Human brain activities relating to mouth movements perception

Kensaku Miki

Doctor of Philosophy

Department of Physiological Sciences
School of Life Sciences
The Graduate University for Advanced Studies

2003

Abbreviations

ALITS	Ascending limb of ITS
ANOVA	Analysis of variance
BESA	Brain electric source analysis
ECD	Equivalent current dipole
EEG	Electroencephalography
EOG	Electrooculogram
ERP	Event related potential
FFT	Fast Fourier transform
fMRI	Functional magnetic resonance imaging
fT	femto Tesla
fT/cm	femto Tesla per centimeter
GoF	Goodness of fit
HG	Heschl's gyrus
ITS	Inferior temporal sulcus
MEG	Magnetoencephalography
MST	Middle superior temporal area
MT	Middle temporal area
MTG	Middle temporal gyrus
nAm	nano ampere meter
PA	Pre-auricular points
PCA	Principal component analysis
PCITS	Posterior continuation of ITS
PET	Positron emission tomography

PLSD	Protected least significant difference
RMS	Root mean square
RV	Residual variance
SQUID	Superconducting quantum interference device
STG	Superior temporal gyrus
STPa	Anterior superior temporal polysensory area
STS	Superior temporal sulcus
V1	Primary visual cortex
V5	5^{th} visual cortex

Abstract

We used magnetoencephalography (MEG) to investigate how the human brain is activated during perception of mouth movement by performing the following two studies:

(1) The temporal and spatial characteristics of neural responses elicited by viewing mouth movements.

(2) The influence of viewing mouth movement while simultaneously hearing speech sound on auditory cortical activity.

In the first study, we focused on differences in responses to mouth opening and closing movements by apparent motion, using an averted eyes condition as a control. A large clear MEG component, 1M (mean peak latency of approximately 160 ms) was elicited by both mouth movements. The MEG component in the left hemisphere was significantly smaller than 1M during eye aversion. We modeled the neural sources using the brain electric source analysis (BESA) method and placed sources around: (1) the occipito-temporal border, at human MT/V5, (2) primary visual cortex (V1), and (3) fusiform gyrus. The calculated activity of Source (1) was large while the activity of others was small or negligible. Source (1), as calculated separately for mouth closing and opening movements, showed similar locations, compared to those during eye aversion. Our results indicate that human MT/V5 is active in the perception of both mouth and eye motion. Viewing mouth movements elicits smaller MEG responses relative to that elicited by eyes movement; however, responses to mouth opening and closing movements did not differ in size. These data suggest

that MT/V5 may process human motion differently from the superior temporal sulcus (STS).

In the second study, we used speech movement to investigate temporal and spatial characteristics of the M100 component, which is generated in the auditory cortex peaking at around 100 ms following auditory stimulation. There was no significant difference in M100 latency and amplitude between auditory stimulus alone and paring of auditory and visual motion stimulus. The estimated sources of M100 from auditory stimulus only and form paring of auditory and visual motion stimulus were located in the Heschl's gyrus (HG), the auditory cortex, and there was no significant difference in the location and moment of estimated sources. This finding shows the auditory cortex processes the characteristics of auditory stimuli without any influences from visual perception of motion.

General introduction

In daily life, it is very important to perceive the mouth movement. For example, we have to see and perceive mouth movement to recognize facial expression and speech of the other persons. Actually, when we talk with other persons, we naturally focus on their mouth movement as well as eye movement (or gaze). In psychological study, McGurk and MacDonald (1976) reported that we usually perceive /da/ when auditory stimulus /ba/ and visual stimulus /ga/ are presented simultaneously. It is suggested that the integration of visual and auditory stimuli is a very important aspect of speech perception. Therefore, it seems very important and interesting to understand how the human brain is activated by viewing mouth movement. However, only a few studies have reported using functional magnetic resonance imaging (fMRI) (Puce et al. 1998) and event-related potentials (ERP) by averaging electroencephalography (EEG) (Puce et al. 2000).

In this study, we chose MEG as a neuroimaging modality because of its ability to study temporal processing of information to millisecond accuracy, which fMRI does not have, and advantages when localizing active neuronal sources, because spatial smearing effects caused by cerebrospinal fluid, skull and scalp are minimal, which EEG does not have.

Therefore, the main objective of this study was to thoroughly investigate human brain activities relating to mouth movement perception using MEG. For this objective, we performed the following two studies;

(1) The temporal and spatial characteristics of neural responses elicited by

viewing mouth movements.

(2) The influence of viewing mouth movement while simultaneously hearing speech sound on the auditory cortex activity.

Experiment 1

Magnetoencephalographic study of occipitotemporal activity elicited by viewing mouth movements

Introduction

Recent neuroimaging studies examining brain responses to viewing the actions of others indicate that in addition to such regions as the superior temporal sulcus (STS) and middle temporal gyrus (MTG), MT/V5 also play a prominent role in processing these complex motion stimuli (Bonda et al 1996; Puce et al 1998; Watanabe et al 2001; Kourtzi & Kanwisher 2000). Particularly intriguing is the finding that MT/V5 is active when *static* images depicting implied motion are presented (Kourtzi & Kanwisher 2000). The suggestion that MT/V5 exhibits specialization, in addition to general motion processing, is not new (Toottell et al 1995b; Watanabe et al 2001), however, the exact nature of this specialization is not known. Previously, fMRI activity to observing eye movements over and above that seen to motion in general has been reported in MT/V5 (Puce et al 1998), leading to speculation that specialized cortical regions for movement of facial parts might be present in MT/V5 (Watanabe et al 2001). This activity occurs at around 200 ms post-motion onset, as indicated by both ERP (Puce et al 2000, 2003) and MEG studies (Watanabe et al 2001).

The idea that there may be multiple motion sensitive regions in the human brain that deal with processing facial movements makes sense given that humans are social primates and routinely 'read' facial expressions and movements of others. Additionally, it is known that information relating to faces is processed by both ventral and lateral temporal regions (Puce et al., 1995, 1999; Allison et al., 1999; McCarthy et al., 1999; Watanabe et al., 2003).

Previous MEG studies including our own demonstrate that reliable responses from ventral and lateral temporal cortex can be elicited to the onset of static faces (Lu et al., 1991; Sams et al., 1997; Swithenby et al., 1998; Linkenkaer-Hansen et al., 1998; Watanabe et al., 1999a&b, 2003; Sato et al., 1999; Halgren et al., 2000; Ioannides et al., 2000; Nakamura et al., 2001; Taylor et al., 2001; Terasaki et al., 2002). Furthermore, we have recorded robust MEG responses to viewing facial eye movements (Watanabe et al., 2001) at around 170 ms post-motion onset (1M). Although our equivalent current dipole (ECD) of 1M was located around MT/V5, consistent with previous MEG data on general motion perception (e.g. Bundo et al., 2000; Kawakami et al., 2002), the ECD to eye movements was more inferior and posterior than that to motion in general (Watanabe et al., 2001).

In this study, we studied both temporal and spatial characteristics of MEG responses elicited by viewing mouth movements (opening and closing), as compared to viewing control movement types such as an eye aversion movement and motion in general. Puce *et al.* (2000 and 2003) and Wheaton *et al.* (2001) have previously reported that ERPs to mouth opening and closing occur at around 170ms post-motion onset (N170), and that N170s to mouth opening are in general larger than those seen to mouth closing. The N170 activity occurred over the bilateral posterior temporal scalp, and could conceivably have been generated by neural sources in MT/V5 and in the STS. In these studies the neural sources of the ERP activity were not localized. In fMRI and electroencepharography (EEG)

studies, either temporal or spatial characteristics were mainly studied. So, in this MEG study, we aimed both to localize the sources for neural activity elicited to two types of mouth movements and to prove the time course of the sources compared to our control conditions. Therefore, as well as the previous study (Watanabe et al., 2001), we used apparent motion, which is perceived by the same mechanism as real motion (e.g. Kaneoke et al., 1997).

Methods

Subjects

We studied seventeen right-handed normal volunteers (4 females, 13 males) ranging in age from 24 to 43 years (mean age, 32.2 years) with normal or corrected visual acuity. All subjects gave informed consent to participate in the experiment, which was approved by the Ethical Committee at the National Institute for Physiological Sciences.

Visual stimulation

We presented stimuli previously used in ERP studies (Puce et al., 2000) in an experimental paradigm that was modified slightly for MEG studies (Fig. 1). The stimulus consisted of a color facial image, superimposed on a radial pattern of three concentric black, white, and gray rings. We used four stimulus types:

(1) *Face with closed mouth*: The mouth is closed and eyes gaze straight at the observer and the radial background remains unchanged.

(2) *Face with open mouth*: This stimulus is similar to (1) with the only difference being that the mouth is open.

(3) *Face with deviated eyes*: This stimulus is similar to (1) with the only difference being that the eyes deviate to the observer's right.

(4) *Changed background*: The face remains unchanged while the radial background's rings were altered by replacing a ring's color with that of its outer neighbor.

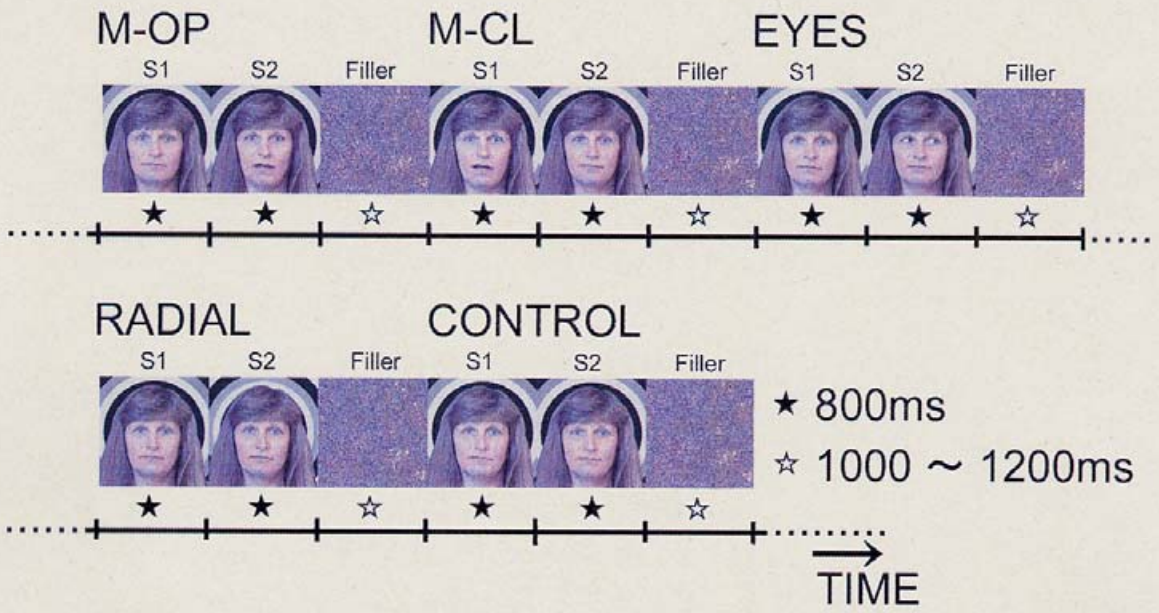


Fig. 1: Examples of five stimulus conditions and their timing during the experiment.

We used these four stimulus types in combination to construct a series of apparent motion conditions, where the first stimulus, S1, was replaced by a second stimulus, S2, with no inter-stimulus interval (Fig. 1):

M-OP (Mouth opening): S1 was *Face with closed mouth* and S2 was *Face with open mouth.*

M-CL (Mouth closing): S1 was *Face with open mouth* and S2 was *Face with closed mouth.*

EYES (Eye aversion): S1 was *Face with closed mouth* where the eyes gazed directly at the observer and S2 was *Face with deviated eyes.*

RADIAL (General motion): For example, S1 was *Face with closed mouth* and S2 was *Changed background.* In this condition, the face remained unchanged, but the observer viewed an inwardly moving radial stimulus.

CONTROL: S1 and S2 were the same e.g. *Face with closed mouth*

presented twice in succession. This condition was used to confirm that MEG responses were generated by viewing movements and not by stimulus onset.

A sixth stimulus type (Fig. 1: Filler) was presented between stimulus conditions, so as to avoid large luminance and contrast changes during the experiment. This interval stimulus consisted of a scrambled image of the face on the radial background. It was generated by taking a two-dimensional Fast Fourier transform (FFT) of the face image on the radial background, randomizing its phase spectrum, and taking an inverse FFT. This produced an image without recognizable form, but with comparable luminance, contrast and spatial frequency to that of the face in radial background stimulus. Hence, there were no *overall* luminance and contrast differences between S1, S2 and the Filler.

In all five conditions, S2 followed S1 with no interstimulus interval. S1 was shown for 800 ms, as was S2. The five stimulus conditions (consisting of S1+S2 combinations) were presented randomly during the course of the experiment. The filler stimulus was presented for a random interval of 1,000-1,200 ms between each stimulus condition, producing a period between successive trials of 2,600-2,800 ms.

All subjects reported experiencing a percept of motion from the four apparent motion conditions, but not for the CONTROL. By using an interval stimulus with equal overall luminance to the five stimulus conditions, onset responses to S2, predominantly generated in earlier visual regions, could be minimized. In addition, as subjects saw a continuously

present face between S1 and S2, activity in the fusiform gyrus, known to occur to face onset, could be minimized. Therefore, cortical responses to movement could be predominantly isolated.

