Exploring visuo-tactile interactions through simple and complex visual stimuli using Fast Periodic Stimulation

Roxane De Keyser

Thèse présentée en vue de l'obtention du grade de docteur en Sciences biomédicales et pharmaceutiques

Promoteur

Valéry Legrain

Exploring visuo-tactile interactions through simple and complex visual stimuli using Fast Periodic Stimulation

Roxane De Keyser

Thèse présentée en vue de l'obtention du grade de docteur en Sciences biomédicales et pharmaceutiques

Promoteur

Valéry Legrain

Comité d'encadrement

Alexandre Zenon (Université de Bordeaux)
André Mouraux (UCL)
Bruno Rossion (UCL)

Président

André Mouraux (UCL)

Comité d'accompagnement et jury

Alexandre Zenon (Université de Bordeaux)
Valérie Goffaux (UCL)
Durk Talsma (Ghent University)
Emiliano Macaluso (Université Lyon 1)

Ce travail est l'aboutissement de quatre années de recherche, parsemé de nombreux défis. Il me sera très difficile de remercier tout le monde car c'est grâce à l'aide de nombreuses personnes que j'ai pu mener cette thèse à son terme.

Je tiens à remercier le professeur Valéry Legrain pour m'avoir donné la possibilité de me lancer dans ce travail de thèse. Je tiens également à remercier le professeur André Mouraux pour toujours avoir pris le temps pour répondre agréablement à toutes mes questions approfondies lié aux analyses EEG et signal EEG

Je remercie également ceux qui ont accepté de participer à tous mes comités d'encadrement : le professeur Bruno Rossion, le professeur Alexandre Zenon et le professeur André Mouraux. Merci pour votre participation scientifique ainsi que le temps qu'ils ont consacré à ma recherche.

Je remercie également le professeur Durk Talsma, le professeur Valérie Goffaux et le professeur Emiliano Macaluso pour avoir accepté d'etre dans mon jury de thèse, pour le temps qu'ils ont consacré à la lecture critique de mon travail et à leur contribution à l'amélioration de cette thèse.

Je m'est impossible d'oublier de remercier particulièrement Diana pour toutes nos discussions et ses conseils qui m'ont accompagné tout au long de ce travail.

Un tout grand merci à toute l'équipe nocions et Cécile, qui ont assuré une ambiance de travail excellente tout au long de ces quatre années. J'ai eu l'honneur de travailleur dans une équipe exceptionnelle, qui étaient toujours disponible pour discuter de recherche (et de non recherche à l'ad hoc). Je les remercie pour leur soutien qui m'a été précieux afin de mener ce travail à bon port.

Je tiens particulièrement à remercier les *Valerettes*, pour leur amitié et leur bonne humeur. Ce fut un plaisir de partager les bons et moins bons moments avec vous. Je tiens surtout à remercier Lieve, pour toute l'aide que j'ai reçu, tout au long de ces quatre années. Merci pour tes conseils et corrections précieuses dans l'écriture de ma thèse.

En dernier, je tiens à remercier mes colocs au fil de ces quatre dernières années, Gax, Mo, Tine et Caro, d'avoir supporté mes bonnes et mauvaises humeurs. Un merci particulier à Martijn, qui a été tellement patient et rassurant les derniers mois de ce travail.

Table of contents

List of abbreviations

AIP	anterior intraparietal cortex
BOLD	Blood oxygen level dependent
EEG	Electroencephalogram
ERP	event-related potential
fMRI	functional magnetic resonance imaging
FPS	Fast periodic stimulation
FPVS	Fast periodic visual stimulation
GBO	Gamma band oscillations
IPL	Inferior parietal gyrus
IPS	anterior intraparietal sulcus
MEP	Motor evoked potential
MT	middle temporal
LED	Light emitting diode
PMd	Dorsal premotor cortex
PMv	Ventral premotor cortex
PPC	Posterior parietal cortex
PZ	polysensory zone
S1	primary sensorimotor cortex
SC	superior colliculus
SMG	Supramarginal gyrus
SSEP	Steady state evoked potential
SSVEP	Steady state visual evoked potential
TOJ	Temporal order judgement
TMS	Transcranial magnetic stimulation

V1	primary visual cortex
VIP	Ventral intraparietal cortex

GENERAL INTRODUCTION & AIMS

General Introduction

In our everyday life, we are constantly confronted with sensory events occurring in the space close to our body, to which we can or need to react to. As such, studying how we process those stimulations has been of great interest. While research initially focussed on the processing of unimodal stimuli, i.e. the processing of one sensory modality at the time, it has shifted to the processing of multisensory events, i.e. events which can be coded in multiple sensory modalities. Indeed, objects in our surroundings are almost always processed through different sensory modalities, i.e. if we see a wasp, and feel it on our hand, we perceive it as a unique integrated perception. To achieve this, it appears that some cognitive processes are specifically devoted to integrate the different sensory inputs to build a coherent representation of our environment. This representation of the body in space is important for monitoring the space surrounding the body, for action or defensive purposes.

Information originating from a common source facilitates perception and processing. For example, we can easily identify the person speaking when hearing a voice, even in a crowded environment, by matching visual and auditory information (Arons, 1992). Similarly, it is not a problem to coordinate our actions to chase away an insect on our body, independently of the position of our body in space. Behaviourally, this facilitation was demonstrated by faster reaction times for multimodal stimulation that are presented at the same location as compared to stimulation that are not congruent in space (Driver, 1996; Kennett, Eimer, Spence, & Driver, 2001; Kennett, Taylor-Clarke, & Haggard, 2001; Teder-Sälejärvi, McDonald, Di Russo, & Hillyard, 2002). Also, congruent multimodal stimulation modulated evoked brain potentials (Eimer & Driver, 2000; Eimer, Velzen, & Driver, 2002; Kennett, Spence, & Driver, 2002), as well as metabolic signal (Brozzoli,

Gentile, Petkova, & Ehrsson, 2011; Huang, Chen, & Sereno, 2018; Macaluso & Driver, 2001; Macaluso, Frith, & Driver, 2000), i.e. it has been demonstrated that paying attention to visual stimuli can influence the processing of tactile stimuli when they are located at the same location in space. Consequently, information from different modalities appear to interact under certain conditions.

The interaction between different sensory modalities complies to three main rules: the spatial rule, which states that it is more likely to observe an interaction between two stimulations if they occur close in space (Meredith & Stein, 1996; Spence, 2013); the temporal rule, which states that it is more likely for two stimulation to interact if they occur close in time (Meredith, Nemitz, & Stein, 1987), and the rule of inverse effectiveness, which states that two stimulations are more likely to interact if evoking weak activity separately (Holmes, 2007).

Different mechanisms can explain this interaction between the different sensory modalities: multisensory integration, crossmodal interaction between primary sensory cortices, or a crossmodal link in spatial attention.

Multisensory integration can be defined as a high-level phenomenon where information from different sensory modalities converges in higher order areas to form one coherent multisensory representation of space. Multisensory integration has been extensively studied in experimental and controlled settings: evidence for multisensory integration ranges from single cell recordings in cats and monkeys to modulation of brain signals in humans (Bremmer, Schlack, Shah, et al., 2001; Brozzoli et al., 2011; Colby, Duhamel, & Goldberg, 1993; Huang et al., 2018; Meredith & Stein, 1983; Wallace, Meredith, & Stein, 1992). As such, neurons responding to multiple modalities (called bimodal neurons, or multisensory neurons) have been discovered in several brain structures, i.e. superior colliculi,

parietal and prefrontal cortex, indicating that information from unimodal areas is integrated in those areas.

Crossmodal interaction points to mechanism where information from one modality will influence the processing of a different sensory modality, often at early processing sites. Examples from this phenomenon is the fact that a tactile event can modulate the processing of visual stimulations, in the visual cortex (Macaluso et al., 2000; Macaluso, Frith, & Driver, 2002). While the term crossmodal interactions suggests a direct interaction between primary sensory cortices, it is possible that this modulation also occurs through feedback connection from multisensory cortices to primary sensory cortices. Multisensory integration and crossmodal interactions are therefore not mutually exclusive.

A third possibility to explain the interaction between sensory modalities is the existence of a crossmodal link in spatial attention. Indeed, a stimulation in a certain modality can attract attention to its spatial location, simultaneously facilitating processing of different sensory information at the same location (Spence, 2010). For example, when presenting a cue in the tactile modality in one side of space, a visual stimulus presented at the same location, will be processed differently. This is evidenced by faster reaction times and modulations of deflections in the event-related potentials [ERPs] (Eimer & Driver, 2000; Kennett, Eimer, et al., 2001; Kennett et al., 2002).

Different mechanisms could thus underlie this interaction. However, all mechanisms have to overcome certain issues to be able to form a coherent representation of the space around us. First of all, primary processing of sensory information occurs according to anatomic coordinate systems, which can differ between sensory modalities and is defined by the anatomical projections from sensory receptive fields to distinct groups of sensory neurons in the cortex**.** As

such, visual stimuli are initially mapped according to where they are projected on the retina (Tootell, Hadjikhani, Mendola, Marrett, & Dale, 1998). Conversely, touch is mapped according to its localization on the body (Roux, Djidjeli, & Durand, 2018). Ultimately, the spatial coordinates of visual and tactile information need to be transformed to be mapped in a common frame of reference. This frame of reference allows taking the surrounding environment and the position of the body into account to enable action. This final representation is referred to as the peripersonal space, which includes the space of the body and the space near the body. The peripersonal space can refer to different concepts. Indeed, it has been used to designate the representation of the space around the body, in which we can reach. On the other hand, the peripersonal space can also refer to the multisensory representation of the body and the space around the body. Studying the latter interpretation of the peripersonal space is especially effective through vision and touch, as those stimuli associate the space of the body and the space near the body. It is critical to note that throughout this thesis, we use the term peripersonal space to label the representation of the body and the adjacent space. Importantly, this interpretation allows for representations centred on different limbs i.e. the peri-hand space

The spatial organisation of this remapping has been studied using transient stimuli. However, real life events usually last longer than a brief flash, and often occurs in a noisy environment. Additionally, sustained stimulations are a crucial for feedback signalling about ongoing actions. Ecologically, the brain has to select all relevant information, discard information that does not belong to the same source, and group it together to reach a stable perception. Consequently, it is timely to study the interaction between the tactile and visual modality using multiple sustained trains of stimulations simultaneously.

There is an ongoing debate on how to differentiate between multisensory integration, crossmodal interaction and a crossmodal link in spatial attention, if possible at all. Using Fast Periodic Stimulation [FPS] (also referred to as Steady-State Evoked Potentials [SSEP]), it becomes possible to present multiple sustained trains of stimulation simultaneously, and to disentangle the corresponding neural responses. FPS consists of a relatively fast periodic repetition of stimulations, whose activity is reflected at exactly the same frequency in the frequency spectrum of the electroencephalogram [EEG]. Additionally, a measure reflecting multisensory integration has been proposed using FPS: crossmodulation frequencies, which is believed to reflect the integration of two stimulation streams. Accordingly, FPS can be used to study the mechanisms underlying visuo-tactile interactions, and their spatial organisation, overcoming some methodological shortcomings of current research. Consequently, this thesis is articulated around the question whether it is possible to observe interaction between visual and tactile inputs in the neural signal using FPS. If this appears to be the case, the question evolves to by which spatial configuration visual and tactile stimulation interact to represent the surrounding space.

This present work consists of a first introductory chapter, which will articulate the research question in a more extensive manner. Second, we will proceed to the general outline of the empirical chapters and the associated aims and hypothesises. This will be followed by two empirical chapters. The first one aims at replicating results frequently found in the literature, stating that an interaction between visual and tactile stimulation exists, using FPS. A second empirical chapter aims at separating low-level from higher-order visual activity, using images of tools and then aims at demonstrating an interaction between tactile processing and higher-order visual processing. In a last part, we will discuss the results, introduce future studies and formulate concluding remarks.

We will first introduce concepts and literature relevant to the field of visuo-tactile interactions, while highlighting the gaps giving rise to the present thesis. While it has been established that multimodal stimulation can facilitate perception and processing, it is not fully understood along which spatial organization this occurs using sustained stimulations. As such, we will go over the frames of references necessary to fully understand the possible organization of visuo-tactile interactions. We will then concentrate on the three principles which have been proposed to govern visuo-tactile interactions. Subsequently, we will discuss animal evidence for the spatial organisation of visuo-tactile interaction. We will end the first section of the introduction discussing human studies focussing on this visuo-tactile remapping in the peripersonal space. The second section of this chapter will focus on the specific object category of tools and why this category is relevant for visuo-tactile interactions in the peripersonal space. Third, we will concentrate on the FPS response, which is the electrophysiological outcome measure throughout this thesis. Its use in multisensory integration research, advantages and origins will be discussed subsequently. Last, we will discuss the aim and outlines of the experiments.

1. Multisensory integration, crossmodal interaction and spatial attention

1.1 Frames of reference for spatial perception

The space in which we can act is easily perceived as a unified construct, where information from different senses converges to form a coherent representation of the surrounding space. In the brain, space is a complex

construct, where the spatial aspects of incoming information can be processed in different ways (Cléry, Guipponi, Wardak, & Ben Hamed, 2015). Importantly, sensory information needs to be integrated to interact with our environment, and to guide our actions successfully, such as to avoid or manipulate an object (Colby, 1998; Colby & Duhamel, 1996; Fogassi et al., 1992). Consequently, depending on the different behaviours that can be performed, the spatial position of sensory information should be processed differently. Indeed, it has been demonstrated that the spatial position of a stimulation can be mapped according to different spatial frames of reference (Colby, 1998; Colby & Duhamel, 1996; Gross & Graziano, 1995; Previc, 1998; Vallar & Maravita, 2009), which are spatial representations depending on the anatomical projections from sensory receptive fields to distinct groups of sensory neurons in the cortex.

Spatial frames of reference rely on a network of separate brain regions and cognitive mechanisms, in which one of the key regions appears to be the parietal cortex (Colby & Goldberg, 1999; Vallar & Maravita, 2009). Indeed, patients suffering from lesions in the parietal cortex often develop pathologies, such as hemispatial neglect, which is characterized by an important deficit in spatial perception and action (Kerkhoff, 2001), demonstrating the link between the parietal cortex and the ability to construct a representation of the surrounding space and its perception. For example, some patients with unilateral neglect will demonstrate body centred neglect symptoms: i.e. they neglect the left side of space with respect to their body. When those patients are asked to draw a scene, they only draw objects on the right of the model drawing. On the other hand, some patient display object-based neglect symptoms: i.e. they neglect the left side of objects. Those patients will typically draw all objects, but will only draw the right side of the object. This demonstrates that space can be perceived according to different reference frames: body-centred or object-centred in this case.

Several other dissociations can be made between different spatial frames of reference (e.g. Caramazza & Hillis, 1990; Colby, 1998; di Pellegrino, Làdavas, & Farnè, 1997; Guariglia & Antonucci, 1992; Halligan, Fink, Marshall, & Vallar, 2003; Halligan & Marshall, 1991; Landis, 2000; Smania & Aglioti, 1995). Accordingly, a distinction can be made between anatomical vs. spatiotopic frames of reference (where information is mapped according to the anatomical projections of the modality or according to its position in external space, and peripersonal vs. extrapersonal frames of reference (which refers to the space close to the body in which actions are possible, or the space which we can only explore using eye movement of by displacing the whole body).

1.1.1 Anatomical and spatiotopic spatial frames of reference

Incoming sensory information is first processed according to different coordinate systems constrained anatomically by the corresponding sensory system. For instance, visual information is conveyed and mapped in a retinotopic manner. The organisation of the receptive fields [RF] in primary visual cortex [V1] reflects its place on the retina. This means that the map in V1 reflects the retina (Figure 1). This is called a retinotopic coordinate system. Similarly, somatosensory information is mapped on the so-called homunculus in the primary somatosensory cortex [S1] (Roux et al., 2018). The spatial functional organisation of S1 reflects the body's surface (Figure 2). Both retinotopic and somatotopic coordinate systems are examples of anatomical frame of references.

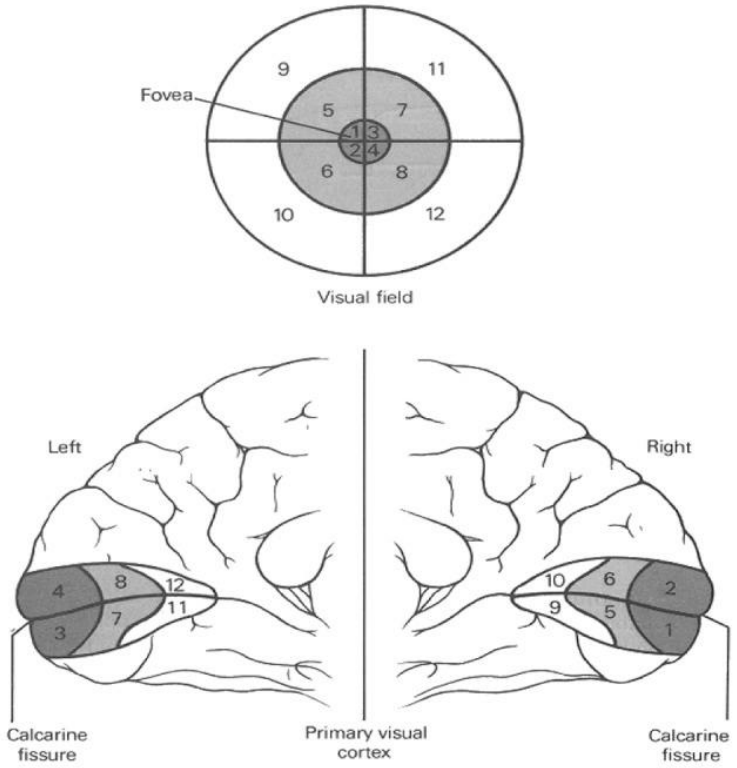

Figure 1. Retinotopic organisation: Schematic representation of the visual field of the retina (top) and its representation in the primary visual cortex (bottom). The incoming visual information is mapped in the primary visual cortex depending on its location on the retina, and the spatial position of a visual stimulus is consequently processed according to a retinotopic organization (From Askenasy & Lehmann, 2013).

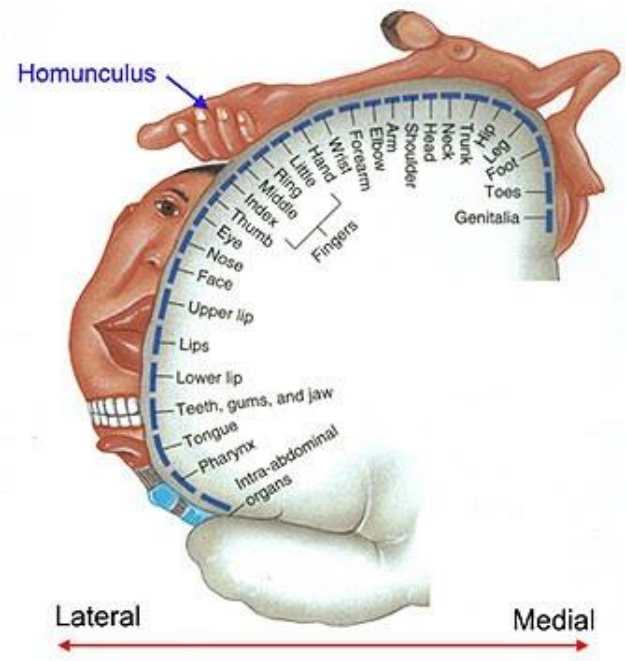

Figure 2: The homunculus depicts the representation of the body in the primary somatosensory cortex. As such, the spatial location of incoming tactile information are processed depending on their location on the body in a somatotopic coordinate system (from Penfield, 1950).

However, relying solely on the anatomical frames of reference is not sufficient to adequately react to events, as it does not reflect the spatial location of events in the external space. Additionally, disregarding the information about the position of the body (i.e. proprioception) would lead to missing a vital part of information. Indeed, in order to form a representation of our environment, we must take proprioceptive information, or the position of the body and body parts in space into account, in addition to the anatomical information (Figure 3). As such, the anatomical coordinate system needs to be remapped towards a coordinate system representing the external space: the spatiotopic frame of reference. Evidence for the distinction can be observed when both reference frames bring

conflicting evidence about the location of a stimulus. Such a situation can be created by asking participants to cross their hands over the body midline, in the so-called 'crossed hands' experiments.

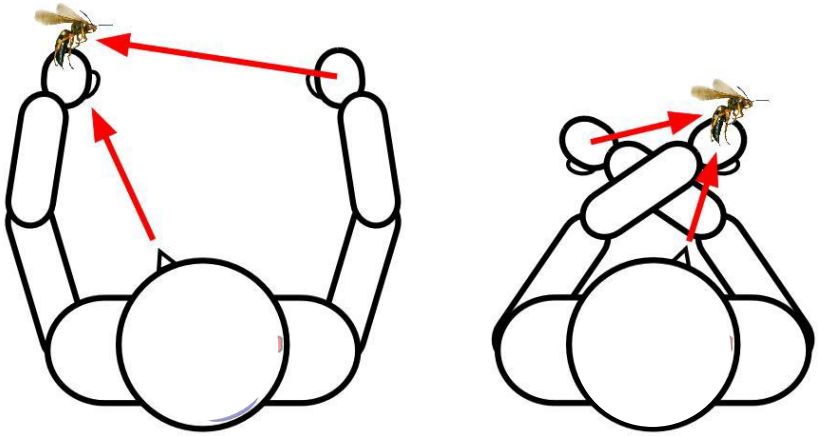

Figure 3: This figure illustrates the necessity of remapping information from anatomical frames of references to a spatiotopic frame of reference. Indeed, while the bee is on the left hand and the left hand is in the left space (left pannel), it is easy to perform the adequate movement to chase the bee away. However, when crossing the hands over the body midline (i.e. the left hand is now on the right of the body midline – the right pannel) the tactile stimulus on the left hand is still projected to the right hemisphere, but is no longer located in the left side in the external space. As such, only detecting the bee on the arm i.e. in the anatomical frame of reference, is not sufficient to react adequately to this situation. One consequently needs to take the position of the limb in external space i.e. in a spatiotopic frame of reference, into consideration. This demonstrates that to interact correctly with the environement, it is important to take the position of the stimuli on the body as well as in external space into account.

Indeed, 'crossed hands' experiments have the particularity of creating a dealign1ment between different reference frames. There is no conflict between different reference frames when feeling a somatosensory stimulus on left hand in an uncrossed position, while the left hand is also in the left side of space (the spatial position of the tactile stimulation consequently coincides on the spatiotopic frame of reference and on the anatomical frame of reference). However, crossing the hand over the body midline introduces conflicting information. While a somatosensory stimulation will be mapped i.e. on the left hand in the anatomical frame of reference, this same stimulation (as is the left hand) is now situated in the right side in the spatiotopic frame of reference. In this case, the stimulation mapped in the anatomical frame of reference carries information about the spatial location of the stimulation on the body, but does not allow for an adequate reaction, as the location of the stimulation in external space is not conveyed by this frame of reference. In this regard, the spatiotopic frame of reference allows for the localization of the stimulation in external space.

Interestingly, such a conflict has been shown to affect our perception of tactile temporal order judgement negatively [TOJ]. In a TOJ task, participants are presented with two stimuli, separated by a variable time interval. They are then asked to judge the temporal order of the two stimuli by reporting which of the two stimuli is perceived as first presented. This results in a measure in time which reflects the interval necessary for the temporal order to be reported correctly. Participants judge the temporal order correctly when the hands are uncrossed when separated by as little 70ms. However, when crossing the hands and consequently introducing a conflict between anatomical and spatiotopic frames of reference, participants need the stimulations to be separated by 124 ms to be able to report them correctly (Shore, Spry, & Spence, 2002, see also Yamamoto & Kitazawa, 2001). These results can be explained by conflicting reference frames.

The larger interval for correct performance during crossed hands conditions could be due to the transformation or realignment of the anatomical and spatiotopic frame of reference for a correct perception.

Further evidence for the dissociation between the anatomical and spatiotopic frame of reference comes from patients with brain lesions suffering from tactile extinction. It was shown that patients that were unable to perceive tactile stimulations on their left hand in an uncrossed posture can improve their performance by crossing their hand, bringing left hand in the right side of space (Smania & Aglioti, 1995). This shows that the performance on a tactile detection test was not only determined by the position of the tactile stimulus on the body, but also in external space.

1.1.2 Peripersonal and extrapersonal frames of reference

A second distinction relevant for this work is between peripersonal and extrapersonal space. Peripersonal space refers to the space around the body in which actions can be performed (Figure 4). It is a coordinate system which integrates the space of the body and the space near the body. It is especially proficient for integration between somatic (e.g. somatosensory, nociceptive) and extra-somatic stimuli (e.g. visual, auditory), and hence acts as an interface between the body and its direct environment (Cardinali, Brozzoli, & Farnè, 2009). Consequently, it is essentially a multisensory coherent representation of the space surrounding the body (Azañón, Stenner, Cardini, & Haggard, 2015; Heed, Buchholz, Engel, & Röder, 2015; Shore et al., 2002). The peripersonal space is principally important for optimizing interactions with our close environment, and probably serves to interact and initiate actions optimally within this environment (Cardinali et al., 2009; Rizzolatti, Fadiga, Fogassi, & Gallese, 1997). In addition, the

peripersonal space has been hypothesises to serve defence mechanisms (Graziano & Cooke, 2006; Graziano, Gross, Taylor, & Moore, 2014), as evidenced by flinch response evoked through electrical stimulation in the cortical regions underlying the peripersonal space in monkeys (Graziano & Cooke, 2006). The extrapersonal reference frame refers to the far to reach space, where objects can only be explored visually, or by moving the body (Cléry et al., 2015). This space is proficient for integration between auditory and visual stimulations (Cardinali et al., 2009).

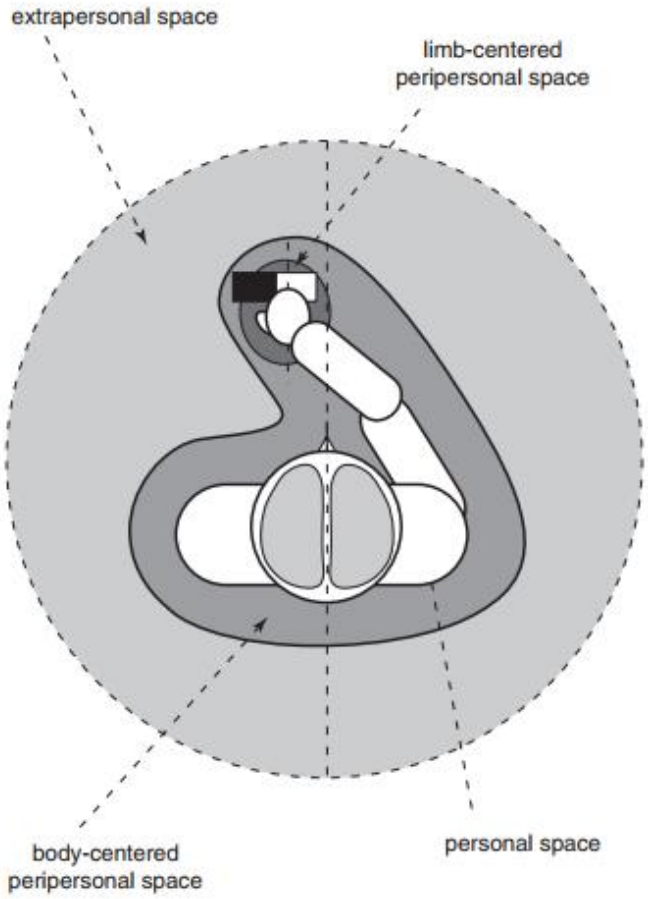

Figure 4: This figure depicts the different frames of reference depending on the distance from the body. The personal space refers to the space of the body. Inputs to the personal space can be coded in an anatomical coordinate system (i.e. according to the anatomical projections of each modality), or in a spatiotopic coordinate system (which takes proprioceptive information into account i.e. when the right hand is on the left of the body midline). The dark-grey zone illustrates the peripersonal space and is a multisensory construct in which we can perform actions. Indeed, stimuli from somatic and extra-somatic modalities can easily interact to guide our actions. As such the peripersonal pace acts as an interface between the

body and the space right next to the body, giving the possibility to for stimuli from different modalities to interact. The peripersonal space can be body centred, but also body part centred, i.e. limb centred. The outer circle depicts the extrapersonal space, or the space which is too far to act in directly (from Legrain & Torta, 2015).

Dissociation between peripersonal and extrapersonal space was observed in patients with hemispatial neglect, a syndrome consecutive to stroke. Neglect is a perceptual syndrome where patients do not perceive a part of space and their body contralateral to the damaged cortical hemisphere, usually the left side. Using line bisection tasks, a patient described by Halligan & Marshall (1991) showed a severe visuospatial deficit when bisecting lines in the peripersonal space, but not when the lines where presented in the extra-personal space. Conversely, Cowey, Small, & Ellis (1994) observed the reverse, with visuospatial deficits in the extrapersonal space, but not in the peripersonal space.

In conclusion, space is a complex percept, requiring different coordinate systems, depending on the actions and goals. Spatial perception can also be studied using multimodal stimuli, which inherently cause a remapping of the spatial position of the different stimuli, e.g. using visual and tactile stimulations in their corresponding reference frames. The use of tactile and visual stimuli is particularly relevant to study the peripersonal space. Indeed, tactile stimuli stimulate the space of the body, whereas visual stimuli allow to study the space external to the body. Therefore, while a large part of research in the field of the interaction between modalities has focussed on uncovering mechanisms underlying the phenomenon, this interaction can also be used to study the organization of space, and the remapping of bodily and external space into one common representation of pace.

1.2 Three principles governing the interaction between the senses

We have now covered the spatial organisation of sensory inputs and why multimodal stimuli can shed a light onto the reference frame which is used to remap information. It is now important to take a look at the different mechanisms available to form a coherent representation of the environment in which we can act.

Simple sensory neurons usually have specific response properties and respond to certain features of one modality (i.e. a neuron in the primary visual cortex responds to certain contrast configurations). Bimodal neurons, or multisensory neurons, are special in the sense that they can be activated through stimulation in different modalities, e.g. touch and vision (Meredith & Stein, 1986). The three main principles governing multisensory integration were derived from the firing patterns of multisensory neurons in the superior colliculi of the anesthetized cat (Stein, Huneycutt, & Meredith, 1988). Importantly, these rules should give rise to the perception that the different stimulation share a common source (Fujisaki, Kitazawa, & Nishida, 2012), and more generally apply more broadly to the interaction between the senses.

The first principle is the temporal rule defining that two stimuli have more chance to interact if they are presented close in time. Meredith et al. (1987) observed that bimodal neurons exhibited super-additive behaviour (i.e. the activity was larger than the sum of the activity when the stimulations were presented separately) or sub-additive behaviour (the bimodal activity is smaller than the sum of the unimodal activity) when stimuli are presented close in time. To observe an interaction between the senses, stimuli thus need to be presented close in time. It is important to make a distinction between three different types of times: the

event time, i.e. the time at which the event is presented; the brain time, i.e. the time course of the brain activity elicited by the event; and thirdly, subjective time, i.e. the time at which the event is perceived by the participant. This is important because stimulation from different modalities might have the same event time, but be processed by the brain at a different time point. However, as differences in timing are naturally present in our environment, multisensory integration and crossmodal interactions have to be robust against these time asynchronies to a certain extent (Fujisaki et al., 2012).

The second principle concerns the spatial aspects. The spatial rule states that two stimuli have more chance to interact if they are presented close in space. The super-additive or sub-additive activity is also observed when the stimulations are presented close in space, indicating an at least partial overlap of receptive fields of the sensory modalities projecting to these neurons (Holmes & Spence, 2005). Behavioural evidence in humans suggests that two stimulations from a different modality interact better when they are in close spatial proximity, but the spatial rule does not systematically imply a behaviourally advantage. Spence (2013) argues that the applicability of the spatial rule in humans seems to depend from the distinction of the 'where' and 'what' processing streams in the brain, initially described in the visual system. While a large part of behavioural data from human experiments corroborates the spatial rule, it still appears that the effect is larger when space is made relevant in the experiment (Spence, 2013). Experiments requiring spatially orienting responses from participants, where space is a relevant feature, or in which spatial attention could be involved, with or without eye movement, show an effect of spatial proximity between two stimuli from different modalities (Diederich, Colonius, Bockhorst, & Tabeling, 2003; Favril, Mouraux, Sambo, & Legrain, 2014; Ngo & Spence, 2010; Sambo et al., 2013). However, the effect is less strong or even absent when participants focus on non-spatial aspects,

such as stimulus identification or temporal perception (Keetels & Vroomen, 2008; Vroomen & Keetels, 2006). Importantly, this indicates that the spatial rule is not a general constraint that always leads an interaction between sensory modalities.

Lastly, multisensory integration is governed by the law of inverse effectiveness. Indeed, it is more likely to observe the phenomenon of multisensory integration if at least one of the stimuli is too weak to elicit activity on its own. However, it appears that when behavioural evidence is found for this rule (Serino, Farnè, Rinaldesi, Haggard, & Làdavas, 2007), alternative explanations related to statistical methodology cannot be excluded (Holmes, 2007). Altogether, it appears that the law of inverse effectiveness also depends from task, stimulus and response demand.

In conclusion, while the evidence for the three principles of multisensory integration seems robust, their application in humans appears to be a matter of debate with alternative explanations sometimes being more suitable (Holmes, 2007).

1.3 Visuo-tactile interactions

The three above-mentioned principles are believed to govern the interaction between modalities. This interaction between the senses can be explained by different three mechanisms: multisensory integration, a crossmodal interaction between sensory cortices, or a crossmodal link in spatial attention. Multisensory integration was first introduced by neurophysiologists to refer neurons responding to two or more modalities, believed to reflect the convergence of the different sensory modalities (Spence, Senkowski, & Röder, 2009). Second, crossmodal interaction between sensory cortices refers to situations in which information of one sensory modality has the ability to influence the processing or

the perception of information originating from a different sensory modality (Spence et al., 2009). This could occur through direct connections between sensory cortices, or through feedback connection originating in multisensory cortices. While the term crossmodal interactions does not imply at what level this interaction takes place, multisensory integration is clearly defined as a hierarchical higher-order process, where information from different senses converge in multisensory cortices to integrate information origination in the primary sensory cortices. Multisensory integration and crossmodal interaction could thus co-exist and are not mutually exclusive. Third, visuo-tactile interactions can also be explained by a crossmodal link in spatial attention, which states that a stimulation can attract attention to its spatial location, and subsequently facilitate the processing of a stimulation in a different modality at the same spatial location. Although this distinction can easily be made clear in neurophysiological studies reporting activity of different types of neurons, the distinction between multisensory integration, crossmodal interaction and a spatial attention is more ambiguous in human studies, whether they rely on electrophysiological, behavioural or perceptual measures.

In the first part of this section, we will review the animal evidence for a network underlying remapping in the peripersonal space, in higher order areas. We will then assess the evidence for a crossmodal interaction between sensory cortices from animal studies. We will finish with human evidence for the spatial organisation of visuo-tactile interactions.

1.3.1 Evidence from animal studies for interaction between the senses

1.3.1.1 Multisensory integration

While research about multisensory integration started with the discovery of bimodal cells in the superior colliculus, the RF of those bimodal cells were mostly mapped with respect to a retinotopic coordinate system, which implies that the location of the RFs moves when the eye moves (Munoz, Pelisson, & Guitton, 1991; Stein, Stanford, Wallace, Vaughan, & Jiang, 2004). Consequently, multisensory integration at this stage does not allow to build a coherent representation of external space, or to study the coordination of spatial reference frames. Later on, bimodal neurons were discovered in a network thought to reflect the ability to map stimuli within a peripersonal frame of reference, including the parietal cortex (area 7b or the inferior parietal lobule [IPL] and the ventral intraparietal cortex [VIP]) (Duhamel, Colby, & Goldberg, 1998), which is generally considered to be involved in spatial perception, and the ventral premotor cortex [PMv] in the prefrontal cortex (F4, located in the area 6, which also sometimes referred to as the polysensory zone) (Graziano et al., 2014)(Figure 5). This system of interconnected multimodal areas is believed to underlie the monitoring of our direct environment in which actions are possible, as well as controlling flinching and other defensive behaviours (Graziano & Cooke, 2006).

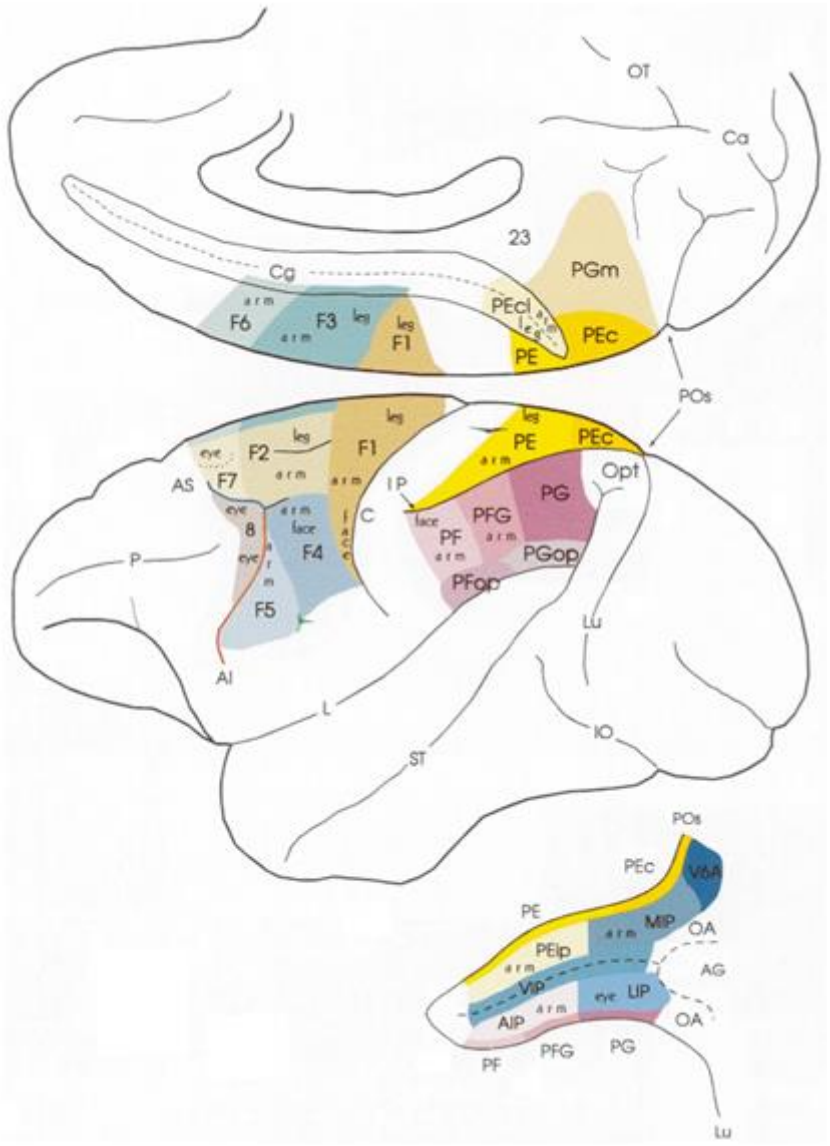

Figure 5: A schematic representation of the macaque cortex, with the main areas involved in multisensory perception of the peripersonal space (VIP, F4, 7b – as indicated by PG/PFG), as well as visuo-motor transformation (AIP, F5 and area 7b (PG/PFG) (From Rizzolatti et al., 1998).

Visuo-tactile bimodal cells in the VIP respond to both visual and tactile stimuli. The visual and tactile RFs of those bimodal neurons partly overlap, extending the bodily space into the external space. Notably, neurons in the VIP are sensitive to motion and selective for the motion direction (Colby et al., 1993; Cook & Maunsell, 2002). The selectivity for motion direction is the same for the visual and the tactile RFs of neuron (Schaafsma & Duysens, 1996). Additionally, in this area, the majority of the visual RFs are locked to the retina to some extent, consequently shifting with eye movements (Duhamel, Bremmer, BenHamed, & Graf, 1997). Importantly, the visual RFs do not exactly match a retinotopic coordinate system. Indeed, while gaze shift causes the visual RFs to shift in the same direction, the size of the shift in RF does not always match the size of the gaze shift (Avillac, Denève, Olivier, Pouget, & Duhamel, 2005; Duhamel et al., 1997). Also, a minority of the visual RFs remain stable, locked to the tactile RF (Duhamel et al., 1997). This implies that space is already taken into account to some extent (Avillac et al., 2005) and that this phase might reflect a transformation where the retinotopically mapped information is remapped according the body in external space, to form a multisensory representation of the peripersonal space.

Visuo-tactile neurons have also been described in area F4 (of the non-human primates) in the PMv (Graziano, Hu, & Gross, 1997; Graziano, Yap, & Gross, 1994; Rizzolatti, Scandolara, Matelli, & Gentilucci, 1981a, 1981b). Those bimodal neurons generally also respond to visual and tactile stimulations (Fogassi et al., 1996; Graziano et al., 1997, 1994; Rizzolatti et al., 1981b). The tactile RFs are located on the face, arm, shoulder or upper torso (Fogassi et al., 1996; Graziano et al., 1997, 1994; Rizzolatti et al., 1981a). The majority of the tactile RFs are located on the contralateral side of the body. The visual RFs of the PMv neurons are located in the space directly next to the tactile RFs, extending the space immediately adjacent to the limb (Figure 6).

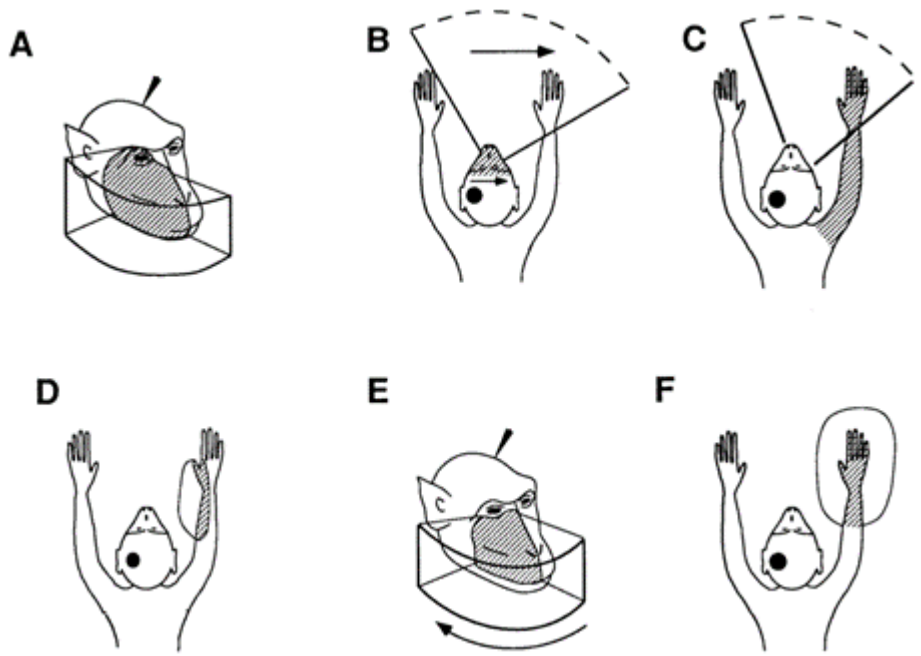

Figure 6: This figure shows the tactile RFs (striped area) and the visual RFs (boxed area) of visuo-tactile bimodal neurons in the PMv. The visual RF matches the tactile RF in locations and extends to the area adjacent to it. Tactile and visual RFs can have preferred directions (B & E). Striped lines indicate when the visual RF extends beyond 1m from the monkey. The tactile and visual RFs are often located on the face near the mouth, or on the limbs and body of the monkey (From Graziano et al., 1997*)*

Interestingly, the anchoring of the visual RFs differs from the ones in the VIP in the parietal cortex (Fogassi et al., 1992; Gentilucci, Scandolara, Pigarev, & Rizzolatti, 1983; Graziano et al., 1997, 1994). Indeed, visual RFs appear to be locked to the tactile RF (Figure 7). This implies that when the arm moves, the visual RF moves with it, independent of gaze shift or direction (Graziano et al., 1997, 1994). Subsequently, it can be hypothesised that these neurons serve to monitor the location and movement in the peripersonal space, and constitute at least a part of

the neural underpinnings of a defensive system (di Pellegrino & Làdavas, 2015; Graziano et al., 2014; Maravita, Spence, & Driver, 2003). In fact, when electrically stimulating bimodal neurons in this area, the animal displays defensive behaviour such as flinching and withdrawing the associated limb (Graziano, Taylor, & Moore, 2002). Altogether, this evidence substantiates the definition of peripersonal space, as it appears to underlie a modality unspecific representation of the limb-centred peripersonal space (Schlack, Sterbing-D'Angelo, Hartung, Hoffmann, & Bremmer, 2005).

