

Vocalization-associated magnetic fields in humans

Atsuko Gunji

Doctor of Philosophy

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

2000 (School Year)

**Vocalization-associated magnetic fields
in humans**

Atsuko Gunji

Doctor of Philosophy

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

and

Department of Integrative Physiology,
National Institute for Physiological Sciences,
Myodaiji, Okazaki, Aichi, 444-8585, Japan.

2000 (School Year)

Contents

1. Abbreviation	3
2. Abstract	5
3. Introduction	7
4. Methods	10
4.1. Subjects	
4.2. Experimental protocol	
4.2.1. VRCF/VRCP recordings	
4.2.2. Triggering for VRCF/VRCP recordings	
4.2.3. AEF recording	
4.2.4. MEG recording	
4.2.5. EEG recording	
4.3. Data analysis	
4.3.1. Brain electric source analysis	
4.3.2. RMS analysis and difference waveforms	
4.3.3. Analysis of single equivalent current dipole (ECD)	
4.3.4. EEG data analysis	
5. Results	25
5.1. VRCF during a simple vowel vocalization	
5.2. VRCF under the control and masking conditions	
5.3. AEF according to one's own voice	
5.4. VRCP under the control and masking conditions	
6. Discussion	48
6.1. Cortical activities just before and after the onset of vocalization	
6.2. The masking effect during a simple vocalization	

1. Abbreviations

AEF	Auditory evoked field
AEP	Auditory evoked potential
ANOVA	Analysis of variance
BESA	Brain electric source analysis
BP	Bereitschaftspotential
ECD	Equivalent current dipole
EEG	Electroencephalography
fMRI	Functional magnetic resonance imaging
F0	Voice fundamental frequency
GFP	Global field power
GOF	Goodness of fit
MEG	Magnetoencephalography
MF	Motor field
MP	Motor potential
MRCF	Movement related cortical field
MRCP	Movement related cortical potential
M1	primary motor cortex
M100	The magnetic response peaking at about 100 ms
N1	The first negative potential
PA	Pre-auricular point
PET	Positron emission tomography
P1	The first positive potential
RMS	Root mean square
RP	Readiness potential
RV	Residual variance

SMA	Supplementary motor area
SQUID	Superconducting quantum interference device
VRCF	Vocalization related cortical field
VRCP	Vocalization related cortical potential
1M	The first magnetic response

2. Abstract

Cortical activities during vocalization were recorded in humans, to identify the motor and auditory components in the vocalization related cortical field (VRCF) by averaging magnetoencephalography (MEG).

Ten right-handed healthy volunteers were examined. For VRCF recording, they were instructed to vocalize a simple vowel (/u/) under two conditions: (1) no masking (control) and (2) masking of the subject's own voice by weighted-white noise during vocalization. Auditory evoked magnetic field (AEF) was also recorded following stimulation of a speech sound applied by voice-recorder to compare the AEF response with the auditory component of VRCF.

Generators of cortical activities just before and after the onset of a simple vowel vocalization were analyzed using software capable of analyzing multiple sources, brain electric source analysis (BESA). Two sources in the primary motor cortex (M1), laryngeal and truncal areas, were identified in bilateral hemispheres, and these were activated prior to the vocalization onset. Sources in the bilateral auditory cortices were also activated just after the vocalization onset. The source localization and time course of their activities were similar in each hemisphere.

The difference waveform, which was obtained by subtracting the VRCF of the masking condition from those of the control, showed a 1M component after the vocalization onset. AEF

following sound stimulus also showed M100 component. There was no significant difference between the peak latencies of 1M and M100 components. The location of the equivalent current dipole (ECD) for the 1M component was close to that of M100 in the auditory cortex in each hemisphere, but, in the left hemisphere, the ECD source for 1M was located slightly but significantly medially to that for M100.

An adequate masking condition could separate the auditory response to one's own voice from the VRCF complex. Although the auditory response in VRCF might be generated in the auditory cortex, the difference in the ECD localization between 1M and M100 in the left hemisphere may indicate a different neural process for auditory feedback after the vocalization onset from that for simple hearing.

3. Introduction

Vocalization is a complex performance which involves sensorimotor preparatory process, motor acts and auditory responses to one's own voice. Especially, the auditory response to voice after the vocalization is concerned as a feedback process for precise vocalization. Since McAdam and Whitaker (1971) reported the Bereitschaftspotential (BP) prior to the speech onset, studies of speech mechanisms have been carried out in neurophysiological field using electroencephalography (EEG). Vocalization-related cortical potential (VRCP), which has been considered to reflect the cortical activities for vocalization, involves sensory, motor and auditory components. Although several characteristics of VRCP have been reported (Grabow et al., 1974; Grözinger et al., 1979, 1980; Deecke et al., 1986; Wohlert and Larson, 1991; Kikuchi and Kita, 1993; Wohlert, 1993; Gunji et al., 1998b), the nature of their generators is not fully understood. From the results of studies in normal subjects using positron emission tomography (PET) (McGuire et al., 1996; Price et al., 1996; Hirano et al., 1997; Murphy et al., 1997; Perry et al., 1999) and functional magnetic resonance imaging (fMRI) (Wildgruber et al., 1996), and of lesion studies in patients using computerized axial tomography (CT) and magnetic resonance imaging (MRI) (Dronkers, 1996), responsible areas for vocalization are considered to involve primary motor cortex (M1), supplementary motor area (SMA), Broca's area, auditory cortex and insula. However, the radiological studies, such as PET and

fMRI, revealed little about the temporal course of activities in those areas.

Magnetoencephalography (MEG), which has been recently developed and is a non-invasive tool to record the cortical neural activities, has several theoretical advantages in identifying the localization of generator sources with high spatial and temporal resolution (Kakigi et al., 2000). Although some MEG studies (Salmelin et al., 1994; Sasaki et al., 1995; Kuriki et al., 1999) have demonstrated the probable brain generators engaged in the language production processes, a basic and systematic study focusing on a simple vocalization has not been carried out. Since I have surmised that vocalization related-cortical field (VRCF) comprises multiple generator sources, the analysis using a brain electrical source analysis (BESA) (Scherg and Berg, 1995), which is a strategy for the spatiotemporal multiple dipole analysis, would be suitable for the study of the VRCF. The BESA attempts to identify the localization of multiple equivalent current dipoles (ECDs) and to image the current waveform of each ECD. It is another matter of my interest whether the “auditory response” for one’s own voice during a vocalization is similarly processed to that for the external sound or not.

Therefore, I focused on and analyzed the sensori-motor and auditory processes before and after the onset of a simple vowel vocalization. The two objectives of the present study were (1) to evaluate the generator sources of VRCF and to investigate the spatio-temporal relationship among those activities during a simple vocalization, and (2) to disclose the auditory evoked

response to one's own voice during vocalization, demonstrating the generator location and the temporal course of its activities of the auditory response. This is the first systematic study to clarify the spatial and temporal characteristics of sensori-motor and auditory activities relating to vocalization.

4. Methods

4.1. Subjects

Ten healthy right-handed subjects (7 males and 3 females; mean age 34 years, range 24-46 years) participated in the experiment. Their informed consent to participate in the experiment, which was first approved by the Ethical Committee of the National Institute for Physiological Sciences, Okazaki, Japan, were obtained prior to the study.

4.2. Experimental protocol

4.2.1. VRCF/VRCP recordings

During the task, subjects rested comfortably on their side on a bed in a magnetically shielded room. Subjects were asked to perform the vocalizing tasks while keeping their body relaxed, jaw opened slightly and tongue rested on the floor of their mouth, with eyes fixed on a target point. They were instructed to utter a specific vowel (/u/) repeatedly, with a self-paced random interval of about 5.0 s. A specific vowel (/u/) was used for the following two reasons; (1) it is a vowel with no effect caused by consonants, and (2) the activities of oral and tongue muscle for pronouncing (/u/) are minimal among the vowels.

The responses were recorded under two conditions (Figure 1A): (1) the control condition without masking and (2) the masking condition. In the masking condition, weighted-white noise (AA-71 audiometer, RION) was continuously delivered at 75 dB SPL or more through a pair of plastic tubes and ear-pieces

A. Vocalization Task

control



masking



B. Hearing Task

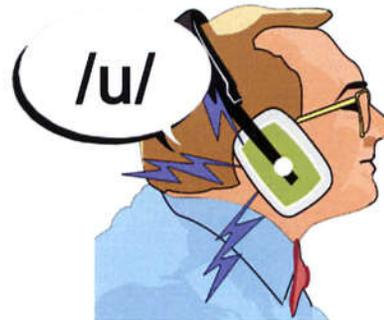


Figure 1. An experimental protocol for the vocalization task and the hearing task. In the vocalization task, subjects were asked to pronounce a specific vowel (/u/) repeatedly, with a self-paced random interval of about 5.0 s. The MEG and EEG were recorded under two conditions: (A) the control condition without masking and (B) the masking condition. In the masking condition, weighted-white noise (AA-71 audiometer, RION) was continuously delivered at 75 dB SPL or more through a pair of plastic tubes and ear-pieces (for MEG recording) or headphones (for EEG recording), which attenuated air conducted signals from each subject's own voice.

(for MEG recording) or headphones (for EEG recording), which could minimize air and bone conducted signals from each subject's own voice.

Prior to recording, the subjects practiced the vocalization guided visually by a 14-inch color monitor screen on which sound waveforms of their own voice were presented. For the performance during the masking condition, subjects were trained to keep their voice with constant and similar intensity level to that of the control condition while experimenters view a sound level meter, since they tended to vocalize at higher intensity level when their voice is masked (the Lombard effect). During both conditions, the experimenter continued to check the sound pressure of the subject's voice and instructed the subject to raise or lower the vocalization sound when the sound pressure was inappropriate.

One session in MEG recording consisted of thirty trials of the vocalization. Four sessions for the control condition and four sessions for the masking condition were alternately tested with a short rest between each session for one subject. One hundred and twenty (30 x 4) trials for both control and masking conditions were thus recorded.

One session in electroencephalography (EEG) recording consisted of twenty-five trials of vocalization. Eight sessions for the control condition and eight sessions for the masking condition were alternately tested with a short rest between each session for one subject. Two hundred (25 x 8) trials for both control and masking conditions were thus recorded.

4.2.2. Triggering for VRCF/VRCP recordings

To obtain an appropriate trigger for averaging epochs, the subject's sound waveforms of vocalization were picked up from a microphone placed close to their mouth (Figure 2). The trigger for averaging the epoch was generated from the vocalization sound wave. In order to obtain an appropriate trigger, two microphones were used to pick up the subject's vocalization. One microphone (Mic.1) was placed close to the subject's mouth, and the other (Mic.2) was placed 30 cm above the mouth. The vocalization sound wave was amplified with a time constant of 0.3 s and a high-frequency cutoff less than 1000 Hz. Mic.1 obtained a sharp onset of the sound waveform, but it also detected the noise from breathing or other sounds. Mic.2 recorded only the vocalization sound, but the onset of the waveform was relatively dull. Thus, the trigger point was decided at the waveform onset from Mic.1 for the epoch, in which the sound waveforms were recorded by both Mic.1 and Mic.2. Using a signal processor (SIGAVG, CED, UK), the sound waveform was rectified, and a trigger pulse was obtained by matching amplitude to the threshold level.

4.2.3. AEF recording

A vowel (/u/) sound of each subject was presented by voice-recorder (Figure 1B). The duration of the vowel sound was approximately 100 ms including 5 ms rise and fall times. The parameters of the sound for recording AEF were similar to

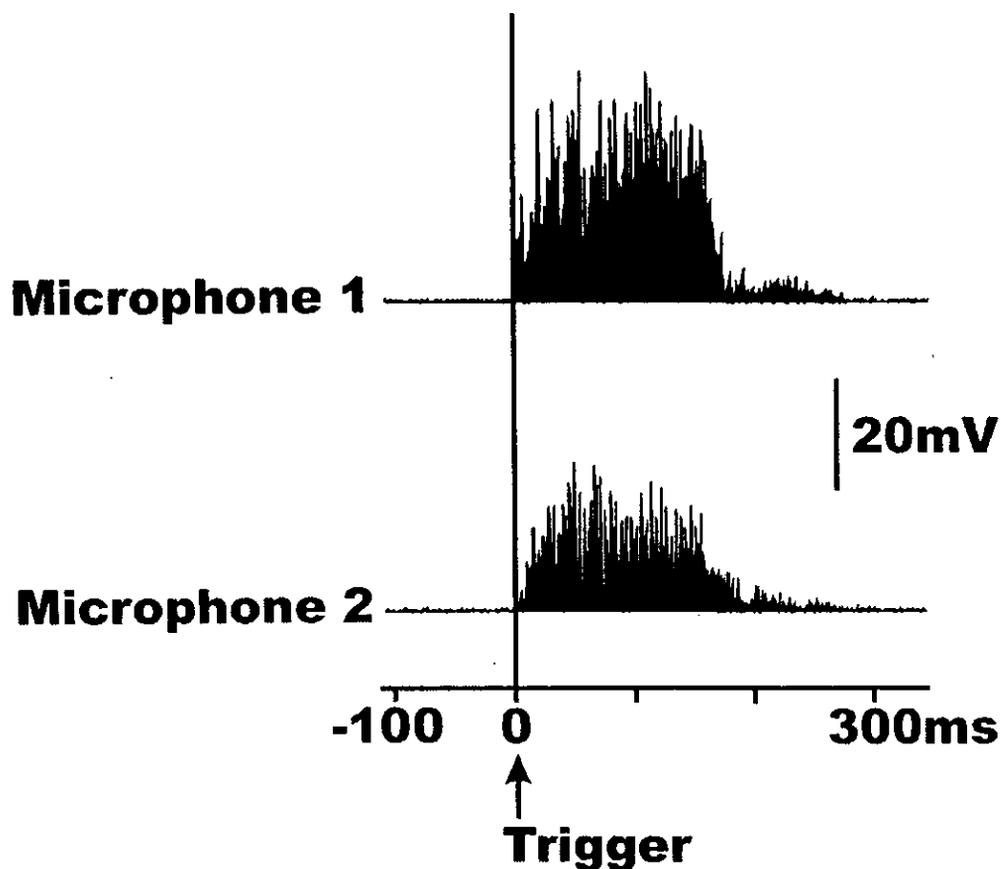


Figure 2. Sound waveforms of a vowel /u/ pronounced by a subject, recorded from two microphones (Mic.1 and Mic.2). In order to obtain an appropriate trigger, we used two microphones to pick up the subject's vocalization. Microphone 1 (Mic.1) was placed close to the subject's mouth, and microphone 2 (Mic.2) was placed about 30 cm above the mouth. Mic.1 obtained a sharp peak onset of the sound waveform, but it also detected the noise from breathing or other sounds. Mic.2 recorded only the vocalization sound, but the onset of waveform was relatively dull. Thus, the sound waveforms recorded by Mic.1 and Mic.2 were compared, and it was decided as trigger point at the onset of waveform recorded by Mic.1.

those of the vowel sound. The averaged frequencies of the first, second and third formants were 403 Hz, 617 Hz and 2489 Hz, respectively, for both sounds. Only the recorded sound stimulus was filtered with a high frequency cut-off of 4000 Hz but this difference was not meaningful. The recorded sound stimuli were delivered to subjects at 65 dB SPL through a pair of plastic tubes and earpieces with a random interval of between 450 and 550 ms. One hundred fifty trials were recorded.