Stimuli were presented by a personal computer (PC, IBM) and video projector (LP-9200, Sanyo, Japan) housed outside of a magnetically shielded room (Vacuumschmerze GmbH, Germany). Mean illuminance of the room was 0.2cd/m^2. The time delay for sending the image between the PC and drawing the image on the projector was monitored by oscilloscope. The mean value was 11.0 ms and hence we corrected the zero time point by subtracting 11.0 ms from the timepoints in the recorded waveforms.

Stimuli were projected on the ceiling using a mirror system. Subjects were lay supine on a bed, with head resting on the probe. The distance between the subject's eyes and the display was 148 cm. Stimuli were projected centrally, and subtended a visual angle of 11.6 x 11.6 degrees. Subjects were asked to maintain their gaze at a point at the top of the nose and between the eyes of the stimulus face. The mean luminance of the center (fixation) point of the face was 14.0cd/m^2.

MEG recording

MEG was measured with a 37-channel biomagnetometer (Magnes, Biomagnetic Technologies Inc., San Diego, CA). The detection coils of the biomagnetometer were arranged in a uniformly distributed array in concentric circles over a spherically concave surface. Thus, all of the

sensor coils were equally sensitive to weak magnetic signals from the brain. The device was 144 mm in diameter and its radius of the curvature was 122 mm. The outer coils were 72.5^0 apart. Each coil was 20 mm in diameter, and the centers of the coils were 22 mm apart. Each coil was connected to a superconducting quantum interference device (SQUID).

Subjects lay supine on the bed and the MEG probe was positioned at the back of the subject's head overlying the occipito-temporal area of one hemisphere (centering around T6 and T5 (10/20 system) in the right and left hemisphere, respectively) in all subjects (Fig. 2(a)), with a separate recording from the other hemisphere also made in the same recording session as well as our previous studies (Watanabe et a. 1999a, b, 2001, 2003). To minimize head movements we taped the subject's forehead to the probe with surgical adhesive tape. The probe sampled activity from primary visual cortex and the occipito-temporal junction. The left and right hemispheres were studied independently, with counterbalanced order across subjects. All subjects were studied with this configuration. Puce et al. (1998), using fMRI, reported that both MT/V5 and STS were activated by viewing mouth movements. Activated region of STS was located approximately 20mm more anterior, 10 mm more lateral, and 4 mm upper in mean than that of MT/V5. Although the standard position in this study could completely cover such activated site in STS recorded by fMRI (Puce et al., 1998), we also placed MEG probe in more anterior site (Fig. 2(b)), mainly STS for 3 subjects who showed reliable BESA results and 1 representative subject who didn't show reliable MEG response to know

activity in more anterior site. Additionally, we also recorded MEG by placing the probe approximately 3 cm anterior to the standard configuration overlying the right STS in 4 subjects in a subsequent recording session (Fig. 2(b)). Three of these were representative subjects with robust MT/V5 reliable MEG responses and the fourth was a subject who showed negligible activity in MT/V5 despite a technically adequate study. This indicates that the occipito-temporal area mainly activates by four conditions (M-OP, M-CL, EYES, and RADIAL) but not STS. So, in this study, the MEG probe was positioned at the location of Fig. 2(a).

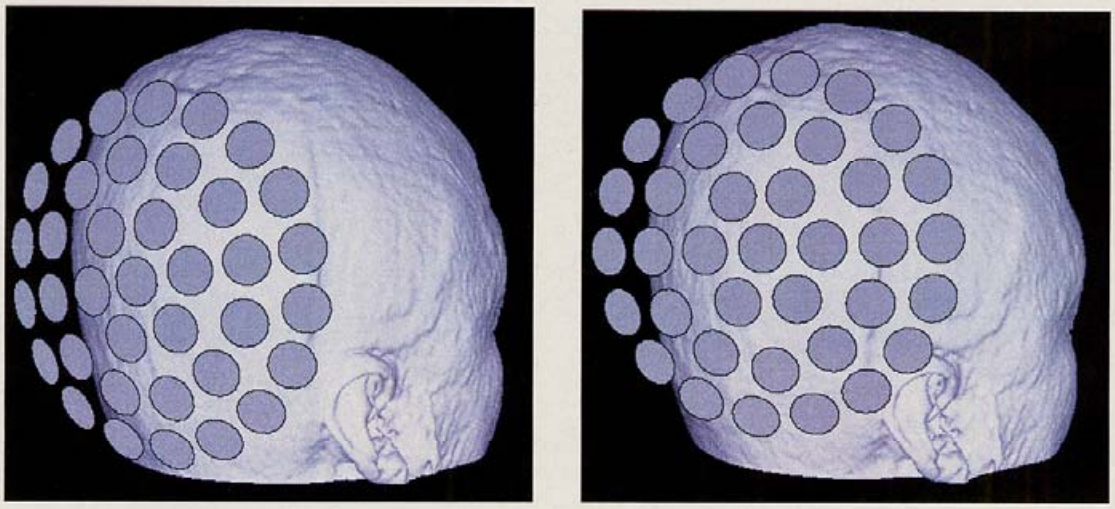


Fig. 2: MEG sensor placement. (a) A standard sensor placement for a right hemisphere study is shown, centered over the occipito-temporal cortex, centering around T6 (10/20 system). (b) An extra placement of the sensor probe, which was used for 4 representative subjects. The center of the probe was placed approximately 3 cm in front of a standard position to record activity in more anterior sites, focusing mainly on the STS.

To minimize subject habituation and drowsiness, we used 10 short-term recording sessions for each subject. Each recording session

included 50 trials for five stimulus conditions i.e. 10 trials for each condition, with a total duration of 130-140 sec (2600-2800 ms x 50 trials) for each recording. A short break of around 2 minutes occurred between each recording session.

MEG and vertical and horizontal electrooculograms (EOGs) were simultaneously recorded with a bandpass of 0.1-50 Hz and digitized at a sampling rate of 520.8 Hz. Epochs in which signal variations were larger than 3pT in the MEG and 80μV in EOG were excluded from the averages. After finishing 10 recording sessions and excluding epochs with large artifacts, we averaged trials for each condition separately.

The analysis window of 1,500 ms was divided into two sections: 800 ms after S1 onset and 700 ms after S2 onset. A 300 ms pre-stimulus baseline was used for responses to S1 and S2. Amplitude of recognizable components was measured as a root mean square (RMS) value across the 37 channels of averaged response data in the order of fT. Peak latency was measured at the point with the largest RMS at visible peaks of each component.

Data analysis

We used a multi-dipole model, BESA (Scherg and Bucher, 1993) (Neuroscan, McLean, VA), computation of theoretical source generators in a three-layer spherical head model. The BESA was modified for the use of our 37-channel biomagnetometer. This allows multiple sources to activate simultaneously for defined intervals. The location and orientation were

calculated by an interactive least-square fit. This method allows the spatio-temporal modeling of multiple simultaneous sources over defined intervals. Before commencing modeling, a principal component analysis (PCA) was applied to determine how many principal components explain the significant variance. We then accepted the model of dipoles using two criteria: 1.The residual variance (% RV) indicated the percentage of data that could not be explained by the model. The goodness of fit (GOF) was expressed as a percentage (100 - % RV). GOF values larger than 85% were defined as indicating adequate multiple dipole models (see Watanabe et al., 1999a&b, 2003). An increase in the dipole number mathematically increases the GOF, since a greater number of dipoles will account for more variance (Watanabe et al., 1999a&b, 2003), however, the generated solutions may not always be physiologically plausible. 2. Sources estimated in grey matter after overlaying on MRI.

For four stimulus conditions (M-OP, M-CL, EYES and RADIAL) except for CONTROL, taking our previous studies (Watanabe et al., 1999a&b, 2001, 2003) and other neuroimaging studies (e.g. Puce et al., 1995, 1998) into account and calculated by PCA, we made a 4-source model as follows; Source 1: the occipito-temporal junction, MT/V5 homologue in humans, Source 2: left primary visual cortex (V1), Source 3: right V1, Source 4: the fusiform gyrus. Our MEG sensor locations in the present study covered V1 bilaterally, as well as the lateral occipitotemporal cortex in one hemisphere. Initially, we placed and each of the 4 sources around each corresponding region. The BESA calculation allows some

change in the initial location and freedom in the orientation of each source on PCA, so it is possible for each source to move from initial placement to a near by location, if better fit to the data result. We have computed the best location and orientation of each source separately in each condition. Finally, we also analyzed the data using 5 sources, by also including one source in the STS (Puce et al. 1998) and analyzed results in all conditions.

The origin of the coordinate system was the point exactly halfway between the pre-auricular points (PA). The x-axis was a line extending through right PA and left PA, with positive values toward the right PA. The y-axis was a line extending through the origin and nasion, with positive values emerging from the head at the nasion. The z-axis was a line extending through the origin and the vertex (Cz of the 10-20 International System), with positive values occurring dorsally. This axis was perpendicular to the plane formed by the left and right PA and nasion.

We used an analysis of variance (ANOVA) with post-hoc tests of Fisher's protected least significant difference (PLSD), or paired t-tests ($P<0.05$) to assess significant differences between conditions.

MRI overlay

T1-weighted contiguous coronal, axial and sagittal magnetic resonance (MR) images with a 1.5 mm slice thickness were acquired on a Magnex 150XT 1.5T system (Shimadzu, Kyoto, Japan). The anatomical landmarks used for the MEG head-based 3D coordinate system (nasion and entrance of the auditory meatus of the left and right ear) were tagged with

high-contrast cod liver oil capsules (3 mm diameter) whose short relaxation time provides a high-intensity signal in T1-weighted images. The common MEG and MR anatomical landmarks enabled a straightforward transformation of the head-based 3D coordinate system used by the MEG source analysis to the MRI. Sources locations calculated by BESA converted into pixels using MRI transformation matrix and over laid onto corresponding MRI.

Results

The number of averaged trials per subject was 95.4 ± 2.9, 94.4 ± 3.0, 94.7 ± 3.3, 94.6 ± 3.6 and 94.7 ± 4.0 for M-OP, M-CL, EYE, RADIAL and CONTROL, respectively.

We analyzed MEG activity following both S1 and S2 independently. We present the results for S2, or motion onset, in detail as this was the focus of our study.

1. MEG activity following S2

The Right hemisphere

(1) Waveform characteristics

In all subjects S2 onset elicited the largest MEG activity in the right hemisphere. The most prominent component, 1M, was observed in all conditions with apparent motion (M-OP, M-CL, EYES and RADIAL) in 12 out of 17 subjects (Fig. 3). 1M peak latency, which was measured at the point with the largest RMS at visible peaks of each component, was similar in the face motion conditions (Table 1), peaking around 160 ms, and was around 20 ms shorter for RADIAL. One-way ANOVA showed a significant main effect for stimulus type ($F=4.35$; $P<0.05$). Post-hoc paired comparisons using Fisher's PLSD confirmed that the latency to RADIAL was significantly shorter than that to the facial motion conditions ($P<0.05$). There were no significant differences in 1M latency between the two mouth motion types (M-OP and M-CL), nor did they differ relative to EYES (see Table 1).

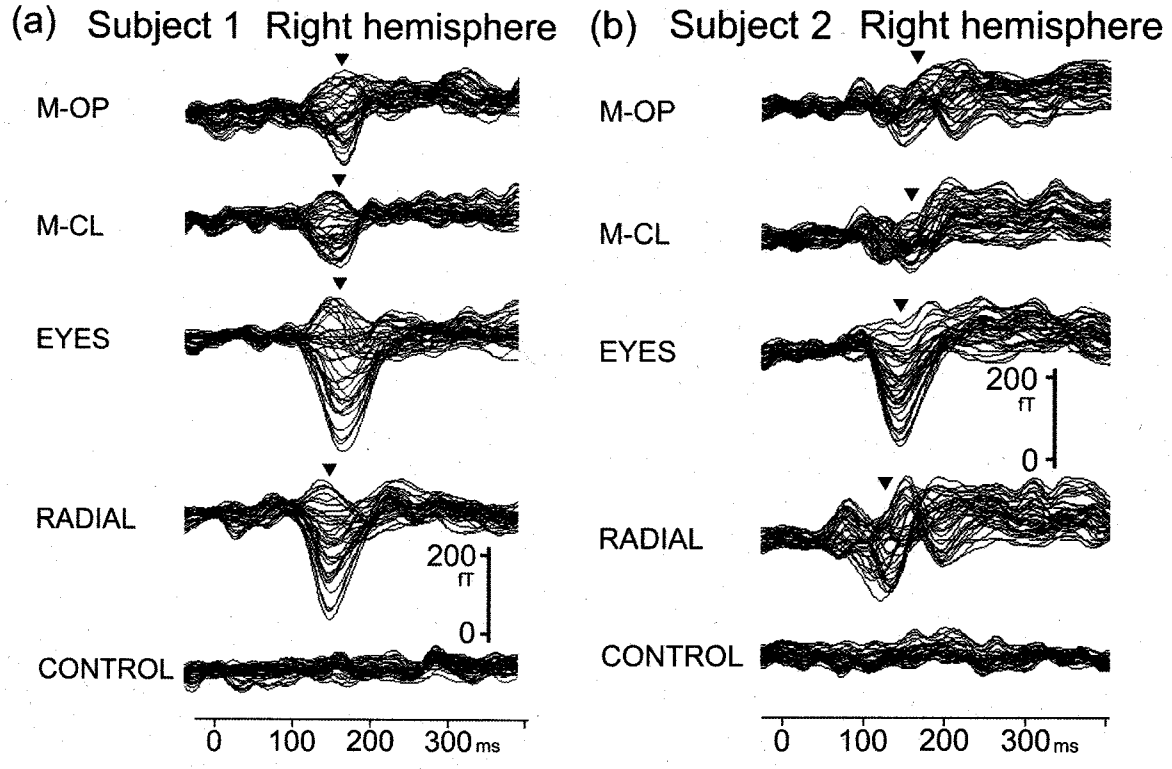


Fig. 3: Right hemisphere S2 MEG activity shown in a 37 channel superimposed display to all conditions. **(a)** In Subject 1 1M peak latency was 154.8, 156.7, 162.5 and 148.1 ms to M-OP, M-CL, EYES and RADIAL, respectively. Associated maximum RMS values were 62.6, 66.7, 122.0 and 119.4 fT. **(b)** Subject 2's data show a similar pattern for 1M, however, peak latencies for M-OP and M-CL were longer than for EYES and RADIAL (168.3, 160.6, 143.3 and 127.9 ms for M-OP, M-CL, EYES and RADIAL, respectively). Its maximum RMSs for M-OP and M-CL were relatively smaller than EYES and RADIAL (48.1, 40.9, 127.9 and 66.5 fT for M-OP, M-CL, EYES, and RADIAL, respectively). Neither subject showed a response to CONTROL.

