Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Cognitive Brain Research 20 (2004) 510–518

COGNITIVE BRAIN RESEARCH

www.elsevier.com/locate/cogbrainres

Research report

Learned audio-visual cross-modal associations in observed piano playing activate the left planum temporale. An fMRI study

Takehiro Hasegawa[a], Ken-Ichi Matsuki[b], Takashi Ueno[c], Yasuhiro Maeda[d], Yoshihiko Matsue[d], Yukuo Konishi[a], Norihiro Sadato[e,f,*]

[a] *Department of Infants' Brain and Cognitive Development, Tokyo Women's Medical University, Japan*
[b] *Faculty of Education and Regional Studies, Fukui University, Japan*
[c] *Division of Life Sciences, Kansei and Human Welfare Research Center, Tohoku Fukushi University, Japan*
[d] *Graduate School of Education, Tohoku University, Japan*
[e] *National Institute for Physiological Sciences, Okazaki, Japan*
[f] *JST (Japan Science and Technology Corporation)/RISTEX (Research Institute of Science and Technology for Society), Kawaguchi, Japan*

Accepted 12 April 2004
Available online 9 June 2004

Abstract

Lip reading is known to activate the planum temporale (PT), a brain region which may integrate visual and auditory information. To find out whether other types of learned audio-visual integration occur in the PT, we investigated "key-touch reading" using functional magnetic resonance imaging (fMRI). As well-trained pianists are able to identify pieces of music by watching the key-touching movements of the hands, we hypothesised that the visual information of observed sequential finger movements is transformed into the auditory modality during "key-touch reading" as is the case during lip reading. We therefore predicted activation of the PT during key-touch reading. Twenty-six healthy right-handed volunteers were recruited for fMRI. Of these, 7 subjects had never experienced piano training (naïve group), 10 had a little experience of piano playing (less trained group), and the remaining 9 had been trained for more than 8 years (well trained group). During task periods, subjects were required to view the bimanual hand movements of a piano player making key presses. During control periods, subjects viewed the same hands sliding from side to side without tapping movements of the fingers. No sound was provided. Sequences of key presses during task periods consisted of pieces of familiar music, unfamiliar music, or random sequences. Well-trained subjects were able to identify the familiar music, whereas less-trained subjects were not. The left PT of the well-trained subjects was equally activated by observation of familiar music, unfamiliar music, and random sequences. The naïve and less trained groups did not show activation of the left PT during any of the tasks. These results suggest that PT activation reflects a learned process. As the activation was elicited by viewing key pressing actions regardless of whether they constituted a piece of music, the PT may be involved in processes that occur prior to the identification of a piece of music, that is, mapping the complex sequence structure of hand movements onto the sequence of sounds.

Theme: Motor systems and sensorimotor integration
Topic: Cortex

Keywords: Auditory; Visual; Cross-modal matching; Integration; Planum temporale

1. Introduction

The human planum temporale (PT) occupies the superior temporal plane posterior to Heschl's gyrus [15]. It represents auditory association cortex, being engaged in the analysis of many types of complex sound. In the left hemisphere, most definitions of Wernicke's area include part of PT [37] and the human PT has traditionally been viewed as a language processor [22,24]. Recently, Calvert et al. [6] found that lip reading activates the planum temporale (PT), which suggests that the PT may integrate visual to auditory cross-modal information. As auditory deprived subjects do not consistently show this activation pattern [21], audio-visual cross modal

---

[*] Corresponding author. Section of Cerebral Integration, Department of Cerebral Research, National Institute for Physiological Sciences, 38 Myodaiji, Okazaki, Aichi, 444-8585, Japan. Tel.: +81-564-55-7874; fax: +81-564-52-7913.

E-mail address: sadato@nips.ac.jp (N. Sadato).

0926-6410/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.cogbrainres.2004.04.005

integration during lip reading may be the result of a long-term learning process. Activation of the planum temporale has also been reported during reading of written language [25] and musical scores [24], both of which involve transforming visually presented symbols into the corresponding sounds of spoken language or music. As both language and score reading require specific training to attain proficiency, the observed PT activation may reflect learned audio-visual integration. To test this hypothesis, we used functional magnetic resonance imaging (fMRI) to image subjects during observation of key pressing finger movements on a piano keyboard, or "key-touch reading". This approach was motivated by the fact that well-trained pianists are able to identify pieces of music by watching the key-touching movements of the hands. We hypothesised that the visual information of sequential finger movements is transformed into the auditory modality during "key-touch reading" as during lip reading, and hence we predicted the activation of the PT during this task.

2. Materials and methods

2.1. Subjects

Twenty-six healthy volunteers, 20 women and 6 men, aged 22.3. ± 5.7 years (mean ± S.D.), were recruited for fMRI. Seven of them had never received piano training. The remaining 19 subjects had received piano training for 1 to 36 years (9.5 ± 7.6 years, mean ± S.D.), starting between the ages of 3 and 18 years (6.4 ± 3.7). Duration of the piano training was positively correlated with the performance of identification of the familiar music ($r = 0.764$, $P < 0.01$, $N = 19$). Hence we divided the piano trained group into two by the average duration of the training (8.0 years): a well trained group who had received at least 9 years of training (14.2 ± 8.6) and a less trained group who had received less than 9 years of training (5.2 ± 2.9). Between the two groups, there was no significant difference in the starting age (5.2 ± 4 years old for well trained, 7.5 ± 4.7 years old for less trained group). The two groups showed significant difference of their piano repertoire. Using the grading system of pianist's repertoire with 1 to 15 scale [34], which is the modified version of Hinson [18], the well-trained subjects were able to play more difficult piano pieces (11.2 ± 2.7, mean ± S.D.) than less-trained subjects (4.0 ± 3.8) ($p = 0.001$, Mann–Whitney test). The subjects were all right-handed according to the Edinburgh handedness inventory [26]. The protocol was approved by the ethical committee of Tohoku Fukushi University, and all subjects gave written informed consent for the study.

