clearly peaking at 9 Hz and was dominant at central and frontal sensors bilaterally. We consider the time-window of the significant effect as prestimulus (prior to the tactile stimulus) even if it exceeds

the presentation of the tactile stimulus, since the ERF of the tactile stimulus had its first peak only at 85 ms on primary somatosensory cortex (Figure 2A). The spatial pattern of the sensor topography is also reflected at source level (Figure 3B). The main cortical generator of the difference between INT and SEG is localized in the right hemisphere (contralateral to tactile stimulus) in the inferior parietal lobule (rIPL) (BA 40, max: [60, -40, 40] MNI coordinates). This cluster propagates towards sensorimotor cortices (BA 2, 3, 4). Another cluster of activation is localized centrally over superior parietal lobules (SPLs) (BA7). Finally, ipsilateral (with respect to touch) activation is localized in lIPL and similarly to contralateral activation, this cluster expands towards sensorimotor cortices (BA 2 and 3) reaching the frontal cortex (BA6).

Functional connectivity analysis

The maximum grid point of the right IPL cluster described above (MNI coordinates [60 -40 40]) was chosen as a seed and functional connectivity between the seed and the rest of the brain was conducted in the time-window of interest (-290 until 40 ms) for the frequency of interest (9 Hz). As shown in Figure 3C, three brain regions - all located in the right hemisphere and contralateral to the upcoming tactile stimulus - showed altered functional connectivity to rIPL during INT with respect to SEG. In particular, rIPL showed more coupling with the right prefrontal cortex (BA 10) and motor-sensory regions (size: 35 voxels, max at MNI [30 50 20]) and with the right visual cortex (BA 18, BA 19) (size: 16 voxels, max at MNI [40 -90 20]) prior to the INT with respect to SEG condition. Interestingly, the coupling between rIPL and the right auditory belt area (BA 20, size: 20 voxels, max at MNI [60, -10, 10]) was reduced in INT with respect to SEG condition. For each of these three clusters we calculated the mean PLV value across voxels and normalized it to the values of the baseline, i.e. prior to the sound onset, as shown in Figure 3D. A coupling in condition SEG drives the effect in auditory belt area, while a decoupling in condition SEG and a coupling in condition INT drive the other effects.

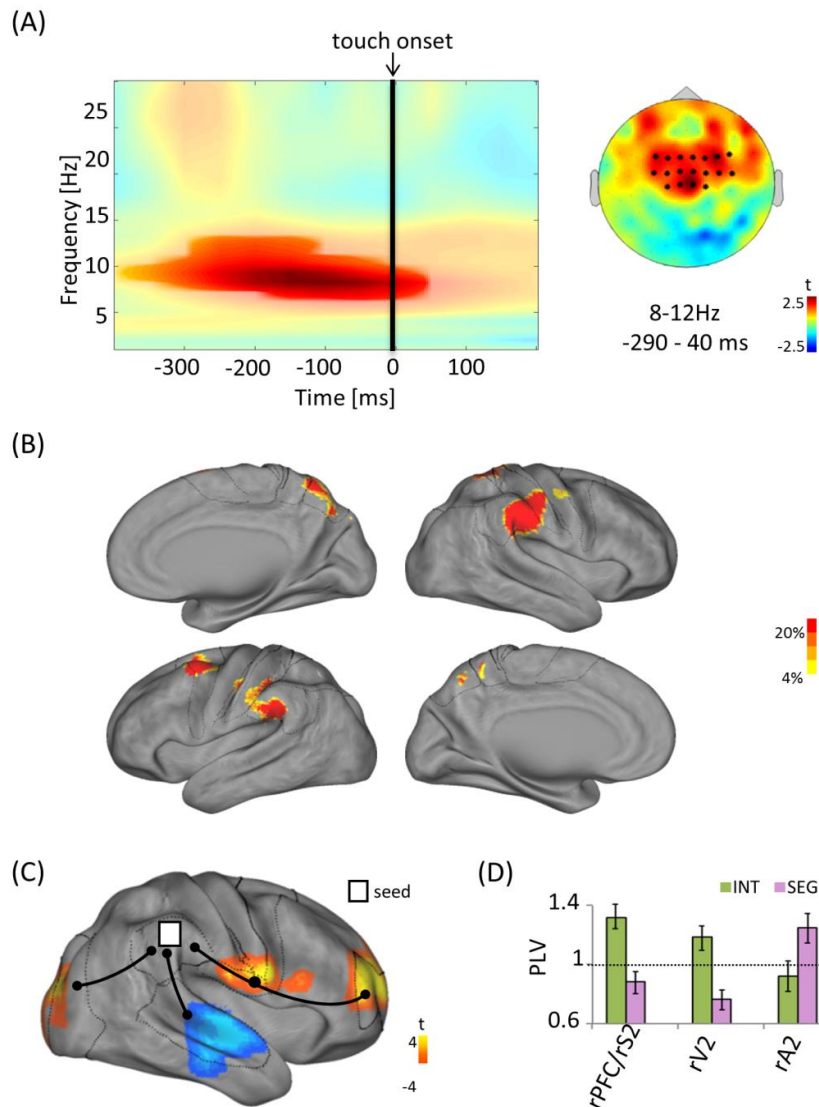


Figure 3

(A) Results of the statistical comparison at sensor level for condition INT vs condition SEG. Red color indicates relatively more power during condition INT. Time 0 ms indicates the onset of touch.

Left: time-frequency representation (TFR) of the statistically significant sensors (cluster- $p < 0.05$) for the interval pre- and post- tactile stimulus. Non-significant values have been reduced to 40% opacity.

Right: Topographical representation of the statistical results for the time-window -290 to 40 ms in the frequency band 8-12 Hz. Black dots represent the sensors that gave statistically significant results at sensor level and whose TFR are plotted on the left side of the figure.

(B) Projection in source space of the grand average of INT versus SEG for the time-frequency window and frequency of the significant results at sensors level.

(C) Phase synchrony between the seed region, rIPL ([60 – 40 40], MNI coordinates), and the whole-brain volume. T-values of INT versus SEG percept masked for statistical significance are displayed. Dashed lines represent the borders of Brodmann areas of interest. rA2 exhibited significant decoupling with the rIPL, while rPFC/rS2 and rV2 showed increased coupling with rIPL during integrated compared to segregated percept.

(D) We calculated PLV of each significant cluster showed in (C) normalized to the baseline prior to the sound onset in order to evaluate if the differences reported in (C), originate by, respectively, a coupling and a decoupling in both conditions with respect to baseline or, on the contrary, the effect is driven only by a coupling (or decoupling) in one condition. PLV greater than 1, indicated by the dashed line, indicates higher coupling with respect to baseline and vice versa. Barplots show differences of PLV in the rPFC/rS2 and rV2 originate from an increase of coupling for INT and a decrease of coupling for SEG with respect to the baseline, while for the rA2 the difference is primarily driven by an increase of coupling for the SEG percept

Discussion

In the present MEG study, we investigated how prestimulus brain-states predispose integrated or segregated percepts of an ongoing audio and an upcoming tactile stimulus. Since the temporal relationship between stimuli of different modalities is a critical factor for their perceptual integration [Calvert et al., 2004; Welch, 1999], we manipulated the temporal relation between the offset of a dynamic (approaching) sound of a mosquito and the onset of a touch (resembling the landing of the mosquito). By stimulating around the threshold SOA, at which stimuli were perceived as either emanating from the same source (i.e. the mosquito approaching the head and landing on it; condition INT) or as two distinct events (condition SEG), perception of physically identical multisensory stimuli fluctuated between trials from integrated to segregated and vice versa. This approach is powerful as it allows direct comparison of physically identical but differently perceived stimuli. We hypothesized that differences in brain-states prior to upcoming tactile stimulus can predispose different multisensory percepts. Alpha power was relatively increased in a pre-tactile stimulus time-window of about 330 ms for subsequent perceptions of integrated versus segregated. The difference was localized in IPLs and SPLs bilaterally. Moreover, the functional connectivity pattern was modulated by different percepts in the same time window.

Prior to integrated percept, rIPL was more coupled with the right secondary somatosensory and visual cortices and more decoupled from secondary auditory cortex as compared to the period prior to segregated percept.

IPLs and SPLs involved in multisensory interactions: evidence from monkeys and humans

Comparisons between the human and primate auditory cortices [Galaburda and Sanides, 1980; Hackett et al., 2001; Sweet et al., 2005] reveal that human IPL seems to be a close homologue of the caudomedial (CM) belt area of primates' auditory association cortex, which consistently shows auditory and somatosensory stimuli convergence with a bias for the skin surfaces of the head and neck [Brosch et al., 2005; Fu et al., 2003; Giard and Peronnet, 1999; Kayser et al., 2005; Schroeder et al., 2001; Schroeder and Foxe, 2002].

Our findings are consistent with the IPL being the human homologue of the multisensory area CM in monkeys [Calvert et al., 2004; Hackett et al., 2007], since the integrated percept of the looming sound and the touch were preceded by higher alpha power exactly in bilateral IPLs. In previous literature, IPLs and SPLs have been described as being activated during a localization task for both auditory and tactile stimuli [Renier et al., 2009], and relevant in localization tasks for stimuli of all modalities [e.g. Bushara et al. 1999; Reed et al., 2005]. Furthermore, rIPL appears to be involved in perception of time [e.g. Harrington et al. 1998; Rao et al. 2001] and the detection of temporal synchrony between auditory and visual stimuli [Adhikari et al., 2013; Bushara et al., 1999; Dhamala et al., 2007], and in tasks where multisensory temporal and spatial information need to be evaluated together [Assmus et al., 2003; Assmus et al., 2005].

