UvA-DARE (Digital Academic Repository)

Positive emotions preferentially engage an auditory-motor "mirror" system

Warren, J.E.; Sauter, D.A.; Eisner, F.; Wiland, J.; Dresner, M.A.; Wise, R.J.S.; Rosen, S.; Scott, S.K.

Link to publication

Behavioral/Systems/Cognitive

Positive Emotions Preferentially Engage an Auditory–Motor "Mirror" System

Jane E. Warren,[1,3] **Disa A. Sauter,**[4] **Frank Eisner,**[6] **Jade Wiland,**[4] **M. Alexander Dresner,**[2] **Richard J. S. Wise,**[1,3] **Stuart Rosen,**[5] **and Sophie K. Scott**[1,6]

[1]Magnetic Research Council Clinical Sciences Centre and [2]Robert Steiner Magnetic Resonance Imaging Unit, Imaging Sciences Department, Hammersmith Hospital Campus, Imperial College London, London W121 0NN, United Kingdom, and [3]Division of Neuroscience and Mental Health, Departments of [4]Psychology and [5]Phonetics and Linguistics, and [6]Institute of Cognitive Neuroscience, University College London, London WC1E 6BT, United Kingdom

Social interaction relies on the ability to react to communication signals. Although cortical sensory–motor "mirror" networks are thought to play a key role in visual aspects of primate communication, evidence for a similar generic role for auditory–motor interaction in primate nonverbal communication is lacking. We demonstrate that a network of human premotor cortical regions activated during facial movement is also involved in auditory processing of affective nonverbal vocalizations. Within this auditory–motor mirror network, distinct functional subsystems respond preferentially to emotional valence and arousal properties of heard vocalizations. Positive emotional valence enhanced activation in a left posterior inferior frontal region involved in representation of prototypic actions, whereas increasing arousal enhanced activation in presupplementary motor area cortex involved in higher-order motor control. Our findings demonstrate that listening to nonverbal vocalizations can automatically engage preparation of responsive orofacial gestures, an effect that is greatest for positive-valence and high-arousal emotions. The automatic engagement of responsive orofacial gestures by emotional vocalizations suggests that auditory–motor interactions provide a fundamental mechanism for mirroring the emotional states of others during primate social behavior. Motor facilitation by positive vocal emotions suggests a basic neural mechanism for establishing cohesive bonds within primate social groups.

Key words: emotion; auditory; motor cortex; vocal expression; mirroring; fMRI

Introduction

The ability to generate appropriate behavioral responses to visual and auditory communication signals is fundamental to social intercourse in many animal species. Increasing evidence suggests that perceptual–motor interaction plays a key role in visual aspects of primate social behavior (Preston and de Waal, 2002; Adolphs, 2003). In nonhuman primates, so-called visuomotor mirror neurons, neurons that discharge during the observation or execution of a particular movement (Gallese et al., 1996), have been implicated in the processing of communicative gestures (Ferrari et al., 2003). Although human visuomotor mirror responses have been demonstrated from neuronal recordings only rarely (Krolak-Salmon et al., 2006), functional neuroimaging studies have demonstrated cortical-level "mirror" responses to the observation and generation of facial expressions of emotion (Carr et al., 2003; Leslie et al., 2004; Hennenlotter et al., 2005). In the auditory domain, auditory–motor mirror neurons, responsive to observing an action and hearing the sound of the same

action, have been identified in nonhuman primates (Kohler et al., 2002; Keysers et al., 2003), and there is evidence of interplay between auditory and motor systems within the specialized domain of human speech processing (Fadiga et al., 2002; Watkins et al., 2003; Watkins and Paus, 2004; Wilson et al., 2004), including the processing of affective prosody (Hietanen et al., 1998). However, a generic role for auditory–motor interaction in the communication of nonverbal information, such as emotion, is yet to be established in primates.

In this functional magnetic resonance imaging (fMRI) study, we investigated cortical regions responsive to both the perception of human vocalizations and the voluntary generation of facial expressions. In four auditory–perceptual conditions, subjects listened passively, without overt motor response, to nonverbal emotional vocalizations conveying two positive-valence emotions, amusement and triumph, and two negative-valence emotions, fear and disgust (Ekman, 1992, 2003). Use of nonverbal, rather than verbal, vocalizations optimized recognizability of emotional content (Scott et al., 1997) and avoided confounds of phonological and verbal content (Hietanen et al., 1998; Fadiga et al., 2002; Watkins et al., 2003; Hauk et al., 2004; Watkins and Paus, 2004; Wilson et al., 2004). In a facial movement condition, subjects performed voluntary smiling movements in the absence of auditory input. We hypothesized that cortical regions showing combined auditory–perceptual and motor responses would be located within premotor and motor cortical regions.

Furthermore, because emotional valence and arousal are widely considered to be critical factors in models of the processing and representation of emotional signals (Russell, 1980), we investigated the effect of these stimulus properties on hemodynamic responses within cortical regions demonstrating auditory–motor mirror responses.

