NeuroImage 48 (2009) 475–485

Contents lists available at ScienceDirect

NeuroImage

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

The role of actions in auditory object discrimination

Marzia De Lucia [a,c], Christian Camen [c], Stephanie Clarke [c], Micah M. Murray [a,b,c,d,*]

[a] *Electroencephalography Brain Mapping Core, Center for Biomedical Imaging of Lausanne and Geneva, Switzerland*
[b] *The Functional Electrical Neuroimaging Laboratory, Neuropsychology and Neurorehabilitation Service, University Hospital Center and University of Lausanne, Nestlé Hospital, 5 Av. Pierre-Decker, 1011 Lausanne, Switzerland*
[c] *The Functional Electrical Neuroimaging Laboratory, Radiology Service, University Hospital Center and University of Lausanne, Nestlé Hospital, 5 Av. Pierre-Decker, 1011 Lausanne, Switzerland*
[d] *Department of Hearing and Speech Sciences, Vanderbilt University, Nashville, TN, USA*

ARTICLE INFO

Article history:
Received 27 February 2009
Revised 12 June 2009
Accepted 16 June 2009
Available online 24 June 2009

Keywords:
Action
Auditory evoked potential (AEP)
Electrical neuroimaging
Event-related potential (ERP)
Mirror neuron
Object recognition

ABSTRACT

Action representations can interact with object recognition processes. For example, so-called mirror neurons respond both when performing an action and when seeing or hearing such actions. Investigations of auditory object processing have largely focused on categorical discrimination, which begins within the initial 100 ms post-stimulus onset and subsequently engages distinct cortical networks. Whether action representations themselves contribute to auditory object recognition and the precise kinds of actions recruiting the auditory–visual mirror neuron system remain poorly understood. We applied electrical neuroimaging analyses to auditory evoked potentials (AEPs) in response to sounds of man-made objects that were further subdivided between sounds conveying a socio-functional context and typically cuing a responsive action by the listener (e.g. a ringing telephone) and those that are not linked to such a context and do not typically elicit responsive actions (e.g. notes on a piano). This distinction was validated psychophysically by a separate cohort of listeners. Beginning ~300 ms, responses to such context-related sounds significantly differed from context-free sounds both in the strength and topography of the electric field. This latency is >200 ms subsequent to general categorical discrimination. Additionally, such topographic differences indicate that sounds of different action sub-types engage distinct configurations of intracranial generators. Statistical analysis of source estimations identified differential activity within premotor and inferior (pre)frontal regions (Brodmann's areas (BA) 6, BA8, and BA45/46/47) in response to sounds of actions typically cuing a responsive action. We discuss our results in terms of a spatio-temporal model of auditory object processing and the interplay between semantic and action representations.

Introduction

Recognition of visual, linguistic, and auditory stimuli can be influenced by associated actions (Gibson, 1979; Rizzolatti et al., 1996, Grèzes et al., 2003; Aziz-Zadeh et al., 2004; Barraclough et al., 2005; Pizzamiglio et al., 2005; Pulvermüller, 2005; Tettamanti et al., 2005; Hauk et al., 2006, 2008; Lahav et al., 2007; Pazzaglia et al., 2008), and distinct neuronal response patterns or networks can be observed for objects linked to actions. In the case of sounds, these networks can include premotor and (pre)frontal cortices often, but not exclusively, attributed to the so-called audio–visual mirror neuron system (Kohler et al., 2002; Keysers et al., 2003). More generally, such activations are consistent with current anatomical models of the auditory 'what' pathway (Rauschecker, 1998; Rauschecker and Tian, 2000) that includes projections from auditory regions of the superior temporal cortex ultimately to prefrontal and premotor regions (Romanski et al., 1999a,b; Kaas and Hackett, 2000). In agreement, functional imaging studies have documented reliable activations within prefrontal cortices and elsewhere in response to environmental sounds and vocalizations (e.g. Lewis et al., 2005; Fecteau et al., 2005; Murray et al., 2006). One implication of this construct is that action representation is itself operating in concert with and perhaps guiding object recognition processes. However, the precise spatio-temporal relationship between object and action-related processes remains poorly understood, particularly with regard to sounds of environmental objects, and was the focus of the present electrical neuroimaging study that capitalized on the high temporal resolution of scalp-recorded electroencephalography as well as recent improvements in source estimations (Michel et al., 2004).

Auditory object recognition has been shown to include categorical discrimination, such that sounds of living (including the sub-category of vocalizations) and man-made environmental objects, for example, can engage distinct brain networks (Belin et al., 2000; Fecteau et al., 2005; Lewis et al., 2005; Altmann et al., 2007) at post-stimulus latencies as early as 70 ms and with different durations of activity

---

[*] Corresponding author. Electroencephalography Brain Mapping Core, CIBM Radiology, BH08.078, CHUV, rue du Bugnon 46, 1011 Lausanne, Switzerland. Fax: +41 21 3144645.
E-mail address: micah.murray@chuv.ch (M.M. Murray).

1053-8119/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2009.06.041

(Murray et al., 2006; reviewed in Murray and Spierer, 2009). In particular, stronger responses have been observed to sounds of man-made objects within premotor and prefrontal cortices (Lewis et al., 2005; Murray et al., 2006), raising the possibility that sounds of man-made objects have stronger associations with action representations than sounds of living objects, which may instead have stronger associations with visual representations (e.g. Murray et al., 2004, 2005; Amedi et al., 2005). While the precise type(s) of actions necessary to elicit response modulations in these premotor and prefrontal regions remain undefined, it is noteworthy that the stimuli in the abovementioned studies included a wide variety of man-made objects, including tools (Lewis et al., 2005) as well as a mixture of musical instruments, household items/appliances, and alarms (Murray et al., 2006). Still others have documented responses within lateral ventral (pre)frontal cortices in response to vocalizations (Fecteau et al., 2005 for functional magnetic resonance imaging results; Murray et al., 2006 for electrical neuroimaging data).

This pattern of results led us to hypothesize that there is general activity within mirror neuron regions in response to sounds of objects that may in turn modulate as a function of sub-types of associated actions. One line of support for this hypothesis comes from research examining responsiveness of neurons within ventral (lateral) prefrontal cortices (vPFC) to animal vocalizations. These neurons differentially responded to vocalizations referring to food discovery vs. other communicative situations, irrespective of the quality of the foods to which they referred (Cohen et al., 2006). Such results are suggestive of a dichotomy in the responsiveness within vPFC (and perhaps elsewhere) between sound categories that may reflect their social and/or functional context as well as their cuing of the listener to react in a specific manner (e.g. partake in the discovered food vs. greetings). The present study considers two sub-groups of sounds of actions: those conveying a specific social and/or functional context often cuing listeners to act in response and those sounds not forcibly linked to a specific context and not cuing a responsive action. We use the terms ‘context-related’ and ‘context-free’, respectively, as shorthand to refer to this distinction (see Seyfarth et al., 1980; Hauser, 1998 for similar varieties of distinctions).

Links between modulated activity within the mirror system and action representations elicited by sounds have been established (e.g. Kohler et al., 2002; Keysers et al., 2003; Hauk et al., 2006, 2008; Kaplan and Iacoboni, 2006; Galati et al., 2008). For example, in their study that first described the responsiveness of ventral premotor mirror neurons to sounds of actions, Kohler et al. (2002) presented macaque monkeys with sounds of actions (e.g. paper ripping or a stick hitting the floor), animal vocalizations, and noise bursts. They found that while these neurons reliably responded to sounds of actions they failed to exhibit robust responses to sounds of vocalizations or noise bursts (see also Keysers et al., 2003). In addition, these authors reported a near-perfect correspondence between a single neuron's selectivity for a given action when presented as a sound and when presented visually. This selectivity and inter-sensory correspondence would suggest that these modulations are not reflecting simple semantic analysis. However, specification of the spatio-temporal brain dynamics of semantic and action-related processes remains to be fully established.

Investigations in humans that studied the interplay between environmental sound recognition and action representations are relatively rare and have thus far generated discordant conclusions regarding the temporal dynamics of these processes. On the one hand, Pizzamiglio et al. (2005) reported effects starting at ~300 ms post-stimulus onset using a masked repetition priming paradigm with sounds produced by human beings or not (e.g. hands clapping vs. water boiling). By contrast, Hauk et al. (2006) reported effects as early as ~100 ms post-stimulus onset using an adaptation of a multi-deviant mismatch negativity (MMN) paradigm and comparing responses to clicks produced by the finger or tongue both with each other and also with respect to acoustically controlled synthetic variants. While the use of an MMN paradigm allowed Hauk et al. to also assess whether action representations are accessed pre-attentively, a potential limitation of their contrast, which the authors themselves acknowledge, is that complex spectral features were only presented in the naturalistic stimuli and could have elicited larger MMNs than the control sounds (though such would not account for the topographically distinct and somatopic effects they observed between finger and tongue sounds). That is, larger MMNs have been reported for meaningful than for meaningless control stimuli (e.g. Frangos et al., 2005; also Hauk et al., 2006 for discussion). More generally, the difference in the latency of effects reported in these studies could stem from numerous sources, including but not limited to task-related effects (i.e. explicit discrimination of actions in Pizzamiglio et al. 2005 vs. passive listening in Hauk et al., 2006). As such, it remains unresolved both when action representations are accessed, in particular with respect to ordinate-level object discrimination, and whether such access occurs pre-attentively.