4.2.4. MEG recording

The MEG was measured with dual 37-channel biomagnetometers (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 magnetometer's entire surface was 144 mm in diameter and its radius of curvature was 122 mm. The outer coils in each device were 72.5 degrees apart. Each coil was connected to a superconducting quantum interference device (SQUID).

With the MEG system, dual biomagnetometers are centered at C3 and C4 positions of the International 10-20 EEG system, which covered in all instances the left and right hemispheres, in each subject (Figure 3). These sites were measured at the same time. Movement-related cortical potentials (MRCP), averaged EEG, have been usually recorded with a wide-band high-pass

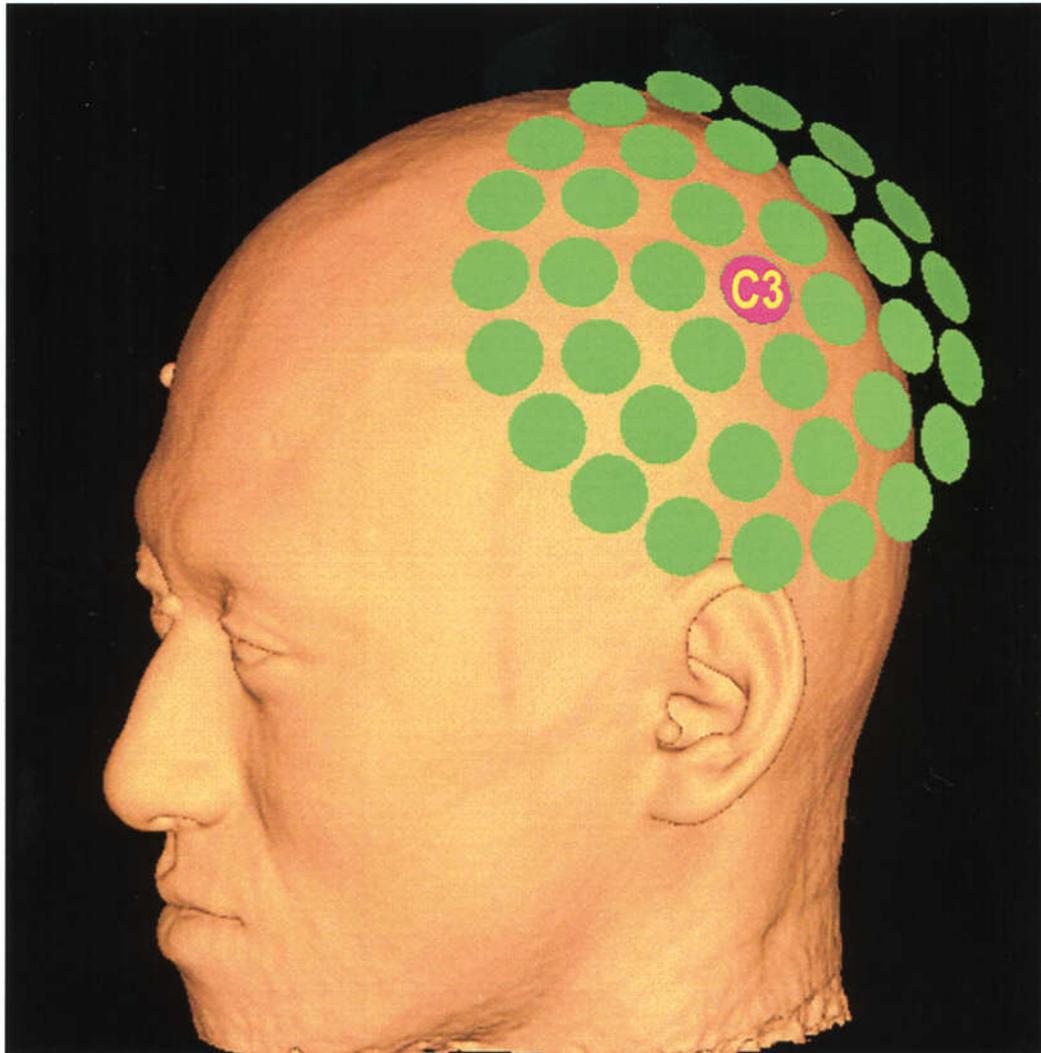


Figure 3. Placement of the MEG devices. The devices were centered on around C3 (left hemisphere) and C4 (right hemisphere) of the international 10-20 system in each subject. This figure shows the MEG device placement in the left hemisphere.

filter, which is close to DC, to record BP. However, since the baseline of an MEG is largely shifted when using such a wide-band filter, and I wanted to focus on activities just before and after the vocalization onset rather than BP, the magnetic fields were recorded with a relatively narrow-band filter of 0.1-100Hz.

Epochs with artifacts of eye movement were rejected before analysis. Epochs without artifacts from each of the subject's four sessions were respectively collected and averaged for control and masking conditions in each subject. The epochs were digitized during an analysis time from 1400 ms before to 400 ms after the trigger in each condition at a sampling rate of 1024 Hz. The baseline was corrected (DC offsets) for each channel according to the mean value of the signal between 1400 ms and 1000 ms before the trigger.

4.2.5. EEG recording

The EEG was recorded using Ag/AgCl disk electrodes. Electrodes were placed at Fz, Cz, Pz, C3, C4, T3 and T4 on the scalp according to the International 10-20 system. The reference was linked earlobes. Electrooculogram (EOG) of the left eye was also recorded with the electrodes placed 1.5 cm above the superior orbicular edge and 2 cm outside of the lateral angle of the eye.

The impedance between exploring and referential electrodes was less than 5 k Ω . EEG was amplified with a band-pass filter, 0.1-100 Hz. EOG was recorded with a band-pass filter,

0.06-100 Hz.

The data for EEG, EOG and sound waveforms were transferred on-line and stored with a sampling rate of 1000 Hz by a signal processor (SIGAVG, CED, UK). All data were also stored on floppy disk for later analysis.

Epochs with artifacts of eye movement were rejected before analysis. Epochs without artifacts from each of the subject's eight sessions were respectively collected and averaged for control and masking conditions in each subject. The EEG epochs in each condition were averaged from 1500ms before to 500ms after the trigger in each condition. DC-offset was done using the mean value from 1500 ms to 1250 ms before the onset of vocalization.

4.3. Data analysis

4.3.1. Brain electric source analysis

In the control condition, VRCFs for eight subjects were analyzed using the brain electric source analysis (BESA*) (Scherg and Berg, 1995) software package (NeuroScan, Inc., Mclean, VA) for computation of theoretical source generators in a 3-layer of spherical head model. BESA was modified for our 2x37 channel magnetometers placed in bilateral hemispheres. This method allows the spatiotemporal modeling of multiple simultaneous sources over defined intervals. A six-step strategy for localizing the generators of the VRCF was applied independently to the waveforms from each subject. At each step, the location and orientation of dipoles were calculated by an

iterative least-square fit. The residual variance (%RV) indicated the percentage of data which could not be explained by the model. The goodness of fit (GOF) was expressed in % as $100 - \%RV$, and signal epochs for source analysis were defined on the basis of the global field power (GFP) (Lehman et al., 1986).

The period from 150 ms before to 150 ms after the vocalization onset was selected for analyzing the spatiotemporal dipole modeling, because this interval was considered to cover the motor potentials for vocalization and the auditory feedback regarding sufficient vocalizing. At first, dipole model was tentatively made from; sources 1 and 2 located in the laryngeal area of M1 in the left and right hemispheres, respectively, and sources 3 and 4 in the auditory cortex in the left and right hemispheres, respectively, but this four dipole model was not appropriate to account for the recorded waveforms. Therefore, we added sources 5 and 6 in the truncal area of M1 in the left and right hemispheres, respectively.

In step 1, a dipole solution for the activity in the time range from 150 ms to 0 ms before vocalization onset was calculated in each hemisphere. Sources 1 and 2 were put in the left and right laryngeal motor area respectively, and then the orientations of sources 1 and 2 were determined with temporary symmetrical location (step 1a). The location constraint was released, and bilateral locations were fitted independently retaining the orientations obtained in step 1a (step 1b). Then, the orientations of sources 1 and 2 were fitted again retaining the location obtained in step 1b (step 1c).

In step 2, a dipole solution for the activity in the time range from 0 to 150 ms after vocalization onset was calculated in each hemisphere. I put sources 3 and 4 around the left and right auditory cortices, respectively, since they were suspected to be activated in response to the vocalization sound. During the fitting of sources 3 and 4, the location and orientation of sources 1 and 2 were fixed and switched off. Sources 3 and 4, which were restricted in location, were free of the orientation of the ECDs (step 2a). These locational constraints were based on a PET study reported by Hirano et al. (1997), as discussed later. After we determined the orientations of sources 3 and 4, the locations of these sources were calculated again, keeping the orientation constraints of step 2a, in the same manner as step 1 (step 2b). The orientations of both sources were then fitted with the new locational constraint (step 2c). Finally, sources 1 and 2 were switched back on. All four sources were refitted in the time interval between 150 ms before and 150 ms after the vocalization onset (step 3).

In step 4, the sources responsible for the activity in the time range from 150 ms before to 150ms after the vocalization onset were calculated in each hemisphere. At this time, sources 5 and 6 were put in the left and right truncal motor area, respectively, since they would be also activated for vocalization. During the fitting of sources 5 and 6, sources 1, 2, 3 and 4 were constrained to the location and orientation of each source in step 1 or 2 in each hemisphere and were switched off. The orientation and location of sources 5 and 6 were calculated in the same

manner as in step 1 and 2 (step 5). The source 5 and 6 locations were constrained to be mirror images. Finally, the four sources (1 to 4) were switched back on and refitted the locations and orientations of all six sources which were obtained in above steps (step 6).

In step 6, GOF using the locations and orientations of the sources obtained were calculated for the period from 150 ms before to 150 ms after the vocalization onset. I also tried to put other sources to increase GOF in step 7 without any locational or orientational constraints. The adaptation of the dipoles could be acceptable when the GOF (%) was more than 90%.

4.3.2. RMS analysis and difference waveforms

The maximum amplitude of the outgoing and ingoing magnetic fields of each component was evaluated by root mean square (RMS) for the VRCF waveforms under the control and the masking conditions. Using a paired t-test, maximum RMS values were compared between the control and the masking conditions.

Subtracting the waveform in the masking condition from the control, difference VRCF/VFCP waveforms were obtained. The peak latency of the RMS was calculated for the subtracted waveform.

4.3.3. Analysis of single equivalent current dipole (ECD)

For the AEF and the waveforms, a spherical model was fitted to the digitized head-shape of each subject, and the location

(x, y and z positions) (Figure 4), the orientation and the magnitude of a single equivalent current dipole (ECD) were computed at a peak latency of RMS. The zero point of the three-dimensional measurement was the point exactly halfway between the pre-auricular points (PAs). The x-axis indicates the antero-postero direction, with positive x coming out of the head at the nasion. The z-axis indicates the ventral direction, with positive values toward the upper side. The y-axis indicates the vertical direction with positive values toward the left PA. A single-ECD analysis based on the non-linear inverse problem reported by Sarvas (1987). The mean correlation value indicates how closely the measured values correspond to the theoretical field generated by the model.

A single ECD was also estimated at the peak latency of RMS for the AEF while the subject heard vowel sound stimuli. The latency and the dipole locations (x, y and z positions) were compared between those of the different waveforms and the AEF waveforms using analysis of variance (ANOVA). When comparing activities between each hemisphere using ANOVA, the dipole moment value ($|Q|$) indicating the dipole strength in nAm was analyzed, since RMS was variable depending on the placement of the MEG device and on the depth of the ECD in each hemisphere.

Magnetic resonance imaging (MRI) scans (Shimadzu Magnex 150 XT 1.5T, Shimadzu, Kyoto, Japan) were obtained for all subjects. The T1-weighted coronal, axial and sagittal images with continuous slices 1.5 mm in thickness were adopted

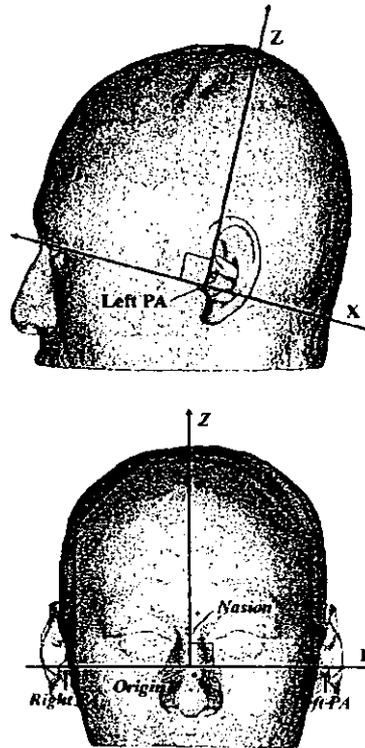


Figure 4. The head-based coordinate system. The origin was the point exactly halfway between the pre-auricular points (PAs). The x-axis indicated a line extending through the origin and the nasion, with positive x coming out of head at nasion. The z-axis was a line extending through the origin and the top of the head, with positive values toward the upper side. This axis was perpendicular to the plane formed by the left and right PAs and nasion. The y-axis was a line perpendicular to the x-axis and z-axis extending through the origin and the sides of the head, with positive values toward the left PA.

for overlays with the ECD sources detected by the different (control – masking) waveforms related to vocalizing and the AEF waveforms related to hearing speech. The same anatomical landmarks were used to create the MEG head-based three-dimensional (3D) coordinate system based on the location of the nasion and bilateral pre-auricular points, and the landmarks were visualized in the MR images by affixing to these points.