Materials and Methods

Auditory stimuli. Two positive-valence emotions, amusement and triumph, and two negative-valence emotions, fear and disgust, were selected for investigation (Ekman, 1992, 2003). Nonverbal vocal expressions of triumph, amusement, fear, and disgust were collected from two male and two female native British English speakers. None of the speakers was a trained actor. Speakers were presented with written stories describing scenarios relating to each emotion label and asked to generate appropriate vocal responses. Speakers were instructed not to produce verbal responses (e.g., "yuck" or "yippee") but otherwise were not given explicit guidance as to the precise sort of sounds they should generate; for example, they were not given examples to mimic. Vocal responses were recorded in an anechoic chamber and then digitized.

Emotion content, valence, and arousal properties of vocal sounds were assessed by normal subjects who did not subsequently take part in the fMRI experiment. To assess the recognizability of the emotion expressed by each sound, a group of 20 subjects (mean ± SEM age, 21.4 ± 0.7 years) performed a forced-choice categorization task. Subjects were provided with a list of verbal labels (the names of each emotion category) and asked to select the label that best described each sound. The 20 sounds from each emotion category that were most consistently labeled with the correct emotion were chosen for the fMRI stimulus set; the final stimulus set for each emotion category contained several sounds from each speaker.

Sounds in the fMRI stimulus set were rated for their emotional valence and arousal properties by an additional group of 20 normal subjects (mean age, 24.5 ±.2.2 years) To minimize response bias associated with inclusion of zero or negative numbers in rating scales, each rating task consisted of judging the extent to which each sound expressed the given dimension on a seven-step scale, with 1 denoting the minimum and 7 the maximum. In the case of arousal, a score of 1 indicated no arousal and a score of 7 denoted maximum arousal. In the case of valence, a score of 1 denoted a strongly negative emotion and a score of 7 denoted a strongly positive emotion. After data collection, arousal and valence scores for all stimuli from a given emotion category were averaged across subjects, resulting in an overall mean arousal and valence rating for each of the four categories. For ease of interpretation, mean arousal and valence scores were transformed from the original 1–7 rating scales; mean arousal ratings were transformed to a scale ranging from 0 (no arousal) to 6 (maximum arousal), and mean valence ratings were transformed to a scale ranging from −3 (maximum negative valence) to +3 (maximum positive valence). Transformed mean valence and arousal ratings for each emotion category are shown in Figure 1 and supplemental Table 1 (available at www.jneurosci.org as supplemental material). To ensure that low-valence and low-arousal stimuli were not systematically rated more variably than high-valence and high-arousal stimuli, we investigated correlations between the mean and SEM values of rating scores within each emotion category. Across emotion categories, there were no significant correlations between the magnitudes of the SEM and the mean of the arousal ratings (Spearman's $\rho = 0.8$; $p = 0.2$) or between the magnitude of the SEM and the absolute (unsigned) magnitude of the valence ratings (Spearman's $\rho = -0.6$; $p = 0.4$), indicating that there was no significant relationship between rating variability and score magnitude for either valence or arousal.

Stimuli for the baseline condition were created by spectral rotation (~2 kHz) of a selection of the original vocalizations from each category, resulting in a set of 15 rotated stimuli. Spectral rotation was performed using a method described previously (Blesser, 1972; Scott et al., 2000). Stimuli were first equalized with a filter (essentially high-pass) that gave the rotated signal approximately the same long-term spectrum as the original. The equalized signal was then amplitude-modulated by a sinusoid at 4 kHz, followed by low-pass filtering at 3.8 kHz. This acoustic
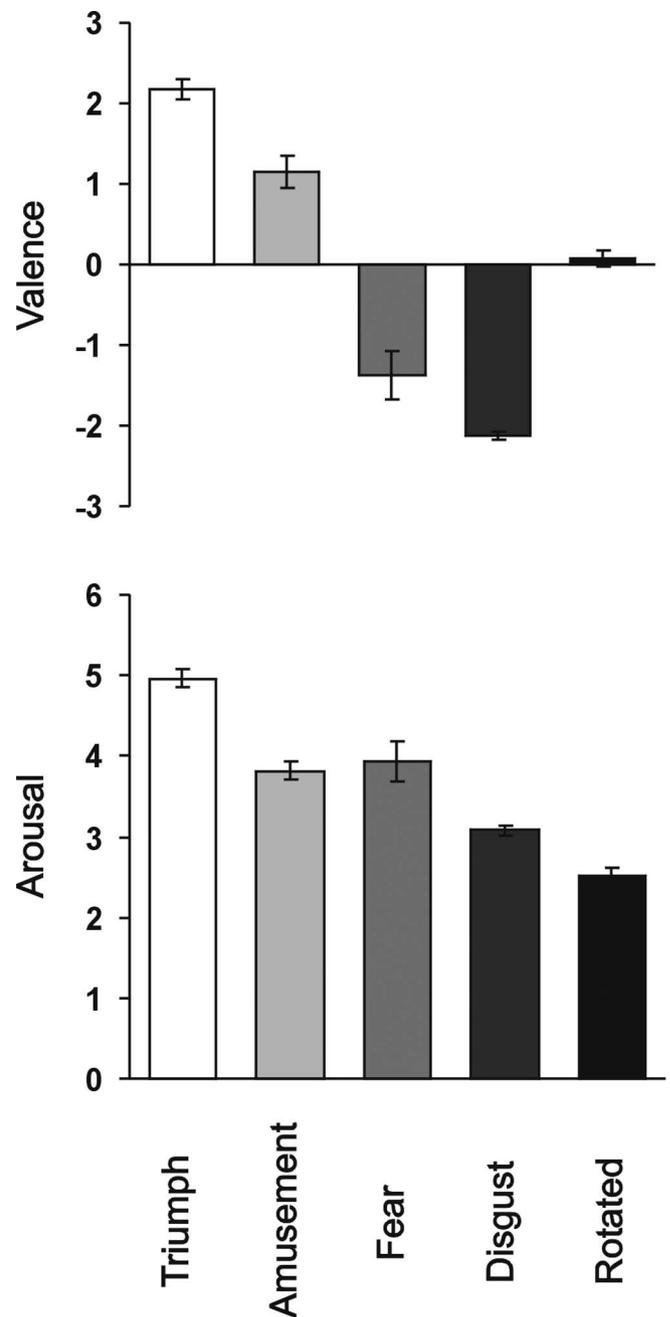