In our paradigm, we manipulated the temporal relationship between the two stimuli. Given that the sound is looming, time is translated into the location of the sound source so that the integration/segregation task also had a spatial component, i.e. the localization of the flying mosquito, based on the loudness of its buzzing sound, at the time the tactile stimulus is delivered. Although the role of IPLs and SPLs described above has been determined based on post-stimulus activities, in our experiment the difference in IPLs and SPLs prior to the application of the tactile stimulus between conditions INT and SEG suggests that the ongoing brain state affects upcoming computations of spatio-temporal relations between stimuli of different modalities (here, audio and tactile). Our results propose for the first time that local and inter-areal synchronization properties of IPL *prior* to the stimuli crucially determine whether an AT stimulus will be perceived as integrated or not.

AT interactions in humans: previous findings and the linear additive model

In humans, AT interactions have been localized in regions in the vicinity of IPLs, i.e. in secondary somatosensory cortex, the superior temporal gyrus/sulcus and

the posterior parietal cortex at different latencies [Beauchamp et al., 2008; Foxe et al., 2000; Foxe et al., 2002; Gobbelé et al., 2003; Hofer et al., 2013; Kassuba et al., 2013; Lütkenhöner et al., 2002; Murray et al., 2005; Schürmann et al., 2006], but the results are not consistent. This could be partially explained by the fact that crossmodal interactions are currently estimated using the linear additive model. This model is derived from observations on the cat's superior colliculus and defines interaction between senses as the difference between the activity elicited by the bimodal stimulus and the sum of the activity elicited by its unisensory components (i.e. here: interaction = $AT - [A+T]$). Given that the relative saliency of the stimuli ("inverse effectiveness" rule, Stein and Meredith 1993) affects the interaction, it is likely that part of the variance found in the results in previous works, i.e. results defined through the linear additive model, reflects differences in the type of stimuli used, i.e. in their relative saliency (see Gobbelé et al. 2003 for a discussion of the topic). It is also important to note that the model is sensitive only to non-linear observations, while interactions have been documented as well in a linear fashion [e.g. Stanford et al. 2005]. The model is also potentially sensitive to several confounds if experiments are not carefully designed, such as differences in attentional load between multimodal and unimodal conditions [Besle et al., 2004].

In previous works, interactions reflected the co-occurrence of stimuli of different modalities regardless whether they elicited an integrated or segregated percept. As we compared bimodal physically identical stimuli, the lack of statistically significant difference in the post-stimulus phase suggests that, at least initially, both conditions elicit the same amount of activation. Although we cannot exclude that the lack of significance is due to insufficient power or too conservative statistical methods. However, given that identical bimodal stimuli are differently perceived, either as integrated or as segregated, this paradigm ensured that our results purely reflect our interest, i.e. neural activity/connectivity patterns related to the different percepts.

Integrated percepts are preceded by relative increase of alpha activity in IPLs and SPLs regions

In the interval prior to tactile stimulus, bilateral SPLs and IPLs showed an increased alpha power in cases of subsequent integration with respect to segregation. According to the current literature, alpha band power is believed to be crucial for the inhibitory-excitatory balance of the brain: higher alpha in a certain brain region is related to local inhibition, while lower alpha reflects states of heightened excitability [Jensen and Mazaheri, 2010]. Likewise, in recent years, there has been cumulative evidence that power fluctuations of neuronal activity in this frequency band, prior to stimulation, substantially influence subsequent perception of weak stimuli, with alpha power inversely correlating to detection and discrimination performance. For instance, local fluctuations of alpha activity in the prestimulus period, modulate the perception of near-threshold (NT) stimuli [e.g. Weisz et al. 2014], namely low levels of alpha

power preceded “hits” (the NT stimulus is perceived), while high levels of alpha power preceded misses (the NT stimulus is not perceived) [for similar results see e.g. Thut et al. 2006; Hanslmayr et al. 2007; Schubert et al. 2008; van Dijk et al., 2008; Romei et al., 2010; Jones et al. 2010]. Interestingly, recent studies show that low alpha is related not only to higher probability of NT stimuli detection, but also to identification of more than one super-threshold stimuli of the same modality presented with very close temporal proximity [Baumgarten et al. 2014] or in the context of visual illusions [Lange et al. 2013].

Although we cannot exclude that high alpha simply biased participants towards a particular judgment, this is highly improbable for the following reasons. Our findings are localized in areas characterized as multisensory in the literature (see first section of the Discussion). The ambiguous temporal relation of audio and tactile stimuli in our paradigm allows considering that the level of alpha in those specific areas was what determined the upcoming INT or SEG perception. High excitability of these areas predisposed segregation of AT stimuli, which is in line with the aforementioned findings, where also low alpha was connected to perception of two distinct stimuli [Baumgarten et al. 2014; Lange et al. 2013]. We can therefore argue first that excitability of IPLs and SPLs is critical for perceptual binding, and second, that alpha plays a role in crossmodal perception on relevant areas.

Engagement/disengagement of relevant brain areas into a distributed network predispose distinct multisensory percepts

The brain connectivity analysis revealed a specific pattern of functional connections between distant cortical areas preceding different multisensory percepts. For the connectivity analysis, we used rIPL as a seed from which we computed connectivity to all other voxels of the brain. rIPL was chosen because it yielded the maximum difference between INT and SEG in the prestimulus time window and was therefore assumed to reveal major differences in a network level. rIPL exhibited increased coupling with rPFC, right secondary somatosensory and visual cortices, and decreased coupling with secondary auditory cortex, prior to the integrated percept with respect to the segregated percept.

The relationship between prestimulus activity patterns and perception has been recently investigated [Hanslmayr et al., 2007; Keil et al., 2012; Ploner et al., 2010]. These studies not only showed local power differences over relevant regions, but also that the inclusion or exclusion of these regions into a globally distributed functional network critically influences subsequent perception [Weisz et al. 2014]. Our data essentially confirm the relevance of both local and global aspects of brain-states in the prestimulus for subsequent perception: local alpha power levels as well as long-range connections appear to predispose whether subsequent AT stimuli are integrated or not.

The absence of post-stimulus effects suggests that, rather than a specific multisensory area, integrated/segregated perception is predisposed by a network

comprising "multisensory integration" areas with unisensory areas. The precise network configuration is possibly modulated by local inhibitory/excitatory balance. Indeed, inhibition is believed to be essential for the establishment of long-range networks [Jonas and Buzsaki, 2007] and one might hypothesize that the higher alpha power observed in IPLs and SPLs in our study for the integrated condition reflect the formation of the network relevant for multisensory integration.

The notion of a network that predisposes integrated or segregated percepts offers advantages and proposes a complementary mechanism to the traditional hierarchical model. The latter assumes that each sense is first processed in isolation and interaction takes place in multisensory "heteromodal" brain regions where feedforward convergence from the sensory-specific cortices occurs [Stein and Meredith, 1993]. However, this fails to explain interactions already at early latencies and in areas close to primary cortices [e.g. Foxe et al. 2000], and multisensory interactions that create novel percepts such as the "Hearing Hands Effect" in which gently touching a vibrating probe dramatically changes the perception of the sound intensity of a probe tone [Schürmann et al., 2004], or the "Parchment-Skin Illusion" [Jousmäki and Hari 1998]. Contrary to the hierarchical model, large-scale interactions provide flexibility and efficiency in a network [Varela et al., 2001], which, when it comes to the brain, is fundamental to the efficient integration of inputs from a dynamic and multifaceted world along with endogenous activity. Thus, functional dynamic long-scale interactions seem to be a plausible mechanism to be added to the way multisensory interactions are thought about, using the hierarchical model which is primarily based on unchanged anatomical pathways.

The fact that the differentiated network comprised even visual areas despite the presentation of AT stimuli, suggests a non-modality specific integration/segregation network. Indeed, vision plays a dominant role in localization tasks [Eimer, 2004]. For example, in the famous Ventriloquist Effect, auditory [Alais and Burr, 2004] and tactile [Caclin et al., 2002] stimuli were wrongly detected towards a simultaneous visual stimulus at a different location. Interestingly, it has also been shown that by blurring the visual stimulus, the effect is reversed with vision being captured by audition thus showing how multisensory perception must rely on a flexible mechanism. In our experiment, participants had no visual information that could bias their percept and no post-stimulus effects were found in visual areas. It seems that visual cortices form part of a distributed cortical network that predisposes integration/segregation possibly for exploiting any upcoming visual information.

Conclusions

Previous research has already pointed out the role of prestimulus brain-states and cortical networks in multisensory processing [Hipp et al., 2011; Keil et al., 2012; Senkowski et al., 2008]. In the present MEG study, we investigated the neural correlates of multisensory percepts. We enhanced our results by exploiting a paradigm

that allowed direct comparison of physically identical AT stimuli that elicited different percepts. Locally, our results showed how relative inhibition prior to the upcoming tactile stimulus of areas involved in temporal and localization tasks across different modalities leads to an AT integrated percept. The results suggest a functional role of these areas in comparing and separating stimuli of different modalities. Globally, we observed prestimulus differences in a cortical network that encompasses secondary areas of different modalities and prefrontal cortex suggesting a non-modality specific integration/segregation network.

Acknowledgments

The contribution of CL and NW was supported by the European Research Council (ERC StG 283404).