2.2. MRI

A time-course series of 130 volumes was acquired using T2*-weighted, gradient echo, echo planar imaging (EPI) sequences with a 1.5 tesla MR imager (Magnetom Vision Plus, Siemens, Erlangen, Germany). Each volume consisted of 19 slices, with a slice thickness of 4 mm without gaps, to include the entire cerebral cortex. The time-interval between two successive acquisitions of the same image was 3000 ms, and the echo time was 66 ms. The field of view (FOV) was 22 cm. The in-plane matrix size was 64 × 64 pixels with a pixel dimension of 3.44 × 3.44 mm.

For anatomical reference, T2-weighted fast spin echo images were obtained from each subject for the same slice locations used for the EPIs. In addition, high-resolution whole-brain MRIs were obtained with a conventional T1-weighted, gradient echo sequence.

2.3. Tasks

The task was to identify the names of pieces of music by observing piano playing finger movements. Prior to the fMRI session, subjects were informed that after the experiment they would be asked to name the pieces. An fMRI session consisted of alternating periods of task and rest (with a total of three task and three rest periods, each 60 s in duration). The visual stimulation was projected using a LCD projector (VL-1035SB, NEC, Tokyo, Japan) connected to a video player (HR-B9, JVC, Tokyo, Japan) onto a semi-transparent screen hung 10 cm from the subject's eyes. The screen was viewed by the subjects through a mirror. The visual angle of the screen was approximately 20°. It was confirmed that the subjects were not able to see their own both hands during the task. During task periods, subjects were required to view the hand movements of a piano player playing one piano piece per task period. The hand movements of the piano player were recorded from above (Fig. 1). During rest periods, subjects viewed the same hands sliding backwards and forwards across the piano keys without any key pressing movements of the fingers. No sound information was provided. In the familiar music session (FM), the viewed hands played three pieces of piano music that are typically learnt by all piano players in early piano lessons. In the unfamiliar music session (UM), the hand played three pieces of piano music which are now rarely performed. In the random movement session (RM), three patterns of random key presses were used. The location of the key-press was randomly selected for each hand. Prior to the experiment, we had confirmed that the random sequences did not contain any musical structure by listening to them. All subjects underwent fMRI in the order of FM, followed by UM and RM. After all fMRI sessions were completed, the subjects were asked to name the pieces presented during the sessions. For the RM session, the correct response was "the finger movements did not represent music".

2.4. Data analysis

The first 10 volumes of each fMRI session were discarded to allow for stabilisation of the magnetisation,

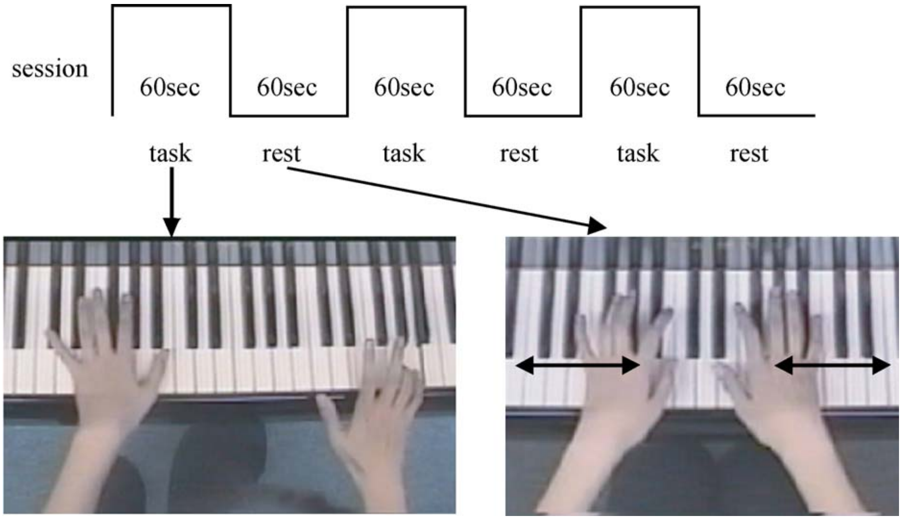

Fig. 1. Experimental setup. (Top) An fMRI session consisted of alternating periods of task and rest (six blocks in total, each 60 s in duration). (Bottom, left) During task periods, subjects were required to view the hand movements of a piano player playing piano pieces (one piece per task period), recorded from above. (Bottom, right) During the rest period, subjects viewed the same hands moving from side to side without key pressing movements of the fingers.

and the remaining 120 volumes per session, a total of 360 volumes per subject, were used for analysis. The data were analysed using statistical parametric mapping (SPM99, Wellcome Department of Cognitive Neurology, London, UK) implemented in Matlab (Mathworks, Sherborn, MA, USA) [9–11]. Following realignment, all images were anatomically normalized to a standard stereotaxic space [8], then filtered using a Gaussian kernel of 8 mm (full width at half maximum) in the x, y, and z axes.

2.5. Statistical analysis

Statistical analysis in the present study was conducted at two levels. First, individual task-related activation was evaluated. Second, the summary data for each individual were incorporated into the second level analysis using a random effects model [13] to make inferences at a population level.