4.3.4. EEG data analysis

The peak amplitude and latency from baseline to peak of VRCP and the amplitude at the vocalization onset were measured in the control and masking conditions. Using two-way analysis of variance (ANOVA) of conditions and electrodes (General linear model in SYSTAT 7.0 SPSS Inc.), the values were compared between the control and masking conditions. The difference waveform was calculated between the control and masking conditions, the peak latency and amplitude of the deflections were measured in the different waveforms.

5. Results

All subjects successfully performed the vocalization tasks with minimal face movement and blinking, and the VRCF and VRCP were recorded before and after the vocalization onset under both control and masking conditions. All of the subjects reported that they were unable to hear their own voice under the masking condition. The artifact-free 124 ± 18 and 143 ± 8 MEG epochs for the control and masking conditions, 140 ± 32 and 138 ± 36 EEG epochs for the control and masking conditions were collected. The latency before and after the vocalization onset (zero point) were described as minus and plus values, respectively. For example, latencies 200 ms before and 150 ms after the vocalization onset are described as -200 ms and +150 ms, respectively.

5.1. VRCF during a simple vowel vocalization

In the control condition, the 37 superimposed waveforms recorded from C3 and C4 positions are shown in Figure 5. The VRCF was clearly identified from approximately -150 ms, and a peak of reversed-phase deflection was recognized at about +90 ms. This trend of waveforms was similarly recorded in the bilateral hemispheres in all subjects.

Low %RV with six ECDs was obtained during the period from -150 to +150 ms. Figures 6, 7 and 8 show the current waveform in each source (left) and the source localization and orientation of six sources estimated on the spherical head model

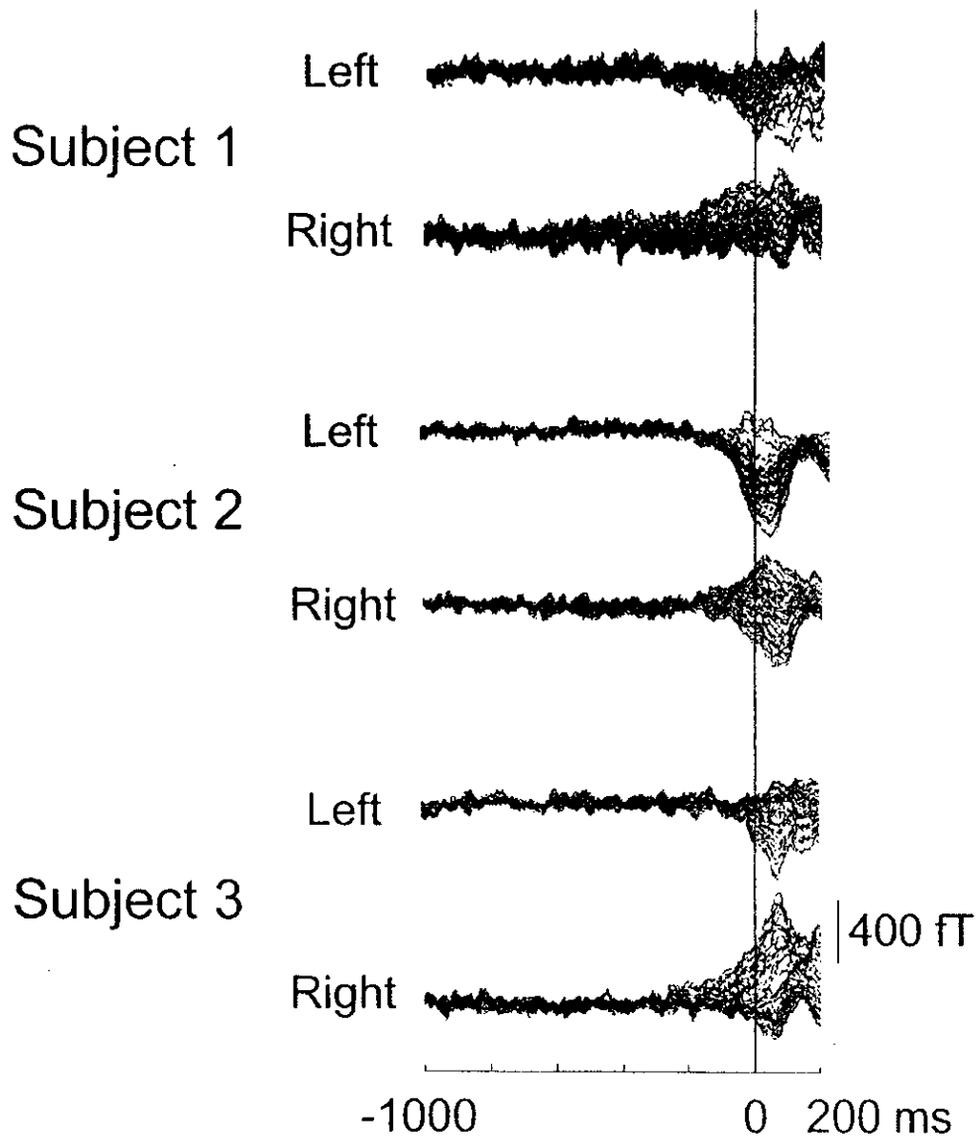


Figure 5. Vocalization related cortical field (VRCF) in the control condition following the vocalization for three subjects. The waveforms recorded from 37 channels at the left and right hemispheres (around C3 and C4) were superimposed. Activities before the vocalization in the left hemisphere appear similar than those in the right hemisphere in subject 3, but this difference was not in the other subjects.

Subject 1 (GOF=91.4%)

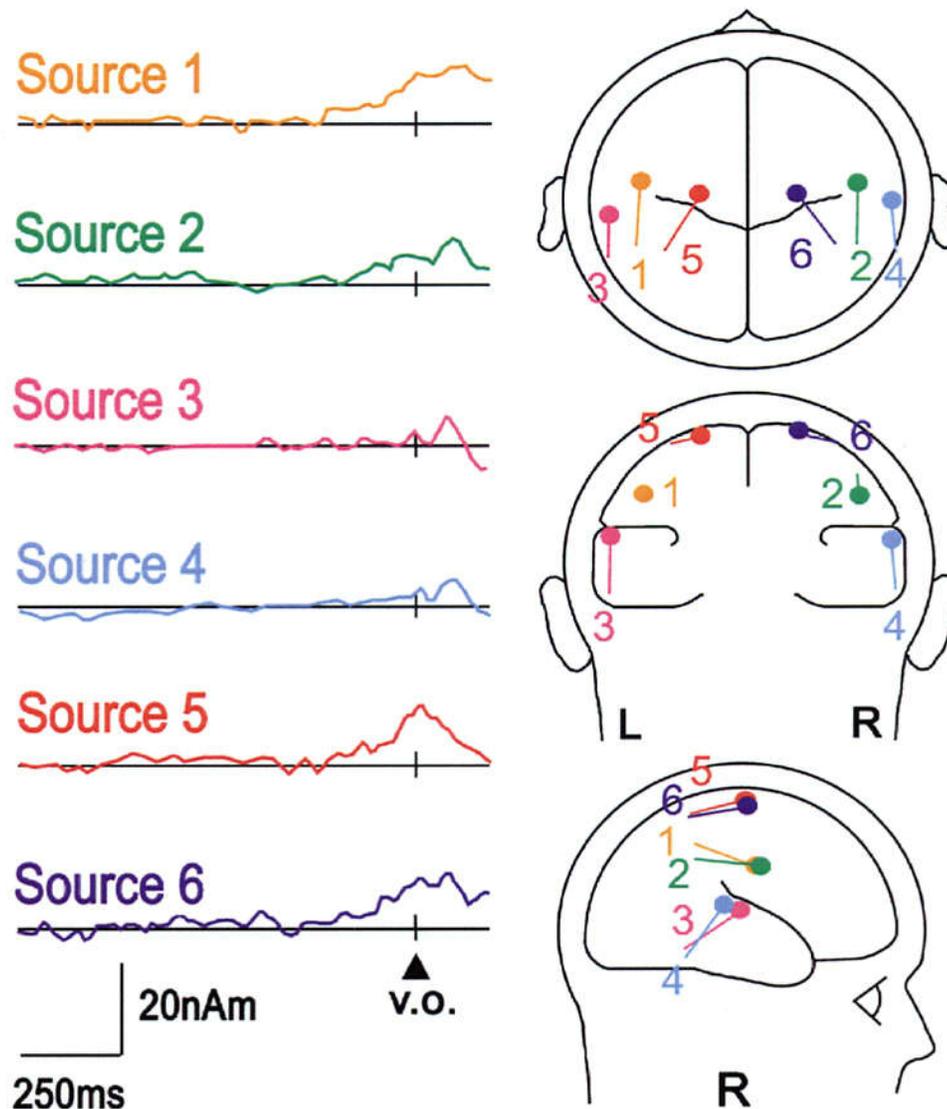


Figure 6. The result of the source analysis for subject 1. Left: temporal activity of each source obtained by spatiotemporal source analysis (BESA). Right: the localization and the orientation of the dipole on the spherical head model. The line from each point indicates the direction of the dipole current. Sources 1 and 2 were located in the laryngeal motor areas. Sources 3 and 4 were located in the auditory areas, and sources 5 and 6 were in the truncal motor areas. The sources in the motor areas (sources 1, 2, 5, 6) were activated approximately 100 ms prior to the vocalization onset (V.O.), while the activity of the auditory sources (sources 3 and 4) appeared after the vocalization onset. All six sources were temporally overlapped after the vocalization onset.

Subject 2 (GOF=91.6%)

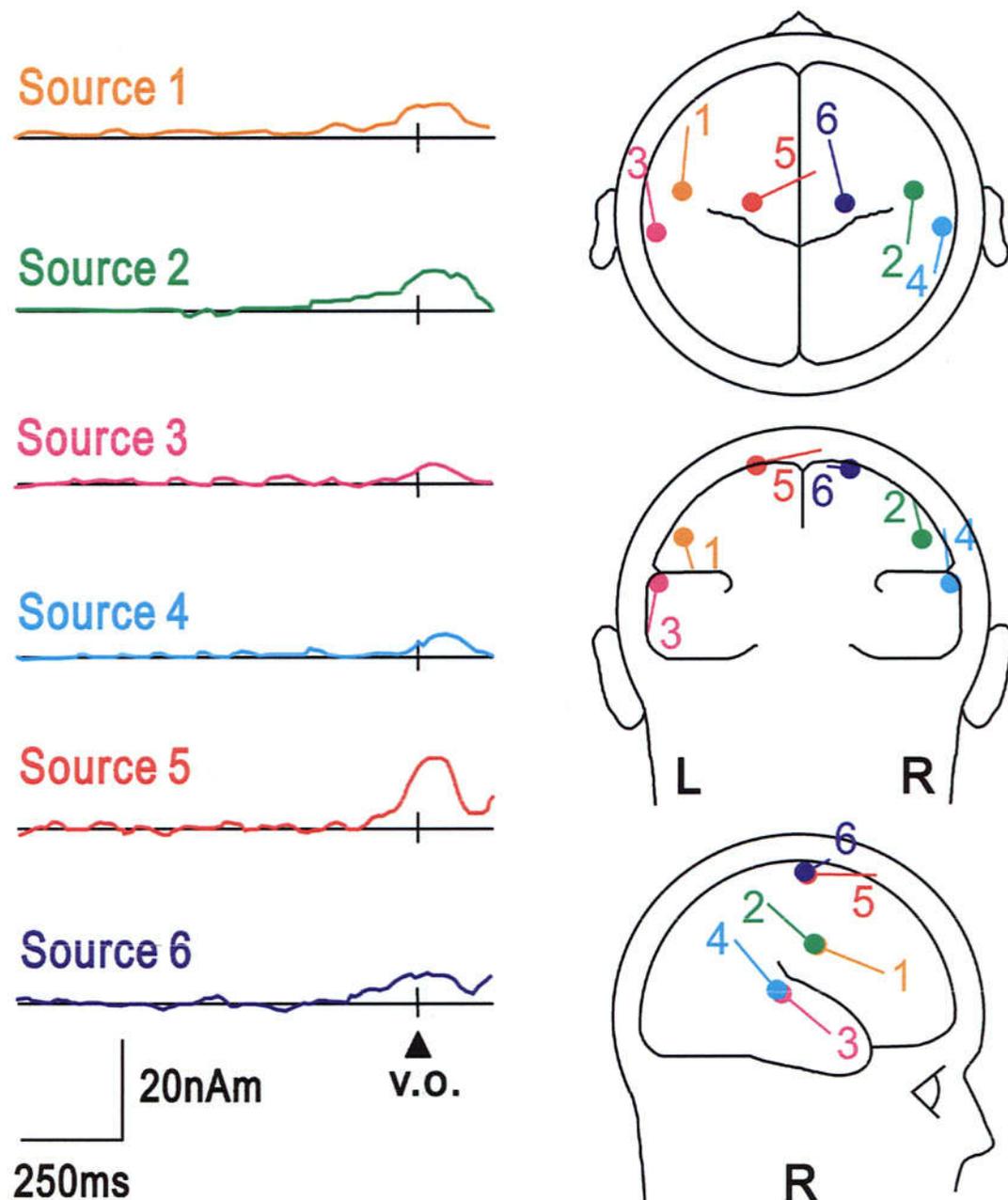


Figure 7. The result of the source analysis for subject 2 (see Fig.5 legend for details). The temporal changes of activity were basically similar to those for subject 1, although some inter-individual differences were recognized.

