

## Separate neural systems for processing action- or non-action-related sounds

L. Pizzamiglio,<sup>a,b,\*</sup> T. Aprile,<sup>a,b</sup> G. Spitoni,<sup>a,b</sup> S. Pitzalis,<sup>a,b</sup> E. Bates,<sup>c,✱</sup>  
S. D'Amico,<sup>a</sup> and F. Di Russo<sup>b,d</sup>

<sup>a</sup>Department of Psychology, University of Rome "La Sapienza," Rome, Italy

<sup>b</sup>Fondazione Santa Lucia IRCCS, Rome, Italy

<sup>c</sup>Department of Cognitive Science, UCSD, La Jolla, CA, USA

<sup>d</sup>Istituto Universitario di Scienze Motorie (IUSM), Rome, Italy

Received 15 June 2004; revised 3 September 2004; accepted 17 September 2004

Available online 24 November 2004

**The finding of a multisensory representation of actions in a premotor area of the monkey brain suggests that similar multimodal action-matching mechanisms may also be present in humans. Based on the existence of an audiovisual mirror system, we investigated whether sounds referring to actions that can be performed by the perceiver underlie different processing in the human brain. We recorded multichannel ERPs in a visuoauditory version of the repetition suppression paradigm to study the time course and the locus of the semantic processing of action-related sounds. Results show that the left posterior superior temporal and premotor areas are selectively modulated by action-related sounds; in contrast, the temporal pole is bilaterally modulated by non-action-related sounds. The present data, which support the hypothesis of distinctive action sound processing, may contribute to recent theories about the evolution of human language from a mirror system precursor.**

© 2004 Elsevier Inc. All rights reserved.

*Keywords:* Neural systems; Sounds; Actions

---

### Introduction

Impaired understanding of meaningful environmental sounds has been described as an isolated deficit (Spreen et al., 1965); however, it is most often associated with language deficits in aphasic patients (Clark et al., 2000; Pinard et al., 2002; Saygin et al., 2003; Varney, 1980; Vignolo, 1982). Also, lesions producing

pure auditory agnosia have been described in the right temporal lobe (Fujii et al., 1990; Spreen et al., 1965; Taniwaki et al., 2000).

Recent neuroimaging studies describe a distributed network involved in processing spoken language and environmental sounds: The common activation for sounds and words is described in the left inferior frontal lobe and in the right and left temporal cortex (Thierry et al., 2003). When word processing was contrasted with sounds, left anterior and posterior activations were found in the superior temporal gyrus (Giraud and Price, 2001; Humphries et al., 2001). However, the contrasting of sounds and words did not show greater right hemisphere involvement (Giraud and Price, 2001; Humphries et al., 2001) when subjects were exposed to passive listening. Further, a clear activation of the right posterior superior temporal cortex was found when the task involved a semantic judgment (Thierry et al., 2003). It was concluded that the semantic system has a dual access for verbal material and for sounds involving the left and right hemispheres, respectively.

In a different domain, many studies have focused on the functional organization of brain motor areas and their role in recognizing other individuals' actions. Recently, a single cell recording study reported a discharge of neurons located in the monkey F5c premotor area when the animal performed an action and also when it only perceived the same action-related sound (Kohler et al., 2002). These audiovisual mirror neurons apparently fired when the animal ripped a piece of paper apart by itself or when it perceived the experimenter ripping the paper apart, causing the appropriate sound, and also when this sound was the only stimulus. However, the same neurons did not respond when a non-action-related sound was presented, such as a burst of white noise or a monkey call. Involvement of the prefrontal cortex in integrating the auditory domain was also found in primates (Romanski and Goldman-Rakic, 2002).

The aim of this study was to investigate the neural processing of complex sounds in humans by integrating the findings from different research domains, considering in particular the recent

---

\* Corresponding author. Dipartimento di Psicologia, Università di Roma "La Sapienza" Via dei Marsi 78, 00185 Roma, Italy.

E-mail address: luigi.pizzamiglio@uniroma1.it (L. Pizzamiglio).

✱ Deceased December 13, 2003.

Available online on ScienceDirect ([www.sciencedirect.com](http://www.sciencedirect.com)).

findings on auditory semantic processing in both the human and the monkey brain.

In the present study, we hypothesized that two distinct neural substrates exist for the identification of sound meaning. One processes the meaning of action-related sounds (i.e., hands clapping) when an audiovisual mirror system is activated together with a specific motor action program. The other distinct neural system is involved when sound identification has to rely solely on the acoustic and perceptual properties of the sound itself (i.e., water boiling) without the possibility of activating any action-related discharge.

The strategy used in the present experiment combines a semantic priming paradigm with the phenomenon of repetition suppression. A prime stimulus, either semantically or perceptually related to the target, facilitates a speeded response to a subsequent target, that is, a faster reaction time. Miller and Desimone (1991) and Desimone (1996) showed that, at the neurophysiological level, the repetition of a stimulus reduces the neural activity of stimulus-specific cells compared to the same stimulus when it is presented non-repetitively. Recently, Schnyer et al. (2002) showed that masked word priming increased the fMRI signal in a region of the cortex associated with perceptual identification of the target, while visible priming decreased activity in the same area. This demonstrates that repetition-related changes may result in different patterns of signal modulations depending on whether or not the prime is identified or rapidly learned (Dobbins et al., 2004).

The masked priming paradigm, using identical or congruent stimuli, successfully modulated electrical as well as BOLD brain responses in studies by Dehaene et al. (1998, 2001) and Naccache and Dehaene (2001), in which brain activation was investigated during the coding of numerical quantities in the parietal lobe. The same authors stressed that the priming method proved to be more sensitive than the fMRI subtraction method for identifying brain areas involved in the cognitive task they used. For all the above reasons, the use of subliminal priming paradigms to study the neural phenomenon of repetition suppression is becoming a successful strategy for measuring activity at the level of small columns of neurons in both event-related fMRI and ERP studies.

We recorded ERPs and reaction times to two categories of action and non-action-related sounds, preceded by a brief visual prime. The prime (written word) could be semantically congruent (repeated) with the following sound or incongruent (nonrepeated):

The very short presentation of the prime followed by a mask prevented the subjects from ever identifying the meaning of the written word (Fig. 1). The subject's task was to press one of two buttons, depending on whether the sound could be produced by a human being (i.e., hand clapping) or not (i.e., water boiling).

In order to generalize the potential influence of the action-related sound, we performed two separate experiments using meaningful sounds produced by motor actions of two body areas (hands and mouth) as target stimuli in contrast to various nonaction sounds.

Finally, to anatomically localize the task-related activity for action and non-action-related sounds, we combined ERP data with cortical surface reconstruction of a single subject's brain.