A further complication for generating a synthesis in terms of the necessary conditions for observing response modulations within the human auditory mirror neuron system is that action-related differences between stimuli are often confounded by semantic differences. For example, response differences between the sound of paper being ripped and a non-speech vocalization may either reflect action-related processes and/or man-made vs. living categorization. The present study sought to circumvent this confound by comparing different subtypes of sounds of man-made environmental objects that were further sorted between ‘context-related’ and ‘context-free’ actions. Specifically, we applied electrical neuroimaging analyses (Murray et al., 2008a) to auditory evoked potentials (AEPs) in response to distracter trials during a living vs. man-made discrimination task in order to identify the spatio-temporal mechanism whereby representations of responsive actions impact sound discrimination and situate such with respect to current models of auditory object processing (Griffiths and Warren, 2004; Murray and Spierer, 2009).

Materials and methods

Subjects

Ten healthy, right-handed individuals (7 female), aged 21–34 years participated. All subjects provided written, informed consent to participate in the study, the procedures of which were approved by the Ethics Committee of the University of Geneva. None had a history of neurological or psychiatric illnesses, and all reported normal hearing. None were musicians. Data from these individuals have been previously published in an investigation of living versus man-made categorical discrimination (Murray et al., 2006). The present study is thus a further analysis of a subset of these data (i.e. the AEPs to specific subclasses of man-made stimuli).

Stimuli

Auditory stimuli were complex, meaningful sounds (16 bit stereo; 22,500 Hz digitization) of common environmental objects (cf. Table 1 in Murray et al., 2006 for a full listing, including details on the acoustic attributes as well as psychometrics concerning these stimuli). There were 120 different sound files in total, 60 of which represented sounds of living objects (3 exemplars of 20 different referent objects) and 60 of which represented sounds of man-made objects (3 exemplars of 20 different reference objects). Each sound was 500 ms in duration, which included an envelope of 50 ms decay time that was applied to the end of the sound file to minimize clicks at sound offset. All sounds were further normalized according to the root mean square of their amplitude. The stimuli at the focus of the present investigation were restricted to those of man-made objects, which were further sorted between objects associated with a specific socio-functional context and objects not associated with such (hereafter 'context-related' and 'context-free', respectively). The 10 sounds of context-free actions included exemplars of notes being played on the following musical instruments (3 exemplars per object): accordion, flute, guitar, harmonica, harp, organ, piano, saxophone, trumpet, and violin (i.e. both string and brass instruments involving mouth and hand actions). We would emphasize that these stimuli were neither rhythmic nor melodic in character and were not perceived as music, but rather in terms of the instrument generating the sound. The 10 sounds of context-related actions included exemplars of the following objects (3 per object): bicycle bell, car horn, cash register, cuckoo clock, doorbell, closing door, glass shattering, police siren, church bell, and telephone (i.e. sounds that typically trigger a responsive action upon being heard, as supported by our psychophysical experiment below). Three distinct exemplars of each object were used during the experiment, yielding a total of 60 sound files.

Our prior work has demonstrated that the sounds used in this study were all highly familiar as well as reliably identified with a high level of confidence (see Table 1 in Murray et al., 2006). However, to more empirically assess whether or not the context-related versus context-free distinction we invoke in this study was indeed experienced by listeners, we asked a cohort of 18 healthy individuals aged 22–36 years (12 women), none of whom were musicians and none of whom participated in the EEG portion of the study, to listen to each of the 60 sound files and rate each on two attributes using a 5-point Likert scale. First, we asked them to indicate whether hearing the sound would typically elicit a responsive action on their part (1 referring to no such tendency and 5 referring to a strong tendency). Second, we asked them to indicate whether the sound was emotionally charged (1 indicating no emotional content and 5 a very strong emotional content). The mean ratings for each question for context-related and context-free sounds are shown in Fig. 1. Context-related sounds were reliably more highly rated as typically eliciting a responsive action (2-tailed paired *t*-test; $t_{(17)}=4.85$; $p<0.0002$). By contrast, no differences were observed between the sounds in terms of their emotionality ($t_{(17)}=0.97$; $p>0.30$). To assess whether mental imagery of the heard action itself varied between sounds, a subset of these individuals ($N=8$) rated the sounds according to the extent to which they could imagine producing the action required to generate the sound. That is, one could contend that mental imagery of the sounds of context-free actions could be higher than that with sounds of context-related actions because the former included only musical instruments that could be readily pantomimed (though we would remind the reader that the sounds were 500 ms in duration and entailed notes without rhythm or melody). Contrary to this possibility, context-related sounds resulted in significantly higher ratings of imagining the actions necessary to produce the sound than context-free sounds (2.99 ± 0.17 vs. 1.85 ± 0.13; $t_{(7)}=5.08$; $p<0.002$).

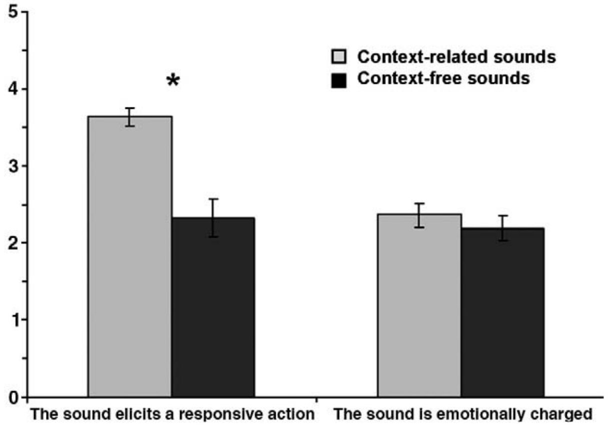

Fig. 1. Psychophysical evaluation of stimuli. The bar graph displays the group-average ($N=18$; s.e.m. shown) ratings on each of two questions for context-related and context-free sounds (gray and black bars, respectively). Ratings were based on a 5-point Likert scale, which larger values indicating stronger positive responses. Asterisk indicates significant pair-wise differences.

To assess whether these sounds differed acoustically, we statistically compared the spectrograms (defined with Matlab's spectrogram function with no overlapping and zero padding), using a time-frequency bin width of ~5 ms and ~80 Hz. Statistical contrasts entailed a series of non-parametric *t*-tests based on a bootstrapping procedure with 1000 iterations in order to derive an empirical distribution against which to compare the actual difference between the mean spectrograms from each sound category (Manly, 1991). To partially correct for multiple contrasts and autocorrelation a significant difference at a given time-frequency bin was only considered reliable if all 8 of its immediately adjacent bins also yielded *p*-values ≤0.00625 (i.e. a 3×3 bin threshold was applied). Statistically reliable differences between the spectrograms from each group of sounds were limited to low frequencies (below ~200–300 Hz; see Supplementary Fig. 1). The sounds were likewise compared in terms of their mean harmonics-to-noise ratio (HNR), which was calculated using PRAAT software (http://www.fon.hum.uva.nl/praat/). Briefly, the HNR provides an index of the ratio of the energy contained in the harmonics of a sound to that contained in the non-harmonics (i.e. noise). The mean (±s.e.m.) HNR for the 30 context-free sounds was 12.9 ± 1.5 (range -0.1 to 33.5) and that for the 30 context-related sounds was 9.2 ± 1.8 (range -3.0 to 28.5). These HNR values did not significantly differ ($t_{(28)}=1.59$; $p>0.10$).

Procedure and task

Participants performed a living versus man-made 'oddball' detection paradigm, such that on a given block of trials 'target' stimuli to which subjects pressed a response button occurred 10% of the time. The use of sounds of living and man-made objects as target stimuli was counterbalanced across blocks. The remaining 90% of stimuli ('distracters') were comprised of the other (i.e. non-target) sound category. Stimuli were blocked into a series of 300 trials (~18 min) with an inter-stimulus interval of 3.4 s. Each participant completed 4 blocks of trials (2 where man-made sounds were targets and 2 where living sounds were targets) and took a 5–10 min break between blocks in order to minimize fatigue. Both behavioral as well as EEG data were collected from all conditions throughout the length of the experiment, and STIM (Neuroscan, Inc., El Paso, Texas, USA) was used to control stimulus delivery and to record behavioral responses. Audiometric quality insert earphones (supplied by Neuroscan, Inc.) were used for stimulus delivery.