Subject 3 (GOF=91.2%)

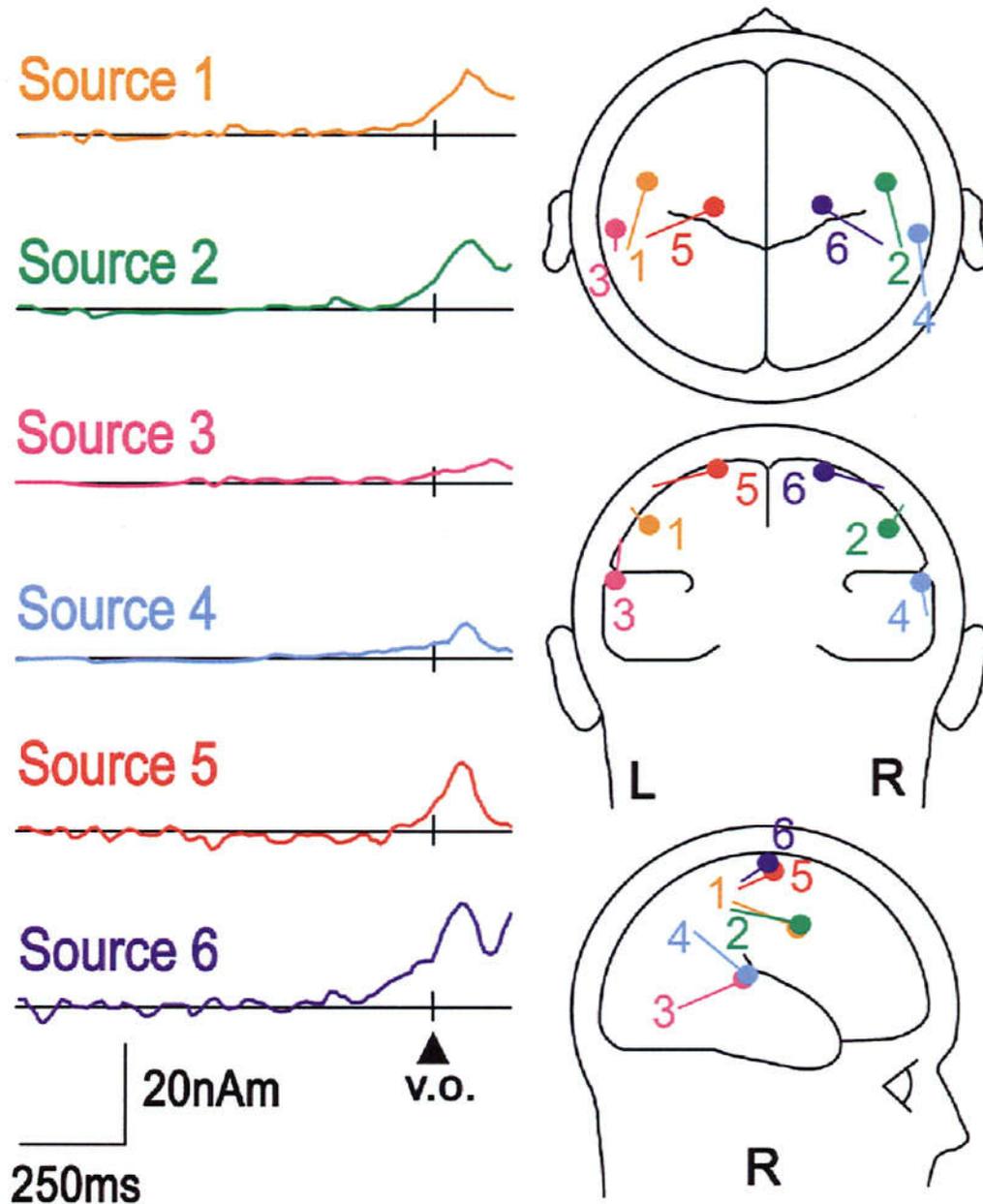


Figure 8. The result of the source analysis for subject 3 (see Fig.5 legend for details). The temporal changes of activity were basically similar to those for subject 1, although some inter-individual differences were recognized.

by BESA (right) in three representative subjects. The dipole direction shown in the spherical head models means the positive direction of the waveforms. Sources 1 and 2 were located in the area inferior to the hand motor area of the left and right hemispheres, respectively. Sources 3 and 4 were located in the auditory area, the upper part of the temporal lobe, of the left and right hemispheres, respectively. The locations of sources 5 and 6 were in the motor area around truncal motor areas of the left and right hemispheres. The RV for the six-dipole model was low enough, i.e., 9.0 ± 1.76 (mean \pm SD) in seven subjects, but one subject shows relatively high RV (15.4 %). We tried to find, but no source showed an appropriate waveform or improved the RV.

Table 1 shows the onset and peak latency of the waveforms of the six sources, and the three-dimensional location of the sources. The time courses of the activities of sources 3 and 4 were clearly different from those of the other sources. The major peak of sources 3 and 4 appeared after the vocalization onset, while the other sources' onset was 150-200 ms before the vocalization onset. However, the mean peak latency of all 6 sources overlapped approximately 50-100 ms after the vocalization onset, indicating that activities in the motor cortices continued after the vocalization onset.

5.2. VRCF under the control and masking conditions

The VRCF waveform under each condition for a subject is shown in Figure 9. In the control condition, activities in the

Table 1. Mean onset latency and mean peak latency of each source, and the three-dimensional dipole localization. The latency before and after the vocalization onset are described as minus and plus values, respectively. The zero point of the three-dimensional measurement was the point exactly halfway between the pre-auricular points (PAs). The x-axis indicates the antero-postero direction, with positive x coming out of the head at the nasion. The z-axis indicates the ventral direction, with positive values toward the upper side. The y-axis indicates the vertical direction with positive values toward the left PA.

	Motor area (laryngeal)		Auditory area		Motor area (truncal)	
	Source 1 Left	Source 2 Right	Source 3 Left	Source 4 Right	Source 5 Left	Source 6 Right
Onset (ms)	-244±93	-196±93	32±56	28±25	-173±47	-168±72
Peak (ms)	94±36	70±40	88±28	95±26	48±33	62±29
X (cm)	3.6±0.8	3.5±0.9	1.2±0.6	1.2±0.6	3.7±0.8	3.7±0.8
Y (cm)	5.7±0.5	-6.0±0.3	6.9±0.4	-7.2±0.4	2.2±0.2	-2.4±0.4
Z (cm)	7.9±0.6	7.9±0.4	6.4±0.5	6.3±0.5	11.5±0.4	11.6±0.5

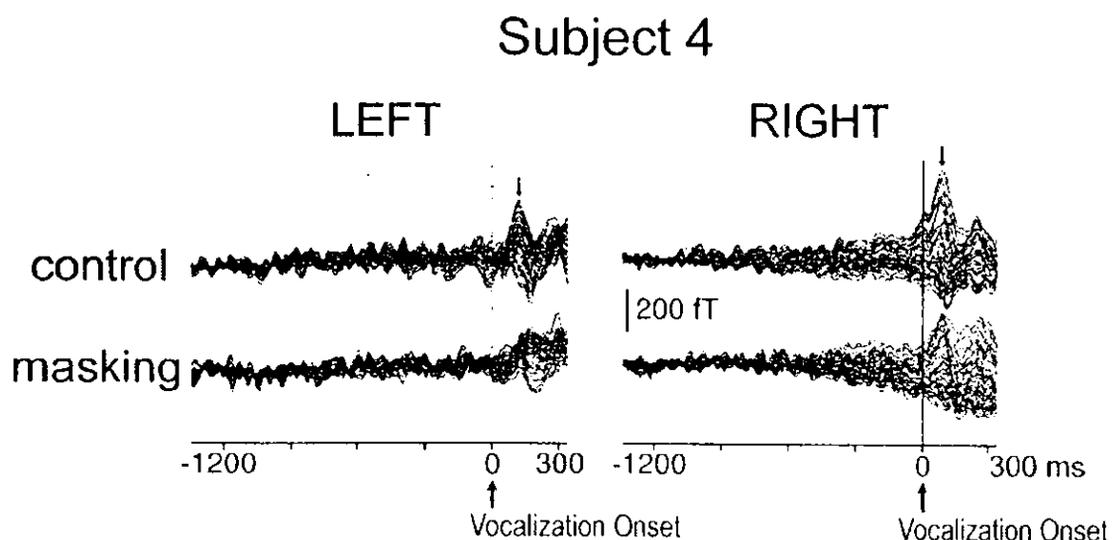


Figure 9. VRCF recorded under the control and the masking conditions for subject 4. The VRCF was identified from approximately -150 ms, and a peak of reversed-phase deflection was recognized after the vocalization onset (V.O.) in both conditions. The deflection after the vocalization onset was significantly larger in the control condition than in the masking condition. Activities before the vocalization in the left hemisphere appear similar than those in the right hemisphere in this subject, but this difference was not in the other subjects.

VRCF were identified from approximately -150 ms, and large phase-reversed deflections were recognized after the vocalization onset. A similar trend of waveforms was recorded in each hemisphere in all subjects. In the masking condition, on the other hand, phase-reversed deflections after the vocalization onset were markedly small, compared to those in the control condition. The maximum RMS values were 156.4 ± 68.4 (mean \pm SD) fT at 76.7 ± 26.9 ms in the control condition and 119.1 ± 41.1 fT at 55.8 ± 48.9 ms in the masking condition (Figure 10). The RMS value in the control condition was significantly large, compared to the masking condition ($P < 0.05$).

Difference waveforms were obtained by subtracting the VRCF waveforms for the masking condition from those for the control condition (Figure 11 and 12). In nine of ten subjects, the first component (1M) with the outgoing and ingoing magnetic fields was recognized approximately at +85 ms, but one subject did not show any components. There was no consistent component before the vocalization onset. The correlation between the recorded measurements and the values expected from the ECD estimate was calculated. The mean value of the correlation of the 1M was 0.93. There was no significant difference of latency, locations and dipole moment between hemispheres (Table 2).

5.3. AEF according to one's own voice

Waveforms of AEF following vowel sound stimuli were obtained. These AEF components appeared at 60.0 ± 13.6 ms

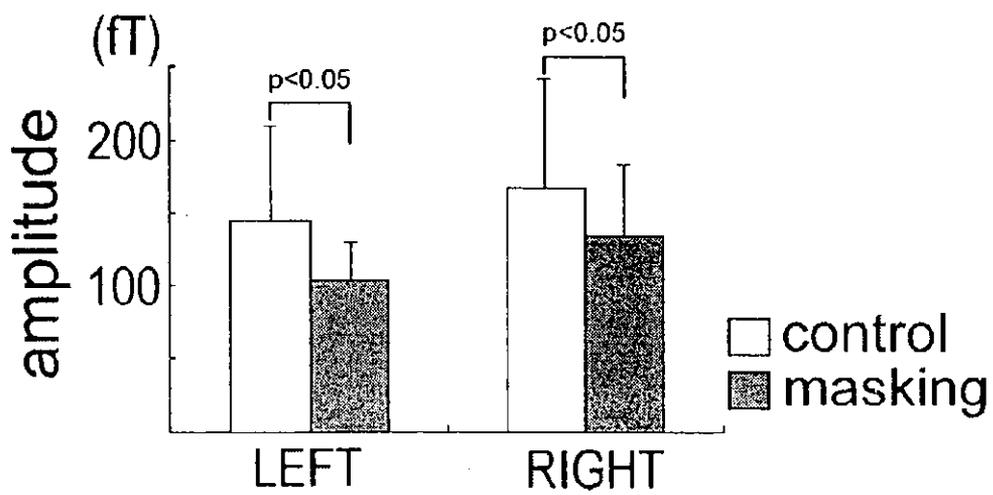


Figure 10. Mean peak RMS amplitudes of a component after the vocalization onset in the control (open columns) and masking (hatched columns) conditions. The amplitude in the control condition was larger than that in the masking condition ($p < 0.05$).

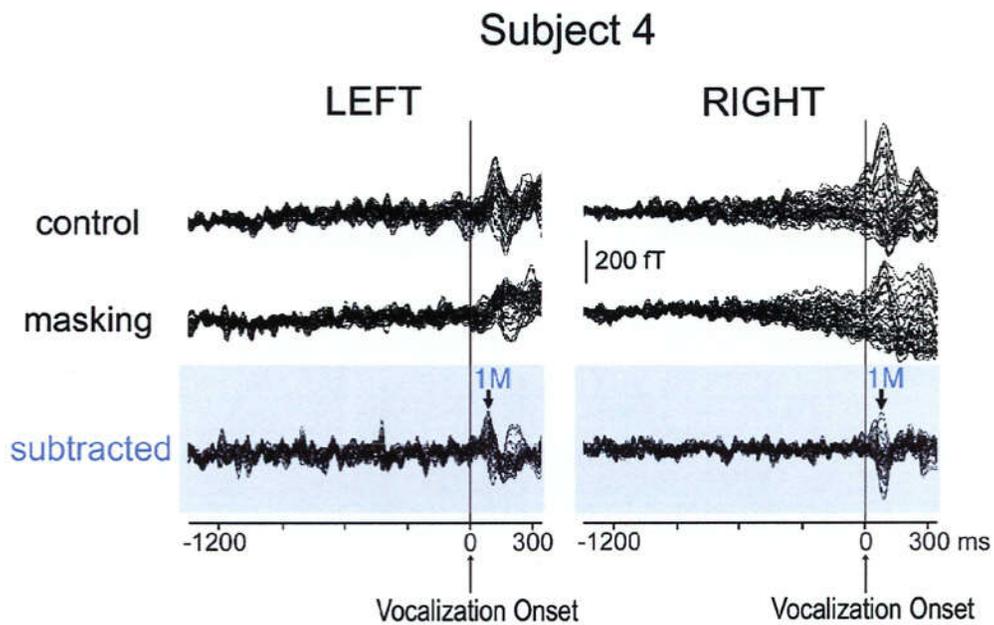


Figure 11. VRCF recorded under the control and the masking conditions, and the subtracted waveform for subject 4. The subtracted waveform was obtained by subtraction of the masking condition waveform from the control condition waveform. The first magnetic response (1M) was recognized at approximately +80 ms. There was no consistent response before the vocalization onset.

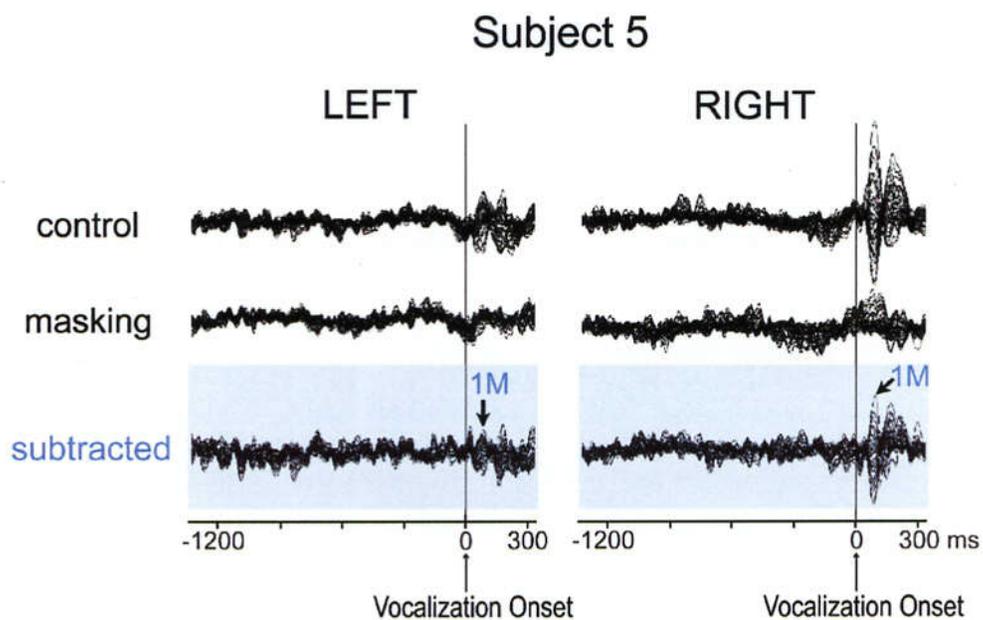


Figure 12. VRCF recorded under the control and the masking conditions, and the subtracted waveform for subject 5 (see Fig.11 legend for details). 1M was also recognized at approximately +80 ms.