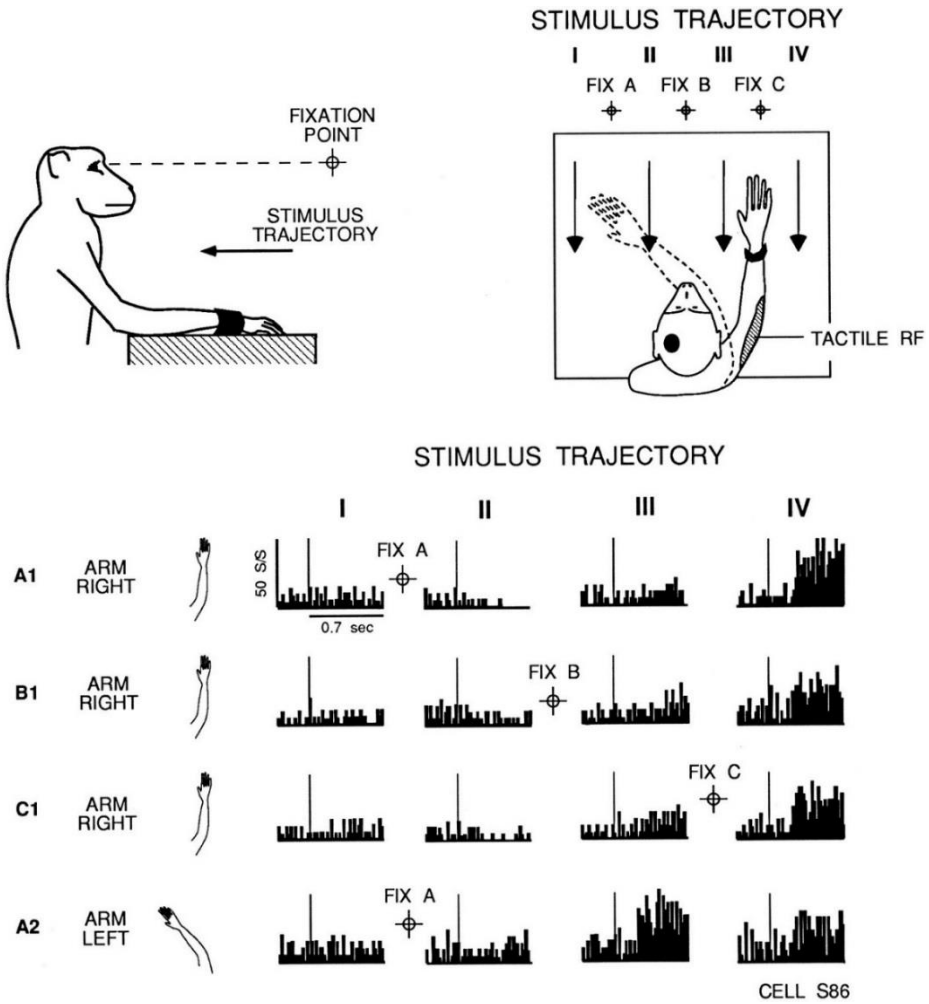

Figure 7: The neural responses of the visuo-tactile RF to an approaching visual stimuli in the PMv. The upper panel shows the experimental set-up: the monkey fixated one of the three fixation points (FIX A, FIX B, FIX C), while a visual stimulation was moved along one of the four stimulus trajectories (I, II, III or IV). The tactile RF is indicated by the stippled area on the arm of the monkey. Additionally, the arm of the monkey could be fixed to the right or to the left. The bottom panel shows the neural response of a visuo-tactile neuron. The first three rows indicate that the visual response is not modulated by the fixation point of the monkey,

demonstrating that the visual response is anchored to the tactile RF (i.e. there is most neural activity when the visual stimulus approached the spatial location of the tactile RF, independently of the fixation point). Additionally, the bottom line shows that when the position in space of the tactile RF moves (i.e. moving the arm to the left), the location of the strongest visual response moves accordingly. The visual response is the strongest when the visual stimuli approaches the tactile RF, which is now located at the end of trajectory III. Overall, these results show that the visual RF extends the tactile RF, irrespectively of fixation and spatial location i.e. the visual RF is anchored to the tactile RF (from Graziano et al., 1997).

Bimodal neurons can also be found in the parietal are 7b (Graziano & Gross, 1995; Hyvärinen, 1981; Hyvärinen & Poranen, 1974; Leinonen, Hyvärinen, Nyman, & Linnankoski, 1979; Leinonen & Nyman, 1979). This area is less well studied in comparison to the F4 and the VIP. The area 7b bimodal neurons are characterized by large tactile RFs, covering large parts of the body (Hyvärinen, 1981). In some cases the tactile RF covers the entire body (Hyvärinen, 1981). About 35% of these neurons also respond to visual stimulation (Graziano & Gross, 1995). However, their RFs only roughly match the location of the tactile RFs, but the response is strongest when the visual stimuli is delivered near the body (Graziano & Gross, 1995). Similarly to the VIP, the majority of the visual RFs are to some extend anchored to the location on the retina (Graziano & Gross, 1995). Again, this gives the impression to be an intermediate stage where the retinal coordinate system being remapped towards a coordinate system fixed to the body or limb surface.

1.3.1.2 Crossmodal interaction

Research into the sensory modalities has long focussed on unisensory processing. When research moved from unisensory processing to multisensory

processing, which started with the findings of Meredith & Stein (1983), multisensory integration was thought to be a bottom-up process, where unisensory information projected onto higher-order areas, in which the information converges. However, more recently, it came to attention that senses can interact as early as at sensory cortices through crossmodal interactions. This crossmodal interactions can refer to existent feedback connections from higher-order cortices to primary cortices, as well as direct connections between primary cortices, termed as a crossmodal interaction between sensory cortices.

Subthreshold multisensory neurons have been described in sensory cortices (Lakatos, Karmos, Mehta, Ulbert, & Schroeder, 2008; Lakatos et al., 2009; Lakatos, Chen, O'Connell, Mills, & Schroeder, 2007). Subthreshold neurons are neurons that respond to their preferential modality (e.g. vision in primary visual cortices, audition in the auditory cortex etc.). These neurons do not respond when presented a stimulus from a different modality, i.e. a subthreshold neuron in the auditory cortex will not respond to a somatosensory stimulus. Interestingly, the activity of those neurons can be modulated by stimulations in a different sensory modality when presented together with their preferential modality (Allman, Keniston, & Meredith, 2008). For example, neurons in the auditory cortex preferentially respond to auditory stimulation. However, the response of auditory subthreshold neurons will be different when a tactile stimuli is presented simultaneously with an auditory stimuli (Lakatos et al., 2007). While such connections between auditory and visual areas have frequently been demonstrated, evidence for a similar link between vision and touch is less extensive: a direct link between V1 an S1 has been demonstrated in rodents (Miller & Vogt, 1984; Wang & Burkhalter, 2007) and connections between sensory cortices and visual cortices have been shown in the marmoset (Cappe & Barone, 2005).

1.3.2 Human evidence for an interaction between the senses

While a clear distinction between multisensory integration and crossmodal interaction can be made in animal studies, the line between multisensory integration, crossmodal interaction and a crossmodal link in spatial attention is less clear in human studies. The critical factor that has been proposed is related to timing (Fujisaki et al., 2012). Indeed, if a modulation of the outcome variable is observed and the stimulations in the different modalities were delivered simultaneously, the results are often interpreted in favour of multisensory integration or crossmodal interaction. If there is a short delay between the presentation of the different stimuli, results are often interpreted as a crossmodal of spatial attention (McDonald, Teder-Sälejärvi, & Ward, 2001). This approach is, however, debated. As discussed earlier (see section about the temporal rule), a distinction can be made between different types of time. Considering different conduction velocities for different sensory modalities, the physiological time and the brain time will not necessarily be the same, even if stimulation were presented simultaneously. As such, maximal integration in a multisensory neuron is often predicted when discharges in response to the stimuli from different modalities overlap in time, rather than synchronous stimulation in the external world. However, this also implies that processes such as multisensory integration and crossmodal interaction should be robust against a certain asynchrony between stimulations. Asynchronies up to 100ms between visual and tactile stimulation still yield activity in bimodal neurons (Meredith et al., 1987). Consequently, the distinction based on timing arguments is not sufficient to make a difference between multisensory integration, crossmodal interaction and a crossmodal link in spatial attention.

While multisensory integration and crossmodal spatial attention could be closely related (Macaluso, 2012; Talsma, Senkowski, Soto-Faraco, & Woldorff,

2010), there is evidence that they rely on different mechanisms. First, multisensory integration has initially been shown in anesthetized animals (Meredith & Stein, 1983), where spatial attention could not have played a role. Second, the ventriloquist effect, a well-known audio-visual multisensory illusion is known to happen preattentively and independently of both voluntary and involuntary spatial attention shifts (Bertelson, Vroomen, De Gelder, & Driver, 2000; Vroomen, Bertelson, & de Gelder, 2001). Third, spatial attention can influence the perception, but not abolish it (Driver, 1996), again showing that multisensory processing is possible without the influence of spatial attention.

While there is evidence for a separation between multisensory integration and a crossmodal link in spatial attention, it remains difficult to disentangle the two, as has been shown by the numerous critical reviews evaluating the principles of multisensory integration in humans (Holmes, 2007; Spence, 2013). It is also not impossible that multisensory integration, crossmodal interaction and a crossmodal link in spatial attention all influence behavioural and neurophysiological results in humans. As such, we consider lesion studies and studies in healthy volunteer in which activity in multisensory areas was uncovered as evidence for multisensory integration, and studies in which modulation in sensory cortices was demonstrated as evidence for a crossmodal interaction. Behavioural evidence from crossmodal spatial attentional task will be discussed with respect to a crossmodal link in spatial attention.

1.3.2.1 Multisensory integration

The first studies exploring multisensory integration in humans using functional magnetic resonance imaging [fMRI] aimed at uncovering a areas that responded to multiple modalities, similar to areas with bimodal neurons in the

animal studies. Presenting dynamic unimodal stimulations in different modalities, Bremmer et al. (2001) identified multisensory areas which could be activated above baseline in all conditions. Similar to the multisensory network in non-human primates, areas in the anterior intraparietal sulcus [IPS] and PMv matching the location of the VIP and the PMv were located, theorized to integrate sensory and proprioceptive information (Bremmer, Schlack, Shah, et al., 2001). Since then, studies have extended this work, exploring the areas in the brain that are activated by visual stimulations close to the hand (Makin, Holmes, & Zohary, 2007). As such, the authors demonstrated that the anterior IPS and the PMv are especially active when a visual stimulus is close to the hand, while the posterior IPS was involved in the visual processing of the position of the hand (Makin et al., 2007). Correspondingly, Brozzoli et al. (2011) investigated the implication of the same network when a visual stimuli approached the right hand. They found that when a visual stimulus approaches the right hand, this resulted in activity in the anterior part of the IPS, in the supramarginal gyrus [SMG] in the parietal lobe and the PMv and dorsal premotor cortex [PMd] in the frontal lobe, corresponding to the network uncovered by the initial work of Bremmer et al. (2001). Importantly, this activation did disappear when the visual stimulation was presented in far space (100 cm from the hand), or when the hand was retracted, indicating the necessity of the spatial congruence between visual and tactile stimulation. As such, this is evidence that these areas code for the peripersonal space, centred on the hand (Brozzoli et al., 2011). Extending these results by studying visuo-tactile interactions, Gentile, Petkova, & Ehrsson (2011) show activation in the left anterior IPS, the insula, the PMd and in the putamen, arguing that these regions underlie the integration of visual and tactile information for the multisensory representation centred on the hand. Complementary work has been done on the peripersonal space centred on the head. Huang et al. (2018) delivered lateralized

tactile and lateralized looming visual stimuli to the face. Crucially, the visual and the tactile stimuli could be congruent in space (i.e. delivered to the same side in space) or incongruent. As such, the authors could uncover the areas involved in the processing of spatially congruent visuo-tactile stimulations, which yielded activations in the middle temporal area [MT] and V6A (Huang et al., 2018). Overall, there is evidence for a network for remapping visual and tactile stimuli in a common spatial reference frame. Crucially, these representation appear to be centred on the body part, and not on the trunk.

Evoked brain potentials have also been used to study multisensory integration, with the aim to uncover the precise timing of these mechanisms. Multisensory integration research in monkeys have shown super-additive or sub-additive activity in multisensory neurons (Stein et al., 1988), this has been generalized to human studies. Researchers have inferred that if the activity related to the presentation of a multimodal stimuli is larger or smaller than the sum of the two unimodal stimulations, this is proof for multisensory integration in human studies as well. As such, multisensory integration as evidenced by super additive activity have been shown in crossmodal studies, which are most audio-visual (Barth, Goldberg, Brett, & Di, 1995; Foxe et al., 2000; Giard & Peronnet, 1999; Gondan & Röder, 2006; Greenwood & Goff, 1987; Hay & Davis, 1971; Klucharev, Möttönen, & Sams, 2003; Lam et al., 1999; Molholm, Ritter, Javitt, & Foxe, 2004; Raij et al., 2010; Talsma, Doty, & Woldorff, 2006; Talsma, Slagter, Nieuwenhuis, Hage, & Kok, 2005; Teder-Sälejärvi et al., 2002; Wright, Pelphrey, Allison, McKeown, & McCarthy, 2003). These conclusions are drawn on the observations that activity related to a multimodal stimulation is smaller or larger than the sum of the activity related two unimodal presented separately. However, this approach has been criticised (Besle, Fort, & Giard, 2004; Calvert & Thesen, 2004; Teder-Sälejärvi et al., 2002). Indeed, deflections in the ERPs do not only reflect modality

specific components, but also more general processes such as attention, saliency (the ability to automatically attract attention) and orienting, which are known to greatly modulate and even explain components such as the vertex component, which is a large deflection characterized by a central topography (Goebel & Van Atteveldt, 2009; Näätänen & Gaillard, 1983; Teder-Sälejärvi et al., 2002). However, activity related to such more general processes is present in the brain potentials of both modalities, but only once in the bimodal stimulation condition. If one compared the sum of a visual stimulation (V) and an auditory stimulation (A) to the bimodal audio-visual stimulation (AV), this general activity is therefore counted double in the sum of the unimodal stimulations, naturally biasing this comparison. Consequently, it is impossible to make the difference between an effect of multisensory integration or a bias using the above mentioned paradigm.

1.3.2.2 Crossmodal interaction

A visuo-tactile crossmodal spatial interaction was demonstrated by comparing blood oxygen level dependent [BOLD] metabolic signal from spatially congruent visuo-tactile stimulation (i.e. presented in the same spatial location), to incongruent visuo-tactile stimulation (Macaluso et al., 2000). Using this paradigm, the authors observed increased activity in the lingual gyri in the visual cortex for multimodal stimulation, but only when visual and tactile stimulation were spatially congruent, i.e. delivered in the same spatial location. The lingual gyri had been considered visual unimodal up to that point. Albeit the tactile and visual stimulation were short in duration and delivered simultaneously, the authors argue for the case of a crossmodal link in spatial attention, which might happen through feedback connections from multisensory areas. This interpretation has however been disputed, notably by De Gelder (2000) and McDonald, Teder-Sälejärvi and Ward (2001), who both interpret the results as multisensory integration, with the

reason of the bimodal stimulation being presented simultaneously. Indeed, McDonald argues that with stimulations being presented simultaneously, this is enough to exclude spatial attention. However, it should be kept in mind that the time resolution of fMRI is approximate at best, leaving the possibility for both spatial attention and multisensory integration as an underlying mechanism, or an interplay between the two processes. Nonetheless, this study shows that a vibrotactile stimulation can alter the processing of visual stimulation when presented at the same spatial location

While this case already brings compelling evidence for crossmodal interactions, the results might still be interpreted in terms of hemispheric correspondence. This theory states that stimulations treated in the same hemisphere are facilitated as compared to when processed by left and right hemisphere (Kinsbourne, 1970). In the previously explained experiment, the somatosensory and visual stimulations project to their respective somatosensory and retinotopic frame of reference, which are located in the same hemisphere. This thus does not inform us by which reference frames we perceive the surrounding space. In order to further investigate this question, Macaluso, Frith, & Driver (2002) investigated the effect of gaze shift on the involvement of the lingual gyri in the processing of spatially congruent visual and somatosensory stimulations. Using the same paradigm, participants were asked to place the stimulated hand right in front of them (Figure 8). Depending on the condition, participants were asked to fixate a point which was located either to the left or to the right of the hand, bringing the congruent visual stimuli either in the left or right hemifield, while the somatosensory stimuli always projects to the same hemibody. This procedure allows to test if the visuo-tactile interaction depends on the congruence of the somatosensory and visual stimuli in external space or on a match between their respective anatomical coordinate systems. Again the authors observe an

increase in visual response when the location in external space is congruent between visual and tactile stimulation. This demonstrates that the altering in visual processing can be attributed to the spatially congruent vibrotactile stimulation, irrespectively from initial anatomical coordinate systems, and consequently for the spatial rule in visuo-tactile interactions. Looking at the timing of these effects using ERPs, Zimmer & Macaluso (2007) observed a first modulation that was due the hemispheric congruence, while a later modulation could be attributed to the congruence in external space.

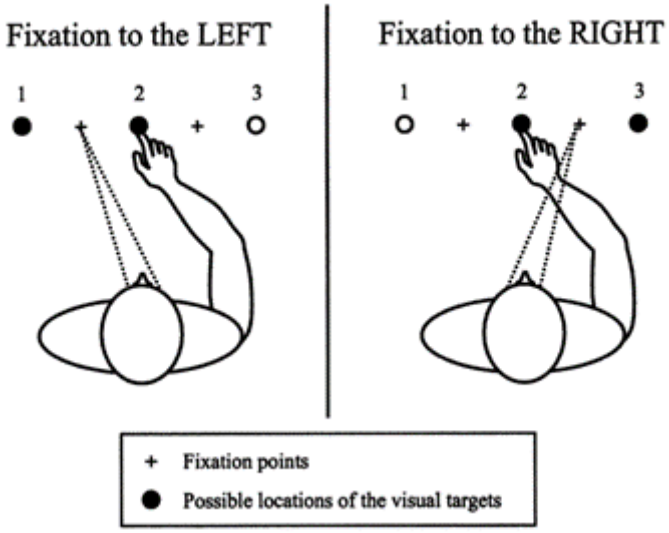

Figure 8: Schematic representation of the experiment procedure for the influence of gaze on visuo-tactile interaction. In the left panel, the somatosensory and the visual stimulation are both projected to their respective anatomical coordinate systems (i.e. somatotopic and retinotopic) which are anatomically aligned i.e. projected to the same hemisphere. However, in the right panel, by shifting the gaze of the participants, the visual stimuli changes hemifield, while the projections of the somatosensory stimuli, still applied to the same hemibody, are unchanged. The respective anatomical coordinate frames are consequently no longer anatomically aligned. Crucially however, the visual and tactile stimuli are still congruent in external space. This allows to study by what spatial principles visuo-tactile interaction is organised and shows the necessity of a spatiotopic frame of reference to realign the visual and somatosensory information according to their location in external space (from Macaluso, 2002).

1.3.2.3 Visuo-tactile crossmodal spatial attention

A third part of evidence for visuo-tactile interaction comes from spatial attention paradigms, measured in behaviour and ERPs. Spatial attention is known to operate across different modalities. Indeed, attending a certain stimulus in one modality is likely to spread to stimuli of different sensory modalities at that same spatial location. Visuo-tactile crossmodal spatial attention has been studied using crossmodal sustained spatial attentional paradigms and crossmodal spatial cueing paradigms. Crossmodal sustained spatial attentional paradigms usually request the participants to pay attention during a whole stimulation block to stimuli of one particular sensory modality at one spatial location, and to ignore same modality stimuli presented at another location and all stimuli of the other modality. In such studies, visuo-tactile interactions have been studied, without looking at their spatial remapping. However, the spatial organisation of visuo-tactile interaction has been studied using crossmodal spatial cueing paradigms, in which attention is attracted in a trial-to-trial fashion to the location of a stimulus in one modality by the preceding presentation of a stimulus of the other modality (i.e. the cue) at the same location.

Using a crossmodal sustained spatial attentional paradigm, a unidirectional crossmodal link between vision and touch was shown. Indeed, a stream of intermingled visual and tactile stimuli were presented, at the left and right of the body midline of the participant (Eimer & Driver, 2000). Three different conditions were tested (Figure 9). In a first condition, participants were asked to report visual target stimuli occurring at the attended location and ignore all other stimulation (‘judge vision’). In a second condition, participants were required to report tactile target stimuli occurring at the attended location, and ignore all other stimulation (‘judge touch’). In a third and last condition, participants were asked to report

tactile targets, which were infrequent but could occur on both sides, in addition to reporting more frequent visual target at the attended location ('vision primary'). This paradigm allowed to study the effect of sustained attention on stimuli in a different modality. As such, the authors showed that paying attention to the visual stimulation leads to an enhanced ERP for the attended visual stimulation, as compared to the unattended (Figure 9). Paying attention to visual stimulation at a particular spatial location did not influence the processing of the tactile stimulations at the same spatial location. Similarly, paying attention to the tactile stimuli led to an enhanced early components of the ERP for attended tactile stimuli as compared to unattended stimuli. Interestingly, paying attention to the tactile stimuli at a given spatial location did also enhance the processing of the visual stimulations occurring at that same spatial location. Crucially, the results of the third condition, where the tactile modality was made task relevant, showed that the tactile response was modulated when it was in the attended location for visual target. As such, it appeared that visual spatial attention can be decoupled from tactile spatial processing. The two processes appear to interact when the tactile stimulation are made of interest as well (Eimer & Driver, 2000). These results have been replicated between touch and vision, and vision and audition (Eimer, Cockburn, Smedley, & Driver, 2001; Eimer & Driver, 2001).

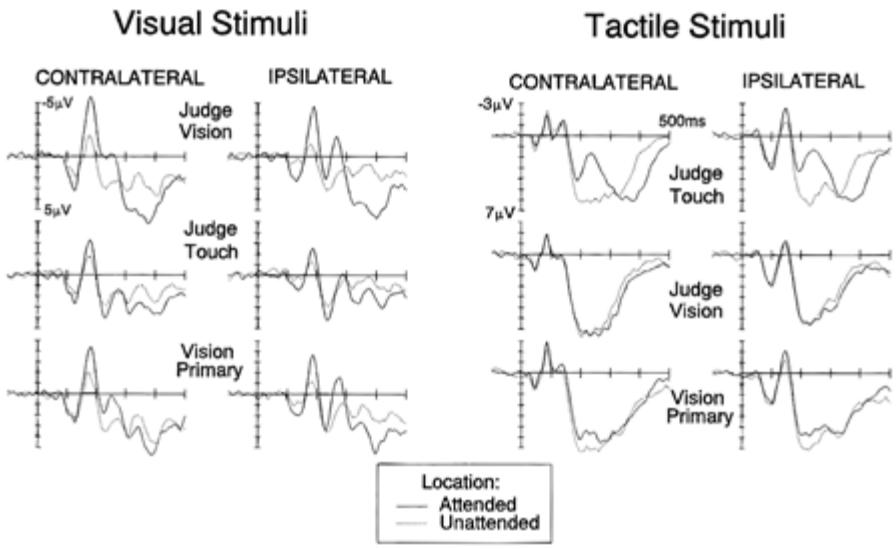

Figure 9: ERPs for a sustained spatial attention paradigm. The visual ERP are modulated when their spatial location is attended versus unattended, as represented by the black and grey lines respectively. However, visual ERPs at the attended spatial location are also modulated when participants are required to report tactile stimulations at the same location. This particularly affects the first positive and negative components (N1 and P1) of the visual ERP. Tactile ERP show a similar modulation when comparing attended versus unattended, However, there is no modulation of the tactile ERP when it appears at the same attended location as the visual stimulation when the visual stimulation are the target. This implies that when tactile information is not necessary for spatial processing, this information can be ignored. On the contrary, visual information appears to be automatically processed for spatial information. Interestingly, when the tactile stimulation also carries spatial information of importance, the tactile ERP is also modulated depending on attended or unattended location. This implies that tactile stimulations are not automatically taken into account in space processing (From Eimer et al., 2002).

In crossmodal cueing paradigms, the cue can be predictive for the spatial location of the stimulation in the second modality, or it can be non-predictive (i.e., participants receive a tactile cue on one of both hands, which they are told to ignore as it does not predict the side of the visual target to which they have to respond). Using such a paradigm to study the visuo-tactile link in spatial attention, participants were asked to respond to a visual target, which was preceded by a non-predictive tactile stimulation (Kennett, Eimer, et al., 2001). Participants' hands could be uncrossed or crossed over the body midline. In a first psychophysiological experiment, the authors assessed reaction times to visual targets. Behavioural data revealed that participants reacted faster when the visual target is preceded by a tactile stimulation on the same side of space as compared to when the tactile stimulation was presented in the other side of space, irrespectively of posture. Additionally, the authors found converging evidence in the ERP data. Indeed, the ERP to the visual stimulation displayed an enhanced early negativity on the occipital electrodes when the visual stimuli was preceded by a tactile stimulation on the same side of space as compared to when the tactile stimulation was delivered on the other hand. Interestingly, the authors tested different time intervals between cue and target, and noted that this facilitator effect of a spatially congruent vibrotactile stimulation stronger when the interval between the tactile and visual stimulation was short (150ms vs. 300 ms). This demonstrates that the ability of a spatially congruent tactile cue to direct attention is short lived and that stimulations have to be presented close in time to be able to interact. This is also additional evidence, showing a remapping between modalities to a spatiotopic reference frame.

While such a cueing paradigm has been used several times to study crossmodal spatial attention, this is one of the few studies exploring the influence of touch on vision. Indeed, most studies have looked at the influence of vision on

touch (Maravita et al., 2003). Using the crossmodal congruency, they explored the influence of a cue in one modality on the discrimination performances in a different modality. Participants were asked to hold a sponge cube in both their hands, with two vibrotactile stimulators and two visual stimulators on each cube, at different elevations. Participants are then asked to judge the elevation of unpredictable vibrotactile stimuli delivered to their thumb or index finger (up or down), while fixating a central point. A visual stimulus is delivered previous to the target at a random location on each trial. These studies show that distractors that are incongruent in elevation delay the tactile discrimination performance. Additionally, this effect is larger when the visual distractor is delivered near the stimulated hand. Interestingly, results show that it is the congruence in external space that determines this effect, when crossing the hands over the body midline.

Similar results have been obtained using visuo-nociceptive TOJ, studying the effect of nociceptive cues on visual TOJ (Filbrich, Alamia, Blandiaux, Burns, & Legrain, 2017; Filbrich, Alamia, Burns, & Legrain, 2017; Filbrich, Halicka, Alamia, & Legrain, 2018; Vanderclausen, Filbrich, Alamia, & Legrain, 2017). Indeed, the TOJ measure is biased towards the side opposite of the nociceptive cue, to the uncued side. Interestingly, this effect is stronger when the visual stimuli are located close to the hands, as compared to when the visual stimuli are at a certain distance from the stimulated hands. This is observed irrespectively from the distance between the body and the stimulated limb. As such, these results indicate the existence of a system representing the space, centred around stimulated limb, and not the central body. Additionally, it is the proximity between the visual and nociceptive cue in external space that drives these results, even if they are anatomically mapped in opposite hemispheres. While these results are based on nociceptive cues, there is no reason to believe that visuo-tactile interactions behave in a different manner concerning their integration in the peripersonal space. Overall,

this indicates that the mechanisms underlying these interactions are based on a spatial representation that is centred on the limb.

Measuring EEG to sustained visual and tactile stimulations, Colon, Legrain, Huang, & Mouraux (2015) demonstrated an increase in the amplitude of the tactile FPS response in response to approaching and receding visual FPS. However, the amplitude of the vibrotactile FPS response only increased when the visual stimulation was over the hand, showing the necessity of spatial proximity to observe visuo-tactile interaction. Additionally, they showed that this effect is preserved when participants crossed the hand over the body midline; indeed, it is the proximity on external space that causes the increase in amplitude of the vibrotactile FPS response.

Further evidence for the spatial organisation visuo-tactile interactions can be found in patients with brain damage (Brozzoli, Demattè, Pavani, Frassinetti, & Farnè, 2006). These patients sometimes display extinction behaviour. Indeed, they are able to report unilateral stimuli independently of the side of space, but encounter problems when the stimuli are applied concurrently on both side of space simultaneously. In such case, patients only report the stimulations in the ipsilesional side, and it is said the perception of the contralesional stimulus is extinguished by the simultaneous presentation of the ipsilesional stimulus (Brozzoli et al., 2006). This paradigm can also be used in a multimodal setting, when the stimulation delivered concurrently on both side are from a different modality. As such, it has been shown that when a tactile and a visual stimulation are delivered concurrently, the visual stimulation delivered in the ipsilesional side extinguishes the tactile stimulation in the contralesional side as much as a tactile stimuli delivered in the ipsilesional side (Làdavas et al., 1998). Such results have been shown robustly (Farnè, Demattè, & Làdavas, 2005; Ladavas, Zeloni, & Farnè, 1998). Importantly, this visuo-tactile extinction appeared to be more efficient when close to the body (in the peripersonal space) as compared to distant from the body (in extrapersonal space) (Làdavas et al., 1998). However, it is important to note that it was the proximity between the stimulated limb and

the visual stimulation that was driving this interaction, more than the proximity to the trunk. These results show that both tactile and visual stimulation are integrated into the peripersonal space.

In conclusion, visuo-tactile interaction appear to be robust and organised in a space based reference frame centred on the limb. However, it is unclear along what spatial frame of reference vision and touch interact when using sustained stimulations and how sustained tactile stimulation influence the processing of visual sustained stimulations. Sustained stimulation might serve an important role in haptic and visual feedback mechanisms during movement and action. As such, it might be interesting to use mode ecological visual stimuli to stimulate these underlying integrative mechanisms.

2. Tools and multisensory integration

The studies described in the previous sections relied on basic visual stimulations, believed to mainly activate V1 (Norcia, Appelbaum, Ales, Cottereau, & Rossion, 2015; Srinivasan, Bibi, & Nunez, 2006), to demonstrate an interaction between the sensory modalities. However, it should also be possible make touch and vision interact in the peripersonal reference frame using more complex visual stimuli, implying an involvement of higher-order visual areas. Such stimuli could be tools, which carry a special meaning as they automatically activate associated motor schemes to initiate actions, and might be particularly relevant to study visuo-tactile interactions with sustained stimuli. Similarly to the multisensory network that underlies the peripersonal space, there appears to be a network in the brain, processing the visuo-motor transformation involved in grasping objects in our direct environment. Additionally, there is evidence that spatial frames of

references are also used with tool use and grasping, potentially indicating that there is an interplay between the network of multisensory integration and tool use. In this part, we will first discuss the neural underpinnings of the visuo-motor transformation implied in object perception and the higher-order areas integrating this visual information to form the appropriate preparation for a motor action. In a second part, we will discuss the human evidence for the spatial effect of visuo-motor tool use.

2.1 Neural underpinnings of tool perception and grasping

2.1.1 Spatial perception in the visual system

A dissociation between different networks in the visual system based on experimental work was first proposed by Ungerleider & Mishkin (1982), established by lesion studies in macaque rhesus monkeys. They observe that the ability to discriminate between different objects based on features such as form, colour and texture was strongly altered by a lesion in the infero-temporal cortex. Interestingly, the monkeys were still able to localize food based on local spatial cues. Conversely, they observe that monkeys were still able to discriminate between different objects but could not localize food based on spatial cues anymore when a lesion is applied in the posterior parietal cortex [PPC]. Consequently, the authors argue that the visual system extending towards the temporal lobe (the ventral visual stream) is involved in object recognition, and the visual system extending towards the parietal lobe (the dorsal visual stream) is involved in the spatial perception of visual information. This distinction led to the more common naming of the 'what/how' and 'where' stream in visual processing system, respectively. After the observations in the macaque monkey, lesion studies in human have led to the proposition of a similar system in human (Goodale &

Milner, 1992). They postulate that the two distinct visual streams serve different behavioural aims: perception or action (Goodale & Milner, 1992). Indeed, patients with lesions in the occipito-temporal region are often unable to recognise and describe visual objects, but can navigate in their everyday environment without much trouble. Conversely, patient with lesions in the PPC, suffering from optic ataxia, are unable to reach and grasp visual objects, but have no difficulty recognising objects (Haaland, Harrington, & Knight, 2000; Johnson-Frey, 2004). It thus appears that the PPC is involved in the visuo-motor transformations, or how we transform perception into action.

2.1.2 Tools are special

While visual objects of all categories activate the ventral visual stream for semantics and categorization, they do not always activate the dorsal visual stream. Tools, however, have been shown to automatically activate both the dorsal visual stream and the ventral visual stream. Importantly, this activation is independent of the intention to handle (Grafton, Fadiga, Arbib, & Rizzolatti, 1997; Lewis, 2006). This special activation pattern has been explained through the theory of affordances (Gibson, 1979) which states that we perceive our environment through the potential actions. Seeing a tool would in this case thus potentiate the different actions we can perform with it, and this inevitable requires a visuo-motor transformation, transforming our perception into an action, and in consequence the activation of the associated motor schemes (Creem-Regehr & Lee, 2005). Indeed, to be able to successfully interact with our environment, we need to adapt our behaviour.

Human neuroimaging studies have shown that the observation of tools, which have strong and specific action associations, activates a left-lateralized

network of brain regions including the AIP (Chao & Martin, 2000; Chouinard & Goodale, 2012; Valyear, Cavina-Pratesi, Stiglick, & Culham, 2007), ending up in the left dorsal and ventral premotor cortex, where perception is transformed into action (Figure 10). The same left-lateralized network has been associated with viewing, hearing, executing, planning and pantomiming tool use (Lewis, 2006). It has also been postulated that tools automatically grab selective attention, potentially influencing aforementioned conclusions to some extent (Adamo & Ferber, 2009; Handy, Grafton, Shroff, Ketay, & Gazzaniga, 2003).

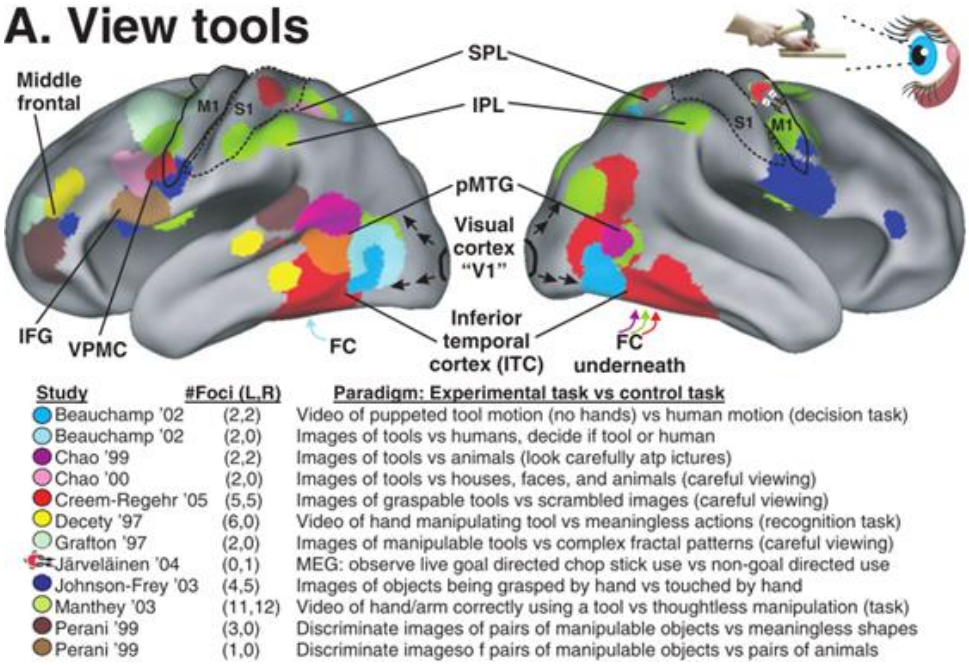

Figure 10: Network of areas activated when viewing static tools (without hand) or observing movies of tool use (with and without hand). The primary visual cortex (V1), the primary somatosensory cortex (S1) and primary motor cortex (M1) are indicated. The arrow depict the hierarchical flow of visual information through the ventral and dorsal visual pathways. Viewing tools is characterized by activity in the temporal cortex, but also in the parietal and frontal areas, demonstrating the

automatic activation of both pathways by observing tools. Differences between studies could arise through the choice of contrasting conditions (from Lewis, 2006).

2.2 Tool perception, grasping & space perception

Even more importantly, there is behavioural evidence in humans showing that the simple observation of tools invokes motor representations of actions afforded by the tools (Grèzes, Armony, Rowe, & Passingham, 2003; Humphreys, Riddoch, Forti, & Ackroyd, 2004; Masson, Bub, & Breuer, 2011; Tucker & Ellis, 1998). Indeed, it has been shown that when response hand and tool orientation are congruent (i.e. the tool can be naturally grasped by the right hand and responses are carried out by the right hand), this speeds up reaction time. In this study, participants had to judge if the presented object was upright or inverted. Additionally and unrelated to the task, the objects were congruent, i.e. their orientation was favourable for a grasping movement. The authors demonstrate that when the orientation of the object was congruent with the response hand, participants responded faster as compared to when the object was not congruent to the response hand (Tucker & Ellis, 1998). This appears to happen automatically and supports the claim that affordances are automatically activated. Interestingly, the same has been shown for grasp types. While being asked to performing either a power grip or a precision grip to classify objects as man-made or natural, the time needed to initiate a response is speeded up if the response grip type is afforded by the object (Tucker & Ellis, 2001).

Importantly, this facilitation of motor actions by afforded actions takes space into account. Cardellicchio et al. (2011) asked participants to observe objects which were in or out of reach, as such making the difference between the peripersonal and extrapersonal space. The objects were either a tool or a

geometric 3D shape. While observing those objects, which could be either graspable or not, and reachable or not, participants received TMS-pulses to measure the motor evoked potential [MEP]. The authors showed that tool in the reachable space gave rise to increased MEP as compared to a non-graspable tool within reach, and as compared to graspable and non graspable objects outside the reachable space. They consequently conclude that the effect of affordance only appears to be valid in the peripersonal space, the space in which a person can act (see also Costantini et al. (2010)). However, the fact that the theory of affordances only holds for graspable objects has been criticized as this facilitation can be explained by the saliency (i.e. the quality of automatically grabbing attention) of the asymmetry of the graspable object i.e. a handle of a cup could automatically grab the attention of the participant (Anderson, Yamagishi, & Karavia, 2002). However, Buccino et al. (2009) showed that the facilitation of affordance only holds when the handle of the cup they use as stimuli is whole, and that the effect is annihilated when the cup is presented with a broken handle. This demonstrates that the effect of affordance cannot be explained by saliency alone and that it is the 'toolness' that activates the associated motor scheme.

2.3 Multisensory integration and tool perception?

Also, it has been shown that the interaction between a tactile and task-irrelevant visual stimulation increases when a grasping movement is initiated, as compared to when no movement is initiated. These results support the notion that the peripersonal space is involved in object-oriented actions (Brozzoli, Pavani, Urquizar, Cardinali, & Farnè, 2009). Incrementing these results, the authors show that there is a substantial difference in visuo-tactile interaction when pointing or grasping, which both involve reaching behaviour, but a different final motor part and goal (Brozzoli, Cardinali, Pavani, & Farnè, 2010). Interestingly, the visuo-tactile

interaction was indistinguishable in the reaching part, but was further enhanced when participants executed the grasping behaviour as compared to the pointing behaviour.

In conclusion, the visuo-motor system serves the interactions we can have with objects our direct environment, the circuit involved in multisensory integration and peripersonal space serves to coordinate and monitor actions in our direct environment and our defensive system (Graziano & Cooke, 2006). Consequently, an interplay between the two systems is necessary to be able to interact correctly with our environment, which is also supported by the above mentioned evidence. Additionally, sustained stimulations might be particularly important for feedback mechanisms in i.e. tool-use. As such, tools might be particularly relevant for visuo-tactile interactions using sustained stimulations.

3. Fast Periodic Stimulation

While very precise timing is one of the advantages of ERPs, this method has the disadvantage of being very affected by noise and artefacts, such as eye blinks and muscular activity (Gratton, 1998). Consequently, it requires averaging a great number of trials to obtain a stable response. Additionally, it is impossible to disentangle responses to different stimuli in an ERP response when the stimulations are delivered close in time or simultaneously, as responses naturally overlap in time. This drawbacks affect the way ERPs can be used in multisensory research. A second possibility is to analyse the EEG signal in the frequency domain. Stimulation at a strict periodic rate is reflected by activity at that exact same frequency in the EEG data, and sometimes harmonics, called FPS response or SSEP. This has proven to be a very robust method, especially in visual research, giving rise to a fast periodic visual stimulation [FPVS] response or steady state visual evoked potentials [SSVEP] (Regan, 1966). To summarize, Regan (1989) introduced the term of SSEP as the response to a gentle shake of the system at a fixed repetition rate, whereas ERP were described as 'the response to a kick in the system'. However, the terms SSEP and SSVEP refer to a much debated neural entrainment due to repetitive periodic stimulation. Consequently, the term of FPVS, which refers to the stimulation more than to the type of response, will be used throughout the rest of this thesis to refer to periodic visual stimulation.

3.1 Description of the methodology

Stimulating at a strict periodic rate yields a response at the same frequency in the frequency spectrum of the EEG data. This periodic response is characterized by its amplitude and phase, appearing as a peak at the frequency of stimulation in

the frequency spectrum (Figure 12). Each periodic response is also characterized with a corresponding topography, reflecting the distribution of amplitude of that frequency bin over the scalp electrodes. The FPVS response could reflect two different processes. On one hand, the response could reflect an entrainment, a resonance of a particular neural population responding to the stimulus at the frequency of stimulation. Evidence for this view is the fact that a FPS response sometimes continues, even when the stimulation per se has already ended (Herrmann, 2001; Vialatte, Maurice, Dauwels, & Cichocki, 2010). On the other hand, the FPS response could be the linear superposition of single evoked responses to each succeeding separate stimulation to constitute a train of evoked response (Bohórquez & Özdamar, 2008; Capilla, Pazo-Alvarez, Darriba, Campo, & Gross, 2011).

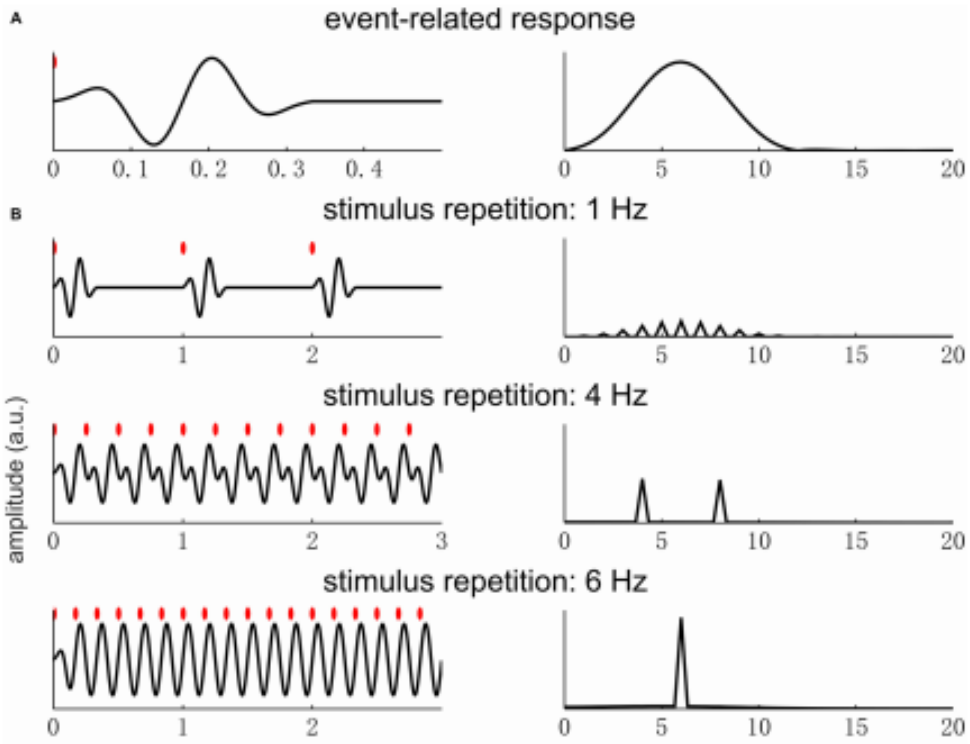

Figure 11: The relation between stimulation (red dots in the left panel), the signal in the time domain (left panel) and in the frequency domain (right panel). The signal in the frequency domain is related to the signal in the time domain. First, the signal depends on the waveform of the evoked response in the time domain (i.e. a pure sinusoidal waveform is reflected as one single peak at the frequency of stimulation in the frequency domain. If the signal is not purely sinusoidal in the time domain, this will be translated in the addition of harmonics at multiples of the frequency of stimulation in the time domain. Second, the stimulus needs to be repeated a certain number of times to reach a stable and focussed signal in the frequency domain (from Zhou et al., 2016)

3.2 Advantages of the methodology

As compared to ERPs, the FPS response offers interesting advantages. The FPS response is characterized by a high signal to noise ratio, lowering the total experiment time and repetition of stimuli needed to obtain a stable signal (Colon, Legrain, & Mouraux, 2012; Meigen & Bach, 1999; Noss, Boles, & Yingling, 1996; Vialatte et al., 2010). This high signal-to-noise ratio is attributable to the periodic response which is concentrated at the frequency bin of stimulation and harmonics. The FPS response has desirable features as compared to other neurophysiological methods, when it comes to studying multisensory integration. As the FPS response is narrowly concentrated at the frequency of stimulation, it is possible to disentangle the sources of different stimulation presented simultaneously, as long as the frequency and harmonics do not overlap. Second, while it is necessary to use short and transient stimulation to obtain ERPs due to phase-locking, sustained trains of stimulations can be used in the FPS technique, to circumvent short term attentional effects. Third, the frequency of stimulation can be used in order to target specific neural populations. Indeed, as it is believed that FPS entrain neural

population to resonate (Galambos, 1980; Herrmann, 2001), and that neural populations have inherent preferential firing rates, specific regions or populations can be targeted by using certain frequencies of stimulation (Alonso-Prieto, Belle, Liu-Shuang, Norcia, & Rossion, 2013; Gauthier, Eger, Hesselmann, Giraud, & Kleinschmidt, 2012; Koenig-Robert & VanRullen, 2013; Vialatte et al., 2010). This is also evidenced by different topographies characterizing different frequencies (Herrmann, 2001; Srinivasan et al., 2006).

3.3 Sources of FPVS

The origins of the FPVS response depend from the type of stimulation used. However, it is safe to say that a large part of the activity can be related back to the periodic fluctuations of low-level visual features such as contrast, luminance etc., and therefore originates in the visual cortex (Ding, Sperling, & Srinivasan, 2005; Regan, 1989; Srinivasan et al., 2006). In the context of this thesis, visual stimulations were constructed either by periodically modulated the luminance in the experiments in which we used light emitting diodes [LED], either by periodically modulating the contrast of the image. It has been shown that the modulation of luminosity using strobe light, originate mainly in V1. While the source of the activity greatly varies with the method of stimulation use, V1 is a predominant source, present in all methods of visual stimulation, such as strobe light, checkerboard or Gabor gratings (Vialatte et al., 2010).

3.4 An objective measure for visuo-tactile integration

While the interaction between the different sensory modalities have often been studied with behavioural paradigms, or ERP, FPS could bring significant

progress in this field. Indeed, FPVS allows for presenting multiple stimulations from different sensory modalities simultaneously.

It is theoretically possible to obtain an 'objective' measure of multisensory integration through the observation of crossmodulation frequencies. Crossmodulation frequencies are defined as activity as the sum or subtraction of the frequencies of two unimodal stimulations. For example, if a visual stimulation is delivered at f1 and a tactile stimulation is delivered at f2, an activity at the crossmodulation frequency could be expected either at f1 + f2, either at f1 – f2 or other combinations of the two frequencies (Figure 13).