Figure 1. Mean valence and arousal ratings for emotional vocalizations and spectrally rotated stimuli used in the fMRI and EMG experiments. Arousal scales range from 0 (no arousal) to 6 (maximum arousal); valence scales ranged between +3 (maximum positive valence) and −3 (maximum negative valence). Error bars indicate ±SEM.

manipulation produced unintelligible sounds that lacked the human vocal quality of the original stimuli but maintained a comparable level of acoustic complexity (Blesser, 1972) and was used in preference to emotionally neutral vocalizations because rotated vocalizations are unpronounceable and therefore should not map easily onto articulatory motor representations. Ratings of emotion content, valence, and arousal were obtained for the rotated sounds from a group of five normal subjects (mean age, 30.0 ±. 1.0 years). So as not to bias results, these subjects were not also asked to rate the sounds from the emotion categories. As with the emotional vocalizations, recognizability of the emotion expressed by each rotated sound was assessed using a forced-choice categorization task; however, in this case, the list of verbal labels subjects selected from contained the option "none of the above" in addition to the name of each emotion category. The label most commonly applied to the rotated

sounds was "none of the above" (57.3% of categorizations) compared with "triumph" (9.3%), "amusement" (14.7%), "disgust" (10.7%), and "fear" (8%): thus, rotated stimuli did not consistently convey any of the emotion categories under investigation. Emotional valence and arousal ratings were obtained for the rotated stimuli by the same methods used for the emotional vocalizations. Transformed mean ± SEM valence and arousal ratings for rotated stimuli are shown in Figure 1 and supplemental Table 1 (available at www.jneurosci.org as supplemental material). Rotated stimuli were rated as possessing a neutral emotional valence and had lower arousal ratings than the emotional vocalizations, confirming the validity of these sounds as an appropriate baseline for comparison with the affective vocal stimuli.

All auditory stimuli were scaled to the same peak amplitude and edited to a length of 2.4 s. Stimulus examples are provided in supplemental Audio Files 1–5 (available at www.jneurosci.org as supplemental material).

Subjects and fMRI procedures. Twenty right-handed subjects with no history of significant neurological illness or hearing impairment participated in the fMRI study (12 females; mean age, 32.9 ± 2.4 years) after giving informed consent. All subjects had normal or corrected-to-normal vision. Seventeen subjects were native English speakers; the remaining three subjects were highly proficient non-native speakers who had acquired English early in childhood. None of the subjects had participated in behavioral testing of the stimuli.

MRI data were obtained on a Philips (Best, The Netherlands) Intera 3.0 Tesla MRI scanner using Nova Dual gradients, a phased array head coil, and sensitivity encoding (SENSE) with an undersampling factor of 2. Functional MRI images were obtained using a T2*-weighted gradient-echo echoplanar imaging (EPI) sequence with whole-brain coverage (repetition time, 10.0 s; acquisition time, 2.0 s, echo time, 30 ms; flip angle, 90°). Thirty-two axial slices with a slice thickness of 3.25 mm and an interslice gap of 0.75 mm were acquired in ascending order (resolution, 2.19 × 2.19 × 4.0 mm; field of view, 280 × 224 × 128 mm). Quadratic shim gradients were used to correct for magnetic field inhomogeneities within the anatomy of interest. T1-weighted whole-brain structural images were also obtained in all subjects.

Functional data were acquired using a sparse sampling protocol (Hall et al., 1999), in which stimuli were presented during the 8 s intervals between image acquisition periods, to avoid interference from scanner noise during listening trials and motion-related artifact during facial movement trials. Stimuli were presented using E-Prime software (Psychology Software Tools, Pittsburgh, PA) run on an IFIS-SA system (Invivo Corporation, Orlando, FL). The IFIS-SA package uses pneumatic sound delivery to headphones contained within ear defenders. All subjects used the same sound volume level and wore cannulated earplugs that provided additional shielding against scanner noise but allowed transmission of headphone output into the external auditory canal. For each of the four emotion conditions and the baseline condition, subjects listened, without overt motor response, to three randomly selected tokens from the appropriate stimulus set, presented at 2.5 s intervals during the 8 s interscan silent period, while the instruction "LISTEN" was presented on a video monitor. During each trial of the motor condition, subjects were cued to initiate a voluntary smiling movement at the beginning of the interscan silent interval by the appearance of the instruction "SMILE" on the video monitor. Two brief periods of relaxation, followed by immediate resumption of smiling movements, were cued 2.5 and 5.0 s after trial onset by the serial appearance of two exclamation marks at the end of the "SMILE" instruction on the video monitor. Cessation of each trial was cued by the replacement of the written instruction with a fixation cross at the end of the interscan silent period. Before the scanning session, subjects were instructed to listen attentively to auditory stimuli, regardless of whether sounds were meaningful or not, and were trained in performance of the motor task. Data were acquired in two consecutive scanning runs, each involving 96 whole-brain images. Each run used a different randomized trial order, with run order randomized across subjects. Consecutive trials were always from different conditions.

