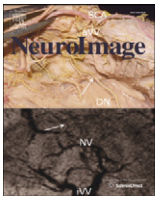
NeuroImage 66 (2013) 110–118

Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg



Task-demands and audio-visual stimulus configurations modulate neural activity in the human thalamus

Björn Bonath [a,b,*], Sascha Tyll [a,f], Eike Budinger [c,d,e], Kerstin Krauel [b], Jens-Max Hopf [c,d], Tömme Noesselt [a,e]

[a] Institute of Biological Psychology, Otto-von-Guericke University, 39106, Magdeburg, Germany
[b] Department of Child and Adolescent Psychiatry, Psychotherapy and Psychosomatics, Otto-von-Guericke University, 39120, Magdeburg, Germany
[c] Department of Neurology, Otto-von-Guericke University, 39120, Magdeburg, Germany
[d] Leibniz Institute for Neurobiology, 39118 Magdeburg, Germany
[e] Center for Behavioral and Brain Sciences, 39120, Magdeburg, Germany
[f] Institute of Cognitive Neurology and Dementia Research, Otto-von-Guericke-University, 39120 Magdeburg, Germany

ARTICLE INFO

Article history:
Accepted 12 October 2012
Available online 22 October 2012

Keywords:
Multisensory interplay
Temporal and spatial task-demands
fMRI
Thalamus
Human

ABSTRACT

Recent electrophysiological studies have reported short latency modulations in cortical regions for multisensory stimuli, thereby suggesting a subcortical, possibly thalamic origin of these modulations. Concurrently, there is an ongoing debate, whether multisensory interplay reflects automatic, bottom-up driven processes or relies on top-down influences. Here, we dissociated the effects of task set and stimulus configurations on BOLD-signals in the human thalamus with event-related functional magnetic resonance imaging (fMRI). We orthogonally manipulated temporal and spatial congruency of audio-visual stimulus configurations, while subjects judged either their temporal or spatial congruency. Voxel-based fMRI results revealed increased fMRI-signals for the temporal versus spatial task in posterior and central thalamus, respectively. A more sensitive region of interest (ROI)-analysis confirmed that the posterior thalamic nuclei showed a preference for the temporal task and central thalamic nuclei for the spatial task. Moreover, the ROI-analysis also revealed enhanced fMRI-signals for spatially incongruent stimuli in the central thalamus. Together, our results demonstrate that both audio-visual stimulus configurations and task-related processing of spatial or temporal stimulus features selectively modulate thalamic processing and thus are in a position to influence cortical processing at an early stage.

Introduction

Real-world events often stimulate more than one sense and merging information across senses may increase the reliability of our representation of the outside world; the neural mechanisms underlying multisensory interplay have been investigated extensively for the past years (Driver and Noesselt, 2008; Stein and Stanford, 2008). Converging evidence suggests that multisensory integration (MSI) can occur at many levels of the central nervous system.

Among them are classical multisensory association cortices [e.g. superior temporal sulcus (STS), posterior parietal cortex (PP), dorsolateral prefrontal cortex (DLPFC), see Driver and Noesselt, 2008; Ghazanfar and Schroeder, 2006; Kaas and Collins, 2004; Kayser and Logothetis, 2007 for review], but also sensory specific low-level auditory areas (for audio-tactile stimulation see e.g. Foxe et al., 2002; Kayser et al., 2005; Lakatos et al., 2007; Murray et al., 2005; and for audio-visual stimulation Bonath et al., 2007; Brosch et al., 2005; Kayser et al., 2007; Noesselt et al., 2007; van Atteveldt et al., 2007) and in low-level visual areas (for audio-visual modulations see Mishra et al., 2007; Noesselt et al., 2007; Watkins et al., 2006; and visuo-tactile stimuli see e.g. Macaluso et al., 2000). Several subcortical structures are also involved in multisensory processing. In addition to the well-established multisensory modulations in the superior colliculi (e.g. Stein and Stanford, 2008), these also include the basal ganglia (Graziano and Gross, 1993), sensory-specific thalamic structures as the medial (MGB) and lateral geniculate body (LGB) (Budinger et al., 2006; Hackett et al., 2007a, 2007b; Komura et al., 2005; Noesselt et al., 2010; see Tyll et al., 2011 for review) as well as sensory non-specific central and posterior thalamic nuclei (Cappe et al., 2009; Hackett et al., 2007a, 2007b; Naumer and van den Bosch, 2009; see Cappe et al., 2012 for review).

