ISSN 1749-8023 (print), 1749-8031 (online)
International Journal of Magnetic Resonance Imaging
Vol. 01, No. 01, 2007, *pp*. 003-014

Perceptual Gain and Perceptual Loss: Distinct Neural Mechanisms of Audiovisual Interactions*

Yi Jiang, Shihui Han +

Department of Psychology and Functional Imaging Center, Peking University, Beijing, P. R. China

(*Received 10 January 2007, accepted 13 March 2007*)

Abstract. What we see can be influenced by what we hear. However, the neural basis of the audiovisual interactions, and whether common or distinct mechanisms can be involved, remains unclear. Here we showed that auditory beeps can either induce illusory light flashes (perceptual gain) or reduce perceived number of physical flashes (perceptual loss), when beeps mismatched with simultaneous flashes. Moreover, we reported neuroimaging data on the neural substrates of perceptual gains and perceptual losses in visual perception. Illusory gains in perceived flashes produced activations in the left supramarginal gyrus, the left prefrontal cortex, and the right cerebellum, reflecting a neural network associated with integrative processes in working memory. Illusory losses of perceived flashes, however, were associated with activations in the medial occipital cortex and thalamus, linking with brain regions associated with early visual processing. The results suggest that distinct neural systems underlie distinct audiovisual interactions in the human brain.

Keywords: MCSpan /TTD 8.9 BT1 0 TW o TOT 0.025 Tw 22.8 0(w)-4(e)2(l-known exa)-3(m)-(ple)-3(0)-5(f visual infor

of visual stimuli can also be affected by simultaneous auditory events
mpanied by more than one beep, it can be perceived to flash twice or
erception are particularly salient when stimuli are displayed in the
nterval between the visual and auditory events [3] Moreover, visual
visual perceptual gains, recorded at the occipital area, are similar
hes [9] sugg esting a possible common neural mechanism for the
ashes. However, the limitation of spatial resolution of ERPs makes it
nvolved in such auditory-visual interactions. In addition, it is unclear
ese bidirectional cross-modal interactions are the same or different.

In the present paper we first extend our knowledge of audiovisual interactions by demonstrating a new illusion of visual-perceptual loss, supplementing prior evidence on visual-perceptual gains [7]. When brief visual flashes are accompanied by simultaneous auditory beeps, the number of perceived flashes is

* This work was supported by grants from National Natural Science Foundation of China (Project 30630025), the Ministry of Science and Technology of China (Project 2002CCA01000), and the Chinese Academy of Sciences (KGCX2-SW-101).

+ Corresponding author. *E-mail address*: shan@pku.edu.cn.

influenced by the beeps: there is a perceptual gain, i.e., auditory beeps induce illusory light flashes when there are more beeps than flashes [7], and a perceptual loss, i.e., auditory beeps reduce perceived number of physical flashes when there are fewer beeps than flashes. We then employed functional magnetic resonance imaging (fMRI) to investigate the neural substrates of these effects. We compared brain activations in conditions where the visual and auditory stimuli mismatched with baselines where there were only visual

[redacted]

Fig.1: Illustrations of the stimuli and procedure used in the current study. 3F5B: 3 flashes accompanied with 3 beeps, 5F3B: 5 flashes accompanied with 3 beeps.

Subjects were asked to remain fixated on the fixation cross during the experiments. In the behavioral experiment, subjects were required to judge the number of the flashing disks on each trial. They wrote down the perceived number of flashes on an answer sheet and then pressed a key to start the next trial. There were 16 trials in each of the six conditions which were presented in a random order. In the fMRI experiment that employed a box-car design, subjects were not required to make any responses during the scan procedure. Each trial was presented within the 500-ms delay during each TR. Six scans of 164 s were first obtained from each subject. Each scan consisted of eight 20-second epochs of 10 trials, alternating randomly between the 3F, 5F, 3F5B, and 5F3B conditions. The first 4 seconds of each scan were excluded from statistical analysis to obtain a steady baseline. Another three scans of 124 s were obtained with the identical foregoing parameters and three conditions, i.e., 3F5B, 5F3B, and 5F5B. Additional two scans were also obtained, which included only three conditions, i.e., 3 or 5 beeps with fixation and fixation with no beeps. The epoch of beeps (including 10 trials) or baseline also lasted for 20 s.