2.6. Individual analysis

The signal was proportionally scaled by setting the whole-brain mean value to 100 arbitrary units. The signal time course for each subject was modelled using a box-car function convolved with a haemodynamic response function and temporally high pass filtered using a filter of 246 s width. To test hypotheses about regionally specific condition effects, estimates for each model parameter were compared with linear contrasts. First, we delineated the areas activated during the FM, UM and RM compared with the rest periods of the same session. The resulting set of voxel values for each contrast constituted a statistical parametric map (SPM) of the t statistic (SPM$\{t\}$). The SPM$\{t\}$ was transformed to the unit normal distribution (SPM$\{Z\}$). The threshold for SPM$\{Z\}$ was set at $P<0.05$ with a correction for multiple comparisons at the voxel level for the entire brain [12].

2.7. Group analysis using a random effects model

The weighted sum of the parameter estimates in the individual analyses constituted "contrast" images, which were used for the group analysis [13]. Contrast images obtained via individual analyses represent the normalised task-related increment of the MR signal of each subject. To evaluate the group effect and familiarity effect, a one-way ANOVA was performed using all contrast images, i.e., FM, UM, and RM of the well trained, less trained, and naïve groups. The threshold for SPM$\{Z\}$ was set at $P<0.05$ corrected for multiple comparisons at the cluster level with a threshold of $Z>3.09$ [12].

3. Results

3.1. Performance

Well-trained subjects were able to identify the familiar music more accurately than less-trained subjects (well trained: 77.8 ± 0.7% correct responses; less trained: 26.7 ± 0.8%, naïve: 0%). Unfamiliar pieces of music were not well identified by either trained group (3.3 ± 0.3% for less trained, 18.5 ± 0.9% for well trained, difference not statistically significant, naïve: 0%). Well-trained subjects were able to correctly identify random finger movements as such (88.9 ± 0.1%), whereas less-trained (50.0 ± 1.7%) or naïve subjects (28.6 ± 0.2%) were less accurate.

3.2. Group analysis

Task related neural activations for each group (well trained, less trained and naïve) are summarised in Fig. 2 and Table 1.

In the well trained group, viewing random finger movements activated the bilateral dorsal premotor cortex (PMd) extending to the inferior prefrontal cortex, superior parietal lobule (LPs), left middle occipital gyrus (GOm), superior temporal gyrus (GTs), thalamus and preSMA. During familiar and unfamiliar music conditions, similar activation patterns were observed. When activations were pooled across all three task conditions, the well-trained subjects showed activation in the bilateral superior and inferior parietal lobules, dorsal and ventral premotor cortex, occipito-parietal junction, thalamus, inferior frontal gyrus, SMA and the left planum temporale in the superior temporal gyrus (Fig. 2). These areas showed no significant gender effect or gender x group effect (two-way ANOVA). They showed significantly greater activation in the well trained group than the other groups (Fig. 3). The left planum temporale of the well-trained subjects ($x = -62$, $y = -34$, $z = 14$) was equally activated by familiar music, unfamiliar music and random movement, whereas this region was not strongly activated by any condition in the naïve and less trained groups (Fig. 4).

In the less trained group, weaker activation patterns were observed: viewing random finger movement activated the left LPi and LPs. During the familiar music condition, bilateral LPi and LPs, and left precentral gyrus were activated. During the unfamiliar music condition, bilateral LPs and right posterior GTs were activated. When pooled across all three conditions, activation was observed in the bilateral LPs, left LPi, and right GTm.

In the naïve group, viewing random finger movement did not elicit any significant activation. During the familiar music condition, left postcentral gyrus, right middle temporal gyrus (GTm), PMd and LPi were activated. During the unfamiliar music condition, left LPi, right PMd and posterior GTm were activated. Pooled across all the conditions, activation was observed in the left LPs and LPi, right PMd and right GTm.

Representative individual data for each group are shown in Fig. 5.

4. Discussion

4.1. Performance and expected neural activations

The performance measurement revealed that, as expected, the well-trained piano players were able to identify the pieces