Table 1: 1M peak latency, which was measured at the point with the maximum RMS at visible peaks of each component, and maximum RMS values for S2. Means and standard deviations to M-OP, M-CL, EYES, and RADIAL in the right and left hemispheres. One-way ANOVA showed a significant main effect for stimulus type (Right hemisphere: $F= 4.35$; $p<0.05$, Left hemisphere: $F=12.73$; $p<0.05$) in 1M peak latency and maximum RMS (Right hemisphere: $F=5.51$; $p<0.05$, Left hemisphere: $F=6.30$; $p<0.05$).

		Right (n=12)	Left (n=11)
M-OP			
	Latency (ms)	159.8 ± 17.3 *	162.4 ± 11.6 **
	RMS (fT)	62.5 ± 23.5 **	56.0 ± 28.1 ** #
M-CL			
	Latency (ms)	161.9 ± 15.0 *	160.9 ± 9.8 **
	RMS (fT)	59.1 ± 21.8 ** #	50.1 ± 17.5 ** ##
EYES			
	Latency (ms)	161.2 ± 18.9 **	164.6 ± 14.2 **
	RMS (fT)	82.5 ± 32.7	87.3 ± 42.7
RADIAL			
	Latency (ms)	140.1 ± 18.0	138.4 ± 9.0
	RMS (fT)	98.7 ± 29.7	99.4 ± 33.0

** P<0.01, * P<0.05: Comparison with results of RADIAL

P<0.01, # P<0.05: Comparison with results of EYES

While signal strength of 1M, as measured by maximum RMS, tended to be more variable across conditions, a one-way ANOVA analysis did demonstrate a main effect of stimulus type ($F=5.51$; $P<0.05$). Post-hoc paired comparisons using Fisher's PLSD showed that the maximum RMS for the mouth movement conditions (M-OP and M-CL) did not differ significantly. However, M-OP was significantly smaller than RADIAL

($p<0.05$), but was not different from EYES. M-CL was significantly smaller than EYES ($p<0.05$) and RADIAL ($p<0.01$). There were no significant differences between EYES and RADIAL.

(2) The source analysis using BESA

At first, the location, orientation, and amplitude of MEG responses were estimated with a single-dipole analysis (Sarvas 1987), but it was difficult or impossible to estimate reliable sources in most subjects, probably contaminated by some activities in V1 and fusiform gyrus to some degree in such subjects. Therefore, we used multi-dipole model (BESA) in this study. BESA results for M-OP and M-CL fulfilled our strict analysis criteria in 11 of the 12 subjects whose 1M was observed in all conditions with apparent motion (M-OP, M-CL, EYES and RADIAL). In the 11 subjects who fulfilled GOF criteria for 4 stimulus types, Source 1, located in the lateral temporal region, around MT/V5, was very large in amplitude, relative to the small amplitudes of the other 3 sources (Fig. 4). In one of the other subjects not in this group of 11, the BESA 4-source model did not fulfill the criteria for all conditions, again probably due to low S/N ratio. Despite this, for the stimulus conditions that the BESA model could compute, Source 1 was very large as compared with the other 3 sources.

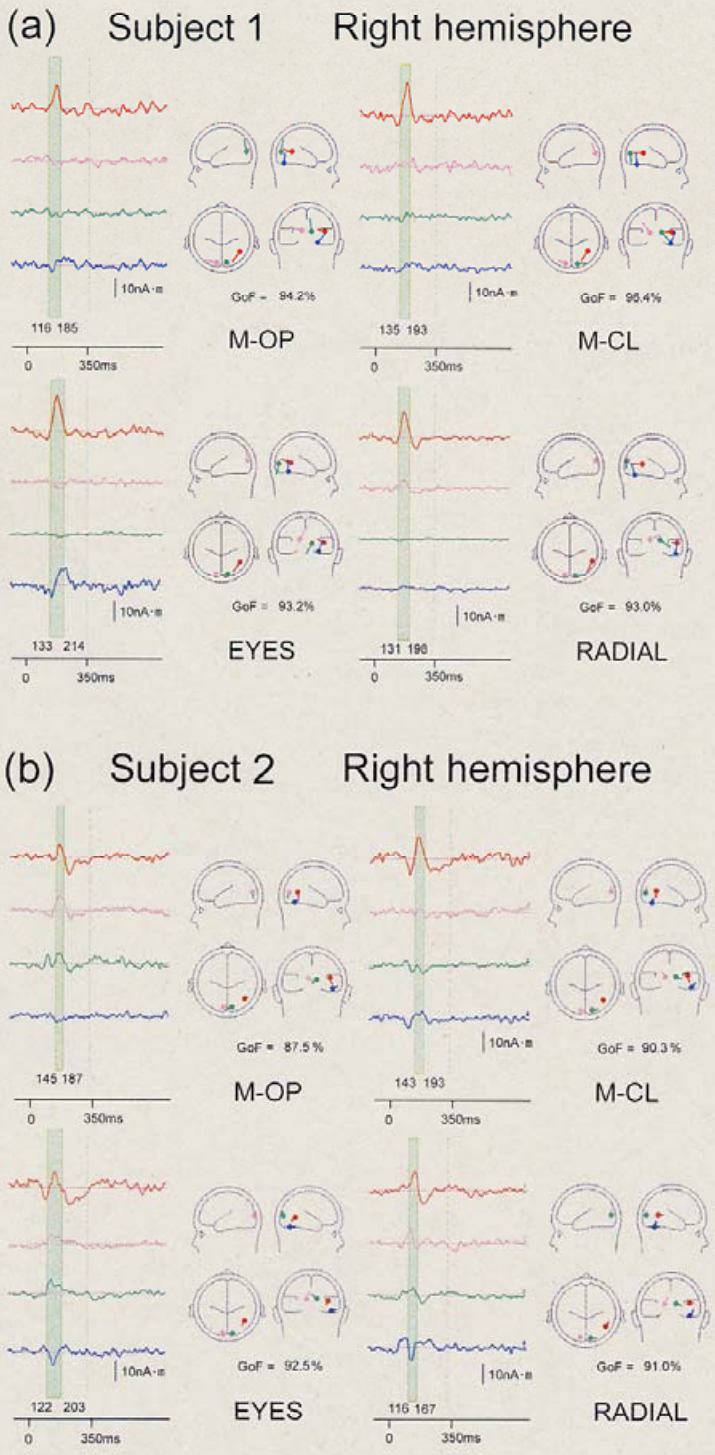


Fig. 4: Right hemisphere 4-source BESA model for S2 for the time interval highlighted in green. Same subjects as in Fig. 3. Source 1 (red) was located in human MT/V5, Source 2 (pink) and 3 (green) in the left and right primary visual fields (V1) and Source 4 (blue) in the right fusiform gyrus. Overall GOF values are also displayed. **(a)** Subject 1. The activity of Source 1 was very large relative to the other sources, and this pattern was observed in the majority of subjects. **(b)** Subject 2 shows clear activity to Source 1 but also to other Sources.

We used the anatomical landmark criteria reported by Dumoulin et al. (2000) to confirm the Source 1's localization in MT/V5. Their MT/V5 fMRI activation to moving random checkerboard patterns was located

along the inferior temporal sulcus (ITS), which they separated into 3 parts: (1) posterior ITS; (2) the ascending limb of ITS (ALITS); and (3) the posterior continuation of ITS (PCITS). In the present study, the Source 1s to all movement conditions in all subjects were located in one of these three defined regions: in 2 subjects in or adjacent to (1), 5 subjects in or adjacent to (2) and other 4 subjects in or adjacent to (3).

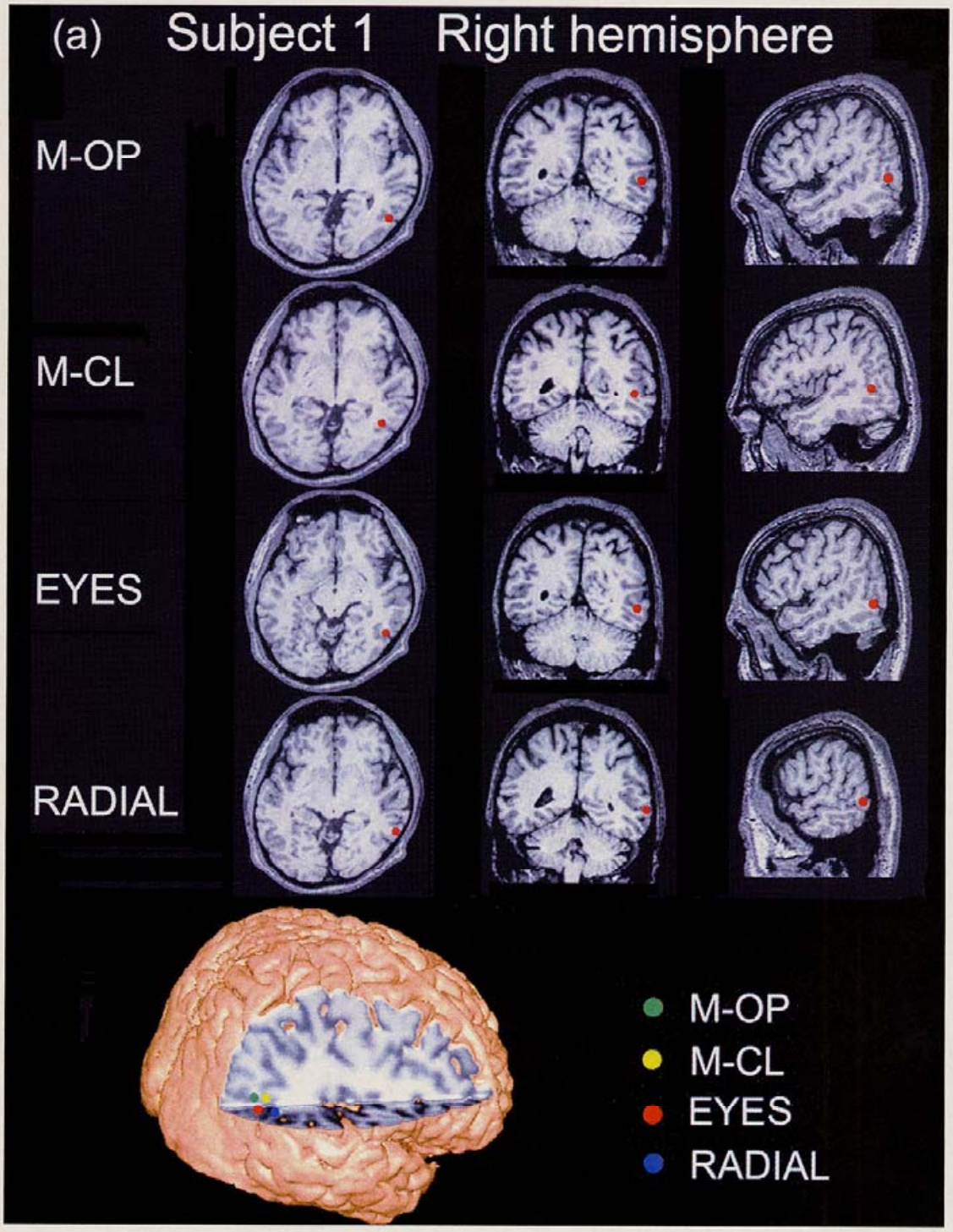
The locations of Source 1 for M-OP and M-CL, as indicated by their x, y and z coordinates were similar (Table 2, Fig. 5). Source 1 to both M-OP and M-CL appeared to be located more anteriorly and superiorly relative to EYES, and more medially relative to RADIAL. However, these differences were not significant when the x, y and z coordinates were submitted to a one-way ANOVA with a factor of stimulus type.

Table 2: Source 1 locations for S2. Means and standard deviations of x, y, and z coordinates (mm) (means and standard deviations) to M-OP, M-CL, EYES, and RADIAL in the right and left hemispheres.

	Right (n=11)			Left (n=10)		
	X (mm)	Y (mm)	Z (mm)	X (mm)	Y (mm)	Z (mm)
M-OP	42.5±6.9	-26.2±9.5	59.5±7.6	-35.3±5.3	-30.9±8.6	62.9±5.7
M-CL	43.1±6.3	-25.2±10.0	60.1±7.4	-35.6±6.8	-30.1±8.5	60.7±5.8
EYES	41.4±8.2	-30.6±8.5	58.5±7.2	-35.6±7.2	-34.4±4.0	59.4±5.1
RADIAL	44.8±8.2	-27.0±7.9	60.8±5.9	-33.3±6.4	-32.3±4.9	63.3±6.0

We placed Source 5 in the STS region where Puce et al. (1998) reported in all subjects for both standard and anterior position, but no reliable activity

was identified in any subjects using our GOF criteria.