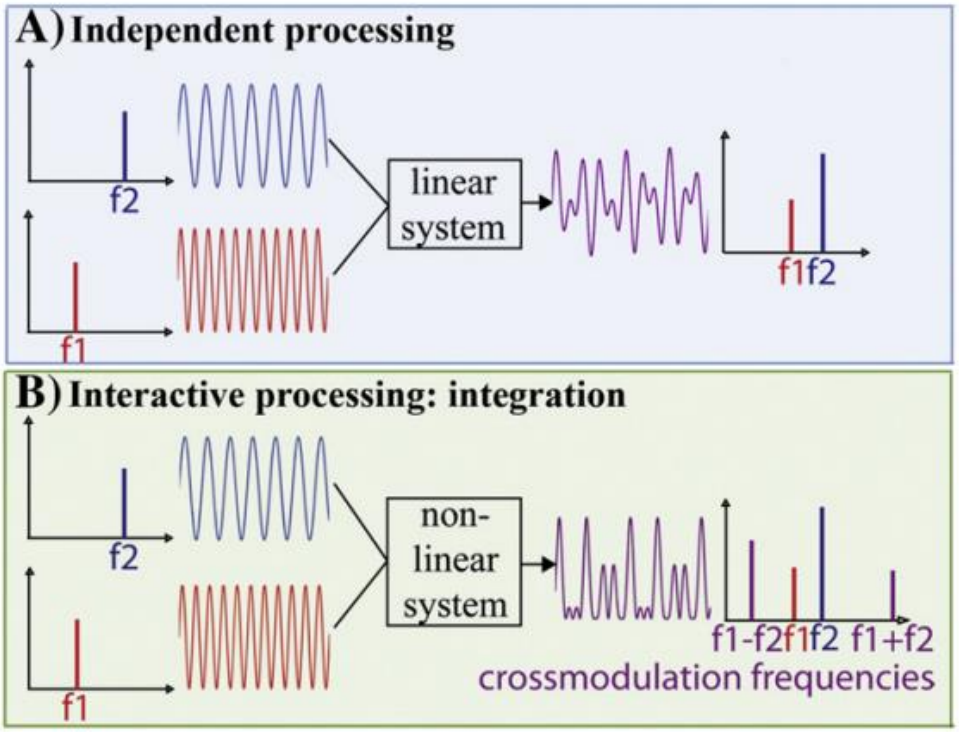

Figure 12: the left side of panel A and B shows the signal evoked by two stimulations in the frequency and time domain. The right side shows the result in

the time and frequency domain if both are presented together. Stimulations can interact in a linear manner, which results in the signal at the two stimulation frequencies. However, if integration takes place, this results in crossmodulation frequencies (from Giani et al., 2012).

While this activity is theoretically possible, the evidence is sparse in the literature. Initially described by Regan (1995), these crossmodulations have only been replicated by a few authors. Nozaradan et al. (2013) identified cross-modulation activity between auditory beat perception and hand tapping. They hypothesised that this activity could be due to a non-linear integration process through which information from the different senses could converge. In addition, Hertz and Amedi (2010) showed that by delivering auditory stimulation at 0.037 Hz and visual stimulation at 0.051 Hz, it is possible to observe activity at the crossmodal frequency of 0.014 Hz. However, this also coincides with every time the visual and auditory stimulation was presented simultaneously. While the potential of this approach has long been recognised, studies which have been able to demonstrate activity at crossmodal crossmodulation frequencies are scarce, probably implicating the low success rate of such studies. Furthermore, it remains unknown under which conditions it can be observed, probably due to the difficulty to observe this phenomenon. It is however worth to note that it is possible to observe activity at crossmodulation frequencies within one modality; e.g. by stimulation visually at two different frequencies (Appelbaum, Wade, Pettet, Vildavski, & Norcia, 2008; Boremanse, Norcia, & Rossion, 2014; Giani et al., 2012; Sutoyo & Srinivasan, 2009; Victor & Conte, 2000; Zemon & Ratliff, 1984). Overall, while the FPS methodology has the potential to produce great advances in the field, its use as described above still remains to be proven. However, one has to take into account, that by using different stimulation frequency, this could impact the perception and timing which is extremely relevant for multisensory integration.

In case of this thesis, crossmodulation frequencies offer the possibility to research visuo-tactile interactions. Indeed, depending on reference frame, the tactile stimuli are supposed to interact with different visual stimuli at a different frequency, which can be disentangled from the crossmodulation frequencies, if present in the frequency spectrum.

3.5 Modulation of activation

While is it unlikely to observe activity at the crossmodal crossmodulation frequencies, an alternative resides in the modulation of the amplitude of the FPS response. The modulation of visual FPVS by attention is quite well known. Indeed, attention does increase the amplitude of the FPVS response (Morgan, Hansen, & Hillyard, 1996; Müller & Hillyard, 2000). This increase is believed to result from an increase phase synchronisation between ongoing oscillatory activity when a stimulation is attended (Porcu, Keitel, & Müller, 2013). While the studies of Muller and colleagues show a modulation of the FPVS amplitude through intramodal attention, Porcu et al. (2013) observed an effect of crossmodal attention, as they define it. Indeed, they show that the amplitude of signal increases when the tactile or visual modality is attended, as compared to when the other modality is attended, using visual FPS and tactile FPS. However, they observe a differential effect on the different harmonics. Indeed, while the effect is present on the base frequency and the first harmonic of the tactile frequency, the effect of attention is only observed on the base frequency of the visual stimulation frequency, and not on the first harmonic. This is evidence that different harmonics potentially have different neural generators (Pastor, Valencia, Artieda, Alegre, & Masdeu, 2007; Porcu et al., 2013). Overall, this implies that attention does not affect all modalities equally. With regards to crossmodal interaction, Colon et al. (2015) showed a modulation of vibrotactile FPS by visual stimulation which was either approaching

or receding. Indeed, they observe an increase in the amplitude of vibrotactile FPS, but only when the visual stimulus was making an approaching movement, and was on the hand. This can be considered FPS evidence for the peripersonal space, where stimulations are only integrated or able to influence each other when it is needed to construct a multisensory representation of space.

In conclusion, FPS offers great advantages over other methods to study the spatial organisation of multisensory integration. Also, while it is unlikely to observe activity at crossmodal crossmodulation frequencies, a modulation of activity by a different modality could already indicate that information is integrated into a multisensory representation of the surrounding space.

Through this introduction, it has become clear that while the spatial aspects of visuo-tactile interactions are well-known, it is unclear what happens when using sustained stimulations simultaneously. Indeed, simultaneous stimulation might mimic a more natural environment better, and sustained stimulations are of interest through their role in feedback mechanisms during motor control. Because of this, the use of images of tools might be a better approach to study visuo-tactile interactions with sustained stimulations. Last, untangling effects of multisensory integration and spatial attention remains an issue in current research but might be possible using FPS, through activity at crossmodulation frequencies.

4. Aim & outline of experiments

The neural underpinnings of the spatial organisation of multisensory integration has extensively been studied in human, trying to uncover similar mechanisms as shown in monkeys. It is known that stimulations that occur congruently in space and in time, are able to influence each other's processing in a positive way, such as speeding up reaction time, enhancing BOLD responses, etc. It remains unclear how a coherent maps of the environment is. To deal with this issue, we propose to study this visuo-tactile interaction with sustained trains of stimulation, where it has traditionally been studied with brief transient stimuli. Taking advantage of the high signal to noise ratio of the FPS technique, it is also possible to present multiple stimulations simultaneously, and it is not limited in time or number. In consequence, this thesis aims at extending the knowledge about the reference frames underlying the interaction between the different sensory modalities, and more precisely visuo-tactile processing, with the aim of witnessing activity at crossmodulation frequencies or a modulation of activity at fundamental frequencies. We generally hypothesise that when tactile stimulation is spatially congruent to visual stimulation, this will be evidenced by crossmodulation frequencies at the sum of the two stimulation, of by a modulatory effect on the amplitude of the visual stimuli i.e. the processing of this visual stimulation will be enhanced as compared to when this tactile information is not spatially congruent. Each objective is addressed in a separate study.

This thesis consisted of two main parts. In the first part, the objective was to demonstrate in which conditions vision and touch interact, using periodic modulations in the luminance of visual stimulations and in the amplitude of the vibrotactile stimulations. We expected visual and touch to interact when they were

congruent in space, reflecting the visuo-tactile interaction in the peripersonal frame of reference, representing the external space. A first experiment aimed at demonstrating that visual processing can be modulated by a spatially congruent vibrotactile stimulation, when the task involved both visual and tactile stimulations. We delivered two visual and one vibrotactile stimulation stream simultaneously, at different stimulation frequencies. Crucially, one visual stimulation was always spatially congruent to the vibrotactile stimulation. As such, we explored the effect of a spatially congruent vibrotactile stimulation on visual processing. Participants were asked to pay attention to all sides of space and all modalities. Indeed, by making all stimulations task-relevant, we investigate if visual and tactile stimuli interact when dividing attention between spatial locations and between modalities.

A second experiment aimed at demonstrating that visual processing was modulated by a spatially congruent vibrotactile stimulation, when participants could ignore the tactile stimulation, and to only pay attention to the two visual stimulations to meet task demands. This experiment used the exact same set-up as the first experiment. However, participants were no longer required to pay attention to different modalities. Indeed, paying attention to the visual modalities was sufficient to perform the task. As such, spatial attention was divided over two locations, but no longer within modalities, as was the case in the first experiment.

A third experiment again aimed at demonstrating a visuo-tactile interaction when visual and tactile stimulation were spatially congruent. As in the first experiment, participants were required to pay attention to both visual and tactile stimulations. Trying to optimise conditions to witness visuo-tactile interactions, participants were instructed to pay attention to one side of space, ignoring all stimulation in the opposite side of space. As such, the set-up and stimulations were

identical to the first two experiments. However, in one half of the blocks, participants were instructed to only pay attention to stimuli occurring in the left side of space, while in the other half, their attention was redirected to the stimuli occurring on the right side. As such, we expected to see an effect on spatial congruence when both visual and vibrotactile stimuli were in the attended side of space, as compare to when there was only a visual stimuli in the attended side of space. A comparative overview of the three experiments can be found in table 1.

Table 1: overview of the three experiments indicating where attention was directed to in space and which modalities were attended. All was modulated through task demand.

	Attended spatial locations	Attended modality
Experiment 1	Two: Left and right	Visual and tactile
Experiment 2	Two: Left and right	Visual
Experiment 3	One: Left or right	Visual and/or tactile

Subsequently, we looked to evidence visuo-tactile interactions using more complex stimuli. A second part looked at the differential visuo-tactile interaction, making a difference between the effect of vibrotactile stimulation on higher-order tool-selective visual processing and low-level visual processing. More particularly, we looked at the higher-order activity related to tools. Tools are believed to automatically activate motor schemes, to which sustained stimulation might be extremely relevant to. The main objective of the second study was to study the possibility to separate tool-selective processing, which is special through its activation of both ventral and dorsal visual stream, from visual processing common to all categories. The subsequent objective of this second study was to demonstrate that it was possible to separate activity related to tools from general image processing which is shared between tools and other image categories using

a variation of FPVS. The main objective of the third study was to build up a sound methodology to study visuo-tactile interactions using pictures of tools, therefore stimulating higher-order brain areas and making the visual stimuli relevant for sustained stimulations. This studies aim was to obtain a modulation of the tool-selective activity, depending on spatially congruent vibrotactile stimulation and the position of the hand in relation to the visual stimulation.

Studying the neural underpinnings of the spatial aspects of multisensory integration using FPVS will allow to uncover aspects of multisensory integration which were impossible to study with the traditional EEG methodologies. This thesis will thus shed a light and lay the basis of studying multisensory integration and its organisation in space using sustained periodic visual stimulation. The basic understanding of how sensory modality organisation is integrated and remapped according to a certain reference frame may advance our basic understanding of longer-lasting visuo-tactile interactions.

PART 1: DOES SPATIALLY CONGRUENT VIBROTACTILE STIMULATION MODULATE VISUAL PROCESSING?

Study 1: Does spatially congruent vibrotactile stimulation modulate visual processing?

Abstract

Covertly directing attention to the location of a stimulus of one sensory modality can affect the processing of the stimulus of another modality when presented at the same location. In a series of three experiments, we explored this visuo-tactile interaction using fast periodic stimulation, suggested to reflect (at least partially) brain processing in primary sensory cortices. Throughout all experiments, two sustained trains of visual stimulations were concomitantly presented in both side of space. At the same time a train of vibrotactile stimulation was applied on the participant's hand. In different blocks, the hand was positioned close to either the left or the right diode. Three different frequencies of stimulation were chosen to tag the specific EEG responses to the three different stimuli (visual left, visual right, tactile). Throughout the three experiments, we modulated which stimulations needed to be attended, by requiring participants to report interruptions, which could occur in the tactile and visual modality (experiment 1 & 3), or only in the visual modality (experiment 2). Specifically, in experiment 3, we modulated which side of space should be attended, by asking participants to only report interruptions in that side of space. Surprisingly, throughout all three experiments, the results show no modulation of the amplitude of the FPVS response when congruent with the vibrotactile stimulation. This suggest that the crossmodal mechanisms known to modulate the ERP components, might not be involved in the FPVS response.

1. Introduction

In everyday life, we are continually required to react to events happening in our direct environment. Most of the time, those events are encoded by more than one sensory modality. In consequence, it is of primary interest that information conveyed by the different senses interact or are even integrated together to enable adequate reactions. Humans are very capable of doing so, and the brain is well adapted to this end Spence, Senkowski, & Röder, 2009). That information conveyed by different sensory modalities presents an advantage has been evidenced through the years. Indeed, reaction times become faster when the stimulation is bimodal, ERP components are enhanced, and BOLD signal is changed as well, showing an interaction between touch and vision (Kennett, Eimer, et al., 2001; Macaluso & Driver, 2001; Macaluso et al., 2000; Macaluso & Maravita, 2010; Volcic, van Rheede, Postma, & Kappers, 2008). Using transient stimuli, it has been shown that these interactions occur when the visual and tactile stimuli share a common spatial location, independently of the position of the body in space (Kennett, Eimer, et al., 2001; Macaluso, 2000). Indeed, we need to be able to localize stimulations in external space, irrespectively of the how the body is positioned in space. Nevertheless, some positions of the body introduce a dealignement of the anatomical frame of references of the different modalities. Interestingly, it remains unclear by what reference frames vision and touch interaction to form a common representation of external space when using sustained trains of stimulation. Sustained visual and tactile stimuli might be relevant for feedback information during action in the peripersonal space.

These visuo-tactile interaction are believed to be, at least partially, dependent on a network of brain areas, containing bimodal neurons (neurons which respond to more than one modality). This network appears to integrate and

remap the different modalities in a multisensory representation of external space (Graziano & Cooke, 2006). Evidence for visuo-tactile interactions has been evidenced in human participants as well. Indeed, visuo-tactile stimuli yielded an enhanced BOLD signal in the visual cortex, but only when the tactile and visual stimulus were spatially congruent, i.e. congruent in space, potentially implying the same multisensory network (Macaluso, Frith, & Driver, 2000). Here again, the authors showed that this enhancement was independent of gaze shift, demonstrating that it was truly the spatial congruence in external space driving this crossmodal interaction (Macaluso, 2000; Macaluso, Driver, van Velzen, & Eimer, 2005; Macaluso et al., 2002). However, it is not clear which processes underlie these interactions in humans. Indeed, results could be explain by multisensory integration, or by a crossmodal link in spatial attention (Kennett, Eimer, et al., 2001). However, measuring the EEG signal to periodic stimuli could be used to look for an objective signature of multisensory integration. Regan introduced the concept of Fast Periodic Visual Stimulation [FPVS] (traditionally referred to as SSEP), where a stimulus at a given frequency causes the energy related to that train of stimuli to be concentrated at the same frequency in the frequency spectrum (Regan, 1966, 1977). This methodology has the advantage of enabling sustained trains of stimuli and frequency tagging (Regan, 1977). Frequency tagging allows to disentangle the processing of simultaneously presented stimuli, as long as they are presented at different frequencies. Interestingly, it has been proposed that it is possible to obtain an 'objective' measure of multisensory integration using FPS. Indeed, an objective measure of multisensory integration would consist of activity concentrated at the sum or subtraction of the two used frequency (f1+f2 or f2-f1) (Regan, 1977). Alternatively, a visuo-tactile interaction could be evidenced by the modulation of visual processing, depending on the presence of a spatially congruent vibrotactile stimulus. We hypothesise that a vibrotactile and a visual

stimulus which are spatially congruent will interact, evidenced through activity at crossmodulation frequencies or through influencing the visual processing as compared to a visual stimulus without spatially congruent vibrotactile stimulus.

Throughout the three experiments of this study, we looked at the interaction between visual and tactile stimuli, and more precisely the effect of vibrotactile stimulus on visual processing depending on the proximity between the visual stimulus and the limb on which the tactile stimulus is applied. In order to do so, the participant was presented with two visual stimuli, one in each side of space (i.e. left- vs. right-sided) with two respective stimulation frequencies, concomitantly with one vibrotactile stimulus applied on the hand at a third stimulation frequency. The hand on which the vibrotactile stimulus was applied was placed close to either the left or the right visual stimulus while the non-stimulated hand rested on the lap of the participant. We tested several factors throughout the three experiments. The main factor of interest in all experiments is the congruence in external space between the visual and vibrotactile stimulation (i.e. the visual stimulation is at the same spatial location as the vibrotactile stimulation). Indeed, we aimed at showing an interaction, due to spatial proximity of vibrotactile stimulation.

Throughout all experiments, we hypothesised that we would have an enhanced response to the visual stimulus that was spatially congruent to the tactile stimulus as compared to the visual stimulus that was not spatially congruent. Indeed, we expect a larger visual response when the vibrotactile stimulation is at the same location in external space, as compared to the same visual stimulation, but when the vibrotactile stimulation is close to the other visual stimulation. The first experiment involved three factors: The visual stimulation could either be spatially congruent with the vibrotactile stimulation (Congruence: CONGRUENT

and INCONGRUENT). Additionally, there were two different frequencies for the visual stimuli (Frequency: 6 Hz and 6.2 Hz). Last, the posture of the arm of the participant was either uncrossed or uncrossed, allowing testing for the configuration of visuo-tactile interactions (Body posture: UNCROSSED and CROSSED). Crossing the hands over the body midline allows testing if the effect is truly due to a congruence in external space, or if it can be attributed to congruent anatomical frames of reference. The second experiment had exactly the same experimental factors. The difference between the first and the second experiment lied in the task demand. In the first experiment, participants were asked to pay attention to both vibrotactile and visual stimulation. This task aimed at modulating the attentional load. Indeed, in the first experiment, the participant had to divide his attention over two modalities, but also in space. This manipulation was done to avoid that participants would focus their spatial attention to one side of space. In the second experiment, participants only had to pay attention to the visual stimulations, and could ignore the vibrotactile stimulations. As such, their attention had to be divided across space, but not across modalities. Indeed, it has been shown that the vibrotactile stimulation does not have to be relevant to observe visuo-tactile interactions (Kennett, Eimer, et al., 2001; Macaluso, 2000). As such, these two experiments allowed studying the contribution of attention to visuo-tactile interactions. In the third experiment, we added a fourth factor: side of space spatial attention was directed at. Consequently, stimuli could be attended or unattended. This results in a fourth variable to control the contribution of spatial attention (Attention: ATTENDED and UNATTENDED). We hypothesised that a visuo-tactile interaction would be observed when both stimulations were in the attended side of space.

2. Experiment 1

2.1 Material & methods

2.1.1 Participants

Fifteen healthy volunteers (8 male, 14 right handed (self-reported), 27.9 ± 4.5 years old ranging from 23 to 38 years old) participated in the study. All had normal or corrected-to-normal vision, no prior history of neurological, psychiatric, or chronic pain disorder, and no current use of psychotropic or analgesic drugs. Experimental procedures were approved by the ethics committee of the Université catholique de Louvain and conformed to the latest revision of the declaration of Helsinki. Written informed consent of each participant was obtained.

2.1.2 Stimuli

2 blue Light Emitting Diodes (LED) were used for visual stimulation (LED OVS-3303, multicomp, 540 mcd, 120° diffusion angle). Both LEDs were activated simultaneously at different frequencies to allow frequency tagging. As such, the activity at the defined frequencies can be related back to the processing of a stimulus at that same frequency. LED stimulations followed a sinusoidal modulation of intensity of luminance fluctuating between 0 and 3 volt at one of the following frequency: 6 Hz and 6.2 Hz.

Vibrotactile stimuli were generated by a recoil-type vibrotactile transducer driven by a standard audio amplifier (Haptuator, Tactile Labs Inc., Canada) and positioned on the dorsal surface of the hand. The vibrotactile stimulus lasted 16 seconds and consisted of a 255 Hz carrier frequency periodically modulated at 21Hz. The participants perceived the stimulation trains as a continuous vibration.

2.1.3 Procedure

Participants were comfortably seated at a table in a dark room. The haptuator was placed on the dorsum of either left or right hand (counterbalanced across participants). Participants listened to white noise set at a comfortable levels during the experiment to mask sounds of the haptuator. The experiment consisted 4 blocks of 14 trials, leading to a total of 56 trials. A trial consisted of two visual stimuli, one on each side of space (with a fixation cross in the middle), and a vibrotactile stimuli delivered to one single hand. Stimuli started simultaneously and lasted 16 sec. Four different conditions were tested. Across blocks, the hand on which the vibrotactile stimulus was applied was placed in an uncrossed vs. crossed posture so that it was near either the ipsilateral or the contralateral visual stimuli (Figure 1).

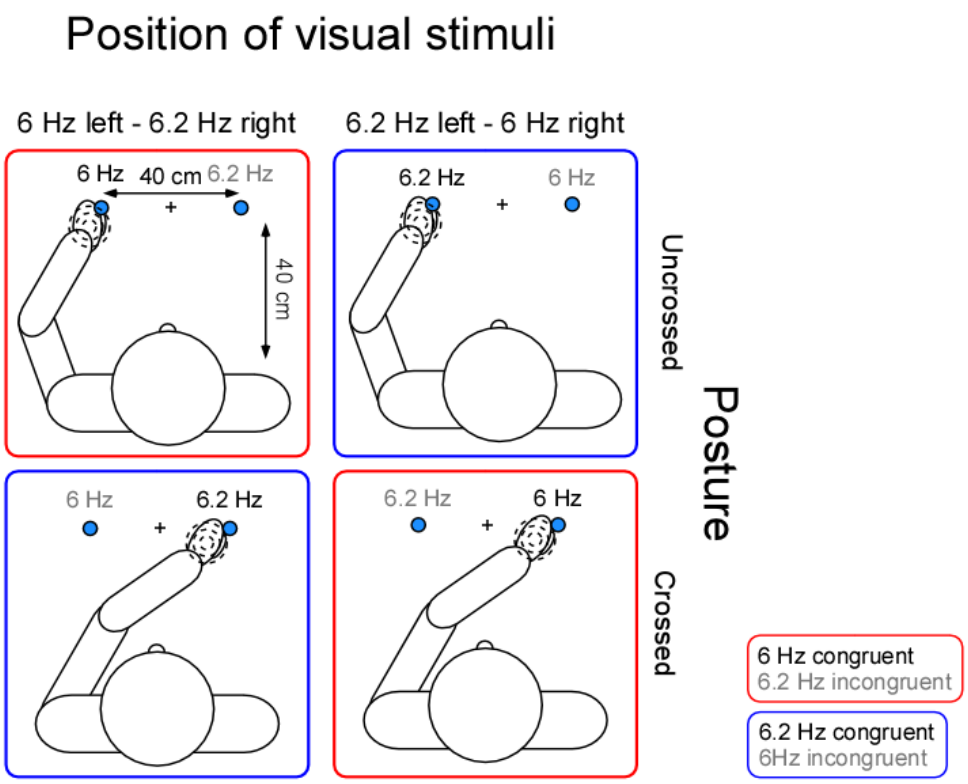

Figure 1: Experimental set-up of experiment 1 and 2. Participants were simultaneously presented with three trains of stimulation. One vibrotactile applied on the hand dorsum (21 Hz) and two visual stimulations (6 Hz and 6.2 Hz). The stimulated hand was placed next to one of the two visual stimuli, in a crossed or uncrossed position. Each frequency could be presented either on the left or the right of the body midline. Within one trial, the vibrotactile stimulation was always spatially congruent with one of the visual stimuli, and spatially incongruent with the other. Consequently, we extract two measures from each trial: one at 6 Hz and one at 6.2 Hz. We expected to see larger activity for 6 Hz when congruent with the vibrotactile stimulus as compared to when incongruent (i.e. red frame > blue frame). Correspondingly, we would expect to observe larger activity for 6.2 Hz when

spatially congruent with the vibrotactile stimulus as compared to when incongruent (blue frame > red frame).

The position of the stimulated hand was counterbalanced across blocks. The visual stimuli were presented at 6 Hz and 6.2 Hz either in the left and right sides of space respectively, or in the reverse order, randomized across trials (Figure 1). During 4 trials out of the 14 trials of each block, stimulation streams were interrupted by a 200-ms duration gap in the vibrotactile stimuli and in one of the visual stimuli. These stimuli were considered as target trials as participants were instructed to detect the interruptions and to report whether the tactile and the visual gaps occurred simultaneously or not. Half of the interruptions occurred in a simultaneous fashion, while the other half were asynchronous. Target stimuli were presented only in order to encourage participants to pay attention to the stimulation streams, and were discarded from further analyses. It resulted in 40 non-target trials in total considered for further analyses, 10 for each of the 4 conditions resulting from the combination of congruence (CONGRUENT and INCONGRUENT to the stimulated hand) and body posture (UNCROSSED and CROSSED) (both frequencies were present in each trial).

2.1.4 Measures

Behavioural results (correct or incorrect) consisted of the reporting of interruptions and the percentages correct per condition were used to make sure that the participant was paying attention towards the stimuli and not randomly reporting interruptions. A response was considered as correct if participants reported the timing of the interruptions correctly. Participants with a performance below 80% were excluded from further analysis.

The EEG was recorded using 64 Ag–AgCl electrodes placed on the scalp according to the international 10/10 system (Waveguard64 cap, Cephalon A/S, Denmark). Electrode impedances were kept below 10 kΩ. Ocular movements and eye blinks were recorded using two additional surface electrodes placed at the upper-left and the lower-right sides of the right eye. Signals were amplified and digitized using a sampling rate of 1000 Hz, using an average reference (64-channel high-speed amplifier, Advanced Neuro Technology, The Netherlands).

All EEG processing steps were carried out using Analyzer 1.05 (Brain Products, Germany), Letswave (https://www.letswave.org/; see also (Mouraux and Iannetti, 2008)) and Matlab (The MathWorks, USA). Statistical analyses were performed using R3.1.2 and RStudio 0.98.1087. Non-overlapping EEG epochs were obtained by segmenting the recordings from -2 to 18 s relative to the onset of each stimulus train, thus yielding a total of 40 epochs per participant (10 stimulation trains× 2 body posture [uncrossed - crossed] x 2 proximity [congruent-incongruent]). Continuous EEG recordings were filtered using a notch filter on 50 Hz first, and subsequently a 0.5 Hz – 100 Hz band pass Butterworth zero-phase filter (order=4) to remove slow drifts and high frequencies in the recorded signals non relevant to the present study. Artefacts due to eye blinks or eye movements were then removed using a validated method based on an independent-component analysis (FastICA algorithm) (Hyvärinen & Oja, 2000). For each participant, components linked to either horizontal or vertical eye movements were manually removed and kept to a minimum. We removed maximum two components per participant. All epochs were then cropped from 1 second to 16 seconds. The first second of the train was removed to avoid the evoked potential elicited by the stimulation (Nozaradan et al., 2015). For each participant, posture and frequency group, EEG epochs were averaged such as to attenuate the

contribution of activities non phase-locked to the stimulation train. Finally, the obtained average waveforms were transformed in the frequency domain using a discrete Fourier Transform (FFTW) (Frigo & Johnson, 1998), yielding an amplitude spectrum (μv) ranging from 0 to 500 Hz with a frequency resolution of 0.067 Hz (Bach & Meigen, 1999). Within the obtained frequency spectra, the signal amplitude at the frequencies of 6 Hz and 6.2 Hz was measured. These signal amplitude values may be expected to correspond to the sum of the stimulus-evoked steady-state response (i.e. the FPS response, if present) and unrelated residual background noise. To determine which harmonics to include in the analysis, we performed a z-score baseline correction (Alonso-Prieto et al., 2013) to the 20 adjacent bins on the average of all channels and all participant. Subsequently, we did the same per participant to see if the overall significant frequencies stood out in each participant. Parallel, to obtain valid estimates of the magnitude of the recorded FPVS responses, the contribution of this residual noise was removed by subtracting, at each electrode and at each frequency bin, the average amplitude of the signal measured at neighbouring frequencies (four frequency bins ranging from −0.67 to −0.87 Hz and four frequency bins ranging from +0.67 to +0.87 Hz relative to each frequency bin) (Mouraux, Diukova, Lee, Wise, & Iannetti, 2011). We then extracted all amplitudes at the frequencies and harmonics per participant from this noise-subtracted frequency spectrum which yielded a z-score of 1.64 or larger. Due to large interparticipant variability, activity was averaged over all posterior electrodes (TP7, CP5, CP3, CP1, CPz, CP2, CP4, CP6, TP8, P7, P5, P3, P1, Pz, P2, P4, P6, P8, PO7, PO5, PO3, POz, PO2, PO4, PO8, O1, Oz, O2). Only trials without interruptions were included in the analysis. As a results of the z-score signal to noise ratio test, we summed up the amplitudes of the frequencies which reached the selection criterion for each participant: 6 Hz, 12 Hz,

18 Hz and 6.2 Hz, 12.4 Hz and 18.6 Hz, as harmonics can be considered as part of the response (Zhou, Melloni, Poeppel, & Ding, 2016).

2.1.5 Analyses

We tested for a main effect of frequency (6 Hz & 6.2 Hz), body posture (UNCROSSED & CROSSED) and congruence (CONGRUENT & INCONGRUENT), as well as for interactions between these factors using a repeated measure ANOVA. The factor body posture indicates if the hand was crossed over the body midline. The factor congruence indicates if the visual stimulation was congruent with the vibrotactile stimulation. We explore significant interactions with paired-sample Student t-tests when justified.

2.2 Results

All participants performed above 80% correct on perceiving and reporting the interruptions.

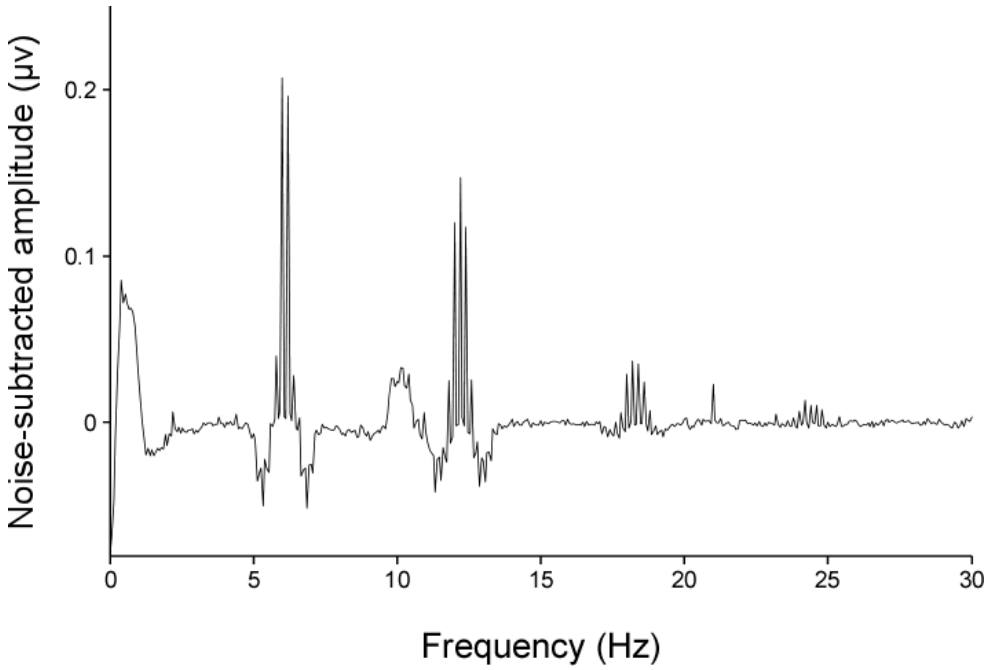

Figure 2: Frequency spectrum of the grand average over all participants and all electrodes. Fundamental frequencies at 6 and 6.2 Hz as well as side bands, harmonics (2F & 3F) and cross modulations can be seen in the spectrum.

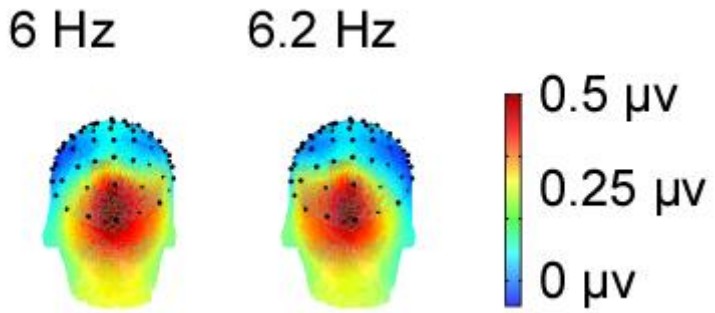

Figure 3: Topographies related to the base frequency of stimulation, 6 Hz and 6.2 Hz. Both topographies are characterized by occipital activation

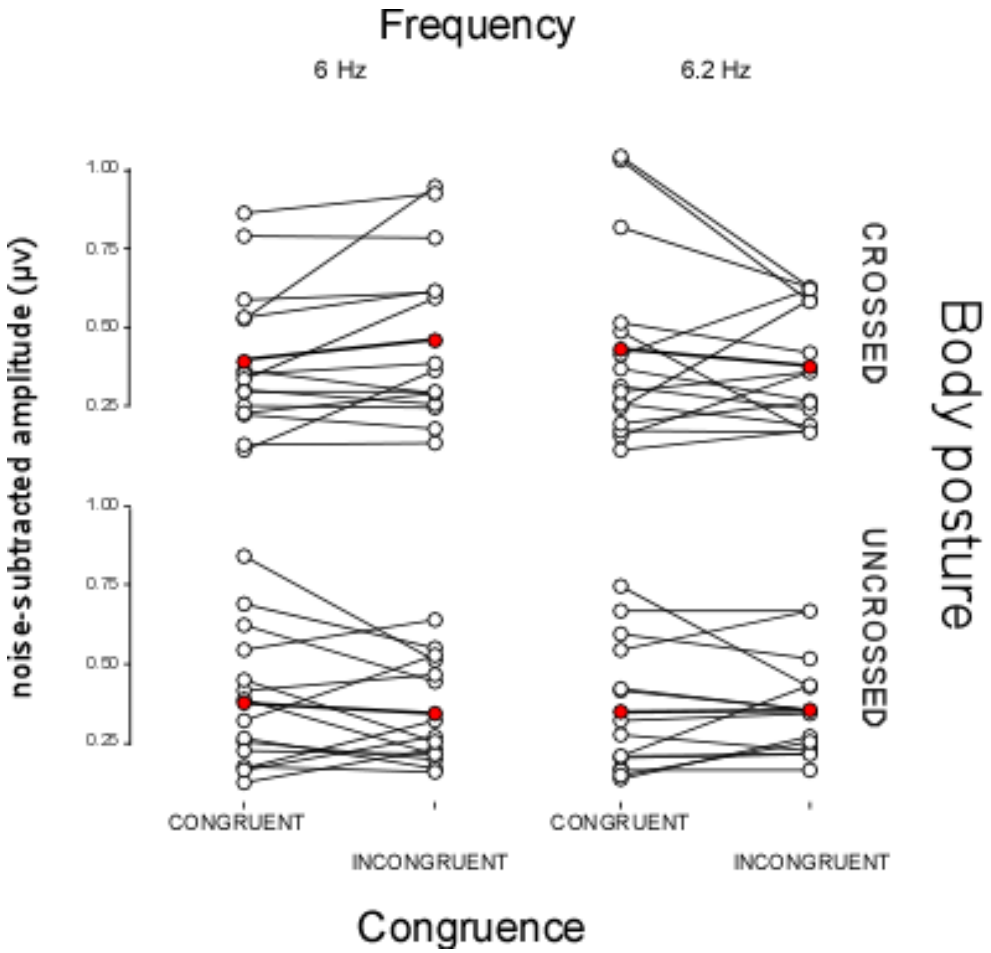

Figure 4: Amplitudes to the visual stimulation depending on condition (congruence, frequency and body posture). Each line depicts the amplitude of an individual when the visual stimulation is congruent (close to the hand) as compared to incongruent (not close to the hand). Individual values are denoted by thin lines and white connectors. The red dots and black thick lines shows the group average.

Table 2: Repeated measure ANOVA with three factors: frequency (6 Hz & 6.2 Hz), congruence (CONGRUENT & INCONGRUENT) and body posture (UNCROSSED & CROSSED). The repeated measure shows a main effect of body posture.

Effect	DFn	DFd	F	p	Ges
frequency	1	14	2.71	0.12	0.001
congruence	1	14	0.01	0.93	0.000
Posture	1	14	5.02	0.04	0.018
frequency:congruence	1	14	0.99	0.34	0.002
Frequency:posture	1	14	2.12	0.17	0.000
congruence: posture	1	14	0.44	0.52	0.001
frequency:congruence: posture	1	14	2.21	0.16	0.009

EEG results

We obtain a frequency spectrum with clear responses at the frequency of interest (6 & 6.2 Hz) and harmonics up to F3 (Figure 2). Interestingly, we observe large within modality interactions reflected at 12.2 for example. Topographies to the visual stimuli are characterized by large occipital activity (Figure 4). We also observe a clear response to the vibrotactile stimulation at 21 Hz. The repeated measures ANOVA yielded a significant main effect of body posture ($F(1,14)=5.02$, $p=0.04$, $\eta^2_p = 0.018$) (Table 2). Indeed, amplitudes of the FPVS response were significantly larger when in the hand was crossed over the body midline as compared to when the hand is uncrossed (0.41 µv ± 0.24 and 0.36 µv ± 0.18 respectively). All other p-values were larger than 0.1 (Figure 3).

2.3 Interim discussion

In the present experiment, we investigated the effect of a spatially congruent vibrotactile stimulation on the processing of the visual stimulations,

relying on the theory of multisensory integration, which states that two stimulations delivered in different modalities have more chance to be integrated or interact when they are close together in space (Meredith & Stein, 1996). We show that, through FPVS, we are not able to demonstrate an effect of spatial congruence. Indeed, amplitude of the FPVS response was not modulated by the presence of a spatially congruent vibrotactile stimulation. In the present experiment, we asked participants to pay attention to both modalities to be able to fulfil the task, with the idea that taking both modalities into account enhanced the chance to integrate the information. As such, participants had to divide their attention between vibrotactile and visual stimulation, but also in space, as visual stimulations occurred in the left and right side of space. However, there is evidence that paying attention to the stimulation in both modalities is not required to show the effect of the spatial rule behaviourally and with ERPs (Eimer & Driver, 2000). Also, in exogenous cueing paradigms, participants have to react to target stimulations, which are preceded by non-predictive cues. These kind of paradigm often demonstrated that cues, albeit non-predictive, can affect the processing of the target stimulus (Kennett, Eimer, et al., 2001). Using this paradigm, Kennett, Eimer, et al. (2001)shows that participants were faster to react to visual stimuli, when it is preceded by a short tactile stimulation at the same spatial location. Participants were asked to report the spatial location of the target stimulus. They demonstrate a facilitation effect, but only when the tactile and visual cue are at the same spatial location. However, this most likely is related to a crossmodal link in spatial attention, which does not imply that two stimulations are processed further or interact. This study showed that paying attention to both sensory modalities was not necessary to benefit from the spatial rule. In the next experiment, all experimental parameters were kept exactly the same, but participant had to report interruptions in the visual modality, and ignore the vibrotactile stimulation.

3. Experiment 2

3.1 Material & methods

3.1.1 Participants

Ten healthy volunteers (3 male, 9 right handed (self-reported), 23.5 ± 3.8 years old ranging from 19 to 32 years old) participated in the study. All had normal or corrected-to-normal vision, no prior history of neurological, psychiatric, or chronic pain disorder, and no current use of psychotropic or analgesic drugs. Experimental procedures were approved by the ethics committee of the Université catholique de Louvain and conformed to the latest revision of the declaration of Helsinki. Written informed consent of each participant was obtained.

3.1.2 Stimuli

Visual and vibrotactile stimuli were the same as in the first experiment.

3.1.3 Procedure

Participants were comfortably seated at a table in a dark room. The haptuator was placed on the dorsum of either left or right hand (counterbalanced across participants). Participants listened to white noise set at a comfortable levels during the experiment to mask sounds of the haptuator. The experiment consisted 4 blocks of 10 trials, leading to a total of 40 trials. A trial consisted of two visual stimuli, one on each side of space, and a vibrotactile stimuli delivered to one single hand. Stimuli started simultaneously and lasted 20 sec. The conditions were the same as in experiment 1 (Figure 1). Conversely to the previous study, participants had to report the temporal order of the 2 interruptions (200ms) in the visual

stimuli, which always occurred in the last 4 seconds of each trial. The timing of the interruptions in the visual stimuli was random. Possible outcomes were thus left first, right first, or simultaneous. Vibrotactile stimuli could be disregarded and were not relevant for the task. Interruptions were presented in order to encourage participants to pay attention to the stimulation streams, and the four last seconds of each trial were discarded from further analyses. It resulted in 40 non-target trials of 16 sec in total considered for further analyses, 10 for each of the 4 conditions resulting from the combination of congruence (CONGRUENT and INCONGRUENT to the stimulated hand) and body posture (UNCROSSED and CROSSED) (Figure 1).

3.1.4 Measures

Behavioural results, EEG recording and processing were the same as in the first experiment.

3.1.5 Analyses

Analyses were identical to the ones of the first experiment.

3.2 Results

All participants performed above 80% correct on perceiving and reporting the interruptions.

We obtain a frequency spectrum with clear responses at the frequency of interest (6 & 6.2 Hz) and harmonics up to F2 (Figure 5). Interestingly, we observe large within modality interactions reflected at 12.2 for example. Topographies to the visual stimuli are characterized by large occipital activity (Figure 6). In the

repeated measure ANOVA on the amplitude of the FPVS response, we observe a interaction between frequency and posture ($F(1,9)= 8.79$, $p= 0.01$, $\eta^2_p= 0.002$). All other effects were non-significant. Subsequently testing the interaction with paired t-test reveals no further differences (Figure 7).

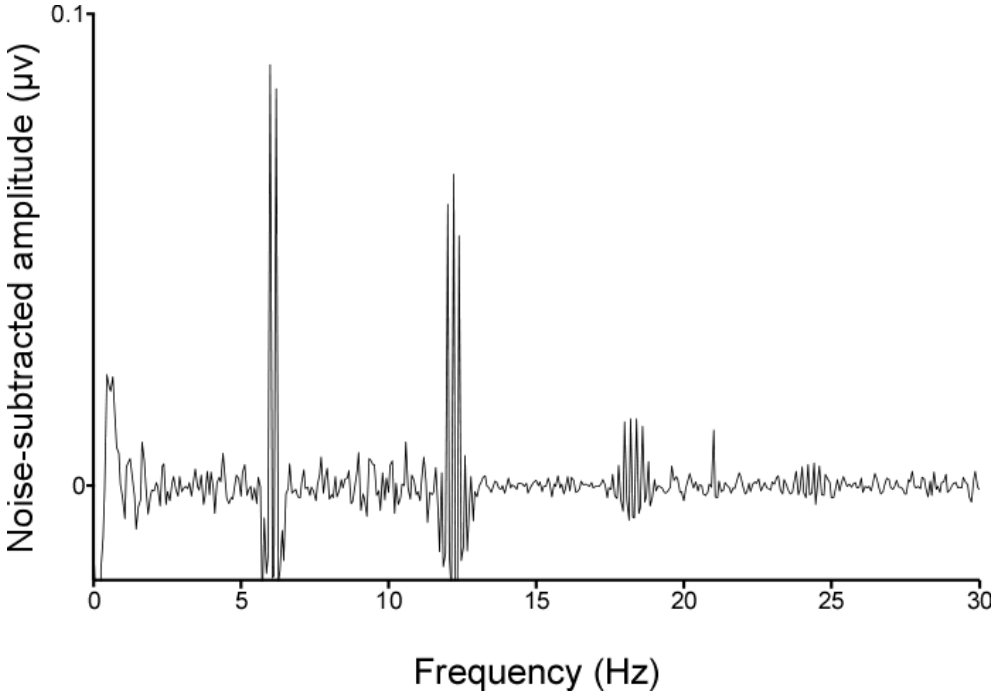

Figure 5: The frequency spectrum shows a peak at the fundamental frequencies (6 and 6.2), as well as harmonics and cross modulations. The vibrotactile stimulation is reflected at 21 Hz in the frequency spectrum.

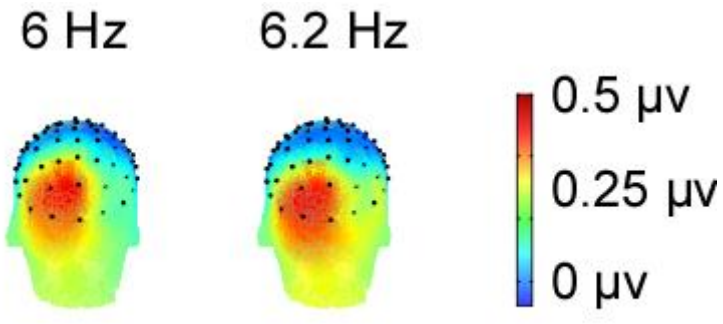

Figure 6: Topographies to 6 and 6.2 Hz are characterized by an occipital activity.

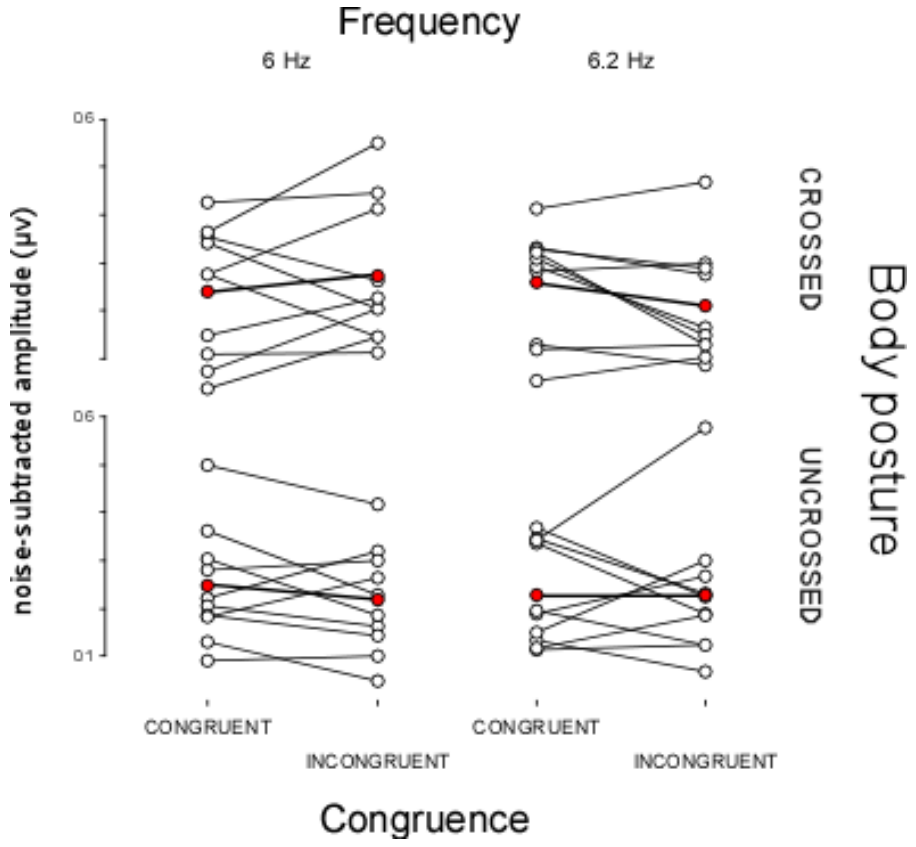

Figure 7: Individual values depending on condition. The amplitude of the FPVS signal is not modulated by frequency used, body posture, or by the presence of spatially

congruent vibrotactile stimuli. Individual values are denoted by thin lines and white connectors. The red dots and black thick lines show the group average.