References

- Adhikari BM, Goshorn ES, Lamichhane B, Dhamala M (2013): Temporal-order judgment of audiovisual events involves network activity between parietal and prefrontal cortices. *Brain Connect* 3:536–45.
- Alais D, Burr D (2004): The ventriloquist effect results from near-optimal bimodal integration. *Curr Biol* 14:257–62.
- Algazi VR, Duda RO, Thompson DM, Avendano C (2001): The CIPIC HRTF database. *Proc IEEE Work Appl Signal Process to Audio Electroacoust*:99–102.
- Assmus a, Marshall JC, Noth J, Zilles K, Fink GR (2005): Difficulty of perceptual spatiotemporal integration modulates the neural activity of left inferior parietal cortex. *Neuroscience* 132:923–7.
- Assmus A, Marshall JC, Ritzl A, Noth J, Zilles K, Fink GR (2003): Left inferior parietal cortex integrates time and space during collision judgments. *Neuroimage* 20:S82–S88.
- Beauchamp MS, Yasar NE, Frye RE, Ro T (2008): Touch, sound and vision in human superior temporal sulcus. *Neuroimage* 41:1011–1020.
- Baumgarten TJ, Schnitzler A, Lange J (2014): Prestimulus alpha power influences tactile temporal perceptual discrimination and confidence in decisions. *Cereb Cortex*:1–13.
- Besle J, Fort A, Delpuech C, Giard M-H (2004): Bimodal speech: early suppressive visual effects in human auditory cortex. *Eur J Neurosci* 20:2225–34.
- Brosch M, Selezneva E, Scheich H (2005): Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. *J Neurosci* 25:6797–806.
- Bruce C, Desimone R, Gross CG (1981): Visual properties of neurons in a polysensory area in the superior temporal sulcus of the macaque. *J Neurophysiol* 46:369–384.
- Bushara KO, Weeks R a, Ishii K, Catalan MJ, Tian B, Rauschecker JP, Hallett M (1999): Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. *Nat Neurosci* 2:759–66.
- Caclin A, Soto-Faraco S, Kingstone A, Spence C (2002): Tactile “capture” of audition. *Percept Psychophys* 64:616–30.
- Calvert GA, Spence C, Stein BE. 2004. *The handbook of multisensory processes*. Cambridge (MA): MIT press.
- Canzoneri E, Magosso E, Serino A (2012): Dynamic sounds capture the boundaries of peripersonal space representation in humans. *PLoS One* 7:e44306.

- Cox RW (1996): AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res* 29:162–73.
- Cox R, Hyde J (1997): Software tools for analysis and visualization of fMRI data. *NMR Biomed* 10:171–8.
- Dhamala M, Assisi CG, Jirsa VK, Steinberg FL, Kelso J a S (2007): Multisensory integration for timing engages different brain networks. *Neuroimage* 34:764–73.
- Eimer M (2004): Multisensory Integration: How visual experience shapes spatial perception. *Curr Biol* 14:R115–R117.
- Foxe JJ, Morocz IA, Murray MM, Higgins BA, Javitt DC, Schroeder CE (2000): Multisensory auditory-somatosensory interactions in early cortical processing revealed by high-density electrical mapping. *Brain Res Cogn Brain Res* 10:77–83.
- Foxe JJ, Wylie GR, Martinez A, Schroeder CE, Daniel C, Guilfoyle D, Ritter W, Murray MM, Butler JS, Molholm S, Fiebelkorn IC, Mercier MR, Schwartz TH (2002): Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. *J Neurophysiol* 88:540–543.
- Fu KG, Foxe JJ, Murray MM, Higgins BA, Javitt DC, Schroeder CE (2001): Attention-dependent suppression of distracter visual input can be crossmodally cued as indexed by anticipatory parieto – occipital alpha-band oscillations. *Cogn Brain Res* 12:145–152.
- Fu K-MG, Johnston T a, Shah AS, Arnold L, Smiley J, Hackett TA, Garraghty PE, Schroeder CE (2003): Auditory cortical neurons respond to somatosensory stimulation. *J Neurosci* 23:7510–5.
- Galaburda a, Sanides F (1980): Cytoarchitectonic organization of the human auditory cortex. *J Comp Neurol* 190:597–610.
- Ghazanfar A, Schroeder CE (2006): Is neocortex essentially multisensory? *Trends Cogn Sci* 10:278–85.
- Giard MH, Peronnet F (1999): Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J Cogn Neurosci* 11:473–90.
- Gobbelé R, Schürmann M, Forss N, Juottonen K, Buchner H, Hari R (2003): Activation of the human posterior parietal and temporoparietal cortices during audiotactile interaction. *Neuroimage* 20:503–511.
- Graziano MS, Reiss L a, Gross CG (1999): A neuronal representation of the location of nearby sounds. *Nature* 397:428–30.
- Gross J, Kujala J, Hämäläinen M, Timmermann L, Schnitzler A, Salmelin R (2001): Dynamic imaging of coherent sources : Studying neural interactions in the human brain. *Proc Natl Acad Sci U S A* 98:694–699.
- Hackett TA, Preuss TM, Kaas JH (2001): Architectonic identification of the core region in auditory cortex of macaques, chimpanzees, and humans. *J Comp Neurol* 222:197–222.
- Hackett TA, Smiley JF, Ulbert I, Karmos G, Lakatos P, de la Mothe LA, Schroeder CE (2007): Sources of somatosensory input to the caudal belt areas of auditory cortex. *Perception* 36:1419–1430.
- Hansen PC, Kringelbach ML, Salmelin R. 2010. MEG. An introduction to methods. New York: Oxford University Press.
- Hanslmayr S, Aslan A, Staudigl T, Klimesch W, Herrmann CS, Bäuml K-H (2007): Prestimulus oscillations predict visual perception performance between and within subjects. *Neuroimage* 37:1465–73.
- Harrington DL, Haaland KY, Knight RT (1998): Cortical networks underlying mechanisms of time perception. *J Neurosci* 18:1085–95.
- Hipp JF, Engel AK, Siegel M (2011): Oscillatory synchronization in large-scale cortical networks predicts perception. *Neuron* 69:387–96
- Hoefler M, Tyll S, Kanowski M, Brosch M, Schoenfeld M a, Heinze H-J, Noesselt T (2013): Tactile stimulation and hemispheric asymmetries modulate auditory perception and neural responses in primary auditory cortex. *Neuroimage* 79:371–82.

- Hummel F, Andres F, Altenmu E, Dichgans J, Gerloff C (2002): Inhibitory control of acquired motor programmes in the human brain. *Brain* 125:404–20.
- Jensen O, Gelfand J, Kounios J, Lisman JE (2002): Oscillations in the alpha band (9-12 Hz) increase with memory load during retention in a short-term memory task. *Cereb Cortex* 12:877–82.
- Jensen O, Mazaheri A (2010): Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Front Hum Neurosci* 4:186. Jonas P, Buzsaki G (2007): Neural inhibition.
- Jonas P, Buzsaki G. (2007): Neural inhibition. *Scholarpedia* 2: 3286.
- Jones SR, Kerr CE, Wan Q, Pritchett DL, Hämäläinen M, Moore CI (2010): Cued spatial attention drives functionally relevant modulation of the mu rhythm in primary somatosensory cortex. *J Neurosci* 30:13760–5.
- Jousmäki V, Hari R (1998): Parchment-skin illusion: sound-biased touch. *Curr Biol* 8:190.
- Kassuba T, Menz MM, Röder B, Siebner HR (2013): Multisensory interactions between auditory and haptic object recognition. *Cereb Cortex* 23:1097–107.
- Kayser C, Petkov CI, Augath M, Logothetis NK (2005): Integration of touch and sound in auditory cortex. *Neuron* 48:373–84.
- Keil J, Müller N, Ihssen N, Weisz N (2012): On the variability of the McGurk effect: audiovisual integration depends on prestimulus brain states. *Cereb Cortex* 22:221–31.
- Kitagawa N, Zampini M, Spence C (2005): Audiotactile interactions in near and far space. *Exp Brain Res* 166:528–37.
- Klimesch W, Sauseng P, Hanslmayr S (2007): EEG alpha oscillations: the inhibition-timing hypothesis. *Brain Res Rev* 53:63–88.
- Lachaux J-P, Rodriguez E, Martinerie J, Varela FJ (1999): Measuring phase synchrony in brain signals. *Hum Brain Mapp* 8:194–208.
- Lakatos P, Chen C-M, O’Connell MN, Mills A, Schroeder CE (2007): Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron* 53:279–92.
- Lange J, Oostenveld R, Fries P (2013): Reduced occipital alpha power indexes enhanced excitability rather than improved visual perception. *J Neurosci* 33:3212–20.
- Larson-Prior LJ, Oostenveld R, Della Penna S, Michalareas G, Prior F, Babajani-Feremi a, Schoffelen J-M, Marzetti L, de Pasquale F, Di Pompeo F, Stout J, Woolrich M, Luo Q, Bucholz R, Fries P, Pizzella V, Romani GL, Corbetta M, Snyder a Z (2013): Adding dynamics to the Human Connectome Project with MEG. *Neuroimage* 80:190–201.
- Laurienti PJ, Burdette JH, Wallace MT, Yen Y-F, Field AS, Stein BE (2002): Deactivation of sensory-specific cortex by crossmodal stimuli. *J Cogn Neurosci* 14:420–9.
- Lütkenhöner B, Lammertmann C, Simões C, Hari R (2002): Magnetoencephalographic correlates of audiotactile interaction. *Neuroimage* 15:509–22.
- Maris E, Oostenveld R (2007): Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164:177–90.
- Medendorp WP, Kramer GFI, Jensen O, Oostenveld R, Schoffelen J-M, Fries P (2007): Oscillatory activity in human parietal and occipital cortex shows hemispheric lateralization and memory effects in a delayed double-step saccade task. *Cereb cortex* 17:2364–74.
- Murray MM, Molholm S, Michel CM, Heslenfeld DJ, Ritter W, Javitt DC, Schroeder CE, Foxe JJ (2005): Grabbing your ear: rapid auditory- somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cereb Cortex* 15:963–974.
- Nolte G (2003): The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors. *Phys Med Biol* 48:3637–52.
- Ploner M, Lee MC, Wiech K, Bingel U, Tracey I (2010): Prestimulus functional connectivity determines pain perception in humans. *Proc Natl Acad Sci U S A* 107:355–60.
- Rao SM, Mayer AR, Harrington DL (2001): The evolution of brain activation. *Nat Neurosci* 4:317–323.