EEG acquisition and pre-processing

Continuous 64-channel EEG was acquired through Neuroscan Synamps (impedances <5 kΩ), referenced to the nose, band-pass filtered 0.05–200 Hz, and digitized at 1000 Hz. Peri-stimulus epochs of continuous EEG (-100 ms to 900 ms) from distracter trials (i.e. those not requiring a motor response) were averaged from each subject separately to compute auditory evoked potentials (AEPs). EEG responses to target trials were not analyzed as they were too few for sufficient signal quality, though the behavioral results reported below refer to these trials. The average (±s.e.m.) number of accepted EEG epochs in response to context-free sounds and context-related sounds was 189 ± 17 and 189 ± 15, respectively ($t_{(9)}=0.092$; $p>0.90$). Trials with blinks or eye movements were rejected off-line, using horizontal and vertical electro-oculograms, and were excluded from AEPs, using a threshold of ±60 μV at these channels. An artifact criterion of ±100 μV was applied at all other electrodes and each EEG epoch was also visually evaluated prior to its inclusion in the AEP. Data from artifact electrodes from each subject and condition were interpolated and down-sampled to a common 61-channel montage that excluded the bipolar EOG channels (3-dimensional spherical splines; Perrin et al., 1987). Following this procedure and prior to group-averaging, each subject's data were 40 Hz low-pass filtered, baseline corrected using the −100 ms pre-stimulus period, and recalculated against the common average reference.

EEG analyses and source estimation

Effects were identified with a multi-step analysis procedure, which we refer to as electrical neuroimaging, examining both local and global measures of the electric field at the scalp. These analyses have been detailed elsewhere (Murray et al., 2004, 2008a; Michel et al., 2004). Briefly, it entails analyses of response strength and response topography to differentiate effects due to modulation in the strength of responses of statistically indistinguishable brain generators from alterations in the configuration of these generators (viz. the topography of the electric field at the scalp), as well as latency shifts in brain processes across experimental conditions. In addition, we utilized the local autoregressive average distributed linear inverse solution (LAURA; Grave de Peralta Menendez et al., 2001, 2004) to visualize and statistically contrast the likely underlying sources of effects identified in the preceding analysis steps.

Electrical neuroimaging analyses have several advantages over canonical waveform analyses. For example, waveform analyses are reference-dependent and statistical outcomes will thus change with the choice of the reference electrode (Murray et al., 2008a). Additionally, such analyses often require the pre-selection of time windows and electrode sites that will be submitted to statistical tests. Such a procedure unnecessarily introduces experimenter bias. Nonetheless, in order to minimize the possibility of missed effects, we further analyzed average-reference waveform data from all electrodes as a function of time post-stimulus onset in a series of pair-wise *t*-tests with correction for temporal autocorrelation at individual electrodes through the application of a 15 consecutive data point criterion for the persistence of differential effects (Guthrie and Buchwald 1991), which is equivalent to 15 ms duration in the present study (i.e. a 1000 Hz sampling rate was used). While this analysis gives a visual impression of effects within the dataset; our conclusions are based solely on global measures of the electric field at the scalp.

Changes in the strength of the electric field at the scalp were assessed using GFP (Lehmann and Skrandies 1980) from each subject and experimental condition. Values at each time point were compared with a paired *t*-test, as above. To statistically identify periods of topographic modulation, we calculated the global dissimilarity (Lehmann and Skrandies 1980) between responses for each time point and applied a Monte Carlo bootstrapping analysis procedure that is colloquially referred to as topographic ANOVA or “TANOVA” (detailed in Murray et al., 2008a). Because electric field changes are indicative of changes in the underlying generator configuration (Lehmann 1987), this analysis provides a statistical means of determining if and when brain networks mediating responses to context-related and context-free sounds differ.

A clustering analysis of the AEP topography at the scalp identified time periods of stable topography. This approach is a data-driven means for defining AEP components. The most dominant scalp topographies appearing in the AEPs of the group-averaged data were identified with a modified agglomerative hierarchical clustering (Murray et al., 2008a), and the optimal number of topographies or ‘template maps’ that accounted for the whole data set (i.e. the post-stimulus periods of both conditions, collectively) was determined by a modification of the Krzanowski–Lai (KL) criterion (Tibshirani et al., 2005; see Murray et al., 2008a for details). The pattern of template maps identified in the group-averaged data was then statistically tested in the data of each individual subject. For each time point, the AEP topography was compared to template maps and was labeled according to the one with which it best correlated spatially. The output is a measure of relative map presence that is in turn submitted to a repeated measure ANOVA with factors of condition and map. In conjunction with the aforementioned TANOVA, this procedure reveals whether AEPs from a given condition are more often described by one map versus another, and therefore whether different intracranial generator configurations better account for AEPs from each condition.

We estimated the sources in the brain underlying the ERPs from each condition using a distributed linear inverse solution (ELECTRA) applying the local autoregressive average (LAURA) regularization approach to address the non-uniqueness of the inverse problem (Grave de Peralta Menendez et al., 2001). The inverse solution algorithm is based on biophysical principles derived from the quasi-static Maxwell's equations; most notably the fact that independent of the volume conductor model used to describe the head, only irrotational and not solenoidal currents contribute to the EEG (Grave de Peralta Menendez et al., 2001, 2004). This is further supported by empirical evidence showing that secondary macroscopic currents are several orders of magnitude higher than primary macroscopic currents in excitable tissues (Plonsey, 1982). The consequence of this property is that the inverse problem can be reformulated based on a first kind Fredholm integral equation. Despite this solid theoretical and empirical underpinning, uniqueness to the inverse problem requires the application of a regularization strategy. As mentioned above, we here used LAURA, which applies the physical principle concerning the spatial relationships between the estimated potentials at neighboring voxels, such that potentials decay as a function of the square distance to the source. In our study, homogenous regression coefficients in all directions and within the whole solution space were used. LAURA uses a realistic head model, and the solution space included 4024 nodes, selected from a 6×6×6 mm grid equally distributed within the gray matter of the Montreal Neurological Institute's average brain (courtesy of R. Grave de Peralta Menendez and S. Gonzalez Andino; http://www.electrical-neuroimaging.ch/). As an output, LAURA provides current density measures, the scalar values of which can be statistically analyzed at each node. Prior basic and clinical research has documented and discussed in detail the spatial accuracy of this inverse solution (e.g., Gonzalez Andino et al. 2005a,b; Grave de Peralta Menendez et al. 2004; Michel et al. 2004). Regarding the precision of the source estimations, activity within relatively superficial cortical locations is on the order of the grid size of the solution space (i.e., 6×6×6 mm) and the resolution kernels are correctly centered around the right source location with little interference from distant sources (little blurring; see e.g. Martuzzi et al., 2009).

The topographic pattern analysis was used to define the time period (s) over which sources were estimated for each subject and both conditions. The mean difference between source estimations was used as a basis for identifying clusters for subsequent statistical analysis. Specifically, we first identified the position(s) of maximal differences throughout the 4024 solution points. In the case of the present study, 4 such maxima were observed (see Fig. 4c). We then calculated the mean and standard error across subjects at these local maxima. For inclusion within a given cluster, the mean difference at any solution point was required to be within 1 standard error of the maximal difference for that cluster. Moreover, to be considered robust each cluster was required to include at least 12 solution points (see also Toepel et al., 2009 for a similar spatial criterion). In the case of the present study, clusters ranged in size from 13 to 44 solution points. Statistical contrasts were then performed for each cluster separately using the mean values across solution points for each subject and condition.

Results

Behavioral results

Participants accurately performed the target detection task (see Murray et al., 2006 for details). In terms of their performance with context-free and context-related sounds, the mean (± s.e.m.) percentage of correct responses were 94.0 ± 3.1% and 93.5 ± 1.9%, respectively, and did not significantly differ ($t_{(9)} = 0.25$; $p > 0.8$). Likewise, reaction times to context-free sounds versus context-related sounds were 883 ± 41 ms and 903 ± 38 ms, respectively, and did not significantly differ ($t_{(9)} = 0.76$; $p > 0.45$). These reaction times are consistent with previous studies of environmental sound recognition where reaction times on the order of ~1 s were obtained (Lebrun et al., 1998; Saygin et al., 2003; Bergerbest et al., 2004; De Lucia et al., 2009). Thus, behavioral differences cannot readily account for any AEP modulations. In addition, we would remind the reader that all AEP analyses were based on distracter trials (i.e. those without a motor response), minimizing (if not altogether removing) the role of any response-related activity in the effects we obtained.