fMRI data analysis. Image preprocessing and analysis was performed with the SPM2 software package (http://www.fil.ion.ucl.ac.uk/spm). Image preprocessing involved realignment of EPI images to remove the effects of head movement between scans, coregistration of the T1-weighted structural image to the mean EPI image, normalization of EPI images into Montreal Neurological Institute (MNI) standard stereotactic space using normalization parameters derived from the coregistered T1-weighted image, and smoothing of normalized EPI images using an 8 mm full-width at half-maximum Gaussian filter. Analysis of imaging data were conducted using a random-effects model. At the first level, individual design matrices were constructed for each subject, modeling each of the six experimental conditions in two scanning runs and including movement parameters derived from the realignment step as nuisance variables. Blood oxygen level-dependent responses were modeled using a finite impulse response model of length 2.0 s (i.e., equal to the acquisition time) and order 1. Contrast images for each subject for each of the contrasts of interest were created at the first level and entered into second-level analyses. Preliminary random-effects analyses demonstrated no significant differences in activation between male and female subjects for any of the conditions of interest, so all subsequent analyses modeled subjects as a single group.

Brain regions showing significant modulation of auditory–perceptual activation on the basis of emotional category were identified using a second-level voxelwise ANOVA; separate contrast images for each emotion class versus auditory baseline were entered into a one-way repeated-measures ANOVA model (using a nonsphericity correction for repeated measures). An *F* contrast of effects of interest identified brain regions in which the magnitude of perceptual responses to heard vocalizations was significantly modulated by emotional category. Regions demonstrating significant activation during the facial movement task (vs the baseline condition) were identified using a second-level one-sample voxelwise *t* test model. To identify brain regions that showed both auditory–perceptual modulatory effects and significant activation during facial movement, ANOVA *F* contrast was masked inclusively by the motor task contrast, so that the resulting statistical parametric map showed only those voxels demonstrating a significant response to both contrasts. Second-level voxelwise linear regression (correlation) analyses, using mean valence and arousal ratings of vocal stimuli as covariates of interest, were used to investigate positive and negative relationships between emotional valence and arousal scores and the measured hemodynamic responses for each category of vocal emotion. For all contrasts of interest, the threshold for significance was set at $p < 0.05$ adjusted for multiple comparisons using the false discovery rate (FDR) correction (Genovese et al., 2002) (or the equivalent *T* value in the case of masking contrasts) with a cluster extent threshold of 10 voxels.

EMG study. To determine whether subjects were involuntarily generating overt facial movements in response to hearing emotional vocalizations, we conducted an additional experiment to record facial EMG responses. Because we were unable to record EMG responses during scanning, this experiment was conducted outside the MRI scanner, using a group of five normal subjects (four males; mean age, 30.0 ± 1.0 years) who did not take part in the fMRI study and who had no previous exposure to any of the stimuli. Subjects sat in front of a video screen in an isolation chamber and were presented with stimuli from one run of the fMRI experiment. Six experimental conditions (triumph vocalizations, amusement vocalizations, fear vocalizations, disgust vocalizations, rotated vocal sounds, and voluntary facial movement) were presented using the same stimuli, visual cues, timings, and stimulus order as the fMRI experiment. As in the fMRI experiment, each trial began with a 2.5 s period in which subjects simply viewed a fixation cross in the center of the video screen. However, whereas in the fMRI experiment this fixation interval coincided with the 2 s scan acquisition period, in the EMG experiment this fixation period was silent. Auditory stimuli were presented via headphones. An additional "null" condition was added to provide a resting baseline. In the null condition, subjects viewed a fixation cross in the center of the video screen for the entire 10 s duration of the trial, without hearing any auditory stimuli and without any instruction to produce a movement: 16 trials of this condition were added pseudorandomly to the trial order. Thus, subjects in the EMG study essentially performed one run of the fMRI experiment, with the addition of a rest condition.

During the experiment, responses were recorded via bipolar surface electrodes from two sites on the right side of the face. The lower facial electrode was situated over the zygomaticus major muscle, whereas the upper facial electrode was situated over the corrugator supercilii muscle (Dimberg et al., 2000). Subjects were unaware of the purpose of the experiment before participation: all were explicitly told that the facial electrodes were intended to record sweat responses during the experiment. Subjects were informed of the true purpose of the study during debriefing after completion of the EMG recordings: no subject deduced the true purpose of the electrodes before this point. A chin rest, adjusted to a comfortable height for each individual subject, was used to support the head and minimize unwanted movement. EMG responses were recorded continuously for the entire duration of the experiment, digitized, and then stored at a sampling frequency of 100 Hz. Pulsed signals conducted from the stimulus delivery computer indicated the onset times of stimulus presentation within each trial; onset markers were stored within the same files as the EMG data.