- Reed CL, Klatzky RL, Halgren E (2005): What vs. where in touch: an fMRI study. *Neuroimage* 25:718–26.
- Renier L a, Anurova I, De Volder AG, Carlson S, VanMeter J, Rauschecker JP (2009): Multisensory integration of sounds and vibrotactile stimuli in processing streams for “what” and “where”. *J Neurosci* 29:10950–60.
- 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? *J Neurosci* 30:8692–7.
- Ruhnau P, Hauswald A, Weisz N (2014): Investigating ongoing brain oscillations and their influence on conscious perception – network states and the window to consciousness. *Front Psychol* 5:1230.
- Schlack A, Sterbing-D’Angelo SJ, Hartung K, Hoffmann K-P, Bremmer F (2005): Multisensory space representations in the macaque ventral intraparietal area. *J Neurosci* 25:4616–25.
- Schroeder CE, Foxe JJ (2002): The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. *Brain Res Cogn Brain Res* 14:187–98.
- Schroeder CE, Lindsley RW, Specht C, Marcovici A, Smiley JF, and Javitt DC (2001): Somatosensory input to auditory association cortex in the macaque monkey. *J Neurophysiol* 85:1322–1327.
- Schulz M, Chau W, Graham SJ, McIntosh AR, Ross B, Ishii R, Pantev C (2004): An integrative MEG-fMRI study of the primary somatosensory cortex using crossmodal correspondence analysis. *Neuroimage* 22:120–33.
- Schürmann M, Caetano G, Hlushchuk Y, Jousmäki V, Hari R (2006): Touch activates human auditory cortex. *Neuroimage* 30:1325–31.
- Schürmann M, Caetano G, Jousmäki V, Hari R (2004): Hands help hearing: facilitatory audiotactile interaction at low sound-intensity levels. *J Acoust Soc Am* 115:830–832.
- Senkowski D, Schneider TR, Foxe JJ, Engel AK (2008): Crossmodal binding through neural coherence: implications for multisensory processing. *Trends Neurosci* 31:401–9.
- Stanford TR, Quessy S, Stein BE (2005): Evaluating the operations underlying multisensory integration in the cat superior colliculus. *J Neurosci* 25:6499–508.
- Stein BE, Meredith MA. 1993. *The merging of the senses*. Cambridge, MA: MIT Press.
- Sweet R a, Dorph-Petersen K-A, Lewis D a (2005): Mapping auditory core, lateral belt, and parabelt cortices in the human superior temporal gyrus. *J Comp Neurol* 491:270–89.
- Thut G, Nietzel A, Brandt S a, Pascual-Leone A (2006): Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *J Neurosci* 26:9494–502.
- Van Dijk H, Schoffelen J-M, Oostenveld R, Jensen O (2008): Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. *J Neurosci* 28:1816–23.
- Van Veen BD, van Drongelen W, Yuchtman M, Suzuki a (1997): Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Trans Biomed Eng* 44:867–80.
- Varela F, Lachaux J, Rodriguez E, Martinerie J (2001): The brainweb: phase synchronization and large-scale integration. *Nat Rev Neurosci* 2:229–239.
- Weisz N, Wühle A, Monittola G, Demarchi G, Frey J, Popov T, Braun C (2014): Prestimulus oscillatory power and connectivity patterns predispose conscious somatosensory perception. *Proc Natl Acad Sci U S A* 111:E417–25.
- Welch RB. 1999. Meaning, attention, and the “unity assumption” in the intersensory bias of spatial and temporal perceptions. *Advances in psychology*, 129: 371-387

APPENDIX

AT interactions have been shown to depend on many factors (see s 1.3). This behavioral study focuses on the investigation of their spatial properties. Its main goal was to obtain the most information possible for preparing an MEG study. In this perspective, we explored factors potentially critical for spatial interactions.

In particular, we investigated AT interactions in the peri-head space and examined their sensitivity to the spatial arrangement of the stimuli (i.e. the relative spatial position to which the A and T stimuli occur), we then analyzed how they are modulated by the distance between the auditory stimulus and the body (by using moving auditory stimuli) and by whether the auditory stimulus is perceived as approaching or receding.

Moreover, we designed an experiment that can be easily replicated in an MEG room, given that this experimental environment imposes many limitations. For example, we choose to spatially localize the auditory stimuli in a virtual manner so as to deliver the sound through headphones rather than utilizing spatially distributed loudspeakers.

Given their preliminary nature, these results are shown in the appendix rather than in the thesis' main body.

Introduction

Sensitivity to spatial arrangements between stimuli of different modalities is a debated topic in the field of multisensory perception. Neurophysiological studies in cats' Superior Colliculus (SC) found that multisensory integration is largest when stimuli are spatially aligned in their external positions (neurons' RFs overlap), while for spatially misaligned stimuli no facilitation or even response depression can be observed. This rule has long been considered as one of the three fundamental general principles (see s. 1.9), holding not only for direct recordings in SCs, but in the whole field of multisensory research. However, it is not yet clear to which extent it is possible to directly transpose these (and the others) multisensory principles, based on neural observations in SC to behavioral results or to neocortex recordings (see Spence 2013 for a review).

Contrary to what has been reported for other modalities pairings, a previous EEG study (Murray et al. 2005) failed to reveal any modulation, at behavioral and at cortical levels, of AT interactions in function of the spatial arrangement of the stimuli. In this study static auditory stimuli were presented aligned/misaligned (left/right) to a touch in the space of the hand. However, evidences already highlighted in this thesis (s 1.3) suggest that by presenting the stimuli in the space close to the head, different findings could be revealed. In fact, sensitivity to spatial arrangement for AT was shown in successive behavioral studies in the space of the head (Kitagawa et al., 2005; Tajadura-Jiménez et al., 2009). For these reasons we conducted a study on spatial arrangement between auditory and tactile modalities in the peri-head space. Moreover, we used different types of auditory stimuli, i.e. instead of static sounds, we used dynamic sounds. Dynamic sounds possess higher ecological valence and are particularly relevant (see s. 1.4) compared to sounds presented at fixed locations. In particular, as typology of sounds we used both

possible typologies of dynamic sounds: looming and receding. These two types of sounds have shown asymmetrical behavior being the approaching ones especially salient. Moreover, in the space of the head sensitivity to spatial arrangement was modulated by distance between the sound and the body. Interactions were more pronounced when the sound was presented near the body (20 cm) than far from it (70 cm). Dynamic auditory stimuli offer the advantage to investigate interactions along a continuum of perceived distances with respect to the body.

A similar behavioral study on AT interactions, i.e. using dynamic (receding and looming) sounds has been conducted by the group of Serino (Canzoneri et al. 2012). In this study a tactile stimulus was administered at the hand at different delays from the onset of an approaching or of receding sound. Their results showed that the moving auditory stimulus speeded up the processing of the tactile stimulus as long as the sound was perceived at a limited distance from the hand for both sounds. Moreover, when comparing looming vs receding sounds, the approaching one seemed to have a stronger effect in speeding up the processing of a tactile stimulus.

Here, we collected RTs to a tactile stimulus (while auditory stimuli were task-irrelevant). The touch has been presented to the participants' head (right perioral area). Moreover, touch was delivered at different time points with respect to the sound onset, thus when the moving (looming or receding) auditory stimulus was perceived at various distances from the body. Auditory stimuli were virtually manipulated so that the two sounds (looming or receding), although binaurally presented through headphones, could be perceived either as coming from the left side or from the right side, i.e. spatially congruent or incongruent with the tactile stimulus.

We expected that when the auditory stimuli are perceived close to the body, spatial arrangement of the stimuli affects RTs. Possibly, when stimuli are perceived in this space, also the typology of sounds, i.e. whether it is receding or approaching

may influence behavior. On the contrary, when stimuli are perceived far from the body we expected to not see behavioral differences between different spatial arrangements or sound types.

Methods

Subjects

Twenty healthy subjects (6 males, age mean \pm standard deviation 21.8 ± 1.5 years) participated in the study. All participants reported normal hearing and touch. All subjects gave their written informed consent to participate in the study, which was approved by the Ethical Committee of the University of Trento and was performed in accordance with the Declaration of Helsinki amended in October 2008. Participants received credits as a reimbursement in return for taking part in the study.

Audio stimuli

The sound used in the experiment was a mosquito sound (22050 Hz sampling rate, 16 bits resolution). The original sample can be found in the following webpage: <http://www.acoustics.org/press/132nd/3pab10.html>. The sound stimulus (~1.4 sec) was edited in duration and amplitude. The duration was truncated to 1 sec. The amplitude was attenuated or amplified in various parts of the sound stimulus in order to obtain a stimulation signal that was as much as possible constant in root mean square amplitude over time. For the experimental conditions testing the looming and the receding auditory stimuli, the sound equalized for stimulation energy was further modulated in amplitude in two different manners: with an exponential envelope increasing from 35 to 65 dB over the sound's duration, so to give an impression of an approaching sound (looming), or viceversa decreasing from 65 dB to 35 dB so to give the impression of a

receding sound. The two sounds (looming and receding) were filtered with HRTF coming from the CIPIC HRTF Database (Algazi et al., 2001). This database, freely available in internet (<http://interface.cipic.ucdavis.edu/sound/hrtf.html>), contains HRTF obtained from recordings in 50 individuals of the impulse responses measurements collected over the sphere surrounding individuals' heads at 25 different azimuths and 50 different elevations. Convolution of a sound with the binaural impulse responses obtained from the recording at a defined azimuth and elevation angle, the final sound used for the stimulation possess all the cues (ILT, IAD and others) representing the selected location. In particular, we used the two HRTF obtained from measurements from a KEMAR dummy head recorded respectively at an elevation of 34°, at an azimuth angle of 65° and at an elevation of 34° and an azimuth angle of -65° so that the sounds source appeared with respect to the subject's gaze direction at the same elevation but coming either from the left or right side. Thus, in total four types of auditory stimuli, representing the buzzing of a mosquito, were used: looming from the left/right side or receding from the left /right side. Sounds were delivered via headphones.