Table 2. Dipole locations (x, y and z positions), dipole moment ($|Q|$) and peak latency (ms) of the 1M of the difference waveforms and M100 of the AEF waveforms. Mean and one standard deviation are shown. There was no significant difference of the peak latency, dipole locations and dipole moment between each hemisphere. There was no significant difference of peak latency, x and z positions and dipole moment between 1M and M100. The y value of 1M was significantly smaller than that of M100 in the left hemisphere ($p < 0.05$). Bold indicates that the 1M was located more medial than the M100.

	Left Hemisphere		Right Hemisphere	
	M100	1M	M100	1M
Dipole location				
x (cm)	1.0±0.8	1.0±1.2	1.4±1.0	2.1±1.3
y (cm)	5.5±0.7	4.5±1.1	-5.9±0.4	-5.5±1.4
z (cm)	6.3±0.7	6.8±1.2	5.8±0.6	5.7±0.9
Dipole moment $ Q $ (nAm)				
Peak latency (ms)	12.4±6.4	23.9±17.3	14.4±7.1	17.4±16.9
Correlation	0.98±0.02	0.94±0.03	0.98±0.02	0.92±0.08

(M50) and 94.3 ± 18.4 ms (M100) after the stimuli. These values were based on the mean peak latency of each component recorded in the left and right hemispheres. The M50 component was not determined in some subjects. Figure 13 shows the AEF waveforms for vowel sound stimuli and the subtracted waveforms of VRCF between the control and masking conditions in two subjects. To indicate the peak latency of M100 component, we calculated the RMS of AEF waveforms. The mean value of the correlation of the M100 was 0.98. There was also no significant difference of the peak latency, x, y and z positions and dipole moment between hemispheres (Table 2).

Comparing the 1M for the subtracted waveforms with the M100 for the AEF waveforms, there were no significant differences of peak latency, x and z positions. The y value of the 1M was significantly smaller than that of the M100 in the left hemisphere ($p < 0.05$), indicating that the 1M was located more medial than the M100.

The ECDs of the M100 for the AEF waveforms and the 1M for the subtracted waveforms were estimated around the auditory cortex in each hemisphere. Their locations were very similar to each other on the MRI, but the 1M was located more medial than the M100 in the left hemisphere (Figure 14).

5.4. VRCF under the control and masking conditions

The characteristic patterned VRCF in the control and masking conditions for a subject is shown in Figure 15.

In the control condition, VRCF showed a negative slope,

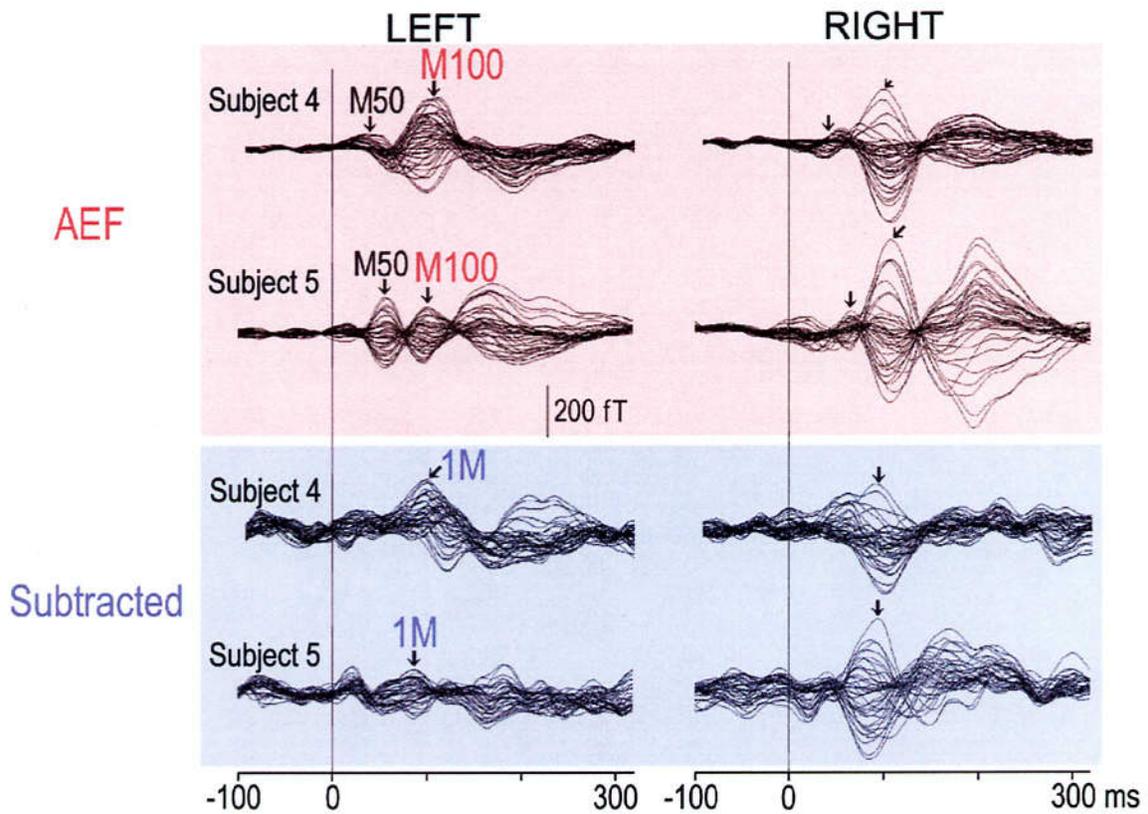


Figure 13. The subtracted waveforms of VRCF and the AEF waveforms for vowel sound stimuli for subjects 4 and 5. The subtracted waveforms are the same as in Figures 11 and 12. The AEF components appeared at 60.0 ± 13.6 ms (M50) and 94.3 ± 18.4 ms (M100). The latencies of the M100 were similar to those of the 1M.

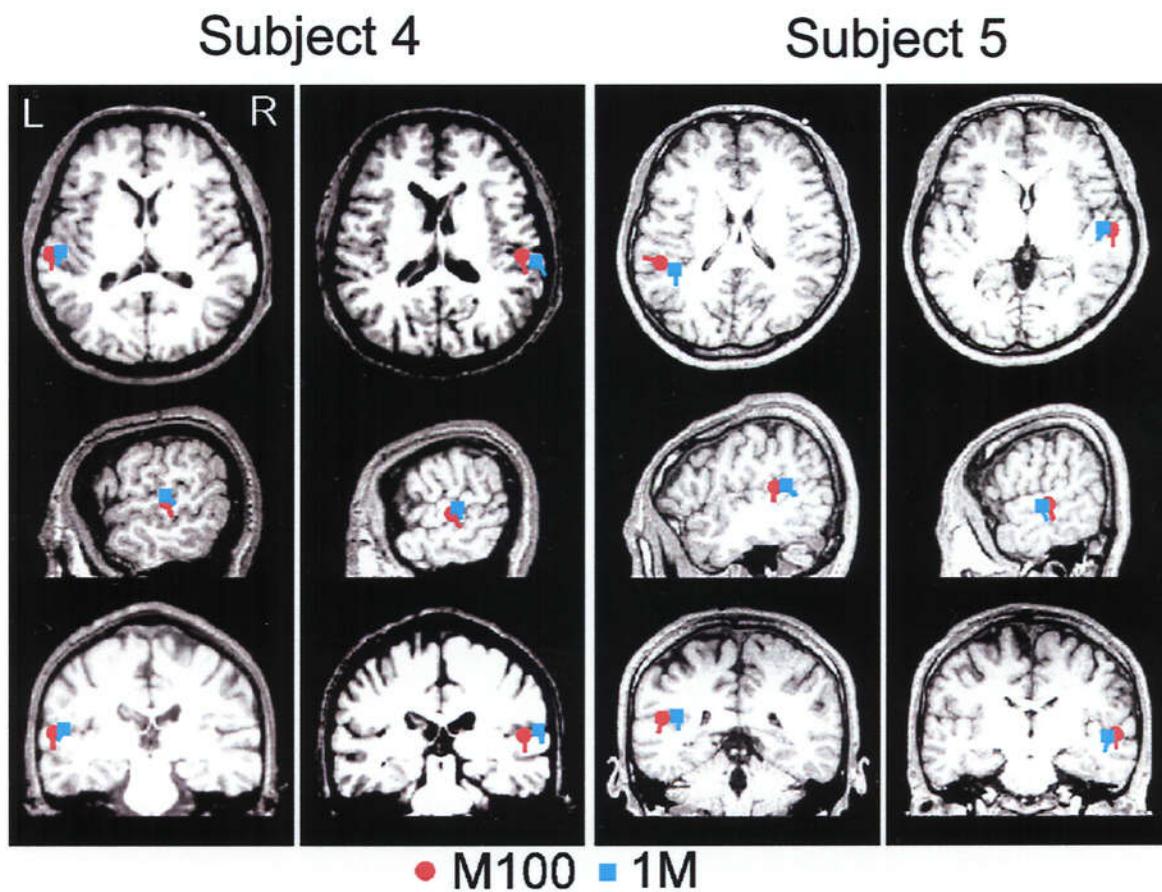


Figure 14. Location of ECDs of the M100 (red) and the 1M (blue) components for subjects 4 and 5. The ECDs of M100 for the AEF waveforms and 1M for the subtracted waveforms were located similarly in the auditory cortex, but the 1M was located more medial than the M100 in the left hemisphere

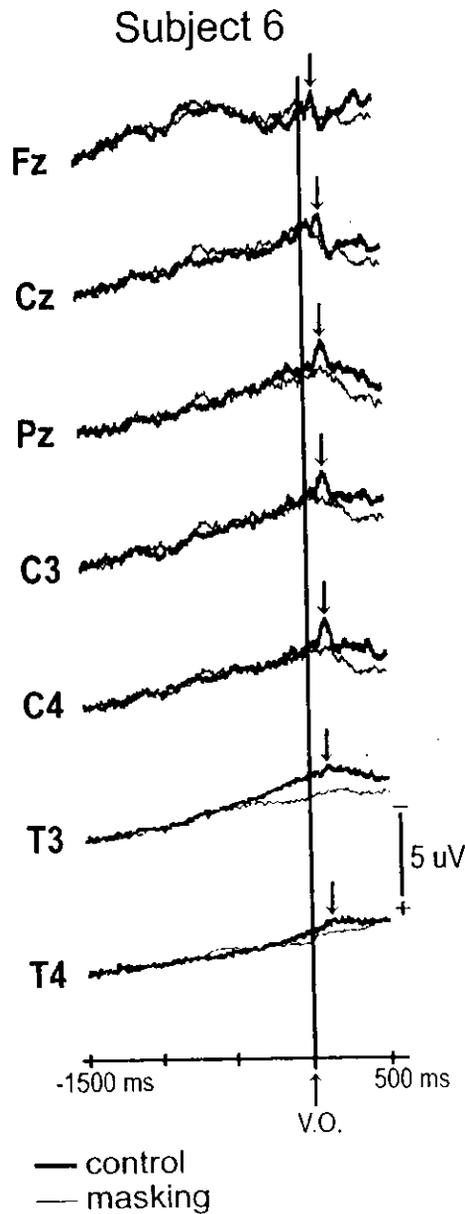


Figure 15. Vocalization related cortical potentials (VRCP) recorded from one subject. The bold lines indicate the VRCP in the control condition (no masking) and the standard lines those in the masking condition. The negative slope started approximately 1.0 sec before vocalization onset at the vertex and the central electrodes (Cz and Pz) in both conditions. The negative potential with a relatively sharp peak indicated by the arrows was identifiable after vocalization onset only in the control condition. In the masking condition, however, the negative potential decreased just after the vocal onset without producing a sharp peak.

which started about 1.0 sec before the vocalization onset, being dominant in the vertex and central areas, and a relatively sharp peak with a maximum value at approximately +80 ms. In temporal areas, the sharp negative peak was not evident although the negative slope peaked at approximately +80 ms.

In the masking condition, the negative slope preceding vocalization was similarly dominant in the vertex and central areas. However, the waveform lacked the sharp peak after the vocalization, and the peak of the negative slope was just after the vocalization onset. There was no consistent difference of waveform in both conditions between each hemisphere.

Figure 16 shows the grand average waveforms of VRCP recorded during vocalization under the control and masking conditions in all subjects. Although there were inter-individual differences, all subjects showed similar changes of VRCP for each condition. A sharp peak of VRCP was recognizable with a maximum value at $+78.9 \pm 20.9$ (mean \pm SD) ms and -2.0 ± 22.6 ms for the control and masking conditions. ANOVA in peak latency of VRCP revealed a condition main effect ($p < 0.001$). The difference was significant after the vocalization onset, but not before. Figure 17 shows the amplitude of the negative peak at each electrode for the control and masking conditions. It was significantly larger in the control condition at all electrodes than in masking condition ($p < 0.02$), but the electrode main effect did not reach significant level.