Several electrophysiological studies on the temporal dynamics on multisensory interplay in macaques and humans reported very short latencies for interactions of audio-tactile and audio-visual stimulation (Brosch et al., 2005; Giard and Peronnet, 1999; Murray et al., 2005; though see Teder-Salejarvi et al., 2002), thus hinting at the possibility that some multisensory interactions might already occur at the thalamic level prior to the subsequent cortical processing (Driver and Noesselt, 2008; Lakatos et al., 2007; Senkowski et al., 2011).

However, it remains to be elucidated whether all of the structures reported to be instrumental in multisensory interplay genuinely

---

[*] Corresponding author at: Department of Child and Adolescent Psychiatry, Psychotherapy and Psychosomatics, Otto-von-Guericke University, 39120, Magdeburg, Germany.

E-mail address: bjoern.bonath@ovgu.de (B. Bonath).

1053-8119/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neuroimage.2012.10.018

integrate information from different modalities or rather provide a feature-specific or attentional framework for the task at hand. Most of the cortical and subcortical structures mentioned above are not exclusively associated with multisensory interplay: stimulus-driven responses within parietal regions are known to reflect attentional control processes (Corbetta et al., 2000; Greenberg et al., 2011) and the posterior STS as part of the temporo-parietal junction (TPJ) may also be modulated by attention (Karnath et al., 2002; see also Hein and Knight, 2008 for a review of bottom-up and top-down functions represented in the STS).

With regard to subcortical structures potentially involved in multisensory interaction processes, the superior colliculi are known to be a key structure for the control of eye movements (Krebs et al., 2011; Sparks, 2002; Wurtz and Goldberg, 1971) and are involved in pupil dilation as part of the orienting reflex (Wang et al., 2012). For thalamic structures, it has been shown that central and posterior nuclei of the thalamus are involved in task-dependent processes (Arend et al., 2008), eye-movement-control (Tanaka, 2005, 2006, 2007; Tanibuchi and Goldman-Rakic, 2005; Wurtz et al., 2011), distractor filtering (Kastner et al., 2004; Smith et al., 2009; Snow et al., 2009; Strumpf et al., in press) and/or attentional processes (Michael and Buron, 2005; Wilke et al., 2010; Yantis et al., 2002). For instance, using purely visual stimuli, Arend et al. (2008) reported that attention to spatial or temporal unisensory visual features is impaired in patients with central or posterior thalamic nucleus lesions, respectively. Finally some of the thalamic structures may be well adapted to modify sensory-specific responses in a task-set and/or attention-dependent manner (Benevento and Miller, 1981; Hulme et al., 2011; Kastner et al., 2004; Strumpf et al., in press). Moreover, thalamic neurons may also be involved in cortico-cortical control (for review see e.g. Sherman, 2007), thereby acting as neuronal hub, for example to synchronize oscillations between cortical areas (for review see also Shipp, 2003; Siegel et al., 2012; Engel et al., 2012).

Thus, given the potential multi-functionality of these regions, it is conceivable, that some of the observed modulations within multisensory structures are rather driven by task-demands than by stimulus-driven multisensory interplay per se. So far however, most neurophysiological studies on multisensory processing in humans focused on one single aspect of audio-visual interplay in isolation, for instance, on spatial configurations using an auditory localization task (Bonath et al., 2007) or temporal judgments (Bushara et al., 2001; McDonald et al., 2005; Mishra et al., 2007; Watkins et al., 2006) or semantic content discrimination (e.g. Doehrmann and Naumer, 2008) or they investigated cross-modal attention effects on unisensory and multisensory stimulus configurations (Busse et al., 2005; Donohue et al., 2011).