## Materials and methods

### Subjects

Two age- and sex-matched groups of 20 subjects (10 females in each group, mean age  $23.9 \pm 3.2$  and  $24.5 \pm 2.8$  years, respectively) participated in the study. A subset of four of these subjects (two females, mean age 26.8 years SD = 4.1) also received anatomical MRI scans. All subjects were right handed; they had normal vision and hearing and no history of psychiatric or neurological disease. Written informed consent was obtained from all participants after the procedures had been fully explained to them; all participants were paid.

### Stimuli and task

Subjects were seated in a comfortable chair 114 cm away from a display. They were asked to perform an auditory semantic categorization task by pressing a response key with the middle or the index finger of the right hand when they perceived a sound typically produced by human hand or mouth actions or a nonaction sound, respectively. The subjects were not informed that each target sound was preceded by a masked visual prime. Subjects were trained to maintain stable fixation on a dot located  $1.5^\circ$  below the visual stimulus. Half of the trials were "repeated", the prime and the target produced the same meaning (e.g., prime = bell; target = bell sound); the other half were "nonrepeated," the prime

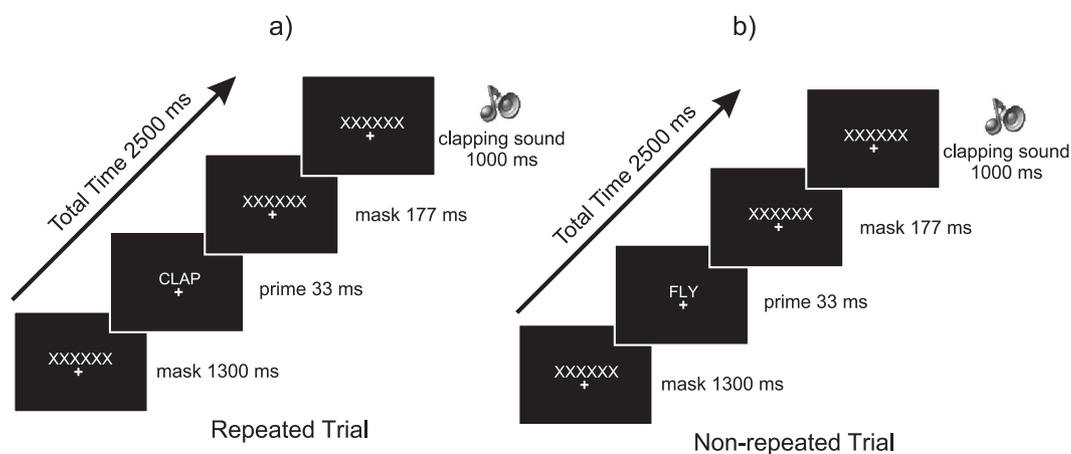


Fig. 1. Experimental design. For each target sound, the participants performed two kinds of trials: (a) semantically repeated prime-target pairs, and (b) nonrepeated prime-target pairs. In the nonrepeated trials, the prime was unrelated to the target sound.

and the target were items from different categories (prime = hand bell; target = boiling sound). Each group performed one experiment in which target sounds could be either meaningful hand- or mouth-related sounds as opposed to nonaction sounds. In each experiment, the stimulus set consisted of eight pairs of prime and target items. In the experiment with hand-related sound stimuli, the prime of each pair consisted of an Italian word spelled out, such as “SUONA” (ringing a hand bell), “BATTE” (clapping), “BOLLE” (boiling), “MOSCA” (fly), and the target consisted of four auditory stimuli correctly representing the sounds of the written words. In the experiment using mouth-related sound stimuli, the prime target pairs were “FISCHIA” (whistling), RIDE (laughing), PIOGGIA (rain), and SIRENA (ambulance siren). For one group, the semantic categorization task was hand action sounds vs. nonaction sounds. For the second group, the semantic categorization task was mouth action sounds vs. nonaction sounds. The prime stimulus consisted of a white word (size:  $3 \times 1.25^\circ$ , luminance:  $15.9 \text{ cd/m}^2$ ) flashed onto a black background (luminance:  $0.65 \text{ cd/m}^2$ ) for 33 ms, and immediately followed by a rectangular mask ( $5.75 \times 5.75^\circ$ , luminance:  $11.1 \text{ cd/m}^2$ ); the mask remained on the screen for the entire duration of the trial; the target sound was bilaterally delivered through earphones for a duration of 0.8–1.2 s, 177 ms after the prime: the intensity was 80 dB. The stimulus onset asynchrony was 2.5 s (see Fig. 1).

A semistructured interview at the end of the experimental session assessed whether the subjects could detect or name the masked priming words. Participants were basically asked to describe the items seen during the experimental recordings; two participants were able to correctly read and name the masked primes, so their data were excluded from analysis.

#### *EEG recording and data processing*

The EEG was recorded using the 10–10 system montage (see Di Russo et al., 2002). All scalp channels were referenced to the left mastoid (M1) and linked to ground by CPz. Horizontal eye movements were monitored with a bipolar recording from electrodes at the left and right outer canthi. Blinks and vertical eye movements were recorded with an electrode below the left eye, which was referenced to site Fp1. The EEG from each electrode was digitized at 250 Hz with an amplifier band-pass of 0.01–80 Hz together with a 50 Hz notch filter and was stored for off-line averaging. Electrical activity was recorded using a BrainVision 64-channel system. Artifact rejection was performed prior to signal averaging in order to discard epochs in which deviations in eye position, blinks, or amplifier blocking occurred. On average, about 13% of the trials were rejected for violation of these artifact criteria. Blinks were the most frequent cause for rejection. About 150 artifact-free trials were collected for each target sound. The EEG was segmented for each target stimulus giving epochs of 1500 ms (from –500 to 1000 ms). The baseline was calculated from 500 to 200 ms prior to the target. ERPs obtained in the two experimental conditions (repeated vs. nonrepeated) were compared sample-by-sample using a two-tailed *t* test with a consecutive 20 ms criterion, showing significant differences at  $P > 0.05$  on a temporal window from 200 to 500 ms after target onset; Bonferroni correction was employed.

Following the logic of the repetition suppression paradigm, differential ERPs were obtained by subtracting the nonrepeated pairs from the repeated ones. We focused on ERPs in the 200–500 ms temporal window after target onset, during which previous

studies found ERP signatures of semantic processing of auditory and visual stimuli (Dehaene et al., 1998, 2001; Naccache and Dehaene, 2001). Scalp voltage mapping and modeling of the dipolar sources underlying the repetition effect were obtained for three different categories of sounds, that is, non-action-related, hand action-related, and mouth action-related sounds.