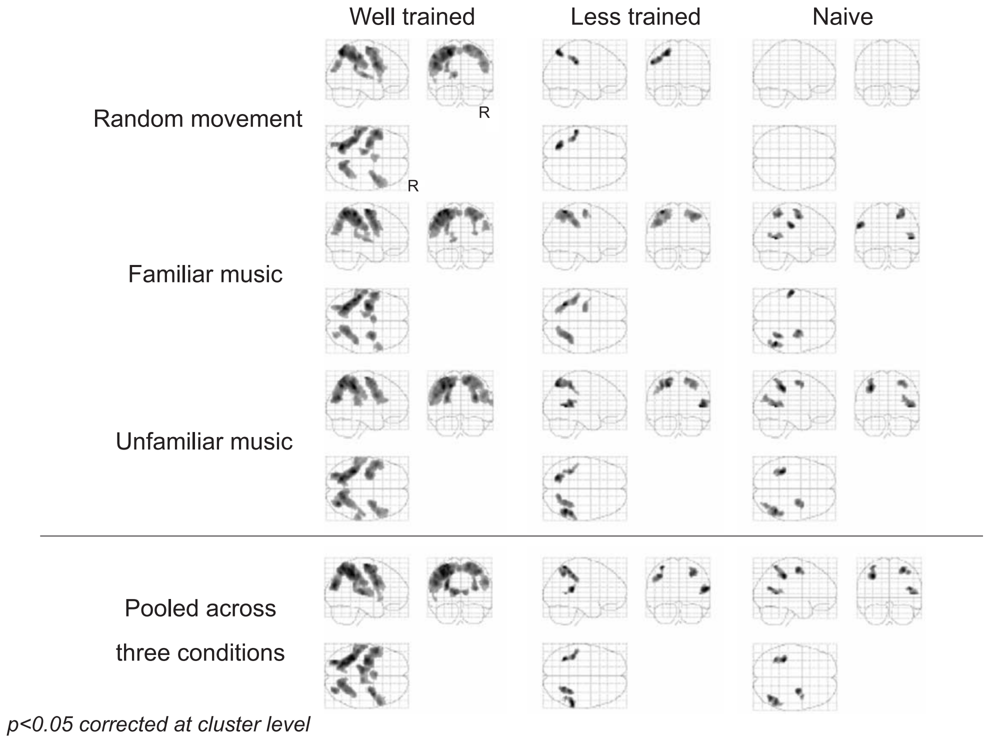

Fig. 2. Statistical parametric maps showing the group analysis of fMRI activity in well trained (left column), less trained (middle column), and naïve groups (right column), during random movement (first row), familiar music (second row), and unfamiliar music (third row). The results of pooling all conditions are shown in the bottom row. Each map is shown in standard anatomical space. The three-dimensional information was collapsed into two-dimensional sagittal, coronal and transverse images (i.e., maximum intensity projections viewed from the right, back, and top of the brain). The statistical threshold was $P < 0.05$, corrected for multiple comparisons at the cluster level.

Table 1
Task-related activation: pooled across all conditions

Cluster	Talairach co-ordinates						
Size	*Z*	*x*	*y*	*z*		Location	BA
Well trained							
3774	5.84	−36	−42	48	L	LPi	40
	5.63	−24	−62	62	L	LPs	7
	5.63	−28	−80	24	L	GO	19
	4.01	−62	−34	14	L	GTs	22
	3.75	−54	−28	26	L	LPi	40
1934	5.59	−32	−12	66	L	GPrC	6
	5.27	−52	6	30	L	GFi	44/6
	4.89	−52	2	42	L	GFm	9
	4.51	−44	−2	50	L	PMd	6
	3.41	−58	12	8	L	PMv	6
1161	5.29	40	0	54	R	PMd	6
	4.66	54	12	30	R	GFi	44
799	5.19	−12	−20	8	L	Thalamus	
	4.49	8	−22	12	R	Thalamus	
1486	5.18	18	−62	64	R	PCu	7
	4.21	40	−38	54	R	LPi	40
342	4.87	32	−80	18	R	GOm	19
297	4.47	−2	4	58	L	SMA	6
Less trained							
326	4.13	−34	−56	64	L	LPs	7
	3.83	−40	−36	44	L	LPi	40
264	4.11	58	−38	16	R	GTs	22
	3.48	46	−54	8	R	GTm	21
	3.11	50	−58	10	R	GTm	39
408	3.87	32	−44	52	R	LPs	7
Naïve							
424	4.63	−34	−36	44	L	LPi	40
	4.09	−34	−54	64	L	LPs	7
276	4.37	30	−8	56	R	GPrC	6
293	4.33	42	−64	18	R	GTm	39

GFi; inferior frontal gyrus, GFm; middle frontal gyrus, GO; occipital gyrus, GOm; middle occipital gyrus, GPrC; precentral gyrus, GTm; middle temporal gyrus, GTs; superior temporal gyrus, LPi; inferior parietal lobule, LPs; superior parietal lobule, PCu; precuneus, PMd; dorsal premotor cortex, PMv; ventral premotor cortex, SMA; supplementary motor cortex.

of music, whereas less-trained players and naive subjects were not. In accordance with the performance data, fMRI activation patterns were quite different across the groups. The well trained group revealed extensive activation during task performance, whereas less prominent activation was observed in the less trained and naïve groups. This difference may be attributable to the different mental processes that would take place during the tasks across the groups. For example, transformation of the visual information of sequential finger movements into the auditory information required to identify the pieces of music would occur more frequently in the well trained group. On the other hand, there was no significant difference in activation across the different task conditions in any subject group. Therefore, the activation differences between the well trained and other groups may be related to processes that are common to all three task conditions, e.g., transformation of visual input into the motor or auditory modalities. The following discussion will focus on the group differences in activation pooled across all task conditions.

4.2. Well-trained subjects only

4.2.1. Parieto–premotor–prefrontal network (including P–O junction)

The present study revealed that well-trained pianists activate the bilateral dorsal premotor cortex extending to the ventral premotor cortex, inferior frontal cortex, parietal cortex, and SMA during observation of hand movements.

It has previously been shown that area F5, the macaque homologue of Broca's area, is active during both execution and observation of movement. It has been proposed that activity of F5 is related to matching observed actions to similar, internally generated actions, forming a link between the observer and the actor, and hence constituting a 'mirror' system [30]. Human neuroimaging studies support this finding: the left inferior frontal cortex (BA 44) is active during imitation of hand movement [20] and bilateral BA 6 and 44 are somatotopically activated during observation of mouth, hand, and foot movements [5]. Furthermore, observing object related actions additionally activates parietal regions, prompting the suggestion that the mirror system may extend beyond the premotor cortex [5].