Electrical neuroimaging results

The first level of analysis focused on determining the onset of response differences between sounds, which entailed a millisecond-by-millisecond paired *t*-test across the entire electrode montage. Fig. 2 displays group-averaged AEP waveforms from exemplar midline electrodes (FPz, Fz, Cz, Pz, and Oz). These waveforms exhibit prototypical AEP components that are maximal over fronto-central scalp locations — i.e. P50, N1, P2, and N2 (e.g. Picton et al., 1999). A response modulation is evident over the approximately 300–350 ms post-stimulus interval. Fig. 3a displays the results of this analysis across the entire electrode montage. Temporally-sustained and statistically reliable differences were observed across the majority of the montage over the approximately 300–600 ms post-stimulus interval.

The second level of analysis focused on determining the underlying neurophysiologic basis of the differential responses observed across the electrode montage, using global measures of the AEP. Paralleling our observations at the individual electrode level, the millisecond-by-millisecond analysis of the group-averaged global field power (GFP) waveforms revealed sustained differences between responses over the 298–359 ms and 422–562 ms post-stimulus periods (Fig. 3b). Second, global dissimilarity between conditions tested on a millisecond-by-millisecond basis whether the topographies of the AEPs differed between conditions. Sustained topographic differences were first observed over the 295–359 ms post-stimulus period and also thereafter (Fig. 3c).

A hierarchical topographic cluster analysis was conducted to identify time periods of stable electric field topography both within and between experimental conditions. This analysis, first performed at the group-averaged AEP level, is a means of identifying AEP components and for determining whether the above topographic modulation identified via TANOVA follows from a singular and stable topographic difference or rather from multiple configuration changes (see Murray et al., 2008b for discussion). The global explained variance of this clustering for the concatenated group-averaged dataset from both experimental conditions was 97.72%. This analysis indicated that one map predominated responses to context-free sounds whereas another predominated responses to context-related sounds over the 313–360 ms post-stimulus period (Fig. 4a). Specifically, responses to context-free sounds were characterized by a right-lateralized fronto-central negativity, whereas responses to context-related sounds were characterized by a bilateral fronto-central negativity. The topographic clustering at the group-average AEP level was statistically evaluated using a measure of map presence that is based on the spatial correlation between the template maps identified in the group-averaged AEPs and single-subject data. Over the 313–360 ms period there was a significant interaction between factors of experimental condition and template map ($F_{(1,9)} = 12.945$; $p = 0.006$).

Analyses to this point indicate that AEP responses to context-related and context-free sounds first differed both in their strength

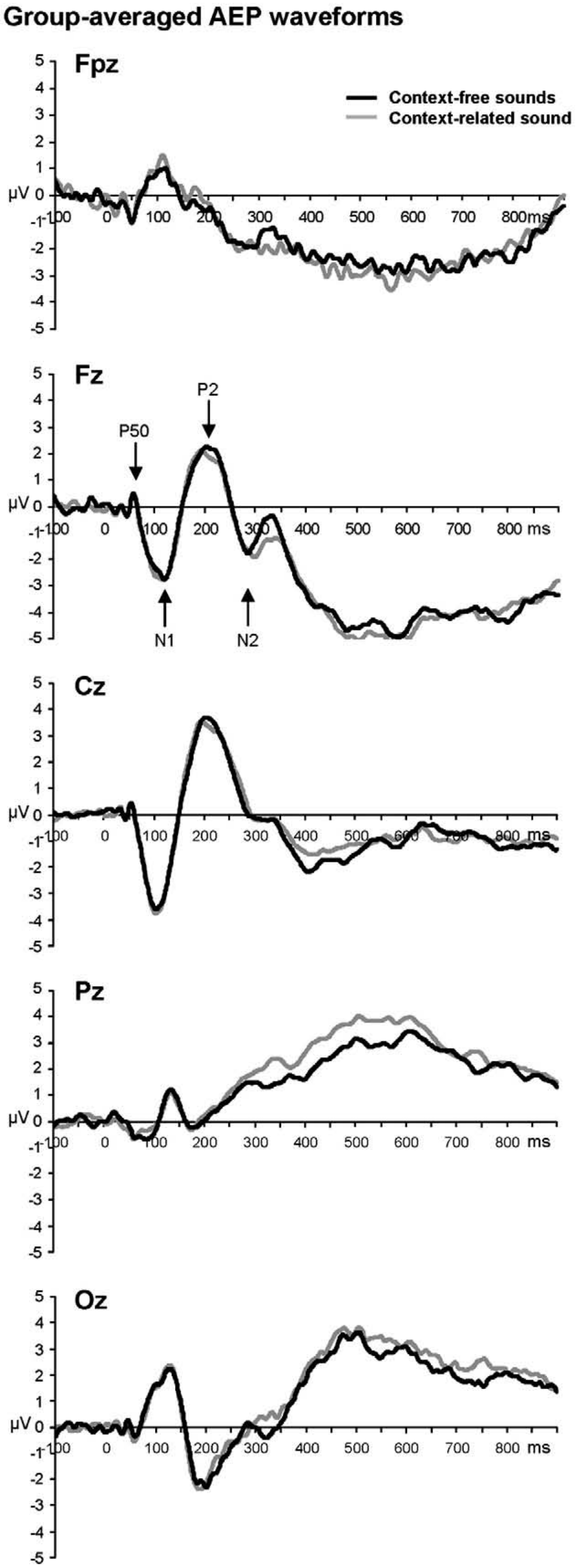

Fig. 2. Exemplar AEP waveforms from a set of midline electrodes. These group-averaged ($N = 10$) AEP waveforms exhibit prototypical peaks. Response modulations are visually apparent from ~300 ms post-stimulus onset. See Fig. 3 for a detailed statistical analysis of these waveforms on a millisecond-by-millisecond basis.