Difference waveforms obtained by subtracting the VRCP in the masking condition from the control condition were shown in

Grand average (n=9)

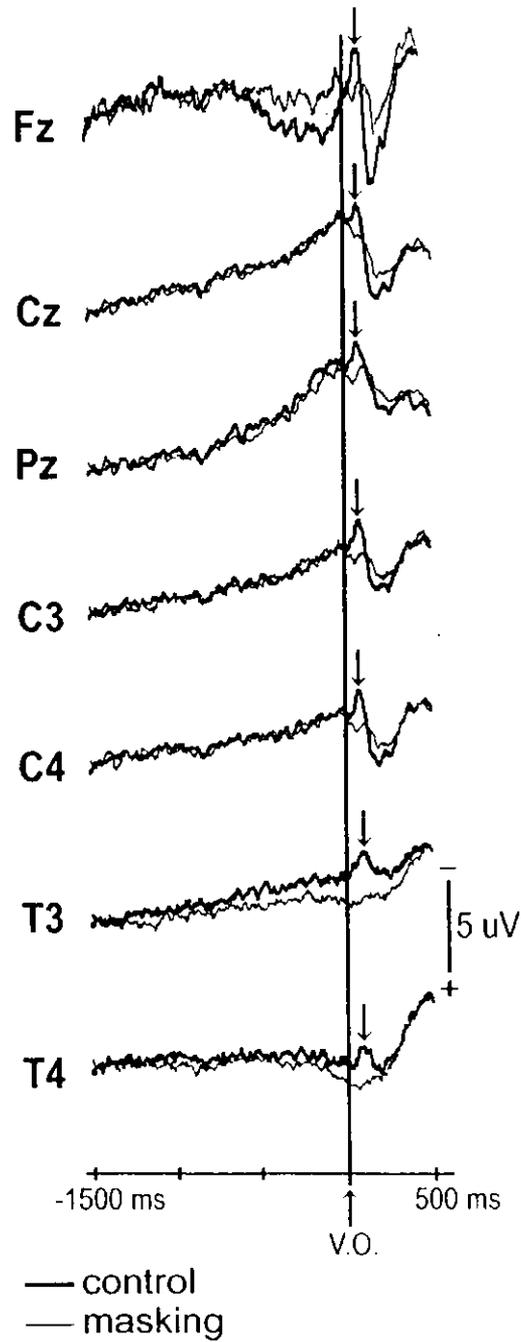


Figure 16. Grand average waveform of VRCP in the control (bold line) and the masking (standard line) conditions for all subjects.

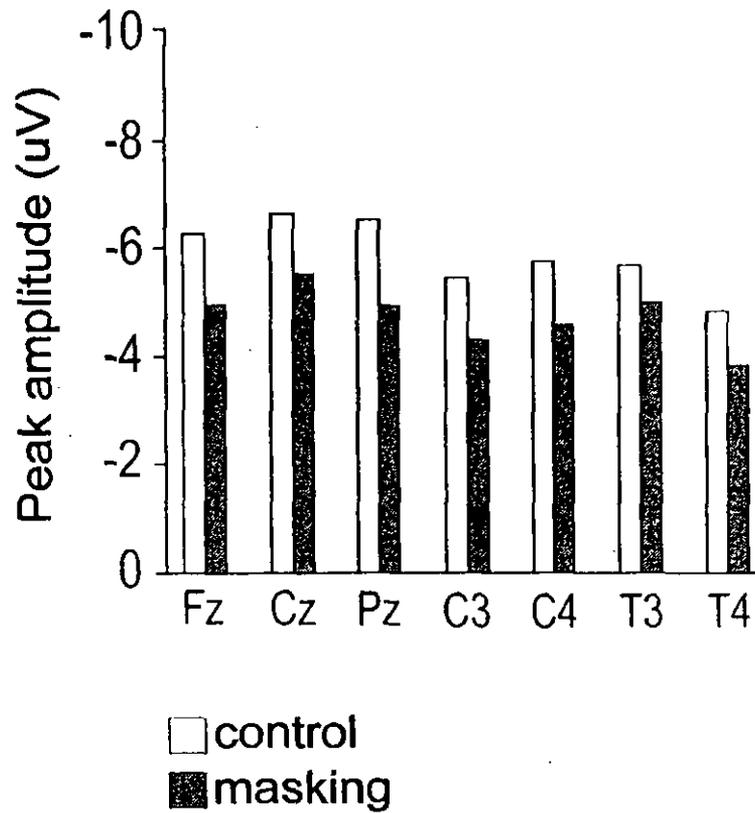


Figure 17. Mean amplitudes of the negative potentials in the control (open columns) and masking (hatched columns) conditions. The amplitude in the control condition was larger than that in the masking condition at each electrode ($p < 0.02$).

Figure 18. One clear biphasic (negative and positive, N1 and P1, respectively) deflection was recognized at Fz, Cz, Pz, C3 and C4 electrodes, and its onset was just after the vocalization onset. At the T3 and T4 electrodes, a negative peak after the vocalization was identified but a positive deflection following the negative peak was not clearly recognized. The latency and amplitude of the N1 and P1 are shown in Table 3.

Grand average (n=9) of
the subtracted waveforms

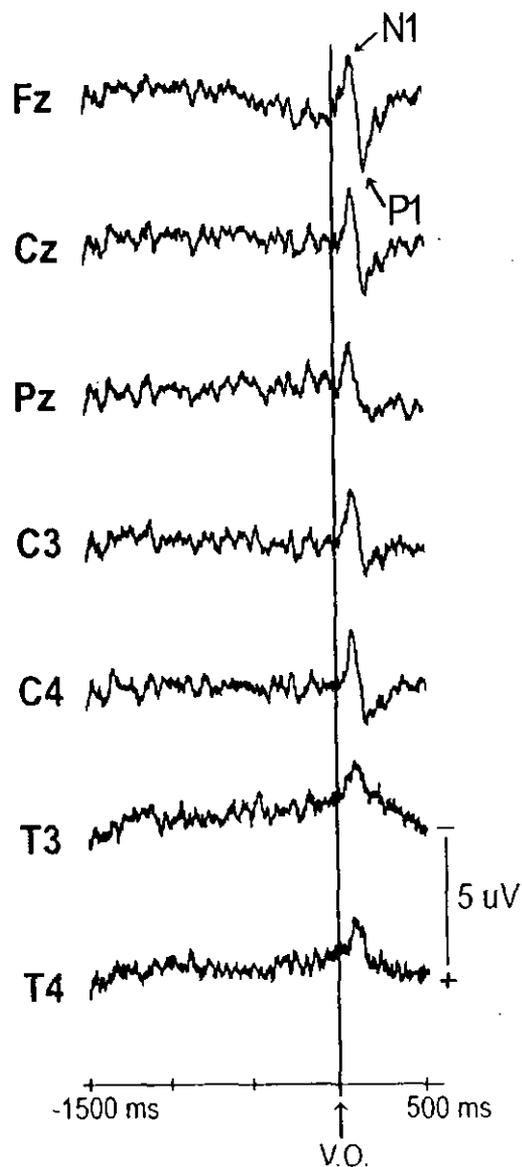


Figure 18. Grand average "subtracted waveform" of VRCP obtained by subtraction of the masking condition waveform from the control condition waveform. Biphasic potentials (N1-P1) were recorded at each electrode after the vocalization onset (V.O.) except for the temporal electrode where only N1 was identified. There was no consistent response before vocalization onset.

Table 3. Peak latencies and amplitudes (mean and standard deviation) of the N1 and P1 components at each electrode obtained by subtracting the VRCP in the masking condition from the control condition (see Fig. 18). The P1 component was not clearly identified in all subjects at the T3 and T4 electrodes.

	Latency (ms)		Amplitude (μ V)	
	N1	P1	N1	N1-P1
Fz	77.6 \pm 24.6	167.0 \pm 31.1	2.5 \pm 2.4	4.7 \pm 2.3
Cz	78.7 \pm 21.9	168.8 \pm 38.1	3.2 \pm 3.0	4.0 \pm 2.0
Pz	77.8 \pm 24.9	169.4 \pm 33.7	2.1 \pm 2.9	2.6 \pm 1.0
C3	80.7 \pm 29.2	163.4 \pm 31.5	2.3 \pm 2.5	3.5 \pm 2.1
C4	80.6 \pm 21.6	161.6 \pm 28.8	2.4 \pm 2.7	3.6 \pm 1.4
T3	75.5 \pm 26.5	-	1.3 \pm 1.3	-
T4	84.7 \pm 24.1	-	2.2 \pm 2.5	-

6. Discussion

The VRCF and VRCP were, in general, successfully recorded both in the control and the masking conditions in the present study. Components of VRCF and VRCP appeared before the vocalization onset and showed a peak after the vocalization onset. I considered that the components before the vocalization onset consisted of the preparatory brain activities which probably included components similar to BP and motor field/potential (MF/MP) in movement related cortical field (MRCF) (Nagamine et al., 1994; Hoshiyama et al., 1997) and movement related cortical potential (MRCP) (Deecke et al., 1976; Shibasaki et al., 1980). However, unlikely to the simple motor act, brain activities after the vocalization onset include the auditory response to one's own voice, which is one of the issues focused on in the present study.

Human vocalization is produced by movements of articulation related organs, i.e. tongue, lips, larynx and respiratory muscles. Cortical activities recorded from the surface often suffer from artifacts of muscle activities and/or rapid impedance change due to the movement occurred at vocalization of related organs (Szirtes and Vaughan, 1977; Brooker et al., 1980). Since the present study used a simple vowel (/u/), these artifacts could be minimized.

6.1. Cortical activities just before and after the onset of vocalization

The spatiotemporal change of the VRCF under the control condition was analyzed, and a six-dipole model in VRCF was identified using BESA. I confirmed that the VRCF involves (1) the activities in the motor cortices which were probably responsible for the initiation and preparation for vocalization, and (2) the activities in the auditory cortices related to one's own voice during self-vocalization. The ECDs estimated in the source model were in the laryngeal and truncal areas of the 'homunculus' of the sensori-motor cortices (Penfield and Boldley, 1937), and in the auditory cortices.

The reason for the restricted location of sources 1 and 2 given by the BESA is based on the evidence for bilateral laryngeal activation of the motor cortices during vocalization studied by PET (McGuire et al., 1996; Hirano et al., 1997) and fMRI (Wildgruber et al., 1996). The peak latencies of these components were extended slightly after the vocalization onset. In studies of MRCF (Nagamine et al., 1994; Hoshiyama et al., 1997) and MRCP (Kristeva et al., 1991), the peak latency of motor activity was just after the movement onset. In this study, the subject could articulate a fixed vowel (/u/) without the movements of their lips. Thus, sources 1 and 2 were considered to involve the motor activity of laryngeal muscle and/or vocal cords mainly.

Small components were observed in the auditory cortices, on the inferior bank of the Sylvian fissure (sources 3 and 4). The responses of those sources started just after the vocalization onset, while the activation of the sources in motor cortices appeared

from about 250-150 ms before the vocalization onset. The peak latencies of sources 3 and 4 were both about 90 ms after the vocalization onset. The previous studies (Lawson and Gaillard, 1981; Eulitz et al., 1995; Kuriki et al., 1995; Näätänen et al., 1997; Gross et al., 1998; Hosokawa et al., 1998, 1999; Tiitinen et al., 1999; Gootjes et al., 1999; Nummien and Curio, 1999; Shtyrov et al., 1999; Curio et al., 2000) reported that the peak latency of the brain response after hearing speech sounds was approximately 100 ms (M100/N1). In addition, the locations of sources 3 and 4 were similar to the auditory areas estimated in these studies. Therefore, I considered that the activities of sources 3 and 4 corresponded to the activity related to the response of subjects' central auditory system to their own voice, probably including M100/N1 activity.

Each activity in motor areas showed a peak after the vocalization onset, because of sustained motor activity in each area. The temporal changes of the activities in sources 5 and 6 appeared to be very similar to those of sources 1 and 2, and their locations seemed to be in the truncal part in the homunculus of the motor cortex in the bilateral hemispheres. This finding indicated that the truncal muscles including the abdominal muscle contribute to produce the voice (Higgins et al., 1998; Hough and Klich, 1998). The activation of a similar motor area was reported in a PET study (Price et al., 1996) which substantially supports our present results. Hirano et al. (1997), in a PET study, reported the activities in the visual cortex, Broca's area, SMA, left Heschl's gyrus and cerebellum during the vocalization

of sentences, in addition to the activity of the motor cortex. Although they did not describe clearly the detailed anatomical site of the motor cortex, their findings suggested activities in both truncal and laryngeal areas. Some investigators (Nakamura et al., 1998; Itomi et al., 2000) have recorded somatosensory-evoked magnetic field (SEF) following the stimulation of trunk using mostly the same subjects as in the present study. Compared the location of source for SEF with sources 1, 2, 5 and 6 for VRFCF, y and z positions were very stable in each source. The y position of VRFCF was more frontal than that of SEF in each hemisphere, indicating that sources 1, 2, 5 and 6 for VRFCF were located in trunk and laryngeal areas of motor cortices. Therefore, the locations of sources 1 to 6 for VRFCF under natural vocalization in the present study are generally compatible with the activated areas reported by Hirano et al. (1997), except for the visual area, Broca's area and SMA.

No definite evidence of SMA activity was obtained during the period analyzed. It is generally accepted that the SMA plays an important role during the preparation period of movement (Cheyne and Weinberg, 1989; Kristeva et al., 1991) and speech (Grözinger et al., 1979, 1980; Wohlert, 1993). However, the previous MEG studies failed to show SMA activity during voluntary movements (Kristeva et al., 1991; Hoshiyama et al., 1997). There may be at least three possible reasons for the lack of SMA activity. One possibility is that it is difficult to record the SMA activity by MEG, because the SMA has bilateral functions, that is, the dipoles in the SMA of each hemisphere are

very close to each other and their directions are opposite each other. Therefore, the magnetic fields generated in the SMA of each hemisphere may cancel each other out. Second, the SMA activity may not be dominant during simple vocalization. Several authors reported that the SMA contribution was less during simple movement than during complex movement (Simonetta et al., 1991; Rao et al., 1993; Tanji and Mushiake, 1996). Therefore, the SMA activity might be minimal during the simple task, which was the vocalization of a fixed vowel repeatedly, in the present study. The third possibility is based on the analysis period and sampling condition of the present study, 150 ms before to 150 ms after the vocalization onset with a bandpass filter of 0.1-100 Hz. Although, the SMA is thought to play an important role in preparation for movement, the SMA is the main contributor for generating BP, which may be recorded with much broad low-cut filter, e.g. DC-0.001 Hz, rather than activities just before the movement.