Tactile stimulus

Tactile stimulation was delivered through a customized device. A piston was moved back and forth by means of a double-acting pneumatic cylinder working with compressed air with a pressure of 0.7 bars and driven electrically by a valve via a custom-built relay box. The terminal part of an optic fiber system was attached to the piston. One fiber of the system was connected to an infra-red light and the other to a light sensor (Keyence series FS-N, Neu-Isenburg, Germany). Thus, approaching the skin during stimulation resulted in an increase of reflected light until the fiber has reached the skin and indicated the timing of the delivery of the touch to participants. The pneumatic system was placed outside the

experimental room, so that the noise produced by the compressed air device was not perceivable. The terminal part of the stimulation equipment containing the pneumatic valve was mounted on a metal pedestal (having no direct contact with the participant). The participants' head was placed on a chinrest in order to minimize head movement. Tactile stimulation was delivered always to the right peri-oral area. Prior to each block, the stimulator was adjusted so that the point and the intensity of the tactile stimulus was constant among blocks. We accurately controlled that participants executed the experimental task (see next section) only on the basis of the perceived touch, and not on the bases of other cues, such as the noise of the stimulator or vision of the moving tactile stimulus. To this aim, the participants wore foam earplugs under the headphones in order to cancel the audition of the weak sound produced by the tactile stimulator placed outside the experimental room. Moreover, participants' view of the final part of the tactile stimulator, the moving piston, was covered. In this way, the tactile stimulus was totally "invisible" and "silent".

Paradigm

During the experiment, on each trial, the participants listened to one of the four sounds described above, i.e. a looming or a receding sound coming from either the left or right side with respect to body midline. Since the tactile stimulus was always localized on the right peri-oral area, in half of the experimental conditions it was spatially congruent with the auditory stimulus ($A_{loom}T_{con}$, $A_{rec}T_{con}$), while in half of the experimental conditions tactile and auditory stimuli were spatially separated ($A_{loom}T_{inc}$, $A_{rec}T_{inc}$) (see Figure 4.1A). Along with the auditory stimulation, in the 90% of trials subjects received a tactile stimulus, as described above. In 10% of the trials, one of the four types of sounds was presented but no tactile stimulation was delivered. This type of trials was introduced in order to

minimize expectations about an upcoming tactile stimulus. Five different stimulus onset asynchronies (SOA) between the sound and the touch were used, i.e. the touch could occur either after 100 ms, 600 ms, 1100 ms, 1600 ms, or 2100 ms from sound onset (see Figure 4B). Since the a specific time during the presentation of the looming and receding sound carries subjective information about the distance of the sound source from the subject, in this way, the touch might occur when sounds were perceived at different locations (T1 and T2), or when the sound was already off. Each sound was preceded by a period of silence lasting randomly between 400 ms \pm 200 and followed by a random period of silence of 2700ms \pm 500. The SOAs and the type of sound could vary from trial to trial and were randomly presented, together with the condition of only sound, within a block. Each block consisted of 88 trials. 80 were “bimodal” trials (20 for each type of sound, 4 for each different SOAs) and 8 only sound (2 for each type of sound). Each participant underwent four blocks. The participants’ task was to respond as fast as possible to touch through bottom presses, while ignoring the co-occurring auditory stimulus (detection task). During the trial, participants were instructed to fixate on a cross on the display screen.

As additional measures of the efficacy of the stimuli, after the experiment each sound was presented ten times and behavioral ratings about various features of the sounds presented (i.e., pleasantness, remoteness, motion, arousal, power, distance, duration, ecological validity, threat) were collected.

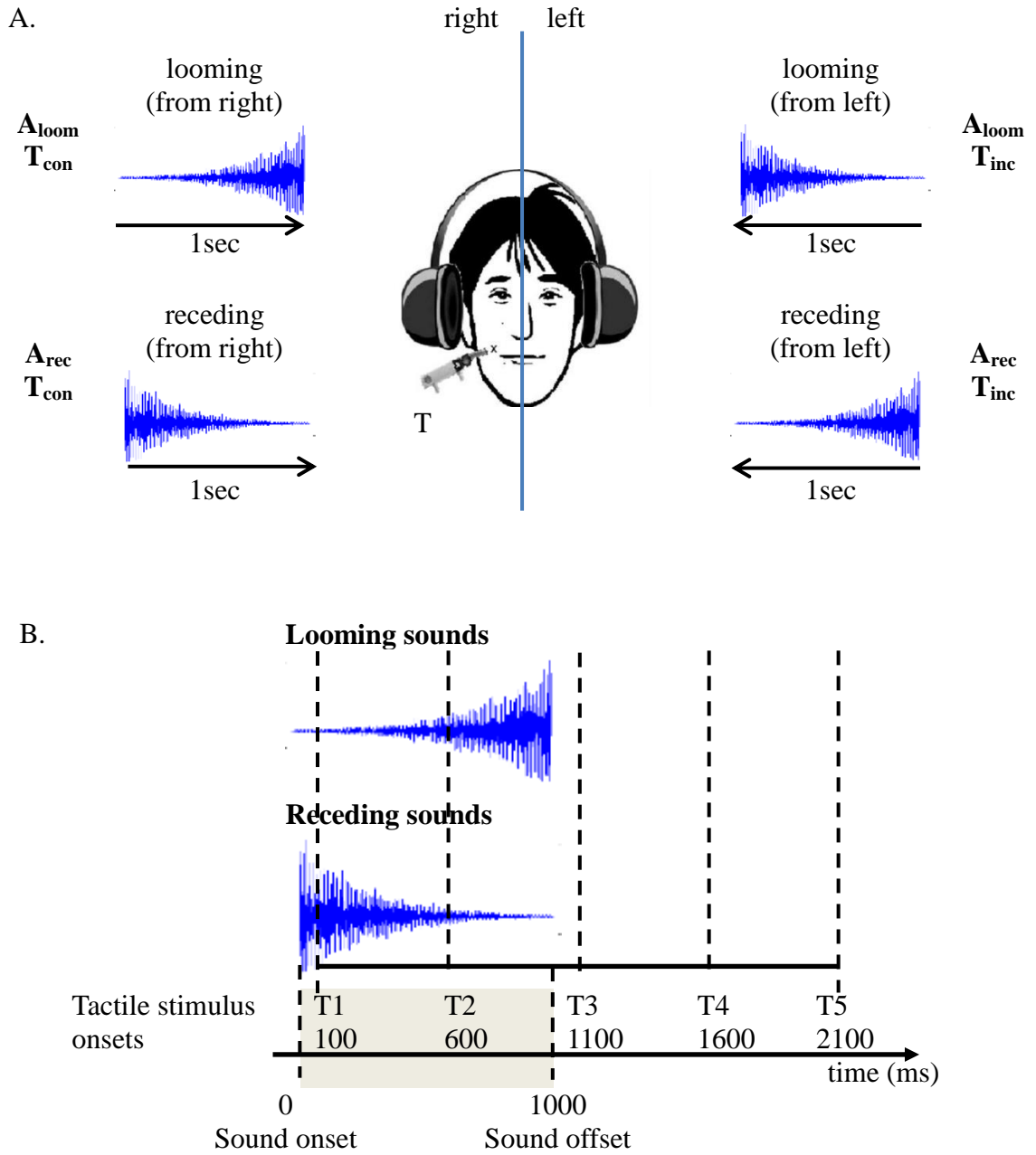


Figure 1.

A. Experimental setup. Subjects received a tactile stimulus at the right perioral area while task-irrelevant sounds either approached to or receded from the left or right side (4 auditory stimuli).

B. Procedure. From trial to trial tactile stimuli were delivered at different temporal delays from sound onset (from T1 to T5). In this the touch way occurred when sounds were perceived at a different location (T1 and T2), or when the sound was already off. Participants had to respond as fast as possible to touch, while they had to ignore the sound.

Results

Reaction times analysis

Unimodal auditory trials were discarded from the analysis. Mean RTs to tactile targets were calculated for every temporal delay, (i.e., T1 to T5), separately for $A_{loom}T_{con}$, $A_{loom}T_{inc}$, $A_{rec}T_{con}$, $A_{rec}T_{inc}$. RTs exceeding more than 2 standard deviations from the mean RT were considered outliers and eliminated from the analyses. The RTs to the tactile target for the different temporal delays at which the tactile stimulus was administered (from T1 to T5) and the different conditions (i.e., $A_{loom}T_{con}$, $A_{loom}T_{inc}$, $A_{rec}T_{con}$, $A_{rec}T_{inc}$) is represented in Figure 2.