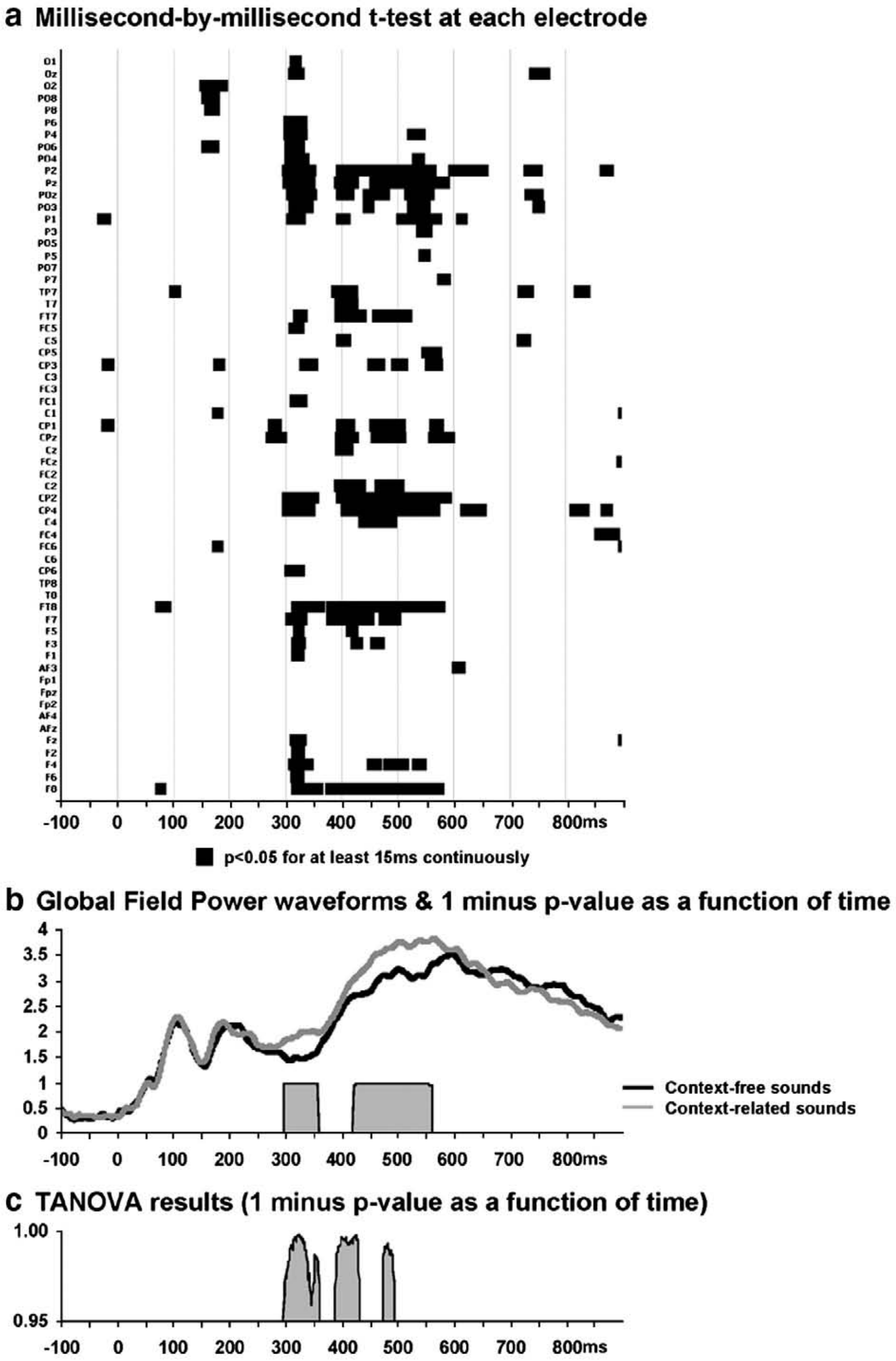

Fig. 3. Local and global electric field analyses. (a) The results of millisecond-by-millisecond paired *t*-tests at each of the scalp electrodes from the group-averaged ($N = 10$) AEP waveforms are shown as an intensity plot (only $p < 0.05$ for a minimum of 15 consecutive milliseconds are shown). (b) Modulations in response strength were identified using global field power (GFP). Group-average GFP waveforms are displayed along with the results of millisecond-by-millisecond paired *t*-tests that include a temporal criterion of 15 consecutive milliseconds. (c) Topographic modulations between conditions were assessed using global dissimilarity. The results of the TANOVA procedure are illustrated as a function of time (1 minus *p*-value shown; 15 ms temporal criterion, as above).

and topography over the 313–360 ms period and that this difference followed from singular and stable topographies better accounting for each condition. By extension, such topographic differences result from changes in the configuration of the underlying intracranial sources, which were in turn estimated for each condition and participant over the 313–360 ms period and then group-averaged (Fig. 4b). Both conditions included prominent sources within the superior and middle temporal lobes of the left hemisphere, extending posteriorly towards the temporo-parieto-occipital junction. In addition, context-related sounds included sources within the right middle temporal cortex as well as left lateral inferior frontal cortex and dorsal prefrontal cortices, bilaterally. Fig. 4c displays the mean difference of these source estimations. From this mean difference 4 clusters were identified (see Materials and methods for details), which were located in Brodmann's Areas (BA) 45/46/47 of the left hemisphere, BA6 of the left hemisphere, and BA8 bilaterally. The cluster within BA45/46/47 of the left hemisphere included 44 solution points, and the maximal difference was located at −41, 34, −5 mm using the coordinate system of Talairach and Tournoux (1988). The cluster within BA6 of the left hemisphere included 13 solution points and the maximal

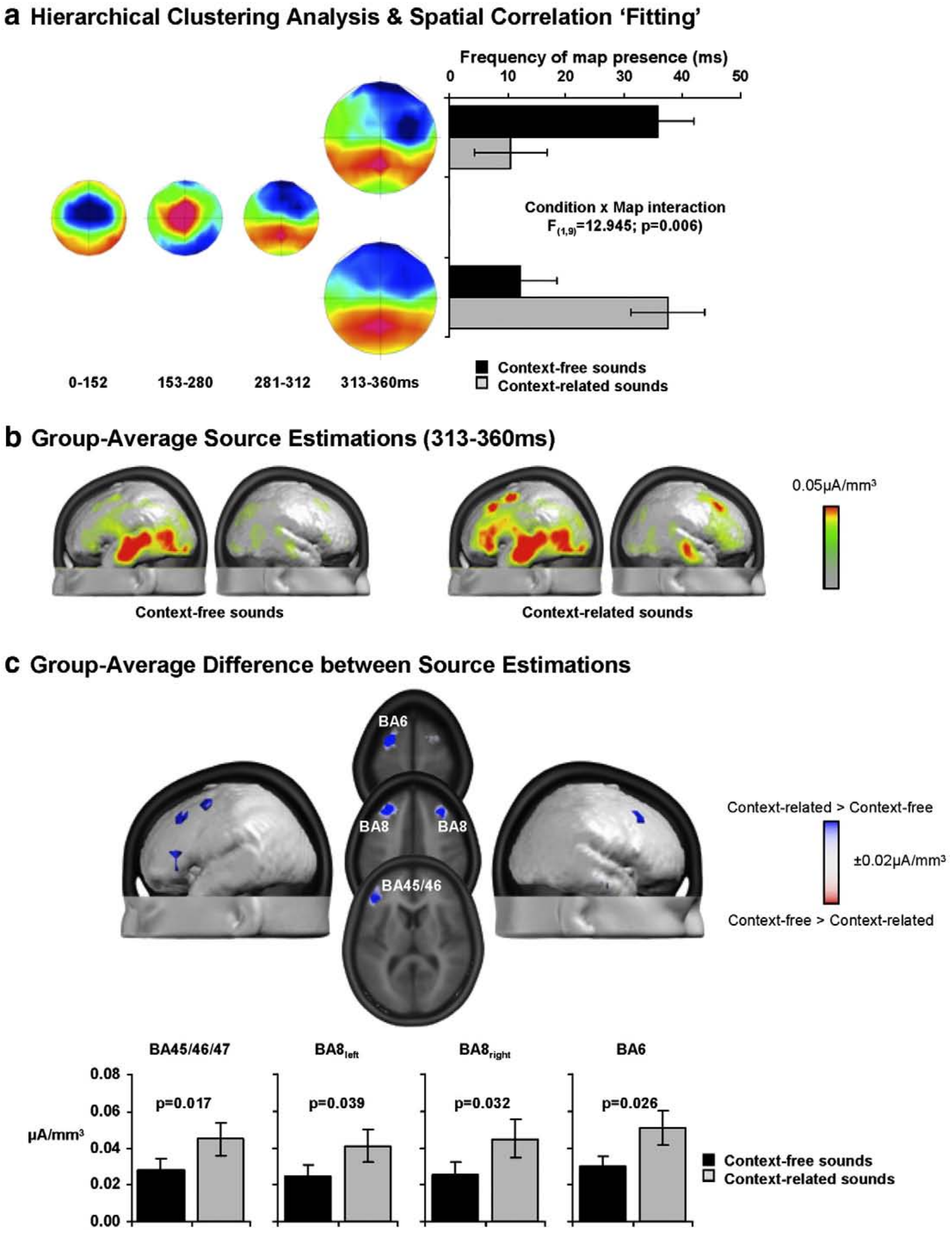

Fig. 4. Topographic analyses and source estimations. (a) The hierarchical clustering analysis identified different electric field topographies in response to each condition over the 313–360 ms post-stimulus period. An identical sequence of topographies was identified for both conditions over earlier time periods. The bar graph illustrates the results of the spatial correlation 'fitting' analysis applied to the single-subject data from each condition (topographies are shown with the left hemiscalp on the left and the nose upwards; time period of observation indicated). (b) Group-averaged distributed linear source estimations were calculated over the 313–360 ms post-stimulus period for each experimental condition (scale indicated). (c) The mean difference of these source estimations are rendered on the average brain of the Montreal Neurological Institute. Left- and right-sided views are shown in addition to axial slices at $z=9$ mm, 48 mm, and 57 mm. The bar graphs illustrate the mean (s.e.m. indicated) scalar value within each cluster for each condition (BA = Brodmann's Area).

difference was located at −29, 3, 62 mm. The cluster within BA8 of the left hemisphere included 22 solution points and the maximal difference was located at −29, 31, 44 mm. The cluster within BA8 of the right hemisphere included 32 solution points and the maximal difference was located at 35, 31, 44 mm. Means values within each of these clusters were calculated for each subject and condition and analyzed with a paired *t*-test (bar graphs in Fig. 4c). In each cluster source estimations were significantly stronger in response to context-related sounds than to context-free sounds.

Discussion

We identified the timing and neurophysiologic mechanism by which sounds of man-made environmental objects, all of which involve actions for their generation, are discriminated from one another. To determine the role of action representations in object discrimination, we focused here on the impact of whether or not the sounds typically cue the production of an action in response by the listener. Electrical neuroimaging analyses revealed that AEPs to context-related sounds (i.e. those that also typically cue the production of an action in response) differed both in their strength and topography beginning ~300 ms post-stimulus onset, indicative of the recruitment of distinct configurations of brain networks. Source estimations identified regions within the premotor and inferior (pre)frontal cortices as responding significantly more strongly to context-related than context-free sounds, consistent with the role of these areas in the audio–visual mirror neuron system, as will be discussed below. This differential processing is subsequent to several varieties of object processing, including the initial categorization of sounds of living and man-made objects observed over the 70–119 ms period (Murray et al., 2006; see also Supplementary Fig. 2), task-related modulations at 100 ms (Murray et al., 2006), and repetition-induced plasticity in auditory object processing over the 156–215 ms period (Murray et al., 2008b). Our results refine existing spatio-temporal models of auditory object processing to suggest that action representations differentially affect object discrimination only at relatively late stages (Murray and Spierer, 2009). They also further expand current understandings of object representations within the premotor and (pre) frontal cortices to highlight the interplay between object and action representations.