3.3 Interim discussion

In the second experiment, we aimed at modulating the processing of the visual stimulations depending on the presence of congruent vibrotactile stimulation. We showed that, through FPVS, we were not able to demonstrate a modulation of the visual processing. Indeed, the amplitude of the FPVS response was not modulated by the presence or absence of a spatially congruent vibrotactile stimulation. However, in the present study, we showed that the amplitude of the activity due to the FPVS is modulated differently depending on frequency and posture with p-values close to 0.05 and small eta squares. Subsequently exploring this interaction did not reveal any significant effects. The posture and frequency interaction is of no interest to the present hypothesis.

One way of knowing if it is even possible to modulate the FPVS response is by directing spatial attention to one side of space. Previous studies have often concluded that this cross-modal link might exist in spatial attention. As crossmodal attentional effects occur when attention is redirected to one side of space by a stimulation of one modality, which in turn modulates/facilitates the processing of the stimulation in the other modality (Eimer & Driver, 2000). By explicitly directing spatial attention to one side of space, we expected to observe a visuo-tactile interaction, but only when spatial attention was directed to it. Responses were expected to be larger for visual stimulation when it was congruent with the vibrotactile stimulation and in the attended side of space.

4. Experiment 3

4.1 Material & methods

4.1.1 Participants

Eight healthy volunteers (2 male, 8 right handed (self-reported), 23.8 ± 2.1 years old ranging from 21 to 27 years old) participated in the study. All had normal or corrected-to-normal vision, no prior history of neurological, psychiatric, or chronic pain disorder, and no current use of psychotropic or analgesic drugs. Experimental procedures were approved by the ethics committee of the Université catholique de Louvain and conformed to the latest revision of the declaration of Helsinki. Written informed consent of each participant was obtained.

4.1.2 Stimuli

Visual stimuli and vibrotactile stimuli were the same as in the first two experiments.

4.1.3 Procedure

Participants were comfortably seated at a table in a dark room. The haptuator was placed on the dorsum of either left or right hand (counterbalanced across participants). Participants listened to white noise set at a comfortable levels during the experiment to mask sounds of the haptuator. The four same conditions as in experiment 1 and 2 were tested, with the factor frequency, congruence and posture. Additionally, in half of the blocks, participants were asked to only attend the stimulations occurring in the left side of space, while in the other blocks, they only had to attend the stimulations occurring in the right side of space. The side

they had to pay attention to was randomised across blocks (Figure 8.). This results in an experiment of 8 blocks of 13 trials, leading to a total of 104 trials. As in experiment 1, participants were asked to report interruptions in the vibrotactile and the visual modality. Trials with interruptions were considered as target trials and participants were instructed to only detect the interruptions and to only report whether the tactile and the visual gaps occurred simultaneously or not when they occurred in the attended side of space. As in experiment one, the interruptions always took place in the vibrotactile stimulation, and one of the visual stimulations. Their occurrence was unrelated to the side participants had to pay attention to. For example, when the stimulated hand was next to the right LED and the participant was asked to attend the right side, one interruption occurred in the vibrotactile stimulation, which was attended, and a second interruption occurred in one of the visual stimuli (the attended or unattended one). If the interrupted visual stimulus was attended, the subject then had to evaluate if the interruptions in the vibrotactile and visual modality were simultaneously or not. In the other case (i.e. the second interruption occurred in the unattended visual stimulus), the subject only reported the interruption in the vibrotactile stimulation, and had to ignore the interruption in the unattended visual stimulus. Following this logic, when the stimulated hand was next to the right LED but the participant was instructed to pay attention to the left, he only reported an interruption if the interruption occurred in attended LED. He had to disregard the interruption in the vibrotactile stimulation, as well as the interruption in the unattended visual stimulus. 3 out of 13 trials were target trials. This resulted in 80 non-target trials in total considered for further analyses, 10 for each of the 8 conditions resulting from the combination of the frequency (6Hz and 6.2 Hz), congruence (CONGRUENT and INCONGRUENT to the stimulated hand) and body posture (CROSSED and UNCROSSED), and attention (ATTENDED and UNATTENDED).

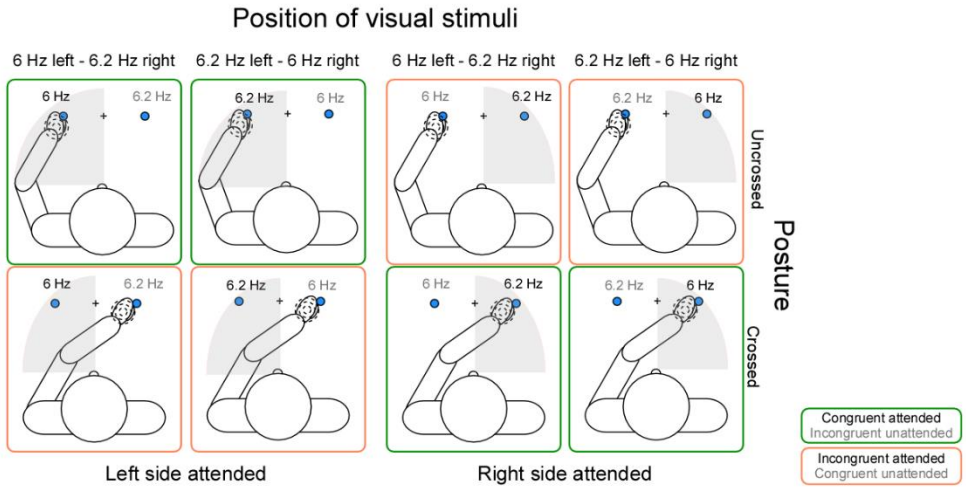

Figure 8: Participants were presented with three stimulations: two visual stimulations (6 Hz and 6.2 Hz) and a vibrotactile stimulus applied on the hand (21 Hz). The stimulated hand was placed next to one of the visual stimuli, in a crossed or uncrossed posture. As such, one of the two visual stimuli was spatially congruent with the vibrotactile stimulus in each condition. Additionally, subjects were asked to pay attention to one side of space (either left of right). Participants were required to report interruptions, which occurred in the vibrotactile stimulus and one of the two visual stimulus. I.e. when the vibrotactile and visual stimuli are congruent and attended, one interruption would occur in the vibrotactile stimulus, and one in one of the two visual stimuli (not necessarily the attended one). Conversely, when the stimulated hand is not in the attended side of space, participants only had to report an interruption of the visual stimulus, if it occurred in the attended side of space. We expected a visuo-tactile interaction when the vibrotactile and visual stimuli are congruent and attended (green frames), but not when the vibrotactile and visual stimuli are congruent but unattended (orange frames)..

4.1.4 Measures

Behavioural results, EEG recording and processing were the same as in the previous experiment.

4.1.5 Analyses

We performed a repeated measures ANOVA with four factors with two levels each: congruence (CONGRUENT and INCONGRUENT), body posture (UNCROSSED and CROSSED), frequency (6 Hz and 6.2 Hz) and attention (ATTENDED and UNATTENDED). We explored significant interactions with paired-sample Student t-tests when justified.

4.2 Results

All participants performed above 80% correct on perceiving and reporting the interruptions.

We obtain a frequency spectrum with clear responses at the frequency of interest (6 & 6.2 Hz) and harmonics up to F2 (Figure 9). Again, we observe large within modality interactions reflected at 12.2 for example. Topographies to the visual stimuli are characterized by large occipital activity (Figure 10). Running the repeated measure ANOVA on the amplitude of the FPVS response, we observe an effect of attention ($F(1,7)=13.34$, $p= 0.008$, $\eta^2_p = 0.104$) as well as an interaction between attention and congruence ($F(1,7)=6.26$, $p= 0.041$, $\eta^2_p = 0.012$), as expected. We subsequently tested this interaction using paired t-test. We observe a significant difference between the attended and the unattended visual stimuli when the vibrotactile and visual stimulations have a different spatial location ($t(7)= 4.86$, $p=0.002$), i.e. when the vibrotactile stimulation is not spatially congruent to

the visual stimulus. This difference does not appear when the spatial location of the visual and the vibrotactile stimulations is congruent ($t(7)=2.03$, $p=0.08$). However, the expected difference (a difference between congruent and incongruent when attended) could not be shown ($t(7)=-1.28$, $p=0.24$). Additionally, there is also no effect of congruence when the stimulations occur in the unattended side of space ($t(7)=-1.04$, $p=0.33$) (Figure 11).

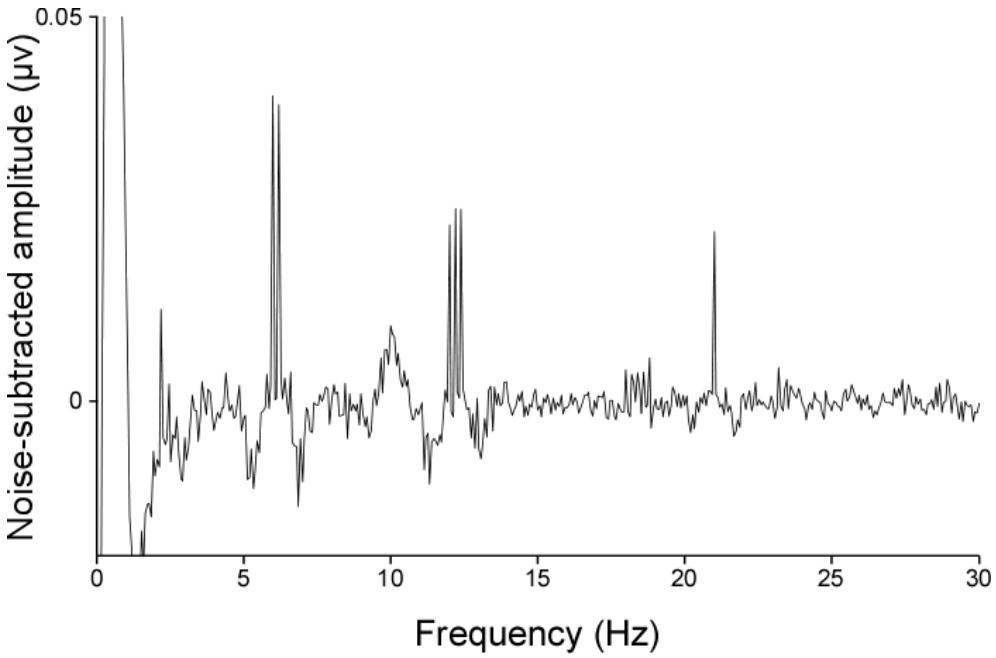

Figure 9: The frequency spectrum shows activity located at the fundamental frequencies (6 and 6.2 Hz), as well as at the harmonics and cross modulations. Vibrotactile activity at 21 Hz also yields clear activity in the frequency spectrum.

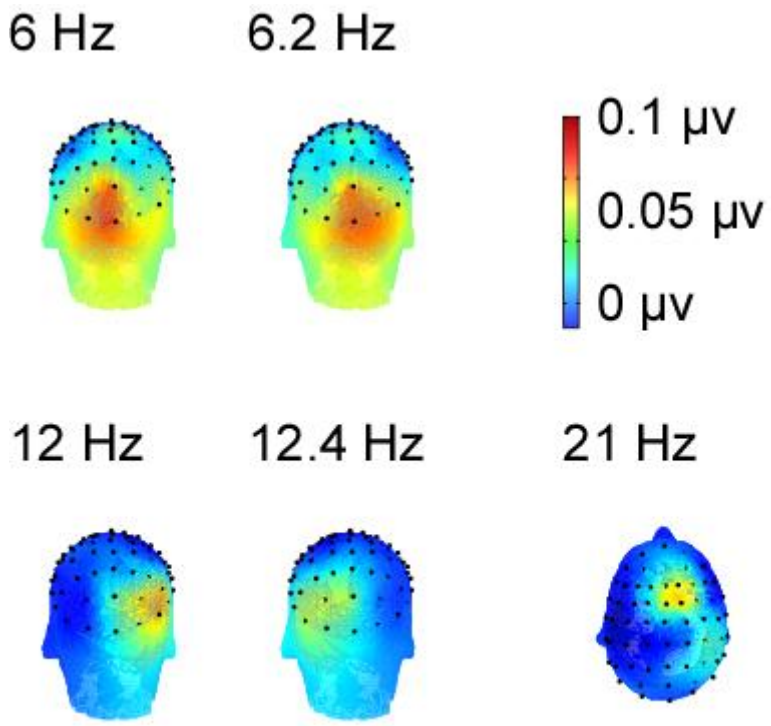

Figure 10: Topographies at 6 and 6.2, as well as on the first harmonic of both fundamental frequencies. Whereas topographies at the fundamental frequencies are occipital, the topographies of the first harmonic is more lateralized, contralateral to the side of stimulation. The vibrotactile stimulation at 21 Hz yields a clear topography reflecting activity in S1.

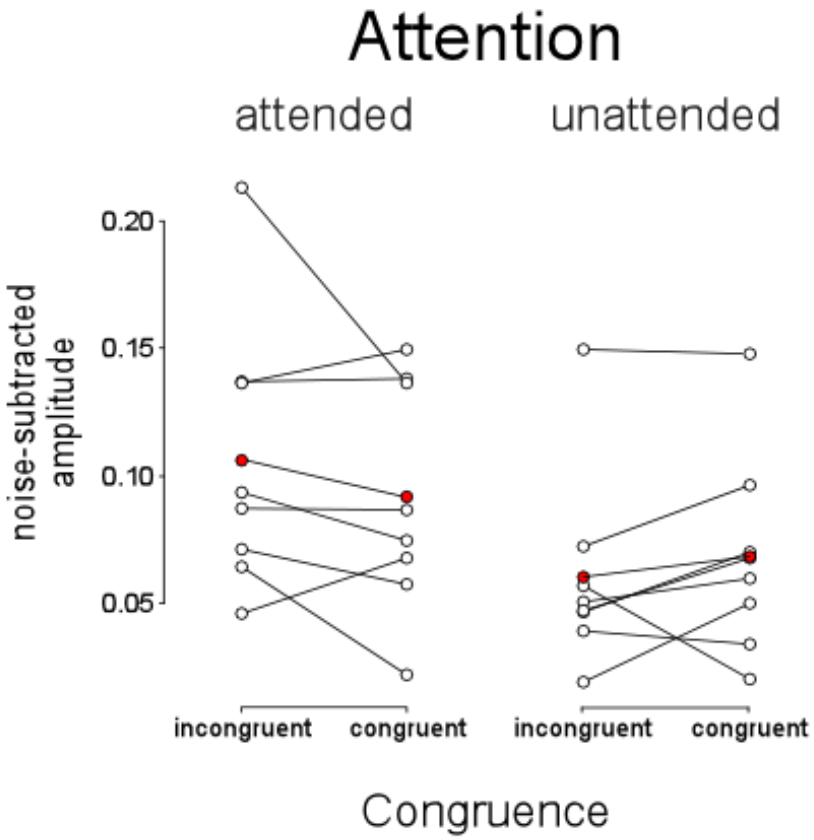

Figure 11: Pooled individual values showing the interaction between attention and congruence. The only significant difference is between congruent and incongruent, when unattended. Individual values are denoted by thin lines and white connectors. The red dots and black thick lines shows the group average.

4.3 Interim discussion

We observed an effect of attention, with a general larger amplitudes when the visual stimuli was attended as compared to unattended, but only when there is no vibrotactile stimulation congruent to the visual stimuli. Indeed, when the vibrotactile stimulus was spatially congruent to the visual stimuli, this attentional effect did not reach significance. We expected an effect of congruence when the stimuli were in the attended side of space. The theory of a crossmodal link in spatial attention implies that when paying attention to a stimulation in a particular location, this also enhances the processing of stimulation of different sensory modalities, demonstrating a crossmodal link. Indeed, it has been shown that paying attention to a spatial location to tactile and visual stimulation can enhance both visual and tactile ERPs (Eimer & Driver, 2000). This has however never been shown with FPVS paradigms using sustained stimulations. It thus appears that accessing space as a multisensory construct holds for short lasting events, but not with long-lasting effects. Similarly, attention has been shown to affect the amplitude of FPVS reliably (Müller, Picton, et al., 1998; Müller & Hillyard, 2000). In this study, we demonstrate an effect of attention, but only when there is no spatially congruent vibrotactile stimulation. It thus appears that adding a congruent vibrotactile diminishes this effect, and in this case. An alternative explanation resides in the task. Indeed, when the vibrotactile was not spatially congruent, participants were only required to pay attention to the visual stimulus. However, when the vibrotactile stimulation was spatially congruent, the task demand changed. Participants had to pay attention to both attended stimulation stream, and pay attention to their timing to perform the task correctly. As such, it is possible that participants paid less attention to the visual stream, which led to a diminished effect of attention on the visual processing.

5. Discussion

Interactions between different sensory modalities presented at the same location in external space has been shown multiple times through different methodologies, behaviourally as well as with ERP or fMRI technologies (Diederich et al., 2003; Eimer & Driver, 2001; Hartcher-O'Brien, Gallace, Krings, Koppen, & Spence, 2008; Macaluso et al., 2000; Macaluso, Frith, & Driver, 2001; Spence, 2010; Spence, McDonald, & Driver, 2004; Zhang, Hong, Gao, Gao, & Roder, 2011). Albeit results being substantial, the frame of reference by which this interaction occurs has not yet been studied using multiple simultaneous sustained stimulations. Additionally, the processes underlying this phenomenon are still debated. The interaction can be explained by three different processes. Multisensory integration states that information from different senses converges to form a coherent representation of the external space (Spence et al., 2009). Second, crossmodal interaction states that information from different senses are able to influence each other's processing at early stages, through direct or feedback connection (Spence et al., 2009). Lastly, the interaction between different senses can be explained by a crossmodal link in spatial attention, where the stimuli are not processed together but are their individual processing is facilitated by the relocation of attentional resources to a shared location in external space (Spence, 2010).

While visuo-tactile interactions are traditionally obtained using transient stimulations, it is impossible to make the difference between the three theories. A way to potentially differentiate between the theory of a crossmodal link in attention and the two others is by using sustained trains of stimulations, as attentional effects are believed to be short-lived (Kennett, Eimer, et al., 2001). As such, FPS presents multiple advantages over the other more traditional

methodologies to study visuo-tactile interactions through frequency tagging, the possibility to present multiple sustained stimulations simultaneously (Colon, Legrain, et al., 2012). Altogether, FPS offers the possibility to overcome the shortcomings of studies using transient stimulation, and is therefore potentially more suited to study the interaction between the sensory modalities.

The aim of the experiments described above was to show that an interaction between visual and vibrotactile stimulation could be obtained using FPS and to study by which spatial frame of reference this interaction was coordinated. In the first experiment, we aimed at showing an interaction between visual and spatially congruent tactile stimuli. Indeed, we expected to see larger activity for cortical processing, as indexed by FPS amplitude, elicited by the visual stimulus at the location spatially congruent with that of the hand on which the tactile stimulus was applied, as compared to activity elicited by the visual stimulus presented at the opposite side. Participants were required to pay attention to all stimulation streams simultaneously to be able to comply with the task, i.e., to report interruptions. We show an effect of body posture, with larger amplitudes in response to visual stimuli when the arm was crossed over the body midline as compared to when the body was in an uncrossed position.

In the second experiment, we again aimed at showing a modulation of the visual processing by the spatial proximity of the vibrotactile stimulation using exactly the same parameters as in the first experiment. The only change was the task demand. Here, participants were asked to only pay attention to the visual stimulation to fulfil task requirements and report interruptions. We report an interaction between the frequency and body posture. However, exploring this interaction with paired t-tests, no significant effects were revealed.

In the third and last experiment, we aimed at showing the interaction between vision and touch through spatial proximity using the same parameters as the first experiment. However, we also modulated spatial attention, as participants were required to only pay attention to stimuli occurring in one single side of space. We showed an interaction between congruence and attention. Subsequently testing this interaction, we reveal that paying attention to a visual stimulus only increases the amplitude when there is no spatially congruent vibrotactile stimulation.

Despite the effect of posture in one study, or an interaction in the second one, not replicating findings throughout experiments with a same set-up indicates very small effect size, lack in power, or false positives, due to small sample sizes. Consequently, the lack of reproducibility and the lack of effect of congruence could be due to a lack of power, and chances of observing multisensory integration or crossmodal interactions could increase using a larger sample size. However, by testing the same hypothesis over three experiments, we never succeed to show an effect of congruence.

Overall, the experiments did not meet the expectations set studying an effect of congruence using FPVS through the modulation of amplitude. Additionally, we expected to observe an 'objective' measure of multisensory integration i.e. crossmodulation frequencies (i.e. between visual and vibrotactile which was expected to show at $f_{visual} + f_{vibrotactile}$). Indeed, we expected to see activity at crossmodulation frequencies (i.e. the sum of the frequency of vibrotactile stimulation, and the frequency of thee visual stimulation which was congruent in external space). However, crossmodulation frequencies were absent in all three experiments. This is in accordance with the results of Giani et al. (2012) which shows crossmodulation frequencies only between stimuli of the same

sensory modality, and failed to show such activity between stimuli of different modalities.

Our results are in accordance with the results of Colon, Legrain, Huang, & Mouraux, (2015), which have shown a modulation of the amplitude of the vibrotactile FPS signal depending on the spatial location of a visual stimulus. However, they failed to show a modulation of the visual signal. It seems therefore possible to observe an effect of visual stimulation on vibrotactile processing, but more difficult to observe the inverse using multiple sustained stimulations.

The effect of attention processes are largely investigated using FPVS (Ding et al., 2005; Malinowski et al., 2007; Morgan et al., 1996; Müller & Hillyard, 2000; Müller, Teder-Sälejärvi, & Hillyard, 1998; Teder-Sälejärvi et al., 2005). Focusing participant's attention on a stream of visual stimulation usually induced an increase of the amplitude of the FPVS response as compared to conditions during attention is focused on a different stimulation. Interestingly, in the third experiment, we showed that the expected effect of attention was only obtained when there was no vibrotactile stimulation spatially congruent to the visual stimuli i.e. spatial attention does modulate the amplitude of the visual FPVS response, but only when there is no spatially congruent vibrotactile stimulation present.

Different stimulation frequencies can tag different networks, and attention has been shown to differently affect different frequencies. This demonstrates that the attentional effect can be interfered with. Indeed, Ding, Sperling, & Srinivasan (2005) showed differential effects of attention on random disc search arrays depending on frequency of stimulation. They showed a marked increase in visual response when the stimuli are attended as compared to unattended at frequencies in delta band and in the upper alpha band. Interestingly, they showed the inverse effect (i.e. attention diminished the amplitude of the FPVS response) for lower

alpha band. Of particular interest here, they state that the FPVS power was almost indistinguishable from noise in the beta band, including 6HZ, and was not modulated by attention. In contradiction with these results, we showed a modulation of FPVS response at both 6HZ and 6.2Hz by attention. However, the absence of attentional modulation when the vibrotactile stimuli is congruent with the visual stimuli might be related to the fact that the attentional gain is not so robust at these frequencies and thus more sensible to interferences from different stimulations in different sensory modalities at different frequencies. We thus speculate that at frequencies in the range of 6 Hz, the vibrotactile stimulation represents a distractor and thus annihilates the effect of attention on the visual stimuli.

It should be noted that the present studies presented some downsides which could explain the absence of results. Several factors can be invoked to explain the absence of modulation of the visual response in general: preferential frequencies of multimodal cortices, multiple stimulations simultaneously, long stimulation trains etc. Those will be discussed in the general discussion. At this point, it is worth discussing what visual areas are stimulated by periodic modulation of the luminance of LED lights and its consequence for visuo-tactile interactions.

FPVS have been used in multiple settings, using simple visual stimuli (i.e. flickering checkerboards, LED lights, Gaussian noise) relying on periodic modulation of low-level visual features such as luminance and contrast. Sources of the visual processing related to these types of stimulation appear to depend on the frequency of stimulation, with a broad spectrum of frequency finding resonance in primary visual areas, and stimulation type. Herrmann (2001) tested frequencies up to 100Hz with flickering lights. They showed that FPVS response to flickering lights

can be obtained up to 90HZ in the frequency spectrum, but only discuss the topography elicited by stimulation at 10 Hz, which is characterized by a positivity over the occipital cortex with a negativity over the frontal cortex. The authors therefore suggested that the topography obtained at 10 Hz most likely originates from occipital generators, as has also been demonstrated using fMRI (Hillyard & Anllo-Vento, 1998). Müller, Teder, & Hillyard (1997) investigated the sources of FPVS using MEG and reported that the source of the flickering light FPVS at 6Hz to be posterior occipital. It thus appears that sources of FPVS responses for simple visual stimuli can vary but that the primary visual cortices are always involved.

By using complex images as visual stimuli, it is nonetheless possible to also elicit activity related to higher-order visual processes or even cognitive processes using FPVS. While the processing of low-level visual properties are probably originates in the primary visual cortices, these regions are not specifically involved in extracting semantic content from an image or processing this semantic content. In line with this reasoning, Alonso-Prieto, Belle, Liu-Shuang, Norcia, & Rossion (2013) compared the FPVS response responses testing two conditions at a range of different frequencies. They aimed at looking at the involvement of higher-order processes in FPVS responses. They built a FPVS sequence with identical faces (no higher-order visual activity as the identity remained the same throughout the entire sequence), or faces with different identities (higher-order visual activity related to the processing of the different identities). They observed that while the response to both conditions is different for stimulation frequencies up to 10Hz, the response became identical for higher frequencies. This indicates that there can be a component related to higher-order visual processing in the FPVS response at frequencies lower than 10 Hz.

However, demonstrating a visuo-tactile interaction on such FPVS signal might still be challenging, as the part of the signal related to the periodical modulations of low-level visual features can be expected to be more robust and larger than the higher-order visual processing (and the potential effect consequently minimal). As such, uncovering an effect of spatially congruent vibrotactile stimulation on visual processing might still be challenging, as the effect might be concealed by the response elicited by low-level visual processing. However, as congruent visuo-tactile stimulation are believed to also activate higher-order multisensory areas, it might be more straightforward to demonstrate visuo-tactile interaction on higher-order visual processing selectivity. As such, this is to be considered in future studies attempting to look at visuo-tactile interaction using FPVS.

In conclusion, we showed that studying multisensory integration or crossmodal interaction through FPS might not be as straightforward as expected. Indeed, throughout all three studies, we were not able to show a modulation of the visual processing by a spatially congruent vibrotactile stimulation. To further study multisensory integration with FPS, it would be of interest to selectively target higher-order visual processing, specifically brain areas which are involved in multisensory integration, to study if activity originating from these areas can be modulating by spatial congruent vibrotactile stimulation. We therefore propose to study the modulation of visual processing using a variation of FPVS. This paradigm has the ability to separate low-level visual processing from higher-order processing. In the next study, we propose to study the ability of this paradigm to separate higher-order tool-selective activity from activity which is common to all categories, including the low-level visual features. Visuo-tactile interactions using sustained stimulations might be of interest in feedback mechanisms during actions.

As such, using tools to encourage these interactions appears to be an ecological decision.

PART 2: DOES SPATIALLY CONGRUENT VIBROTACTILE STIMULATION MODULATE TOOL-SELECTIVE PROCESSING?

Study 2: Fast Periodic Visual Stimulation to study tool-selective processing in the human brain

Based on De Keyser R., Mouraux A., Quek G. L., Torta D. M., Legrain V. (2018). Fast periodic visual stimulation to study tool-selective processing in the human brain. Experimental Brain research, 1-13. DOI: 10.1007/s00221-018-5331-2.

Abstract

Tools are often considered as a specific object category that associates perceptual and motor properties. Their neural processing has been studied extensively by comparing the cortical activity elicited by the separate presentation of tool and non-tool objects, assuming that observed differences are solely due to activity selective for processing tools. Here, using a fast periodic visual stimulation (FPVS) paradigm, we isolated EEG activity selectively related to the processing of tool objects embedded in a stream of non-tool objects. Participants saw a continuous sequence of tool and non-tool images at a 3.7 Hz presentation rate, arranged as a repeating pattern of four non-tool images followed by one tool image. We expected the stimulation to generate an EEG response at the frequency of image presentation (3.7 Hz), reflecting activity common to the processing of tool and non-tool images. Most importantly, if tool and non-tool images evoked different neural responses, we expected this differential activity to be reflected at the frequency of tool images (3.7Hz/5=0.74 Hz). The periodic insertion of tool-stimuli within a stream of non-tool stimuli elicited a significant EEG response at the tool-selective frequency and its harmonics. We conclude that FPVS is a promising technique to selectively measure tool-related activity.

1. Introduction

Recognizing and identifying objects according to specific properties and meaning is a task the brain achieves easily and automatically (VanRullen & Thorpe, 2001). Studies about the visual system has identified cortical areas responsible for the categorization of objects and their semantic processing, mostly in the ventral part of the temporal cortex (Ishai, Ungerleider, Martin, & Haxby, 2000; Peelen & Caramazza, 2012). In addition to this ventral stream, a dorsal stream was described, which is involved in the spatial processing of visual information (Goodale & Milner, 1992; Ungerleider & Haxby, 1994). Tools are physical items that can be grasped and manipulated for the purpose of action, and therefore represent a particular object category as they associate perceptual and motor properties. Accordingly, the image of a tool can activate both the ventral and dorsal streams (Chao & Martin, 2000; Creem-Regehr & Lee, 2005). Viewing tools activates the ventral stream for the identification of their specific meaningful category but also the parietal and premotor areas of the dorsal stream, related to the preparation of actions towards tools (Rizzolatti, Fogassi, & Gallese, 2002). Interestingly, premotor activation does not require the intention to handle – passive viewing of a tool is sufficient to generate activity in the left dorsal premotor cortex (Chao & Martin, 2000; Grafton et al., 1997), as well as the parietal cortex (Johnson-Frey, 2004). As such, studying how tools are processed allows investigating how perceptual inputs are translated into actions.

How the visual system processes tool objects has been studied extensively using non-invasive neuroimaging and neurophysiological techniques, such as fMRI and EEG. An inherent drawback of previous approaches is that they rely on assessing the differences between the brain responses elicited by the separate presentation of tool vs. non-tool objects by subtraction, assuming that the

differences capture the cortical activity selective for the processing of tool objects. However, differences in the neural activity elicited by the comparison between of tools vs. non-tool objects could, at least in part, be due to differences in low-level features distinguishing at least a subset of tool objects vs. non-tool control objects. For example, Grafton and colleagues (1997) showed premotor activation during simple observation of tools by comparing tools and simple fractals. It is extremely delicate to determine whether the observed differential activity is due to the observation of tools, or influenced by differences in low-level visual features. Separating neural correlates of low-level visual processing from higher-order processing is consequently a major obstacle in visual research (Koenig-Robert & VanRullen, 2013). Additionally, the tool-selective activity could also be explained by differences in the brain processes consequential to the identification of tool vs. non-tool objects.

Here we address the issue of untangling low-level visual processes from higher-order processing by directly indexing the differential response to tool objects embedded in a stream of non-tool objects, using a novel paradigm called Fast Periodic Visual Stimulation (FPVS) (Rossion, Torfs, Jacques, & Liu-Shuang, 2015). It relies on the fact that periodic sensory stimulation delivered at a given frequency can elicit a periodic EEG response at the frequency of stimulation and its harmonics, (also called steady-state evoked potentials) (Regan, 1977). The FPVS approach presents several advantages over the standard recording of event-related potentials time-locked to the onset of a transient sensory stimulus. First, FPVS often yields robust EEG responses with a high signal to noise ratio (Norcia et al., 2015; Vialatte et al., 2010). Second, it has the advantage of isolating and concentrating the response at well-defined frequencies. Third, responses can be obtained by periodically modulating low-level features such as visual luminance or contrast, but also by periodically modulating high-level visual features such as facial

identity, facial expression, or object category (Dzhelyova, Jacques, & Rossion, 2017; Guillaume, Mejias, Rossion, Dzhelyova, & Schiltz, 2018; Lochy et al., 2018; Lochy, Van Belle, & Rossion, 2015; Quek & Rossion, 2017; Xu, Liu-Shuang, Rossion, & Tanaka, 2017). Finally, it was recently shown that a variation of the FPVS can be used to disentangle cortical activity related to high-order visual processes from cortical activity related to the processing of low-level visual features, by concentrating both responses at distinct frequencies. In this specific paradigm, visual stimuli belonging to two different categories are presented as a periodic pattern of n base stimuli belonging to a first category, followed by one contrasting stimulus belonging to a second category. With this approach, cortical activity triggered by both types of stimuli can be expected to generate a periodic EEG signal at the base frequency of visual stimulus presentation, whereas cortical activity differentially triggered by the two categories of stimuli can be expected to generate a periodic EEG signal at the frequency of the critical stimulus. To ensure that the activity at the frequency of critical stimulation is not due to systematic unaccounted differences in low-level visual features between the two different stimulus categories, the exact same sequence can be shown, replacing each image by its phase-scrambled version. Phase-scrambling is a frequently used methodology to create control images in visual research as it removes the semantic content of image while preserving low-level visual features such as spatial frequency, contrast and luminance. Using this FPVS paradigm, Liu-Shuang et al., (2014) showed that the periodic presentation of face and non-face stimuli elicits activity at the base frequency originating primarily from low-level visual areas, but also activity at the frequency isolating face-selective cortical processes (see also Retter and Rossion 2015; Quek and Rossion 2017). Using the same methods, similar studies have succeeded to dissociate EEG activities specific to words from

those elicited by presentation of non-word strings of characters (Lochy et al., 2015; Lochy, Van Reybroeck, & Rossion, 2016).

As such, we hypothesised that if tools and non-tool objects are processed differently in the human brain, the presentation of tool images embedded in a periodic stream of non-tool images should lead to a periodic EEG signal at the frequency of the tool stimulus, resulting in a direct measure of the cortical activity selective to the processing of tool images.

Importantly, we hypothesised that if this activity results from high-order visual processes, this response would be absent in the EEG responses elicited by phase-scrambled versions of the same tool and non-tool images.

2. Material & methods

2.1 Participants

Eleven right-handed healthy participants (6 women and 5 men, aged 22 ± 2 years, range = 20-26 years) took part in the experiment. These numbers have been shown to be sufficient to show significant responses using similar FPVS paradigm (Liu-Shuang et al., 2014; Lochy et al., 2015; Rossion et al., 2015). Inclusion criteria were the absence of neurological or psychiatric disease, no regular drug use, normal or corrected-to normal vision, and having slept at least 6 hours during the previous night. Approval for the experiment was obtained from the local ethical committee in agreement with the Declaration of Helsinki. All participants signed an informed consent form and received financial compensation for their participation.

2.2 Stimuli

We collected 285 highly variable natural photographs accessible on the internet: 98 photographs of manually handled tools (hammer, saw, pencil, cutlery

etc.) and 187 photographs of non-tool images (animals, buildings, plants, furniture etc.). Tools were defined as man-made objects that can be manipulated to perform a motor action, e.g. a pen to write, a cup to drink, etc. The images were transformed to grayscale and cropped/scaled to 256x256 pixels. Importantly, the images varied strongly in terms of background, viewpoint, size, luminance, spatial frequency and contrast, and were not equalized in terms of spatial frequency, luminance or contrast, such that the full natural image set varied widely in terms of low level properties, as well as higher-level properties such as viewpoint, background, lighting, shape, texture (Rossion et al., 2015). Therefore, 'toolness' was the only consistent difference between tool and non-tool images.

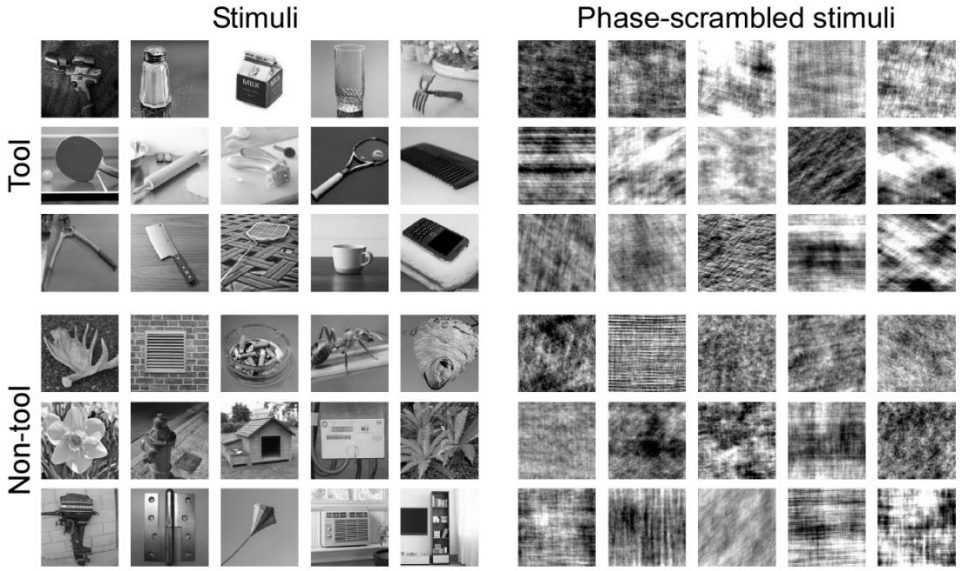

Figure 1: Examples of stimuli (tools and non-tools), and their phase-scrambled counterparts.

We further created a second separate image set by phase scrambling each tool and non-tool image. Phase scrambling was achieved by replacing the phases of each image by random coefficients. Both sets are available upon request (Figure 1).

2.3 Procedure

Participants were seated 58.5 cm in front of a 32” Display ++ IPS LCD Monitor (1920x1080 resolution, Cambridge research System, Rochester, UK, 100 Hz refresh rate) with their head stabilized using a chinrest. The images appeared on a light grey background, subtending a visual angle of ~14.7°. Each trial consisted of a central blue fixation cross (30x30 pixels, visual angle of 1.7°') that appeared for 1-2 sec before the start of the periodic visual stimulation, and remained visible throughout the entire trial. Participants were asked to fixate the cross throughout the trial, but were allowed to blink. Throughout the trial, images appeared at a strict periodic base rate of 3.7 Hz, that we achieved by modulating the contrast of each image sinusoidally from 0%-100%-0%, creating a continuous flow of information. This periodic stimulation had a fade-in period of 5 s during which the maximal contrast was progressively increased (Figure 2), followed by 60 seconds of stimulation during which maximum contrast remained constant at 100%, and then a 5 s of fade-out of contrast. Fade-in and fade-out periods were included to avoid an abrupt onset and offset of stimulation, as well as to reduce eye movements and blinks triggered by stimulation onset (Alonso-Prieto et al., 2013). A trial consisted therefore of a sequence of 259 images. At the end of the periodic stimulation, the fixation cross remained on the screen for 1-2 seconds. The onset of each trial was self-paced.

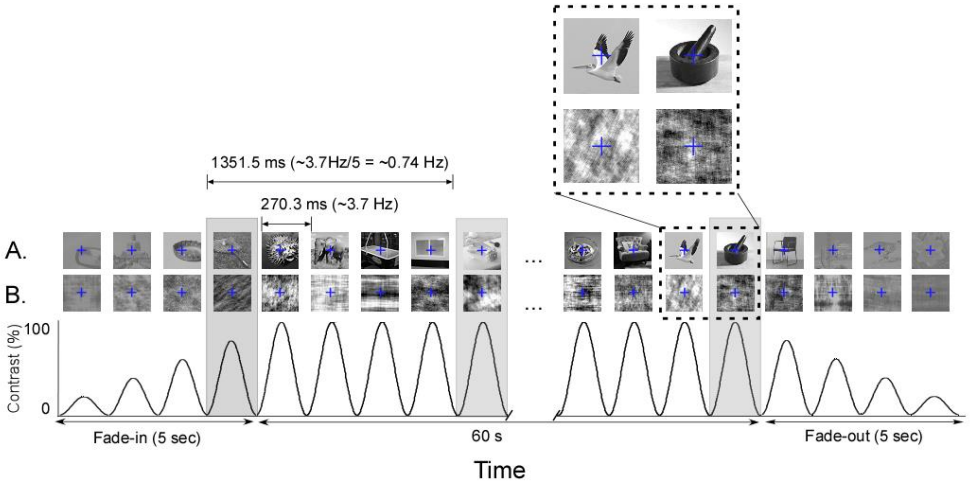

Figure 2: Fast periodic visual stimulation using non-tool and tool images (A) and phase-scrambled non-tool and tool images (B). The periodic visual stimulation was achieved by a sinusoidal modulation of contrast at a strict periodic rate of 3.7 Hz. The stimulation sequence consisted of a pattern of four non-tool images followed by a tool image, resulting in the periodic presentation of a tool image at 0.74 Hz.

Images were presented as a pattern consisting of four non-tool images (or phase-scrambled non-tool images) followed by one tool image (or phase-scrambled tool image). Therefore, tool images (or phase-scrambled tool images) were presented at a frequency of 3.7Hz/5=0.74 Hz. As such, there were two frequencies of interest: a base response at 3.7Hz which corresponds to the presentation rate of both tool and non-tool images, and the embedded tool-selective frequency at 0.74Hz, which corresponds to the presentation rate of the tool images. Three conditions were tested. In the first condition (TARGET IMAGE), participants were presented a target image at the beginning of each trial, and were requested to count the number of occurrences of this target image. Importantly, the target image, which changed every trial, was always a non-tool image. In the

second condition (FIXATION CROSS), participants were asked how many times the fixation cross changed colour during the trial. In these trials, the fixation cross changed to red during 200 ms, 2-8 times per trial. These two conditions showing tool images embedded in a stream of non-tool objects were used to examine whether the elicited responses were modulated by attending to the content of the non-tool images. Finally, in a third condition (SCRAMBLED), each non-tool and tool image was replaced by their respective phase-scrambled versions. As in the FIXATION CROSS condition, participants were asked to count the number of colour changes of the fixation cross, and provide a verbal response at the end of each trial. This third condition showing phase-scrambled tool images embedded in a stream of phase-scrambled non-tool images was used to examine whether the tool-selective response observed in the first two conditions could be due to systematic unaccounted differences in global low-level properties of tool and non-tool images (luminance, contrast, spatial frequency spectrum). Eight trials belonging to each condition were presented in a randomised fashion in four blocks of six trials. The presentation was programmed using SinStim, a custom-made Java program (Bruno Rossion, Face Categorization lab, Université catholique de Louvain, Belgium).

2.4 Measures

The EEG was recorded using 64 Ag-AgCl electrodes placed on the scalp according to the international 10/10 system (Waveguard64 cap, Cephalon A/S, Nørresundby, Denmark). Electrode impedances were kept below 15 kΩ. Signals were amplified and recorded using a sampling rate of 1000 Hz and an average reference (64-channel high-speed amplifier, Advanced Neuro Technology, Enschede, the Netherlands). The EEG data was analysed offline using Letswave 6 (https://www.letswave.org/) running in MATLAB 2014a. The data was first

segmented from 2 s before the fade-in period (with a duration of 5 s) to 2 s after fade-out period (with a duration of 5 s), resulting in 24 epochs of a total length of 74 s (2 s + 5 s fade-in + 60 s stimulation + 5 s fade-out + 2 s). We then applied a Butterworth 50 Hz notch filter followed by a bandpass-filter (0.1-100 Hz).

Frequency-domain analysis. All epochs were segmented to the stimulation length during which the maximum contrast was 100%, i.e. from +5 to +65 s relative to fade-in onset. Eye blink and movement artefacts were removed using an independent component analysis (ICA, Hyvärinen and Oja 2000; Colon et al. 2012, 2015; Nozaradan et al. 2012b). Separate 60-s average waveforms were computed for each condition and participant. Subsequently, we performed a Fast Fourier transform (FFT) to transform the data into the frequency domain. Then, the contribution of background noise was removed by subtracting, at each bin of the frequency spectra, the average amplitude measured at neighbouring frequency bins (four frequency bins ranging from −0.03 to −0.08 Hz and four frequency bins ranging from +0.03 to +0.08 Hz relative to each frequency bin (Colon et al., 2015; Mouraux, Iannetti, et al., 2011; Nozaradan, Peretz, & Mouraux, 2012a)). The validity of this subtraction procedure relies on the assumption that, in the absence of a strong periodic signal, the signal amplitude at any given frequency bin should be similar to the signal amplitude of the mean of the surrounding frequency bins (Mouraux, Iannetti, et al., 2011; Retter & Rossion, 2015). Two measures were extracted from the EEG frequency spectra. The first was a measure of the amplitude of the periodic EEG signal at the base frequency (i.e., 3.7 Hz and its harmonics). The second was a measure of the amplitude of the periodic EEG signal at the tool-selective frequency (i.e., 0.74 Hz and its harmonics). To select the statistically significant harmonics to include in the estimation of response amplitude, we expressed the amplitude at each harmonic frequency as a z-score relative to the mean and standard deviation of the amplitude of the 20

neighbouring bins of the FFT, ranging from -0.03 to - 0.18 Hz and from +0.03 to +0.18 Hz. A z-score greater than 1.64 (i.e., $p<0.05$, one-tailed) was the criterion to consider the amplitude at that frequency significantly greater than the amplitude at neighbouring frequency bins (Dzhelyova et al., 2017; Lochy et al., 2015; Quek & Rossion, 2017; Rossion et al., 2015). This selection of harmonics was performed on the grand average frequency spectrum, averaged over all participants, conditions and scalp electrodes.

