Neuropsychologia 50 (2012) 85–89

Contents lists available at SciVerse ScienceDirect

Neuropsychologia

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

The neural correlates of coloured music: A functional MRI investigation of auditory–visual synaesthesia

J. Neufeld[a,c,*], C. Sinke[a,c], W. Dillo[a], H.M. Emrich[a,c], G.R. Szycik[a], D. Dima[b], S. Bleich[a], M. Zedler[a]

[a] *Hannover Medical School, Clinic for Psychiatry, Social Psychiatry and Psychotherapy, Germany*
[b] *Kings's College London, Institute of Psychiatry, Department of Psychosis Studies, Section of Neurobiology of Psychosis, UK*
[c] *Center of Systems Neuroscience, Hannover, Germany*

ARTICLE INFO

Article history:
Received 1 August 2011
Received in revised form 31 October 2011
Accepted 1 November 2011
Available online 7 November 2011

Keywords:
Synaesthesia
Auditory–visual
fMRI
Inferior parietal
Multi-sensory

ABSTRACT

In auditory–visual synaesthesia, all kinds of sound can induce additional visual experiences. To identify the brain regions mainly involved in this form of synaesthesia, functional magnetic resonance imaging (fMRI) has been used during non-linguistic sound perception (chords and pure tones) in synaesthetes and non-synaesthetes. Synaesthetes showed increased activation in the left inferior parietal cortex (IPC), an area involved in multimodal integration, feature binding and attention guidance. No significant group-differences could be detected in area V4, which is known to be related to colour vision and form processing. The results support the idea of the parietal cortex acting as sensory nexus area in auditory–visual synaesthesia, and as a common neural correlate for different types of synaesthesia.

1. Introduction

In synaesthesia the perception of a certain stimulus ('inducer') results automatically in an additional internally generated sensation ('concurrent'). The main characteristics of synaesthesia are its consistency (Baron-Cohen, Wyke, & Binnie, 1987; Cytovic, 2002; Simner & Logie, 2007) and automaticity (Lupianez & Callejas, 2006; Mills, Boteler, & Oliver, 1999): one inducer always triggers the same concurrent sensation, which cannot be suppressed or altered voluntarily. The most investigated form is grapheme-colour synaesthesia, in which achromatic letters, words or numbers are perceived in specific colours. In auditory–visual synaesthesia, all kinds of sound (e.g. music or single tones) can induce additional visual experiences, as for example colours, forms and textures (Cytovic, 2002; Ward, Huckstep, & Tsakanikos, 2006). Investigating the neural basis of acoustically induced synaesthesia is of particular interest as it is a condition, in which visual experiences can be elicited without any external visual input. Further, a better understanding of this phenomenon may help to shed more light on the mechanisms of audio–visual integration.

Two types of models are discussed to explain the mechanisms of synaesthesia: a model of direct cross-activation (Ramachandran & Hubbard, 2001) and a disinhibited feedback model (Grossenbacher & Lovelace, 2001). The former model propagates a direct linkage between the areas of inducer- and concurrent-representation, e.g. in grapheme-colour synaesthesia the area of grapheme representation and the adjacent colour processing region V4 in the fusiform gyrus, while the latter proposes an unusual activation of concurrent-areas via disinhibition of feedback coming from a "multisensory nexus" area.

There is growing evidence from several neuro-imaging studies for an involvement of V4 (Brang, Hubbard, Coulson, Huang, & Ramachandran, 2010; Hubbard, Arman, Ramachandran, & Boynton, 2005; Nunn et al., 2002) and the parietal cortex (Rouw & Scholte, 2007, 2010; van Leeuwen, Petersson, & Hagoort, 2010; Weiss & Fink, 2009; Weiss, Zilles, & Fink, 2005) in grapheme-colour synaesthesia. However, it is not clear if the same neural mechanisms are involved in different types of synaesthetes or different forms of synaesthesia. One possibility is that there is one common factor and additional variable factors which depend on individual differences (Hubbard, 2007; Rouw & Scholte, 2010). A recent investigation using dynamic causal modelling (DCM) of fMRI data (van Leeuwen, den Ouden, & Hagoort, 2011) found evidence for different mechanisms underlying synaesthesia depending on the type of synaesthetes. According to their data, disinhibited feedback from the parietal cortex to V4 is more relevant for associator synaesthetes (who perceive synaesthetic colours evoked by graphemes

---

[*] Corresponding author at: Hannover Medical School, Clinic for Psychiatry, Social Psychiatry and Psychotherapy, Carl-Neuberg-Straße 1, 30625 Hannover, Germany. Tel.: +49 511 532 6658.
E-mail address: neufeld.janina@mh-hannover.de (J. Neufeld).