2.3. fMRI Data Acquisition and Analysis

Scanning was performed on a 3T Siemens Trio system using a standard head coil at Beijing MRI Center for Brain Research. Twenty-four axial slices of functional images that covered the whole brain were acquired using a gradient-echo echo-planar pulse sequence (64×64×24 matrix with 3.4×3.4×6-mm spatial resolution, TR=2000 ms, TE=30 ms, FOV=220 mm, flip angle=90º). Anatomical images were obtained using a standard 3D T1-weighted sequence (resulting in a 256×256×176 matrix with 0.938×0.938×1.3-mm spatial resolution, TR=1600 ms, TE=3.93 ms). Subjects' heads were immobilized during the scanning sessions using pieces of foam.

SPM99 (the Wellcome Department of Cognitive Neurology, UK) was used for data processing and analysis. Following correction for differences in the timing of slice acquisition within a volume, the functional images were realigned to the first scan to correct for the head movement between scans. The structural image was coregistered with the mean functional image produced during the process of realignment. All images were normalized to a 2×2×2 mm^3 Montreal Neurological Institute (MNI) template in Talairach space [10] using bilinear interpolation. Functional images were spatially smoothed using a Gaussian filter with a full-width at half maximum (FWHM) parameter set to 8 millimeters. The image data were modelled using a box-car function. Contrasts were defined to compare the effect of sound on visual perception. Regions preferentially engaged in visual illusion were defined as areas more activated by 3F5B was used to identify neural substrates for the processing of simple auditory stimuli. Random effect analyses were then conducted across the group of subjects based on statistical parameter maps from each individual subject to allow population inference. Areas of significant activation were identified using a voxel-based t-test and a significance threshold of $P < 0.005$ (uncorrected). The SPM coordinates for standard brain from MNI template were converted to Talairach coordinates using a non-linear transform method (http://www.mrc-cbu.cam.ac.uk/Imaging/mnispace.html).

3. Results

3.1. Behavioural Data

The reported numbers of visual flashes in each condition are shown in Fig.2. There were both perceptual gains and perceptual losses in visual perception when the visual and auditory stimuli mismatched. Perceptual gains were demonstrated by the increased number of flashes reported in Condition 3F5B than in Condition 3F (3.58 vs. 3.25, $t(9) = 3.86$, $p<0.005$) and the increased number of flashes reported in Condition 3F5B than in Condition 3F3B (3.58 vs. 3.01, $t(9) = 3.50$, $p<0.007$). Perceptual losses were also apparent. There were fewer reports of visual flashes in Condition 5F3B compared with the baseline, Condition 5F (3.06 vs. 3.83, $t(9) = 4.84$, $p<0.001$), and fewer reports of visual flashes in Condition 5F3B than in Condition 5F5B (3.06 vs. 3.83, $t(9) = 4.97$, $p<0.001$). The magnitudes of these two illusions were assessed in a repeated measure analysis of variance with Flash Number (3 or 5 flashes) and Beep (presence or absence of beeps) as independent variables and performed on 3F, 3F5B, 5F, and 5F3B conditions. There was a significant interaction between Flash Number and Beep, due to the face that the effect of perceptual loss was stronger than that of perceptual gain ($F(1, 9) = 19.0$, $p<0.002$). These differences were not simply due to subjects reporting the number of beeps, since more flashes were reported in condition 5F5B than in c273 Td[)7( )]78B6 Tw -34