#### *Source analysis*

Estimation of the dipolar sources of the difference between the nonrepeated and repeated condition was carried out using Brain Electrical Source Analysis (BESA 2000 v. 5). The BESA algorithm estimates the location and the orientation of multiple equivalent dipolar sources by calculating the scalp distribution that should be obtained for a given dipole model (forward solution) and comparing it to the original ERP distribution. Interactive changes in the location and orientation of the dipole sources led to minimization of the residual variance (RV) between the model and the observed spatiotemporal ERP distribution. In these calculations, BESA assumed a realistic approximation of the head with the radius obtained from the group average (90 mm). Single dipoles were fit sequentially over specific latency ranges to correspond with the distinctive components in the difference waveform. Dipoles accounting for the earlier portions of the waveform were left in place as additional dipoles were added. The reported dipole fits remained consistent as a function of starting position. A spatial digitizer was used to record the three-dimensional coordinates of each electrode and of the three fiducial landmarks (the left and right preauricular points and the nasion). A computer algorithm was used to calculate the best-fit sphere encompassing the array of electrode sites and to determine their spherical coordinates. The mean spherical coordinates for each site averaged across all subjects were used for the topographic mapping and source localization procedures. In addition, individual spherical coordinates were related to the corresponding digitized fiducial landmarks and to fiducial landmarks identified on the standardized finite element model (FEM) of BESA 2000. The standardized FEM was created from an average head using 24 individual MRIs in Talairach space.

#### *MR data processing*

##### *Imaging parameters*

For each subject, we acquired two high-resolution ( $1 \times 1 \times 1 \text{ mm}$ ) T1-weighted images of the whole brain, using a 3-D Magnetization Prepared Rapid Gradient Echo (MPRAGE) sequence (TR = 11.4 ms, TE = 4.4 ms, flip angle =  $10^\circ$ ,  $256 \times 256$  matrix,  $1 \times 1 \text{ mm}$  in-plane resolution, 220 contiguous 1 mm coronal slices). These structural scans, tuned to optimize the contrast between grey and white brain matter, were acquired in a single session using a head coil for full head coverage.

##### *Cortical flattening*

Processing of anatomical images was performed using FreeSurfer (Fischl et al., 1999; Dale et al., 1999; <http://surfer.nmr.mgh.harvard.edu>). For each subject, the two high-resolution anatomical images were used to generate a patch of flattened cortex containing the entire brain surface. The flattened cortex was generated using Dale and Sereno's (1993) procedure and is outlined here. First, the skull was stripped off automatically by a stiff deformable template onto the brain images; the high contrast

between grey and white matter in the reference anatomical images permitted the automated detection of the white matter outline in each slice. The slices were combined into a single three-dimensional volume, and the surface of the white matter was constructed by “shrink-wrapping” the brain. This surface was then inflated, with minimized distortion, to eliminate folding of the gyri and sulci. Finally, the inflated brain was cut along the calcarine fissure and four other standard anatomical regions to completely flatten the cortical surface. The flat patch provides a convenient way of examining activity variation across the three-dimensional surface of the cortex in a single view.

#### Talairach coordinates

Correspondences between ERP activity and underlying anatomical areas in the flat patch were assessed using Talairach and Tournoux (1988) stereotaxic coordinate system. Talairach coordinates were calculated through an automatic nonlinear stereotaxic normalization procedure (Friston et al., 1995), performed using the SPM99 software platform (Wellcome Department of Cognitive Neurology, London, UK), implemented in MATLAB (The MathWorks Inc., Natick, MA, USA). The template image was based on average data provided by the Montreal Neurological Institute (Mazziotta et al., 1995).

## Results

### Behavioral results

Reaction times to target sounds were submitted to a  $2 \times 2$  ANOVA with sound type (action-related vs. non-action-related) and semantic prime (repeated vs. nonrepeated) as repeated factors. Results showed a significant main effect of the semantic-prime factor ( $P < 0.001$  for both hand and mouth sounds) and a main effect of the sound-type factor ( $P < 0.002$  only for mouth sounds). The interaction effect between the two factors was not significant. Mean reaction time on the semantic categorization task for hand action sounds was 495 ms in the repeated and 523 ms in the nonrepeated condition; for mouth action sounds, it was 482 ms in the repeated and 512 ms in the nonrepeated condition. For nonaction sounds, the mean reaction time was 507 ms in the repeated and 532 ms in the nonrepeated condition for the hand sound experiment and 502 ms in the repeated and 527 ms in the nonrepeated condition for the mouth sound experiment. The behavioral repetition effect for all sound categories was very similar, ranging from 25 to 30 ms.

### Electrophysiological results and source localization

The general shape of the potentials was the same in the two conditions (repeated vs. nonrepeated) and for all kind of sounds; the ERPs included the typical auditory N1, P2, N2, and P3 components and peaked around 100, 175, 250, and 300 ms, respectively.

The top of Fig. 2 shows the ERPs for the nonaction sounds from representative electrodes. The first priming effect occurred at about 280 ms after the stimulus onset and lasted for about 80 ms (peak 320 ms). The repetition effect was statistically significant between 300 and 340 ms in the frontal (F1, Fz, and F2) and temporoparietal (T7, TP7, T8, and TP8) electrodes ( $p < 0.01$ ). As showed in the 3-D maps of Fig. 2, the repetition effect topography

of the non-action-related sounds revealed negative activity on the frontal areas and bilateral positive activity on the temporoparietal cortex. The bottom of Fig. 2 shows the repetition effect source localization of the nonaction sounds superimposed on a standard MRI. The differential activity from 300 to 340 ms was best fit by a symmetrical dipole pair located at the tip of the temporal pole of both hemispheres. The Talairach coordinates of these dipole sources were  $\pm 26, 12, -21$ . The dipole pair explained 95.4% of the data variance in the interval used for fitting. The source waveforms (dipole moment) of the modeled dipoles showed bilateral peak activity at 320 ms resembling the surface differential ERPs.

The top of Fig. 3 shows the ERPs for hand action sounds. This category of sounds showed a repetition effect at about 240 ms after the stimulus onset and lasted for about 80 ms (peak 280 ms). For the hand action sounds, the repetition effect was statistically significant between 260 and 310 ms in the central (C1, Cz, and C2) and left frontocentral (FC7 and T7) electrodes ( $P < 0.01$ ). As showed in the 3-D maps of Fig. 3, the repetition effect topography of these sounds was negative activity on the central areas and a left positive activity on the temporofrontal cortex. The bottom of Fig. 3 shows repetition effect source localization of the hand action sounds superimposed on a standard MRI. The differential activity from 260 to 310 ms was best fit by two single dipoles with slightly different source activity. The first dipole, located in the superior temporal sulcus (STS) of the left hemisphere (Talairach coordinates:  $-35, -12, 20$ ) had a peak source activity at 280 ms; the second dipole, located in the left premotor cortex (Talairach coordinates:  $-32, 15, 43$ ), had a peak source activity at 300 ms. The two dipoles explained 95.8% of the data variance in the interval used for fitting.