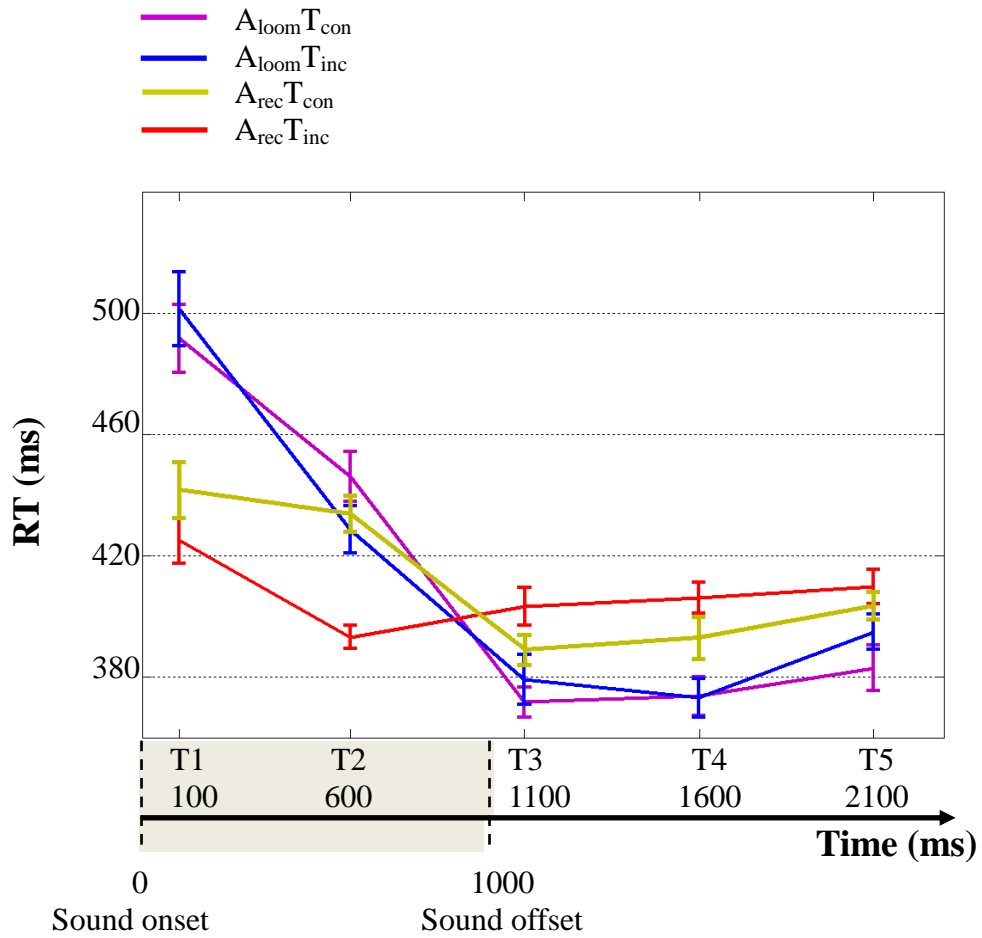


Figure 2

Mean RTs (and S.E.M.) to the tactile target at different temporal delays for the different experimental conditions. The light brown region indicates the duration of the sound

A three way repeated measure ANOVA on tactile RTs with the within subjects factors of Sound (loom, rec), Temporal Delay (T1, T2, T3, T4, T5) and Spatial Congruency (Tcon Tinc) was performed. A significant three-way Sound x Temporal Delay x Spatial Congruency interaction was found. Results are shown in Table 1. In order to investigate the different aspects of this 3way-interaction we performed two different post-hoc analyses.

Effect	DF	F	P
Sound	1	2,292	0,146
spatial congruency	1	0,317	0,58
temporal delay	4	43,525	0
sound x spatial congruency	1	1,911	0
sound x temporal delay	4	58,778	0,183
spatial congruency x temporal delay	4	9,835	0,005
sound x spatial congruency x temporal delay	4	6,424	0,02

Table 1: results of the three-way ANOVA

Post-hoc Analysis 1

The data were analyzed separately for each of the five Temporal Delays, therefore producing five distinct ANOVAs with two factors, Sound (A_{loom}/A_{rec}) and Space Congruency (T_{con}/T_{inc}).

For T1, T3 and T4 we found a main effect of Sound (T1: $F(1,32.14)$, $p < 0.0001$; T2: $F(1,20)$, $p < 0.00001$; T3: $F(1,20)$, $p < 0.00001$). For these temporal

delays, only the type of sound and not the spatial congruency between the moving sounds and tactile stimuli, influenced significantly the participants' responses. In particular, participants' responses were slower for looming than for receding sounds at T1 (respectively $494\text{ms} \pm 18$ vs $432\text{ms} \pm 14$, mean \pm SEM), while participants' responses were faster for looming than for receding sounds at T3 and T4 (respectively, T3: 374 ± 13 vs 393 ± 11 ; T4: 371 ± 13 vs 400 ± 14).

Furthermore, we found a significant Sound x Spatial Congruency interaction ($F(1,5.65)$, $p=0.028$) and a main effect of Sound ($F(1,9.126)$, $p<0.007$ uncorrected) at T2. In order to disentangle this result, two paired t-test Bonferroni corrected between A_{loom} and A_{rec} , with spatially congruent touches ($A_{\text{loom}}T_{\text{con}}$ vs $A_{\text{rec}}T_{\text{con}}$ $p = 0.151$, respectively 446 ± 17 and 434 ± 14) and spatially incongruent touches ($A_{\text{loom}}T_{\text{inc}}$, $A_{\text{rec}}T_{\text{inc}}$, $p= 0.003$ uncorrected, respectively 429 ± 14 and 393 ± 13) thus revealing that only for incongruent touches whether the sound is looming or receding influence RTs in T2 (one can argue that at this points sounds are more or less localized at the same distance). Space significantly influence RTs for receding sounds and $A_{\text{rec}}T_{\text{con}}$ vs $A_{\text{rec}}T_{\text{inc}}$ ($p<0.001$ uncorrected) but not for looming sounds ($A_{\text{loom}}T_{\text{con}}$ vs $A_{\text{loom}}T_{\text{inc}}$ $p=0.013$ uncorrected).

For T5 we found significant main effects for Sound ($F(1,15.46)$, $p=0.001$) and Space ($F(1, 11.354)$, $p=0.003$). Two paired t-test Bonferroni corrected were conducted between the factor Sound ($A_{\text{loom}}T_{\text{con}}$ vs $A_{\text{rec}}T_{\text{con}}$, $p=0,001$ uncorrected; $A_{\text{loom}}T_{\text{inc}}$ vs $A_{\text{rec}}T_{\text{inc}}$, $p=0,038$ uncorrected) revealed sound being a significant factor. Interpretation of the main effect space ($A_{\text{loom}}T_{\text{con}}$ vs $A_{\text{loom}}T_{\text{inc}}$, $p= 0.02$ uncorrected; $A_{\text{rec}}T_{\text{con}}$ vs $A_{\text{rec}}T_{\text{inc}}$ $p= 0.2$ uncorrected) shows how spatial factor is significant only in the case of looming sound .

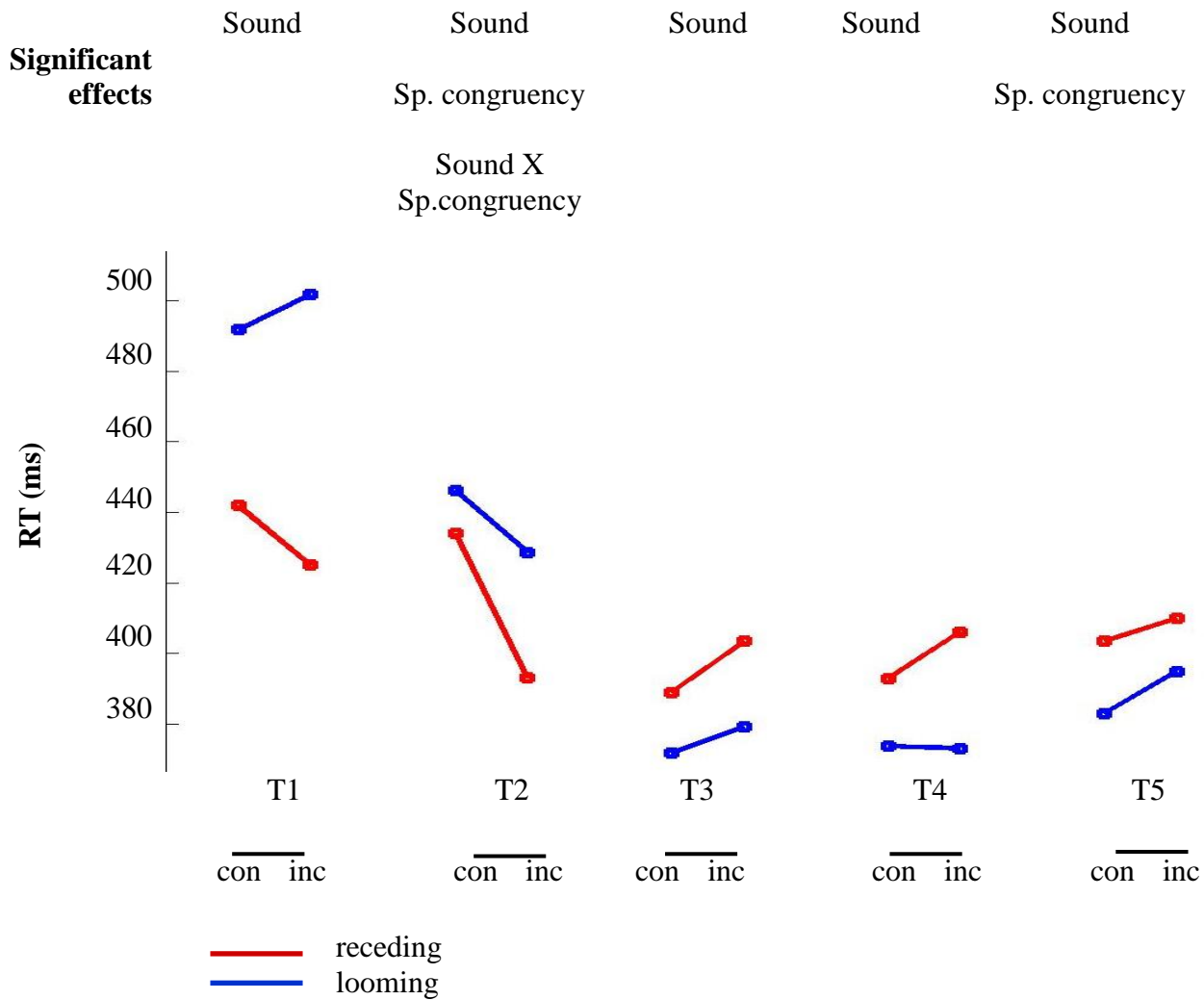


Figure3

The plot presents the mean of the RTs for spatially congruent or incongruent for all five different time delays, for looming (blu lines) and receding (red lines) sounds. The results of the two-ways ANOVAs on single time delay made to explore the three-way interaction are shown above each temporal delay.