Data from the upper and lower facial electrodes were analyzed separately using the same method. For the purpose of analysis, each 10 s trial was divided into four 2.5 s quarter-trials, and each quarter-trial was divided into five 500 ms time bins. EMG amplitude was measured within each quarter-trial using the root mean square (RMS) of the recorded responses (expressed as microvolts). The first quarter-trial always corresponded to a silent fixation period. The second, third, and fourth quarter-trials corresponded to stimulus presentations: written instructions to smile, sound stimuli accompanied by the written instruction to listen, or, in the null trials, a fixation cross alone. In each trial, the RMS values from all five time bins in the fixation period were averaged to provide a trial-specific measure of baseline EMG activity. In each trial, the difference between the RMS value within a given time bin and the baseline (fixation) RMS value was calculated for the second, third, and fourth quarter-trials; these RMS differences were averaged across a given trial to give a trial-specific measure of stimulus-related EMG responses. Initially, we examined EMG responses to the motor condition in all five time bins and found that EMG responses were maximal within the third time bin (1000–1499 ms from stimulus onset): therefore, all subsequent analyses were confined to data from this 1000–1499 ms time bin. Because we found that EMG activity carried over into the subsequent fixation period after trials of the voluntary facial movement condition, precluding the use of these fixation periods to measure baseline (resting) EMG activity, trials immediately after the motor trials were excluded from additional analysis. For each subject, the remaining data were collapsed across trials to give a measure of mean EMG amplitude for each of the experimental conditions. At the group level, mean EMG responses from each of the auditory conditions were then compared directly with responses from the motor and null conditions using paired-sample t tests.

Results
fMRI study
Initially, we determined all regions in which perceptual responses to emotional vocalizations were significantly modulated by emotional category, and all regions significantly activated during per-
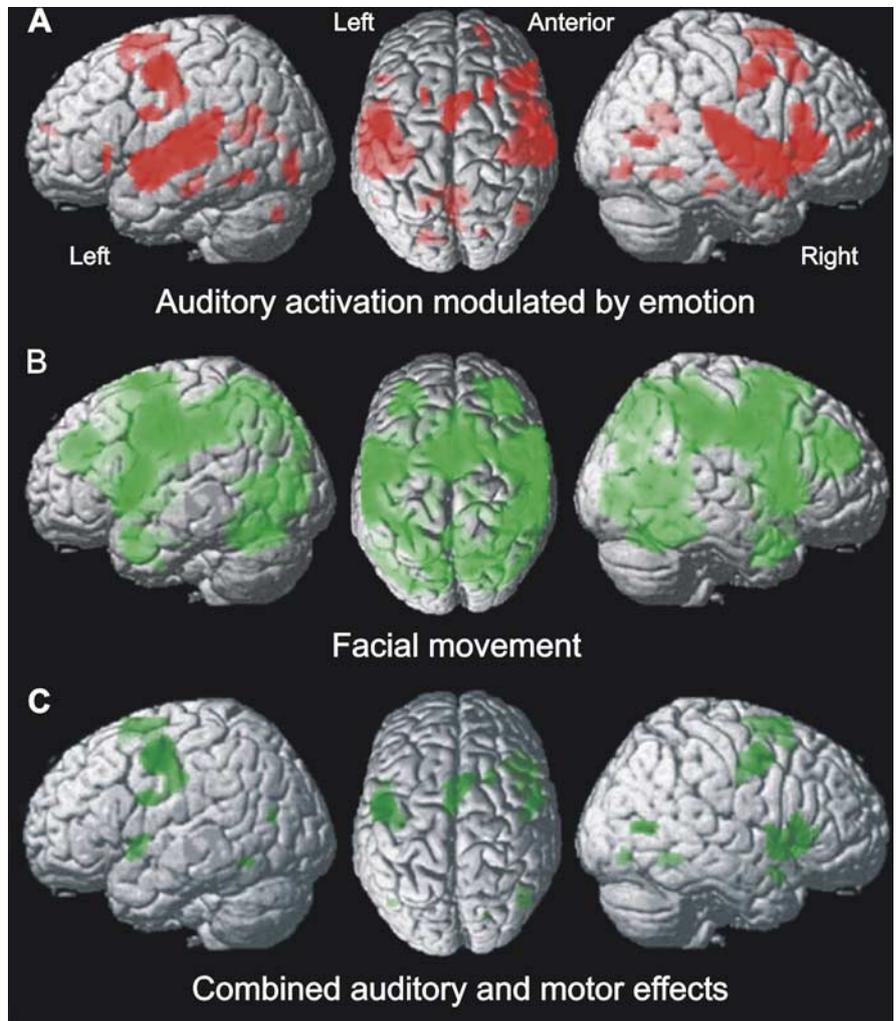