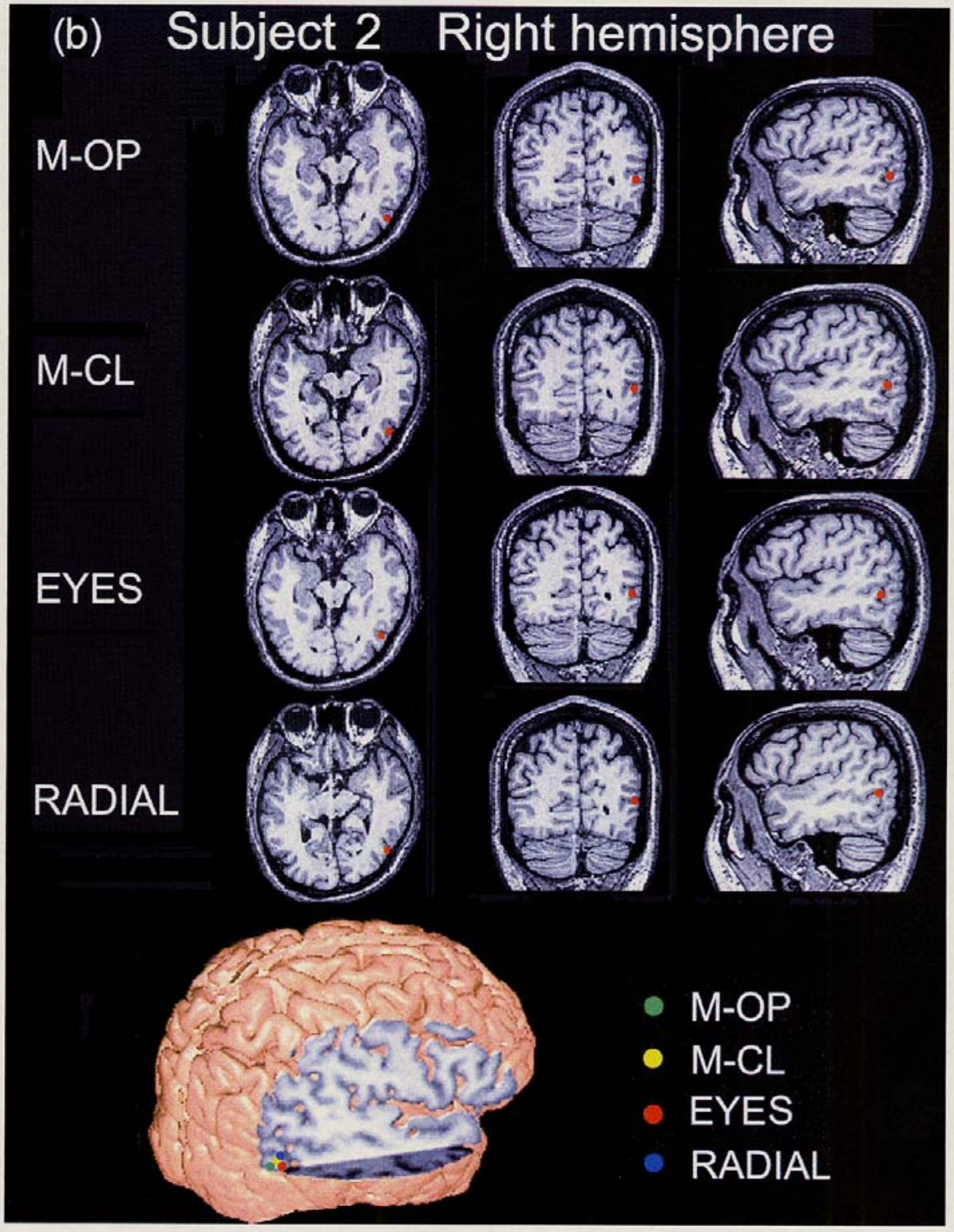


Fig. 5: Right hemisphere locations for S2 Source 1 ECDs for apparent motion conditions overlaid on axial, coronal, and sagittal MRI slices, and the volume rendered brain of each subject. **(a)** Subject 1. **(b)** Subject 2.

The left hemisphere

(1) Waveform characteristics

In the left hemisphere, as in the right, the most prominent MEG component was 1M (Fig. 6, Subject 1), which was observed to all apparent motion conditions in 11 out of 17 subjects. No response was observed to CONTROL in any subject. A one-way ANOVA showed a significant main effect for stimulus condition (Table 1, $F=12.73$; $P<0.05$). Post-hoc paired comparisons using Fisher's PLSD showed that the latency of the 1M to RADIAL was significantly shorter than to all other conditions ($P<0.01$). There were no significant latency differences between the mouth movement conditions (M-OP and M-CL), or between these conditions and EYES.

The maximum RMS of the 1M also differed significantly across stimulus condition, as indicated by a one-way ANOVA (Table 2, $F=6.30$; $P<0.05$). Post-hoc paired comparisons using Fisher's PLSD showed that the maximum RMS to M-OP and M-CL were significantly smaller than those of EYES ($P<0.01$) or RADIAL ($p<0.01$), but showed no significant differences amongst themselves. 1M RMS values to EYES and RADIAL did not differ significantly.

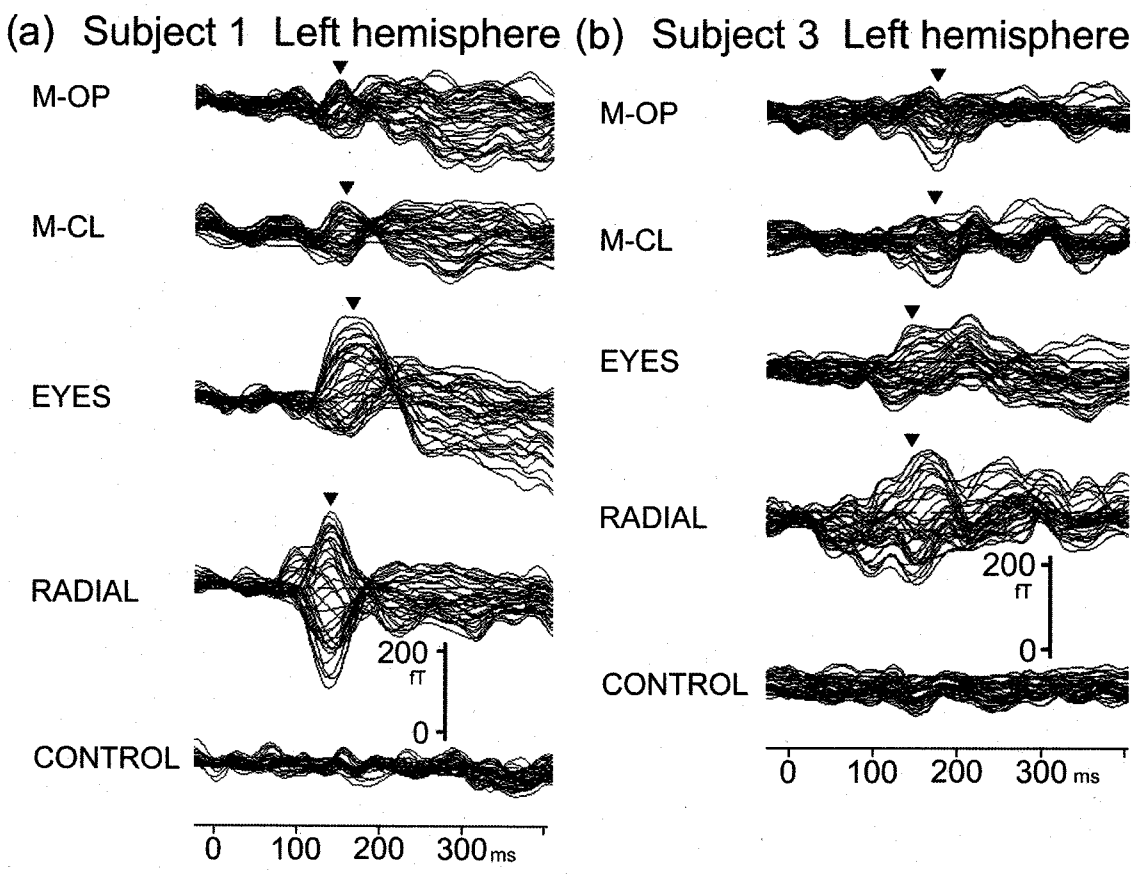


Fig. 6: Left hemisphere S2 MEG activity shown in a 37 channel superimposed display. **(a)** Subject 1. **(b)** Subject 3. In both subjects, a clear 1M was identified for M-OP, M-CL, EYES, and RADIAL, and was absent for CONTROL. 1M for M-OP and M-CL was relatively longer in latency and smaller in amplitude than EYES and RADIAL, as found in the right hemisphere.

(2) The source analysis using BESA

Using our GOF criteria, BESA results for M-OP and M-CL fulfilled the strict criteria in 10 subjects out of the 11 subjects who showed clear elicited MEG activity. In all 10 subjects who fulfilled the GOF criteria for the apparent motion stimulus conditions, Source 1 located in the lateral temporal region around MT/V5, was very large in amplitude, but the other 3 Sources were very small in amplitude or showed no significant activity

(Fig. 7). In the eleventh subject, the BESA 4 source model did not fulfill the criteria for M-CL condition, again probably due to a low S/N ratio. As found in the right hemisphere, the left hemisphere Source 1s to all movement conditions in all subjects were located in or adjacent to the three anatomical regions defined earlier: in 2 subjects for region (1), 5 subjects for region (2) and other 3 subjects for region (3). Therefore, our findings were in agreement with Dumoulin et al. (2000), and other imaging studies of MT/V5, for both the left and right hemispheres.

The locations, as measured by x, y and z co-ordinates, of Source 1 to M-OP and M-CL were similar (Table 2, Fig. 8) and relative to EYES, appeared to be located more superiorly, and compared with RADIAL were located more posteriorly (Table 2). Statistical testing, however, indicated that these apparent differences in location of Source 1 were not significant.

Comparison of data from right and left hemispheres

1M peak latencies in both hemispheres were compared using paired t-tests in 9 subjects who showed clear 1M components in each hemisphere. There were no significant inter-hemispheric differences of latency (Table 1). We did not compare RMS values between the hemispheres, as sensor placement over each hemisphere may not have been perfectly symmetrical and hence the distance between MEG sensors and brain may not have been the same across hemispheres.