In the current study, well-trained pianists showed activation in the bilateral premotor–parietal network during all task conditions. To correctly identify pieces of music, online processing of structured sequences of key presses and integration of this information with the location of the fingers on the piano keyboard is essential. Schuboz and von Cramon [32] revealed that the complexity of structured movement sequences at a perceptual level was related to activity of the premotor cortex, and serial prediction of movements was related to activity in the premotor and parietal cortices. They argued that to predict how a target will move, the motor system generates a "blue-print" of the observed motion that is the basis for target motion prediction. Activation in these areas has also been related to the execution of complex sequential finger movement [17].

4.2.2. Planum temporale

As expected, activation in the posterior part of BA 22, posterior to the Heschl's gyrus, in a region corresponding to the planum temporale (PT), was seen in the well-trained subjects but not in the less-trained or naïve subjects. This finding suggests that activation of the PT by observation of hand motion reflects a learned process. McIntosh et al. [23] reported auditory to visual cross-modal likage through learning: the activation of the visual cortex by learning that an auditory stimulus signaled a visual event. Present study showed that the reverse may occur: visual to auditory cross-modal linkage is learning dependent. As the activation was elicited with or without a music component to the hand motion, PT activation may be related to processes that occur prior to identification of pieces of music. Furthermore, the left

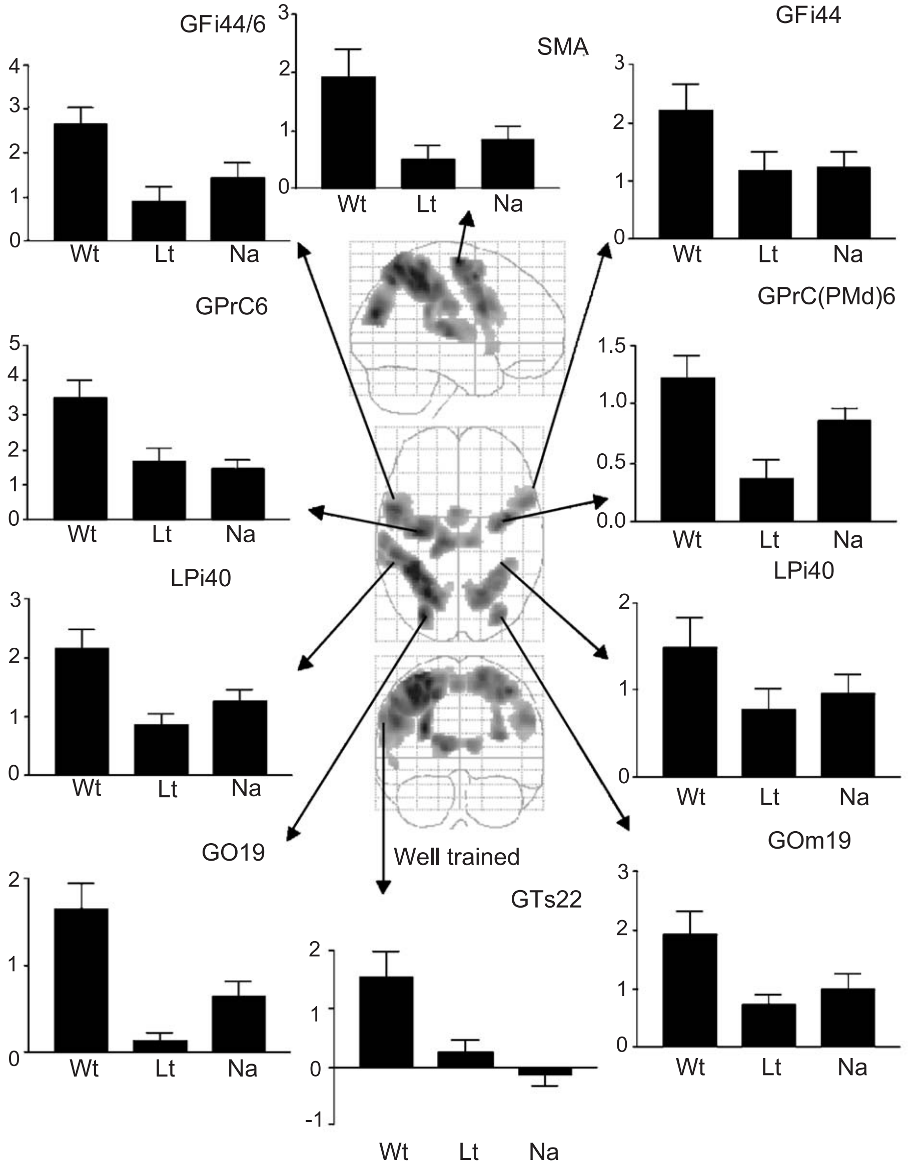

Fig. 3. Percent signal change in the well trained (Wt), less trained (Lt), and naïve (Na) groups. Regions were defined by pooling results across all conditions (Fig. 2, the bottom row.) in the well trained (upper middle column), and less trained (lower middle column) groups. GFi, inferior frontal gyrus; LPi, inferior parietal lobule; GO, occipital gyrus; GOm, middle occipital gyrus; GPrC, precentral gyrus; GTm, middle temporal gyrus; GTs, superior temporal gyrus.

lateralized activation may suggest special function of the left PT rather than "internal hearing of sounds".

Anatomically, the left PT is significantly larger, has greater cell size and density and more functionally distinct columnar systems per surface unit than the right PT [14]. Functionally, the left PT is involved in word detection and generation which have been attributed to its ability to process rapid frequency changes [2,33], whereas the right homologue is specialised for the discrimination of melody, pitch, and sound intensity [2,38]. Left lateralised activity in the PT may be related to the functional asymmetry noted in auditory processing. The left auditory cortical areas have a higher degree of temporal sensitivity, making them optimal for speech discrimination, which is highly dependent on detecting rapidly changing broadband sounds. Whereas the right auditory cortical areas, with greater spectral sensitivity, are optimal for processing the tonal patterns of music, in which small and precise changes in frequency are important [40]. Hence, activation of the left PT in the well-trained subjects may be related to coupling of the rapidly alternating position of the fingers with the sequence of sounds corresponding to the touched keys.