Activity originating from ventral and dorsal streams may be expected to contribute predominantly to the EEG signals recorded from occipito-temporal and central electrodes, respectively. Therefore, separate single-subject estimates of base and tool-selective EEG responses were computed by averaging the responses obtained at central (FC1, FCz, FC2, C1, Cz, C2, CP1, CPz and CP2) and occipito-temporal (P7, P5, P6, P8, PO3, PO4, PO5, PO6, PO7, PO8, POz, Oz, O1 and O2) channels. We then summed the noise-subtracted amplitudes at the base frequency and its significant harmonics, and at the tool-selective frequency and its significant harmonics (excluding tool-selective frequencies overlapping with the base frequency), to obtain a single measure of the base and tool-selective response for each participant and condition (Quek & Rossion, 2017; Retter & Rossion, 2015).

Time-domain analysis. To assess the time course of the periodic EEG responses, the filtered signals were segmented in 1350 ms epochs, corresponding to a complete stimulation sequence (four non-tool stimuli and one tool stimulus), yielding a total of 416 epochs per participant and condition. Such as for the frequency domain analysis, eye blink and movement artefacts were removed using an independent component analysis. Epochs containing deflections larger than 80 μV were removed. Separate averages were then computed for each participant and condition. To separate the base response from the tool-selective response, we

filtered out the respective unwanted frequency components (Dzhelyova et al., 2017; Quek & Rossion, 2017). Specifically, to visualize the base response in the time domain, we filtered out the tool-selective response using a series of FFT multi-notch filter at the frequencies which were significant in the frequency domain analysis, i.e. 0.74, 1.48, 2.22 and 2.96 Hz. Similarly, to obtain the tool-selective response, we removed the base response using a series of notch filters at the frequencies which were significant in the frequency domain analysis, i.e. 3.7, 7.4, 11.1, 14.8, 18.5, 22.2 and 25.93 Hz.

2.5 Analyses

The sum of the amplitudes of the base and tool-selective EEG responses obtained in the different conditions were compared using a repeated-measures analysis of variance (ANOVA) with three factors: electrode (OCCIPITO-TEMPORAL, CENTRAL), condition (TARGET IMAGE, FIXATION CROSS, SCRAMBLED) and response type (BASE, TOOL-SELECTIVE). Effect sizes were measured using partial Eta squared. Greenhouse-Geisser corrections of degrees of freedom were performed when necessary. Significant interactions were followed up by contrast analyses using partial ANOVA or paired-sample Student t-tests. Significance level was set at $p \leq .05$ and corrected using the Bonferroni correction procedure.

3. Results

Frequency-domain results. In all experimental conditions, a base response at 3.7 Hz and its harmonics was observed, clearly standing out from neighbouring bins (Figure 3). This increase was significant for the first seven harmonics (Figure 4). The topographical distribution of this base response was maximal over occipital electrodes, and symmetrically distributed over the two hemispheres (see Figure 5, panel A). In both the TARGET IMAGE and the FIXATION

CROSS conditions, the EEG frequency spectra revealed clear peaks at the frequency of tool presentation and its harmonics (Figure 3). Significant peaks were present up to the fourth harmonic in the TARGET IMAGE condition, and up to the ninth harmonic in the FIXATION CROSS condition (Figure 4). It should be noted that a different amount of harmonics does not necessarily reflect a greater response in the FIXATION CROSS condition, but a difference in the time course of the periodic signal (Zhou et al., 2016). No tool-selective responses reached significance in the SCRAMBLED condition. As compared to the occipital scalp topography of the base response, the scalp topography of tool-selective EEG response observed in the TARGET IMAGE and the FIXATION CROSS conditions extended towards occipito-temporal and central electrodes (Figure 5, Panel B).

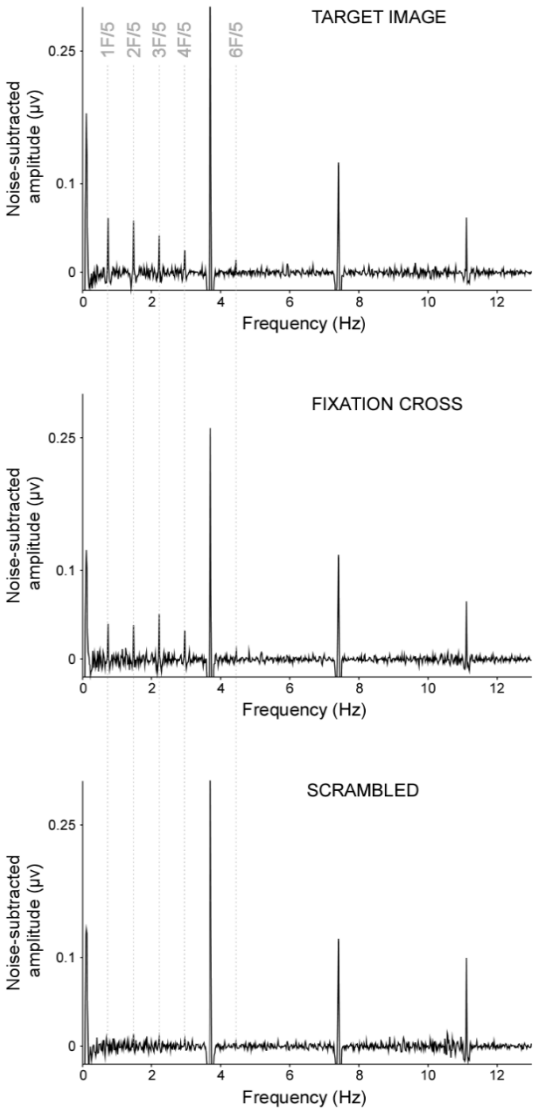

Figure 3: Group-level average EEG frequency spectrum for each condition, averaged across all scalp channels. Tool-selective responses are indicated by the dashed vertical lines. Clear tool-selective responses are evident in the TARGET IMAGE and the FIXATION CROSS conditions. In contrast, no tool-selective response is observed in the SCRAMBLED condition. The base response is present at 3.7 Hz and harmonics in all three conditions.

The repeated-measures ANOVA on the sum of the noise-subtracted amplitudes showed a main effect of response type ($F(1,10)=110.9$, $p<0.001$, $\eta^2_p=0.67$), a main effect of electrodes ($F(1,10)=131.8$, $p<0.001$, $\eta^2_p=0.47$), a main effect of condition ($F(2,20)=4.1$, $p=0.03$, $\eta^2_p=0.046$), an interaction between the factors electrodes and response type ($F(1,10)=80.6$, $p<0.001$, $\eta^2_p=0.27$), an interaction between the factors response type and condition ($F(2,20)=12.0$, p[GG]$=0.002$, $\eta^2_p=0.14$) and a triple interaction between the factors condition, response type and electrode ($F(2,20)=4.46$, $p=0.02$, $\eta^2_p=0.02$).

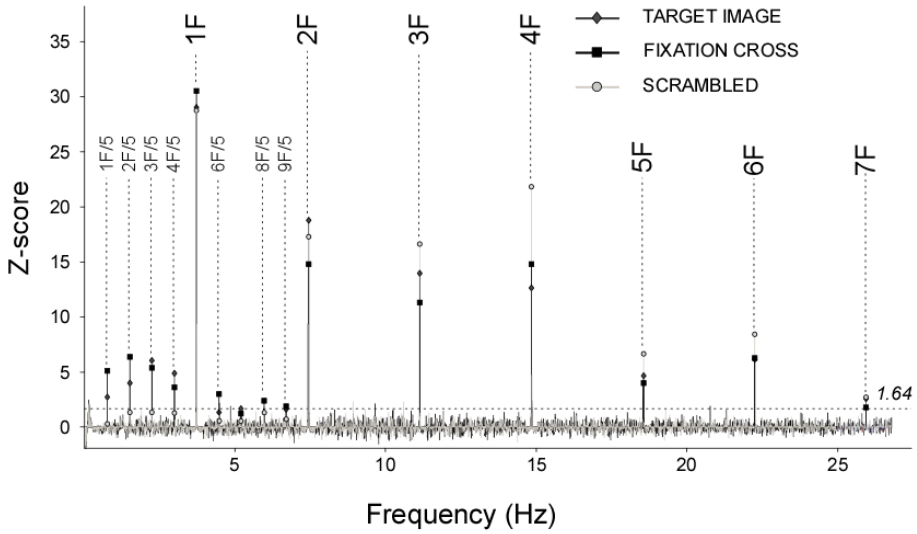

Figure 4: Frequency spectrum of the EEG activity recorded in the TOOL $_{TARGET\ IMAGE}$, TOOL $_{FIXATION\ CROSS}$ and TOOL $_{SCRAMBLED}$ conditions, expressed as z-scores. The dotted line depicts the significance criterion of z<1.64 (one-tailed).

We subsequently followed up this triple interaction by subjecting the base response and tool-selective response to separate repeated-measures ANOVA with two factors: condition (TARGET IMAGE, FIXATION CROSS, SCRAMBLED) and electrode (CENTRAL and OCCIPITO-TEMPORAL). The analysis of the base response showed a main effect of electrode ($F(1,10)=130.56$, $p>0.001$, $\eta^2_p=0.57$) with larger activity at occipito-temporal electrodes as compared to central electrodes. There

was no significant main effect of condition ($F(2,20)=3.92$, p[GG]=0.06, η^2_p =0.09), and no interaction between the two factors ($F(2,20)=1.39$, p[GG]=0.27, η^2_p =0.01) (Figure 6, Panel A). Conversely, the analysis of the tool-selective response showed a main effect of condition ($F(2, 20)$ =38.46, p<0.001, η^2_p =0.57), a main effect of electrode ($F(1, 10)$ = 25.5, p<0.001, η^2_p = 0.37) as well as a significant interaction between the two factors ($F(2, 20)$ =12.1, p<0.001, η^2_p =0.15). To explore the interaction, we first used paired t-test to look at the difference between electrode pools in each condition. There appeared to be a larger activity in the OCCIPITO-TEMPORAL electrodes as compared to the CENTRAL electrodes in both FIXATION CROSS condition ($t(10)$ = -4.58, p = 0.001) and the TARGET IMAGE condition ($t(10)$ = -4.79, p < 0.001), but not in the SCRAMBLED condition ($t(10)$= -1.7, p = 0.12). Post-hoc paired-sample t-tests (alpha = 0.008) showed that the signal at tool-selective frequencies was significantly greater in the TARGET IMAGE condition as compared to the SCRAMBLED condition, both at OCCIPITO-TEMPORAL ($t(10)$ = 8.2, p<0.001) and CENTRAL electrodes ($t(10)$ = 4.57, p= 0.001). The response obtained in the FIXATION CROSS condition was also significantly greater than the signal obtained in SCRAMBLED condition (CENTRAL electrodes: $t(10)$ = 6.63, p<0.001; OCCIPITO-TEMPORAL electrodes: $t(10)$ = 4.98, p<0.001). There was no significant difference in the amplitude of the tool-selective response between the TARGET IMAGE and FIXATION CROSS conditions, both at CENTRAL ($t(10)$ = -1.05, p= 0.32) and OCCIPITO-TEMPORAL electrodes ($t(10)$ = -1.61, p = 0.14) (Figure 6, Panel B).

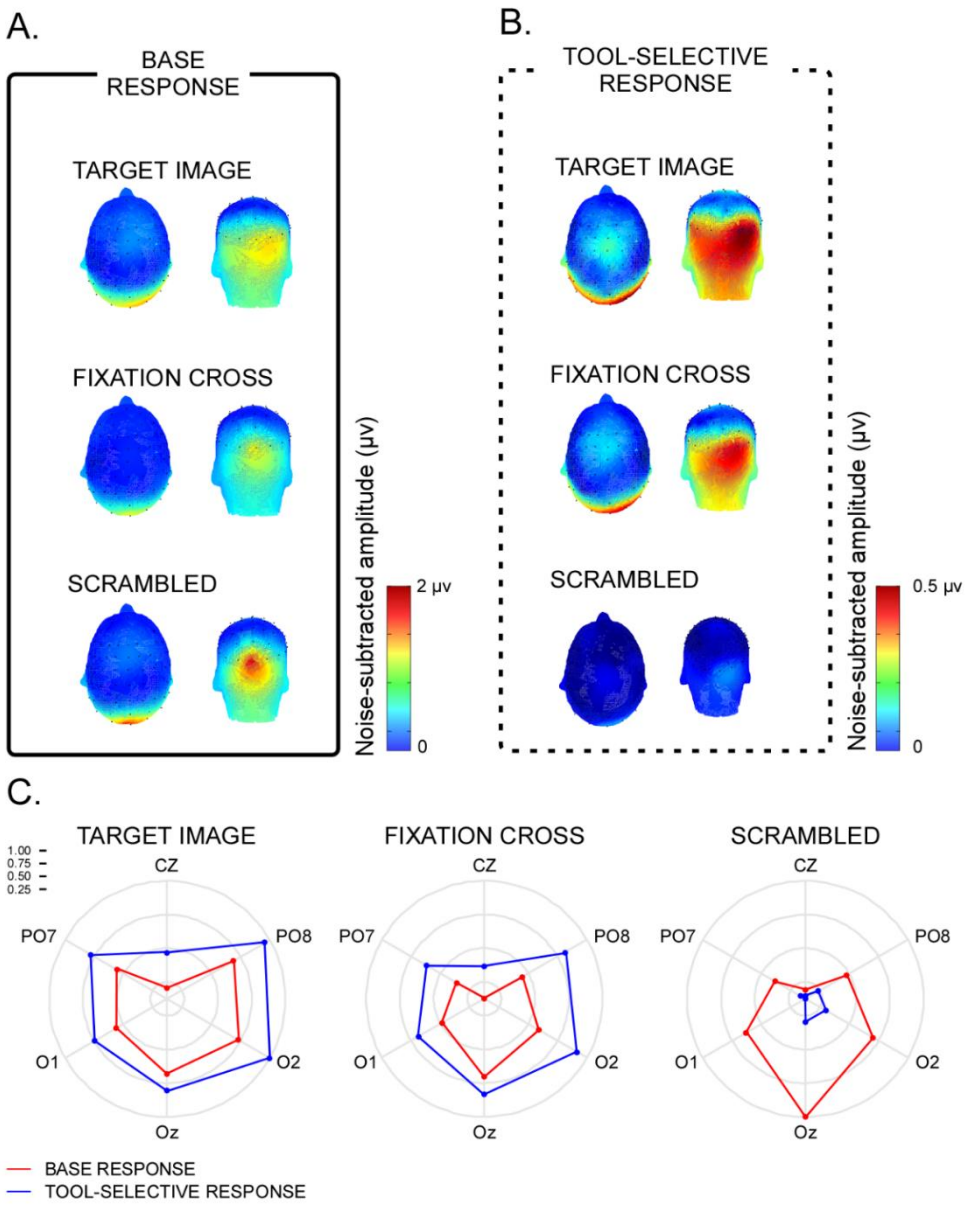

Figure 5: Scalp topography: distribution of the noise-subtracted amplitude over the scalp electrodes of the periodic base response (A) and the tool-selective response (B) obtained in the different experimental conditions. Topographies of the base response were mainly characterized by activity over occipital electrodes. In the

TARGET IMAGE and the FIXATION CROSS conditions, the tool-selective response was predominant over occipito-temporal and central electrodes. No clear activity at the tool-selective frequencies was observed in the SCRAMBLED condition. Panel C shows the base response and tool-selective response over occipital (O1, OZ, O2), occipito-temporal (PO7, PO8) and central (CZ) electrodes for each condition. The data was normalized per response type (base response and tool-selective response) across the three conditions.

Time-domain results. Analysis in the time-domain analysis revealed a pronounced differential tool-selective response in the FIXATION CROSS and the TARGET IMAGE conditions. This differential response disappeared in the SCRAMBLED condition (Figure 7. Panel B). The base response showed the activity to the stimulus repetition which is similar for all condition (Figure 7. Panel A).

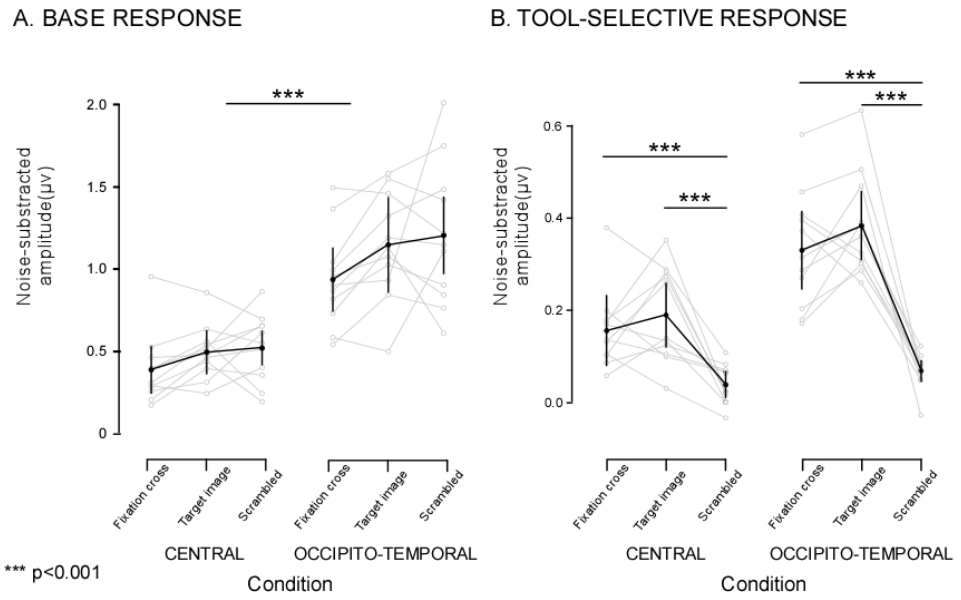

Figure 6: Single-participant and group-level average amplitude of the base response and the tool-selective response obtained in the TARGET IMAGE, FIXATION CROSS

and SCRAMBLED conditions, at occipito-temporal and central electrodes. While the base response was similar across conditions, the tool-selective response was present only in the TARGET IMAGE, FIXATION CROSS conditions. Individual values are shown as grey connecting lines and group-level average values are shown as black connecting lines. The vertical bars represent the 95% confidence interval.

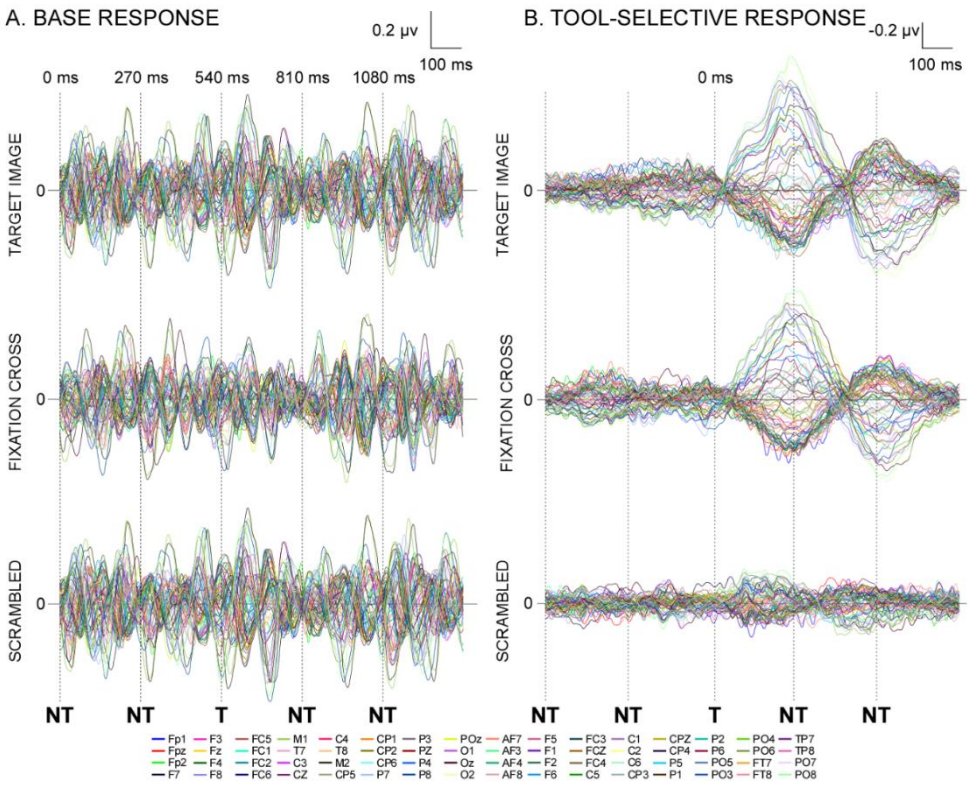

Figure 7: Time-domain EEG waveforms averaged across repetition patterns in the TARGET IMAGE, FIXATION CROSS and SCRAMBLED conditions. The waveforms obtained at each scalp channel are shown as different colours. Panel A. indicates the base response to one full sequence (5 cycles), isolated by filtering out the tool-selective response using notch filters at 0.74 Hz and harmonics. Panel B. shows the tool-selective response to one full sequence (5 cycles), isolated by filtering out the

base response using notch filters at 3.7Hz and harmonics. NT indicates the onsets of the four non-tool images. T indicates the onsets of the contrasting tool image.

4. Discussion

In this experiment, we used Fast Periodic Visual Stimulation (FPVS) to tag EEG activity related to the differential processing of tool images compared to non-tool images. Our results showed that the periodic insertion of a tool image within a periodic stream of non-tool images separates EEG activity related to the high-order processing of visual features distinguishing tool vs. non-tool objects (Figure 7, panel B) from the EEG activity related to the processing of low-level visual features common to the two object categories (Figure 7, panel A). Specifically, we showed that an image of a tool inserted as a periodic contrasting stimulus within a stream of non-tool objects elicits a significant EEG response at the frequency of tool presentation, maximal over occipito-temporal regions, and extending towards central regions. Importantly, this tool-selective response was present regardless of whether participants attended to the content of the images (detecting the occurrence of a pre-specified target image) or performed a completely orthogonal task (detecting the occurrence of a colour change in the fixation cross). In contrast, little or no activity at the frequency of the contrasting stimulus was observed when participants viewed phase-scrambled versions of the images (SCRAMBLED condition) (Figure 6, panel B.)

The usefulness of phase-scrambling comes from its ability to destroy the semantic content of an image whilst preserving most of the global low-level properties of the images (Ales, Farzin, Rossion, & Norcia, 2012). Hence, had the tool-selective response entirely been driven by systematic differences in low-level features distinguishing tools and non-tools that are preserved in the phase-scrambled condition, a tool-selective response of the same magnitude would have

been expected in the SCRAMBLED condition. However, one must take into consideration the fact that phase-scrambling does not flawlessly preserve all low-level visual features of an image (Stojanoski & Cusack, 2014; Thomson, 1999). Hence, the absence of a tool-selective response in the SCRAMBLED condition does not allow us to definitely exclude that some unaccounted low-level features distinguishing tool from non-tool images contributed to the contrast responses observed in the TARGET IMAGE and FIXATION CROSS conditions. However, this seems highly unlikely, considering that both tool and non-tool images varied strongly across exemplars in terms of both their low-level features and their high-level features (luminance, spatial frequency and contrast, background, viewpoint, shape, lightning, etc.).

It is well known that stimulating the visual system at a fixed periodic rate with simple visual stimuli, such as flickering lights, elicits a periodic EEG signal which is maximal at occipital electrodes, and is thought to mainly reflect low-level stages of visual processing occurring in primary visual areas (Di Russo et al., 2007; Müller et al., 1997; Norcia et al., 2015; Pastor, Artieda, Arbizu, Valencia, & Masdeu, 2003). Whereas the scalp topography of the base response was maximal over occipital electrodes, and thus likely reflects low-level stages of visual processing common to the processing of non-tool and tool stimuli, the scalp topography of the tool-selective response was maximal over occipito-temporal and central electrodes. Although one should be cautious to infer conclusions based on topographies (Urbach & Kutas, 2002, 2006), the different topographical distribution of the tool-selective response as compared to the base response suggests that the tool-selective response originated, at least in part, from cortical regions distinct from those activated by the modulation of low-level visual features and captured in the base response. Most importantly, the scalp topography of the tool-selective response is compatible with activation of the ventral and dorsal

visual streams. The ventral stream is located in the temporal lobe, including the fusiform gyrus which is thought to be involved in the visual processing of graspable objects (Chao, Haxby, & Martin, 1999; Creem-Regehr & Lee, 2005; Martin, Wiggs, Ungerleider, & Haxby, 1996). The bilateral occipito-temporal activity that we observed could thus be related to ventral stream processes involved in object semantics (Fabre-Thorpe, 2011). Importantly, the finding that contrasting tool vs. non-tool objects elicits an occipito-temporal tool-selective response suggests that tool and non-tool objects differentially activate this occipito-temporal network. Supporting this hypothesis, Martin et al., (1996a), Chao et al., (1999), Ishai et al., (2000) and more recently Hutchison et al. (2014) showed that different categories of objects activate the semantic network differently within the temporal lobe. The scalp topography of the tool-selective response also extended towards central electrodes. This might be related to a recruitment of the dorsal visual stream and premotor during passive viewing of tool objects (Grafton et al., 1997). Interestingly, this extension towards central electrodes was not observed for category-selective EEG responses elicited by other kinds of visual categories using a similar FPVS paradigm. Indeed, word-selective activity has been shown to be most pronounced over left occipital electrodes (Lochy et al., 2015, 2016), while face-selective activity has been shown to be maximal over right temporo-occipital electrodes (Dzhelyova et al., 2017; Liu-Shuang et al., 2014; Quek & Rossion, 2017; Xu et al., 2017). Similarly, FPVS studies have shown that numerical processing elicits selective activity maximal over medial occipital regions (Guillaume et al., 2018; Park, 2016). As such, higher-order category-specific activity can be measured at occipital electrodes. In other words, the topography of category-selective EEG responses obtained using the FPVS paradigm appear to depend on the eliciting object category. This suggests that this activity is not merely related to the detection of a contrasting event whatever its category.

It has also been demonstrated that stimulating the visual system at different frequencies results in different scalp topographies for the same stimulus type (Lithari, Sánchez-Garciá, Ruhnau, & Weisz, 2016). Therefore, testing the responses elicited by contrasting tool and non-tool objects presented at different frequencies could be of interest in future studies.

The tool-selective response could reflect selective cortical activity related to the differential processing of tool objects. One other option is that the tool-selective response does not reflect the activity related to the tool stimulus, but rather the response to a base stimulus after the presentation of the tool stimulus. However, this option appears unlikely. Indeed, we show that the deviation in signal corresponding to the tool-selective response started shortly after the occurrence of the tool stimulus and, most importantly, before the occurrence of the next base stimulus (Figure 7, panel B.). This indicates that it was predominantly triggered by the tool image, rather than by the occurrence of a non-tool image following the presentation of a tool image.

Additionally, the tool-selective response could also reflect the ability of the brain to recognize regular patterns of inputs across repetitions. Fiser and Aslin (2002) showed that fixed sequences of shapes presented repeatedly during a familiarization session can be distinguished from novel sequences of familiar shapes, indicating the ability to learn temporal sequences implicitly or explicitly. Thus, it is possible that participants in the present experiment did learn the periodic image pattern (four non-tool images followed by one tool image) over the course of the experiment. While we did not test the possible contribution of such sequence learning to the observed tool-selective response, there is good reason to think that the tool-selective response we observe here does not reflect an expectation response. Recently, Quek and Rossion (2017) showed that face

stimuli embedded in a stream of non-face images elicit a similar face-selective response irrespective of whether the faces appeared at temporally predictable or unpredictable intervals (i.e., periodic vs. nonperiodic). Regardless, it should be noted that even if participants did indeed tap into the tool image periodicity, this interpretation remains compatible with the hypothesis that the tool-selective activity is due to automatic image categorization. Indeed, implicit learning of the temporal regularity requires discerning the difference in image content defining tool vs. non-tool images.

Overall, it appears that the FPVS can circumvent the need of subtracting signals from different contrasting conditions by concentrating activity related to high-level and low-level visual processing at distinct frequencies. Additionally, the choice of control or base stimuli can be as diverse as needed, enabling researchers to introduce a lot of variety in the used stimuli, reducing the possibility of a bias due to low level visual features and the attention towards the target stimuli. Moreover, this method eliminates the need to standardise images on features such as spatial frequency, luminance and contrast, as its strength lies in the variation in these features. As such, this approach allows for more ecological visual stimulation.

In conclusion, we show that it is possible to index EEG activity related to the differential processing of tool vs. non-tool objects using FPVS. Furthermore, we show that the tool-selective response related to processing tool objects cannot be explained by attention, relevance or differences in low-level visual features. The proposed approach offers new possibilities to study, in humans, distinct processing by directly tagging and measuring the differential activity.

Study 3: Can tool selective activity be modulated by hand position and vibrotactile stimulation?

Abstract

Stimulations from the tactile and visual modality can interact when presented at the same spatial location. However, it has not yet been possible to demonstrate this interaction using a sustained train of stimulation. We explore the influence of tactile stimuli on visual processing using a variation of Fast Periodic Visual Stimulation. This variation allows to study the effect of a spatially congruent vibrotactile stimulation on low-level visual processing as well as on higher-order visual processing related to tool-selective processing. To test for an interaction between touch and vision through their spatial congruence, the position of the stimulated limb was modulated and could either be close to the visual stimulation, or close to the body. Additionally, to test whether modulation can truly be attributed to the presence spatially congruent tactile stimulation, tactile stimulation was either present or absent. Participants were asked to report the number of times they perceived a color change in the fixation cross. Surprisingly, the results show no effect of vibrotactile stimulation. As such, there is no evidence for a visuo-tactile interaction either on the base response reflecting more general visual processing, or on the tool-selective processing. Additionally, the position of the hand did influence the tool-selective processing but not in the expected direction. Indeed, tool-selective processing appears to be inhibited when the hand is positioned next to the visual stimulation, as such obstructing visual processing instead of enhancing it. These results demonstrate that the processes underlying visuo-tactile interactions as evidenced by reaction times and brain potentials might be not be involved in the response obtained through sustained stimulations.

1. Introduction

In the discussion of the first study of this thesis, we pointed out that trying to show multisensory integration using the periodic modulating of luminance might be suboptimal, as the activity mainly reflects low level visual activity, while multisensory integration refers to a process where unimodal stimulations progressively converge in higher-order associative areas (Spence et al., 2009). In the second study, we established that activity related to viewing a particular complex object category such as tools can be selectively tagged using a variation of FPVS. It was also shown that this activity is unlikely to be related to the low level activity commonly elicited by all visual stimuli. In this third study, we try to integrate the results of the second study and the methodology of the first study. By using the same FPVS paradigm as in the second study, but adding vibrotactile stimuli which was either placed next to the visual stimuli or to the body, we again tried to find evidence of the spatial rule of multisensory integration. However, we can now dissociate between an effect on higher-order visual processing and more general visual processing.

The hypothesis of this experiment, which is to observe a modulation of the visual activity by spatially congruent vibrotactile stimuli, relies on two theories: multisensory integration in the peripersonal space and the affordance theory. Multisensory integration can be said to essentially be a higher-order process (Engel et al., 2012; Laurienti, Perrault, Stanford, Wallace, & Stein, 2005; Spence et al., 2009) and integrates information to form a multisensory representation of the peripersonal space (Avillac, Ben Hamed, & Duhamel, 2007; Fogassi et al., 1996; Rizzolatti et al., 2002). By targeting higher-order visual processing, it can be expected to become more likely to observe this phenomenon through the spatially

congruent vibrotactile stimulation, partly by removing signal related to low level visual processing.

The importance of the semantic content of the images relies on the theory of affordances. Affordances are defined as properties in the environment that are relevant to the person (Gibson, 1979). For example, tools can afford different actions, depending on the goal (Creem-Regehr & Lee, 2005). It also states that the most plausible action to perform on a tool object is automatically activated when seeing an image of a tool object (Tucker & Ellis, 1998). Single cell recordings underlie this theory. Neurons in the F5 area, in the ventral premotor cortex fire specific to grasping objects, according to the necessary hand grip, which may be precision, finger precision or whole-hand precision grip (Jeannerod, Arbib, Rizzolatti, & Sakata, 1995). More interestingly, about 20% to 30% of those cells also respond to the visual presentation of an object, most of the time in relation with the type of grip the neuron affords. Grasping and viewing graspable objects also activate the anterior intra-parietal sulcus (AIP), also often considered as a multisensory area (Colby & Goldberg, 1999; Murata & Ishida, 2007; Murata et al., 2016). In the AIP also, some neurons are activated by the mere sight of a graspable object and during the grasping action.

So far, we introduced why tools might have a special role in multisensory integration by affording actions. Behavioural evidence for the specificity of tools and for the affordance theory has been observed in a human participants. It has been shown that if the object is congruent to the response hand (i.e. the object configuration is suitable for a grasping movement with the response hand), this speeds up reaction times in a task which is unrelated to the variable of interest (Tucker & Ellis, 1998), draws attention (Handy et al., 2003), and increases MEP's (Buccino et al., 2009).

Importantly, visuo-tactile interactions through sustained stimulations might be important for feedback mechanisms for actions. Indeed, haptic feedback is crucial for motor control (Arbib, Bonaiuto, Jacobs, & Frey, 2009). As such, using images of tools might be particularly suited to study visuo-tactile interactions.

Consequently, we hypothesised that higher-order tool-selective processing can be influenced by spatially congruent vibrotactile stimuli. In order to demonstrate the effect of spatially congruent vibrotactile stimuli, this last experiment tested two factors. First, we tested whether vibrotactile stimulation had a modulatory effect on the cortical activity as measured by FPVS activity. Indeed, the presence or absence of the vibrotactile stimuli enables to test whether a potential influence is due to the spatially congruent vibrotactile stimulation i.e. spatially congruent vibrotactile stimulation is necessary to obtain an effect. Second, we manipulated the position of the hand on which vibrotactile stimuli were applied. It was either placed next to the visual stimuli, or at a distance, more specifically on the lap of the participant. This allowed testing for the influence of the spatial congruence per se i.e. we only expected to see an influence of the vibrotactile stimuli when it is spatially congruent with the visual stimuli. As such, we expected an interaction between the position of the hand and the vibrotactile stimuli, showing that tool-selective visual processing was influenced by the presence of vibrotactile stimuli, but only when spatially congruent. However, if a modulation of base response was observed, this would imply that multisensory integration can already be observed at low-level visual activity, as this peak is thought to mainly reflect activity related to changes in low level visual features such as luminance, contrast and spatial frequency. While this appears to be in conflict with the conclusion of the first study, it is important to keep the difference in visual stimuli between the studies in mind (LED vs. images).

2. Material & methods

2.1 Participants

Sixteen healthy participants with normal or corrected-to-normal vision participated to this study. The mean age of the participants (8 men and 8 women, all right-handed) was 24 ± 5, ranging from 20 to 42. Exclusion criteria were psychiatric or neurological disease (including chronic migraine and epilepsy), no regular drug use, normal or corrected-to normal vision, having slept 6 hours during the previous night. Approval for the experiment was obtained from the local ethical committee in agreement with the Declaration of Helsinki and was carried out in accordance with the corresponding guidelines and regulations. All participants signed an informed consent form and received financial compensation for their participation.

Figure 1: Examples of the images of tools and non-tools

2.2 Stimuli

We collected 285 images of the internet (98 images of tools (i.e. hammer, saw, pencil, cutlery...), 187 images of non-tool categories (i.e. animals, buildings,

plants, furniture...). Images of tools were such that the orientation was congruent with the right hand (Figure 1). Images were cut/resized to 256x256 pixels. Importantly, the images varied strongly in terms of background, viewpoint, size, luminance, spatial frequency and contrast, and were not equalized in terms of spatial frequency, luminance or contrast, such that the full natural image set varied widely in terms of low-level properties, as well as viewpoint, background, lighting, etc. In this way, we ensured the tool and non-tool image sets did not differ systematically in their low-level properties (Rossion et al., 2015). Images were presented at a rate of 3.75 HZ (267 ms per image), modulating contrast in a sinusoidal way using SinStim, a program based in Java. Visual stimulation consisted of a strict periodic stream of images of non-tool categories (at 3.75 Hz), with an image of a tool inserted every 5 images (corresponding to 0.75 Hz) (Figure 1). This stimulation protocol is known to elicit a peak in the frequency domain at the base frequency (3.75Hz) as well as at the tool-selective frequency (0.75Hz). The participant was seated at a table made of a horizontal semi-reflective mirror (36 cm wide x 69.5 cm long). The visual stimulation was projected on this semi-transparent mirror by a horizontal 16-to-9 52-inches monitor (60 Hz refresh rate) placed above the semi-reflective mirror, giving rise to the illusion that the images were projected on the hand, placed in the centre on the platform below the semi-reflective mirror (The KINARM robotic system, BKin Technologies, Kingston, ON, Canada) (Figure 2). A blue fixation cross (30x30 pixels, visual angle of 1.7°') was placed in the centre of the images for the entire duration of the trial to maintain the gaze of the participants stable.

Figure 2: The Kinarm Robotic system. This system consists of a screen (top layer), a semi-reflective mirror (middle layer) and a table underneath the semi-reflective mirror (bottom layer). The screen is consequently reflected on the mirror, which is semi-reflective and gives the impression that the image is placed below the mirror, next to the hand.

Vibrotactile stimulation consisted of a sinusoidal carrier wave at 250Hz modulated at 4HZ delivered through a recoil-type tactile transducer (Haptuator, Tactile Labs Inc., Canada) held between the thumb and the index finger of the right hand for maximal receptor activation.

2.3 Procedure

The experiment consisted of 4 blocks of 8 trials. One trial consisted of a blue fixation cross appearing on the screen (between 1-2 sec before the start of the visual stimulation), a fade-in period of 5 sec in which the contrast was progressively increased, reaching 100% after 5 sec, 60 sec of stimulation, 5 sec of fade-out, after which the fixation cross disappeared after 1-2 sec. Participants were asked to fixate the cross through the entire trial, while accomplishing the task. The task consisted of counting the number of occurrences of colour change in the fixation cross. The fixation cross turner red during 200 ms two to eight times per trial. Participants gave a verbal response at the end of each trial, indicating how many colour changes they perceived. Two sources of stimulation were used in this experiment: visual and vibrotactile, at respectively 3.75 Hz and 4 Hz with the aim of tagging the EEG activities respectively elicited by stimuli of each of the two sensory modalities. Four conditions were tested during this experiment. In the first condition, the hand is placed on a box on the right of the visual stimuli, giving the perception that the image and the hand were at the same height, and the vibrotactile stimulus was applied simultaneously to the stream of visual stimulation. In the second condition, the hand is still placed next to the visual stimuli but no vibrotactile stimulation is applied. In the third condition, the participant keeps the hand close to his/her trunk on his knees and the vibrotactile stimulus is applied concomitantly. In the fourth condition, the participant keeps the hand close to the trunk without vibrotactile stimulus (Figure 3). Each condition consisted of 8 repetitions of the trial, leading to a total of 32 trials. One condition was tested per block. Block order was randomizes across participants.

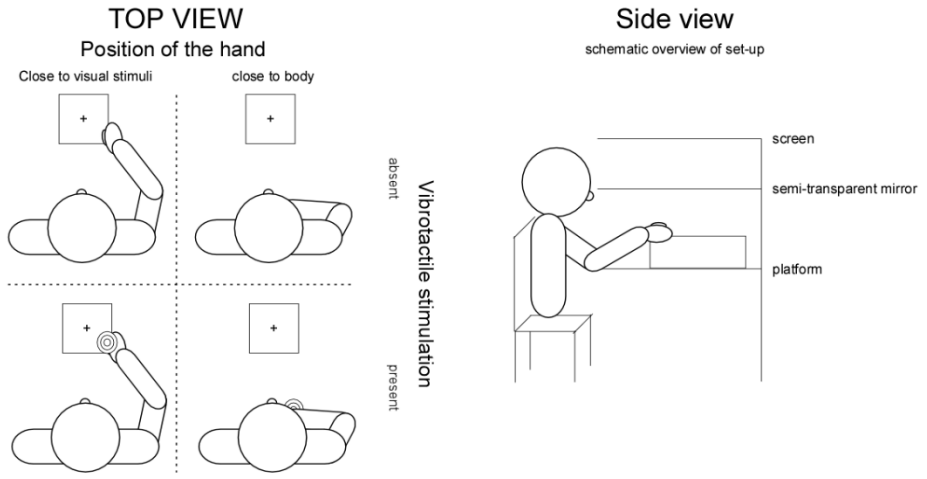

Figure 3: Experimental set-up. The left panel shows the four different experiment conditions, where the factor of vibrotactile stimulation (absent or present) and position of the hand (close to visual stimulation or close to body) are modulated. The right panel shows a schematic overview of the set-up. The image was displayed on the screen, and reflected in the semi-transparent mirror. The hand was placed on a box on a platform to give rise to the illusion that the images were projected at the same height as the hand.

2.4 Measures

Behavioural results consisted of reporting the number of perceived colour changes in the fixation cross. A response was considered as correct if the participants reported number of colour changes interruptions correctly.

The EEG was recorded using 64 Ag-AgCl electrodes placed on the scalp according to the international 10/10 system (Waveguard64 cap, Cephalon A/S, Nørresundby, Denmark). Electrode impedances were kept below 15 kΩ. Signals were amplified and recorded using a sampling rate of 1000 Hz and an average reference (64-channel high-speed amplifier, Advanced Neuro Technology,

Enschede, the Netherlands). The EEG data was analysed offline using Letswave 6 (https://www.letswave.org/) running in MATLAB 2014a. The data was first segmented from 2 sec before the fade-in period to 2 sec after fade-out period, resulting in 32 epochs of a total length of 74 s. We then applied a Butterworth 50 Hz notch filter followed by a bandpass-filter (0.1-100 Hz).

Frequency-domain analysis. All epochs were segmented to the stimulation length during which the maximum contrast was 100%, i.e. from +5 to +65 s relative to fade-in onset. Eye blink and movement artefacts were removed using an independent component analysis (ICA, Hyvärinen & Oja, 2000). Separate 60-s average waveforms were computed for each condition and participant. Subsequently, we performed a Fast Fourier transform (FFT) to transform the data into the frequency domain. Then, the contribution of background noise was removed by subtracting, at each bin of the frequency spectra, the average amplitude measured at neighboring frequency bins (four frequency bins ranging from −0.03 to −0.08 Hz and four frequency bins ranging from +0.03 to +0.08 Hz relative to each frequency bin (Colon et al., 2015; Mouraux, Iannetti, et al., 2011; Nozaradan et al., 2012a)). The validity of this subtraction procedure relies on the assumption that, in the absence of a strong periodic signal, the signal amplitude at any given frequency bin should be similar to the signal amplitude of the mean of the surrounding frequency bins (Mouraux, Iannetti, et al., 2011; Retter & Rossion, 2015). Two measures were extracted from the EEG frequency spectra. The first was a measure of the amplitude of the periodic EEG signal at the base frequency (i.e., 3.75 Hz and its harmonics). The second was a measure of the amplitude of the periodic EEG signal at the tool-selective frequency (i.e., 0.75 Hz and its harmonics). To select the statistically significant harmonics to include in the estimation of response amplitude, we expressed the amplitude at each harmonic frequency as a z-score relative to the mean and standard deviation of the amplitude of the 20

neighbouring bins of the FFT, ranging from -0.03 to - 0.18 Hz and from +0.03 to +0.18 Hz. A z-score greater than 1.64 (i.e., $p<0.05$, one-tailed) was the criterion to consider the amplitude at that frequency significantly greater than the amplitude at neighbouring frequency bins (Dzhelyova et al., 2017; Lochy et al., 2015; Quek & Rossion, 2017; Rossion et al., 2015). This selection of harmonics was performed on the grand average frequency spectrum, averaged over all participants, conditions and scalp electrodes.

Because we expected a modulation due to multisensory integration on the central electrodes, EEG responses were computed by averaging the responses obtained at central (FC1, FCz, FC2, C1, Cz, C2, CP1, CPz and CP2) channels. Indeed, multisensory areas reflecting the peripersonal coordinate systems are mainly situated in the parietal and premotor cortex (Brozzoli et al., 2011; Huang et al., 2018). We then summed the noise-subtracted amplitudes at the base frequency and (3.75 Hz, 7.5 Hz, 11.25 Hz and 15 Hz), and at the tool-selective frequency and harmonics (0.75 Hz, 1.5 Hz, 2.25 Hz, 3 Hz and 4.5 Hz) excluding tool-selective frequencies overlapping with the base frequency), to obtain a single measure of the base and tool-selective response for each participant and condition (Quek & Rossion, 2017; Retter & Rossion, 2015). Additionally, we extracted the amplitudes at 4Hz which reflects EEG response to the vibrotactile stimuli.

2.5 Analyses

We analysed the number of correct responses with a repeated measure ANOVA with two factors: hand position (BODY vs. VISUAL OBJECT) and vibrotactile stimulation (PRESENT vs. ABSENT). The sum of the amplitudes of EEG responses of the tool-selective response obtained in the different conditions were compared using a repeated-measures analysis of variance (ANOVA) with two factors: hand

position (BODY vs. VISUAL OBJECT) and vibrotactile stimulation (PRESENT vs. ABSENT). Additionally, we performed the same analysis on the sum of the amplitudes at the base frequencies. Effect sizes were measured using partial Eta squared. Greenhouse-Geisser corrections of degrees of freedom were performed when necessary. Significant interactions were followed up by contrast analyses using paired-sample Student t-tests. Significance level was set at $p \leq .05$ and corrected using the Bonferroni correction procedure. Lastly, we analysed the effect of the spatial proximity of visual stimulation (i.e. through the position of the hand) on the vibrotactile processing using a paired-sample Student t-tests (BODY vs. VISUAL OBJECT).

3. Results

Frequency-domain results. In all experimental conditions, a base response at 3.75 Hz and its harmonics was observed, with a high signal to noise ratio (Figure 4). The topographical distribution of this base response was maximal over occipital electrodes, and symmetrically distributed over the two hemispheres (Figure 5). The EEG frequency spectra revealed clear peaks at the frequency of tool presentation and its harmonics. Significant peaks were present up to the 8th harmonic (Figure 4).