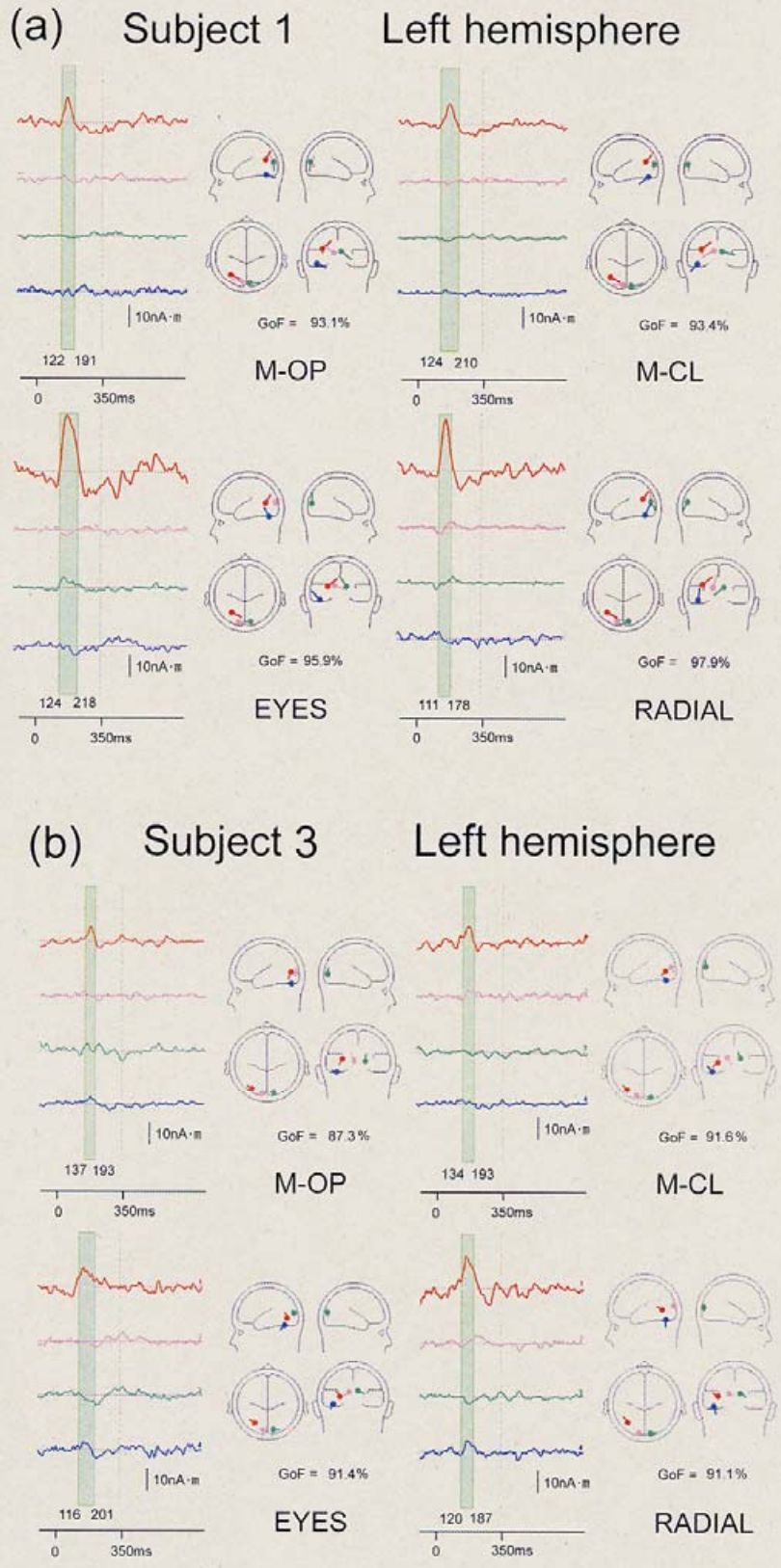
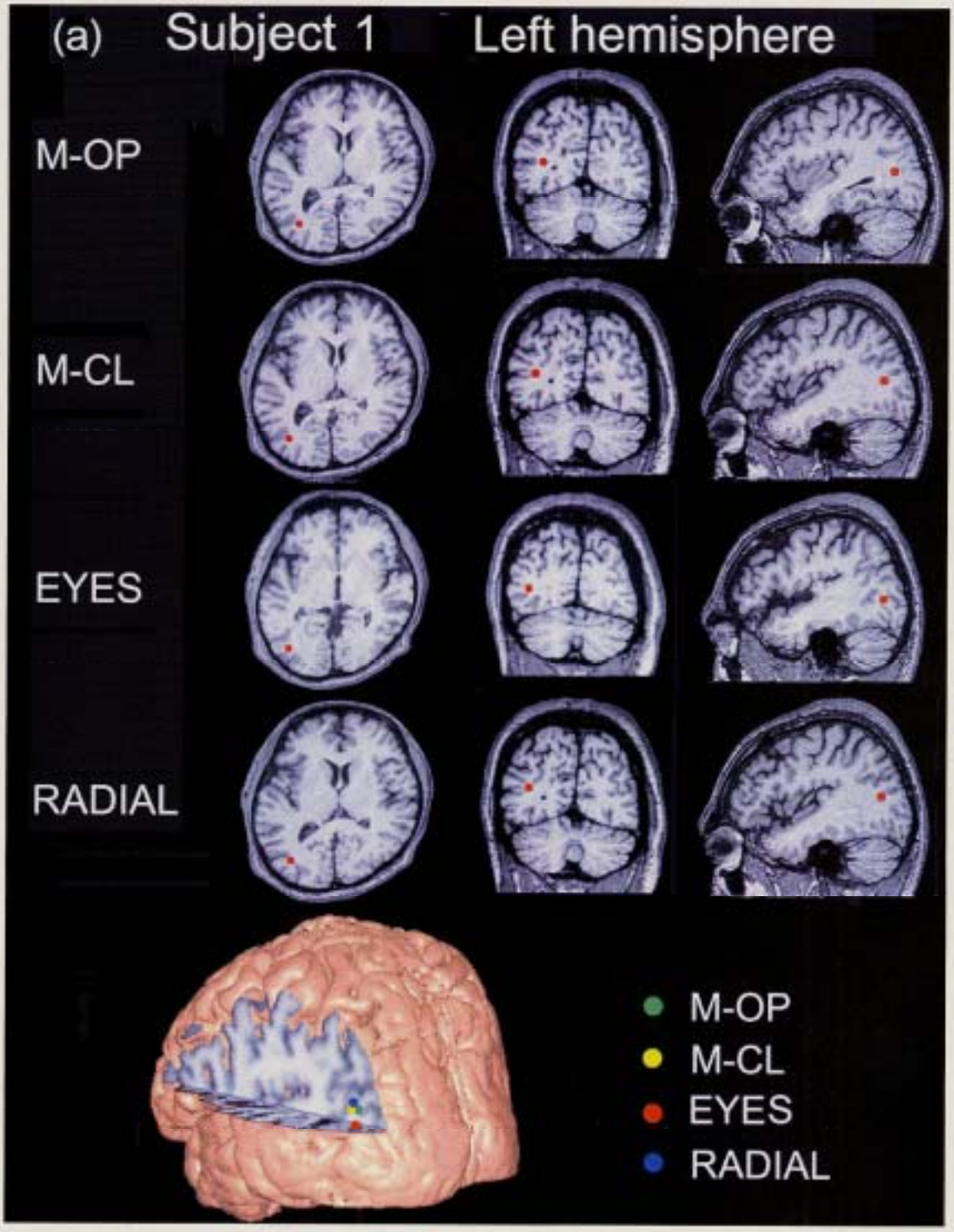


Fig. 7: Left hemisphere 4-source BESA model for S2 for the subjects shown in Fig. 6. **(a)** Subject 1. **(b)** Subject 3. Legend is identical to that of Figure 4.


(a) Subject 1 Left hemisphere
M-OP
M-CL
EYES
RADIAL
M-OP
M-CL
EYES
RADIAL

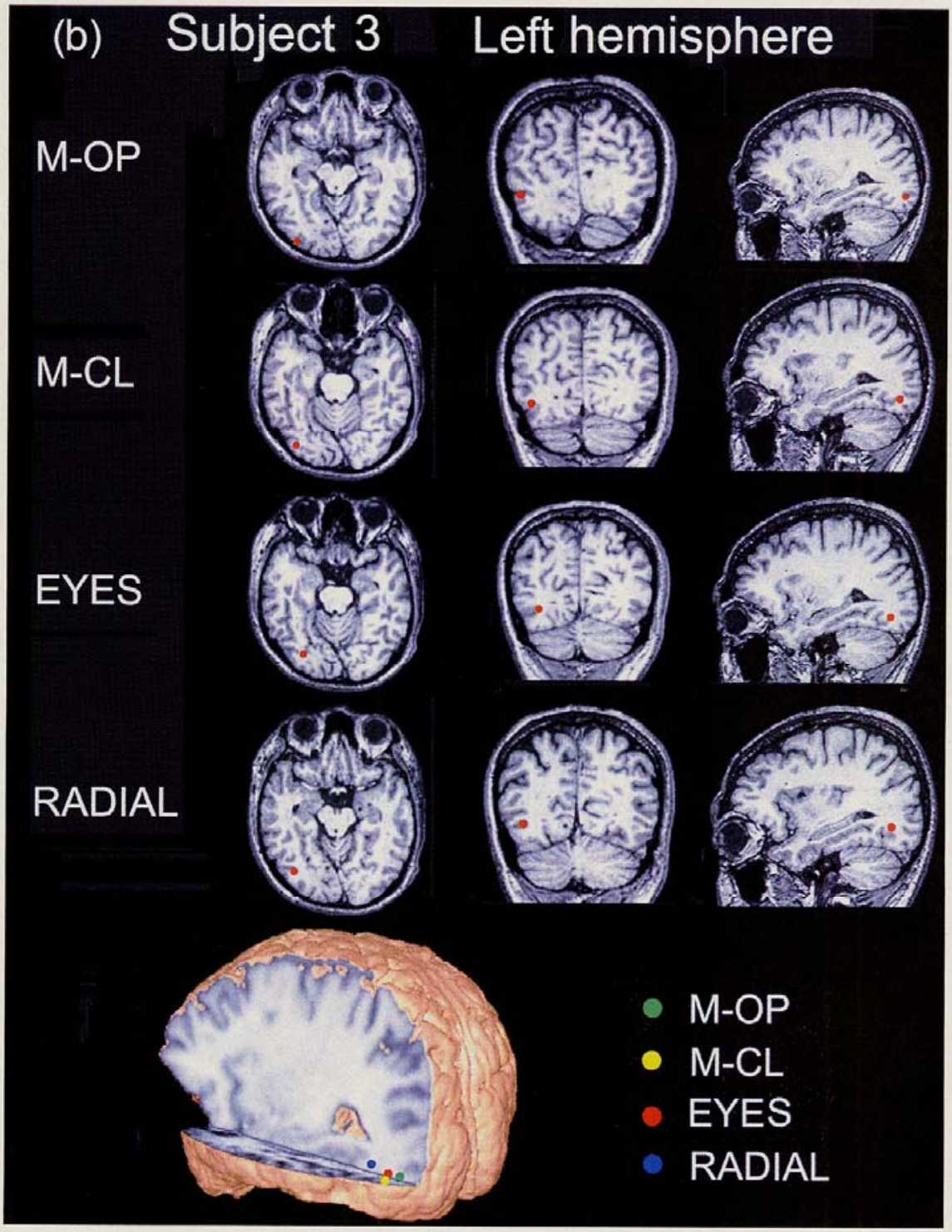


Fig. 8: Left hemisphere locations for S2 Source 1 ECDs for apparent motion conditions overlaid on axial, coronal, and sagittal MRI slices, and the volume rendered brain of each subject. **(a)** Subject 1. **(b)** Subject 3.

2. MEG activity following S1

The Right hemisphere

(1) Waveform characteristics

Similar to S2 onset, right-sided MEG activity to S1 onset was the most prominent, with clear evoked responses occurring to the onset of all stimulus types (see Fig. 9, MEG waveforms for Subject 1). Since we have previously reported ERP and MEG findings to viewing facial motion involving the eyes (Watanabe et al., 1999a, 2001, 2002; Puce et al., 2000), here we contrast only differences in waveforms to viewing the *Face with closed mouth* and *Face with opened mouth*. Two components (M120 and M170) whose peak latency was approximately 120 and 170 ms, respectively were identified bilaterally (Fig. 9, Table 3). This nomenclature was based on our previous reports (Watanabe et al., 1999a, b). The M120 peaked at 129.0 ± 12.5 ms to *Face with closed mouth* and 127.7 ± 10.4 msec to *Face with opened mouth*, and was recorded in 9 of 17 subjects. Peak latency and RMS of the M120 showed no significant difference between *Face with closed mouth* and *Face with opened mouth*. We identified M170 in 14 of 17 subjects from the right hemisphere.

Differences in peak latency and RMS of both M120 and M170 between *Face with closed mouth* and *Face with opened mouth* were compared using paired t-tests, and there were no significant differences between the two stimulus conditions.

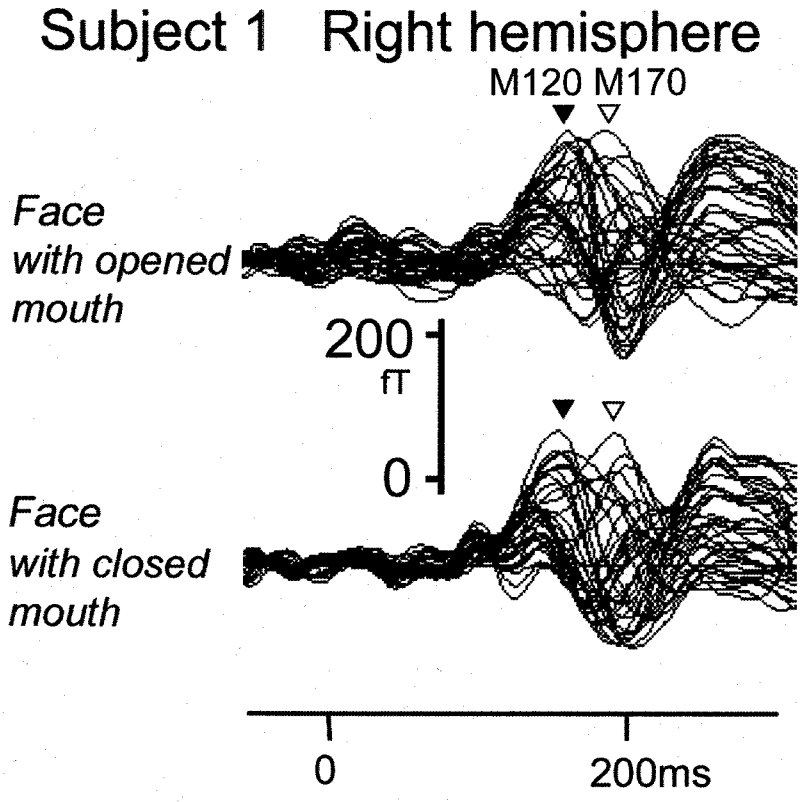


Fig. 9: S1 onset right hemisphere MEG waveforms to *Face with closed mouth* and *Face with opened mouth* from Subject 1. The 37 channel superimposed display shows clear M120 and M170 components. M120 latencies were 127.3 and 134.9 ms to *Face with opened mouth* and *Face with closed mouth*, respectively. M170 latencies were 182.9 and 186.8 ms to *Face with opened mouth* and *Face with closed mouth*, respectively. The maximum RMSs of M120s were 77.8 and 83.6 fT to *Face with opened mouth* and *Face with closed mouth*, respectively. The maximum RMSs of M170s were 84.5 and 83.6 fT to *Face with opened mouth* and *Face with closed mouth*, respectively.

Table 3: M120 and M170 peak latency and maximum RMS values in the right and left hemispheres for S1. Means and standard deviations to *Face with closed mouth* and *Face with opened mouth*. M120s were identified to both *Face with closed mouth* and *Face with closed mouth* in 9 and 6 out of 17 subjects from the right and left hemisphere, respectively. M170s were identified to both *Face with closed mouth* and *Face with closed mouth* in 14 and 12 out of 17 subjects from the right and left hemisphere.