4.3. Lack of activation in the right GTs

There was no significant activation in the right GTs of the well-trained subjects. Based on previous studies [16,39] we had expected activation of the right GTs in the well-trained subjects during familiar or unfamiliar music conditions, as

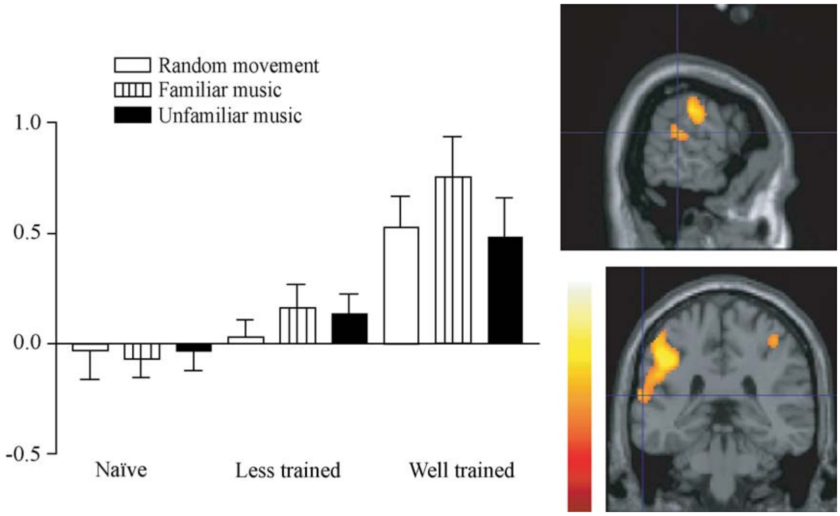

Fig. 4. There was prominent activation in the left planum temporale of the well-trained subjects during "key touch" reading tasks pooled across three conditions (random movement, familiar music, and unfamiliar music). The focus of activation is shown as a pseudocolour functional MRI superimposed on a high-resolution anatomical MRI in the sagittal (bottom left), and coronal (bottom right) planes, as indicated by the blue lines that cross at (− 62, − 34, 14) corresponding to the left planum temporale. The activity level (*t*-value) is as indicated by the colour bar, increasing from red to white. The statistical threshold was $P<0.05$, corrected for multiple comparisons at the cluster level. Top, the percent signal change in the left PT (− 62, − 34, 14) in naïve, less-trained and well-trained subjects during each task condition. During random movement without any music component, activation of the left PT of the well trained group was noted.

retrieval of melodies would be required for correct responses. Halpern and Zatorre [16] reported activation of the right GTs by imagining the continuation of a tune contrasted with re-imaging a sequence that had just been heard. This comparison segregates the retrieval of melody into distinct components. However, Halpern and Zatorre found no activation in the auditory cortices during the re-imaging task contrasted with listening only, instead SMA and other left frontal regions were activated by this comparison and were probably related to the motor code involved in the re-imaging task. In contrast, the present study revealed similar activation patterns across the task conditions with (familiar music, unfamiliar music) or without (random movement) retrieval of melody. This may be due to the fact that throughout the session, subjects were mainly focused on visual-to-auditory transformations rather than retrieval of auditory long-term memory.

4.4. Less trained or naïve groups only

4.4.1. Right MT

Less trained and naïve groups revealed activation in the right middle temporal cortex during the tasks. This area is

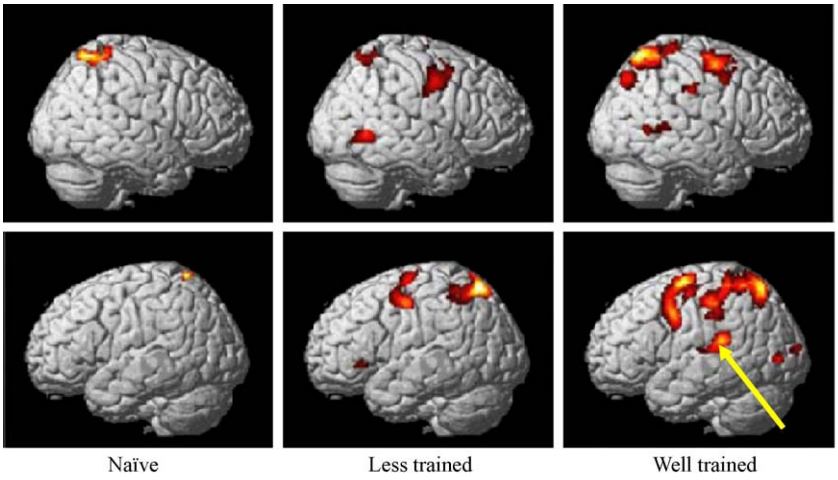

Fig. 5. Statistical parametric maps of individual analysis of fMRI activity in naïve, less trained and well trained groups. A representative case is shown for each group. Increased neuronal activition in naïve, less-trained and well-trained subjects during all tasks (familiar music, unfamiliar music, and random movement) compared with the control condition were superimposed onto surface-rendered high-resolution MRIs viewed from the right (top row) and left (bottom row). The statistical threshold is $P<0.05$, corrected for multiple comparisons at the cluster level. The yellow arrow indicates activation in the left planum temporale of a well-trained subject during the tasks.

close to the posterior rami of the superior temporal gyrus and anterior to the well-documented V5 area. The pattern of activity in the naïve and less trained groups evoked during all task conditions is consistent with the proposed neural substrates subserving the perception of human movement [7]. Activity was observed predominantly in the right V5 and posterior STS regions. The right occipitotemporal junction was activated during all tasks in the naïve and less trained groups. This area is coextensive with the presumed human area V5. Human area V5, homologue to monkey MT/V5, is known to be specifically involved in motion perception [3,19,35,36]. Thus, the activation of V5 is likely due to perception of the movement of the hands.