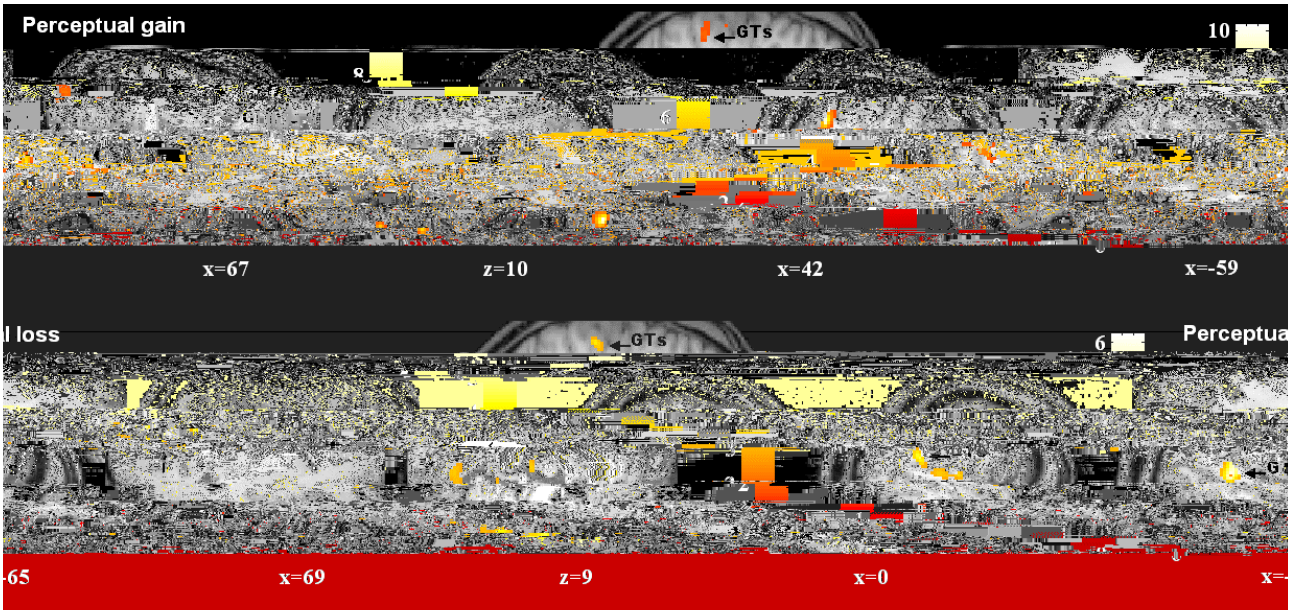

Fig.3: Brain activations associated with sound-induced perceptual gain and perceptual loss. The results of the group analysis from 10 subjects were plotted on MR images of a representative subject. The color bar indicates the scale of z values. Cer = cerebellum; GFm = middle frontal gyrus; Gsm = supramarginal gyrus; GTs = superior temporal gyrus; Occ = occipital cortex; Thm = thalamus.