0028-3932/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuropsychologia.2011.11.001

Table 1
Mean values, standard deviations (SD) and *T*-statistics of demographic data.

	Synaesthetes	Controls	Statistics
N (male)	14 (5)	14 (5)	
Age (SD)	38.00 (13.77)	36.79 (12.64)	$t = 0.215$; $p = 0.832$
MWT-B[a] right answers (SD)	30.79 (3.87)	30.00 (4.13)	$t = 0.519$; $p = 0.608$
OMSI[b] probability in % (SD)	43.45 (27.66)	36.54 (26.76)	$t = 0.671$; $p = 0.508$
Years of music lessons (SD)	7.64 (5.92)	8.36 (8.85)	$t = 0.251$; $p = 0.804$
Years of instrumental training (SD)	8.21 (10.53)	7.36 (9.46)	$t = 0.227$; $p = 0.822$
Handedness: right (left)	13 (1)	13 (1)	
Grapheme-colour synaesthesia[c]	11	–	
Coloured month/week-days[c]	10	–	
Other forms of synaesthesia[c]	11	–	

[a] Mehrfach Wortschatz Test B according to Lehrl et al.

[b] Ollen Musical Sophistication Index according to Ollen et al.

[c] Additionally reported forms of synaesthesia.

in their ‘mind’s eye’), while a cross-activation mechanism between V4 and the ‘letter shape area’ is more relevant for projector synaesthetes (who see the colour directly projected to the written letter). The results found by van Leeuwen et al. fit also to a combined model of synaesthesia (Hubbard, 2007) integrating the cross activation idea together with a parietal ‘hyperbinding’ mechanism (Esterman, Verstynen, Ivry, & Robertson, 2006).

While ‘coloured-hearing’ was a very popular topic of the scientific literature from the 18th to the 20th century (Marks, 1975), few recent studies have investigated auditory–visual synaesthesia depending on non-linguistic inducer stimuli. A recent investigation demonstrated that on the one hand this form of synaesthesia shows the characteristics of genuine synaesthesia (consistency and automaticity) and on the other hand that it seems to recruit mechanisms which are also used in normal cross-modal perception (Ward et al., 2006). A recent electrophysiological event related potential study on auditory–visual synaesthesia revealed differences in early as well as late components between synaesthetes and controls due to tone perception, but no evidence for an auditory evoked potential over occipital sites (Goller, Otten, & Ward, 2009). The authors proposed an involvement of audio–visual integration areas which lay close to regions normally involved in auditory perception.

The current study investigates the neural correlates of auditory–visual synaesthesia induced by single tones and chords. It is to our knowledge the first group study on this form of synaesthesia using fMRI. Our aims were two folded: first to identify brain areas which show differences in activation according to inducing stimuli in synaesthetes compared to controls. Second to test the hypothesis that V4 is involved in auditory–visual synaesthesia, as it has been reported for grapheme-colour synaesthesia, by comparing brain activation in this area as a region of interest (ROI) between groups.

2. Methods

2.1. Participants

Fourteen auditory–visual synaesthetes and fourteen control subjects, who did not report synaesthesia, participated in the study. The controls were matched for age, sex, handedness, IQ as measured by the MWT-B (Mehrfach-Wortschatz-Intelligenztest B) (Lehrl, 1995) and musical expertise as measured by the Ollen Musical Sophistication Index (OMSI) (Ollen, 2006) and years of music lessons (Table 1). The local ethics committee approved the study and written informed consent was obtained from all participants.

Synaesthetes were asked (directly prior to the scanning procedure) to rate the strength of their synaesthesia induced by sounds on a 10 point scale (1 = very weak, 10 = very strong).