Post-Hoc Analysis 2

In order to analyse further aspects of the interaction found in the three-way ANOVA, we performed two separate ANOVAs for each of the two sounds (looming or receding), with two factors, Temporal Delays and Space Congruency.

The ANOVA between $A_{loom}T_{con}$ and $A_{loom}T_{inc}$ revealed a main effect of Temporal Delay ($F(4,5.8), p < 0.0001$). Newman-Keuls post-hoc tests showed that RTs at T1 (when sounds were perceived far from the body; mean RTs ms, S.E.M.) was significantly slower compared to RTs at T2, T3, T4 and T5 (all $ps < 0.005$). T2 was significantly slower than T3, T4 and T5 (all $p < 0.05$). No statistically significant differences were revealed between T3 and T4 and T5 as well as between T4 and T5.

The ANOVA between $A_{rec}T_{con}$ and $A_{rec}T_{inc}$ revealed an interaction SpatialCongruency x Temporal Delay ($F(4,4.91), p < 0.00001$) and a main effect of Temporal Delay ($F(4,8.8), p < 0.00001$). To analyze this interaction we performed two separate ANOVA $A_{rec}T_{con}$ and $A_{rec}T_{inc}$ with Temporal Delays as within-subjects factor.

For $A_{rec}T_{con}$ a significant main effect of Temporal Delay was revealed ($F(4,12), p < 0.0001$). Newman-Keuls post-hoc tests showed that T1 and T2 were significantly slower than T3, T4 and T5 (all $ps < 0.05$). No statistically significant difference was revealed between T1 and T2 and between T3, T4 and T5.

Subjective ratings results

Each sound was rated by the participants according to various dimensions (see Section of Methods) after the experimental session. For each rated dimension, an ANOVA between the four typologies of sound was performed and, when significant, post-hoc paired comparisons were calculated (Bonferroni corrections). Whereas the four sounds were not evaluated as differing for distance at their

Appendix

beginning, they were perceived as significantly differing for the other features (see Table 1). $A_{loom}T_{inc}$ was evaluated significantly more pleasant than $A_{loom}T_c$ and $A_{rec}T_{inc}$ and significantly more powerful than $A_{rec}T_{inc}$. Also, $A_{loom}T_{inc}$ was considered as significantly more moving than both RC and $A_{rec}T_{inc}$ and more ecological than $A_{rec}T_{inc}$. $A_{rec}T_{inc}$ was evaluated as the least arousing sound. $A_{loom}T_{con}$ was considered as significantly closer at the end than both RC and $A_{rec}T_{inc}$ and more threatening than $A_{rec}T_{inc}$. L sounds were also considered as longer than R sounds. The participants were also asked to indicate the direction of motion of each sound on a template (see Table3). The results indicate that C sounds were localized as more right-lateralized than INC sounds (see also Figure xx).

Table 1. Explicit subjective ratings.

* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$.

<i>Dimension</i>	<i>Mean±SD subjective ratings</i>		<i>ANOVA</i>	<i>Post-hoc paired comparisons</i>			
Valence (pleasant — unpleasant)	$A_{loom}T_{con}$	4.39±1.42	p=.003 **	$A_{loom}T_{con}$ vs $A_{loom}T_{inc}$	p<.05*	$A_{loom}T_{inc}$ vs $A_{rec}T_{con}$	p=.06
	$A_{loom}T_{inc}$	5.44±1.10		$A_{loom}T_{con}$ vs $A_{rec}T_{con}$	p=1.00	$A_{loom}T_{inc}$ vs $A_{rec}T_{inc}$	p=.002**
	$A_{rec}T_{con}$	4.50±1.29		$A_{loom}T_{con}$ vs $A_{rec}T_{inc}$	p=1.00	$A_{rec}T_{con}$ vs $A_{rec}T_{inc}$	p=1.00
	$A_{rec}T_{inc}$	4.28±1.18					
Power (weak — strong)	$A_{loom}T_{con}$	4.33±1.19	p=.002 **	$A_{loom}T_{con}$ vs $A_{loom}T_{inc}$	p=.53	$A_{loom}T_{inc}$ vs $A_{rec}T_{con}$	p=.64
	$A_{loom}T_{inc}$	4.78±1.11		$A_{loom}T_{con}$ vs $A_{rec}T_{con}$	p=1.00	$A_{loom}T_{inc}$ vs $A_{rec}T_{inc}$	p=.001** *
	$A_{rec}T_{con}$	4.22±1.56		$A_{loom}T_{con}$ vs $A_{rec}T_{inc}$	p=.17	$A_{rec}T_{con}$ vs $A_{rec}T_{inc}$	p=.27
	$A_{rec}T_{inc}$	3.61±1.14					

Appendix

Arousal (calming — activating)	$A_{loom}T_{con}$	5.22±.88	p<.001 ***	$A_{loom}T_{con}$ vs $A_{loom}T_{inc}$	p=.29	$A_{loom}T_{inc}$ vs $A_{rec}T_{con}$	p=.86
	$A_{loom}T_{inc}$	5.61±.85		$A_{loom}T_{con}$ vs $A_{rec}T_{con}$	p=1.00	$A_{loom}T_{inc}$ vs $A_{rec}T_{inc}$	p=.002**
	$A_{rec}T_{con}$	5.11±1.08		$A_{loom}T_{con}$ vs $A_{rec}T_{inc}$	p=.006* *	$A_{rec}T_{con}$ vs $A_{rec}T_{inc}$	p=.004**
	$A_{rec}T_{inc}$	4.11±.90					
Motion (static — high motion)	$A_{loom}T_{con}$	5.39±1.33	p=.001 ***	$A_{loom}T_{con}$ vs $A_{loom}T_{inc}$	p=1.00	$A_{loom}T_{inc}$ vs $A_{rec}T_{con}$	p=.007**
	$A_{loom}T_{inc}$	5.61±.85		$A_{loom}T_{con}$ vs $A_{rec}T_{con}$	p=.41	$A_{loom}T_{inc}$ vs $A_{rec}T_{inc}$	p=.004**
	$A_{rec}T_{con}$	4.61±1.54		$A_{loom}T_{con}$ vs $A_{rec}T_{inc}$	p=.18	$A_{rec}T_{con}$ vs $A_{rec}T_{inc}$	p=1.00
	$A_{rec}T_{inc}$	4.22±1.63					
Remoteness Begin (far — near)	$A_{loom}T_{con}$	3.67±1.61	p=.79				
	$A_{loom}T_{inc}$	3.39±1.75					
	$A_{rec}T_{con}$	3.83±1.76					
	$A_{rec}T_{inc}$	3.67±1.81					
Remoteness End (far — near)	$A_{loom}T_{con}$	5.22±2.24	p=.001 ***	$A_{loom}T_{con}$ vs $A_{loom}T_{inc}$	p=1.00	$A_{loom}T_{inc}$ vs $A_{rec}T_{con}$	p=.23
	$A_{loom}T_{inc}$	4.89±2.03		$A_{loom}T_{con}$ vs $A_{rec}T_{con}$	p=.04*	$A_{loom}T_{inc}$ vs $A_{rec}T_{inc}$	p=.17
	$A_{rec}T_{con}$	3.50±1.65		$A_{loom}T_{con}$ vs $A_{rec}T_{inc}$	p=.02*	$A_{rec}T_{con}$ vs $A_{rec}T_{inc}$	p=.1.00
	$A_{rec}T_{inc}$	3.50±1.82					
Ecology (artificial — natural)	$A_{loom}T_{con}$	4.83±1.89	p=.005 **	$A_{loom}T_{con}$ vs $A_{loom}T_{inc}$	p=1.00	$A_{loom}T_{inc}$ vs $A_{rec}T_{con}$	p=.45
	$A_{loom}T_{inc}$	5.00±1.37		$A_{loom}T_{con}$ vs $A_{rec}T_{con}$	p=1.00	$A_{loom}T_{inc}$ vs $A_{rec}T_{inc}$	p=.04*
	$A_{rec}T_{con}$	4.33±1.61		$A_{loom}T_{con}$ vs $A_{rec}T_{inc}$	p=.12	$A_{rec}T_{con}$ vs $A_{rec}T_{inc}$	p=.17
	$A_{rec}T_{inc}$	3.89±1.53					
Thread (low — high)	$A_{loom}T_{con}$	4.44±1.76	p=.001 ***	$A_{loom}T_{con}$ vs $A_{loom}T_{inc}$	p=.69	$A_{loom}T_{inc}$ vs $A_{rec}T_{con}$	p=.06
	$A_{loom}T_{inc}$	5.17±1.47		$A_{loom}T_{con}$ vs $A_{rec}T_{con}$	p=.44	$A_{loom}T_{inc}$ vs $A_{rec}T_{inc}$	p=.44