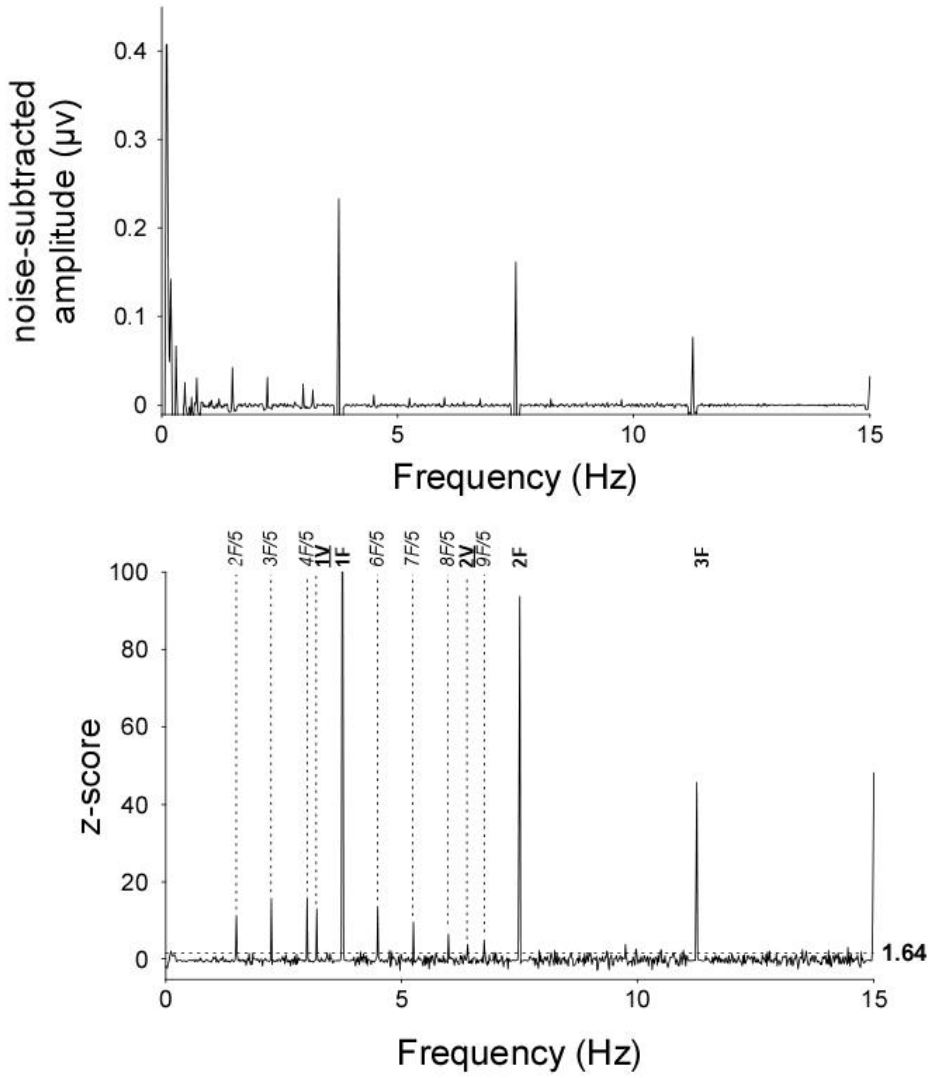

Figure 4: frequency spectra displaying the noise subtracted amplitude and the z-scores depending on frequency bin.

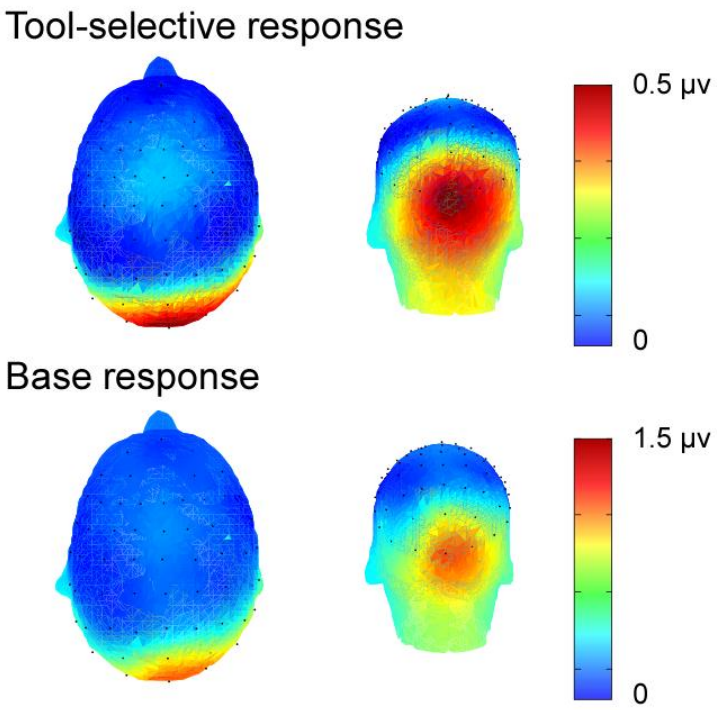

Figure 5: Topographies of tool-selective response and base response, are both categorized by large occipital activity with an additional smaller central component.

Behavioural response: the reported number of colour changes show no modulation by *hand position* ($F(1,15) = 0.04$, $p=0.83$, $\eta^2_p < 0.001$), *vibrotactile stimulation* ($F(1,15) = 0.31$, $p=0.59$, $\eta^2_p < 0.001$), or by the interaction between *hand position* and *vibrotactile stimulation* ($F(1,15) = 0.07$, $p=0.79$, $\eta^2_p < 0.001$).

Tool-selective response: In the main ANOVA, we find a main effect of *hand position* ($F(1,15) = 4.63$, $p=0.047$, $\eta^2_p =0.03$), with larger amplitudes of the FPVS response when the hand is near the body (0.13 ± 0.07) as compared to when the hand is near the visual stimulation (0.1 ± 0.06). The presence of the vibrotactile stimulation does not affects the amplitude of the tool-selective FPVS response (Effect of *vibrotactile stimulation:* $F(1,15)= 0.12$, $p=0.73$, $\eta^2_p < 0.001$), or interact with the position of the hand ($F(1,15)=2.25$, $p=0.15$, $\eta^2_p =0.01$) (Figure 6)

Base response: The repeated measure ANOVA on the base response does not reveal any significant effect of *hand position* ($F(1,15)= 0.05$, $p=0.82$, $\eta^2_p = 0.0$), *vibrotactile stimulation* ($F(1,15)= 0.07$, $p=0.79$, $\eta^2_p < 0.001$), or interaction between *hand position* and *vibrotactile stimulation* ($F(1,15)= 0.37$, $p=0.55$, $\eta^2_p < 0.001$).

Vibrotactile response: There is no effect of hand position on the vibrotactile processing ($t(15)=-0.50$, $p=0.62$)

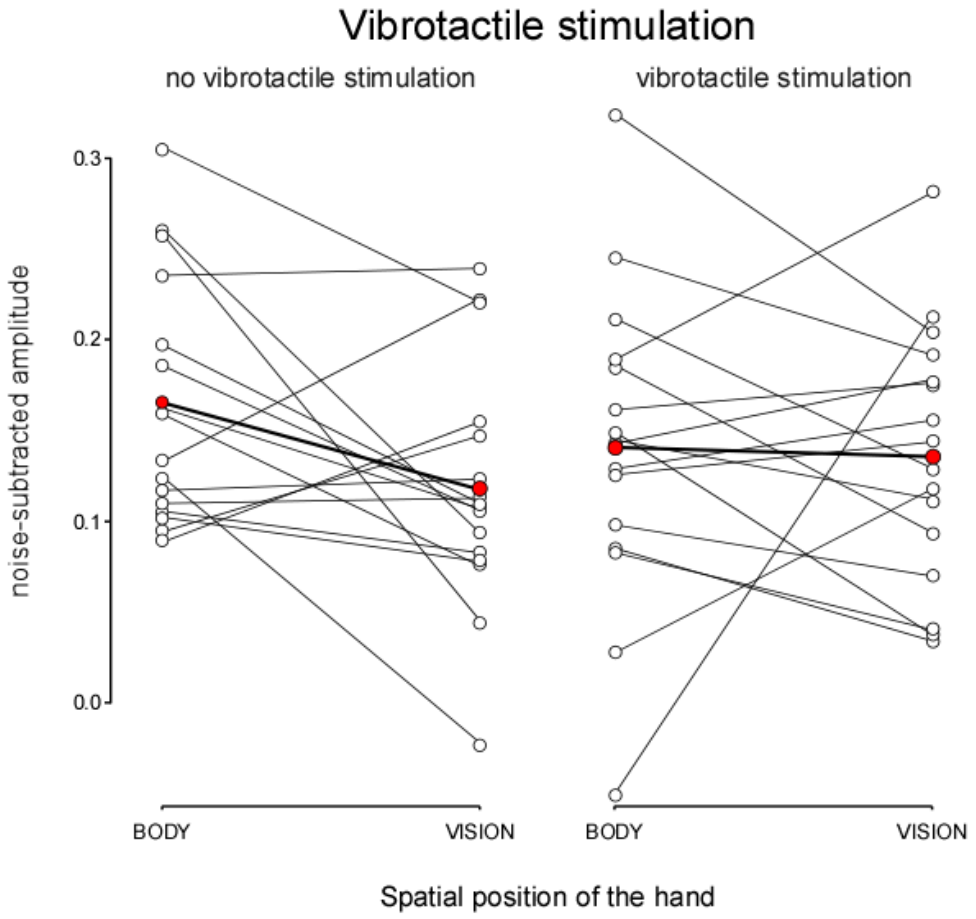

Figure 6: Individual values for the tool-selective response as a function of the presence of vibrotactile stimulation and the position of the hand. The tool-selective activity is not modulated by the presence of the vibrotactile stimuli, but is by the position of the hand, with larger activity when the hand is close to the body as

compared to next to the visual stimuli. Individual values are denoted by thin lines and white connectors. The red dots and black thick lines shows the group average.

4. Discussion

The present experiment aimed at studying the effect of vibrotactile on the processing of visual stimulation using a variation of FPVS. This variation of FPVS allows for separation of low-level visual processing and higher-order tool-selective visual processing, as processing between tools and non-tools differs (Study 2). As it is believed that multisensory integration is a higher-order cognitive process (Spence et al., 2009), we expected vibrotactile stimuli to affect the higher-order processing specifically, while the low-level visual processing would remain unaffected. The explanation of this effect could be based on two mechanisms. First of all, if a difference in visual processing is observed, this would indicate that the vibrotactile stimulation affects the visual processing. This shows that information from both modalities converge and are integrated into a multisensory percept, which does or does not warrant feedback connections, modulating visual processing at a low level. However, it has recently been shown that the processing of vibrotactile stimulation can be enhanced by a spatially congruent visual stimulation approaching the hand (Colon et al., 2015), showing that there can be an interaction between low-level stimulations, using FPVS stimulations.

The present design also relies on a second theoretical concept, the theory of affordance, which states that there is an automatic activation of a motor scheme compatible with the object. Importantly, haptic feedback, through sustained stimulation, is crucial in motor control, and consequently in tool manipulation, linking the theory of multisensory integration and affordances. In line with those two theories, we would thus expect to see larger activity when vibrotactile stimulation is spatially congruent with the visual stimulation. However,

the observed results do not allow to draw this conclusion, as no effect of the spatially congruent vibrotactile stimulation was observed. Furthermore, no effect of vibrotactile stimulation was observed at all, leading to the conclusion that vibrotactile does not impact visual processing, independent of its spatial location. Additionally, we tested the effect of affordance by modulating the position of the hand, which could either be located close to the body, or close to the visual stimulation. We expected that when the hand was located close to the visual stimulation, this would lead to an enhancement due to automatic motor scheme activation. Surprisingly, we observe the inverse. Indeed, tool-selective processing appears to be enhances when the hand is close to the body. As such, it could be that the hand serves as a distractor and not as a facilitator.

Interestingly, the topography of base response and tool-selective response are both characterized by activity at central and occipito-temporal electrodes. It is known that neural activities in cortical areas involved in low-level visual respond to a larger range of frequencies, whereas areas involved in higher-order processing, such as those hypothesized to underlie multisensory processing preferentially respond to lower frequencies and might not respond at all to high stimulation frequencies. By using a base frequency of 3.75 Hz, this is a frequency which can reach higher-order cortices. While we can conclude that the tool-selective activity is very unlikely to reflect low-level visual activity, we cannot conclude that tool-selective activity is not also partly reflected in the base activity. This issue could be remediated by using a higher base frequency. However, in the present experiment, this frequency was chosen as ERP between tools and non-tools start to differ from 250 ms onwards (Proverbio, Adorni, & D'Aniello, 2011).

Multiple reasons can be evoked to explain why no effect of multisensory integration or crossmodal interaction is observed in the present study. Indeed,

parameters such as frequency, length of stimulation might potentially influence this results (Macaluso, 2012), as considered in the general discussion. Amongst all the possibilities, one is related to the unity assumption. Indeed, while stimulus congruence can be defined in time and space, it can also be defined by its semantics (Talsma et al., 2010). The unity assumption refers to results showing that it is more likely to observe multisensory integration or crossmodal interaction if participants perceive the two stimuli of the different modalities as originating from a common source. While the validity of this assumption is still debated and only applies under particular conditions (Chen & Spence, 2017), it is worth noting that all effects that have been obtained by using semantic congruent stimuli, can also be observed by using more simple stimuli such as simple flashes and sounds (Bertelson & De Gelder, 2004). It is thus clear that having the impression that a common source underlays unimodal stimulation can influence performance. This demonstrates that the interaction between sensory modalities can also be influenced by top-down factors, such as this unity assumption. For example, Participants are better at performing temporal order judgement [TOJ] task when audio and visual stimulation are semantically unrelated (e.g. face men, female voice). As such, when sound and vision are congruent, this interferes with the TOJ task (Vatakis & Spence, 2007). This phenomenon has extensively been studied in audio-visual integration, but less in visuo-tactile integration. The few visuo-haptic studies on the unitary assumption show that when visual and haptic information about for example shape, is congruent, a multisensory object is constructed (Tsilionis & Vatakis, 2016). It thus remains difficult to comment on the link between the present study and the unity assumption, even more so because we do not research perception.

In conclusion, our observation does not demonstrate that by targeting higher-order visual processing, multisensory integration becomes more likely to observe.

GENERAL DISCUSSION & CONCLUSION

GENERAL DISCUSSION & CONCLUSION

The spatial organisation of the peripersonal space is not straightforward. The spatial positions of stimuli from different modalities are initially mapped in distinct coordinate systems, which need to be remapped in a shared multisensory reference frame in order to be integrated in the representation of the surrounding space for optimal functioning. This is not only necessary for defensive actions against potential threats, but also to be able to react and interact within the space directly surrounding us. As such, stimuli applied to the personal space, such as tactile stimuli, need to be processed in function of their location on the body but also in function of the position of the stimulated body part in external space. In addition, stimuli in the external space, such as visual stimuli, might signal the imminent contact of a potentially threatening object, especially when that object approaches the body. Importantly, spatial information about both tactile and visual information processing must be coordinated. Consequently, the peripersonal space is essentially a multisensory construct of the surrounding space.

Multiple sustained trains of stimuli allow to study by what spatial principles information from different senses interact and come together to form the peripersonal frame of reference. Research has repeatedly shown that presenting multimodal stimuli is prone to produce advantages as compared to unimodal stimuli, especially when congruent in space and time. Advantages include faster reaction times and modulated neurophysiological measures. However, theories to explain this improvement rely either on theories of multisensory integration, crossmodal interaction between sensory modalities, or a crossmodal link in spatial attention. While transient stimuli have often been used in experiments investigating this phenomenon, this kind of stimuli are especially sensible to exogenous attention and could be explained by the short lasting effects of

exogenous spatial attention. Indeed, a stimulus in one modality can attract the attention to its spatial location, and consequently facilitate the processing of stimuli from another modalities, especially if they are presented at the same spatial location.

Consequently, the main aim of the present work was to investigate the interaction between touch and vision to shed a light on the spatial organisation of their integrated representation. This was attempted using sustained trains of periodic stimulation, with the aim to observe activity at crossmodal frequencies, and potentially circumventing the short lasting effects of exogenous spatial attention, while this is traditionally undertaken using transient stimulations,

In this discussion, the main findings will first be reviewed. We will subsequently discuss the results with respect to the underlying mechanisms of spatial attention, looking at the interplay between attention and multisensory integration. Next, bottom-up and top-down factors which potentially influenced the results will be explored. This will be followed by the implications of these results for the defensive system. As the limitations of the studies will already have been discussed thoroughly at that point, we will finish with discussing potential further studies and concluding remarks.

1. Main findings

The present thesis aimed at studying the spatial organisation of multisensory integration using periodic sustained trains of periodic stimulations. A first step to investigate this spatial organisation was to observe if visual and tactile stimuli can interact when spatially congruent, evidenced by modulatory effects of spatially congruent vibrotactile stimuli on the visual response. Throughout all studies, we did not demonstrate a modulation of the visual processing though the

spatial congruence of vibrotactile stimulation. Additionally, while we were ultimately looking for crossmodal integration frequencies to evidence integration, these were absent in all studies. Despite this absence of results with respect to the question of this thesis, some effects unrelated to the main research question appeared to be significant.

In a first study, which consisted of three experiments, this was attempted using low-level visual and vibrotactile stimulation. While the research questions remained the same throughout the three experiments, several factors changed between each study, such as relevance of the task and attention. As such, in the first experiment, participants were asked to pay attention to vibrotactile and visual stimuli, and to report whether interruptions, which occurred in both the visual and vibrotactile modalities, were perceived as having occurred simultaneously or not. In this first experiment, we found no effect of spatially congruent vibrotactile stimulation on the visual response, as expressed by amplitude of the FPVS signal. As interactions between sensory modalities have been shown when one of the modalities is not task relevant (Eimer & Driver, 2001), a second experiment was conducted where vibrotactile stimulations were not task relevant and could be ignored by participants. In this second experiment, the experimental paradigm was kept exactly the same as in the first experiment. However, participants were asked to only pay attention to the visual stimulus, and to report the temporal order of the visual interruptions. Once more, no modulation of the visual processing by a spatially congruent vibrotactile stimulation could be shown. In a third experiment, spatial attention was directed either to the left or right side of space, by asking participants to only report interruptions (visual and vibrotactile) in that side of space. Spatial attention was divided in the previous experiments, where participants needed to pay attention to both sides of space to meet the task demands. As a crossmodal link in spatial attention is sometimes hypothesised to

underlie the interaction between different sensory modalities (Macaluso, 2012), we hypothesised that spatially congruent vibrotactile stimulation would affect the visual processing, but only when located in the attended side of space. However, we could not demonstrate an influence on the visual processing by the occurrence of the spatially congruent vibrotactile stimulation. Additionally, we expected an effect of attention on the visual processing. Indeed, attending FPVS is known to increase the amplitude of the response in the frequency spectrum. However, we showed an interaction between the factor congruence (i.e. vibrotactile stimuli and visual stimuli are spatially congruent) and attention (i.e. the visual stimuli is in the attended part of space). Contrasts on this interaction shows that there is an increase in visual activity when the visual stimulation is attended as compared to when it is in the unattended side of space. Surprisingly, this effect only appeared on the visual stimulation that was not next to the stimulated hand. Where studies suggested that attention increases the amplitude of the FPVS response, these results suggested that adding a tactile stimulation at the same spatial location was detrimental to the attentional effect. This implies that the presence of the vibrotactile stimulation next to the visual stimulation attenuated the attentional effect. However, this effect can potentially be attributed to the task demand in this experiment. Indeed, when visual stimuli were incongruent, participants only had to pay attention to the visual stimulation, while when the visual and vibrotactile stimulation were congruent, participants had to pay attention to two stimulations from different modalities. Overall, these outcomes suggest that the crossmodal attentional mechanisms which are involved in ERPs might not be involved in the FPVS response.

Therefore, we decided to try to target higher-order visual areas using images of tools, since LED stimuli are known to generate activation originating in primary visual cortices. A response which reflects activity originating in primary

visual cortices might not be suitable to study visuo-tactile interaction, which is believed to also rely on higher-order activation. Also, strong activity elicited by the primary visual cortices might mask activity related to visuo-tactile interaction. Additionally, tools were chosen as this object category might be particularly relevant for sustained stimulations, through feedback mechanisms during action. By using a variation of FPVS, we showed, in study 2, that processing which is specifically related to tool processing can be separated in the frequency domain from processing common to all object categories. This variation of FPVS relies first on the periodic presentation of visual stimuli from a variety of category, but periodically inserting an image of on specific category (i.e. a tool). We show that tools are processed differently, as evidenced by tool-selective activity reflected at the tool presentation rate. Additionally, this approach has the advantage of concentrating all activity other than the tool-selective activity at the frequency of the base response. We show that the tool-selective activity is absent when presenting phase-scrambled versions of the images, and that the tool-selective activity is not modulated by attention. This is of major interest for three reasons.

First, this visual paradigm has been validated using faces (Dzhelyova et al., 2017; Rossion et al., 2015), numbers (Guillaume et al., 2018), different semantic object categories (Stothart, Quadflieg, & Milton, 2017) and symbols (Lochy et al., 2015) as category specific visual stimuli. The present results incremented the current knowledge, as we showed that tools elicit specific activity which is related to a differential temporal activity as compared to other object categories. This activity was characterized by a more centro-parietal component, which is potentially linked to the automatic activation of the dorsal visual stream, related to the spatial processing of visual stimuli.

Second, the same paradigm has proven useful in the research centred on face perception. However, it has been shown that face perception is highly

dependent on visual experience as faces are a very significant visual category. The results of the second study thus strengthened the proofs of this methodology, as tools might have a specific processing, but are probably not as familiar and extensively trained as faces (Lochy et al., 2018).

Third, separating higher-order visual processing from the low-level processes has been a challenge in the field of visual research. Diverse solutions such as phase-scrambling (which removes the semantic content of an image while preserving low-level visual features) have been proposed, but those cannot fully assure identical low-level visual features. This variation of FPVS allows to introduce as much variation in terms of low-level visual features as wanted into the stimuli, while traditionally stimuli are transformed to obtain equal spatial frequencies, luminance and contrast spectra through all stimuli. This is usually done to exclude the possibility that activity can be related to these variations in low-level visual features. Similarly, using FPVS, introducing variation in low-level visual features has the advantage that if a so-called category specific activity is observed, chances of it being related to low-level features are small. Indeed, as the paradigm relies on periodic presentation of stimuli, only periodic variation of the low-level visual features at the same frequency as the presentation rate of the category of interest can influence the category specific activity. Altogether, allowing for natural variation makes for a more ecological approach of the visual system. While traditional methods struggle to separate low-level visual processing from processing related to higher-order features, FPVS does this inherently, without recurring to obligatory post-hoc contrasts between different conditions.

In study 3, the exact same variation of FPVS, as validated in the second study, was used in a set-up to study the interaction between visual and tactile stimuli. This allowed dissociating the effect of spatially congruent tactile stimuli on low-level and higher-order visual signal. The study aimed at showing the

modulation of higher-order visual processing by the occurrence of spatially congruent vibrotactile stimulus. Also, to control for the necessity of the vibrotactile and visual stimulus to be spatially congruent, the hand on which the vibrotactile stimulus was applied was placed either next to the visual stimuli, or close to the body (on the lap). We expected to see an interaction between the two factors (hand position and vibrotactile stimulation), which would potentially show a modulated visual processing by spatially congruent vibrotactile stimuli. However, the only significant factor was the position of the hand, with an increased visual processing observed when the hand was next to the body, as compared to when the hand was next to the visual stimulation stream. Surprisingly, the vibrotactile stimuli did not impact the tool-selective response or the base response significantly. Here again, it appears that the hand, both with or without vibrotactile stimuli, represented some kind of distractor, and instead of being an advantage for the visual processing, appeared to have a detrimental or inhibitory effect.

These findings again are controversial and unexpected. Indeed, the literature shows behavioural facilitation when hand and tool are congruent (i.e. the tool handle can naturally be grasped by the right hand and the participant is asked to respond with the right hand), which we do not observe in the present study. We expected a modulation of the tactile and visual response due to the presence of the hand near the stimuli, as the orientation of the presented tool was congruent to the hand placed next to it. Moreover, we also expected that this effect would be enhanced by the presence of a vibrotactile stimulation due to multisensory integration (i.e. two stimulation which occur close in space have a higher chance to interact). However, we did not observe the expected effects. Instead, we showed that the amplitude of the tool-selective activity is reduced when the hand is placed next to the visual stimuli. Interestingly, this pattern is only observed in the tool-selective response and not in the base response, suggesting a differential effect of

proximity related to general visual processing or to tool-selective processing. Generally, these results lead to think that placing the hand (with or without stimulation) next to a visual stimulation leads to a degradation of the visual signal, especially in the tool-selective signal.

The expected effect of the position of the hand was based on studies of Tucker and Ellis (1998, 2001) who showed faster reaction times when hand and object are spatially congruent. However, there is a major difference between the present study 3 and the studies of Tucker and Ellis (1998, 2001). In those studies, participants were required to make a motor response to an unrelated feature of the presented tool, such as if the object was presented upright or inverted. Importantly, participants responded with the hand that was sometimes congruent to the object orientation (left – right orientation). The advantage of congruence is thus consequently observed on the motor plan. In study 3, participants were not required to respond with a motor action, but reported changes in the colour of the fixation cross. Additionally, participants were also not required to make (unrelated) judgements on the presented images. Not involving motor preparations/actions or judgements could explain the contradictory results, and could be worth taking a look at in future studies.

Overall, the results of the studies aiming at observing a modulatory influence of spatially congruent vibrotactile stimulation on visual processing demonstrated that the paradigm might not be optimal. As a consequence, numerous experimental factors could be looked at and if possibly improved such as stimulation type, frequency, number, length, modality, or task and perception. Indeed, it is possible that our stimulations were not adapted to trigger the multisensory peripersonal space. Last, our results could suggest that the processes underlying the known effects of crossmodal spatial attention do not affect the

FPVS response as it affects ERPs. Even more, it could be that ERPs and FPVS reflect different processes.

2. Bottom-up and top-down factors influencing the interaction between modalities.

The absence of significant results is difficult to interpret as an absence of visuo-tactile interactions. Indeed, absence of evidence should not be confused with evidence of absence. The results of the multisensory integration studies (studies 1, and 3) did not indicate a visuo-tactile interaction, and multiple reasons can be evoked to attempt to explain this absence of results. In this part, reasons related to the experimental paradigm for the absence of evidence will be discussed. These factors can be divided in stimulus driven, or bottom up factors (such as type of visual stimulus, frequency, number of stimulation, duration and the dominance of the visual modality), and top-down factors (such as task demand and perception). However, again, it is important to keep in mind that these results do not imply that visuo-tactile interaction is impossible to evidence using a FPS methodology.

2.1 Bottom-up factors

2.1.1 Type of visual stimulation

The type of visual stimuli used and by which cortical area it is processed could have had an impact on the obtained results. The stimuli used in the first three studies relied on the periodic modulation of luminance, believed to target low-level visual areas. Indeed, simple flashing LED lights are most likely to mainly activate V1 (Herrmann, 2001; Srinivasan et al., 2006; Vialatte et al., 2010). The visual stimuli used in the study 2 aimed at separating low-level visual processing

from higher-order visual processing, consequently targeting additional higher-order regions in the brain along the ventral and dorsal visual pathway.

Theoretically, an interaction between the visual and the tactile system is possible and has been shown before using simple brief flashes (Kennett, Eimer, et al., 2001; Macaluso et al., 2000). As such, whether stimuli relied on periodic modulation of low-level or higher-order visual features should not influence the underlying processes. However, one can posit that if the process of the interaction between sensory modalities is a higher-order process, and if this activity is evoked by higher-order associative cortices, this activity is potentially masked by the activity originating in V1, which is very robust (Di Russo & Spinelli, 2002; Ding et al., 2005; Herrmann, 2001; Srinivasan et al., 2006). Indeed, while the existence of bimodal neurons have been shown in multiple areas in the brain (Duhamel et al., 1998; Graziano & Gross, 1995; Murata et al., 1997; Royal, Carriere, & Wallace, 2009a; Stein et al., 1993, 2004; Wallace et al., 1992), bimodal neurons mostly make up a fraction of that region (Duhamel, 2002; Gentilucci et al., 1983; Graziano et al., 2002; Laurienti et al., 2005; Murata et al., 1997; Rizzolatti et al., 1981a). As such, activity originating from bimodal cells might be expected to only cause a weak EEG signal, if strong enough to evoke any signal at all (Laurienti et al., 2005). Separating low-level visual activity from higher-order visual processing would then increase chances of observing activity related to bimodal neurons, by removing the prevailing low-level visual activity. Study 3 aimed exactly at doing that, by capturing EEG activity related to processing common to all visual stimuli in the base frequency, and tagging activity related to tool-selective processing. We did not expect to see a modulation of the low-level visual activity, in accordance with the results of the first three experiments, where no effect of spatially congruent vibrotactile stimulations was observed on low-level visual processing. In the 3rd study, no modulation of the visual processing at any level (higher-order and low-

level visual processing) could be demonstrated. This implies that if the experimental paradigm succeeds in activating bimodal neurons in higher-order cortices, their activity is potentially not strong enough to be picked up with the EEG methodology.

In conclusion, the type of visual stimulation used should not have influenced the present results. However, selectively targeting higher-order visual areas might be beneficial for studying visuo-tactile interactions by making higher-order multisensory processes more visible from a predominant low-level visual signal.

2.1.2 Stimulation frequency

A second factor that could have influenced the outcome of the experiments is the frequencies used. Indeed, it is possible that some frequencies are more prone to effects of interaction or integration. Frequency related arguments are twofold: First, specific brain regions display preferential frequencies, and second, cognitive processes do not affect all frequencies in the same way.

First, Gauthier (2012) demonstrated that neurons have an inherent preferential frequency, at which neurons are optimally activated. They tested the temporal tuning of the neurons along the ventral visual stream in human participants, by presenting stimuli at different frequencies. In a first experiment, they showed that the peak responses showed at higher frequency ranges in low-level visual cortical areas, and at lower frequency ranges while travelling up the ventral visual stream. This makes sense for two related reasons: more complex stimulations need more processing time, consequently resulting in an optimal low frequency. Second, slow stimulating frequencies allow for large integration

windows, allowing for time differences between incoming sensory signals (Gauthier et al., 2012; Koenig-Robert & VanRullen, 2013). In a second experiment, they aim at studying the effect of an attentional task on the tuning peak values. They found that paying attention to the stimulation stream slows down the tuning peak. Paying attention to stimulation implies that regions preferentially respond to slower frequencies of stimulation. By slowing down the preferential frequency, a longer window between stimulations becomes available. Therefore, when paying attention, stimulations can be integrated or interact together in a longer time window. These results can probably be generalised towards all low-level versus higher-order areas in the brain.

This preference for lower frequency in higher-order areas has been demonstrated in the human observer (Alonso-Prieto et al., 2013; Koenig-Robert & VanRullen, 2013). The visual stimuli in this particular study consisted of a stream of identical faces, or faces with different identities (Alonso-Prieto et al., 2013). Importantly, showing faces with the same identity implies that the face does not need identification at every exemplar, and does not rely on higher-order areas as much as if the identity was changed at every exemplar. Showing faces where the identity is modulated at every cycle does imply higher-order categorizing. The authors tested different frequencies of stimulation and were interested by the frequency at which the processing between the two conditions would become identical, as this would be an indication of the boundary between low-level and higher-order processing (Alonso-Prieto et al., 2013). Indeed, when processing becomes similar, the higher-order visual processing linked to identity processing cannot take place anymore. The authors observe that while the response between those two conditions differs at frequencies up to 10Hz, the response to both becomes identical when frequencies higher than 10Hz are used. This indeed implies that more time is needed for the analysis of facial identities before the next

image comes to interrupt or reset that processing at frequencies higher than 10 Hz (Alonso-Prieto et al., 2013). Likewise, developing and testing a method to disentangle low-level visual processing from higher-order processing, the authors were also interested in the temporal dynamics and limitation of their method (Koenig-Robert & VanRullen, 2013). By testing different frequencies, they argue that the optimal frequency lies between 4Hz to 7 Hz, which is a decision also based on the disappearance of the response related to the higher-order visual processing.

At a cellular level, the spatiotemporal characteristics from multisensory neurons were studied in the cat anterior ectosylvanian sulcus (Royal et al., 2009). The authors demonstrate that temporal characteristics of the response of bimodal neurons differ notably depending on the stimulation (unimodal or bimodal). These bimodal neurons display a higher and longer discharge rate for multisensory stimulations as compared to unisensory stimulations. The authors also argue that this behaviour reflects the longer window in which the multisensory response can take place, and in which stimulations from different sensory modalities can interact. This observation could explain why lower frequencies are preferred in multisensory areas.

Second, different stimulation frequencies can be impacted differently by task. Indeed, depending on stimulation frequency, either the attended or unattended stimulations are boosted (Ding et al., 2005). By displaying two search arrays, one flickering at different frequencies, and the second one flickering non-periodically, these authors tested the effect of attention on different frequencies. Indeed, the search arrays could be attended or unattended. As such, they showed that attention enhances the response when stimulation in the delta (2 Hz - 4 Hz), and upper alpha range (10Hz – 12 Hz). They suggested that attention enhance the phase locking and consequently the signal. Conversely, they argued that when

stimulation frequencies in the low alpha range (8 Hz – 10 Hz) are used, involved cortices preferentially respond to the unattended stimulation (Ding et al., 2005).

Consequently, the frequency of stimulation used in the experiments of this thesis could have influenced the results we obtained in two ways. First, it is possible that the frequency used is not suited or optimal to tag the network we are looking to activate. Even though the stimulation frequencies used were below 10 HZ and theoretically suited to reach higher-order cortices, we cannot ascertain that other frequencies might be more appropriate. Second, it is possible that the mechanisms we are looking to observe do not affect the chosen frequencies.

2.1.3 Number of simultaneous stimulations

A third potential explanation is the role of the additional visual stimulation and the occurrence of the vibrotactile stimulation. While the occurrence of the vibrotactile and visual stimuli was obviously necessary to study visuo-tactile interaction in the present set-up and was expected to act as a facilitator, it could have had the inverse effect and acted as a distractor on the visual processing.

First, concerning multiple visual stimuli, the results could have been influenced by the presentation of two visual stimuli. Participants were required to fixate in between the two visual stimulations. However, it is known that intraocular grouping - perceiving the two streams as one – is evidenced by intramodal crossmodulation frequencies (Giani et al., 2012), which are present in experiment one, two and three of the first study. It is thus possible that the two visual streams were not perceived as two separate streams but as one merged visual stimulation, despite the ability of the participants to indicate where the visual interruption was taking place, Consequently, it is possible that the visual stimulation could not be integrated with the vibrotactile stimulation which was lateralized.

Second, concerning the vibrotactile stimulation, perception is a factor that could have made the difference between a facilitator and distractor effect. Indeed, if visual and tactile stimulation were perceived as belonging together or as originating from a shared source, it is more likely that the vibrotactile would have acted as a facilitator (Macaluso, 2000; Zimmer & Macaluso, 2007). However, by using different frequencies for visual and tactile stimulation, we consequently delivered asynchronous stimulation. As such, it is quite possible that participants never had the perception that the different stimulation streams originated from the same object or shared a common source. Therefore, the vibrotactile stimulation could have acted as distractor and not promote multisensory integration. Evidence for hypothesis comes from the study using FPS exploring intramodal attention (Porcu et al., 2013). By asking participants to either pay attention to the visual or auditory modality, they observe an increase in visual activity when participants payed attention to the visual as compared to when participants payed attention to the auditory modality. They demonstrated that deploying attention to one modality did not necessarily boost the other modality, as paying attention to the visual modality did not boost auditory processing and vice versa.

Importantly, it appears that the facilitator or distractor role of the stimuli in a second modality is not fixed and can be modulated by task demand. Sinnett et al. (2008) showed that this differential facilitation and distractor effect can be demonstrated in reaction to a same bimodal event. In their study, participants were asked to detect targets in one part of the experiment and make a discrimination response in a second part. They were required to respond with a button press to a target, which could be visual, auditory or audio-visual in separate condition. They observed that participants were significantly faster to react to bimodal stimuli as compared to unimodal stimuli. In contrast, when participants

were asked to make a discrimination response depending on the modality in which the target appears, reaction times to bimodal stimulation were markedly slower. While these results indicate that bimodal stimulation can have a differential effect depending on the task demands, it also suggests that applying two congruent stimulation does not always and automatically results in a facilitation.

While there is no direct evidence that placing a (stimulated) limb close to visual stimulation could act as a distractor, it is an effect we found indirectly. Indeed, in experiment 3 of the first study, we show that the effect of attention on visual processing is only significant if the vibrotactile stimulation is not spatially congruent. Similarly, in study 3, we show that the tool-selective visual processing is diminished if the hand is positioned close to the visual stimuli, irrespectively if the hand is stimulated. It can thus well be that, while the vibrotactile stimulation was meant to enhance processing, that it has an inhibitory effect instead.

2.1.4 Duration of stimulation

There are multiple reasons and advantages related to the use of sustained trains of stimulation. First, interactions between sensory modalities have repeatedly been shown using transient stimuli. It has been proposed that these interactions could be explained by exogenous spatial attention, as exogenous spatial attention is short lived, lasting a few hundred milliseconds (Cheal & Lyon, 1991; Müller & Rabbitt, 1989; Nakayama & Mackeben, 1989). Sustained trains of stimulation could be used to limit the influence of exogenous spatial attention. It is however more complicated to limit effects of endogenous attention, which is suggested to last much more longer than the effect of exogenous attention (Müller & Rabbitt, 1989). Limiting the effect of endogenous attention could allow to separate, if possible, between attentional and multisensory effects. As such, by

using sustained trains of stimulations, we aimed at minimizing these short lasting effect. Additionally, we also aimed at diminishing the potential effect of spatial endogenous attention by making participants pay attention to all stimulations simultaneously. However, while we aimed at reducing or controlling spatial attentional effects, it is probably difficult or even impossible to totally get rid of spatial attentional processes.

Second, the main advantage of using FPS in this thesis lies in the frequency-tagging approach. Indeed, by using different frequencies, it is possible to disentangle the neural signal and relate it back to the stimulation, even when multiple stimulations are presented simultaneously. However, it is possible that using different stimulation frequencies presents a disadvantage, especially in relation to multisensory integration and endogenous neural oscillations.

Endogenous neural oscillations are important for cognitive processes (Lakatos et al., 2008; Zoefel, ten Oever, & Sack, 2018). The phase of this endogenous neural oscillations is believed to reflect changes between high and low excitability phases (Buzsaki & Draguhn, 2004). As such, neural oscillations have been assumed to underlie the amplification of important input, by aligning the high-excitability phase with the predicted timing of expected events (Schroeder & Lakatos, 2009). Consequently, the phase of the endogenous oscillations could theoretically play a major role in multisensory integration, perception and the binding of different stimulation, through an alignment of the phase of the endogenous neural oscillations for the predicted timing of synchronised multisensory input.

An intriguing observation is that when presenting an auditory and a visual stimulation with a variable delay, the oscillations in the gamma band [GBO] were the highest when the delays correspond to the cycle length of gamma. This supports the hypothesis that these endogenous neural oscillations (especially in

the gamma band) are important for multisensory integration and perception (Fries, 2005). It can thus be said that these endogenous oscillations are suitable for communication between different sensory areas, depending on the phase.

Moreover, the binding of different stimulations, which is of key importance in multisensory integration, could occur through the endogenous oscillations (Gray, König, Engel, & Singer, 1989; Kaiser, Hertrich, Ackermann, & Lutzenberger, 2006; Senkowski, Talsma, Herrmann, & Woldorff, 2005). Indeed, the phase coherence between two neuronal groups has been shown to predict each other's response strength (Womelsdorf et al., 2007). Additionally, it has been shown that when an auditory and visual stimulation are presented together, this results in an increased GBOs, which have previously been linked with multisensory perception (Senkowski, Schneider, Foxe, & Engel, 2008; Senkowski, Talsma, Grigutsch, Herrmann, & Woldorff, 2007). Potential mechanisms by which multisensory integration or crossmodal interaction occur have also been elucidated. Indeed, studies have shown that somatosensory input can reset the phase of the ongoing neural oscillations in the auditory cortex, potentially preparing for the incoming auditory stimulus (Lakatos et al., 2007). In line with these findings, Senkowksi (2008) argued that multisensory integration was likely to occur because incoming stimuli lead to increased phase coherence in areas which are primarily responsible for unisensory processing. This in turn would increase activity in multisensory regions which receive input from the unisensory regions (Senkowski et al., 2008).

The processes reflected through FPS are still debated. Indeed, some argue that the FPS response reflects a regular repetition from evoked responses (Capilla et al., 2011; Keitel, Quigley, & Ruhnau, 2014) while others claim that FPS entrain the endogenous oscillatory activation (Nozaradan et al., 2015; Spaak, de Lange, & Jensen, 2014; Zoefel, Archer-Boyd, & Davis, 2018; Zoefel, ten Oever, et al., 2018).

Additionally, the ability of FPS to modulate the endogenous neural oscillations is much debated (Zoefel, ten Oever, et al., 2018). However, there are two different lines of evidence showing the link between endogenous neural oscillations and FPS. First, FPS activity has been demonstrated without direct input, such as to imaginary rhythmic activity, or by suggested rhythm (Celma-Miralles, de Menezes, & Toro, 2016; Nozaradan, Peretz, Missal, & Mouraux, 2011; Nozaradan et al., 2015). Indeed, by suggesting a rhythm, Nozaradan (2014) showed FPS activity at the frequency of this suggested rhythm, even though that frequency was not present in the spectrum of the stimulation (Nozaradan, 2014). Interestingly, similar results have been shown in the visual modality, by asking participants to imagine a beat on regular visual stimuli (Celma-Miralles et al., 2016). Second, it has been shown that while looking at the envelope of the different frequency bands, the frequency of stimulation can be uncovered (Colon, Liberati, & Mouraux, 2017). This study demonstrated that the frequency of stimulation could function as a modulating frequency for endogenous oscillatory activity. Indeed, the authors deliver sinusoidal laser stimuli at a frequency of 0.2Hz, and discover activity at 0.2Hz in the envelope of theta, alpha and beta band. Similarly, it has been shown that GBO are modulated at the frequency of stimulations in the attended visual modality (Lakatos et al., 2008). As such, FPS and associated perceptions can induce endogenous neural oscillations at a certain frequency. Additionally, FPS can potentially modulate the ongoing endogenous oscillations at the frequency of the delivered stimulation, showing an interplay between evoked responses and endogenous oscillatory activity.

As such, if a periodic stimulation is indeed able to phase reset the ongoing endogenous oscillatory activity, and if phase coherence across sensory modalities is necessary for multisensory integration, the frequency tagging approach, which is the major conceptual advantage of the FPS approach in multisensory integration,

might actually block the possibility of observing an interaction between touch and vision.

2.1.5 Dynamic stimuli to target the multisensory peripersonal space?

The peripersonal space is theorized to be a interface for somatic and non-somatic stimuli to interact. As such, this enables actions in our direct environment (Cléry et al., 2015), and also defensive actions (Graziano and Cooke (2006)). Interestingly, experiments studying this spatial organisation of the peripersonal space in non-human primates use dynamic visual stimuli (i.e. approaching or receding), and have shown that the multisensory areas are sensible for direction of movement (Bremmer, Schlack, Duhamel, Graf, & Fink, 2001). This has been corroborated by studies in humans. Indeed, it has been demonstrated that a similar multisensory system for the peripersonal space is particularly sensitive to movement (Bremmer, Schlack, Shah, et al., 2001; Brozzoli et al., 2011; Colon et al., 2015; Huang et al., 2018). Interestingly, studies exploring visuo-tactile interactions could thus be separated in two clusters: one cluster using dynamic stimuli, using approaching and receding stimuli (which might activate the multisensory integration network), and one using static stimuli (which might be more sensitive to crossmodal spatial attention effects).

Bremmer et al. (2001) showed activation using the fMRI in the multisensory network of the peripersonal space using moving stimuli. The authors sought to directly test the equivalencies between monkeys and humans concerning multisensory integration in body centred coordinate system by presenting similar stimuli. They found activity in the IPS, PMv and lateral inferior postcentral cortex. Consequently, they conclude that the network is similar between macaque

monkeys and humans. More recently, Huang et al. (2018) showed activation in the multisensory network when visual looming stimuli were matched with tactile stimulations, as air puffs. Participants looked at looming visual stimuli, which were either delivered on the left or right side of a screen. Simultaneously, air puffs were delivered to the participants, either on the left or right side of his face. Thus, air puffs and visual stimulation could be either presented in a congruent fashion (i.e. same side of space) or incongruent (different side of space). The activation was largest when the multimodal stimulation was perceived as delivered simultaneously. Remarkably, the visual stimulation was perceived as simultaneous to the air puff when the air puff was delivered towards the end of the visual looming stimulation. This data corroborates the results found in the study of Colon et al. (2015), where they show an increase in the tactile FPS with approaching and receding visual stimulations, but only when the visual stimulation is making approaching movement and is on the hand. While Colon et al. (2015) interpret their results as a crossmodal link in spatial attention, it is indeed possible that the multisensory integration network in peripersonal space can also explain the results, in view of the similarity of the stimulation (as in approaching or receding) used as in Huang et al. (2018) and Bremmer et al. (2001).

Based on the results of the studies of Bremmer (2001) and Huang (2018), one of the decisive factors to activate multisensory integration could well be this approaching movement. Altogether, it appears that it is more likely to activate the multisensory network of peripersonal space using dynamic stimuli than using static stimuli as we did in the present thesis.

2.1.6 Visual dominance

Another factor that could underlie the absence of results is the dominance of the visual modality. Indeed, through the years, studies have shown that the

visual processing is difficult to modulate in spatial tasks i.e. the spatial position of a visual stimulus remains stable, even with the delivery of a stimulation in a different modality at a different location (Bertelson et al., 2000; Vroomen et al., 2001). Consequently, it has been formulated that the visual modality is dominant as compared to the auditory and tactile modality. The dominance of the visual modality with respect to the auditory or tactile modality finds it ground in the Colavita visual dominance experiment (Colavita, 1974). Participants were asked to look for auditory and visual targets in an audio-visual stream, and press one key when the target was in the visual modality, or another when the target was auditory. However, unknown to the participants, target could also be audio-visual. The authors report that when targets were audio-visual, participants only reported the visual modality. While there are various methodological problems with the initial study, this paradigm has since then been improved and used to study crossmodal extinction (Spence, Parise, & Chen, 2012). Crossmodal extinction is a paradigm where two stimulations from different modalities are presented to the participant. Participants are required to report the stimulations. While participants are capable of reporting the stimulations when they are presented alone, they encounter difficulties when the two stimulations are presented together. In the visuo-auditory case, participants often only report the visual stimuli, which is said to extinguish the auditory stimuli. Indeed, bimodal stimulation are not always reported correctly, and that the visual stimuli can extinguish both auditory or tactile stimulation (Hartcher-O'Brien et al., 2008; Spence et al., 2012).

Visual dominance has also been reported in studies that mimic the ventriloquist effect. Studies researching spatial judgements have demonstrated that the spatial location of auditory stimulation are being misjudged and biased towards the visual stimulation, when there is a spatial separation between visual and auditory (Alais & Burr, 2004). This appears to be a very robust effect, and is not

affected by training nor do the visual stimuli need to be attended or consciously perceived (Vroomen & De Gelder, 2004). Similarly, Eimer (2004) demonstrated that vision can also bias the spatial localization judgement of tactile stimulations (Eimer, 2004). Similar results have been described in endogenous attention. Indeed, results show that while the early components of tactile ERP are affected by paying attention to visual stimulation presented in the same spatial location, visual ERPs are not affected the same way, again demonstrating that visual stimulation can affect the processing of other modalities, but is not or less modulated itself (Eimer & Driver, 2000). From these studies, it has been proposed that the representation of our world is shaped and dominated by the visual input.