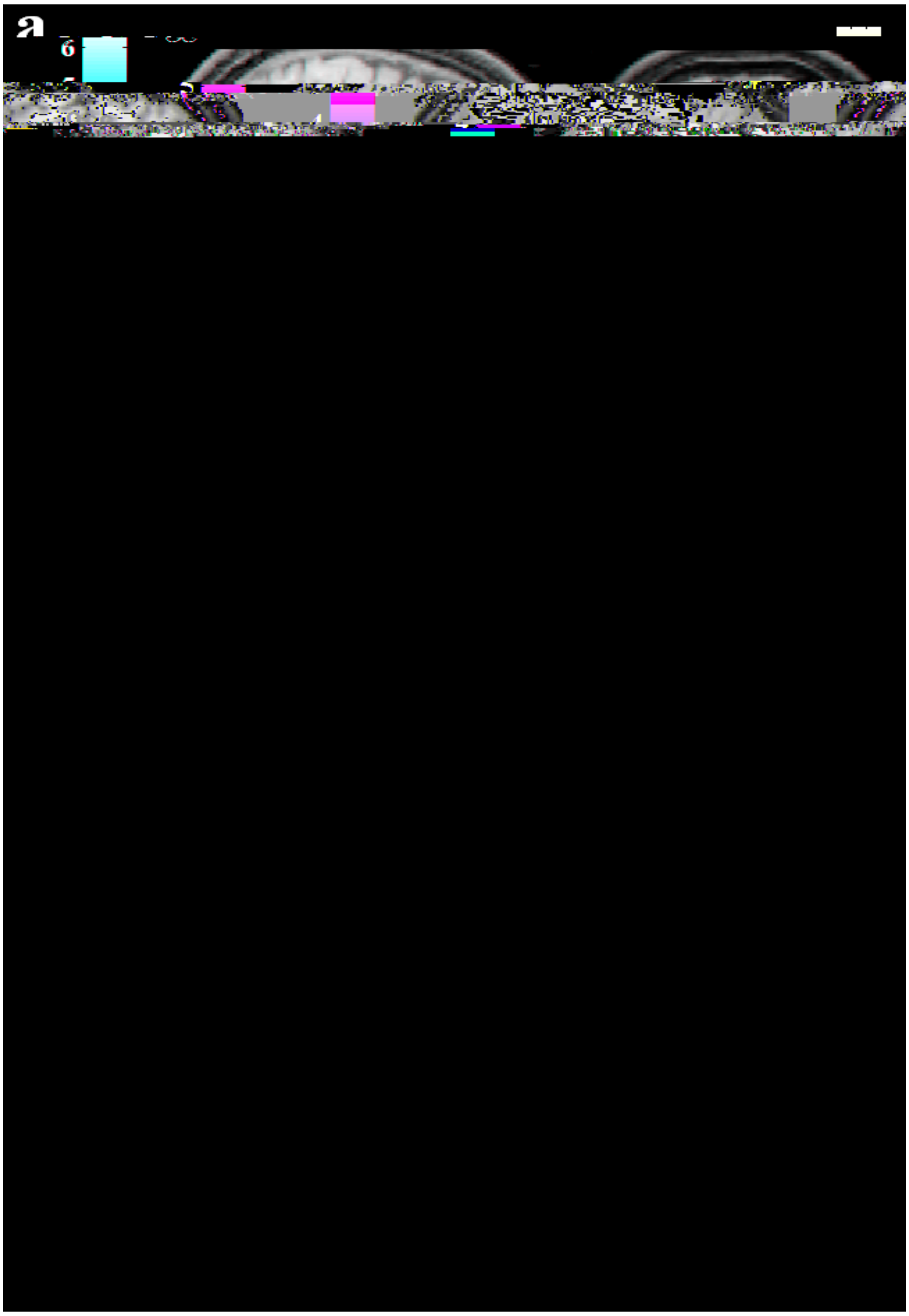

Fig.4: Brain activations associated with lack of perceptual and physical flashes. (a) Brain activation in the medial visual cortex shown in the contrast between 5F3B with 3F conditions; (b) Brain activations in bilateral visual cortices shown in the contrast between 5F with 5F3B conditions; (c) Brain activations in bilateral visual cortices shown in the contrast between 5F with 3F condition; (d) Brain activations in bilateral superior temporal gyri in the contrast between unimodal auditory stimuli and silent baseline conditions. The results of the group analysis from 10 subjects were plotted on MR images of a representative subject. The color bar indicates the scale of z values. GTs = superior temporal gyrus; Occ = occipital cortex.

Finally we examined if the activations observed in bilateral superior temporal cortex associated with perceptual gains and losses reflected early audiovisual interactions that are not specific to the illusions. 20-second epochs of 3F5B, 5F3B, and 5F5B conditions were presented randomly in three scans. Conditions 3F5B and 5F3B were respectively contrasted with Condition 5F5B in which no perceptual gains and perceptual losses occurred. In addition, flashes and beeps were presented simultaneously in the 5F5B condition and thus early audiovisual interactions also occurred, similar to those in the 3F5B/5F3B conditions.