The few studies testing the effects of attending to one versus another modality reported differential ERP-modulations over frontal electrodes for identical stimuli under different attentional conditions (Talsma et al., 2007), or found task-dependent fMRI-modulations in multisensory cortex (Lewis and Noppeney, 2011; Macaluso et al., 2004). Van Atteveldt et al. reported that fMRI responses to audio-visual stimuli in auditory cortex are also modulated by task sets (van Atteveldt et al., 2007, see also Lakatos et al., 2009 for evidence in macaques). Finally, Bushara et al. (2001) observed modulations in the parietal cortex, the insula and most notably the thalamus when attending to temporal stimulus properties using positron emission tomography (PET). But, to our knowledge, no imaging study so far has formally tested how selective attention to spatial versus temporal properties of multisensory objects may modulate neural processing. This question, however, is of utmost importance, because temporal and spatial properties are key determinants of multisensory integration (e.g. Stein and Meredith, 1993) and most previous neuroimaging studies in humans have often investigated MSI using one specific task only. Thus, it remains to be elucidated, whether subcortical and cortical regions potentially involved in multisensory processing are part of a bottom-up driven (automatic) interplay network, or whether these structures form a flexible network that is dependent of the task-demands at hand.

In the present study, we tested how the manipulation of task sets and audio-visual stimulus configurations modulates the neuronal activity in the human brain using event-related fMRI. With regard to the current interest in the thalamic involvement in multisensory interplay (Driver and Noesselt, 2008; Schroeder and Foxe, 2005; Senkowski et al., 2008) and some reports on thalamic involvement in task-related processing of visual and audio-visual stimuli (Arend et al., 2008; Bushara et al., 2001; Coull and Nobre, 1998), we focused on whether the neural responses to task set and audio-visual stimulus configurations differ in the human thalamus.

Materials and methods

Participants

Eighteen right-handed volunteers (eight female, age range 21–33 years) with no history of psychiatric or neurological disorders participated in the experiment. Informed consent was obtained from all the subjects in accord with local ethics, and each was paid for his or her participation. Only highly trained subjects with expertise in visual experiments and the ability to hold fixation during experimental sessions participated in this experiment.

Stimuli and experimental design

Pure sounds (2 kHz, 80 dB, 30 ms duration) were presented from non-visible piezoelectric speakers situated in the left (A_L) and the right (A_R) hemifield (15° eccentricity from center position) above the subjects' head within the scanner (stimuli were presented within silent interscan periods, see below for scanning protocol). Visual stimuli were checkerboards (size: 2.6°, 30 ms duration) presented to the left (V_L) or to the right (V_R) visual hemifield (12° eccentricity) projected on a mirror display mounted above the subject. Audio-visual stimulus combinations (A_LV_L, A_RV_R, A_LV_R, and A_RV_L) could occur synchronously (syn) and asynchronously (asyn). To keep the number of experimental conditions as low as possible, we decided to only use a single asynchronous condition with the sound preceding the visual stimulus at a fixed gap by 300 ms (see Fig. 1A). This particular stimulus combination (AV instead of VA) was chosen based on previous results (Slutsky and Recanzone, 2001) which reported that subjects' accuracy to separate audio-visually presented stimuli into two temporally independent events is well above threshold at a temporal gap of 200 ms. In accord, results from a behavioral pilot study indicated that subjects' sensitivity for temporal and spatial incongruence (here 300 ms and 27°, respectively) was well above threshold. Stimuli were randomized and jittered (mean ISI: 3000 ms [range: 2100–6400 ms, Poisson distributed (Hinrichs et al., 2000)] and were presented using Presentation 9.11 (Neurobehavioral Systems, Inc., CA). Anatomical regions were labeled using the 'Atlas of the Human Brain' (Mai et al., 2008).

Procedure

Participants completed ten sessions of the main experiment and were instructed to discriminate run-wise either spatial or temporal audio-visual stimulus congruence prior to each session (task-order counterbalanced across subjects). To familiarize subjects with both tasks, two practice blocks comprising all stimulus conditions were conducted inside the scanner prior to scanning. For both, the temporal and spatial task, subjects were instructed to hold central fixation. In the spatial task subjects judged whether the checkerboard and the sound occurred within the same or opposite hemifields while ignoring temporal features; in the temporal task they judged whether the stimuli appeared synchronously or asynchronously while ignoring spatial stimulus configurations (see Fig. 1B for stimulus configurations). During the