Perception of movement of the human body consistently activates regions in the posterior STS [3,19,29]. Viewing a representation of a man running compared with random dot motion bilaterally activates the GTs close to the Vpc line [19]. Puce et al. [29] found that viewing eye and non-linguistic mouth movement bilaterally activates a region centred in the posterior STS, posterior to the Vpc line. Bonda et al. [3] showed that the caudal STS, far posterior to the Vpc line, is activated while perceiving hand motion. The upper bank of the rostral STS of the macaque monkey is called the superior temporal polysensory (STP) area, because the neurons there respond to auditory, somatosensory, and visual stimuli [31]. Many of the visual responses in the anterior STP involve motion [4], especially movements of the human body [4,27,28]. Visual information from two distinct pathways processing motion and form are integrated in the anterior STP to establish a coherent representation of the object and direction of motion [27]. Thus, the posterior STS in the human brain may correspond to area STPa in the monkey [7]. These findings suggest that the function of the posterior STS is related to perception of motion of the hands. This area is known to be activated by observation of biological motion in which the form of an object changes [1]. This is consistent with the present finding, as sequential finger movement together with sideways movements of the hands occurred during the task conditions, whereas sideways movements only were shown during the control condition.

Why was the posterior STS not activated by the well-trained subjects? This may be related to associations between sounds and key-presses as opposed to hand movements. As the piano sound is associated with the location of the tapped key rather than the hand movement in which the form of the hand changes, well-trained subjects may have been attending to the spatial and temporal sequence of the tapped keys which then was converted into the temporal sequence of sounds through the parieto-premotor ‘mirror system’ and the left PT. Hence, well-trained subjects are keyboard oriented, whereas less-trained and naïve subjects may be hand-oriented.

In conclusion, the left PT of the trained pianist was activated by “piano key touch reading.” This may represent the neural substrate of association formation between sequential visual codes and the corresponding sounds of pieces of music.

Acknowledgements

This study was in part supported by Grants-in-Aid for Scientific Research (B, No14380370), and the Academic Frontier Promotion Project Foundation (1998–2002) from the Ministry of Education, Science and Culture, Japan.

References

[1] M.S. Beauchamp, K.E. Lee, J.V. Haxby, A. Martin, Parallel visual motion processing streams for manipulable objects and human movements, Neuron 34 (2002) 149–159.

[2] P. Belin, S. McAdams, B. Smith, S. Savel, L. Thivard, S. Samson, Y. Samson, The functional anatomy of sound intensity discrimination, J. Neurosci. 18 (1998) 6388–6394.

[3] E. Bonda, M. Petrides, A. Evans, Neural systems for tactual memories, J. Neurophysiol. 75 (1996) 1730–1737.

[4] C. Bruce, R. Desimone, C.G. Gross, Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque, J. Neurophysiol. 46 (1981) 369–384.

[5] G. Bucchino, F. Binkofski, G.R. Fink, L. Fadiga, L. Fagossi, V. Gallese, R. Seitz, K. Zilles, G. Rizzolatti, H.-J. Freund, Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study, Eur. J. Neurosci. 13 (2001) 400–404.

[6] G.A. Calvert, E.T. Bullmore, M.J. Brammer, R. Campbell, S.C. Williams, P.K. McGuire, P.W. Woodruff, S.D. Iversen, A.S. David, Activation of auditory cortex during silent lip reading, Science 276 (1997) 593–596.

[7] J. Decety, J. Grezes, Neural mechanisms subserving the perception of human actions, Trends Cogn. Sci. 3 (1999) 172–178.

[8] A.C. Evans, M. Kamber, D.L. Collins, D. MacDonald, An MRI-based probabilistic atlas of neuroanatomy, in: S.D. Shorvon (Ed.), Magnetic Resonance Scanning and Epilepsy, Plenum, New York, 1994, pp. 263–274.

[9] K.J. Friston, K.J. Worsley, R.S.J. Frackowiak, J.C. Mazziotta, A.C. Evans, Assessing the significance of focal activations using their spatial extent, Hum. Brain Mapp. 1 (1994) 210–220.

[10] K.J. Friston, J. Ashburner, C.D. Frith, J.D. Heather, R.S.J. Frackowiak, Spatial registration and normalization of images, Hum. Brain Mapp. 2 (1995) 165–189.

[11] K.J. Friston, A.P. Holmes, K.J. Worsley, J.B. Poline, C.D. Frith, R.S.J. Frackowiak, Statistical parametric maps in functional imaging: a general linear approach, Hum. Brain Mapp. 2 (1995) 189–210.

[12] K.J. Friston, A. Holmes, J.-B. Poline, C.J. Price, C.D. Frith, Detecting activations in PET and fMRI: levels of inference and power, NeuroImage 4 (1996) 223–235.

[13] K.J. Friston, A.P. Holmes, K.J. Worsley, How many subjects constitute a study? NeuroImage 10 (1999) 1–5.