visual events (3F), there were activations in bilateral superior temporal gyri, the left supramarginal gyrus, the left prefrontal cortex, and the right cerebellum. These activations mark the neural locus of the conscious reports of increased flashes. The effect does not reflect problems due to the presence of the auditory beeps, since the presence of the beeps has no effect on brain activations outside the primary auditory cortex (BA 41/42). In prior studies of cross-modal interactions (e.g., the McGurk effect), activations in posterior superior temporal gyrus (BA 22) has also been reported [3], and it is highly likely to be involved in the cross-modal synthesis of audiovisual speech [3, 12]. Here we used single tones rather than speech, and audiovisual interaction rather than visual-auditory (as with the McGurk illusion). Nevertheless, if the posterior superior temporal cortex supports the interaction between flashes and tones, and not just speech, then it may play a role in the initial integration of auditory and visual stimuli. This analysis is consistent with our findings that the superior temporal activation was eliminated in the contrast between 3F5B and 5F5B conditions, since early integrative processing of auditory and visual information occurred in both conditions. The supramarginal gyrus may also contribute to the integration of auditory and visual information. However, the integrative process in this brain area may not be audiovisual specific, because other research has shown evidence for the involvement of the supramarginal gyrus in integration of visual and tactile information [13].

In addition to the posterior superior temporal cortex, the perceptual gain condition was associated with activation in the left prefrontal cortex and the right cerebellum. Recent fMRI studies have shown that the network consisting of the left prefrontal cortex and the right cerebellum plays an important role in recalling items from short-term auditory memory [14] and in the storage of precise temporal structures of tones in working memory [15], and thus have been linked to the role of phonological loop in working memory [16]. In the current study, it may be that information from the beeps is stored in the phonological loop of working memory, and then integrated with visual information stored in the visuospatial sketchpad of working memory, possibly through the episodic buffer that stores information held in a multimodal code and binds information from subsystems [17]. When the additional beeps do not match the number of flashes, then matches, created by the beeps, may lead to creation of extra flashes through top-down activation to working memory. This account may also be applied to the McGurk effect because similar activations in the frontal-cerebellum loop were observed in the auditory illusion produced by visual stimuli [3].

Interestingly, our fMRI data showed that sound-induced illusory flashes did not alter neural activities in the striate and extrastriate cortices, suggesting that sound does not produce the illusion by modulating activities in early stages of the visual pathway. Similarly, the activity in the primary and secondary auditory cortices is not changed by visual stimuli that produce the McGurk effect [3]. Taken together, these results suggest that visual and auditory illusions arising from the cross-modal interaction do not necessarily change the activities of the primary sensory and parasensory association cortices. The sound-modulation of VEPs to auditory-induced visual flashes at 170 ms post-stimulus over the occipital region [9] may reflect audiovisual interactions at a later stage of the visual pathway. For instance, the activity in V5 after exposure to audiovisual speech is enhanced relative to activity after a unimodal stimulus [18].

The fMRI data on the perceptual loss effect contrast with the results on perceptual gains. In the perceptual loss condition (5F3B), there was evidence of increased activations in the superior temporal gyrus, the medial occipital cortex, and thalamus compared with the physical-match baseline (5F), and there was reduced activation in the lateral occipital cortex relative to the 5F baseline. These results, in contrast to the data on illusory gains in perception, are consistent with there being additional suppressive effects on visual processing in the occipital cortex, wh l cortex, lat wh which fewer flashes are reported compared with physically matched baseline. The recruitment of the posterior thalamus in the perceptual loss effect fits with other research showing the involvement of the thalamus when information from multiple modalities is integrated [3, 21]. However, the thalamus activation might not be specific to perceptual losses because the thalamus activation was eliminated when auditory beeps were used in the baselines (5F5B) to contrast the perceptual losses (5F3B). It may be that subcortical interactions help to establish the inhibition of the visual signals when perceptual losses take place.

Although there were clear difference in activation between the perceptual gain and loss conditions, there were also common areas involved. In particular, the superior temporal gyrus (BA 22/42) was activated in in