An important distinction of our experimental design was the comparison of AEPs in response to different subclasses of sounds of actions (i.e. two sub-groups of sounds of man-made objects). That is, many (if not all) prior studies intertwined the distinction between action and inaction with that between different supra-ordinate categories of objects, such as sounds of living and man-made objects, which have been repeatedly shown to themselves engage distinct brain networks (e.g. Lewis et al., 2005; Altmann et al., 2007; Doehrmann et al., 2008); parts of which include regions attributed to the mirror neuron system. Another possibility is that our effects reflect function-related categorical processes within higher-order auditory cortices along the recognition or 'what' functional pathway (see also Kaas and Hackett, 2000; Sugihara et al., 2006; Romanski, 2007). Such categorization effects would be similar to those described by Cohen et al. (2006) who presented rhesus monkeys with species-specific vocalizations and found that vPFC neurons differentially responded according to the functional category of the vocalization. Such a form of categorical discrimination is not in diametric opposition with attributes ascribed to the mirror neuron system. More fully resolving the extent of interplay between the identity of the sound's referent and its related action representations will require additional investigation. Nonetheless, the present study highlights that modulated responses within premotor and (pre)frontal cortices are reliably attained – albeit at relatively late latencies – when contrasting subclasses of sounds of man-made objects, all of which entail actions by the agent producing the sound.

It could be contended that our effects simply reflect semantic categorization of the stimuli at a level subordinate to the living vs. man-made distinction that was task-relevant, because all of the context-free sounds were musical instruments, whereas the context-related sounds were more semantically varied (at least in terms of the objects generating the sounds). While we cannot unequivocally refute this alternative interpretation, there was no evidence for selective activity in response to context-free sounds (i.e. the group of musical instruments failed to elicit stronger responses). Such would have been expected had these stimuli constituted a specialized (sub)category of objects. For example, con-specific vocalizations reliability elicit stronger responses, even when compared to sounds of other living objects or vocalizations from other species (e.g. Belin et al., 2000; von Kriegstein et al., 2007; Petkov et al., 2008). Rather, only the converse was observed here (i.e. larger responses to the semantically less homogenous condition). Likewise, that context-free sounds, here notes played on musical instruments, led to significantly weaker responses (Figs. 3b and 4c) is in strong agreement with recent fMRI findings where only learned melodies, but not individual notes, sufficed to activate the mirror neuron system (Lahav et al., 2007). Of particular interest, therefore, will be to investigate how the present effects are linked to learning appropriate responsive actions. The link between object recognition and accessing action representations has been demonstrated through neuropsychological examinations. Patients with apraxic syndromes can also exhibit parallel impairments in the recognition of environmental sounds that correspond to the impaired action, such that a patient with impaired arm movements will exhibit impaired recognition of sounds produced by arm actions, but not mouth actions (Pazzaglia et al., 2008). Whether or not such patients would nonetheless show differential responses to sounds of actions that elicit a responsive action remains to be investigated and would provide insights on the causal role of premotor and pre(frontal) cortices in auditory object discrimination. One possible direction for future research would be to apply a transcranial magnetic stimulation approach like that in Aziz-Zadeh et al., 2004, while asking participants to perform a recognition task with the sounds. Likewise, given the evidence for the activation of motor areas in a somatotopic fashion in response to action words, pictures, or sounds (Buccino et al., 2005; Hauk et al., 2006, 2008); it will likewise be interesting to investigate any somatotopic organization of the present responses. However, we would note that the sounds examined in this study were both produced with either the mouth or hand. Current efforts in developing single-trial topographic analyses (e.g. De Lucia et al., 2007a, b) will surely assist in addressing the question of potential somatotopic organization of responses to action sub-types in both healthy individuals as well as patients.

Another possibility is that the present effects reflect modulations in mental imagery and/or motor imagery. Recent evidence comparing brain activity during memory retrieval of imagined, pantomimed, observed, and enacted actions only obtained effects over the 800–1300 ms post-period (Senkfor, 2008). The latency of this effect is more than 500 ms later than our own, making it unlikely that we are observing an effect linked to retrieval of these kinds of memories or mental images. More generally with regard to the latency of our effects, it is worth noting reaction times for the living vs. man-made discrimination task of this study were approximately 900 ms and did not reliably differ between conditions. Lastly, it could be reasonably postulated that effects of mental imagery would lead to greater temporal smearing of the AEP, either within and/or across participants, because the imagined motor act would likely be 'triggered' at different instants across sounds and individuals. However, our data show that the AEP difference followed from a temporally delimited modulation in both response strength as well as topography (see Fig. 3).

The premotor and inferior (pre)frontal localization of our differential effects is consistent with regions identified in prior studies as part of the monkey and human audio–visual mirror neuron system (Kohler et al., 2002; Keysers et al., 2003; Aziz-Zadeh et al., 2004; Pizzamiglio et al., 2005; Tettamanti et al., 2005; Gazzola et al., 2006; Hauk et al. 2006; Lahav et al., 2007; Galati et al., 2008). Responses in such regions are obtained not only when the participant is performing an action or observing another doing so, but also when listening to actions performed by others. In the case of auditory mirror neurons, such effects appear to be independent of the acoustic features of the stimuli (which was controlled between categories in the case of the present study; see Supplementary Fig. 1). It is noteworthy, however, that there was no evidence here for differential responses within temporo-parietal structures that have been previously linked to the mirror neuron system. One possibility is that any such differential activity manifests at different latencies than the effects within premotor and prefrontal cortices we identified. Another is that these nodes of the mirror system are not modulated by whether or not an action sound is linked to a subsequent responsive action. Resolving this issue will clearly require additional studies. Such notwithstanding, the present study extends our understanding of the mirror neuron system by comparing responses to different subclasses of heard actions such that any difference was not due to actions, per se, but rather whether or not the sound typically cues a further responsive action on the part of the listener. Consequently, an interpretation of the present effects in terms of simple mirroring or imitation is unlikely, in part because both groups of sounds could be imitated and/or enacted by the listener (i.e. the actions were within their motor repertoire and all the sounds were themselves highly familiar and readily identified). Rather, our results are consistent with a class of mirror neurons described in monkeys as "logically related" (di Pellegrino et al., 1992) or "broadly congruent" (Gallese et al., 1996). These neurons are responsive to the observation of one action and to the execution of another that is the prototypical consequence of the observed action. This interpretation can also be considered alongside the more general proposition that activity of the mirror system reflects an ability to understand the intentions of others (Gallese et al., 1996; Kaplan and Iacoboni, 2006; Fabbri-Destro and Rizzolatti, 2008), such that enhanced responses to our group of context-related sounds would follow from recognition of the situation in which the heard sounds were produced and perhaps the preparation of possible responsive actions.

The timing of the present results suggests that differential activity within the mirror neuron system (at least with the sounds we presented) is subsequent to several other types of object-related processes including living vs. man-made categorization and repetition suppression (see Murray and Spierer, 2009 for review). Such being said, it is essential to note that the regions identified in this study are also involved during these earlier stages of auditory object processing (c.f. Figs. 4 and 5 in Murray et al., 2006). That is, it does not appear to be the case that regions of the mirror neuron system are only or selectively active over the 313–360 ms post-stimulus period identified in this study. Rather, it is important to note that the present results demonstrate that there is differential activity in these regions occurring first over this late time period. Nevertheless, it is still possible that any earlier differential effects within these regions are comparably more subtle than other phenomena (e.g. categorical discrimination) and thus might not have been observed with the paradigm/stimuli we used here. A similar timing of effects has been reported by Pizzamiglio et al. (2005) who used a repetition suppression paradigm to demonstrate modulation of the auditory mirror system. Although responses to different actions were not directly contrasted in their study, repetition effects for different action and non-action stimuli were observed at ~300 ms post-stimulus onset and were localized, using equivalent current dipoles, to the left superior temporal sulcus and left premotor cortex. In addition to their timing, our effects also indicate there to be relative synchronicity across the differential responses, such that no specific area is leading or driving the others, though this remains to be specifically investigated. Still, our electrical neuroimaging analyses do indicate that differential responses entailed a single and stable topographic modulation between conditions over the 313–360 ms interval, rather than multiple topographic differences. Given that robust, but not differential, responses were also observed within superior and middle temporal cortices that are implicated in auditory object discrimination, one possibility is that projections from these regions are driving and/or coordinating responses within premotor and (pre)frontal areas.