		Right	Left
Face with closed mouth			
M120	Latency (ms)	129.0 ± 12.5	126.3 ± 12.8
	RMS (fT)	83.3 ± 23.6	66.2 ± 23.4
M170	Latency (ms)	185.8 ± 16.6	176.0 ± 16.4
	RMS (fT)	92.6 ± 38.0	87.9 ± 29.3
Face with opened mouth			
M120	Latency (ms)	127.7 ± 10.4	126.3 ± 12.1
	RMS (fT)	79.0 ± 32.5	66.9 ± 21.8
M170	Latency (ms)	185.9 ± 16.2	174.1 ± 15.0
	RMS (fT)	108.3 ± 33.2	90.3 ± 27.1

(2) The source analysis using BESA

M120 and M170 source localization, using 4 seeded sources, was performed using BESA (Table 4, Figs. 10 & 11). The time course of Sources 2 and 3 (left and right V1) corresponded to the M120 component, whereas those for Sources 1 (lateral temporal region) and Source 4 (fusiform gyrus) corresponded to M170. The location of each source did not show a significant difference between *Face with closed mouth* and *Face with opened mouth*. Both Source 1 and 4 were the main components of M170. In contrast, the Source 2 and 3 were small in strength.

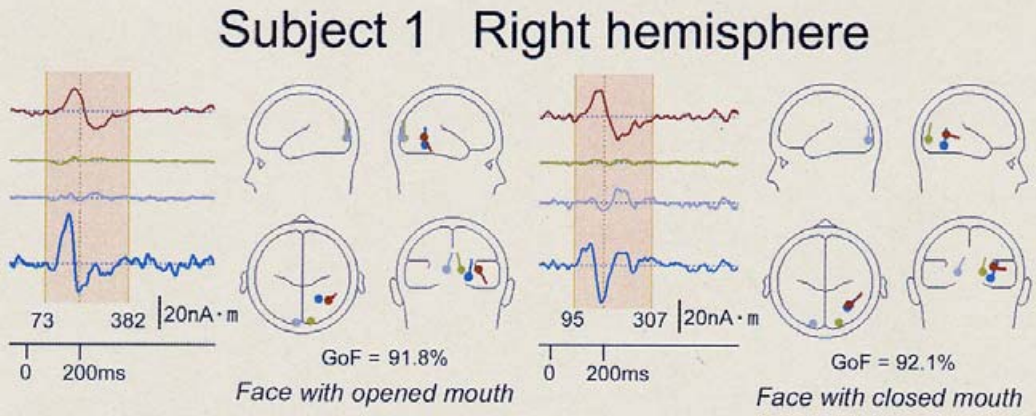


Figure 10: Right hemisphere 4 source BESA model for S1 onset in Subject 1. Legend identical to that of Fig. 4.

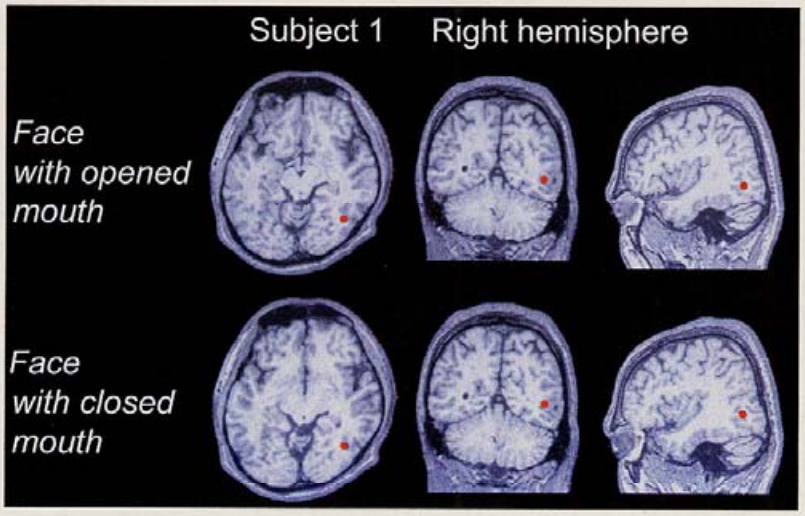


Figure 11: S1 onset ECDs for Source 1 for *Face with closed mouth* and *Face with opened mouth* overlaid on axial, coronal, and sagittal MRI slices in Subject 1's right hemisphere.

The left hemisphere

(1) Waveform characteristics

Left hemisphere data were morphologically similar to those of the right hemisphere, consisting of two components, M120 and M170 (Table 3). The M120 was recorded in 6 of 17 subjects both to *Face with closed mouth* and to *Face with opened mouth*. M120 peak latency and dipole moment showed no significant differences between *Face with closed mouth* and *Face with opened mouth*. M170 was clearly identified in 12 subjects, however, there were no significant differences across stimulus type on peak latency or RMS, when tested using a paired t-test.

(2) The source analysis using BESA

Source localization of M120 and M170 was analyzed by BESA (Figs. 10 & 11) using 4 sources. Sources 2 and 3 (left and right V1) corresponded to the M120 component, and Sources 1 (lateral temporal region) and 4 (fusiform gyrus) corresponded to M170. Localization of each source did not show a significant change between *Face with closed mouth* and *Face with opened mouth*. As well as the right hemisphere, both Source 1 and 4 were main components of M170. In contrast, Sources 2 and 3 were small in amplitude.

Discussion

Our aim was to examine whether viewing mouth movements activate MT/V5, and if so, whether the loci of neural activity elicited to viewing mouth versus eye movements could be differentiated in this region. Clear MEG waveforms were recorded following the movement conditions (S2), as well as to the onset of the face stimulus (S1). In all S2 conditions, a clear 1M occurred in response to all apparent motion conditions in both hemispheres, although the strongest signal was always observed in the right hemisphere. The main source of this activity (Source 1) for all conditions was located in the lateral occipito-temporal region, corresponding to human MT/V5.

The location of the sources to the various types of facial movements did not differ within themselves or relative to the radial background condition. Interestingly, despite this lack of difference in the location of the sources, there were differences in the behavior of 1M across conditions. While 1M showed a peak latency of around 160 ms, the facial motion conditions in general produced longer 1Ms relative to our general motion control – a finding that we have seen previously using eye movements (Watanabe et al., 2001). Additionally, responses to mouth movements were smaller relative to responses observed to the eye aversion or radial background motion controls. These differences in the behavior of 1M suggest that MT/V5 and its surrounds may possess multiple response characteristics. The latency and RMS data to the eye and radial background movements were consistent with our previous study (Watanabe et al., 2001) and will not be discussed

further.

We found no clear amplitude difference between the two mouth movement conditions, unlike the ERP study of Puce et al. (2000 and 2003), who reported an N170 component, corresponding to our 1M, which was significantly earlier in latency and larger in amplitude to mouth opening relative to mouth closing movements. There are a number of possible reasons for the observed differences between the ERP and the present MEG study.

Theoretically, MEG will detect tangential neural sources located just beneath the MEG sensors, and EEG will detect not only radial dipoles located near the electrodes but also more distant sources by volume current conduction, effectively detecting sources directed both tangentially and radially. Regions in relatively close proximity to MT/V5 such as the medial superior temporal area (MST) respond strongly to non-linear motion (Tanaka and Saito 1989), and the anterior superior temporal polysensory area (STPa) respond selectively to the motion of animate objects, including bodies and faces (Oram and Perrett 1994) and are also likely to generate activity to the stimulus types used here. Hence, it is likely that activation to M-OP, M-CL and EYES could be also generated radially in MST and/or STPa, in addition to that produced in MT/V5. As there could potentially be multiple generators in temporoparietal cortex which temporally overlap and may overlap in space partially, it could prove difficult for MEG, or indeed ERP methods, to clearly identify them. Additionally, neural activity, as detected by MEG, in STS might be much smaller relative to that in MT/V5,

so that its contribution would be masked by the net dipole generated by MT/V5 activity. Alternatively, the direction of the net dipole generated in STS was mainly radial, enabling it to be better detected with ERP methods, whereas the direction of the net dipole generated in MT/V5 was mainly tangential. This might be more likely, given that in the ERP studies the activity to general motion controls is smaller than that observed to facial motion (Puce et al 2000, 2003).

Puce et al. (1998), using fMRI, reported that both MT/V5 and STS were activated by viewing mouth movements. Activated region of STS was located approximately 20mm more anterior, 10 mm more lateral, and 4 mm upper in mean than that of MT/V5, so that Euclidian distance between them was about 23 mm, However, we did not detect activity from this STS region, even when the MEG probe was placed more anteriorly to overlie the posterior STS region known to respond to facial motion. In our previous study in which MEG responses were recorded by viewing eye movements (Watanabe et al., 2001), we found activity in only MT/V5 but not in STS, as in the present study.

In this study the recorded MEG activity to the mouth movement stimuli was smaller than that to the eye motion stimuli, potentially making it more difficult to detect this activity. The number of neurons activated by viewing mouth movements could be smaller than those that fire to viewing eye movements. Second, the orientation of the neural generators to mouth movement, while in close proximity to those responding to eye movements, may be less radially oriented. In contrast, in the ERP study (Puce et al.,

2000) reported that N170 amplitudes to eye and mouth movements were not significantly different.

If the stimulated visual fields for the mouth and eye conditions were different, the question arises whether the difference in centers of activation is a function of retinotopy in human MT/V5? Single-unit studies in macaques indicate that MT/V5 contains a complex non-first order transformation with field discontinuities visual field representation (Van Essen et al., 1981; Gattass and Gross, 1981). Specifically: (I) MT/V5 contains the representation of virtually the entire contralateral visual field; (II) the representation of the central 5 degrees of visual field is greatly magnified; (III) the lower visual field shows a greater magnification relative to the upper visual field. Tootell et al. (1995a) were the first investigators to thoroughly study the response properties of human MT/V5 in a systematic series of fMRI experiments. They observed that human MT/V5, unlike that of monkeys: (I) responds robustly to stimulation of the contralateral and ipsilateral visual hemifield; (II) does not have a lower visual field bias. A potential difficulty with fMRI studies is that partial voluming can confound the sampled activation, although this could be dealt with by collecting fMRI data at higher resolution (Tootell et al., 1995a). In our study, mouth and eye movements were presented within the central 5 degrees of visual field, and it is notable that the mouth movements, which were shown in the lower visual field, somewhat unexpectedly showed the smaller amplitude.

Finally, an MEG study (Naito et al., 2000) using a short bar as an

apparent motion stimulus as a function of visual field (90, 135, 180, 225 and 270 degree) and movement direction (inward and outward) did not observe differences between outward and inward movement at 270 degrees. Notably, our stimulus conditions M-OP and M-CL were not unlike the outward and inward movement at 270 degree, respectively, studied by Naito et al (2000). Therefore overall, we believe that visual field effects were unlikely to contribute to the differences seen in our data.

The motion direction is considered to be one of factors influencing activities in MT/V5 in monkeys (Maunsell et al 1983). Since the motion direction for M-OP, M-CL, EYES, and RADIAL was different, the question may arise whether the motion direction might be the main factor for the difference of the activation of all conditions. In the previous MEG study (Maruyama et al 2002), the speed of the movement is the most important factor for activities in human MT/V5, even for the incoherent as well as coherent motion, that is, the motion direction do not affect latency, amplitude, and the estimated origin of the activation in human MT/V5. Therefore, we believe that the motion direction were unlikely to contribute the difference found in our data. Therefore, we think the motion direction is not the main factor for this study.

The locations of sources for mouth opening, mouth closing and eye aversion movements or radial background movements were not different from one another despite significant differences in response latency. MEG source locations indicate centroids of activity, and hence, it may be difficult for MEG to detect difference of activated regions, when regions partially

overlap one another. Similarly, partial voluming effects in PET and fMRI could also make it difficult to detect such small differences in partially overlapped activated regions.

Is there a hemispheric dominance for the 1M to mouth movements? More subjects showed consistent 1M components and reliable BESA models in the right hemisphere, and RMS values in the right hemisphere tended to be larger than the left. Previous neuroimaging studies of human MT/V5 have not reported significant inter-hemispheric differences (Puce et al., 1998 for eye and mouth movement; Puce et al., 2003 for mouth movement; Van Oostende et al., 1997 for random textured patterns). While both hemispheres are clearly activated by viewing facial movements, a right hemisphere bias may exist for face movements.

We implemented this experimental design specifically to reduce activity not related to general motion perception or to face perception, particularly in V1 and the fusiform gyrus. A potential confounding source of activation for our calculated extrastriate 1M could be summed contributions from V1. The generally smaller activity of the extrastriate cortices can be easily masked by V1 activity (Okusa et al., 1998, 2000). Robust activity in V1 can be elicited by both stimulus onset and motion (e.g. Uusitalo et al., 1997). By designing the experiment so that overall luminance and contrast remained unchanged despite regular changes in stimulus display (Fig. 1), we were able to overcome this considerable confound. Hence, not surprisingly, V1 activity was absent or negligible.

A similar logic in stimulus delivery applies to activity in the fusiform

gyrus, a known source of PET and fMRI (e.g. Sergent et al 1992; Puce et al., 1995, 1996, 2003) and MEG (e.g. Watanabe et al., 1999a&b, 2003) activation to face onset. In this study, activity of fusiform gyrus to S2 was negligible – a result was not caused by MEG sensor placement, since our previous studies (Watanabe et al., 1999a&b, 2003) could clearly detect activity generated in fusiform gyrus by placing MEG sensors in the same position as the present study, and activity in the fusiform was observed in response to S1 in the current study.