2.2. Consistency of synaesthesia

All participants performed a consistency test (Eagleman, Kagan, Nelson, Sagaram, & Sarma, 2007), a modification of the offline version of the Synaesthesia Test Battery (http://www.synesthete.org/) of Eagleman and colleagues, on a PC. Thirty-six pure tones of 12 different timbres were presented, with three different pitches for each instrument. Every stimulus was presented separately and three times in a randomized order, so that there were 108 stimuli in total. The stimuli were generated on an apple computer using the EXS24 sampler of the audio sequencer software ‘Logic Audio Pro’ (Emagic). They were edited for length and loudness and presented via standard headphones.

Subjects were instructed to adjust loudness to a comfortable listening level and to select for every stimulus a colour by moving a cross hair cursor over a colour matrix. Synaesthetes were asked to choose the colour which matched their experienced synaesthetic colour induced by the tone best, non-synaesthetes were asked to select the colour which they thought to fit best to the tone. Every tone could be played as often as needed until participants confirmed their choice by pressing a button. Consistency was calculated via the geometric distance of the RGB (red green blue) values between the three repetitions as described in (Eagleman et al., 2007).

2.3. Functional magnetic resonance imaging (fMRI) procedure

2.3.1. Experimental design

To avoid habituation, different sound stimuli were presented in a pseudo randomized order. There were six sound conditions: major, minor and dissonant piano chords and pure piano, sine and bassoon tones (each stimulus group in 12 different pitches: C’, Cis’, D’, Dis’, E’, F’, Fis’, G’, Gis’, A’ Ais’ H’). Stimuli were presented via pneumatic headphones in an event-related design with three sessions and 48 stimuli per session (8 stimuli per condition per block; 24 stimuli per condition in total), with a stimulus duration of 2 s and an inter-stimulus interval of 13 s. Between the sessions the participants had the opportunity to relax to avoid tiring and attention diminishment. To achieve a better signal to noise ratio of acoustic stimuli against the scanner noise, the sound level was adjusted individually for each subject during a test scan in which some of the stimuli were presented and sound level was adjusted until the stimuli were clearly audible and tones and chords were clearly discriminable for the subject. All participants held a response device in their right hand and performed a task during measurement to guaranty that they fully attended the stimuli. They were asked to press the right button (with their right middle finger) when hearing a chord and the left one (with their right index finger) when hearing a tone. Subjects were instructed to keep their eyes closed during the sessions to avoid visual deflection. All synaesthetes reported that they perceive synaesthesia with open as well as with closed eyes.

2.3.2. Image acquisition

Functional images were acquired on a 1.5 T General Electrics scanner (Signa Horizon; GE Medical Systems, Milwaukee, WI) equipped with a standard head coil, at the Institute of Diagnostic and Interventional Neuroradiology, Medical School Hannover. T2* functional scans covering the whole brain were acquired by using a multislice two-dimensional echo-planar imaging (EPI) sequence (acquisition matrix 64 × 64 pixels, 26 axial slices, TR = 3000 ms, echo-time (TE) = 40 ms, Field of view (FOV) = 26 cm, slice thickness = 5 mm, flip angle = 90°). Measurements were acquired in three sessions of 12 min. Each fMRI time series consisted of 244 images; the first 4 of them were discarded to allow the scanner to reach a steady state.

2.3.3. Image processing and data analysis

Image processing and statistical analysis was conducted with spm5 (Statistical Parametric Mapping software version 5, Wellcome department of Imaging Neuroscience, London, UK, http://www.fil.ion.ucl.ac.uk) using MATLAB7.0 release 14 (The Mathworks Inc., Natrick, MA). Images were realigned to the 1st volume to correct for inter-scan movements by means of a rigid body transformation with three rotation and three translation parameters. Further, the EPI volumes were spatially normalized to a standard template of the Montreal Neurological Institute (MNI, Canada), resulting in a voxel size of 2 mm × 2 mm × 2 mm, and smoothed (with a Gaussian smoothing kernel of 8 mm) to create statistical maps of changes in relative regional BOLD responses corresponding to the six experimental conditions.

Data analysis was performed on the single subject level by modelling the six stimulus conditions (major, minor and dissonant piano chords and piano, bassoon