Some studies of MEG and EEG have indicated the activation of Broca's area during several syllables or words vocalization (Grözinger et al., 1979, 1980; Sasaki et al., 1995). The left insula and frontal operculum including Broca's area contribute to the motor planning of speech production and inner speech, which are assumed to occur prior to the vocalization onset (Margolin, 1991; Dronkers, 1996; Kuriki et al., 1999). Kuriki et al. (1999) reported that the activity prior to the vocalization onset in the left frontal area might decrease by a repetition of a fixed word. Since the present study used a simple vocalization task of a fixed

vowel rather than a meaningful syllable or word, the cortical activities were bilaterally identified without any significant activities of Broca's area.

Additional generators were not found to show appropriate waveforms and improvement of the RV. Thus, the first six sources were the major generators contributing to the VRCP during the period analyzed, although activity in other areas such as the SMA could not definitely be excluded.

As for EEG (VRCP) recording, previous studies reported that dominance of VRCP is influenced by various factors of speech related processes. Deecke et al. (1986) reported a negative shift of about 100ms prior to the vocalization of a semantic word starting with "p" followed by vowel, and this negativity was dominant at the left hemisphere. On the contrary, when subjects were asked to utter only a fixed word "pool" repeatedly, the negative shift prior to the vocalization was not dominant in the left hemisphere (Wholert, 1993). Wholert (1993) also reported that the brain activities related to either simple word vocalization or non-speech oral movement (i.e., lips pressed together and lips rounded) were symmetrical without any dominant activation in the left hemisphere. In the present study, there was no definite asymmetry of activity, probably due to a lack of specific activity of a dominant hemisphere during simple vocalization under natural condition.

6.2. The masking effect during a simple vocalization

The VRCP analyzed using BESA showed that the bilateral

motor cortices were activated before the vocalization onset. Above mentioning, there is no evidence of activity in SMA during vocalization under the control condition. However, there is a possibility that the masking condition affected the preparatory motor aspects, such as activities of pre-motor or SMA, for the vocalization due to the lack of auditory feedback. Voice fundamental frequency (F0) appears to be controlled by a process in which encoding a signal of desired F0 is used to actuate the respiratory system and larynx (Hain et al., 2000). Recent studies using the pitch-shifting technique have demonstrated that subjects change their voice F0 in response to a change in voice pitch feedback, the audio-vocal reflex (Burnett et al., 1998; Hain et al., 2000). They suggested that the preceding auditory feedback affected the motor activity for the vocalization followed. Furthermore, Wohlert (1993) reported the readiness potential (RP) before silent whisper, and the potentials showed smaller amplitude than RP before vocalization, probably due to the difference of the muscle involved between silent whisper and vocalization. In the present study, subjects were instructed and trained not to change the motor acts of vocalization under the masking condition. In the subtracted VRCF and VRCP waveforms before the vocalization onset, there was no consistent response. Since the VRCF and VRCP preceding the vocalization onset in the masking condition were very similar to that of the control, it was considered that the effects of the auditory feedback overlapped with the motor acts could be minimized at least during simple vowel vocalization, whether

there was masking of the voice or not.

6.3. The response to one's own voice during a vocalization

After the vocalization onset, VRFCF and VRCP were remarkably changed in the masking condition. The subtracted VRFCF and VRCP waveforms, which were obtained by subtraction of the waveforms of the masking condition from that of the control condition, showed a component with a peak about 85 ms after the vocalization onset. Previous papers on MEG studies (AEF) and EEG studies (Auditory evoked potential: AEP) have reported that a peak about 100 ms was elicited by speech sound stimulation (Lawson and Gaillard, 1981; Eulitz et al., 1995; Kuriki et al., 1995; Näätänen et al., 1997; Gross et al., 1998; Gunji et al., 1998a; Hosokawa et al., 1998, 1999; Tiitinen et al., 1999; Gootjes et al., 1999; Nummien and Curio, 1999; Shtyrov et al., 1999; Curio et al., 2000). In the present study, the trigger point was accurately determined, but the minimal difference of 10-20 ms could not be avoided. Therefore, I considered that the response after the vocalization onset in the subtracted waveform mainly included the auditory response which was basically similar to the AEF and AEP components for the speech sound stimuli with approximately 100 ms latency (M100/N1).

The ECD calculated from the subtracted VRFCF waveform was located in the auditory cortex, the inferior bank of the Sylvian fissure. Analysis using BESA for the VRFCF waveform in the control condition, which was a multiple source model, showed the activities of the laryngeal and truncal regions in the

sensorimotor cortices as the motor components of VRCF, and the activity of auditory cortices after the vocalization onset. The component obtained from the subtracted waveform was considered to correspond to the auditory response elicited by BESA.

The M100 component of the AEF at 94.3 ± 18.4 (mean \pm SD) ms was estimated in the auditory cortices. The generator location of the M100 component following speech sound stimuli was similar to those in other studies (Eulitz et al., 1995; Kuriki et al., 1995; Näätänen et al., 1997; Tiitinen et al., 1999; Gootjes et al., 1999; Curio et al., 2000). Some investigators have reported left hemisphere dominance of the response for hearing speech sounds (Nummien and Curio, 1999; Shtyrov et al., 1999). However, in the present study, there was no consistent difference in the response between hemispheres. The auditory response following a simple vowel might be less asymmetric.

The AEF task produced a response at around +60 ms and +95 ms while the vocalization task only produced a response around +85 ms. To focus attention on auditory responses of one's own voice, we subtracted VRCF waveforms in the masking condition from the control condition. However, it might be possible to reduce the auditory response at around +60 ms, which overlapped with responses related to motor activities just before and after the onset of vocalization.

The ECD of the 1M in the subtracted VRCF waveform was located in a similar area of the auditory cortices to that of the M100 in the AEF following vowel sound stimuli in each hemisphere. Therefore, I considered that the component of

subtracted waveforms corresponded to the auditory activity as the response to subjects' own voice. There was a significant difference in the y value, that is, the 1M was estimated more medially than the M100. This result indicated that the receptive area for subject's own voice during self-vocalization might be different from that to recorded one's own voice, although the difference was small, approximately 1 cm. If this difference has some meaning, the frequency difference between the raw one's own voice and the one's own voice recorded using a voice-recorder should be accounted. The ECD to the high frequency tone, i.e. 4000 Hz, was estimated to lie in a significantly deeper area than that of the low frequency tone, i.e. 250 Hz (Romani and Williamson, 1982; Naka et al., 1999). However, the difference of the y value between 250 Hz and 4000 Hz was approximately 0.5-0.6 cm (Naka et al., 1999), being smaller than the difference found in the present study, and, of course, the difference in frequency between the real and recorded own voice was not considered to be large.

In regard to the hemispheric difference, Curio et al. (2000) reported that the peak of 1M component, which was evoked about 100 ms after the vocalization onset, in the left hemisphere was delayed. However, in the present study, there was no significant difference of peak latency for 1M obtained from the subtracted VRCF waveform between the left and right hemispheres. The peak latency of 1M tended to be shorter than that of M100, although it was not significant. In addition, peak latency for N1 obtained from the subtracted VRCP waveform did

not show significant difference between hemispheres, either. We speculated that since we used a different technique from that of Curio et al. (2000), who analyzed VRCF by specific bandpass filter, this might account for the absence of interhemispheric differences found in the present study. The strategy of the present study minimized the effects of the motor component and identified the auditory component, which were contained in VRCF and VRCP, by subtracting the masking condition from the control condition. Furthermore, the contribution of the superior temporal gyrus activity was reported to the auditory feedback process (Hirano et al., 1997), and another study suggested left hemispheric dominance for such an activity (Wildgruber et al. 1996). If the left hemisphere such as the left superior temporal gyrus was activated during the preparatory period of vocalization, the activity could not be detected in the present study, because all components relating to motor acts for vocalization were subtracted, while Curio et al. (2000) analyzed all superimposed activities of VRCF. The difference in the dipole location of IM may have been due to the specific activity of the auditory cortex to the subject's own voice.

One of the most remarkable findings of EEG (VRCP) recordings was that the N1-P1 potential of the subtracted VRCP waveform was symmetrically recorded in both hemispheres, except for the T3 and T4 electrodes. Two reasons are given for why P1 was unclear at the temporal electrodes. The first possibility is that N1 is the component specific to one's own voice, while P1 is a more generalized component such as the "vertex

potential” which is restricted at the vertex. The second possibility is a low signal-to-noise ratio at the temporal electrodes. As in a previous report (Grözinger et al., 1980), VRCP components before vocalization are small in temporal areas, and EEG recordings contained relatively more artifacts than those from other leads.

6.4. Conclusions

VRCP and VRCP under natural condition and condition with complete masking of subject’s own voice were simultaneously recorded, and cortical activities just before and after the onset of a simple vowel vocalization were analyzed. Using BESA, the sources, which were activated prior to the vocalization onset, were successfully identified in the laryngeal and truncal motor areas, and other sources in the auditory cortex were activated just after the vocalization onset. All sources were found symmetrically in the bilateral hemispheres. These activities are speculated to response for the descending motor volley and the central auditory processes during vocalization. Then, the VRCP and VRCP before the vocalization onset were similarly identical in both control and masking conditions. Therefore, I considered that the masking effect was minimal on VRCP and VRCP for a simple vocalization before its onset, while the effect was crucial to the auditory feedback process. An adequate masking condition could separate and identify the response to one’s own voice from the cortical activities relating to speech production. Both activated areas for a simple one’s own voice

and for an external simple sound might be in a similar area in the auditory cortex, although the ECD source for one's own voice was located slightly medially to that for an external voice.

7. Acknowledgements

I would like to express my deep gratitude to Professor Ryusuke Kakigi. Furthermore, Dr. Minoru Hoshiyama and the collaborators in our department are kindly acknowledged for support. I thank Mr. Osamu Nagata and Mr. Yasuyuki Takeshima for the maintenance of the equipments.

This study was supported by the Daiko Foundation.

8. References

- Brooker, B.H. and Donald, M.W. The search for scalp-recordable speech potentials. In: H. H. Kornhuber and L. Deecke, editors. *Motivation, Motor and Sensory Processes of the Brain; Electrical Potentials, Behavior, and Clinical Use*. Amsterdam: Elsevier, 1980;782-789.
- Burnett, T.A., Freedland, M.B., Larson, C.R. and Hain, T.C. Voice F0 responses to manipulations in the pitch feedback. *J Acoust Soc Am*, 1998;103:3153-3161.
- Cheyne, D. and Weinberg, H. Neuromagnetic fields accompanying unilateral finger movements: pre-movement and movement-evoked fields. *Exp Brain Res*, 1989;78:604-612.
- Curio, G., Neuloh, G., Numminen, J., Jousmaki, V. and Hari, R. Speaking modifies voice-evoked activity in the human auditory cortex. *Hum Brain Mapp*, 2000;9:183-191.
- Deecke, L., Grözinger, B. and Kornhuber, H.H. Voluntary finger movement in man: cerebral potentials and theory. *Biol Cybern*, 1976;23:99-119.
- Deecke, L., Engel, M., Lang, W., and Kornhuber, H.H. Bereitschaftspotential preceding speech after holding breath. *Exp Brain Res*, 1986;65:219-223.

- Dronkers, N.F. A new brain region for coordinating speech articulation. *Nature*, 1996;384:159-161.
- Eulitz, C., Diesch, E., Pantev, C., Hampson, S. and Elbert, T. Magnetic and electric brain activity evoked by the processing of tone and vowel stimuli. *J Neurosci*, 1995;15:2748-2755.
- Gootjes, L., Raij, T., Salmelin, R. and Hari, R. Left-hemisphere dominance for processing of vowels: a whole-scalp neuromagnetic study. *NeuroReport*, 1999;10:2987-2991.
- Grabow, J. D. and Elliot, F. W. The electrophysiologic assessment of hemispheric asymmetries during speech. *Speech Hear Res*, 1974;17:64-72.
- Gross, J., Ioannides, A.A., Dammers, J., Maess, B., Friederici, A.D. and Gärtner, H.W.M. Magnetic field tomography analysis of continuous speech. *Brain Topogr*, 1998;10:273-281.
- Grözinger, B., Kornhuber, H.H. and Kriebel, J. Participation of medial cortex in speech: Evidence from cerebral potentials preceding speech production in man. In: Creutzfeldt, O., Scheich, H. and Schreiner, C., editors. *Hearing Mechanisms and Speech*. *Exp Brain Res, Suppl. 2*. New

York: Springer-Verlag, 1979;189-192.

Grözing, B., Kornhuber, H.H., Kriebel, J., Szirtes, J., and Westphal, K.T.P. The Bereitschaftspotential preceding the act of speaking. Also an analysis of artifacts. In: H. H. Kornhuber and L. Deecke, editors. *Motivation, Motor and Sensory Processes of the Brain; Electrical Potentials, Behavior, and Clinical Use*. Amsterdam: Elsevier, 1980;798-804.

Gunji, A., Hosokawa, M., Shinoda, H. and Ozaki, H. Topographic change of ERP evoked by CV sound with various vowel duration. In: Koga, Y. and Hirata, K., editors. *Brain Topography Today*. Amsterdam: Elsevier, 1998a;233-236.

Gunji, A., Kasumi, A., Shinoda, H. and Ozaki, H. Topography of ERPs preceding a vocalization of syllable or word. In: Hashimoto, I. and Kakigi, R., editors. *Recent Advances in Human Neurophysiology*. Amsterdam: Elsevier, 1998b;614-618.

Hain, T.C., Burnett, T.A., Kiran, S., Larson, C.R., Singh, S. and Kenny, M.K. Instructing subjects to make a voluntary response reveals the presence of two components to the audio-vocal reflex. *Exp Brain Res*, 2000;130:133-141.

Higgins, M.B., Netsell, R. and Shulte, L. Vowel-related

differences in laryngeal articulatory and phonatory function. *J Speech Lang Hear Res*, 1998;41:712-724.

Hirano, S., Kojima, H., Naito, Y., Honjo, I., Kamoto, Y., Okazawa, H., Ishizu, K., Yonekura, Y., Nagahama, Y., Fukuyama, H. and Konishi, J. Cortical processing mechanism for vocalization with auditory verbal feedback. *NeuroReport*, 1997;8:2379-2382.

Hoshiyama, M., Kakigi, R., Berg, P., Koyama, S., Kitamura, Y., Shimojo, M., Watanabe, S. and Nakamura, A. Identification of motor and sensory brain activities during unilateral finger movement: spatiotemporal source analysis of movement-associated magnetic fields. *Exp Brain Res*, 1997;115:6-14.

Hosokawa, M., Gunji, A., Shinoda, H. and Ozaki, H. Topographic change of ERP evoked by the human CV syllable with vowel of varied duration. In: Hashimoto, I. and Kakigi, R., editors. *Recent Advances in Human Neurophysiology*. Amsterdam: Elsevier, 1998;619-623.