Experiment 2

Interaction between auditory and visual stimulus relating to the vowel sounds in the auditory cortex in humans: a magnetoencephalographic study

Introduction

In our daily lives, integration of visual and auditory stimuli relating to speech perception seems very important, since it becomes much easier to understand vowel sounds when we can see the mouth movement of the speaker. Furthermore, it has been shown that we can perceive modified and confusing sounds in psychological studies. For example, McGurk and MacDonald (1976) reported that we usually perceive /da/ when auditoy stimulus /ba/ and visual stimulus /ga/ are presented simultaneously. In both primates (Poremba et al., 2003; Scott et al., 2003) and humans studies using neuroimaging and electrophysiological methods, such as fMRI (Calvert et al., 2003; Lebib et al., 2003; Sekiyama et al., 2003; Wright et al., 2003), PET (Sekiyama et al., 2003) and EEG (Klucharev et al., 2003; Lebib et al., 2003; Pourtois et al., 2000) and MEG (Mottonen et al., 2002; Sams et al., 1991), various interactions have been reported between auditory and visual stimuli. For example, Wright et al. (2003), using fMRI, reported that STS and superior temporal gyrus (STG) demonstrated greater responses to pairing of visual and auditory stimulus than visual or auditory stimulus alone. Therefore, in this study, we used MEG, which has very high temporal and spatial resolution, to investigate whether auditory cortex activity was influenced by visual motion. In particular, we focused on the early processing stage by analyzing the main early component, M100 generated in the auditory cortex, by using apparent motion, which we used in our previous study (Watanabe et al., 2001).

Methods

Subject

We studied ten right-handed normal volunteers (1 female and 9 males) ranging in age from 24 to 49 years (mean age, 32.2 years) with normal or corrected auditory and visual acuities. All subjects gave informed consent to participate in the experiment, which was approved by the Ethics Committee at the National Institute for Physiological Sciences.

Stimulation

One vowel sound, /a/, spoken by a Japanese female speaker was recorded and reconstructed the length of the vowel sound for this study using Wave Editor TWE (YAMAHA, Hamamatsu, Japan). We used three visual stimuli as follows: (1) *Face with closed mouth* (S1, S2a and S3 in Fig. 12). (2) *Face with opened mouth* (S2b in Figure 12). (3) *Filler* (F), which was made by dividing and randomizing the stimulus (1) (Fig. 12).

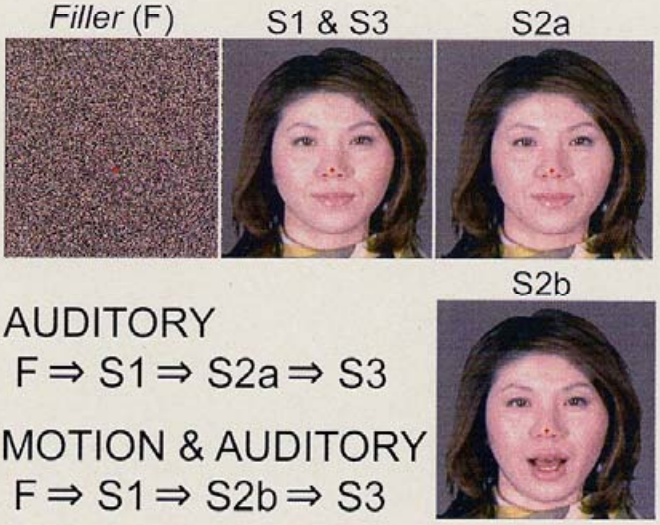


Fig. 12: Two stimulus conditions. (a) AUDITORY (A): S2a was used and all subjects could perceive auditory stimuli /a/ but not speech motion. (b) MOTION & AUDITORY (M & A): S2b was used and all subjects could perceive both auditory stimulus and speech motion as in normal speech.

We used these stimulus types in combination to construct a series of apparent motion conditions, where S1 was replaced by S2a or S2b, then by S3, and then *Filler* with no inter-stimulus interval. S1, S2a and S2b were shown for 800 ms and S3 for 400 ms. *Filler* was presented for 600-800 ms between each stimulus session. Subjects could not predict whether S2a or S2b would be presented after S1. Vowel sound /a/ was presented at S2a or S2b onset was shown and continued for 240 ms. All subjects reported experiencing a percept of clear vowel sound /a/ when either S2a or S2b was presented. We compared two different conditions; (1) **AUDITORY (A)**: S2a was presented, so that subjects did not perceive speech motion. (2) **MOTION & AUDITORY (M & A)**: S2b was presented, so that subjects clearly perceived speech motion.

Visual stimuli were presented by a personal computer (PC, IBM) and video projector (Mirage 2000; CHRISTIE DEGITAL SYSTEM Inc, Kitchener, Canada) housed outside of a magnetically shielded room and the vowel sound was presented to each subject's right and left ears through a plastic tube and ear-pieces (E-A-Rtone 3A; Aero Company, Indianapolis, IN). Visual stimuli were projected on the screen in front of the subject in the magnetically shielded room. The distance between the subject's eyes and the display was 205cm. Stimuli were projected centrally, and subtended a visual angle of 6.7 x 6.7 degrees. Subjects were asked to maintain their gaze at a point at the top of the nose indicated by a small red cross. The mean luminance of the center (fixation) point of the face was 320cd/m^2.

MEG recording

We used a 306-channel biomagneometer, 204 gradiometers and 102 magnetometers, VectorView (Elekta Neuromag Oy; Helsinki, Finland) for recording MEG, but we analyzed results obtained by 204 gradiometers in this study. MEG and vertical and horizontal EOGs were simultaneously recorded with a bandpass filter of 0.1-50 Hz and digitized at a sampling rate of 998 Hz. Epochs in which signal variations were larger than 3pT in the MEG and ±150μV in EOG were excluded from the averages. One hundred ms before and 150 ms after S2 onset, 250 ms in total, was analyzed, and 100 ms before S2 was used as the baseline.

Data analysis

The amplitude of recognizable components was measured as the maximum value in the 204 gradiometers of online-averaged response data in the order of fT/cm. Peak latency was measured at the point with the maximum value at visible peaks of each component. In the source modeling, we used the single ECD modeling (Hamalainen et al., 1993) and estimated dipole location, x, y and z co-ordinates, and dipole moment (strength) in nAm from 14-20 coils around one coil, which showed the maximum value of M100. We evaluated the activity strength of M100 using both the maximum value of the amplitude at one coil (fT/cm) and dipole moment (nAm). We accepted only the dipoles fulfilling the following three strict criteria; (1) Goodness of fit (GoF) values was lager than 95%. (2) Estimated dipoles were located in HG. (3) The dipole

location was stable within 0.5 cm within 5 ms before and after the peak latency. We used a paired t-tests to assess significant differences of peak latency, amplitude, dipole location and moment between the two stimulus conditions in each hemisphere as well as the inter-hemispheric differences of peak latency, amplitude and dipole moment, and $P<0.05$ was considered to be significant.

Results

The most prominent component, M100 peaking at approximately 90 ms, was observed for both conditions in all 10 subjects from the right hemisphere and 9 out of 10 subjects from the left hemisphere (Fig. 13). The M100 showed the maximum amplitude at the same coil in both conditions in all the subjects, though there was a small inter-individual difference of the coil location due to an inter-individual anatomical difference. M100 amplitude and peak latency showed no significant difference between the two stimulus conditions in each hemisphere (Fig. 13), and no significant inter-hemispheric difference (Table 5).

Table 5: M100 peak Latency and Amplitude
There were no significant differences by stimulus condition types or between the right and left hemispheres.

		Right (n=10)	Left (n=9)
A	Latency (ms)	86.3±10.4	90.0± 5.8
	Amplitude (fT/cm)	135.7±25.0	121.5±42.2
M & A	Latency (ms)	88.6± 9.5	92.9± 6.9
	Amplitude (fT/cm)	134.5±23.3	119.0±40.0

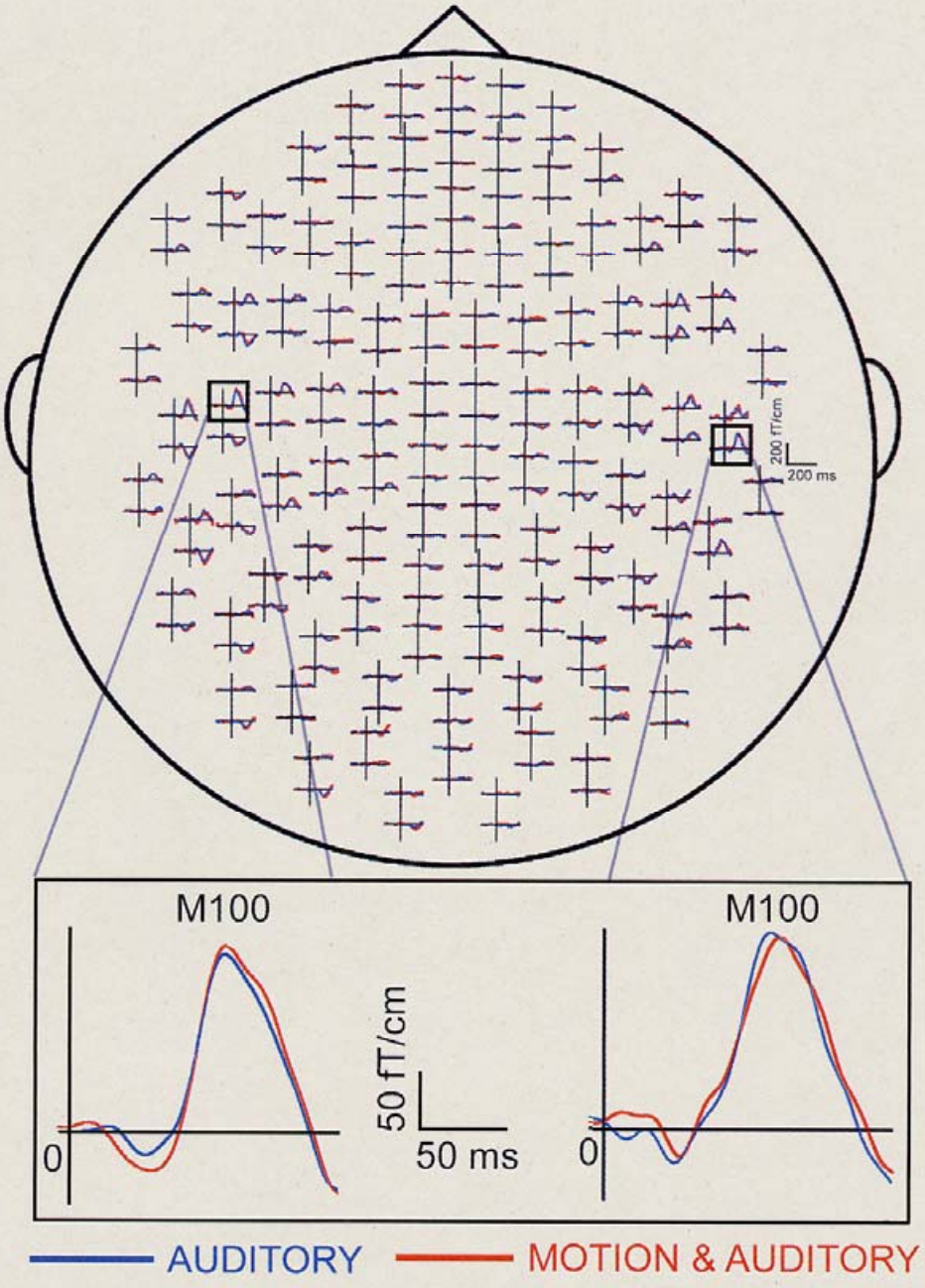


Fig. 13: Auditory evoked magnetic fields of Subject 1 following vowel sound /a/ stimulation. The head is viewed from the top, and in each response pair, the upper trace illustrates the field derivate along the latitude and the lower trace that along the longitude. Waveforms in blue were recorded in AUDITORY condition, that is, S2a (see Fig. 12) was used, and waveforms in red were recorded in MOTION & AUDITORY condition, that is, S2b (see Fig. 12) was used, The lower figure shows the enlarged waveforms recorded at the coil which showed maximum value in each right or left hemisphere. There was no significant difference between two conditions.

Then, we estimated dipole location and moment. However, probably due to our strict criteria, only 9 and 6 subjects fulfilled them from the right and left hemisphere, respectively, so that we adopted only those subjects' results for analyzing dipole location and moment. The M100 was estimated in HG, human auditory cortex, for both conditions in all subjects in each hemisphere (Fig. 14). There was no significant difference in the location and moment of estimated sources (Table 6).

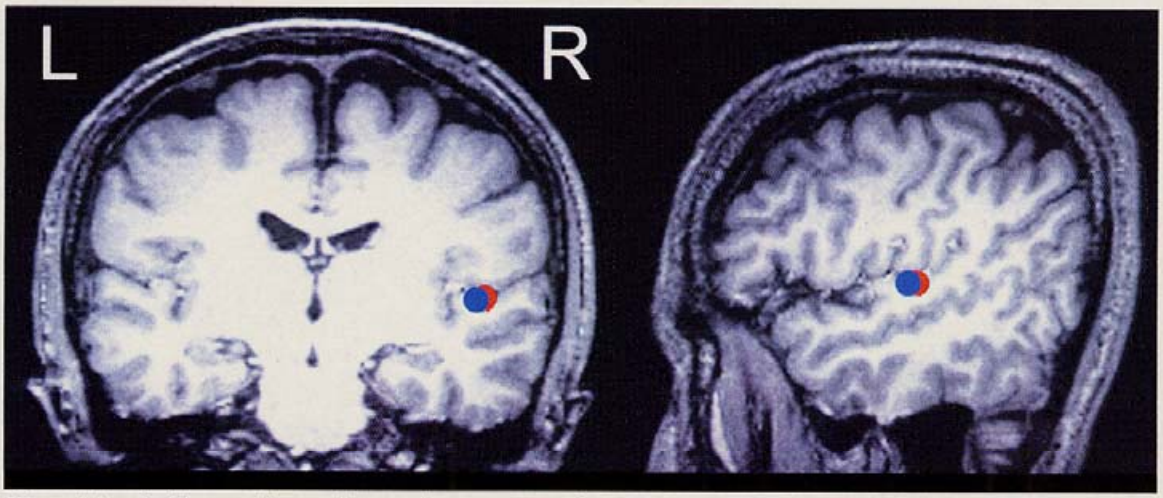


Fig. 14: Subject 1's estimated sources of M100 overlaid on Subject 1's MRI in AUDITORY (blue) and MOTION & AUDITORY (red) condition from the right hemisphere. The sources of both conditions were located in almost the same location of the HG

Table 2: The dipole location and moment of M100 in the right and left hemispheres

There were no significant differences by stimulus condition types in the dipole location and moment and there was no difference between the right and left hemispheres in the dipole moment. X is positive to right, Y to anterior and Z to superior.