To further substantiate the timing of the present effects with respect to auditory object discrimination in general, we contrasted the AEP responses to context-related sounds as well as context-free sounds with AEP responses to a subset of sounds of living objects (i.e. animal vocalizations) from our prior work (Murray et al., 2006). This was done, in part, to assess whether the living vs. man-made modulation at ~70 ms post-stimulus onset is still observable when only half of the stimuli are contributing to the AEP (and thus when signal quality is reduced)[1]. This analysis was also done to assess whether there is sufficient sensitivity in the present data set to detect modulations prior to those we observed at ~300 ms post-stimulus onset. The results of these contrasts (i.e. millisecond-by-millisecond paired *t*-tests with a 15 ms temporal criterion, as above) for both types of actions vs. a subset of living sounds (i.e. animal vocalizations) are shown for each of the 61 scalp electrodes as a function of time peri-stimulus in Supplementary Fig. 2. In both cases, there were differential responses beginning ~70 ms post-stimulus onset. These results thus lend an additional line of support to our claim that differential activity to action sub-types begins after the initial categorical discrimination of sounds.

An important issue that our study unfortunately cannot directly resolve is the precise moment within each sound and by extension the precise latency within sensory-cognitive processing for recognizing action sub-types occurs. For example, Keysers et al. (2003) showed that single neuron responses from the ventral prefrontal cortex of monkeys could reliably discriminate between sounds of actions with 97% reliability, though the latency of their effects (and therefore the likelihood of whether this relates to sound recognition itself or instead to another post-recognition process) were not analyzed. Such being said, our prior research demonstrated there to be an upper limit of 100 ms for the initiation of brain processes mediating living vs. man-made categorization. This was achieved by comparing AEPs elicited by the same sounds when they served as distracters versus when they served as targets (and also served to control for any undetected differences in low-level acoustic features; see VanRullen and Thorpe, 2001 for an application of this approach with visual stimuli). Applying a similar approach to a paradigm wherein the classification of action sub-types was task-relevant would undoubtedly provide insights on the issue of when action sub-type information is treated. Related to this issue is the question of whether each sound we included within the group of context-related sounds would elicit the same type of modulations (and at the same latency). That is, some context-related sounds may be more or less effective in eliciting a responsive action. For example, a ringing telephone may be highly effective, whereas a church bell may be less so. Effects in the present analyses would likely have been dampened by any such variability. Still, and as mentioned above with respect to identifying somatotopic effects in single subjects, continued efforts in developing single-trial topographic analyses will be necessary to fully resolve any graded modulations within the mirror neuron system as a function of the efficacy of the sounds in eliciting responsive actions.

Finally, it is also worth mentioning that the present effects were obtained despite participants performing an orthogonal living vs. man-made discrimination task and despite our analyses being limited to distracter trials. In this way, the present effects suggest that the impact of associated actions can be evaluated implicitly, though it is also worth mentioning that all of the sounds were highly familiar and thus we cannot exclude a role of exposure/learning (see Lahav et al., 2007). On the one hand, such is in keeping with the conditions under which similar studies in non-human primates have been conducted (Kohler et al. 2002; Keysers et al., 2003; Cohen et al., 2006). On the other hand, such is useful for the extension of the present paradigm to clinical populations and/or populations like children in whom demanding tasks cannot be readily performed (Pazzaglia et al., 2008).

In conclusion, our findings support the conceptualization of auditory objects as integrally linked not only to their source and event (Griffiths and Warren, 2004), but also to their associated actions and contexts. Moreover, our application of electrical neuroimaging analyses highlights how these attributes are rapidly processed by widely distributed brain networks.

Acknowledgments

Cartool software (http://brainmapping.unige.ch/Cartool.htm) has been programmed by Denis Brunet, from the Functional Brain Mapping Laboratory, Geneva, Switzerland, and is supported by the Electroencephalography Brain Mapping Core of the Center for Biomedical Imaging

---

[1] We would note, however, that the number of trials contributing to each AEP in the present study (i.e. ~190 per condition per subject) is roughly equivalent to that in Hauk et al. (2006) as well as in Murray et al., (2008b) where AEP modulations were observed at latencies below 200 ms.

(http://www.cibm.ch) of Geneva and Lausanne. Christoph Michel and Jean-François Knebel provided additional analysis tools. Financial support was provided by the Swiss National Science Foundation (grants K-33K1_122518/1 to MDL, 3100A0_118419 to MMM, 3100A0_103895 to SC, and 320030_124897 to SC) and the Leenaards Foundation (2005 Prize for the Promotion of Scientific Research to MMM).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2009.06.041.

References

Altmann, C.F., Doehrmann, O., Kaiser, J., 2007. Selectivity for animal vocalizations in the human auditory cortex. Cereb. Cortex 17, 2601–2608.

Amedi, A., von Kriegstein, K., van Atteveldt, N.M., Beauchamp, M.S., Naumer, M.J., 2005. Functional imaging of human crossmodal identification and object recognition. Exp. Brain Res. 166, 559–571.

Aziz-Zadeh, L., Iacoboni, M., Zaidel, E., Wilson, S., Mazziotta, J., 2004. Left hemisphere motor facilitation in response to manual action sounds. Eur. J. Neurosci. 19, 2609–2612.

Barraclough, N.E., Xiao, D., Baker, C.I., Oram, M.W., Perrett, D.I., 2005. Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. J. Cogn. Neurosci. 17, 377–391.

Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P., Pike, B., 2000. Voice-selective areas in human auditory cortex. Nature 403, 309–312.

Bergerbest, D., Ghahremani, D.G., Gabrieli, J.D.E., 2004. Neural correlates of auditory repetition priming: reduced fMRI activation in the auditory cortex. J. Cogn. Neurosci. 16, 966–977.

Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., Rizzolatti, G., 2005. Listening to action-related sentences modulates the activity of the motor system: a combined TMS and behavioral study. Brain Res. Cogn. Brain Res. 24, 355–363.

Cohen, Y.E., Hauser, M.D., Russ, B.E., 2006. Spontaneous processing of abstract categorical information in the ventrolateral prefrontal cortex. Biol. Lett. 2, 261–265.

De Lucia, M., Michel, C.M., Clarke, S., Murray, M.M., 2007a. Single-subject EEG analysis based on topographic information. Int. J. Bioelectromagn. 9, 168–171.

De Lucia, M., Michel, C.M., Clarke, S., Murray, M.M., 2007b. Single-trial topographic analysis of human EEG: a new 'image' of event-related potentials. Proceedings of the IEEE/EMBS Region 8 International Conference on Information Technology Applications in Biomedicine, ITAB, art. no. 4407353, pp. 95–98.

De Lucia, M., Cocchi, L., Martuzzi, R., Meuli, R.A., Clarke, S., Murray, M.M., 2009. Perceptual and semantic contributions to repetition priming of environmental sounds. Cereb. Cortex in revision.

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., Rizzolatti, G., 1992. Understanding motor events: a neurophysiological study. Exp. Brain Res. 91, 176–180.

Doehrmann, O., Naumer, M.J., Volz, S., Kaiser, J., Altmann, C.F., 2008. Probing category selectivity for environmental sounds in the human auditory brain. Neuropsychologia 46, 2286–2776.

Fabbri-Destro, M., Rizzolatti, G., 2008. Mirror neurons and mirror systems in monkeys and humans. Physiology (Bethesda) 23, 171–179.

Fecteau, S., Armony, J.L., Joanette, Y., Belin, P., 2005. Sensitivity to voice in human prefrontal cortex. J. Neurophysiol. 94 (3), 2251–2254.

Frangos, J., Ritter, W., Friedman, D., 2005. Brain potentials to sexually suggestive whistles show meaning modulates the mismatch negativity. Neuroreport 16 (12), 1313–1317.

Galati, G., Committeri, G., Spitoni, G., Aprile, T., Di Russo, F., Pitzalis, S., Pizzamiglio, L., 2008. A selective representation of the meaning of actions in the auditory mirror system. NeuroImage 40, 1274–1286.

Gallese, V., Fadiga, L., Fogassi, L., Rizzolatti, G., 1996. Action recognition in the premotor cortex. Brain 119, 593–609.

Gazzola, V., Aziz-Zadeh, L., Keysers, C., 2006. Empathy and the somatotopic auditory mirror system in humans. Curr. Biol. 16, 1824–1829.

Gibson, J.J., 1979. The ecological Approach to Visual Perception. Houghton Mifflin, Boston, MA.

Grave de Peralta Menendez, R., Andino, S.G., Lantz, G., Michel, C.M., Landis, T., 2001. Noninvasive localization of electromagnetic epileptic activity: I. Method descriptions and simulations. Brain Topogr. 14, 131–137.

Grave de Peralta Menendez, R., Murray, M.M., Michel, C.M., Martuzzi, R., Gonzalez Andino, S.L., 2004. Electrical neuroimaging based on biophysical constraints. NeuroImage 21, 527–539.

Grèzes, J., Armony, J.L., Rowe, J., Passingham, R.E., 2003. Activations related to "mirror" and "canonical" neurones in the human brain: an fMRI study. NeuroImage 18, 928–937.