[14] R.A.W. Galuske, W. Schlote, H. Bratzke, W. Singer, Interhemispheric asymmetries of the modular structure in human temporal cortex, Science 289 (2000) 1946–1949.

[15] T.D. Griffiths, J.D. Warren, The planum temporale as a computational hub, Trends Neurosci. 25 (2002) 348–353.

[16] A.R. Halpern, R.J. Zatorre, When that tune runs through your head: a PET investigation of auditory imagery for familiar melodies, Cereb. Cortex 9 (1999) 697–704.

[17] D.L. Harrington, S.M. Rao, K.Y. Haaland, J.A. Bobholz, A.R. Mayer, J.R. Binder, R.W. Cox, Specialized neural systems underlying representations of sequential movements, J. Cogn. Neurosci. 12 (2000) 56–77.

[18] M. Hinson, Guide to The Pianist's Repertoire, 3rd ed., Indiana Univ. Press, Bloomington, 1987.

[19] R.J. Howard, M. Brammer, I. Wright, P.W. Woodruff, E.T. Bullmore, S. Zeki, A direct demonstration of functional specialization within motion-related visual and auditory cortex of the human brain, Curr. Biol. 6 (1996) 1015–1019.

[20] M. Iacoboni, R.P. Woods, M. Brass, H. Becckering, J.C. Mazziotta, G. Rizzolatti, Cortical mechanisms of human imitation, Science 286 (1999) 2526–2528.

[21] M. MacSweeney, G.A. Calvert, R. Campbell, P.K. McGuire, A.S. David, S.C. Williams, B. Woll, M.J. Brammer, Speechreading circuits in people born deaf, Neuropsychologia 40 (2002) 801–807.

[22] J.C. Marshall, Planum of the apes: a case study, Brain Lang., (2000) 145–148.

[23] A. McIntosh, R.E. Cabeza, N.J. Lobaugh, Analysis of neural interactions explains the activation of occipital cortex by an auditory stimulus, J. Neurophysiol. 80 (1998) 2790–2796.

[24] T. Nakada, Y. Fujii, K. Suzuki, I.L. Kwee, 'Musical brain' revealed by high-field (3 Tesla) functional MRI, NeuroReport 19 (1998) 3853–3856.

[25] T. Nakada, Y. Fujii, Y. Yoneoka, I.L. Kwee, Planum temporale: where spoken and written language meet, Eur. Neurol. 46 (2001) 121–125.

[26] R.C. Oldfield, The assessment and analysis of handedness: the Edinburgh inventory, Neuropsychologia 9 (1971) 97–113.

[27] M.W. Oram, D.I. Perrett, Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey, J. Neurophysiol. 76 (1996) 109–129.

[28] D.I. Perrett, P.A. Smith, A.J. Mistlin, A.J. Chitty, A.S. Head, D.D. Potter, R. Broennimann, A.D. Milner, M.A. Jeeves, Visual analysis of body movements by neurones in the temporal cortex of the macaque monkey: a preliminary report, Behav. Brain Res. 16 (1985) 153–170.

[29] A. Puce, T. Allison, M. Asgari, J.C. Gore, G. McCarthy, Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study, J. Neurosci. 16 (1996) 5205–5215.

[30] G. Rizzolatti, A. Arbib, Language within our grasp, Trends Neurosci. 21 (1998) 188–194.

[31] K.S. Saleem, W. Suzuki, K. Tanaka, T. Hashikawa, Connections between anterior inferotemporal cortex and superior temporal sulcus regions in the macaque monkey, J. Neurosci. 20 (2000) 5083–5101.

[32] R.I. Schuboz, D.Y. von Cramon, A blueprint for target motion: fMRI reveals perceived sequential complexity to modulate premotor cortex, NeuroImage 16 (2002) 920–935.

[33] J.H. Schwartz, P. Tallal, Rate of acoustic change may underlie hemispheric specialization for speech perception, Science 207 (1980) 1380–1381.

[34] A.Takahashi, Dictionary of Repertoire for Pianist. Shunju-sha, Tokyo (1988) (in Japanese).

[35] R.B.H. Tootel, J.B. Reppas, K.K. Kwong, R. Malach, R.T. Born, T.J. Brady, B.R. Rosen, J.W. Belliveau, Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging, J. Neurosci. 15 (1995) 3215–3230.

[36] J.D.G. Waston, R. Myers, R.S.J. Frackowiak, J.V. Hajnal, R.P. Woods, J.C. Mazziota, S. Shipp, I.S. Zek, Area V5 of the human brain: evidence from a combined study using positron emission tomography and magnetic resonance imaging, Cereb. Cortex 3 (1993) 79–94.

[37] R.J. Wise, S.K. Scott, S.C. Blank, C.J. Mummery, K. Murphy, E. Warburton, Separate neural subsystems within 'Wernicke's area', Brain 124 (2001) 83–95.

[38] R.J. Zatorre, A.C. Evans, E. Meyer, Neural mechanisms underlying melodic perception and memory for pitch, J. Neurosci. 14 (1994) 1908–1919.

[39] R.J. Zatorre, A.R. Halpern, D.W. Perry, E. Meyer, A.C. Evans, Hearing in the mind's ear: a PET investigation of musical imagery and perception, J. Cogn. Neurosci. 8 (1996) 29–46.

[40] R.J. Zatorre, P. Belin, V.B. Penhune, Structure and function of auditory cortex: music and speech, Trends Cogn. Sci. 6 (2002) 37–46.