Figure 2. Brain regions demonstrating auditory–motor mirror responses. *A* shows regions (red) displaying a significant modulatory effect of emotion category on perceptual activation (*F* contrast, one-way repeated-measures ANOVA). *B* shows regions (light green) displaying significant activation during voluntary facial movements (motor > baseline). *C*, *A* masked inclusively in *B* shows regions (dark green) displaying both a significant modulatory effect of emotion category on perceptual activation and significant activation during voluntary facial movements. Voxel-level statistical thresholds for the main and masking contrasts were set at $p < 0.05$, FDR-corrected, with a 10-voxel cluster extent threshold. Statistical parametric maps are displayed on left lateral, superior, and right lateral projections of a canonical brain surface in standard MNI stereotactic space, with color intensity indicating distance from the cortical surface.

formance of the facial movement task (Fig. 2) (supplemental Table 2, available at www.jneurosci.org as supplemental material). Regions in which perceptual responses to emotional vocalizations were significantly modulated by emotion category included superior and inferobasal temporal cortex, precentral and prefrontal cortex, and limbic and mesial temporal cortex in both hemispheres. Regions significantly activated during performance of voluntary facial movements included motor, premotor thalamic, and insular regions in both hemispheres. Brain regions that showed both auditory–perceptual modulatory effects and significant activation during facial movement were identified by inclusive masking of the emotional–modulation contrast (ANOVA *F* contrast) with the motor task contrast, so that the resulting statistical parametric map showed only those voxels demonstrating a significant response to both contrasts (Fig. 2) (supplemental Table 3, available at www.jneurosci.org as supplemental material). Although regions demonstrating combined auditory–perceptual and motor effects were also identified within temporal, prefrontal, and insular cortex, only results pertaining
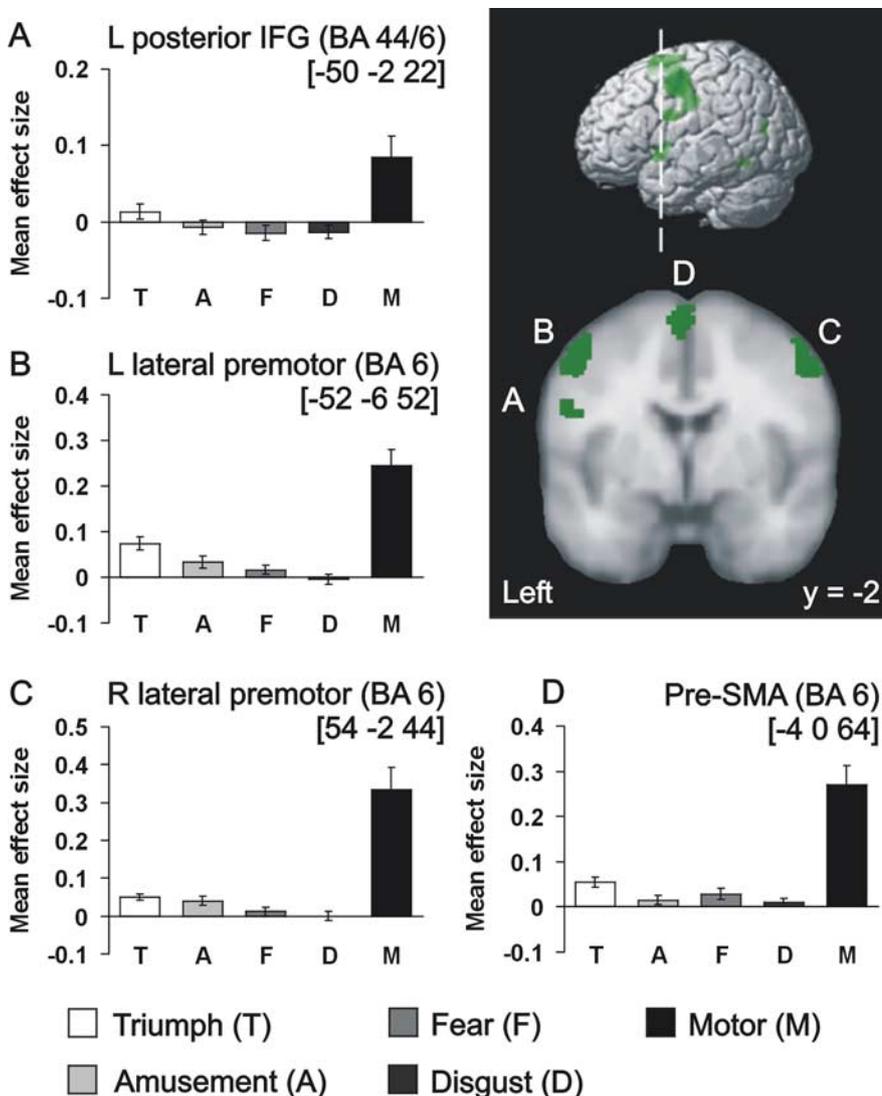
Figure 3. Auditory and motor responses within premotor regions showing demonstrating auditory–motor mirror responses. The top right shows left lateral premotor, right lateral premotor, left posterior IFG, and pre-SMA activation clusters from the masked ANOVA F contrast presented in the bottom section of Figure 2, displayed on a coronal section from a canonical averaged brain image in MNI stereotactic space. Statistical thresholds are the same as in Figure 2. **A–D** demonstrate hemodynamic responses (mean effect sizes, arbitrary units) for each category of emotional vocalization in the four premotor activation clusters. For the purposes of graphical data display, hemodynamic responses for each category of emotional vocalization (mean effect sizes) were extracted in each subject from 10 mm spherical regions of interest centered on the most significant activation peaks in each cluster using the MarsBaR software toolbox within SPM2 (Brett et al., 2002) and then averaged across the group. Brodmann area (BA) locations of each peak are shown in brackets. Coordinates (in millimeters) give the location of the peaks in MNI space. Error bars indicate ±SEM. L, Left; R, right.