However, while most evidence points to a bias towards the visual stimulus, the opposite has been demonstrated as well. Indeed, it is also possible to bias visual perception towards the auditory stimulus (Bertelson & Radeau, 1981). As such, presenting one light with clicking sounds will make the light appear to flicker. Additionally, this improves the performance on visual discrimination task (Berger, Martelli, & Pelli, 2003; Morein-Zamir, Soto-Faraco, & Kingstone, 2003). Similar, if a pulsing sound is presented with a light, this light will also appear to be pulsating (Shams, Kamitani, & Shimojo, 2000, 2002, 2004). The dominance of the visual or auditory modality appears to depend on the information provided by the stimulation. Indeed, visual information has a high spatial information content, while auditory stimulation carry high temporal information (Burr & Alais, 2006; Burr, Banks, & Morrone, 2009). This implies that visual might be 'dominant' in tasks which involves spatial judgements, whereas auditory might be more relevant and thus dominant in paradigms where timing is more important (Kubovy & Van Valkenburg, 2001).

Consequently, in all studies presented as part of this thesis, visual dominance is one of the factors that could explain why no modulation of the visual processing was observed. Overall, multiple factors can explain the absence of evidence of multisensory integration. Importantly, this also indicates that the effect of spatial congruence might not be as robust as previously thought, as it does not survive modulations in stimulus features. In the previous sections, we discussed potential reasons which might explain the lack of positive results in the light of bottom-up or stimulus-driven factors: type of visual stimulation, frequencies, number of stimulations, duration and the dominance of the visual modality. However, top-down or cognitive processes could also influence the results. We thus continue to discuss top-down factors such as task demand and multisensory perception

2.2 Top-down factors

2.2.1 Task demand

Spence (2013) hypothesized the necessity of a task involving space to observe a changed outcome variable through spatial congruence of a stimulation of a different sensory modality in human studies. He argued that to observe an effect of spatial congruence, participants need to make a response relying on spatial orientation, or more generally, make a response which obliges them to take space into account. Contrarily, if the task did not force the participants to take space into account, this effect is greatly diminished, or even disappears. Examples of tasks in which there is no effect of spatial proximity are tasks in which stimuli need to be identified (Doyle & Snowden, 2001), or related to temporal perception (Innes-Brown & Crewther, 2009). As such, in a non-spatial discrimination task, participants were asked to discriminate between an arrow pointing leftwards or

rightwards (Doyle & Snowden, 2001). This visual stimulus was sometimes accompanied by an auditory stimulus, which could be originating either from the centre or the same side as the visual stimuli. The authors show that auditory stimuli did overall facilitate discrimination performance, but did this independently if the sound was spatially congruent with the visual stimuli. An additional example in which the spatial rule did not bring the expected results in a visuo-tactile paradigm it the TOJ experiment carried out by Keetels and Vroomen (2008). Participants were asked to judge the temporal order of two visual stimuli. Tactile stimuli were delivered either simultaneous with the visual stimulation, or before the first and after the second visual stimulation, 'pulling' the visual stimuli apart. Interestingly, they show that this effect also holds when the tactile stimuli are not proximal to the visual stimuli – i.e. the hands not placed near the visual stimuli. They show an interaction between the visual and tactile stimuli, but the spatial congruence between the tactile and visual stimulation does not play a role in the visuo-tactile interaction. Indeed, they merely had to be presented in the same side of space. Importantly, this paradigm focusses on temporal precision and does not explicitly require spatial orienting.

Interestingly, van Atteveldt et al. (2007) also argue that task demand can annihilate bottom-up multisensory integration. They presented participants with sound-letter pairs, which could be congruent or incongruent (i.e. the same letter heard as seen). When participants were listening passively to the stimulation, the authors uncovered an effect in the auditory association cortex that was larger for congruent pairs as compared to incongruent pairs. However, when participants were asked to judge whether letters and sounds were the same of different, this effect disappears. They argue that this can be explained by the fact that the modulation due to congruent audio-visual stimulations can be overruled when

explicitly matching both stimulation, which makes the congruence factor irrelevant (van Atteveldt et al., 2007)

In the present case, participants had to focus on all three stimulations simultaneously (two visual and one vibrotactile stimulation) and report interruptions, which could occur simultaneously or not. As such, while the spatial congruence of the vibrotactile stimulation was modulated, it was not necessary for the participants to take it into account for the task. This raises a potential issue. Indeed, the effect of the spatial rule on the visual processing could potentially have been demonstrated if the participants needed to make a response relying on the spatial aspects of the experiment or if participants were passively observing stimulations.

2.2.2 Unity perception

A top-down principle of multisensory integration, which has not been stated in animal research is the requirement of a unified percept to boost multisensory integration (Welch, 1999). Research about this topic comes from the ventriloquist illusion. In the ventriloquist illusion, auditory and visual stimulation are spatially disparate, however, the auditory stimulation is perceptually displaced to the visual stimuli, giving rise to the illusion of a puppet speaking. The illusion gives the impression that information from different senses are merging, while in reality, there is a spatial disparity between the two stimulation sources. Therefore, it has been proposed that when two stimulation give rise to the perception of a common shared source, this influences the interaction between the sensory processing (Tsilionis & Vatakis, 2016).

Behaviourally, it has been shown that this semantic congruence is necessary for manifestations of multisensory integration, as reflected in audio-

visual discrimination responses (Laurienti, Kraft, Maldjian, Burdette, & Wallace, 2004). Participants were asked to discriminate between red and blue circles. The circles could either be presented alone, or accompanied by the spoken word 'red' or 'blue'. The authors showed that participants were faster when the visual stimulation was matched with the congruent spoken stimuli, demonstrating the effect of semantic congruence in the interaction between the sensory modalities.

Contrary evidence comes from a bimodal TOJ task (Vatakis & Spence, 2007). Participants were asked to judge the temporal order of auditory stimuli, which were paired with congruent or incongruent visual images (i.e. a female voice matched with a female face etc.). Surprisingly, this study shows that participants found it easier to judge the temporal order when the audio-visual pair was incongruent. This demonstrates that the auditory and visual stimuli interact based on their semantic congruence, but that this does not always lead to an advantage.

Interaction between visual and tactile information has been attributed to the fact that participants know that both originate from a common source (Helbig & Ernst, 2007). Indeed, while feeling a rectangular but viewing a square, the final estimation of the shape was biased by the tactile and the visual information the participants received. A subsequent study tested if the effect was indeed due to the perception that both stimulation originated from a common source (Misceo, Jackson, & Perdue, 2014). Participants were asked to estimate the size of a square they were feeling. Simultaneously, participants were looking at a square of half the size. Interestingly, participants perceived the visual and tactile information as having a common source or not (i.e. were under the impression that they were viewing what they were feeling). Results showed that the visual information biases the size estimate of the square, irrespectively of the unity perception. As result, the authors argued that the unity assumption does not influence visuo-tactile

interaction, and that visuo-tactile interaction is mostly driven by bottom-up factors than top-down factors.

While there is evidence that the perception of a common source might have importance in the interaction between the senses, evidence is not totally conclusive for visuo-tactile stimulations. Concerning the present experiments, multiple factors can affect the unified percept principle. First, by using different frequency of stimulation for all three stimulation streams, onset times were unequal for all stimulations, potentially influencing perception. As a consequence, participants maybe did not have the perception that the tactile and visual stimulation shared a common source, leading to segregation instead of integration between the two stimulation streams.

In this part of the discussion we explored bottom-up and top-down factors which could potentially have influenced the absence of positive results in the studies aiming at studying spatial organization of visuo-tactile interactions. While some factors appear to be more probable than others to have contributed to the absence of results, all the above mentioned arguments lead to believe that the paradigm to study multisensory integration could be further improved. However, it is also possible that the processes underlying visuo-tactile interactions using transient stimuli cannot be evidenced using sustained stimuli. In the next part of the thesis, we discuss if the paradigm we used could be less prone to the effect of a crossmodal link in spatial attention. Subsequently, we discuss two theories that integrate spatial attention and multisensory integration.

3. Processes underlying the crossmodal link in spatial attention

It is puzzling that effects of crossmodal spatial attention between touch and vision have been shown (endogenous and exogenous), reflected in ERP, but that this is not replicated using FPVS response. Consequently, it is possible that ERP

and FPVS potentials reflect different processes, or the same process but in a different form. This could then result in a differential modulation of ERPs and FPVS potentials. Conversely, it is possible that different processes underlie crossmodal spatial attention in ERPs and PFVS potentials.

The bilateral frontoparietal network suggested to underlie endogenous spatial attention includes the superior parietal lobule, the intraparietal sulcus and the frontal eye field (Macaluso, 2012; Pessoa, Kastner, & Ungerleider, 2003). Interestingly, this network is activated independent of stimulation modality, suggesting a supramodal control structure of spatial attention (Macaluso & Driver, 2005). Therefore, in the case of crossmodal spatial attention, it can be argued that these supramodal attentional structures operate via the spatiotopic reference frames (Macaluso & Driver, 2005). Consequently, spatial attention and multisensory processing should be closely linked.

Using FPVS, it has been demonstrated that dividing spatial attention is possible and increases the amplitude of the signal, as compared to unattended FPVS (Müller, Malinowski, Gruber, & Hillyard, 2003). This suggests that attentional mechanisms can be divided over space and do not affect processing resources. These findings incite the authors to argue for a dissociation between the principles governing attentional effects involved in sustained and transient stimulations (Müller et al., 2003). Importantly, this implies that asking participants to divide their attention to multiple stimulation trains (as in experiment 1 – experiment 3 in study 1) should not affect the FPVS signal.

Additionally, the link between ERP and FPVS potentials can be questioned. While this is a much debated question, there is evidence that FPVS potentials are a linear superposition of visual ERPs (Capilla et al., 2011). Interestingly, this research question has received much attention in the auditory modality, with changing

results. The authors state that the main reason for divergent results is the estimation method for the ERPs. Additionally, they argue that when using the appropriate method to estimate ERPs, the FPVS signal and the linear superposition of transient ERPs are highly related. While this constitutes compelling evidence for a similarity between ERPs and FPVS signal, one can question whether the ERP components affected by spatial attentions are reflected by their estimation, and consequently reflected in the FPVS response. Additionally, it has been shown that ERP and FPVS responses can be affected differently (Torta et al., 2017). Indeed, while showing a modulation of the ERP related to their experimental procedure, this same modulation could not be demonstrated in the FPVS response. The authors argue that ERP also reflect higher-order processing, which are not reflected in the FPVS response.

Overall, several factors can explain the differential effect of crossmodal spatial attention on ERP and the FPVS response. While the same network is involved, there is evidence that this network is governed by different principles according to the length of stimulation (transient vs. sustained). Additionally, it not guaranteed that components affected by crossmodal spatial attention are reflected in a similar way in the FPVS response.

4. Multisensory integration, crossmodal interaction, and spatial attention

While this thesis attempted to make minimize the effects of spatial attention by using sustained stimulation and dividing attention over all stimulations, it has been argued that those processes interact at different levels and might not be separable (Macaluso, 2012; Talsma, 2015; Talsma et al., 2010). Multisensory integration is the process where the processing of stimulations in different modalities converge, but attention is the process through which this

selection takes place (Macaluso, 2012). Consequently linked, attention and multisensory integration are most likely influencing each other in a bidirectional way. Different frameworks can be considered to model this interaction.

In a first framework, bidirectional interactions between multisensory integration and attention is argued for, acknowledging the independence of multisensory integration in certain cases (i.e. multisensory integration and interaction can occur automatically and preattentively) (Talsma et al., 2010). The authors argue that stimuli from different modalities can automatically be integrated, or interact, when occurring in an environment with low noise (i.e. little competition for processing resources). As such, it has been demonstrated that when presenting a congruent tactile and visual stimulation, this increases processing in the visual cortex (Macaluso, 2000). More recently, it has since been shown that increasing the load of working memory or visuo-spatial attention did not abolish this effect, showing the independence of multisensory integration and other general cognitive effects (Zimmer & Macaluso, 2007).

The framework states that when increasing the noise in the environment or dealing with near threshold stimulations, top-down attention becomes a necessary process to select and attribute processing resources to the relevant stimuli. Conversely, stimuli can grab attention and by this favour multisensory processing. As such, in this framework, the key element for determining the role of top-down and bottom-up attention in multisensory integration is the degree of competition between different stimulations for processing resources.

A second framework that models the interaction between attention and multisensory integration is the predictive coding framework, which states that the brain relies on Bayesian estimates of the environment to form predictions. In this framework, the brain builds internal models of probability for external events

(similar to a Bayesian prior), which influence how the environment is perceived and where attention is directed to. In return, if an event happens that does not match the internal model (i.e. a prediction error), the internal model is than updated in accordance. In this framework, attention is directed in a top-down manner (i.e. depending on expected inputs), but also in a bottom-up manner (i.e. if an unexpected event occurs, this will grab attention to update internal models) (Talsma, 2015). As such, humans are capable of making sense of very noisy information (i.e. talking in a noisy environment, perceiving gestalts, etc.). In this framework, attention is the main mechanisms by which stimuli are selected for further processing based on internal representations and expectations of the environment, as well as the main mechanisms by which these internal representations can be altered by attention-grabbing stimuli.

While the aim of the experiments presented in this thesis was to control the effects of spatial attention by requiring participants to pay attention to a variety of stimulation, and by not performing a spatial task, it is possible that involvement of spatial attention is crucial for visuo-tactile interactions, especially in sustained stimulation paradigms.

5. Future studies

5.1 Improving the FPVS paradigm

While FPS appeared to be a very promising approach to study multisensory integration, and has many advantages such as high signal-to-noise ratio, the possibility to frequency tag, using sustained trains of stimulations and many more, the results of the studies in this thesis demonstrate that fine tuning of experimental parameters is still necessary. For example, the frequency of the FPVS paradigm was chosen based on the time point at which tool and non tool related

ERP begin to differentiate, allowing approx. 250ms between each stimulation (Proverbio et al., 2011). However, this has the drawback that high-level stimulation is also comprised in the base response. To avoid that, the general frequency should increase to rates higher than 10Hz, and images of the critical category could be presented at a lower frequency, to preserve that advantage. This is a point which could be improved in further studies using the FPVS paradigm such as used in study 2 and 3.

5.2 Threatening objects to activate multisensory integration in the peripersonal space?

To further study the spatial organisation of multisensory integration in the peripersonal space, it could be attempted to use images of threatening objects as a critical category in the FPVS paradigm inserted in a stream of non-threatening stimuli. Subsequently, the activity related to threatening stimulation can be attempted to be modulated by the presence of spatially congruent vibrotactile or nociceptive stimulations. As the peripersonal space has been theorized to underlie the defensive system (Graziano & Cooke, 2006), it can be hypothesised that this system is better activated through threatening stimulations as compared to low-level visual stimuli such as LED lights or pictures of tools. Supporting this, it has been shown that the attentional shift to tactile events is larger when the tactile stimulation is preceded by a spatially congruent picture of potential threat as compared to a neutral picture (Poliakoff, Miles, Li, & Blanchette, 2007). However, here again, the effect is more likely to be due to a crossmodal link in spatial attention than pure multisensory integration, as threatening stimulations are naturally more salient. Additionally, it could be an option to introduce dynamic visual stimuli.

5.3 Neural underpinnings of the theory of affordance

5.3.1 FPVS for object orientation

Besides potential future experiments studying visuo-tactile interactions, studies exploring the theory of affordance can be build up relying on the results of study 2. Indeed, in the 2nd study, we separated tool-selective activity from low-level visual activity common to all object categories. We argue that the activation of the two visual streams comes from the theory of affordance, automatically activating motor schemes by observing a tool. Studies have shown that if the object orientation is congruent with the response hand, this leads to a faster reaction times (Tucker & Ellis, 1998). Observing the neural underpinnings of this theory is trickier and again relies on ad hoc subtraction methods. As such, it would be interesting to explore if it is possible to inherently tag the 'pure' affordance, by using tools in a congruent orientation as the category-selective stimuli, intertwined with pictures of tools in different orientations. Participants would be presented with an image of one tool during each trial, but varying in orientation. Ideally, participants would be asked to carry out a motor task, such as responding to colour changes in fixation cross or to a colour change in the object by a key press (collecting reaction times as an additional outcome measure). The presence of activity which would be orientation related would show that there truly is a differential processing for orientation congruent to the hand. To exclude the possibility that this activity is simply related to the periodicity of the orientation, a control condition should be included where an orientation that does not match the hand should be used as critical category. This study would have the ability to study the neural underpinnings of the affordance theory even further, and to access the 'afforded motor scheme' directly through the FPVS paradigm.

5.3.2 FPVS for grip types

The FPVS paradigm could also be used to shed light on the neural underpinnings of a second mechanism related to the visuo-motor transformation. Indeed, it is known that neurons code for specific grasp types that are congruent with the seen object (Murata, Gallese, Luppino, Kaseda, & Sakata, 2000; Murata et al., 2016). Behaviourally, Tucker & Ellis (2001) showed that participants were faster when responding when the grip type was congruent with the observed object, even though the task demand was not related to the afforded grip types. This implies that affordances go as far as grip types. Building a FPVS sequence using grip types could resolve the issue of the underlying neural mechanisms. Indeed, using pictures of objects which e.g. afford a power grip as critical category and pictures of tools affording different grip types, this experiment has the potential to demonstrate that EEG is sensitive to pick up how different grip types are coded for at neural level.

6. Conclusion

This thesis aimed at studying the interaction between vision and touch, using multiple sustained trains of stimuli simultaneously. In a first part, we aimed at modulating the visual processing by a spatially congruent vibrotactile stimulation. In these three experiments, we varied factors such as task demand and spatial attention. In a second part, we designed a FPVS paradigm to separate tool-selective activity from more general low-level visual activity. In a last experiment, we aimed at modulated this tool-selective activity by a spatially congruent vibrotactile stimulation. As such, we could study the influence of a spatially congruent vibrotactile stimulation on simple, more complex and tool related visual processing.

However, throughout all experiments aiming at showing an effect of spatially congruent vibrotactile stimulation, all results showed up negative. Accordingly, we can conclude that having a spatially congruent visual and vibrotactile stimulation does not alter the visual processing, as compared to unimodal visual processing. This conclusion however only holds using the exact same paradigm with exactly the same experimental factors and details. These results could suggest several things. First, it could be that while spatial congruence is a factor influencing ERPs, the FPVS response is less or not sensible to these mechanisms. Additionally, some experimental factors could have influenced the results. Factors such as the type of stimuli, the frequency, the number of stimuli, the duration, and the use of static stimuli, a potentially dominant visual processing, as well as the task demand or perception are worth taking a closer look to in future studies with a similar topic.

REFERENCES

Adamo, M., & Ferber, S. (2009). A picture says more than a thousand words: Behavioural and ERP evidence for attentional enhancements due to action affordances. *Neuropsychologia*, *47*(6), 1600–1608. http://doi.org/10.1016/j.neuropsychologia.2008.07.009

Alais, D., & Burr, D. (2004). Ventriloquist Effect Results from Near-Optimal Bimodal Integration. *Current Biology*, *14*(3), 257–262. http://doi.org/10.1016/S0960-9822(04)00043-0

Ales, J. M., Farzin, F., Rossion, B., & Norcia, A. M. (2012). An objective method for measuring face detection thresholds using the sweep steady-state visual evoked response. *Journal of Vision*, *12*(2012), 1–18. http://doi.org/10.1167/12.10.18.Introduction

Allman, B. L., Keniston, L. P., & Meredith, M. A. (2008). Subthreshold auditory inputs to extrastriate visual neurons are responsive to parametric changes in stimulus quality: Sensory-specific versus non-specific coding. *Brain Research*, *1242*, 95–101. http://doi.org/10.1016/j.brainres.2008.03.086

Alonso-Prieto, E., Belle, G. Van, Liu-Shuang, J., Norcia, A. M., & Rossion, B. (2013). The 6Hz fundamental stimulation frequency rate for individual face discrimination in the right occipito-temporal cortex. *Neuropsychologia*, *51*(13), 2863–2875. http://doi.org/10.1016/j.neuropsychologia.2013.08.018

Anderson, S. J., Yamagishi, N., & Karavia, V. (2002). Attentional processes link perception and action. *Proceedings of the Royal Society B: Biological Sciences*, *269*(1497), 1225–1232. http://doi.org/10.1098/rspb.2002.1998

Appelbaum, L. G., Wade, A. R., Pettet, M. W., Vildavski, V. Y., & Norcia, A. M. (2008). Figure – ground interaction in the human visual cortex. *Journal of Vision*, *8*(9), 1–19. http://doi.org/10.1167/8.9.8.Introduction

Arbib, M. A., Bonaiuto, J. B., Jacobs, S., & Frey, S. H. (2009). Tool use and the distalization of the end-effector. *Psychological Research*, *73*(4), 441–462. http://doi.org/10.1007/s00426-009-0242-2

Arons, B. (1992). A Review of The Cocktail Party Effect. *Journal of the American Voice I/O Society*, *12*, 35–50. http://doi.org/10.1.1.30.7556

Askenasy, J., & Lehmann, J. (2013). Consciousness, brain, neuroplasticity. *Frontiers in Psychology*, *4*, 1–10. http://doi.org/10.3389/fpsyg.2013.00412

Avillac, M., Ben Hamed, S., & Duhamel, J. R. (2007). Multisensory integration in the ventral intraparietal area of the macaque monkey. *The Journal of Neuroscience*, *27*(8), 1922–1932. http://doi.org/10.1523/JNEUROSCI.2646-06.2007

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

Azañón, E., Stenner, M. P., Cardini, F., & Haggard, P. (2015). Dynamic tuning of tactile localization to body posture. *Current Biology*, *25*(4), 512–517. http://doi.org/10.1016/j.cub.2014.12.038

Bach, M., & Meigen, T. (1999). Do's and don'ts in Fourier analysis of steady-state potentials. *Documenta Ophthalmologica*, *99*(1), 69–82. http://doi.org/10.1023/A:1002648202420

Barth, D. S., Goldberg, N., Brett, B., & Di, S. (1995). The spatiotemporal organization of auditory, visual, and auditory-visual evoked potentials in rat cortex. *Brain Research*, *678*(1–2), 177–190. http://doi.org/10.1016/0006-8993(95)00182-P

Berger, T. D., Martelli, M., & Pelli, D. G. (2003). Flicker flutter: is an illusory event as good as the real thing? *Journal of Vision*, *3*(6), 406–12. http://doi.org/10:1167/3.6.1

Bertelson, P., & De Gelder, B. (2004). The psychology of multimodal perception. In C. Spence & J. Driver (Eds.), *Crossmodal space and crossmodal attention* (pp. 141–177). Oxford: Oxford University Press.

Bertelson, P., & Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Perception & Psychophysics*, *29*(6), 578–584. http://doi.org/10.3758/BF03207374

Bertelson, P., Vroomen, J., De Gelder, B., & Driver, J. (2000). The ventriloquist effect does not depend on the direction of deliberate visual attention. *Perception & Psychophysics*, *62*(2), 321–332. http://doi.org/10.3758/BF03205552

Besle, J., Fort, A., & Giard, M.-H. (2004). Interest and validity of the additive model in electrophysiological studies of multisensory interactions. *Cognitive Processing*, *5*(3). http://doi.org/10.1007/s10339-004-0026-y

Bohórquez, J., & Özdamar, Ö. (2008). Generation of the 40-Hz auditory steady-state response (ASSR) explained using convolution. *Clinical Neurophysiology*, *119*(11), 2598–2607. http://doi.org/10.1016/j.clinph.2008.08.002

Boremanse, A., Norcia, A. M., & Rossion, B. (2014). Dissociation of part-based and integrated neural responses to faces by means of electroencephalographic frequency tagging. *European Journal of Neuroscience*, *40*(6), 2987–2997. http://doi.org/10.1111/ejn.12663

Bremmer, F., Schlack, A., Duhamel, J. R., Graf, W., & Fink, G. R. (2001). Space

coding in primate posterior parietal cortex. *NeuroImage*, *14*(1 II), 46–51. http://doi.org/10.1006/nimg.2001.0817

Bremmer, F., Schlack, A., Shah, N. J., Zafiris, O., Kubischik, M., Hoffmann, K.-P., … Fink, G. R. (2001). Polymodal Motion Processing in Posterior Parietal and Premotor Cortex: A Human fMRI Study Strongly Implies Equivalencies between Humans and Monkeys. *Neuron*, *29*(1), 287–296. http://doi.org/10.1016/S0896-6273(01)00198-2

Brosch, M. (2005). Nonauditory Events of a Behavioral Procedure Activate Auditory Cortex of Highly Trained Monkeys. *The Journal of Neuroscience*, *25*(29), 6797–6806. http://doi.org/10.1523/JNEUROSCI.1571-05.2005

Brozzoli, C., Cardinali, L., Pavani, F., & Farnè, A. (2010). Action-specific remapping of peripersonal space. *Neuropsychologia*, *48*(3), 796–802. http://doi.org/10.1016/j.neuropsychologia.2009.10.009

Brozzoli, C., Demattè, M. L., Pavani, F., Frassinetti, F., & Farnè, A. (2006). Neglect and extinction: within and between sensory modalities. *Restorative Neurology and Neuroscience*, *24*(4–6), 217–232.

Brozzoli, C., Gentile, G., Petkova, V. I., & Ehrsson, H. H. (2011). fMRI Adaptation Reveals a Cortical Mechanism for the Coding of Space Near the Hand. *The Journal of Neuroscience*, *31*(24), 9023–9031. http://doi.org/10.1523/JNEUROSCI.1172-11.2011

Brozzoli, C., Pavani, F., Urquizar, C., Cardinali, L., & Farnè, A. (2009). Grasping actions remap peripersonal space. *NeuroReport*, *20*(10), 913–917. http://doi.org/10.1097/WNR.0b013e32832c0b9b

Buccino, G., Sato, M., Cattaneo, L., Rodà, F., & Riggio, L. (2009). Broken

affordances, broken objects: A TMS study. *Neuropsychologia*, *47*(14), 3074–3078. http://doi.org/10.1016/j.neuropsychologia.2009.07.003

Burr, D., & Alais, D. (2006). Combining visual and auditory information. In S. Martinez-Conde, S. L. Macknik, L. M. Martinez, J.-M. Alonso, & P. U. Tse (Eds.), *Progress in Brain Research* (Vol. 155 B, pp. 243–258). Elsevier B.V. http://doi.org/10.1016/S0079-6123(06)55014-9

Burr, D., Banks, M. S., & Morrone, M. C. (2009). Auditory dominance over vision in the perception of interval duration. *Experimental Brain Research*, *198*(1), 49–57. http://doi.org/10.1007/s00221-009-1933-z

Bushara, K. O., Grafman, J., & Hallett, M. (2001). Neural correlates of auditory-visual stimulus onset asynchrony detection. *The Journal of Neuroscience*, *21*(1), 300–4. http://doi.org/21/1/300

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

Calvert, G. A., Campbell, R., & Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Current Biology*, *10*(11), 649–657. http://doi.org/10.1016/S0960-9822(00)00513-3

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

Canzoneri, E., Magosso, E., & Serino, A. (2012). Dynamic Sounds Capture the Boundaries of Peripersonal Space Representation in Humans. *PLoS ONE*, *7*(9),

3–10. http://doi.org/10.1371/journal.pone.0044306

Capilla, A., Pazo-Alvarez, P., Darriba, A., Campo, P., & Gross, J. (2011). Steady-State Visual Evoked Potentials Can Be Explained by Temporal Superposition of Transient Event-Related Responses. *PLoS ONE*, *6*(1), e14543. http://doi.org/10.1371/journal.pone.0014543

Cappe, C., & Barone, P. (2005). Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. *European Journal of Neuroscience*, *22*(11), 2886–2902. http://doi.org/10.1111/j.1460-9568.2005.04462.x

Caramazza, A., & Hillis, A. E. (1990). Spatial representation of words in the brain implied by studies of a unilateral neglect patient. *Nature*, *346*(6281), 267–9. http://doi.org/10.1038/346267a0

Cardellicchio, P., Sinigaglia, C., & Costantini, M. (2011). The space of affordances: A TMS study. *Neuropsychologia*, *49*(5), 1369–1372. http://doi.org/10.1016/j.neuropsychologia.2011.01.021

Cardinali, L., Brozzoli, C., & Farnè, A. (2009). Peripersonal space and body schema: Two labels for the same concept? *Brain Topography*, *21*(3–4), 252–260. http://doi.org/10.1007/s10548-009-0092-7

Celma-Miralles, A., de Menezes, R. F., & Toro, J. M. (2016). Look at the Beat, Feel the Meter: Top–Down Effects of Meter Induction on Auditory and Visual Modalities. *Frontiers in Human Neuroscience*, *10*(108), 1–13. http://doi.org/10.3389/fnhum.2016.00108

Chao, L. L., Haxby, J. V, & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature*

Neuroscience, *2*(10), 913–919. http://doi.org/10.1038/13217

Chao, L. L., & Martin, A. (2000). Representation of manipulable man-made objects in the dorsal stream. *NeuroImage*, *12*(4), 478–84. http://doi.org/10.1006/nimg.2000.0635

Cheal, M., & Lyon, D. R. (1991). Central and Peripheral Precuing of Forced-Choice Discrimination. *The Quarterly Journal of Experimental Psychology Section A*, *43*(4), 859–880. http://doi.org/10.1080/14640749108400960

Chen, Y.-C., & Spence, C. (2017). Assessing the Role of the 'Unity Assumption' on Multisensory Integration: A Review. *Frontiers in Psychology*, *8*(MAR), 1–22. http://doi.org/10.3389/fpsyg.2017.00445

Chouinard, P. A., & Goodale, M. A. (2012). FMRI-adaptation to highly-rendered color photographs of animals and manipulable artifacts during a classification task. *NeuroImage*, *59*(3), 2941–2951. http://doi.org/10.1016/j.neuroimage.2011.09.073

Cléry, J., Guipponi, O., Wardak, C., & Ben Hamed, S. (2015). Neuronal bases of peripersonal and extrapersonal spaces, their plasticity and their dynamics: Knowns and unknowns. *Neuropsychologia*, *70*, 313–326. http://doi.org/10.1016/j.neuropsychologia.2014.10.022

Colavita, F. B. (1974). Human sensory dominance. *Perception & Psychophysics*, *16*(2), 409–412. http://doi.org/10.3758/BF03203962

Colby, C. (1998). Action-Oriented Spatial Reference Frames in Cortex. *Neuron*, *20*(1), 15–24. http://doi.org/10.1016/S0896-6273(00)80429-8

Colby, C., & Duhamel, J. R. (1996). Spatial representations for action in parietal cortex. *Cognitive Brain Research*, *5*(1–2), 105–115.

http://doi.org/10.1016/S0926-6410(96)00046-8

Colby, C., Duhamel, J. R., & Goldberg, M. E. (1993). Ventral intraparietal area of the macaque: anatomic location and visual response properties. *Journal of Neurophysiology*, *69*(3), 902–914. http://doi.org/10.1152/jn.1993.69.3.902

Colby, C., & Goldberg, M. E. M. (1999). Space and attention in parietal cortex. *Annual Review of Neuroscience*, *22*, 319–349. http://doi.org/10.1146/annurev.neuro.22.1.319

Colon, E., Legrain, V., Huang, G., & Mouraux, A. (2015). Frequency tagging of steady-state evoked potentials to explore the crossmodal links in spatial attention between vision and touch. *Psychophysiology*, *52*(11), 1498–1510. http://doi.org/10.1111/psyp.12511

Colon, E., Legrain, V., & Mouraux, A. (2012). Steady-state evoked potentials to study the processing of tactile and nociceptive somatosensory input in the human brain. *Neurophysiologie Clinique/Clinical Neurophysiology*, *42*(5), 315–323. http://doi.org/10.1016/j.neucli.2012.05.005

Colon, E., Liberati, G., & Mouraux, A. (2017). EEG frequency tagging using ultra-slow periodic heat stimulation of the skin reveals cortical activity specifically related to C fiber thermonociceptors. *NeuroImage*, *146*, 266–274. http://doi.org/10.1016/j.neuroimage.2016.11.045

Colon, E., Nozaradan, S., Legrain, V., & Mouraux, A. (2012). Steady-state evoked potentials to tag specific components of nociceptive cortical processing. *NeuroImage*, *60*(1), 571–581. http://doi.org/10.1016/j.neuroimage.2011.12.015

Cook, E. P., & Maunsell, J. H. R. (2002). Dynamics of neuronal responses in

macaque MT and VIP during motion detection. *Nature Neuroscience*, *5*(10), 985–994. http://doi.org/10.1038/nn924

Costantini, M., Ambrosini, E., Tieri, G., Sinigaglia, C., & Committeri, G. (2010). Where does an object trigger an action? An investigation about affordances in space. *Experimental Brain Research*, *207*(1–2), 95–103. http://doi.org/10.1007/s00221-010-2435-8

Cowey, A., Small, M., & Ellis, S. (1994). Left visuo-spatial neglect can be worse in far than in near space. *Neuropsychologia*, *32*(9), 1059–1066. http://doi.org/10.1016/0028-3932(94)90152-X

Creem-Regehr, S. H., & Lee, J. N. (2005). Neural representations of graspable objects: Are tools special? *Cognitive Brain Research*, *22*(3), 457–469. http://doi.org/10.1016/j.cogbrainres.2004.10.006

De Gelder, B. (2000). More to seeing than meets the eye. *Science*, *289*(5482), 1148–1149.

De Paepe, A. L., Crombez, G., & Legrain, V. (2016). What's coming near? The influence of dynamical visual stimuli on nociceptive processing. *PLoS ONE*, *11*(5), 1–17. http://doi.org/10.1371/journal.pone.0155864

di Pellegrino, G., & Làdavas, E. (2015). Peripersonal space in the brain. *Neuropsychologia*, *66*, 126–133. http://doi.org/10.1016/j.neuropsychologia.2014.11.011

di Pellegrino, G., Làdavas, E., & Farnè, A. (1997). Seeing where your hands are. *Nature*, *388*(6644), 730.

Di Russo, F., Pitzalis, S., Aprile, T., Spitoni, G., Patria, F., Stella, A., … Hillyard, S. A. (2007). Spatiotemporal analysis of the cortical sources of the steady-state

visual evoked potential. *Human Brain Mapping*, *28*(4), 323–334. http://doi.org/10.1002/hbm.20276

Di Russo, F., & Spinelli, D. (2002). Effects of sustained, voluntary attention on amplitude and latency of steady-state visual evoked potential: A costs and benefits analysis. *Clinical Neurophysiology*, *113*(11), 1771–1777. http://doi.org/10.1016/S1388-2457(02)00262-6

Diederich, A., Colonius, H., Bockhorst, D., & Tabeling, S. (2003). Visual-tactile spatial interaction in saccade generation. *Experimental Brain Research*, *148*(3), 328–337. http://doi.org/10.1007/s00221-002-1302-7

Ding, J., Sperling, G., & Srinivasan, R. (2005). Attentional Modulation of SSVEP Power Depends on the Network Tagged by the Flicker Frequency. *Cerebral Cortex*, *16*(7), 1016–1029. http://doi.org/10.1093/cercor/bhj044

Doyle, M. C., & Snowden, R. J. (2001). Identification of visual stimuli is improved by accompanying auditory stimuli: The role of eye movements and sound location. *Perception*, *30*(7), 795–810. http://doi.org/10.1068/p3126

Driver, J. (1996). Enhancement of selective listening by illusory mislocation of speech sounds due to lip-reading. *Nature*, *381*(6577), 66–68. http://doi.org/10.1038/381066a0

Duhamel, J.-R. (2002). Multisensory Integration in Cortex. *Neuron*, *34*(4), 493–495. http://doi.org/10.1016/S0896-6273(02)00709-2

Duhamel, J.-R., Bremmer, F., BenHamed, S., & Graf, W. (1997). Spatial invariance of visual receptive fields in parietal cortex neurons. *Nature*, *389*(6653), 845–848. http://doi.org/10.1038/39865

Duhamel, J.-R., Colby, C., & Goldberg, M. E. (1998). Ventral Intraparietal Area of

the Macaque: Congruent Visual and Somatic Response Properties. *Journal of Neurophysiology*, *79*(1), 126–136. http://doi.org/10.1152/jn.1998.79.1.126

Dzhelyova, M., Jacques, C., & Rossion, B. (2017). At a single glance: Fast periodic visual stimulation uncovers the spatio-temporal dynamics of brief facial expression changes in the human brain. *Cerebral Cortex*, *27*(8), 4106–4123. http://doi.org/10.1093/cercor/bhw223

Eimer, M. (2004). Multisensory Integration: How Visual Experience Shapes Spatial Perception. *Current Biology*, *14*(3), R115–R117. http://doi.org/10.1016/S0960-9822(04)00033-8

Eimer, M., Cockburn, D., Smedley, B., & Driver, J. (2001). Cross-modal links in endogenous spatial attention are mediated by common external locations: evidence from event-related brain potentials. *Experimental Brain Research*, *139*(4), 398–411. http://doi.org/10.1007/s002210100773

Eimer, M., & Driver, J. (2000). An event-related brain potential study of cross-modal links in spatial attention between vision and touch. *Psychophysiology*, *37*(5), 697–705. http://doi.org/10.1111/1469-8986.3750697

Eimer, M., & Driver, J. (2001). Crossmodal links in endogenous and exogenous spatial attention: evidence from event-related brain potential studies. *Neuroscience and Biobehavioral Reviews*, *25*(6), 497–511. http://doi.org/S0149-7634(01)00029-X [pii]

Eimer, M., Velzen, J. van, & Driver, J. (2002). Cross-Modal Interactions between Audition, Touch, and Vision in Endogenous Spatial Attention: ERP Evidence on Preparatory States and Sensory Modulations. *Journal of Cognitive Neuroscience*, *14*(2), 254–271. http://doi.org/10.1162/089892902317236885

Engel, A. K., Senkowski, D., & Schneider, T. R. (2012). Multisensory Integration through Neural Coherence. In M. M. Murray & M. T. Wallace (Eds.), *The Neural Bases of Multisensory Processes* (pp. 115–130). Boca Raton (FL): CRC Press/Taylor & Francis.

Fabre-Thorpe, M. (2011). The characteristics and limits of rapid visual categorization. *Frontiers in Psychology*, *2*(243), 1–12. http://doi.org/10.3389/fpsyg.2011.00243

Farnè, A., Demattè, M. L., & Làdavas, E. (2005). Neuropsychological evidence of modular organization of the near peripersonal space. *Neurology*, *65*(11), 1754–1758. http://doi.org/10.1212/01.wnl.0000187121.30480.09

Favril, L., Mouraux, A., Sambo, C. F., & Legrain, V. (2014). Shifting attention between the space of the body and external space: electrophysiological correlates of visual-nociceptive crossmodal spatial attention. *Psychophysiology*, *51*(5), 464–77. http://doi.org/10.1111/psyp.12157

Filbrich, L., Alamia, A., Blandiaux, S., Burns, S., & Legrain, V. (2017). Shaping visual space perception through bodily sensations: Testing the impact of nociceptive stimuli on visual perception in peripersonal space with temporal order judgments. *PLoS ONE*, *12*(8), 1–20. http://doi.org/10.1371/journal.pone.0182634

Filbrich, L., Alamia, A., Burns, S., & Legrain, V. (2017). Orienting attention in visual space by nociceptive stimuli: investigation with a temporal order judgment task based on the adaptive PSI method. *Experimental Brain Research*, *235*(7), 2069–2079. http://doi.org/10.1007/s00221-017-4951-2

Filbrich, L., Halicka, M., Alamia, A., & Legrain, V. (2018). Investigating the spatial characteristics of the crossmodal interaction between nociception and vision

using gaze direction. *Consciousness and Cognition*, *57*(November 2017), 106–115. http://doi.org/10.1016/j.concog.2017.11.011

Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *28*(3), 458–467. http://doi.org/10.1037/0278-7393.28.3.458

Fogassi, L., Gallese, V., di Pellegrino, G., Fadiga, L., Gentilucci, M., Luppino, G., … Rizzolatti, G. (1992). Space coding by premotor cortex. *Experimental Brain Research*, *89*(3), 686–690. http://doi.org/10.1007/BF00229894

Fogassi, L., Gallese, V., Fadiga, L., Luppino, G., Matelli, M., & Rizzolatti, G. (1996). Coding of peripersonal space in inferior premotor cortex (area F4). *Journal of Neurophysiology*, *76*(1), 141–157. http://doi.org/10.1152/jn.1996.76.1.141

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. *Cognitive Brain Research*, *10*(1–2), 77–83. http://doi.org/10.1016/S0926-6410(00)00024-0

Foxe, J. J., & Schroeder, C. E. (2005). The case for feedforward multisensory convergence during early cortical processing. *NeuroReport*, *16*(5), 419–423. http://doi.org/10.1097/00001756-200504040-00001

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, *9*(10), 474–80. http://doi.org/10.1016/j.tics.2005.08.011

Frigo, M., & Johnson, S. G. (1998). FFTW: An adaptive software architecture for the FFT. *Acoustics, Speech and Signal Processing*, *3*, 1381–1384.

Fujisaki, W., Kitazawa, S., & Nishida, S. (2012). Multisensory timing. In B. E. Stein (Ed.), *The new handbook of multisensory processing* (pp. 301–317). Cambridge, MA: MIT Press.

Galambos, R. (1980). Tactile and auditory stimuli repeated at high rates (30-50 per sec) produce similar event related potentials. *Annals of the New York Academy of Sciences*, *338*(1), 722–726.

Gauthier, B., Eger, E., Hesselmann, G., Giraud, A.-L., & Kleinschmidt, A. (2012). Temporal Tuning Properties along the Human Ventral Visual Stream. *The Journal of Neuroscience*, *32*(41), 14433–14441. http://doi.org/10.1523/JNEUROSCI.2467-12.2012

Gentile, G., Petkova, V. I., & Ehrsson, H. H. (2011). Integration of Visual and Tactile Signals From the Hand in the Human Brain: An fMRI Study. *Journal of Neurophysiology*, *105*(2), 910–922. http://doi.org/10.1152/jn.00840.2010

Gentilucci, M., Scandolara, C., Pigarev, I. N., & Rizzolatti, G. (1983). Visual responses in the postarcuate cortex (area 6) of the monkey that are independent of eye position. *Experimental Brain Research*, *50*(2–3), 464–468. http://doi.org/10.1007/BF00239214

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

Giani, A. S., Ortiz, E., Belardinelli, P., Kleiner, M., Preissl, H., & Noppeney, U. (2012). Steady-state responses in MEG demonstrate information integration within but not across the auditory and visual senses. *NeuroImage*, *60*(2), 1478–89. http://doi.org/10.1016/j.neuroimage.2012.01.114

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

Gibson, J. J. (1979). *The Theory of Affordances*. *The Ecological Approach to Visual Perception.* Psychology Press.

Goebel, R., & Van Atteveldt, N. (2009). Multisensory functional magnetic resonance imaging: A future perspective. *Experimental Brain Research*, *198*(2–3), 153–164. http://doi.org/10.1007/s00221-009-1881-7

Gondan, M., & Röder, B. (2006). A new method for detecting interactions between the senses in event-related potentials. *Brain Research*, *1073–1074*(1), 389–397. http://doi.org/10.1016/j.brainres.2005.12.050

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*(1), 20–25. http://doi.org/10.1016/0166-2236(92)90344-8

Grafton, S. T., Fadiga, L., Arbib, M. a, & Rizzolatti, G. (1997). Premotor Cortex Activation during Observation and Naming of Familiar Tools. *NeuroImage*, *6*(4), 231–236. http://doi.org/10.1006/nimg.1997.0293

Gratton, G. (1998). Dealing with artifacts: The EOG contamination of the event-related brain potential. *Behavior Research Methods, Instruments, & Computers*, *30*(1), 44–53. http://doi.org/10.3758/BF03209415

Gray, C. M., König, P., Engel, A. K., & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, *338*(6213), 334–337.

http://doi.org/10.1038/338334a0

Graziano, M. S. A., & Cooke, D. F. (2006). Parieto-frontal interactions, personal space, and defensive behavior. *Neuropsychologia*, *44*(13), 2621–2635. http://doi.org/10.1016/j.neuropsychologia.2005.09.011

Graziano, M. S. A., & Gross, C. G. (1995). The representation of extrapersonal space: A possible role for bimodal, visual-tactile neurons. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 1021–1034). Cambridge, MA: MIT Press.