	Right (n=9)		Left (n=6)	
	A	M & A	A	M & A
X (mm)	51.2±5.0	51.2±5.7	-52.3±5.0	-51.9±4.3
Y (mm)	19.8±6.0	20.7±6.3	12.2±3.2	11.8±3.9
Z (mm)	57.5±7.2	57.5±6.6	60.3±6.2	60.1±6.4
Dipole moment (nAm)	60.5±17.2	58.6±15.9	58.2±17.4	56.2±13.9

Discussion

The obtained findings that there was no difference in the activity of the auditory cortex between only auditory stimulus and pairing of auditory and visual stimuli relating to the vowel sound, at least within 100 ms following stimulation, indicated that its activity was not influenced by visual motion given simultaneously. There have been no MEG studies to our knowledge using a paradigm similar to that used in the present study, but several studies examined related phenomena. Laurienti et al. (2002) reported that no statistical difference was detected in the auditory cortex between a combined visual-auditory stimulus and a pure auditory stimulus using fMRI. Poremba et al. (2003) reported that the auditory cortex of the rhesus monkey corresponding to HG in humans was not activated by visual stimuli but only by auditory stimuli. The results of the present study were consistent with these previous studies.

In conclusion, our findings suggest that HG, the auditory cortex in humans, processes only the acoustic and phonetic characteristics of auditory stimuli in the primary processing period without any influence by visual motion. We presume that the effects take place in the later processing period out of the primary auditory cortex such as STS or STG (Wright et al., 2003).

References

Allison T, Puce A, Spencer DD, and McCarthy G. Electrophysiological studies of human face perception. I: Potentials generated in occipitotemporal cortex by face and non-face stimuli. Cereb Cortex 1999; 9: 415-430.

Bonda E, Petrides M, Ostry D, and Evans A. Specific involvement of human parietal systems and the amygdala in the perception of biological motion. J Neurosci 1996; 16: 3737-3744.

Bundo M, Kaneoke Y, Inao S, Yoshida J, Nakamura A, and Kakigi R. Human visual motion areas determined individually by magnetoencephalography and 3D magnetic resonance imaging. Hum Brain Mapp 2000; 11: 33-45.

Calvert GA, and Campbell R. Reading speech from still and moving faces: The neural substrates of visible speech. J Cognit Neurosci 2003; 15: 57-70.

Dumoulin SO, Bittar RG, Kabani NJ, Baker CL, Jr Le Goualher G, Bruce PG, and Evans AC. A new anatomical landmark for reliable identification of human area V5/MT: A quantitative analysis of sulcal patterning. Cereb Cortex 2000; 10: 454-463.

Gattass R, and Gross CG. Visual topography of striate projection zone (MT) in posterior superior temporal sulcus of the macaque. J Neurophysiol 1981; 46: 621-638.

Halgren E, Raij T, Marinkovic K, Jousmaki V, and Hari R. Cognitive response profile of the human fusiform face area as determined by MEG. Cereb Cortex 2000; 10: 69-81.

Hamalainen M, Hari R, Ilmoniemi R. J, Knuutila J, and Lounasmaa OV. Magnetoencephalography—Theory, instrumentation, and applications to noninvasive studies of the working brain. Rev Mod Phys 1993; 65: 413–497.

Ioannides AA, Liu LC, Kwapien J, Drozdz S, and Streit M. Coupling of regional activations in a human brain during an object and face affect

recognition task. Hum Brain Mapp 2000; 11: 77-92.

Kaneoke Y, Bundou M, Koyama S, Suzuki H, and Kakigi R. Human cortical area responding to stimuli in apparent motion. Neuroreport 1997; 8: 677-682.

Kawakami O, Kaneoke Y, and Kakigi R. Perception of apparent motion is related to the neural activity in the human extrastriate cortex as measured by magnetoencephalography. Neurosci Lett 2000; 285: 135-138.

Kawashima R, Sugiura M, Kato T, Nakamura A, Hatano K, Ito K, Fukuda H, Kojima S, and Nakamura K. The human amygdala plays an important role in gaze monitoring. A PET study. Brain 1999; 122: 779-783.

Kourtzi Z, Kanwisher N. Cortical regions involved in perceiving object shape. J Neurosci 2000; 20:3310-8.

Klucharev V, Mottonen R, and Sams M. Electrophisiological indicators of phonetic and non-phonetic multisensory interactions during audiovisual speech perception. Brain Res Cog Brain Res 2003; 18: 65-75.

Laurienti PJ, Burdette JH, Wallace MT, Yen Y, Field AS, and Stein BE. Deactivation of sensory-specific cortex by cross-modal stimuli. J Cognit Neurosci 2002; 14: 420-429.

Lebib R, Papo D, Bode de S, and Baundonniere P. Evidence of a visual-to-auditory cross-modal sensory gating phenomenon as reflected by human P50 event-related brain potential modulation. Neurosci Lett 2003; 341: 185-188.

Linkenkaer-Hansen K, Palva JM, Sams M, Hietanen JK, Aronen HJ, and Ilmoniemi RJ. Face-selective processing in human extrastriate cortex around 120 ms after stimulus onset revealed by magneto- and electroencephalography. Neurosci Lett 1998; 253: 147-150.

Lu ST, Hamalainen MS, Hari R, Ilmoniemi RJ, Lounasmaa OV, Sams M, and Vilkman V. Seeing faces activates three separate areas outside the occipital visual cortex in man. Neuroscience 1991; 43: 287-290.

Maruyama K, Kaneoke Y, Watanabe K, Kakigi R. Human cortical responses to coherent and incoherent motion as measured by magnetoencepharography. Neurosci Res 2002; 44: 195-205

Maunsell JH. and Van ED. Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. J Neurophysiol 1983; 49: 1127-1147.

McCarthy G, Puce A, Belger A, and Allison T. Electrophysiological studies of human face perception. II: Response properties of face-specific potentials generated in occipitotemporal cortex. Cereb Cortex 1999; 9: 431-444.

McGurk H, and MacDonald J. Hearing lips and seeing voices. Nature 1976; 264: 746-748.

Mottonen R, Krausa CM, Tiippana K, and Sams M, Processing of changes in visual speech in the human auditory cortex. Cog Brain Res 2002; 13: 417-425.

Naito T, Kaneoke Y, Osaka N, and Kakigi R. Asymmetry of the human visual field in magnetic response to apparent motion. Brain Res 2000; 865: 221-226.

Nakamura A, Yamada T, Abe Y, Nakamura K, Sato N, Horibe K, Kato T, Kachi T, and Ito K. Age-related changes in brain neuromagnetic responses to face perception in humans. Neurosci Lett 2001; 312: 13-16.

Okusa T, Kaneoke Y, Koyama S, and Kakigi R. Random dots blinking: A new approach to elucidate the activities of the extrastriate cortex in humans. Neuroreport 1998; 9: 3961-3965.

Okusa T, Kakigi R, and Osaka N. Cortical activity related to cue-invariant shape perception in humans. Neuroscience 2000; 98: 615-624.

Oram MW. and Perrett DI. Responses of anterior superior temporal polysensory (STPa) neurons to "Biological Motion" stimuli. J Cogn Neurosci 1994; 6: 99-116.

Poremba A, Saunders RC, Crane AM, Cook M, Sokoloff L, and Mishikin M. Functional mapping of the primate auditory system. Science 2003; 299: 568-572.

Pourtois G, Gelder de B, Vroomen J, Rossion B, and Crommelinck M, The time-course of intermodal binding between seeing and hearing affective information. Neuroreport 2000; 6: 1329-1333.

Puce A, Allison T, Gore JC, and McCarthy G. Face-sensitive regions in human extrastriate cortex studied by functional MRI. J Neurophysiol 1995; 74: 1192-1199.

Puce A, Allison T, Asgari M, Gore JC, and McCarthy G. Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. J Neurosci 1996; 16: 5205-5215.

Puce A, Allison T, Bentin S, Gore JC, and McCarthy G. Temporal cortex activation in humans viewing eye and mouth movements. J. Neurosci.1998; 18: 2188-2199.

Puce A, and Allison T. Differential processing of mobile and static faces by temporal cortex. Neuroimage 1999; 9(6):S801.

Puce A, Smith A, and Allison T. ERPs evoked by viewing facial movements. Cogn Neuropsychol 2000; 17: 221-239

Puce A, Syngeniotis A, Thompson JC, Abbott DF, Wheaton KJ, Castiello U. The human temporal lobe integrates facial form and motion: Evidence from fMRI and ERP studies. Neuroimage 2003 In Press.

Sams M., Hietanen JK, Hari R, Ilmoniemi RJ, and Lounasmaa OV. Face-specific responses from the human inferior occipito-temporal cortex. Neuroscience 1997; 77: 49-55.

Sams M, Aulanko R, Hamalainen M, Hari R, Lounasmaa OV, Lu S, and Simola J. Seeing speech: visual information from lip movements modifies activity in the human auditory cortex. Neurosci Lett 1991; 127: 141-145.

Sarvas J. Basic mathematical and electoromagnetic concept of the biomagnetic inverse problem. Phys. Med. Biol 1987; 32: 11-22

Sato N, Nakamura K, Nakamura A, Sugiura M, Ito K, Fukuda H, and Kawashima,R. Different time course between scene processing and face processing: a MEG study. Neuroreport 1999; 10: 3633-3637.

Scherg M, Buchner H. Somatosensory evoked potentials and magnetic fields: separation of multiple source activities. Physiol Meas.1993; 14 Suppl 4A:A35-9.

Scott SK, and Johnsrude IS. The neuroanatomical and functional organization of speech perception. Trends Neurosci 2003; 26: 100-107.

Sekiyama K, Kanno I, Miura S, and Sugita Y. Auditory-visual speech perception examined by fMRI and PET. Neurosci Res 2003; 47: 277-287.

Sergent J, Ohta S, and MacDonald B. Functional neuroanatomy of face and object processing. A positron emission tomography study. Brain 1992; 115 Pt 1: 15-36.

Swithenby SJ, Bailey AJ, Brautigam S, Josephs OE, Jousmaki V, and Tesche CD. Neural processing of human faces: a magnetoencephalographic study. Exp Brain Res 1998; 118: 501-510.

Tanaka K, and Saito H. Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. J Neurophysiol 1989; 62: 626-641.

Taylor MJ, George N, and Ducorps A. Magnetoencephalographic evidence of early processing of direction of gaze in humans. Neurosci Lett 2001; 316: 173-177.

Terasaki O, and Okazaki M. Transcallosal conduction time measured by visual hemifield stimulation with face images. Neuroreport 2002; 13: 97-99.

Tootell RB, Reppas JB, Dale AM, Look RB, Sereno MI, Malach R, Brady

TJ, and Rosen BR. Visual motion aftereffect in human cortical area MT revealed by functional magnetic resonance imaging [see comments]. Nature 1995a; 375: 139-141.

Tootell RB, Reppas JB, Kwong KK, Malach R, Born RT, Brady TJ, Rosen BR, and Belliveau JW. Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. J Neurosci 1995b; 15: 3215-3230.

Uusitalo MA, Virsu V, Salenius S, Nasanen R, and Hari R. Activation of human V5 complex and rolandic regions in association with moving visual stimuli. Neuroimage 1997; 5: 241-250.

Van Essen DC, Maunsell JH, and Bixby JL. The middle temporal visual area in the macaque: myeloarchitecture, connections, functional properties and topographic organization. J Comp Neurol 1981; 199: 293-326.

Van Oostende S, Sunaert S, Van Hecke P, Marchal G., Orban G.A. The kinetic occipital (KO) region in man: an fMRI study. Cereb Cortex 1997; 7:690-701.

Watanabe S, Kakigi R, Koyama S, and Kirino E. It takes longer to recognize the eyes than the whole face in humans. Neuroreport 1999a; 10: 2193-2198.

Watanabe S, Kakigi R, Koyama S, and Kirino E. Human face perception traced by magneto- and electro-encephalography. Brain Res Cogn Brain Res 1999b; 8: 125-142.

Watanabe S, Kakigi R, and Puce A. Occipitotemporal activity elicited by viewing eye movements: a magnetoencephalographic study. Neuroimage 2001; 13: 351-363.

Watanabe S, Miki K, and Kakigi R. Gaze direction affects human face perception in humans. Neurosci Lett 2002; 325: 163-166.

Watanabe S, Kakigi. R., Puce. A. The spatiotemporal dynamics of the face inversion effect: A magneto- and electro-encephalographic study. Neuroscience 2003; 116: 879-895.

Wheaton KJ, Pipingas A, Silberstein RB, and Puce A. Human neural responses elicited to observing the actions of others. Vis Neurosci 2001;