The top of Fig. 4 shows the ERPs for the mouth action sounds. The differential waveform showed that the priming effect began at about 270 ms after the stimulus onset and lasted for about 80 ms (peak 300 ms). The repetition effect was statistically significant between 280 and 320 ms in the frontocentral electrodes (FCz and FC1) and left temporoparietal electrode (P7), with a  $P < 0.01$ . As showed in the 3-D maps of Fig. 4, the repetition effect topography of the mouth-related sounds revealed negative activity on the frontocentral areas and left positive activity on the temporoparietal cortex. The bottom of Fig. 4 shows the repetition effect source localization of the mouth action sounds superimposed onto a standard MRI. The differential activity from 280 to 320 ms was best fit by two single dipoles with slightly different source activity. As for the hand sounds, the first dipole was located in the STS of the left hemisphere (Talairach coordinates:  $-38, -20, 18$ ) and had a peak source activity at 295 ms. The second dipole was located in the left premotor cortex (Talairach coordinates:  $-36, 12, 35$ ) and had a peak source activity at 305 ms. The two dipoles explained 95.9% of the data variance in the interval used for fitting.

Fig. 5 shows the dipole localization of the priming effect related to hand-, mouth- and nonaction sounds on the folded, inflated, and flattened left hemisphere on one single subject (GC), selected from the subset that underwent the MR scan. As shown in Fig. 5, there are two pairs of action-related dipoles, indicated by yellow and blue spots (for hands and mouth, respectively). The first pair of action-related dipoles is located in the left hemisphere superior to the transverse gyrus (TRAg) and in between the Sylvian fissure (indicated by dashed lines) and the insula. The second pair of action-related dipoles is located in the premotor cortex, specifically at the most superior end of the middle frontal sulcus and slightly

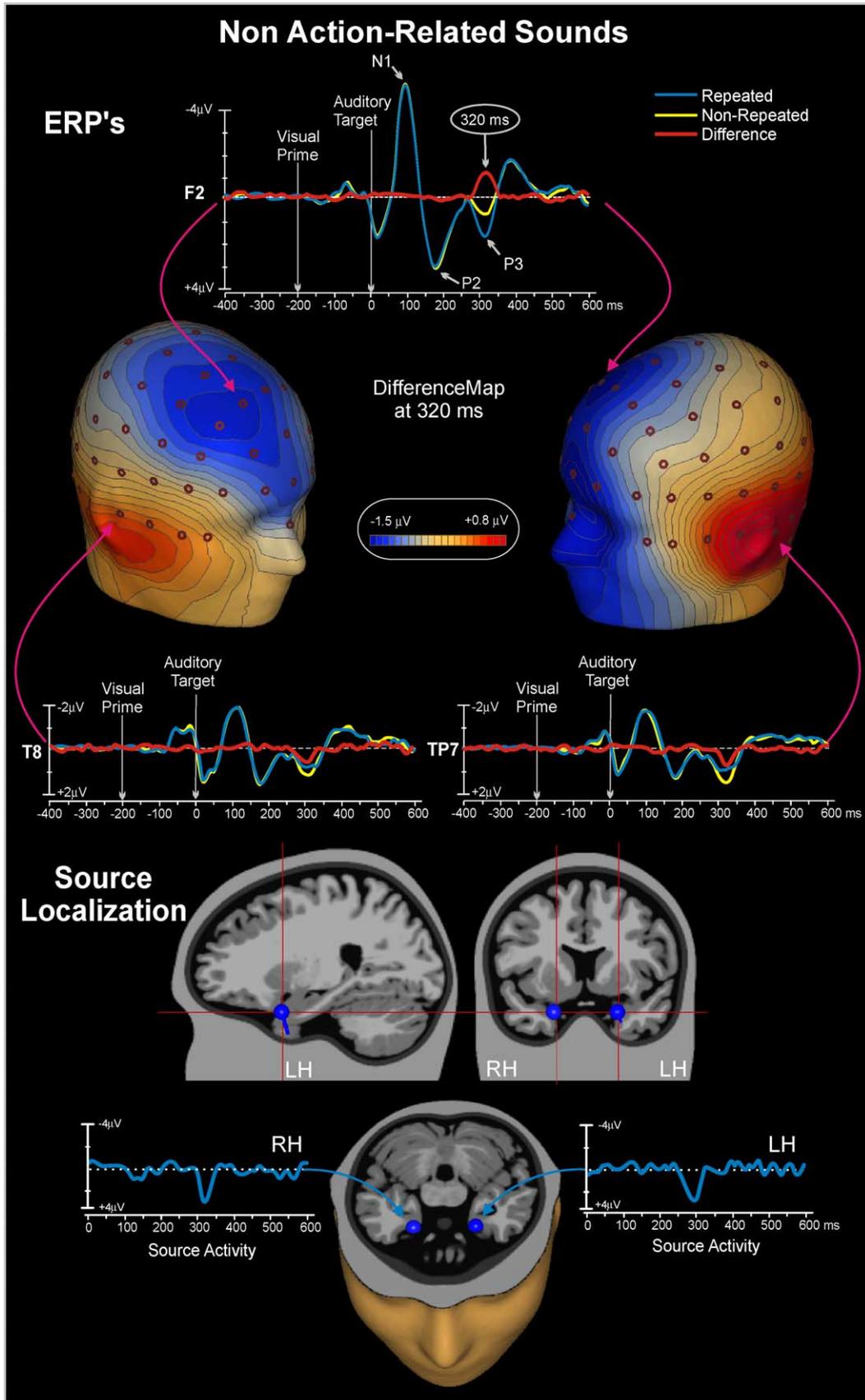


Fig. 2. (Top) Non-action-related sounds. Grand average ERPs for the auditory targets in response to non-action-related sounds and scalp topography of the repetition effect. Superimposed lines represent the ERPs in the two conditions (repeated vs. nonrepeated), and thick red lines the difference wave. (Bottom) Source localization and time course of the nonaction sounds repetition effect on standard MRI.