Graziano, M. S. A., Gross, C. G., Taylor, C. S. R., & Moore, T. (2014). A System of Multimodal Areas in the Primate Brain. In C. Spence & J. Driver (Eds.), *Crossmodal space and crossmodal attention* (pp. 51–67). Oxford: Oxford University Press. http://doi.org/10.1016/S0896-6273(01)00174-X

Graziano, M. S. A., Hu, X. T., & Gross, C. G. (1997). Visuospatial Properties of Ventral Premotor Cortex. *Journal of Neurophysiology*, *77*(5), 2268–2292. http://doi.org/10.1152/jn.1997.77.5.2268

Graziano, M. S. A., Taylor, C. S. R., & Moore, T. (2002). Complex Movements Evoked by Microstimulation of Precentral Cortex. *Neuron*, *34*(5), 841–851. http://doi.org/10.1016/S0896-6273(02)00698-0

Graziano, M. S. A., Yap, G., & Gross, C. (1994). Coding of visual space by premotor neurons. *Science*, *266*(5187), 1054–1057. http://doi.org/10.1126/science.7973661

Greenwood, P. M., & Goff, W. R. (1987). Modification of median nerve somatic evoked potentials by prior median nerve, peroneal nerve, and auditory stimulation. *Electroencephalography and Clinical Neurophysiology/Evoked*

Potentials Section, *68*(4), 295–302. http://doi.org/10.1016/0168-5597(87)90050-5

Grèzes, J., Armony, J. L., Rowe, J., & Passingham, R. E. (2003). Activations related to "mirror" and "canonical" neurones in the human brain: An fMRI study. *NeuroImage*, *18*(4), 928–937. http://doi.org/10.1016/S1053-8119(03)00042-9

Gross, C. G., & Graziano, M. S. A. (1995). REVIEW : Multiple Representations of Space in the Brain. *The Neuroscientist*, *1*(1), 43–50. http://doi.org/10.1177/107385849500100107

Guariglia, C., & Antonucci, G. (1992). Personal and extrapersonal space: A case of neglect dissociation. *Neuropsychologia*, *30*(11), 1001–1009. http://doi.org/10.1016/0028-3932(92)90051-M

Guillaume, M., Mejias, S., Rossion, B., Dzhelyova, M., & Schiltz, C. (2018). A rapid, objective and implicit measure of visual quantity discrimination. *Neuropsychologia*, *111*, 180–189. http://doi.org/10.1016/j.neuropsychologia.2018.01.044

Haaland, K. Y., Harrington, D. L., & Knight, R. T. (2000). Neural representations of skilled movement. *Brain*, *123*(11), 2306–2313. http://doi.org/10.1093/brain/123.11.2306

Halligan, P. W., Fink, G. R., Marshall, J. C., & Vallar, G. (2003). Spatial cognition: Evidence from visual neglect. *Trends in Cognitive Sciences*, *7*(3), 125–133. http://doi.org/10.1016/S1364-6613(03)00032-9

Halligan, P. W., & Marshall, J. C. (1991). Left neglect for near but not far space in man. *Nature*, *350*(6318), 498–500. http://doi.org/10.1038/350498a0

Handy, T. C., Grafton, S. T., Shroff, N. M., Ketay, S., & Gazzaniga, M. S. (2003). Graspable objects grab attention when the potential for action is recognized. *Nature Neuroscience*, *6*(4), 421–427. http://doi.org/10.1038/nn1031

Hartcher-O'Brien, J., Gallace, A., Krings, B., Koppen, C., & Spence, C. (2008). When vision "extinguishes" touch in neurologically-normal people: Extending the Colavita visual dominance effect. *Experimental Brain Research*, *186*(4), 643–658. http://doi.org/10.1007/s00221-008-1272-5

Hay, I. S., & Davis, H. (1971). Slow Cortical Evoked Potentials: Interactions of Auditory, Vibro-Tactile and Shock Stimuli. *International Journal of Audiology*, *10*(1), 9–17. http://doi.org/10.3109/00206097109072535

Heed, T., Buchholz, V. N., Engel, A. K., & Röder, B. (2015). Tactile remapping: From coordinate transformation to integration in sensorimotor processing. *Trends in Cognitive Sciences*, *19*(5), 251–258. http://doi.org/10.1016/j.tics.2015.03.001

Helbig, H. B., & Ernst, M. O. (2007). Knowledge about a common source can promote visual-haptic integration. *Perception*, *36*(10), 1523–1533. http://doi.org/10.1068/p5851

Herrmann, C. S. (2001). Human EEG responses to 1-100Hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. *Experimental Brain Research*, *137*(3–4), 346–353. http://doi.org/10.1007/s002210100682

Hertz, U., & Amedi, A. (2010). Disentangling unisensory and multisensory components in audiovisual integration using a novel multifrequency fMRI spectral analysis. *NeuroImage*, *52*(2), 617–632. http://doi.org/10.1016/j.neuroimage.2010.04.186

Hihara, S., Taoka, M., Tanaka, M., & Iriki, A. (2015). Visual Responsiveness of Neurons in the Secondary Somatosensory Area and its Surrounding Parietal Operculum Regions in Awake Macaque Monkeys. *Cerebral Cortex*, *25*(11), 4535–4550. http://doi.org/10.1093/cercor/bhv095

Hillyard, S. A., & Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(3), 781–787. http://doi.org/10.1073/pnas.95.3.781

Holmes, N. P. (2007). The law of inverse effectiveness in neurons and behaviour: Multisensory integration versus normal variability. *Neuropsychologia*, *45*(14), 3340–3345. http://doi.org/10.1016/j.neuropsychologia.2007.05.025

Holmes, N. P., & Spence, C. (2005). Multisensory Integration: Space, Time and Superadditivity. *Current Biology*, *15*(18), R762–R764. http://doi.org/10.1016/j.cub.2005.08.058

Huang, R.-S., Chen, C., & Sereno, M. I. (2018). Spatiotemporal integration of looming visual and tactile stimuli near the face. *Human Brain Mapping*, *39*(5), 2156–2176. http://doi.org/10.1002/hbm.23995

Humphreys, G. W., Riddoch, M. J., Forti, S., & Ackroyd, K. (2004). Action influences spatial perception: Neuropsychological evidence. *Visual Cognition*, *11*(2–3), 401–427. http://doi.org/10.1080/13506280344000310

Hutchison, R. M., Culham, J. C., Everling, S., Flanagan, J. R., & Gallivan, J. P. (2014). Distinct and distributed functional connectivity patterns across cortex reflect the domain-specific constraints of object, face, scene, body, and tool category-selective modules in the ventral visual pathway. *NeuroImage*, *96*, 216–236. http://doi.org/10.1016/j.neuroimage.2014.03.068

Hyvärinen, A., & Oja, E. (2000). Independent component analysis : algorithms and applications. *Neural Networks*, *13*, 411–430.

Hyvärinen, J. (1981). Regional distribution of functions in parietal association area 7 of the monkey. *Brain Research*, *206*(2), 287–303. http://doi.org/10.1016/0006-8993(81)90533-3

Hyvärinen, J., & Poranen, A. (1974). Function of the Parietal Associative Area 7 as Revealed from Cellular Discharges in Alert Monkeys. *Brain*, *97*, 673–692. http://doi.org/10.1093/brain/97.4.673

Innes-Brown, H., & Crewther, D. (2009). The impact of spatial incongruence on an auditory-visual illusion. *PLoS ONE*, *4*(7). http://doi.org/10.1371/journal.pone.0006450

Ishai, A., Ungerleider, L. G., Martin, A., & Haxby, J. V. (2000). The Representation of Objects in the Human Occipital and Temporal Cortex. *Journal of Cognitive Neuroscience*, *12*(supplement 2), 35–51. http://doi.org/10.1162/089892900564055

Jeannerod, M., Arbib, M. A., Rizzolatti, G., & Sakata, H. (1995). Grasping objects: the cortical mechanisms of visuomotor transformation. *Trends in Neurosciences*, *18*(7), 314–320. http://doi.org/10.1016/0166-2236(95)93921-J

Johnson-Frey, S. H. (2004). The neural bases of complex tool use in humans. *Trends in Cognitive Sciences*, *8*(2), 71–78. http://doi.org/10.1016/j.tics.2003.12.002

Kaiser, J., Hertrich, I., Ackermann, H., & Lutzenberger, W. (2006). Gamma-band activity over early sensory areas predicts detection of changes in audiovisual

speech stimuli. *NeuroImage*, *30*(4), 1376–1382. http://doi.org/10.1016/j.neuroimage.2005.10.042

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

Keetels, M., & Vroomen, J. (2008). Tactile-visual temporal ventriloquism: No effect of spatial disparity. *Perception and Psychophysics*, *70*(5), 765–771. http://doi.org/10.3758/PP.70.5.765

Keitel, C., Quigley, C., & Ruhnau, P. (2014). Stimulus-Driven Brain Oscillations in the Alpha Range: Entrainment of Intrinsic Rhythms or Frequency-Following Response? *The Journal of Neuroscience*, *34*(31), 10137–10140. http://doi.org/10.1523/JNEUROSCI.1904-14.2014

Kennett, S., Eimer, M., Spence, C., & Driver, J. (2001). Tactile-Visual Links in Exogenous Spatial Attention under Different Postures: Convergent Evidence from Psychophysics and ERPs. *Journal of Cognitive Neuroscience*, *13*(4), 462–478. http://doi.org/10.1162/08989290152001899

Kennett, S., Spence, C., & Driver, J. (2002). Visuo-tactile links in covert exogenous spatial attention remap across changes in unseen hand posture. *Perception & Psychophysics*, *64*(7), 1083–1094. http://doi.org/10.3758/BF03194758

Kennett, S., Taylor-Clarke, M., & Haggard, P. (2001). Noninformative vision improves the spatial resolution of touch in humans. *Current Biology*, *11*(15), 1188–1191. http://doi.org/10.1016/S0960-9822(01)00327-X

Kerkhoff, G. (2001). Spatial hemineglect in humans. *Progress in Neurobiology*, *63*(1), 1–27. http://doi.org/10.1016/S0301-0082(00)00028-9

Kinsbourne, M. (1970). The cerebral basis of lateral asymmetries in attention. *Acta Psychologica*, *33*, 193–201. http://doi.org/10.1016/0001-6918(70)90132-0

Klucharev, V., Möttönen, R., & Sams, M. (2003). Electrophysiological indicators of phonetic and non-phonetic multisensory interactions during audiovisual speech perception. *Cognitive Brain Research*, *18*(1), 65–75. http://doi.org/10.1016/j.cogbrainres.2003.09.004

Koenig-Robert, R., & VanRullen, R. (2013). SWIFT: A novel method to track the neural correlates of recognition. *NeuroImage*, *81*, 273–282. http://doi.org/10.1016/j.neuroimage.2013.04.116

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

Làdavas, E., di Pellegrino, G., Farnè, A., Zeloni, G., Pellegrino, G. di, Farnè, A., & Zeloni, G. (1998). Neuropsychological Evidence of an Integrated Visuotactile Representation of Peripersonal Space in Humans. *Journal of Cognitive Neuroscience*, *10*(5), 581–589. http://doi.org/10.1162/089892998562988

Ladavas, E., Zeloni, G., & Farnè, A. (1998). Visual peripersonal space centred on the face in humans. *Brain*, *121*(12), 2317–2326. http://doi.org/10.1093/brain/121.12.2317

Lakatos, P., Chen, C. M., O'Connell, M. N., Mills, A., & Schroeder, C. E. (2007). Neuronal Oscillations and Multisensory Interaction in Primary Auditory Cortex. *Neuron*, *53*(2), 279–292. http://doi.org/10.1016/j.neuron.2006.12.011

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of Neuronal Oscillations as a Mechanism of Attentional

Selection. *Science*, *320*(5872), 110–113. http://doi.org/10.1126/science.1154735

Lakatos, P., O'Connell, M. N., Barczak, A., Mills, A., Javitt, D. C., & Schroeder, C. E. (2009). The Leading Sense: Supramodal Control of Neurophysiological Context by Attention. *Neuron*, *64*(3), 419–430. http://doi.org/10.1016/j.neuron.2009.10.014

Lam, K., Kakigi, R., Kaneoke, Y., Naka, D., Maeda, K., & Suzuki, H. (1999). Effects of visual and auditory stimulation on somatosensory evoked magnetic fields. *Clinical Neurophysiology*, *110*(2), 295–304. http://doi.org/10.1016/S0168-5597(98)00059-8

Landis, T. (2000). Disruption of space perception due to cortical lesions. *Spatial Vision*, *13*(23), 179–191. http://doi.org/10.1163/156856800741199

Laurienti, P. J., Kraft, R. A., Maldjian, J. A., Burdette, J. H., & Wallace, M. T. (2004). Semantic congruence is a critical factor in multisensory behavioral performance. *Experimental Brain Research*, *158*(4), 405–414. http://doi.org/10.1007/s00221-004-1913-2

Laurienti, P. J., Perrault, T. J., Stanford, T. R., Wallace, M. T., & Stein, B. E. (2005). On the use of superadditivity as a metric for characterizing multisensory integration in functional neuroimaging studies. *Experimental Brain Research*, *166*(3–4), 289–297. http://doi.org/10.1007/s00221-005-2370-2

Legrain, V., & Torta, D. M. E. (2015). Cognitive Psychology and Neuropsychology of Nociception and Pain. In G. Pickering & S. Gibson (Eds.), *Pain, Emotion and Cognition: A Complex Nexus* (pp. 3–20). Springer. http://doi.org/10.1007/978-3-319-12033-1

Leinonen, L., Hyvärinen, J., Nyman, G., & Linnankoski, I. (1979). I. Functional properties of neurons in lateral part of associative area 7 in awake monkeys. *Experimental Brain Research*, *34*(2), 299–320. http://doi.org/10.1007/BF00235675

Leinonen, L., & Nyman, G. (1979). II. Functional properties of cells in anterolateral part of area 7 associative face area of awake monkeys. *Experimental Brain Research*, *34*(2), 321–333. http://doi.org/10.1007/BF00235676

Lewis, J. W. (2006). Cortical networks related to human use of tools. *Neuroscientist*, *12*(3), 211–231. http://doi.org/10.1177/1073858406288327

Lithari, C., Sánchez-Garciá, C., Ruhnau, P., & Weisz, N. (2016). Large-scale network-level processes during entrainment. *Brain Research*, *1635*, 143–152. http://doi.org/10.1016/j.brainres.2016.01.043

Liu-Shuang, J., Norcia, A. M., & Rossion, B. (2014). An objective index of individual face discrimination in the right occipito-temporal cortex by means of fast periodic oddball stimulation. *Neuropsychologia*, *52*(1), 57–72. http://doi.org/10.1016/j.neuropsychologia.2013.10.022

Lochy, A., Van Belle, G., & Rossion, B. (2015). A robust index of lexical representation in the left occipito-temporal cortex as evidenced by EEG responses to fast periodic visual stimulation. *Neuropsychologia*, *66*, 18–31. http://doi.org/10.1016/j.neuropsychologia.2014.11.007

Lochy, A., Van Reybroeck, M., & Rossion, B. (2016). Left cortical specialization for visual letter strings predicts rudimentary knowledge of letter-sound association in preschoolers. *Proceedings of the National Academy of Sciences*, *113*(30), 8544–8549. http://doi.org/10.1073/pnas.1520366113

Lochy, A., Zimmermann, F. G. S., Laguesse, R., Willenbockel, V., Rossion, B., & Vuong, Q. C. (2018). Does Extensive Training at Individuating Novel Objects in Adulthood Lead to Visual Expertise? The Role of Facelikeness. *Journal of Cognitive Neuroscience*, *30*(4), 449–467. http://doi.org/10.1162/jocn_a_01212

Macaluso, E. (2000). Modulation of Human Visual Cortex by Crossmodal Spatial Attention. *Science*, *289*(5482), 1206–1208. http://doi.org/10.1126/science.289.5482.1206

Macaluso, E. (2012). Spatial constraints in multisensory attention. In M. M. Murray & M. T. Wallace (Eds.), *The Neural Bases of Multisensory Processes* (pp. 1–18). Boca Raton (FL): CRC Press/Taylor & Francis.

Macaluso, E., & Driver, J. (2001). Spatial attention and crossmodal interactions between vision and touch. *Neuropsychologia*, *39*(12), 1304–1316. http://doi.org/10.1016/S0028-3932(01)00119-1

Macaluso, E., & Driver, J. (2005). Multisensory spatial interactions: a window onto functional integration in the human brain. *Trends in Neurosciences*, *28*(5), 264–271. http://doi.org/10.1016/j.tins.2005.03.008

Macaluso, E., Driver, J., van Velzen, J., & Eimer, M. (2005). Influence of gaze direction on crossmodal modulation of visual ERPS by endogenous tactile spatial attention. *Cognitive Brain Research*, *23*(2–3), 406–417. http://doi.org/10.1016/j.cogbrainres.2004.11.003

Macaluso, E., Frith, C. D., & Driver, J. (2000). Selective spatial attention in vision and touch: unimodal and multimodal mechanisms revealed by PET. *Journal of Neurophysiology*, *83*(5), 3062–3075.

Macaluso, E., Frith, C. D., & Driver, J. (2001). Multimodal mechanisms of attention related to rates of spatial shifting in vision and touch. *Experimental Brain Research*, *137*(3–4), 445–454. http://doi.org/10.1007/s002210000656

Macaluso, E., Frith, C. D., & Driver, J. (2002). Crossmodal spatial influences of touch on extrastriate visual areas take current gaze direction into account. *Neuron*, *34*(0896-6273), 647–658. http://doi.org/10.1016/S0896-6273(02)00678-5

Macaluso, E., & Maravita, A. (2010). The representation of space near the body through touch and vision. *Neuropsychologia*, *48*(3), 782–795. http://doi.org/10.1016/j.neuropsychologia.2009.10.010

Makin, T. R., Holmes, N. P., & Zohary, E. (2007). Is That Near My Hand? Multisensory Representation of Peripersonal Space in Human Intraparietal Sulcus. *The Journal of Neuroscience*, *27*(4), 731–740. http://doi.org/10.1523/JNEUROSCI.3653-06.2007

Malinowski, P., Fuchs, S., & Müller, M. M. (2007). Sustained division of spatial attention to multiple locations within one hemifield. *Neuroscience Letters*, *414*(1), 65–70. http://doi.org/10.1016/j.neulet.2006.12.001

Maravita, A., Spence, C., & Driver, J. (2003). Multisensory integration and the body schema: close to hand and within reach. *Current Biology*, *13*(13), R531–R539. http://doi.org/10.1016/S0960-9822(03)00449-4

Martin, A., Wiggs, C. L., Ungerleider, L. G., & Haxby, J. V. (1996). Neural correlates of category-specific knowledge. *Nature*, *379*, 649–652.

Masson, M. E. J., Bub, D. N., & Breuer, A. T. (2011). Priming of Reach and Grasp Actions by Handled Objects. *Journal of Experimental Psychology: Human*

Perception and Performance, *37*(5), 1470–1484. http://doi.org/10.1037/a0023509

McDonald, J. J., Teder-Sälejärvi, W. ., & Ward, L. M. (2001). Multisensory integration and crossmodal attention effects in the human brain. *Science*, *292*(5523), 1791. http://doi.org/10.1126/science.292.5523.1791a\n292/5523/1791a [pii]

Meigen, T., & Bach, M. (1999). On the statistical significance of electrophysiological steady-state responses. *Documenta Ophthalmologica*, *98*(3), 207–232. http://doi.org/10.1023/A:1002097208337

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

Meredith, M. A., & Stein, B. E. (1996). Spatial determinants of multisensory integration in cat superior colliculus neurons. *Journal of Neurophysiology*, *75*(5), 1843–1857. http://doi.org/10.1152/jn.1996.75.5.1843

Meredith, M., Nemitz, J., & Stein, B. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *The Journal of Neuroscience*, *7*(10), 3215–3229. http://doi.org/10.1523/JNEUROSCI.07-10-03215.1987

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

Miller, M. W., & Vogt, B. A. (1984). Direct connections of rat visual cortex with

sensory, motor, and association cortices. *Journal of Comparative Neurology*, *226*(2), 184–202. http://doi.org/10.1002/cne.902260204

Misceo, G. F., Jackson, S. V. S., & Perdue, J. R. (2014). Again, Knowledge of Common Source Fails to Promote Visual-Haptic Integration. *Perceptual and Motor Skills*, *118*(1), 183–194. http://doi.org/10.2466/24.23.PMS.118k11w0

Molholm, S., Ritter, W., Javitt, D. C., & Foxe, J. J. (2004). Multisensory Visual-Auditory Object Recognition in Humans: A High-density Electrical Mapping Study. *Cerebral Cortex*, *14*(4), 452–465. http://doi.org/10.1093/cercor/bhh007

Morein-Zamir, S., Soto-Faraco, S., & Kingstone, A. (2003). Auditory capture of vision: Examining temporal ventriloquism. *Cognitive Brain Research*, *17*(1), 154–163. http://doi.org/10.1016/S0926-6410(03)00089-2

Morgan, S. T., Hansen, J. C., & Hillyard, S. A. (1996). Selective attention to stimulus location modulates the steady-state visual evoked potential. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(10), 4770–4774. http://doi.org/10.1073/pnas.93.10.4770

Mouraux, A., Diukova, A., Lee, M. C., Wise, R. G., & Iannetti, G. D. (2011). A multisensory investigation of the functional significance of the "pain matrix." *NeuroImage*, *54*(3), 2237–2249. http://doi.org/10.1016/j.neuroimage.2010.09.084

Mouraux, A., Iannetti, G. D., Colon, E., Nozaradan, S., Legrain, V., & Plaghki, L. (2011). Nociceptive Steady-State Evoked Potentials Elicited by Rapid Periodic Thermal Stimulation of Cutaneous Nociceptors. *The Journal of Neuroscience*, *31*(16), 6079–6087. http://doi.org/10.1523/JNEUROSCI.3977-10.2011

Müller, H. J., & Rabbitt, P. M. A. (1989). Reflexive and voluntary orienting of visual attention: Time course of activation and resistance to interruption. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(2), 315–330. http://doi.org/10.1037/0096-1523.15.2.315

Müller, M. M., & Hillyard, S. A. (2000). Concurrent recording of steady-state and transient event-related potentials as indices of visual-spatial selective attention. *Clinical Neurophysiology*, *111*(9), 1544–1552. http://doi.org/10.1016/S1388-2457(00)00371-0

Müller, M. M., Malinowski, P., Gruber, T., & Hillyard, S. A. (2003). Sustained division of the attentional spotlight. *Nature*, *424*(6946), 309–312. http://doi.org/10.1038/nature02148

Müller, M. M., Picton, T. W., Valdes-Sosa, P., Riera, J., Teder-Sälejärvi, W. A., & Hillyard, S. a. (1998). Effects of spatial selective attention on the steady-state visual evoked potential in the 20–28 Hz range. *Cognitive Brain Research*, *6*(4), 249–261. http://doi.org/10.1016/S0926-6410(97)00036-0

Müller, M. M., Teder-Sälejärvi, W. A., & Hillyard, S. A. (1998). The time course of cortical facilitation during cued shifts of spatial attention. *Nature Neuroscience*, *1*(7), 631–4. http://doi.org/10.1038/2865

Müller, M. M., Teder, W., & Hillyard, S. A. (1997). Magnetoencephalographic recording of steady-state visual evoked cortical activity. *Brain Topography*, *9*(3), 163–8. http://doi.org/10.1007/BF01190385

Munoz, D. P., Pelisson, D., & Guitton, D. (1991). Movement of neural activity on the superior colliculus motor map during gaze shifts. *Science*, *251*(4999), 1358–1360. http://doi.org/10.1126/science.2003221

Murata, A., Fadiga, L., Fogassi, L., Gallese, V., Raos, V., & Rizzolatti, G. (1997). Object Representation in the Ventral Premotor Cortex (Area F5) of the Monkey. *Journal of Neurophysiology*, *78*(4), 2226–2230. http://doi.org/10.1152/jn.1997.78.4.2226

Murata, A., Gallese, V., Luppino, G., Kaseda, M., & Sakata, H. (2000). Selectivity for the Shape, Size, and Orientation of Objects for Grasping in Neurons of Monkey Parietal Area AIP. *Journal of Neurophysiology*, *83*, 2580–2601.

Murata, A., & Ishida, H. (2007). Representation of Bodily Self in the Multimodal Parieto-Premotor Network. In S. Funahashi (Ed.), *Representation and Brain* (pp. 151–176). Tokyo: Springer Japan.

Murata, A., Wen, W., & Asama, H. (2016). The body and objects represented in the ventral stream of the parieto-premotor network. *Neuroscience Research*, *104*, 4–15. http://doi.org/10.1016/j.neures.2015.10.010

Näätänen, R., & Gaillard, A. W. K. (1983). The Orienting Reflex and the N2 Deflection of the Event-Related Potential (ERP). *Advances in Psychology*, *10*(C), 119–141. http://doi.org/10.1016/S0166-4115(08)62036-1

Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research*, *29*(11), 1631–1647. http://doi.org/10.1016/0042-6989(89)90144-2

Ngo, M. K., & Spence, C. (2010). Crossmodal facilitation of masked visual target identification. *Attention, Perception & Psychophysics*, *72*(7), 1938–1947. http://doi.org/10.3758/APP.72.7.1938

Norcia, A. M., Appelbaum, L. G., Ales, J. M., Cottereau, B. R., & Rossion, B. (2015). The steady-state visual evoked potential in vision research: A review. *Journal*

of Vision, *15*(6), 4. http://doi.org/10.1167/15.6.4.doi

Noss, R. S., Boles, C. D., & Yingling, C. D. (1996). Steady-state analysis of somatosensory evoked potentials. *Electroencephalography and Clinical Neurophysiology*, *100*(5), 453–461. http://doi.org/10.1016/0168-5597(96)96011-6

Nozaradan, S. (2014). Exploring how musical rhythm entrains brain activity with electroencephalogram frequency-tagging. *Philosophical Transactions of the Royal Society, Series B*, *369*(November), 20130393.

Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the neuronal entrainment to beat and meter. *The Journal of Neuroscience*, *31*(28), 10234–10240. http://doi.org/10.1523/JNEUROSCI.0411-11.2011

Nozaradan, S., Peretz, I., & Mouraux, A. (2012a). Selective Neuronal Entrainment to the Beat and Meter Embedded in a Musical Rhythm. *The The Journal of Neuroscience of Neuroscience*, *32*(49), 17572–17581. http://doi.org/10.1523/JNEUROSCI.3203-12.2012

Nozaradan, S., Peretz, I., & Mouraux, A. (2012b). Steady-state evoked potentials as an index of multisensory temporal binding. *NeuroImage*, *60*(1), 21–28. http://doi.org/10.1016/j.neuroimage.2011.11.065

Nozaradan, S., Zerouali, Y., Peretz, I., & Mouraux, A. (2015). Capturing with EEG the Neural Entrainment and Coupling Underlying Sensorimotor Synchronization to the Beat. *Cerebral Cortex*, *25*(3), 736–747. http://doi.org/10.1093/cercor/bht261

Park, J. (2016). A neural basis for the visual sense of number and its development: A steady-state visual evoked potential study in children and adults.

Developmental Cognitive Neuroscience. http://doi.org/10.1016/j.dcn.2017.02.011

Pastor, M. a, Artieda, J., Arbizu, J., Valencia, M., & Masdeu, J. C. (2003). Human cerebral activation during steady-state visual-evoked responses. *The Journal of Neuroscience*, *23*(37), 11621–11627. http://doi.org/23/37/11621 [pii]

Pastor, M., Valencia, M., Artieda, J., Alegre, M., & Masdeu, J. (2007). Topography of Cortical Activation Differs for Fundamental and Harmonic Frequencies of the Steady-State Visual-Evoked Responses. An EEG and PET H215O Study. *Cerebral Cortex*, *17*(8), 1899–1905. http://doi.org/10.1093/cercor/bhl098

Peelen, M. V., & Caramazza, A. (2012). Conceptual Object Representations in Human Anterior Temporal Cortex. *The Journal of Neuroscience*, *32*(45), 15728–15736. http://doi.org/10.1523/JNEUROSCI.1953-12.2012

Pessoa, L., Kastner, S., & Ungerleider, L. G. (2003). Neuroimaging Studies of Attention: From Modulation of Sensory Processing to Top-Down Control. *The Journal of Neuroscience*, *23*(10), 3990–3998. http://doi.org/10.1523/JNEUROSCI.23-10-03990.2003

Poliakoff, E., Miles, E., Li, X., & Blanchette, I. (2007). The effect of visual threat on spatial attention to touch. *Cognition*, *102*(3), 405–414. http://doi.org/10.1016/j.cognition.2006.01.006

Porcu, E., Keitel, C., & Müller, M. M. (2013). Concurrent visual and tactile steady-state evoked potentials index allocation of inter-modal attention: A frequency-tagging study. *Neuroscience Letters*, *556*, 113–117. http://doi.org/10.1016/j.neulet.2013.09.068

Previc, F. H. (1998). The neuropsychology of 3-D space. *Psychological Bulletin*,

124(2), 123–164. http://doi.org/10.1037/0033-2909.124.2.123

Proverbio, A. M., Adorni, R., & D'Aniello, G. E. (2011). 250ms to code for action affordance during observation of manipulable objects. *Neuropsychologia*, *49*(9), 2711–2717. http://doi.org/10.1016/j.neuropsychologia.2011.05.019

Quek, G. L., & Rossion, B. (2017). Category-selective human brain processes elicited in fast periodic visual stimulation streams are immune to temporal predictability. *Neuropsychologia*, *104*, 182–200. http://doi.org/10.1016/j.neuropsychologia.2017.08.010

Raij, T., Ahveninen, J., Lin, F., Witzel, T., Jääskeläinen, I. P., Letham, B., … Belliveau, J. W. (2010). Onset timing of cross-sensory activations and multisensory interactions in auditory and visual sensory cortices. *European Journal of Neuroscience*, *31*(10), 1772–1782. http://doi.org/10.1111/j.1460-9568.2010.07213.x

Regan, D. (1966). Some characteristics of average steady-state and transient responses evoked by modulated light. *Electroencephalography and Clinical Neurophysiology*, *20*(3), 238–248. http://doi.org/10.1016/0013-4694(66)90088-5

Regan, D. (1977). Steady-state evoked potentials. *Journal of the Optical Society of America*, *67*(11), 1475. http://doi.org/10.1364/JOSA.67.001475

Regan, D. (1989). *Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine*. New York: Elsevier.

Retter, T. L., & Rossion, B. (2015). Global Shape Information Increases but Color Information Decreases the Composite Face Effect. *Perception*, *44*(5), 511–528. http://doi.org/10.1068/p7826

Rizzolatti, G., Fadiga, L., Fogassi, L., & Gallese, V. (1997). The space around us. *Science*, *277*(5323), 190–1. http://doi.org/10.1126/science.277.5323.190

Rizzolatti, G., Fogassi, L., & Gallese, V. (2002). Motor and cognitive functions of the ventral premotor cortex. *Current Opinion in Neurobiology*, *12*(2), 149–154. http://doi.org/10.1016/S0959-4388(02)00308-2

Rizzolatti, G., Scandolara, C., Matelli, M., & Gentilucci, M. (1981a). Afferent properties of periarcuate neurons in macaque monkeys. I. Somatosensory responses. *Behavioural Brain Research*, *2*(2), 125–146. http://doi.org/10.1016/0166-4328(81)90052-8

Rizzolatti, G., Scandolara, C., Matelli, M., & Gentilucci, M. (1981b). Afferent properties of periarcuate neurons in macaque monkeys. II. Visual responses. *Behavioural Brain Research*, *2*(2), 147–163. http://doi.org/10.1016/0166-4328(81)90053-X

Rossion, B., Torfs, K., Jacques, C., & Liu-Shuang, J. (2015). Fast periodic presentation of natural images reveals a robust face-selective electrophysiological response in the human brain. *Journal of Vision*, *15*(1), 18. http://doi.org/10.1167/15.1.18

Roux, F.-E., Djidjeli, I., & Durand, J.-B. (2018). Functional architecture of the somatosensory homunculus detected by electrostimulation. *The Journal of Physiology*, *596*(5), 941–956. http://doi.org/10.1113/JP275243

Royal, D. W., Carriere, B. N., & Wallace, M. T. (2009). Spatiotemporal architecture of cortical receptive fields and its impact on multisensory interactions. *Experimental Brain Research*, *198*(2–3), 127–136. http://doi.org/10.1007/s00221-009-1772-y

Sambo, C. F., Torta, D. M. E., Gallace, a., Liang, M., Moseley, L. G., & Iannetti, G. D. (2013). The temporal order judgement of tactile and nociceptive stimuli is impaired by crossing the hands over the body midline. *Pain*, *154*(2), 242–247. http://doi.org/10.1016/j.pain.2012.10.010

Schaafsma, S. J., & Duysens, J. (1996). Neurons in the ventral intraparietal area of awake macaque monkey closely resemble neurons in the dorsal part of the medial superior temporal area in their responses to optic flow patterns. *Journal of Neurophysiology*, *76*(6), 4056–68.

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*, *25*(18), 4616–4625. http://doi.org/10.1523/JNEUROSCI.0455-05.2005

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

Senkowski, D., Schneider, T. R., Foxe, J. J., & Engel, A. K. (2008). Crossmodal binding through neural coherence: implications for multisensory processing. *Trends in Neurosciences*, *31*(8), 401–409. http://doi.org/10.1016/j.tins.2008.05.002

Senkowski, D., Talsma, D., Grigutsch, M., Herrmann, C. S., & Woldorff, M. G. (2007). Good times for multisensory integration: Effects of the precision of temporal synchrony as revealed by gamma-band oscillations. *Neuropsychologia*, *45*(3), 561–571. http://doi.org/10.1016/j.neuropsychologia.2006.01.013

Senkowski, D., Talsma, D., Herrmann, C. S., & Woldorff, M. G. (2005). Multisensory

processing and oscillatory gamma responses: effects of spatial selective attention. *Experimental Brain Research*, *166*(3–4), 411–426. http://doi.org/10.1007/s00221-005-2381-z

Serino, A., Farnè, A., Rinaldesi, M. L., Haggard, P., & Làdavas, E. (2007). Can vision of the body ameliorate impaired somatosensory function? *Neuropsychologia*, *45*(5), 1101–1107. http://doi.org/10.1016/j.neuropsychologia.2006.09.013

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

Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research*, *14*(1), 147–152. http://doi.org/10.1016/S0926-6410(02)00069-1

Shams, L., Kamitani, Y., & Shimojo, S. (2004). Modulations of visual perception by sound. In G. A. Calvert, C. Spence, & B. Stein (Eds.), *The handbook of multisensory processes* (pp. 27–33). Cambridge, MA: MIT Press.

Shore, D. I., Spry, E., & Spence, C. (2002). Confusing the mind by crossing the hands. *Cognitive Brain Research*, *14*(1), 153–163. http://doi.org/10.1016/S0926-6410(02)00070-8

Sinnett, S., Soto-Faraco, S., & Spence, C. (2008). The co-occurrence of multisensory competition and facilitation. *Acta Psychologica*, *128*(1), 153–61. http://doi.org/10.1016/j.actpsy.2007.12.002

Smania, N., & Aglioti, S. (1995). Sensory and spatial components of somaesthetic deficits following right brain damage. *Neurology*, *45*(9), 1725–1730. http://doi.org/10.1212/WNL.45.9.1725

Spaak, E., de Lange, F. P., & Jensen, O. (2014). Local Entrainment of Alpha

Oscillations by Visual Stimuli Causes Cyclic Modulation of Perception. *The Journal of Neuroscience*, *34*(10), 3536–3544. http://doi.org/10.1523/JNEUROSCI.4385-13.2014

Spence, C. (2010). Crossmodal spatial attention. *Annals of the New York Academy of Sciences*, *1191*(1), 182–200. http://doi.org/10.1111/j.1749-6632.2010.05440.x

Spence, C. (2013). Just how important is spatial coincidence to multisensory integration? Evaluating the spatial rule. *Annals of the New York Academy of Sciences*, *1296*(1), 31–49. http://doi.org/10.1111/nyas.12121

Spence, C., McDonald, J. J., & Driver, J. (2004). Exogenous spatial cuing studies of human crossmodal attention and multisensory integration. In C. Spence & J. Driver (Eds.), *Crossmodal space and crossmodal attention* (pp. 277–320). Oxford: Oxford University Press.

Spence, C., Parise, C., & Chen, Y.-C. (2012). The Colavita Visual Dominance Effect. In M. M. Murray & M. T. Wallace (Eds.), *The Neural Bases of Multisensory Processes*. Boca Raton (FL): CRC Press/Taylor & Francis.

Spence, C., Senkowski, D., & Röder, B. (2009). Crossmodal processing. *Experimental Brain Research*, *198*(2–3), 107–111. http://doi.org/10.1007/s00221-009-1973-4

Srinivasan, R., Bibi, F. A., & Nunez, P. L. (2006). Steady-state visual evoked potentials: distributed local sources and wave-like dynamics are sentitive to flicker frequency. *Brain Topography*, *75*(5), 1781–1791. http://doi.org/10.1007/s11103-011-9767-z.Plastid

Stein, B. E., Meredith, M. A., & Wallace, M. T. (1993). Responsive neuron and

beyond: multisensory integration in cat and monkey. *Progress in Brain Research*, *95*(C), 79–90. http://doi.org/10.1016/S0079-6123(08)60359-3

Stein, B. E., Scott Huneycutt, W., & Alex Meredith, M. (1988). Neurons and behavior: the same rules of multisensory integration apply. *Brain Research*, *448*(2), 355–358. http://doi.org/10.1016/0006-8993(88)91276-0

Stein, B. E., Stanford, T. R., Wallace, M. T., Vaughan, J. W., & Jiang, W. (2004). Crossmodal spatial interactions in subcortical and cortical circuits. In C. Spence & J. Driver (Eds.), *Crossmodal space and crossmodal attention* (pp. 25–50). Oxford: Oxford University Press.

Stojanoski, B., & Cusack, R. (2014). Time to wave good-bye to phase scrambling: Creating controlled scrambled images using diffeomorphic transformations. *Journal of Vision*, *14*(12), 6–6. http://doi.org/10.1167/14.12.6

Stothart, G., Quadflieg, S., & Milton, A. (2017). A fast and implicit measure of semantic categorisation using steady state visual evoked potentials. *Neuropsychologia*, *102*, 11–18. http://doi.org/10.1016/j.neuropsychologia.2017.05.025

Sutoyo, D., & Srinivasan, R. (2009). Nonlinear SSVEP responses are sensitive to the perceptual binding of visual hemifields during conventional “eye” rivalry and interocular “percept” rivalry. *Brain Research*, *1251*, 245–55. http://doi.org/10.1016/j.brainres.2008.09.086

Talsma, D. (2015). Predictive coding and multisensory integration: an attentional account of the multisensory mind. *Frontiers in Integrative Neuroscience*, *09*(March), 1–13. http://doi.org/10.3389/fnint.2015.00019

Talsma, D., Doty, T. J., & Woldorff, M. G. (2006). Selective Attention and

Audiovisual Integration: Is Attending to Both Modalities a Prerequisite for Early Integration? *Cerebral Cortex*, *17*(3), 679–690. http://doi.org/10.1093/cercor/bhk016

Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences*, *14*(9), 400–410. http://doi.org/10.1016/j.tics.2010.06.008

Talsma, D., Slagter, H. A., Nieuwenhuis, S., Hage, J., & Kok, A. (2005). The orienting of visuospatial attention: An event-related brain potential study. *Cognitive Brain Research*, *25*(1), 117–129. http://doi.org/10.1016/j.cogbrainres.2005.04.013

Teder-Sälejärvi, W. ., McDonald, J. J., Di Russo, F., & Hillyard, S. . (2002). An analysis of audio-visual crossmodal integration by means of event-related potential (ERP) recordings. *Cognitive Brain Research*, *14*(1), 106–114. http://doi.org/10.1016/S0926-6410(02)00065-4

Teder-Sälejärvi, W. A., Russo, F. Di, McDonald, J. J., & Hillyard, S. A. (2005). Effects of Spatial Congruity on Audio-Visual Multimodal Integration. *Journal of Cognitive Neuroscience*, *17*(9), 1396–1409. http://doi.org/10.1162/0898929054985383

Thomson, M. G. A. (1999). Visual coding and the phase structure of natural scenes. *Network: Computation in Neural Systems*, *10*(2), 123–132. http://doi.org/http://dx.doi.org/10.1088/0954-898X_10_2_302

Tootell, R. B. H., Hadjikhani, N. K., Mendola, J. D., Marrett, S., & Dale, A. M. (1998). From retinotopy to recognition. *Trends in Cognitive Sciences*, *2*(5), 174–183. http://doi.org/10.1016/S1364-6613(98)01171-1

Torta, D. M. E., Van Den Broeke, E. N., Filbrich, L., Jacob, B., Lambert, J., & Mouraux, A. (2017). Intense pain influences the cortical processing of visual stimuli projected onto the sensitized skin. *Pain*, *158*(4), 691–697. http://doi.org/10.1097/j.pain.0000000000000816

Tsilionis, E., & Vatakis, A. (2016). Multisensory binding: Is the contribution of synchrony and semantic congruency obligatory? *Current Opinion in Behavioral Sciences*, *8*, 7–13. http://doi.org/10.1016/j.cobeha.2016.01.002

Tucker, M., & Ellis, R. (1998). On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 830–846. http://doi.org/10.1037/0096-1523.24.3.830

Tucker, M., & Ellis, R. (2001). The potentiation of grasp types during visual object categorization. *Visual Cognition*, *8*(6), 769–800. http://doi.org/10.1080/13506280042000144

Ungerleider, L. G., & Haxby, J. V. (1994). "What" and "where" in the human brain. *Current Opinion in Neurobiology*, *4*(2), 157–165. http://doi.org/10.1016/0959-4388(94)90066-3

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In *Analysis of Visual Behavior* (pp. 549–586). http://doi.org/10.2139/ssrn.1353746

Urbach, T. P., & Kutas, M. (2002). The intractability of scaling scalp distributions to infer neuroelectric sources. *Psychophysiology*, *39*(6), 791–808. http://doi.org/10.1111/1469-8986.3960791

Urbach, T. P., & Kutas, M. (2006). Interpreting event-related brain potential (ERP) distributions: Implications of baseline potentials and variability with

application to amplitude normalization by vector scaling. *Biological Psychology*, *72*(3), 333–343. http://doi.org/10.1016/j.biopsycho.2005.11.012

Vallar, G., & Maravita, A. (2009). Personal and Extrapersonal Spatial Perception. In *Handbook of Neuroscience for the Behavioral Sciences* (pp. 322–336). Hoboken, NJ, USA: John Wiley & Sons, Inc. http://doi.org/10.1002/9780470478509.neubb001016

Valyear, K. F., Cavina-Pratesi, C., Stiglick, A. J., & Culham, J. C. (2007). Does tool-related fMRI activity within the intraparietal sulcus reflect the plan to grasp? *NeuroImage*, *36*(SUPPL. 2), 94–108. http://doi.org/10.1016/j.neuroimage.2007.03.031

van Atteveldt, N. M., Formisano, E., Goebel, R., & Blomert, L. (2007). Top-down task effects overrule automatic multisensory responses to letter-sound pairs in auditory association cortex. *NeuroImage*, *36*(4), 1345–1360. http://doi.org/10.1016/j.neuroimage.2007.03.065

Vanderclausen, C., Filbrich, L., Alamia, A., & Legrain, V. (2017). Investigating peri-limb interaction between nociception and vision using spatial depth. *Neuroscience Letters*, *654*, 111–116. http://doi.org/10.1016/j.neulet.2017.05.060

VanRullen, R., & Thorpe, S. J. (2001). Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artifactual objects. *Perception*, *30*(6), 655–668. http://doi.org/10.1068/p3029

Vatakis, A., & Spence, C. (2007). Crossmodal binding: Evaluating the “unity assumption” using audiovisual speech stimuli. *Perception & Psychophysics*, *69*(5), 744–756. http://doi.org/10.3758/BF03193776

Vialatte, F.-B., Maurice, M., Dauwels, J., & Cichocki, A. (2010). Steady-state visually evoked potentials: Focus on essential paradigms and future perspectives. *Progress in Neurobiology*, *90*(4), 418–438. http://doi.org/10.1016/j.pneurobio.2009.11.005

Victor, J. D., & Conte, M. M. (2000). Two-frequency analysis of interactions elicited by Vernier stimuli. *Visual Neuroscience*, *17*(6), 959–973. http://doi.org/10.1017/S0952523800176151

Volcic, R., van Rheede, J. J., Postma, A., & Kappers, A. M. L. (2008). Differential effects of non-informative vision and visual interference on haptic spatial processing. *Experimental Brain Research*, *190*(1), 31–41. http://doi.org/10.1007/s00221-008-1447-0

Vroomen, J., Bertelson, P., & de Gelder, B. (2001). The ventriloquist effect does not depend on the direction of visual attention. *Perception And Psychophysics*, *63*(4), 651–659.

Vroomen, J., & De Gelder, B. (2004). Perceptual effects of cross-modal stimulation: Ventriloquism and the freezing phenomenon. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *The handbook of multisensory processes* (pp. 141–151). Cambridge, MA: MIT Press.

Vroomen, J., & Keetels, M. (2006). The spatial constraint in intersensory pairing: No role in temporal ventriloquism. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(4), 1063–1071. http://doi.org/10.1037/0096-1523.32.4.1063

Wallace, M., Meredith, M. A., & Stein, B. (1992). Integration of multiple sensory modalities in cat cortex. *Experimental Brain Research*, *91*(3), 484–488. http://doi.org/10.1007/BF00227844

Wang, Q., & Burkhalter, A. (2007). Area map of mouse visual cortex. *The Journal of Comparative Neurology*, *502*(3), 339–357. http://doi.org/10.1002/cne.21286

Wang, Y., Celebrini, S., Trotter, Y., & Barone, P. (2008). Visuo-auditory interactions in the primary visual cortex of the behaving monkey: Electrophysiological evidence. *BMC Neuroscience*, *9*(1), 79. http://doi.org/10.1186/1471-2202-9-79

Welch, R. B. (1999). Chapter 15 Meaning, attention, and the "unity assumption" in the intersensory bias of spatial and temporal perceptions. *Advances in Psychology*, *129*(C), 371–387. http://doi.org/10.1016/S0166-4115(99)80036-3

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

Wright, T. M., Pelphrey, K. A., Allison, T., McKeown, M. J., & McCarthy, G. (2003). Polysensory interactions along lateral temporal regions evoked by audiovisual speech. *Cerebral Cortex*, *13*(10), 1034–1043. http://doi.org/10.1093/cercor/13.10.1034

Xu, B., Liu-Shuang, J., Rossion, B., & Tanaka, J. (2017). Individual Differences in Face Identity Processing with Fast Periodic Visual Stimulation. *Journal of Cognitive Neuroscience*, *29*(8), 1368–1377. http://doi.org/10.1162/jocn_a_01126

Yamamoto, S., & Kitazawa, S. (2001). Reversal of subjective temporal order due to arm crossing. *Nature Neuroscience*, *4*(7), 759–765. http://doi.org/10.1038/89559

Zemon, V., & Ratliff, F. (1984). Intermodulation components of the visual evoked potential: Responses to lateral and superimposed stimuli. *Biological Cybernetics*, *50*(6), 401–408. http://doi.org/10.1007/BF00335197

Zhang, D., Hong, B., Gao, X., Gao, S., & Roder, B. (2011). Exploring steady-state visual evoked potentials as an index for intermodal and crossmodal spatial attention. *Psychophysiology*, *48*(5), 665–675. http://doi.org/10.1111/j.1469-8986.2010.01132.x

Zhou, H., Melloni, L., Poeppel, D., & Ding, N. (2016). Interpretations of Frequency Domain Analyses of Neural Entrainment: Periodicity, Fundamental Frequency, and Harmonics. *Frontiers in Human Neuroscience*, *10*, 1–8. http://doi.org/10.3389/fnhum.2016.00274

Zimmer, U., & Macaluso, E. (2007). Processing of multisensory spatial congruency can be dissociated from working memory and visuo-spatial attention. *European Journal of Neuroscience*, *26*(6), 1681–1691. http://doi.org/10.1111/j.1460-9568.2007.05784.x

Zoefel, B., Archer-Boyd, A., & Davis, M. H. (2018). Phase Entrainment of Brain Oscillations Causally Modulates Neural Responses to Intelligible Speech. *Current Biology*, *28*(3), 401–408.e5. http://doi.org/10.1016/j.cub.2017.11.071

Zoefel, B., ten Oever, S., & Sack, A. T. (2018). The Involvement of Endogenous Neural Oscillations in the Processing of Rhythmic Input: More Than a Regular Repetition of Evoked Neural Responses. *Frontiers in Neuroscience*, *12*, 1–13. http://doi.org/10.3389/fnins.2018.00095