Appendix

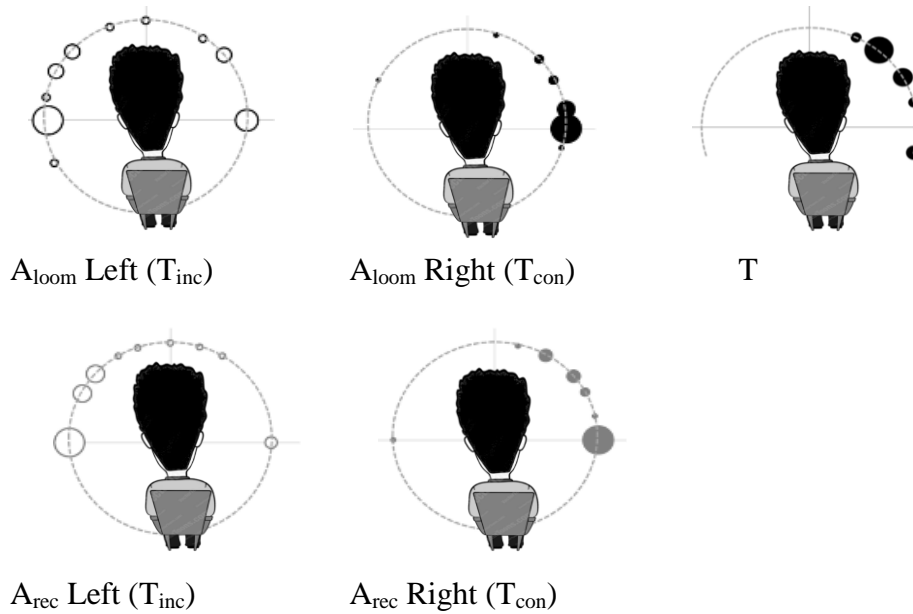
	$A_{rec}T_{con}$	3.72±1.56		$A_{loom}T_{con}$ vs $A_{rec}T_{inc}$	p<.05*	$A_{rec}T_{con}$ vs $A_{rec}T_{inc}$	p=1.00
	$A_{rec}T_{inc}$	3.39±1.50					
Duration (short — long)	$A_{loom}T_{con}$	4.17±.99	p<.001 ***	$A_{loom}T_{con}$ vs $A_{loom}T_{inc}$	p=.24	$A_{loom}T_{inc}$ vs $A_{rec}T_{con}$	p<.001** *
	$A_{loom}T_{inc}$	4.94±1.26		$A_{loom}T_{con}$ vs $A_{rec}T_{con}$	p=.002* *	$A_{loom}T_{inc}$ vs $A_{rec}T_{inc}$	p<.001** *
	$A_{rec}T_{con}$	3.28±.96		$A_{loom}T_{con}$ vs $A_{rec}T_{inc}$	p=.02*	$A_{rec}T_{con}$ vs $A_{rec}T_{inc}$	p=1.00
	$A_{rec}T_{inc}$	3.28±1.07					
Direction of motion	$A_{loom}T_{con}$	4.44±2.57	p<.001 ***	$A_{loom}T_{con}$ vs $A_{loom}T_{inc}$	p=.001* *	$A_{loom}T_{inc}$ vs $A_{rec}T_{con}$	p=.02*
	$A_{loom}T_{inc}$	-1.44±4.71		$A_{loom}T_{con}$ vs $A_{rec}T_{con}$	p=.55	$A_{loom}T_{inc}$ vs $A_{rec}T_{inc}$	p=.87
	$A_{rec}T_{con}$	3.61±2.97		$A_{loom}T_{con}$ vs $A_{rec}T_{inc}$	p<.001 **	$A_{rec}T_{con}$ vs $A_{rec}T_{inc}$	p=.003**
	$A_{rec}T_{inc}$	-2.17±3.85					

Table 2. *Explicit subjective ratings.*
*p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001.

	<i>Dimension</i>	<i>Mean±SD subjective ratings</i>
Tactile	Valence (pleasant — unpleasant)	4.11±1.23
	Power (weak — strong)	3.33±1.14
	Arousal (calming — activating)	4.33±1.46
	Ecology (artificial — natural)	2.61±1.65
	Thread (low — high)	3.61±1.72
	Location	4.17±1.65

Table 3. *Explicit subjective of stimuli position.*

Participants were requested to judge on a circle the perceived position of each stimulus



Discussion

In literature AT multisensory interactions have been shown to be sensitive to a number of features all to some extent related to space: they are sensitive to the typology of sound (whether a moving sound is approaching or receding the participants), to the distance between the auditory stimulus and the body, to the part of the body to which tactile stimulus is delivered and finally they may be sensitive to their relative spatial arrangement.

Given the poor sensitivity of AT interactions to spatial arrangements in the space of the hand (e.g. Zampini et al. 2005; Murray et al. 2005), we

conducted our analysis in the peri-head space, a portion of space that has been shown to be especially sensitive for AT interactions. In this space, we investigated how AT interactions are modulated by a moving auditory stimulus, as dependent from the fact whether this approaches (looming) or recedes the participants' head and whether the sound is perceived close or far from the body. We used approaching and receding sound, since these have been shown to exert significantly different reactions, with the approaching sounds being more salient and activating than the receding (Bach et al., 2009; Tajadura-Jiménez et al., 2010, see also section 1.4). The rationale behind using dynamic stimuli is that we could investigate AT interaction around the head at different positions: by delivering the tactile stimulus at five different time points (T1-T5) with respect to the moving sound, we modified the perceived distance between the auditory stimulus and the body (T1 and T2) and if the effect persisted after auditory stimulus offset (T3 to T5).

By using auditory stimuli oriented in space (looming/receding from left/right), while the tactile stimulus was always placed to the right, we studied the impact of spatial overlap between the stimuli.

Our results (analysis 1) show that the type of sound, i.e. looming or receding, modulates RTs to the tactile stimulus for each time point we tested (though for T2 it interacts with space). Remarkably, receding sounds elicit faster RTs than looming sounds to a tactile stimulus for early temporal delay (T1 and T2), i.e. when the receding sound is perceived closer to the head than the looming sound. For later time points (T3-T5), a reversal occurs, i.e. looming sounds elicit faster RTs than receding sounds. These results confirm the findings reported in previous AT literature (Canzoneri et al. 2012). Interestingly the effect of the type of sound persists also after a sound offset of more than one second, though after sound offset RTs remain constant within condition. Contrary to this, Canzoneri et al. (2012) found RTs following receding sounds, after a first decrease, to return to RT's baseline values (i.e. prior to sound onset). However, this discrepancy may be

due to substantial different paradigms, being the time points at which the tactile stimuli were delivered in Canzoneri's experiment spread over a time window of almost 5 seconds (sounds lasting 3 seconds) against the time window used in the present study of 2.1 seconds (and a sound of 1 second).

Here, when comparing RTs within the different temporal delays (analysis 2) our results showed that looming sound elicited a behavior which is identical, regardless of the different spatial alignment of the stimuli: a decrease in RTs between T1, T2 and T3 while RTs remain stable for successive time points. On the contrary, when it comes to receding sounds, the spatial congruency between the sound and the tactile stimulus affected behavior. In particular, when there is spatial congruency, i.e. the auditory stimulus is receding from the same side of the head where the tactile stimulus is delivered, no RTs' difference is observed between the first two time points (T1 and T2), when the auditory stimulus is still close to the head. The only RTs' decrease is observed after T2. When the auditory stimulus is receding from the opposite side with respect to the side where the tactile event appears, a decrease of RTs occurred between T1 to T2. This dependency on spatial congruency was found only for receding sounds, only for T2, when the sound is localized at an intermediate distance with respect to the body. Surprisingly, at this point RTs are significantly faster for incongruent stimuli (there is a similar trend also in the looming sounds). This result suggest that, at least for receding stimuli, the auditory stimulus on the same side of the tactile stimulus interfere rather than facilitate detection of the tactile stimulus.

A similar pattern of interference is the well-known "Crossmodal congruency effect" (CCE) paradigm. This behavioral task has been used to study the multisensory representation of space. For example, in one of its applications Spence et al. (2000) made a speeded up/down discrimination task for a vibrotactile stimulus, i.e. tactile targets as well as visual distracters were presented to one of two different locations on the hand in a congruent way (same elevation, up/down)

or incongruent (different elevation) while ignoring the visual distractor (that could appear on the same or opposite side). Participants responded significantly slower when the elevation of the vibrotactile distractor was incongruent with that of the visual target than when they were presented from the same (i.e., congruent) elevation. This interference effect was larger when the visual distractor appeared on the same side of space as the vibrotactile target than when they appeared on different sides.

Our results show a modulation of RTs for different spatial arrangements, although this occurred only in the case of receding sounds. In this case, it showed shorter RTs for spatially incongruent, rather than for spatially congruent, AT stimuli. On the contrary, for looming sounds no spatial modulation was observed.

We expected that looming sounds, being more salient, would exhibit enhanced spatial sensitivity so as to exploit this trait to investigate its neural correlates in the MEG, but this did not occur.

In this study, the tactile stimulus was presented always on the right, while the auditory stimuli were virtually localized. A major confounding point was that, as shown in Table 3, virtual localizations of sounds were imprecisely perceived by participants. Furthermore, it has been shown that multisensory enhancement depends on task requirements, e.g. in a visuotactile experiment Girard and colleagues (2010) observed that multisensory integration is sensitive to spatial arrangement of the stimuli only when spatial information is task-relevant, while when spatial information is task-irrelevant, multisensory integration of spatially aligned and misaligned stimuli is equivalent. Provided that only auditory stimuli were differently located in space while the tactile stimulus, i.e. the task-relevant stimulus, was always presented to the right, it might be that auditory spatial information is used in a limited way in executing the task.

Although several results show the prominence of the space of the head for AT interactions, inside this space they are more pronounced in the space behind

(e.g., Farné and Làdavas, 2002, Kitagawa et al. 2005). Although it must be noticed that our stimuli were presented in an area (close to the lips) where vision does not have access, different portions of the peri-head space could be tested by presenting auditory stimuli from behind the head.

Additional references used in the appendix

- Girard, S., Collignon, O., Lepore, F., 2011. Multisensory gain within and across hemispaces in simple and choice reaction time paradigms. *Exp. brain Res.* 214, 1–8. doi:10.1007/s00221-010-2515-9
- Spence, C., Pavani, F., Driver, J., 2000. Crossmodal Links Between Vision and Touch in Covert Endogenous Spatial Attention. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 1298–1319.
- Spence, C., 2013. Just how important is spatial coincidence to multisensory integration? Evaluating the spatial rule. *Ann. N. Y. Acad. Sci.* 1296, 31–49. doi:10.1111/nyas.12121