Griffiths, T.D., Warren, J.D., 2004. What is an auditory object? Nat. Rev., Neurosci. 5, 887–892.

Guthrie, D., Buchwald, J.S., 1991. Significance testing of difference potentials. Psychophysiology 28, 240–244.

Hauk, O., Shtyrov, Y., Pulvermüller, F., 2006. The sound of actions as reflected by mismatch negativity: rapid activation of cortical sensory-motor networks by sounds associated with finger and tongue movements. Eur. J. Neurosci. 23, 811–821.

Hauk, O., Shtyrov, Y., Pulvermüller, F., 2008. The time course of action and action–word comprehension in the human brain as revealed by neurophysiology. J. Physiol. Paris 102, 50–58.

Hauser, M.D., 1998. Functional referents and acoustic similarity: field playback experiments with rhesus monkeys. Anim. Behav. 55, 1647–1658.

Kaas, J.H., Hackett, T.A., 2000. Subdivisions of auditory cortex and processing streams in primates. Proc. Natl. Acad. Sci. U. S. A. 97, 11793–11799.

Kaplan, J.T., Iacoboni, M., 2006. Getting a grip on other minds: mirror neurons, intention understanding, and cognitive empathy. Soc. Neurosci. 1, 175–183.

Keysers, C., Kohler, E., Umilta, M.A., Nanetti, L., Fogassi, L., Gallese, V., 2003. Audiovisual mirror neurons and action recognition. Exp. Brain Res. 153, 628–636.

Kohler, E., Keysers, C., Umilta, M.A., Fogassi, L., Gallese, V., Rizzolatti, G., 2002. Hearing sounds, understanding actions: action representation in mirror neurons. Science 297, 846–848.

Lahav, A., Saltzman, E., Schlaug, G., 2007. Action representation of sound: audiomotor recognition network while listening to newly acquired actions. J. Neurosci. 27, 308–314.

Lebrun, N., Clochon, P., Etevenon, P., Baron, J.C., Eustache, F., 1998. Effect of environmental sound familiarity on dynamic neural activation/inhibition patterns: an ERD mapping study. NeuroImage 8, 79–92.

Lehmann, D., 1987. Principles of spatial analysis. In: Gevins, AS, Remond, A (Eds.), Handbook of Electroencephalography and clinical Neurophysiology, Vol.1: Methods of Analysis of Brain Electrical and Magnetic signals. Elsevier, Amsterdam, pp. 309–354.

Lehmann, D., Skrandies, W., 1980. Reference-free identification of components of checkerboard-evoked multichannel potential fields. Electroencephalogr. Clin. Neurophysiol. 48, 609–621.

Lewis, J.W., Brefczynski, J.A., Phinney, R.E., Janik, J.J., DeYoe, E.A., 2005. Distinct cortical pathways for processing tool versus animal sounds. J. Neurosci. 25, 5148–5158.

Manly, B.F., 1991. Randomization and Monte Carlo Methods in Biology. Chapman & Hall, London, UK.

Michel, C.M., Murray, M.M., Lantz, G., Gonzalez, S., Spinelli, L., Grave de Peralta, R., 2004. EEG source imaging. Clin. Neurophysiol. 115, 2195–2222.

Murray, M.M., Spierer, L., 2009. Auditory spatio-temporal brain dynamics and their consequences for multisensory interactions in humans. Hear. Res. doi: 10.1016/j.heares.2009.04.022 Published online 18 May.

Murray, M.M., Michel, C.M., Grave de Peralta, R., Ortigue, S., Brunet, D., Gonzalez Andino, S., Schnider, A., 2004. Rapid discrimination of visual and multisensory memories revealed by electrical neuroimaging. NeuroImage 21, 125–135.

Murray, M.M., Foxe, J.J., Wylie, G.R., 2005. The brain uses single-trial multisensory memories to discriminate without awareness. NeuroImage 27, 473–478.

Murray, M.M., Camen, C., Andino, S.L.G., Bovet, P., Clarke, S., 2006. Rapid brain discrimination of sounds of objects. J. Neurosci. 26, 1293–1302.

Murray, M.M., Brunet, D., Michel, C.M., 2008a. Topographic ERP analyses: a step-by-step tutorial review. Brain Topogr. 20, 249–264.

Murray, M.M., Camen, C., Spierer, L., Clarke, S., 2008b. Plasticity in representations of environmental sounds revealed by electrical neuroimaging. NeuroImage 39, 847–856.

Perrin, F., Pernier, J., Bertrand, O., Giard, M.H., Echallier, J.F., 1987. Mapping of scalp potentials by surface spline interpolation. Electroencephalogr. Clin. Neurophysiol. 66, 75–81.

Petkov, C.I., Kayser, C., Steudel, T., Whittingstall, K., Augath, M., Logothetis, N.K., 2008. A voice region in the monkey brain. Nat. Neurosci. 11, 367–374.

Picton, T.W., Alain, C., Woods, D.L., John, M.S., Scherg, M., Valdes-Sosa, P., Bosch-Bayard, J., Trujillo, N.J., 1999. Intracerebral sources of human auditory-evoked potentials. Audiol. Neuro-otol. 4, 64–79.

Pazzaglia, M., Pizzamiglio, L., Pes, E., Aglioti, S.M., 2008. The sound of actions in apraxia. Curr. Biol. 18, 1766–1772.

Pizzamiglio, L., Aprile, T., Spitoni, G., Pitzalis, S., Bates, E., D'Amico, S., Di Russo, F., 2005. Separate neural systems for processing action- or non-action-related sounds. NeuroImage 24, 852–861.

Plonsey, R., 1982. The nature of the sources of bioelectric and biomagnetic fields. Biophys. J. 39, 309–312.

Pulvermüller, F., 2005. Brain mechanisms linking language and action. Nat. Rev., Neurosci. 6, 576–582.

Rauschecker, J.P., 1998. Cortical control of the thalamus: top-down processing and plasticity. Nat. Neurosci. 1 (3), 179–180.

Rauschecker, J.P., Tian, B., 2000. Mechanisms and streams for processing of "what" and "where" in auditory cortex. Proc. Natl. Acad. Sci. U. S. A. 97 (22), 11800–11806.

Rizzolatti, G., Fadiga, L., Gallese, V., Fogassi, L., 1996. Premotor cortex and the recognition of motor actions. Brain Res. Cogn. Brain Res. 3, 131–141.

Romankski, L.M., 2007. Representation and integration of auditory and visual stimuli in the primate ventral lateral prefrontal cortex. Cereb. Cortex 17 (Suppl. 1), i61–9.

Romanski, L.M., Bates, J.F., Goldman-Rakic, P.S., 1999a. Auditory belt and parabelt projections to the prefrontal cortex in the rhesus monkey. J. Comp. Neurol. 403, 141–157.

Romanski, L.M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P.S., Rauschecker, J.P., 1999b. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. Nat. Neurosci. 2, 1131–1136.

Saygin, A.P., Dick, F., Wilson, S.M., Dronkers, N.F., Bates, E., 2003. Neural resources for processing language and environmental sounds: evidence from aphasia. Brain 126, 928–945.

Senkfor, A.J., 2008. Memory for pantomimed actions versus actions with real objects. Cortex 44, 820–833.

Seyfarth, R.M., Cheney, D.L., Marler, P., 1980. Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. Science 210, 801–803.

Sugihara, T., Diltz, M.D., Averbeck, B.B., Romanski, L.M., 2006. Integration of auditory and visual communication information in the primate ventrolateral prefrontal cortex. J. Neurosci. 26, 11138–11147.

Talairach, J., Tournoux, P., 1988. Co-planar Stereotaxic Atlas of the Human Brain. Thieme, New York.

Tettamanti, M., Buccino, G., Saccuman, M.C., Gallese, V., Danna, M., Scifo, P., Fazio, F., Rizzolatti, G., Cappa, S.F., Perani, D., 2005. Listening to action-related sentences activates fronto-parietal motor circuits. J. Cogn. Neurosci. 17, 273–281.

Tibshirani, R., Walther, G., Botstein, D., Brown, P., 2005. Cluster validation by prediction strength. J. Comput. Graph. Stat. 14, 511–528.

Toepel, U., Knebel, J.F., Hudry, J., le Coutre, J., Murray, M.M., 2009. The brain tracks the energetic value in food images. NeuroImage 44, 967–974.

VanRullen, R., Thorpe, S.J., 2001. The time course of visual processing: from early perception to decision-making. J. Cogn. Neurosci. 13 (4), 454–461.

von Kriegstein, K., Smith, D.R., Patterson, R.D., Ives, D.T., Griffiths, T.D., 2007. Neural representation of auditory size in the human voice and in sounds from other resonant sources. Curr. Biol. 17, 1123–1128.