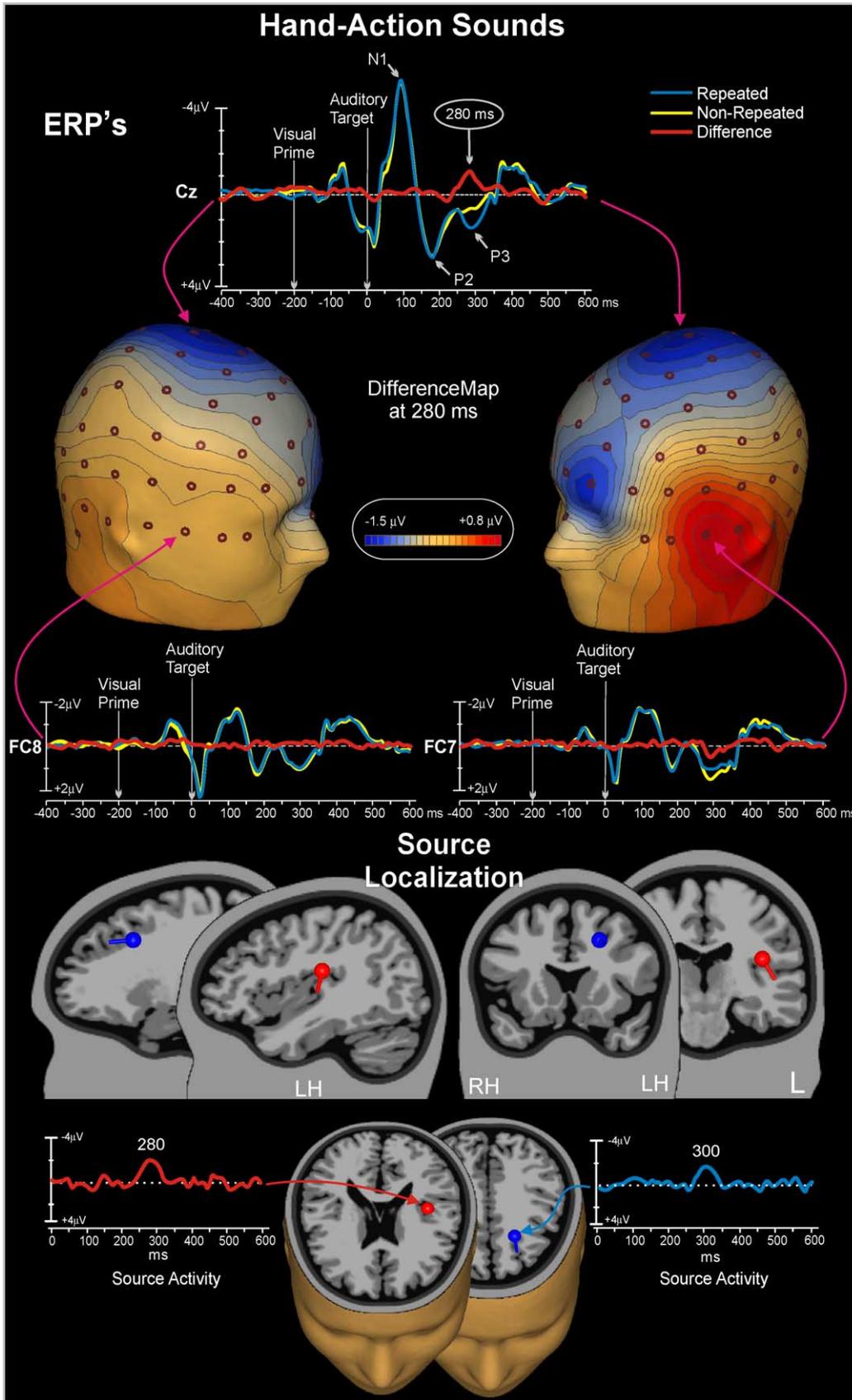


Fig. 3. Hand action sounds. (Top) Grand average ERPs for the auditory targets in response to hand action sounds and scalp topography of the repetition effect. Superimposed lines represent the ERPs in the two conditions (repeated vs. nonrepeated), and thick red lines the difference wave. (Bottom) Source localization and time course of the hand action sounds repetition effect on standard MRI.