tigating significant correlations between mean valence and arousal scores and the measured hemodynamic responses to each category of vocal emotion (Fig. 4). For these analyses, the volume of interest was restricted to regions showing combined perceptual and motor responses by application of a small-volume correction, by application of an FDR correction for multiple comparisons ($p < 0.05$) across a search volume defined as the suprathreshold voxels from the inclusively masked ANOVA F contrast described above (after initial thresholding to ensure inclusion of all possible voxels in the calculation of significance measures). We identified three distinct response patterns within premotor cortex (Fig. 5) (supplemental Table 4, available at www.jneurosci.org as supplemental material). Hemodynamic responses were positively correlated with increasing arousal scores alone in pre-SMA. Responses were positively correlated with increasing positive valence alone in posterior left IFG. Finally, responses were positively correlated with both arousal and valence in left and right lateral premotor regions. No regions showed a significant negative correlation with valence or arousal; specifically, no negative correlations were observed in frontal regions.

During debriefing after scanning sessions, a majority of subjects reported that they initially experienced an involuntary urge to make a facial expression while listening to some emotional categories (most commonly an urge to smile in response to laughter stimuli), but this sensation did not persist beyond the first few trials of the first scanning run in any subject, and no subject reported either overt facial movement or overt vocalization in response to sound stimuli. To formally investigate whether emotional vocalizations produced greater motor activation as a function of the novelty of the stimuli, we conducted random-effects analyses comparing activation related to processing of vocalizations in the first and second experimental runs. These analyses demonstrated no significant differences in any brain regions, indicating that the initial involuntary urges to make responsive facial movements did not significantly influence the magnitude of activation in motor regions.

EMG study

The results of the EMG study are shown in Figure 6 and supplemental Table 5 (available at www.jneurosci.org as supplemental material). During debriefing, several subjects reported that they initially experienced an involuntary urge to make a facial expression while listening to some emotional categories (as with the fMRI subjects, most commonly an urge to smile in response to laughter stimuli), but this sensation did not persist beyond the first few trials in any subject, and no subject reported overt facial

to premotor and motor regions will be discussed further. In the frontal lobes, combined auditory and motor effects were identified in several premotor regions: in the left and right lateral premotor cortices, at the posterior border of the left inferior frontal gyrus (IFG), and in mesial premotor cortex (Fig. 3). Lateral premotor activation extended posteriorly into primary motor cortex. Activation peaks in mesial premotor cortex fell within the anatomical boundaries of the presupplementary motor area (pre-SMA) (Geyer et al., 2000). In each of these regions, responses during facial movement were three to six times greater than responses elicited during perception of emotional vocalizations (Fig. 3).

We then examined the influence of valence and arousal on activation within this auditory–motor mirror network, by inves-

movement. Vocal output was monitored during the experiment: no subject produced overt vocalizations in response to sound stimuli.

As expected, mean EMG amplitude recorded at the lower facial electrode was significantly greater during voluntary facial smiling than during null (resting) trials ($p < 0.03$). Mean EMG amplitude recorded at the lower facial electrode was also significantly greater during the motor condition than when listening to any category of emotional vocalization or to the rotated sound stimuli ($p < 0.03$ in each case). Mean EMG responses recorded at the lower facial electrode when listening to sound stimuli did not differ significantly from responses during null trials (triumph, $p > 0.2$; amusement, $p > 0.2$; fear, $p > 0.4$; disgust, $p > 0.5$; rotated sounds, $p > 0.7$). No significant differences between conditions were identified in EMG responses recorded from the upper facial electrode.

Discussion

This fMRI study demonstrates that passive perception of nonverbal emotional vocalizations automatically modulates neural activity in a network of premotor cortical regions involved in the control of facial movement. Moreover, the degree of activation of specific regions within these premotor regions is determined by emotional valence and arousal properties of affective vocal stimuli. The complementary EMG data clearly demonstrate that these premotor cortical responses do not simply reflect the generation of overt facial movements in response to emotional vocalizations: thus, our findings suggest that listening vocal expressions of positive or arousing emotions automatically engages preparation for responsive orofacial gestures.