Hosokawa, M., Gunji, A., Shinoda, H. and Ozaki, H. Change of categorical perception and ERP topography evoked by single human syllables with various length of vowel. *Jpn J EEG EMG*, 1999;26:241-249.

- Hough, M. S. and Klich, R. Lip EMG activity during vowel production in apraxia of speech: phrase context and word length effects. *J Speech Lang Hear Res*, 1998;41:786-801.
- Itomi, K., Kakigi, R., Maeda, K. and Hoshiyama, M. Dermatome versus homunculus; detailed topography of the primary somatosensory cortex following trunk stimulation. *Clin Neurophysiol*, 2000; 111:405-12.
- Kakigi, R., Hoshiyama, M., Shimojo, M. Naka, D., Yamasaki, H., Watanabe, S., Xiang, J., Maeda, K., Lam, K., Itomi, K. and Nakamura, A. The somatosensory evoked magnetic fields. *Prog Neurobiol*, 2000;61:495-523.
- Kikuchi, Y. and Kita, M. Vocalization related potential and its constituent components. 15th World Congress of Neurology, Vancouver, Canada, 1993.
- Kristeva, R., Cheyne, D. and Deecke, L. Neuromagnetic fields accompanying unilateral and bilateral voluntary movements: topography and analysis of cortical sources. *Electroenceph clin Neurophysiol*, 1991;81:284-298.
- Kuriki, S., Okita, Y. and Hirata, Y. Source analysis of magnetic field responses from the human auditory cortex elicited by short speech sounds. *Exp Brain Res*, 1995;104:144-152.

- Kuriki, S., Mori, T. and Hirata, Y. Motor planning center for speech articulation in the normal human brain. *NeuroReport*, 1999;10:765-769.
- Lawson, E.A. and Gaillard, A.W.K. Evoked potentials to consonant-vowel syllables. *Acta Psychologica*, 1981;49:17-25.
- Lehman, D. and Skrandies, W. Segmentation of EEG potential fields. *Electroencephalogr Clin Neurophysiol*, 1986;38:27-32.
- Margolin D.I. Cognitive neuropsychology. Resolving enigmas about Wernicke's aphasia and other higher cortical disorders. *Arch Neurol*, 1991;48:751-765.
- McAdam, D.W. and Whitaker, H.A. Electroencephalographic localization in the normal human brain. *Science*, 1971;172:499-502.
- McGuire, P.K., Silbersweig, D.A. and Frith, C.D. Functional neuroanatomy of verbal self-monitoring. *Brain*, 1996;119:907-919.
- Murphy, K., Corfield, D.R., Guz, A., Fink, G.R., Wise, R.J.S., Harrison, J. and Adams, L. Cerebral areas associated with motor control of speech in human. *J Appl Physiol*,

1997;83:1438-1447.

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M.,
Huotilainen, M., Iivonen, A., Vainio, M., Paavo, A.,
Ilmoniemi, R.D., Luul, A., Allik, J., Sinkkonen, J. and Alho,
K. Language-specific phoneme representations revealed by
electric and magnetic brain responses. *Nature*,
1997;385:432-434.

Nagamine, T., Toro, T., Balish, M., Deuschl, G., Wang, B., Sato,
S., Shibasaki, H. and Hallett, M. Cortical magnetic and
electric fields associated with voluntary finger movements.
Brain Topogr, 1994;6:175-183.

Naka, D., Kakigi, R., Hoshiyama, M., Yamasaki, T., Okusa, T.
and Koyama, S. Structure of the auditory evoked magnetic
field during sleep. *Neuroscience*, 1999;93:573-583.

Nakamura, A., Yamada, T., Goto, A., Kato, T., Ito, K., Abe, Y.,
Kachi, T. and Kakigi, R. Somatosensory Homunculus as
Drawn by MEG. *Neuroimage*, 1998;7:377-386.

Numminen, J. and Curio, G. Differential effects of overt, covert
and replayed speech on vowel-evoked responses of the
human auditory cortex. *Neurosci Lett*, 1999;272:29-32.

Penfield, W. and Boldrey, E. Somatic motor and sensory

representation in the cerebral cortex of man as studied by electrical stimulation. *Brain*, 1937;60:389-443.

Perry, D.W., Zatorre, R.J., Petrides, M., Alivisatos, B., Meyer, E. and Evans, A.C. Localization of cerebral activity during simple singing. *NeuroReport*, 1999;10:3453-3458.

Price, C.J., Wise, R.J.S., Warburton, E.A., Moore, C.J., Howard, D., Patterson, K., Frackowiak, R.S.J. and Friston, K.J. Hearing and saying the functional neuro-anatomy of auditory word processing. *Brain*, 1996;119:919-931.

Rao, S.M., Binder, J.R., Bandettini, P.A., Hammeke, T.A., Yetkin, F.Z., Jesmanowicz, A., Lisk, L.M. Morris, G.L., Mueller W.M. and Estkowski, L.F. Functional magnetic resonance imaging of complex human movements. *Neurology*, 1993;43:2311-2318.

Romani, G.L. and Williamson, S.J. Tonotopic organization of the human auditory cortex. *Science*, 1982;216:1339-1340.

Salmelin, R., Hari, R., Lounasmaa, O.V. and Sams, M. Dynamics of brain activation during picture naming. *Nature*, 1994;386:463-465.

Sarvas, J. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Physiol Med Biol*,

1987;32:11-12.

Sasaki, K., Kyuhou, S., Nambu, A., Matsuzaki, R., Tsujimoto, T. and Gemba, H. Motor speech centers in the frontal cortex. *Neurosci Res*, 1995;22:245-248.

Scherg, M. and Berg, P. BESA: brain electric source analysis. Version 2.1 handbook, MEGIS, Munich (unpublished). 1995.

Shibasaki, H., Barrett, G., Halliday, E. and Halliday, A.M. Components of the movement-related cortical potential and their scalp topography. *Electroencephlogr Clin Neurophysiol*, 1980;49:213-226.

Shtyrov, Y., Kujala, T., Ilmoniemi, R.J. and Näätänen, R. Noise affects speech-signal processing differently in the cerebral hemispheres. *NeuroReport*, 1999;10:2189-2192.

Simonetta, M., Clanet, M. and Rascol, O. Bereitschaftspotential in a simple movement or in a motor sequence starting with the same simple movement. *Electroencephalogr Clin Neurophysiol*, 1991;81:129-134.

Szirtes, J. and Vaughan, J.R. Characteristics of cranial and facial potentials associated with speech production. *Electroencephalogr Clin Neurophysiol*, 1977;43:386-396.

- Taniguchi, M., Yoshimine, T., Cheyne, D., Kato, A., Kihara, T., Ninomiya, H., Hirata, M., Hirabuki, N., Nakamura, H. and Hayakawa, T. Neuromagnetic fields preceding unilateral movements in dextrals and sinistrals. *NeuroReport*, 1998;9:1497-1502.
- Tanji, J. and Mushiake, H. Comparison of neuronal activity in the supplementary motor cortex. *Brain Res Cogn Brain Res*, 1996;3:143-150.
- Tiitinen, H., Sivonen, P., Alku, P., Virtanen, J. and Näätänen, R. Electromagnetic recording reveal latency differences in speech and tone processing in humans. *Brain Res Cogn Brain Res*, 1999;25:355-363.
- Wildgruber, D., Ackermann, H., Klose, U., Kardatzki, B. and Grodd, W. Functional lateralization of speech production at primary motor cortex: a fMRI study. *NeuroReport*, 1996;7:2791-2795.
- Wohlert, A.B. and Larson, M.A. Cerebral averaged potentials preceding oral movement. *Speech Hear Res*, 1991;34:1387-1396.
- Wohlert, A.B. Event-related brain potentials preceding speech and nonspeech oral movements of varying complexity. *Speech Hear Res*, 1993;36:987-905.

9. Publications

Gunji, A., Hoshiyama, M. and Kakigi, R. Auditory response following vocalization: A magnetoencephalographic study. *Clinical Neurophysiology*, 2001;112:514-520.

Hoshiyama, M., Gunji, A. and Kakigi, R. Hearing the sound of silence: a magnetoencephalographic study. *NeuroReport*, in press.

Yabe, H., Koyama, S., Kakigi, R., Gunji, A., Tervaniemi, M., Sato, Y. and Kaneko, S. Automatic discriminative sensitivity inside temporal window of sensory memory as a function of time. *Brain Research Cognitive Brain Research*, in press.

Koyama, S., Gunji, A., Yabe, H., Yamada, R.A., Oiwa, S., Kubo, R. and Kakigi, R. The masking effect in foreign speech sounds perception revealed by neuromagnetic responses. *NeuroReport*, 2000;11:3765-9.

Koyama, S., Gunji, A., Yabe, H., Oiwa, S., Yamada, R.A., Kakigi, R. and Näätänen, R. Hemispheric lateralization in an analysis of speech sounds: left hemisphere dominance replicated in Japanese subjects. *Brain Research Cognitive Brain Research*, 2000;10:119-24.

Lam, K., Kaneoke, Y., Gunji, A., Yamasaki, H., Matsumoto, E., Naito, T. and Kakigi, R. Magnetic response of human extrastriate cortex in the detection of coherent and incoherent motion. *Neuroscience*, 2000;97:1-10.

Gunji, A., Hoshiyama, M. and Kakigi, R. Identification of auditory evoked potentials of one's own voice. *Clinical Neurophysiology*, 2000;111:214-9.

Gunji, A., Kakigi, R. and Hoshiyama, M. Spatiotemporal source analysis of vocalization-associated magnetic fields. *Brain Research Cognitive Brain Research*, 2000;9:157-163.

Hosokawa, M., Gunji, A., Shinoda, H. and Ozaki, H. Change of categorical perception and ERP topography evoked by single human syllables with various length of vowel. *Jpn J EEG EMG*, 1999;26:241-249.

Gunji, A., Hosokawa, M., Shinoda, H. and Ozaki, H. Topographic change of ERP evoked by CV sound with various vowel duration. In: Koga, Y. and Hirata, K., editors. *Brain Topography Today*. Amsterdam: Elsevier, 1998;233-236.

Gunji, A., Kasumi, A., Shinoda, H. and Ozaki, H. Topography of ERPs preceding a vocalization of syllable or word. In: Hashimoto, I. and Kakigi, R., editors. *Recent Advances in Human Neurophysiology*. Amsterdam: Elsevier,

1998;614-618.

Hosokawa, M., Gunji, A., Shinoda, H. and Ozaki, H. Topographic change of ERP evoked by the human CV syllable with vowel of varied duration. In: Hashimoto, I. and Kakigi, R., editors. Recent Advances in Human Neurophysiology. Amsterdam: Elsevier, 1998;619-623.

Appendix

*** Brain Electric Source Analysis: BESA**

Extracted from BESA: brain electric source analysis. Version 2.1 handbook (Scherg and Berg. 1995)

BESA is a program for the spatio-temporal analysis of EEG, in particular event-related potential (ERP) data, and is specially arranged for own MEG system. The basic idea of BESA is the notion that only a distinct number of brain areas are active during the period or some epoch of an ERP. If we take one equivalent dipole to describe the compound activity in each active, functionally different area, and if we place each dipole at the appropriate equivalent location within an electrical head model, we can obtain a complete and unique model of the scalp activity apart from biological noise. The only requirement is that the number of recording channels is larger than the number of active sources.

In practice, because of noise, probably about twice as many channels as there are sources are required. In principle, the analysis of MEG and EEG data is very similar. As the main difference from the EEG, a different head model is used for MEG analysis, which accounts for the transparency of the skull and scalp and for the lack of a radial dipole component. BESA uses a spherical head model (Sarvas. 1987) to compute the magnetic fields outside an idealized sphere due to a point dipole source. Therefore, the multiple MEG dipole model used in BESA comprises two approximations: a) equivalent point dipole sources, b) approximation of the head or of the local curvature (of the

brain surface) under the sensor array by a sphere.

The task of BESA is therefore threefold:

1. Estimate the number of active sources.
2. Locate appropriate dipoles.
3. Depict the compound activity of each dipole, i.e. the dipole strength as a function of time (dipole source potentials).

The later task is the most important aspect of BESA, because it allows functional imaging, going beyond the mere localization of sources. This means that BESA is not only asking where the sources are located, but also how each structure functions over time.

It turns out that these time functions of local compound activity can be estimated much more robustly than the (equivalent) locations of sources within the head model.

The number of active sources can be estimated by a principal component analysis (PCA). A better estimate can be obtained by testing various hypotheses with an increasing number of dipole. BESA provides all the tools and many examples for these procedures. There are several ways of checking a solution. For example:

1. Addition of a further (test-) dipole will not affect the other dipole source potentials and the test-dipole does not attract any additional compound activity (apart from noise).
2. The solution should be consistent with anatomical and physiological knowledge. Patients with circumscribed vascular brain lesions may provide an independent means of

validating a certain source hypothesis (see Scherg and Cramon 1986).

BESA does not find solutions automatically. It is a tool which needs your anatomical knowledge to be guided among the various hypotheses. It can probably test most of the hypotheses you have formulated. BESA cannot cure data which have a poor single-to-noise-ratio, although you may get some improvements using the filtering facilities (beware of low-frequency noise or DC-shifts).

BESA uses a simple head model. This makes the computer program and therefore the testing of hypotheses relatively fast. Due to individual variation in scalp and skull thicknesses and head geometry the obtained dipole locations may show some interindividual variability in addition to the known anatomical variability. Dipole source potentials are less sensitive to location than they are to orientation: You should therefore not overinterpret location (which may be wrong up to 2 cm in the worst cases). The additional information provided by the dipole orientations can be used to narrow down the range of possible interpretations. In situations with many sources the simple head model may also lead to “apparently” better solutions, for which the residual variance (RV) is a fraction of a percent less, which may be rejected in favor of other solutions satisfying other criteria. In particular, the clear separation of the source processes in the sense of a minimum interaction between the dipole source potentials is the more important criterion to select or reject a hypothesis near the RV minimum.

Sarvas, J. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Physiol Med Biol*, 1987;32:11-12.

Scherg, M. and Berg, P. BESA: brain electric source analysis. Version 2.1 handbook, MEGIS, Munich (unpublished). 1995.

Scherg, M. and von Cramon, D. Evoked dipole source potentials of the human auditory cortex. *Electroenceph clin Neurophysiol*, 1986;65:344-360.