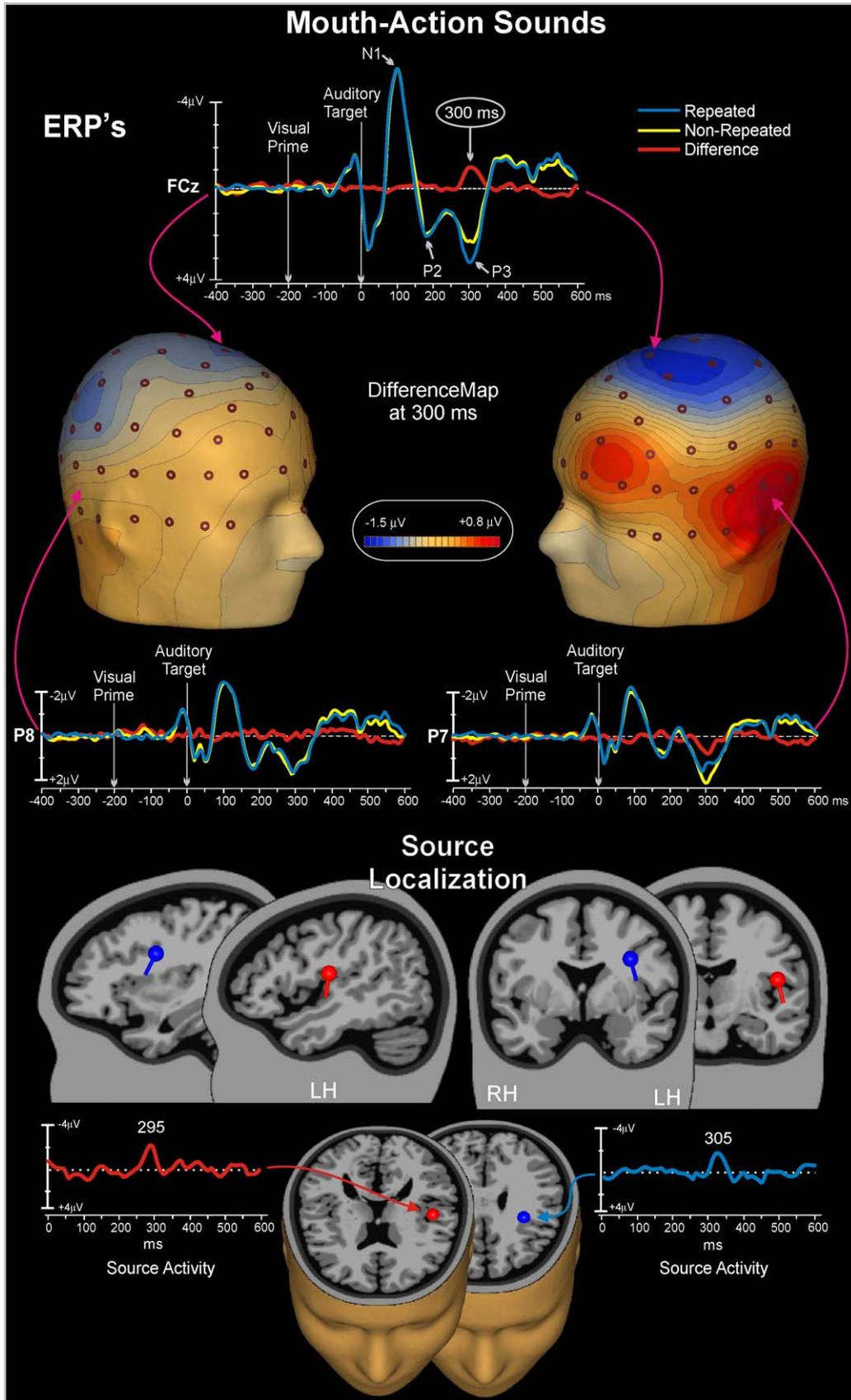


Fig. 4. Mouth action sounds. (Top) Grand average ERPs for the auditory targets in response to mouth action sounds and scalp topography of the repetition effect. Superimposed lines represent the ERPs in the two conditions (repeated vs. nonrepeated), and thick red lines the difference wave. (Bottom) Source localization and time course of the hand action sounds repetition effect.

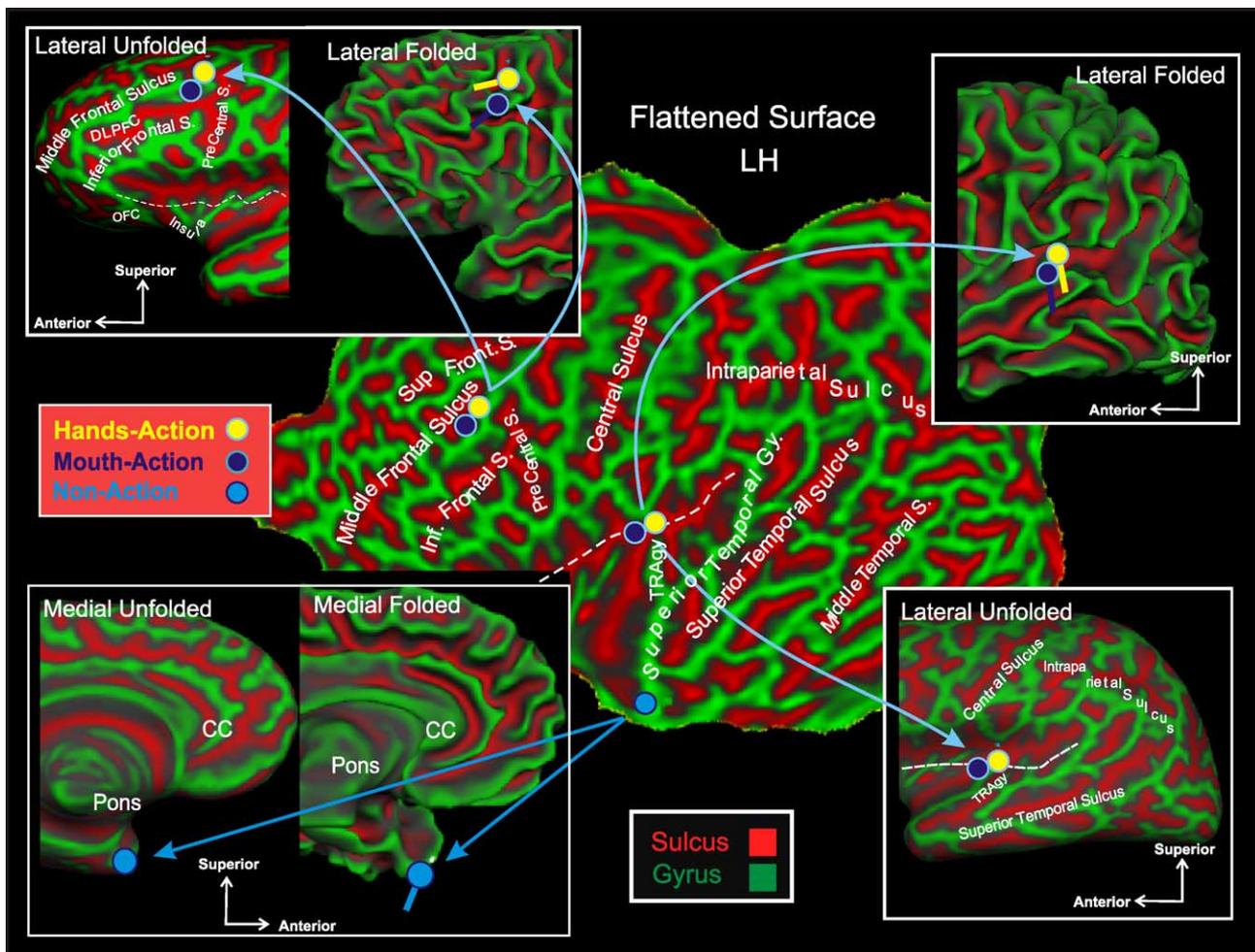


Fig. 5. Single subject analysis. Dipole localization of the priming effect related to hand, mouth, and nonaction sounds on the folded, inflated, and flattened left hemisphere of a single participant (GC). The folded and unfolded surfaces are shown in lateral and medial views (in inset closeup). Main sulci (red) and gyri (green) have text labels: Superior Temporal Sulcus, Superior Temporal Gyrus, Middle Temporal Sulcus, Transverse Gyrus (TRAGy), Central Sulcus, Intraparietal Sulcus, Pons, Corpus Callosum (CC), Superior Frontal Sulcus, Middle Frontal Sulcus, Dorsolateral Prefrontal Cortex (DLPFC), Inferior Frontal Sulcus, Precentral sulcus, Orbitofrontal Cortex (OFC), and Insula. Sulcus from Sylvian fissure to insular cortex is indicated with a dashed line.

inferiorly in the middle frontal gyrus. The non-action-related dipoles, coded by blue/cyan spots, are located bilaterally (data from right hemisphere not shown) on the most anterior tip of the temporal pole.

## Discussion

Meaningful sounds can be understood by using different cognitive processes. We hypothesized that the mirror system may have an important role in identifying the sound category that activates the representation of an action necessary to produce those sounds.

Further, environmental sounds that cannot rely on such activation may depend on a different neural circuit for analysis and identification.

The results of the present experiment clearly outlined two separate systems for understanding action-related compared with non-action-related sounds.

One system involves left hemisphere cortical structures known to be part of the mirror neuron system and will be described first.

However, the system involved in processing non-action-related sound relies on a very distinct circuit and involves bilateral activation of different cortical structures.

### *Mirror neuron system and understanding action sounds*

Dipole modeling of the repetition suppression effect showed that the left posterior superior temporal area and the left premotor cortex were selectively modulated by action-related sounds. The superior temporal area had a peak activity around 290 ms preceding the premotor cortex peak activity of about 10 ms. Several points need to be discussed in this regard.

The first one concerns the activation of two cortical areas in which mirror neurons have been described in humans: the left premotor cortex (Buccino et al., 2001, 2004; Iacoboni et al., 1999) and the STS (Dubeau et al., 2001; Iacoboni et al., 2001). These results match previous animal data on the STS responses in monkeys when they observe biologically relevant visual stimuli (Baylis et al., 1987; Bruce et al., 1981; Mistlin and Perret, 1990). Anatomical studies show reciprocal connections with the parietal area of the mirror system (Seltzer and Pandya, 1994) and, mediated

by the posterior parietal cortex (PPC), with area F5 of the macaque monkey (Rizzolatti et al., 1998).

Second, the results of our experiment show that the STS codes for auditory as well as for visual representations of significant motor action. Our findings constitute the first evidence in support of STS involvement in the human mirror system due to its selective activation in response to a heard action. This result further supports the idea of the multisensory monitoring purpose of STS within the functional system subserving the understanding of sounds and their related actions.

A third comment has to do with the different functional properties of the frontal, parietal, and temporal components of the mirror system. The frontal premotor area is supposed to code the goal of the action to be imitated (Buccino et al., 2001, 2004; Iacoboni, in press), while the STS should provide a visual description of an observed action to be imitated by the parietal mirror neurons. Furthermore, the STS should receive an efferent copy of the command originating in the frontal cortex and preparing the somatosensory information to be used by the parietal cortex (Iacoboni et al., 1999, 2001).

The inferior parietal lobule, in turn, provides the pragmatic aspect of how to perform the observed action. Recently, Buccino et al. (2004) suggested that the activation of the inferior parietal lobule was related to both the activation of previously seen neuron coding motor acts and the involvement of this area in higher order visual functions. The above speculative interpretations, based on fMRI studies in humans, assign a modulatory role to the STS in the interaction between the frontal premotor cortex, where the action originates, and the parietal cortex, where additional somatosensory information is necessary to specify the motor properties of the action. In the present study, the use of ERPs allowed for a precise description of the timing of these interactions: in both the hand and mouth action sound experiments, the STS activity preceded the frontal responses by an interval of about 10 ms. Therefore, STS, which is the first structure to activate the motor program, is likely to identify the meaning of a sound or of a visual stimulus to produce the first general description of the stimulus-related action and to transfer this representation to the premotor cortex. It should be noted that this coding was previously described in the inferior temporal structures (Jeannerod et al., 1995; Milner and Goodale, 1995), a considerably different finding from the present one.

One might speculate that such activation is linguistically mediated, since the prime in our study is a written word. This interpretation is unlikely since the linguistically congruent prime is present also for non-action-related stimuli, without any corresponding activity in the STS. Furthermore, the absence of any ERP activity in the parietal cortex indicates that the meaning activated in STS does not imply any imaginal representation of the object or any detailed specification of the pragmatic aspect of the intended action (Buccino et al., 2001).

Another point concerns the somatotopic organization of the dipoles for the hand- and the mouth-related sounds. In accordance with Dubeau et al. (2001), who found topography for hand fingers and mouth in the STS during processing of biological motion; the present source for hand action sounds in the STS was more posterior, lateral, and ventral than the one found for the mouth action sounds. Regarding the activity localized in the premotor cortex, the source subtending the mouth sounds was more lateral, ventral, and a little more anterior than the source of hand sounds in accordance with the somatotopic organization of the premotor cortex (Hauk et al., 2004; Iacoboni et al., 2001).

A final comment as to do with the nature of the priming affect that we described. Given the semantic priming paradigm used, our results show a bias toward a semantic stage of sound processing. Future studies using different paradigms could clarify the role of premotor and/or STS cortices also during different neural stages of action sounds processing.

#### *Bilateral temporal pole activation for non-action-related sounds*

The non-action-related sounds activated the tip of the temporal pole bilaterally, in agreement with previous studies (Thierry et al., 2003) in which the fMRI activity for recognition of the sound meaning was found to be bilateral. In this study, no distinction was made between classes of sounds at variance with the present paradigm.

It can be suggested that the temporal poles identify the meaning of sounds via an integration of multisensory perceptual features of the stimuli (Zatorre and Belin, 2001). Even considering the less precise localization of the ERP technique, the distance between the anterior and posterior dipoles observed for the two categories of sounds is far too great to be explained by lower spatial resolution of the source localization technique (if compared to other techniques such as fMRI) used in the present study. The segregation of the systems that process the two classes of sounds is further supported by a second strong anatomical finding, that is, environmental sounds, which are nonreproducible by the perceiver's action, require bilateral cooperation of the temporal poles, while the activity of the action-related sounds is lateralized in the frontal and temporal components of the mirror system.

Given the general agreement about the homology between the monkey ventral premotor cortex (F5) and Broca's area in the human brain (for a review, see Rizzolatti et al., 1998), the left laterality for action-related sounds may hold up the intriguing evolutionary hypothesis that our language communication system developed from a mirror system precursor (Rizzolatti and Arbib, 1998).

#### **Acknowledgments**

This research was supported by the Italian Ministero della Salute and MIUR grants and by NIH grant.

#### **References**

- Baylis, G.C., Rolls, E.T., Leonard, C.M., 1987. Functional subdivisions of the temporal lobe neocortex. *J. Neurosci.* 7, 330–342.
- Bruce, C., Desimone, R., Gross, C.G., 1981. Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *J. Neurophysiol.* 46, 369–384.
- Buccino, G., Binkofski, F., Fink, G.R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R.J., Zilles, K., Rizzolatti, G., Freund, H.J., 2001. Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Eur. J. Neurosci.* 13, 400–404.
- Buccino, G., Vogt, S., Ritzl, A., Fink, G.R., Zilles, K., Freund, H.J., Rizzolatti, G., 2004. Neural circuits underlying imitation learning of hand actions; an event-related fMRI study. *Neuron* 42, 323–334.
- Clark, S., Bellmann, A., Meuli, R.A., Assal, G., Steck, A.J., 2000. Auditory agnosia and auditory spatial deficits following left hemispheric lesions: evidence for distinct processing pathways. *Neuropsychologia* 38, 797–807.
- Dale, A.M., Sereno, M.I., 1993. Improved localization of cortical activity

- by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach. *J. Cogn. Neurosci.* 5, 162–176.
- Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis: I. Segmentation and surface reconstruction. *NeuroImage* 9, 179–194.
- Dehaene, S., Naccache, L., Le Clec'h, G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., van de Moortele, P.F., Le Bihan, D., 1998. Imaging unconscious semantic priming. *Nature* 395, 597–600.
- Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D., Mangin, J.F., Poline, G.B., Riviere, D., 2001. Cerebral mechanisms of word masking and unconscious repetition priming. *Nat. Neurosci.* 4, 752–758.
- Desimone, R., 1996. Neural mechanisms for visual memory and their role in attention. *Proc. Natl. Acad. Sci. U. S. A.* 93, 13494–13499.
- Di Russo, F., Martínez, A., Sereno, M.I., Pitzalis, S., Hillyard, S.A., 2002. The cortical sources of the early components of the visual evoked potential. *Hum. Brain Mapp.* 15, 95–111.
- Dobbins, I.G., Schnyer, D.M., Verfaellie, M., Schacter, D.L., 2004. Cortical activity reductions during repetition priming can result from rapid response learning. *Nature* 428, 316–319.
- Dubeau, M.C., Iacoboni, M., Koski, L.M., Markovac, J., Mazziotta, J.C., 2001. Topography for body parts motion in the STS region. *Abstr.-Soc. Neurosci.* 27 Program No. 572.14.
- Fischl, B., Sereno, M.I., Dale, A.M., 1999. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum. Brain Mapp.* 8, 272–284.
- Friston, K.J., Frith, C.D., Turner, R., Frackowiak, R.S., 1995. Characterizing evoked hemodynamics with fMRI. *NeuroImage* 2, 157–165.
- Fujii, T., Fukatsu, R., Watabe, S., Ohnuma, A., Teramura, K., Kimura, I., Saso, S., Kogure, K., 1990. Auditory and agnosia without aphasia following a right temporal lobe lesion. *Cortex* 26, 263–268.
- Giraud, A.L., Price, C.J., 2001. The constraints functional neuroimaging places on classical models of auditory word processing. *J. Cogn. Neurosci.* 13, 754–765.
- Hauk, O., Johnsrude, I., Pulvermuller, F., 2004. Somatotopic representation of action words in human motor and premotor cortex. *Neuron* 41, 301–307.
- Humphries, C., Willard, K., Buchsbaum, B., Hickok, G., 2001. Role of anterior temporal cortex in auditory sentence comprehension: an fMRI study. *NeuroReport* 12, 1749–1752.
- Iacoboni, M., 2004. Understanding others: imitation, language, empathy. In: Hurley, S., Chater, N. (Eds.), *Perspectives on Imitation: From Cognitive Neuroscience to Social Science*. MIT Press. (In press).
- Iacoboni, M., Woods, R.P., Brass, M., Bekkering, H., Mazziotta, J.C., Rizzolatti, G., 1999. Cortical mechanisms of human imitation. *Science* 286, 2526–2528.
- Iacoboni, M., Koski, L.M., Brass, M., Bekkering, H., Woods, R.P., Dubeau, M.C., Mazziotta, J.C., Rizzolatti, G., 2001. Reafferent copies of imitated actions in the right superior temporal cortex. *Proc. Natl. Acad. Sci. U. S. A.* 98, 13995–13999.
- Jeannerod, M., Arbib, M.A., Rizzolatti, G., Sakata, H., 1995. Grasping objects: the cortical mechanisms of visuomotor transformation. *Trends Neurosci.* 18, 314–320.
- Kohler, E., Keysers, C., Umiltà, A., Fogassi, L., Gallese, V., Rizzolatti, G., 2002. Hearing sounds, understanding actions: action representation in mirror neurons. *Science* 297, 846–848.
- Mazziotta, J.C., Toga, A.W., Evans, A., Fox, P., Lancaster, J., 1995. A probabilistic atlas of the human brain: theory and rationale for its development. The international consortium for brain mapping (ICBM). *NeuroImage* 2, 89–101.
- Miller, E.K., Desimone, R., 1991. A neural mechanism for working and recognition memory in inferior temporal cortex. *Science* 254, 1377–1379.
- Milner, A.D., Goodale, M.A., 1995. *The Visual Brain in Action*. Oxford Univ. Press.
- Mistlin, J., Perret, D.I., 1990. Visual and somatosensory processing in the macaque temporal cortex: the role of 'expectation'. *Exp. Brain Res.* 82, 437–450.
- Naccache, L., Dehaene, S., 2001. The priming method: imaging unconscious repetition priming reveals an abstract representation of number in the parietal lobes. *Cereb. Cortex* 11, 966–974.
- Pinard, M., Chertkow, H., Black, S., Peretz, I., 2002. A case study of pure word deafness: modularity in auditory processing? *Neurocase* 8, 40–55.
- Rizzolatti, G., Arbib, M.A., 1998. Language within our grasp. *Trends Neurosci.* 21, 188–194.
- Rizzolatti, G., Luppino, G., Matelli, M., 1998. The organization of the cortical motor system: new concepts. *Electroencephalogr. Clin. Neurophysiol.* 106, 283–296.
- Romanski, L., Goldman-Rakic, P., 2002. An auditory domain in primate prefrontal cortex. *Nat. Neurosci.* 5, 15–16.
- Saygin, A.P., Dick, F., Wilson, S., Dronkers, N., Bates, E., 2003. Neural resources for processing language and environmental sounds: evidence from aphasia. *Brain* 126, 928–945.
- Schnyer, D.M., Ryan, L., Trouard, T., Foster, K., 2002. Masked word repetition results in increased fMRI signal: a framework for understanding signal changes in priming. *NeuroReport* 13, 281–284.
- Seltzer, D.N., Pandya, J., 1994. Parietal, temporal, and occipital projections to cortex of the superior temporal sulcus in the rhesus monkey: a retrograde tracer study. *J. Comp. Neurol.* 343, 445–463.
- Spreen, O., Benton, A.L., Fincham, R.W., 1965. Auditory agnosia without aphasia. *Arch. Neurol.* 13, 84–92.
- Talairach, J., Tournoux, P., 1988. *Co-Planar Stereotaxic Atlas of the Human Brain*. Thieme, New York.
- Taniwaki, T., Tagawa, K., Sato, F., Iino, K., 2000. Auditory agnosia restricted to environmental sounds following cortical deafness and generalized auditory agnosia. *Clin. Neurol. Neurosurg.* 102, 156–162.
- Thierry, G., Giroud, A., Price, C., 2003. Hemispheric dissociation in access to the human semantic system. *Neuron* 38, 499–506.
- Varney, N., 1980. Sound recognition in relation to aural language comprehension in aphasic patients. *J. Neurol., Neurosurg. Psychiatry.* 43, 71–75.
- Vignolo, L.A., 1982. Auditory agnosia. *Philos. Trans. R. Soc. Lond., Ser. B Biol. Sci.* 298, 49–57.
- Zatorre, R.J., Belin, P., 2001. Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* 11, 946–953.