Our results demonstrate the existence of distinct functional subsystems within this auditory–motor mirror network that correspond broadly to known function- and connectivity-based divisions within the primate premotor cortex (Rizzolatti and Luppino, 2001). Premotor responses associated with positive emotional valence were identified at the posterior border of the left IFG. Posterior IFG is the putative human homolog of the nonhuman primate mirror neuron area F5 (Rizzolatti and Arbib, 1998). In addition to neurons responsive to visual perception of hand and orofacial actions (Gallese et al., 1996),
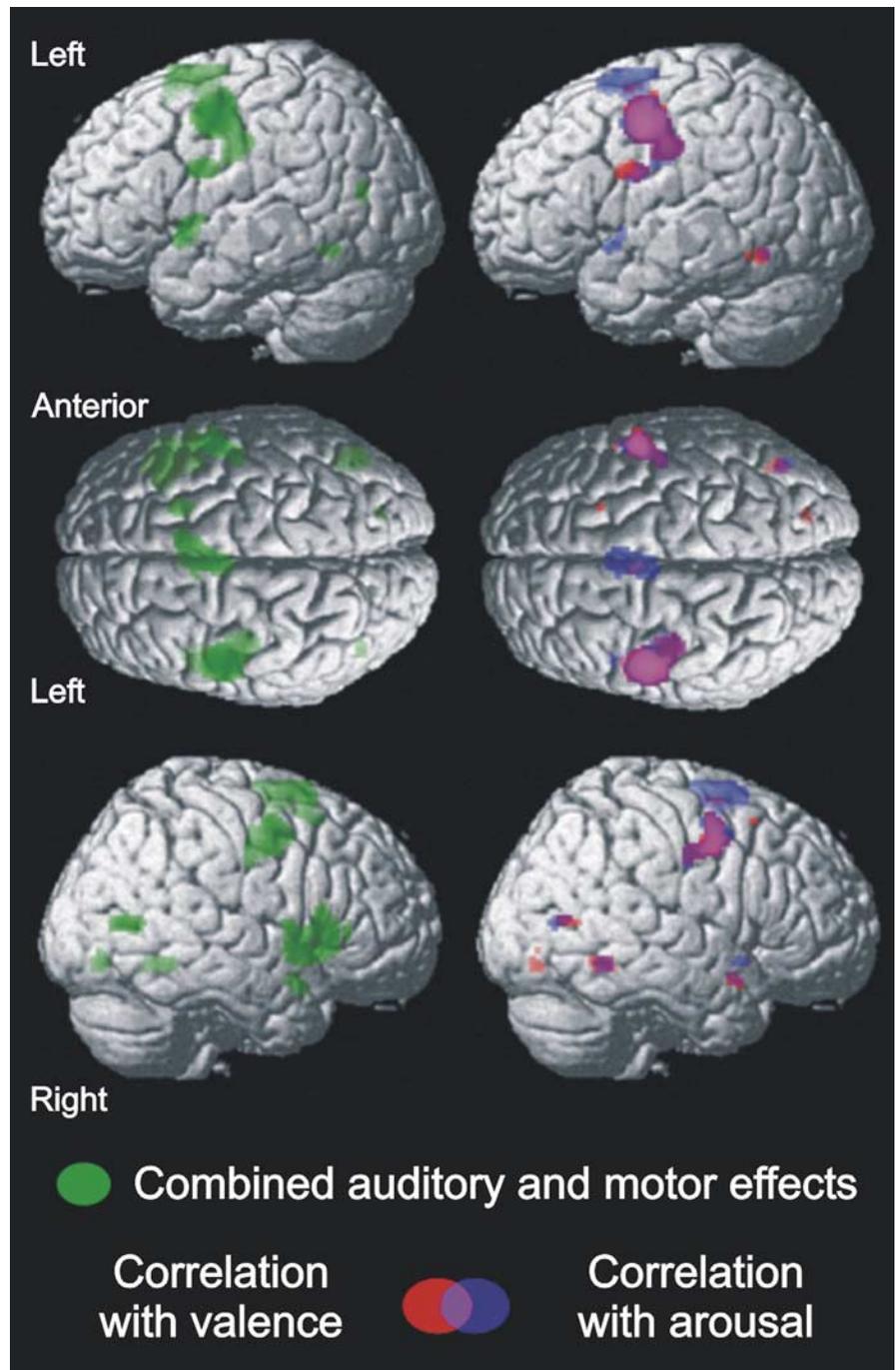

Figure 4. Correlations with emotional valence and arousal in brain regions demonstrating auditory–motor mirror responses. Left, Regions (green) displaying both a significant modulatory effect of emotion category on perceptual activation and significant activation during voluntary facial movements (F contrast, one-way repeated-measures ANOVA, masked inclusively with the contrast of the facial movement condition over baseline), as shown in Figure 2. Statistical thresholds are the same as in Figure 2. Right, Regions demonstrating a significant positive correlation between hemodynamic responses and emotional valence (red), emotional arousal (blue), or both (purple). Statistical thresholds for these contrasts were set at a voxel-level threshold of $p < 0.05$, FDR-corrected across a search volume defined as the suprathreshold voxels from the inclusively-masked ANOVA F contrast (top row images, green). For display purposes, statistical parametric maps of these correlations have been masked to display significant voxels within the search volume only. Statistical parametric maps are displayed on left lateral, superior, and right lateral projections of a canonical brain surface in standard MNI stereotactic space, with color intensity indicating distance from the cortical surface.

including communicative orofacial gestures (Ferrari et al., 2003), a proportion of area F5 neurons also respond when hearing action-related sounds (Kohler et al., 2002; Keysers et al., 2003). Neurons in primate area F5 are thought to encode motor proto-

types: representations of potential actions congruent with a particular stimulus that can be activated either exogenously via sensory projections or endogenously (Rizzolatti and Luppino, 2001). The positive emotions investigated in this study, amusement
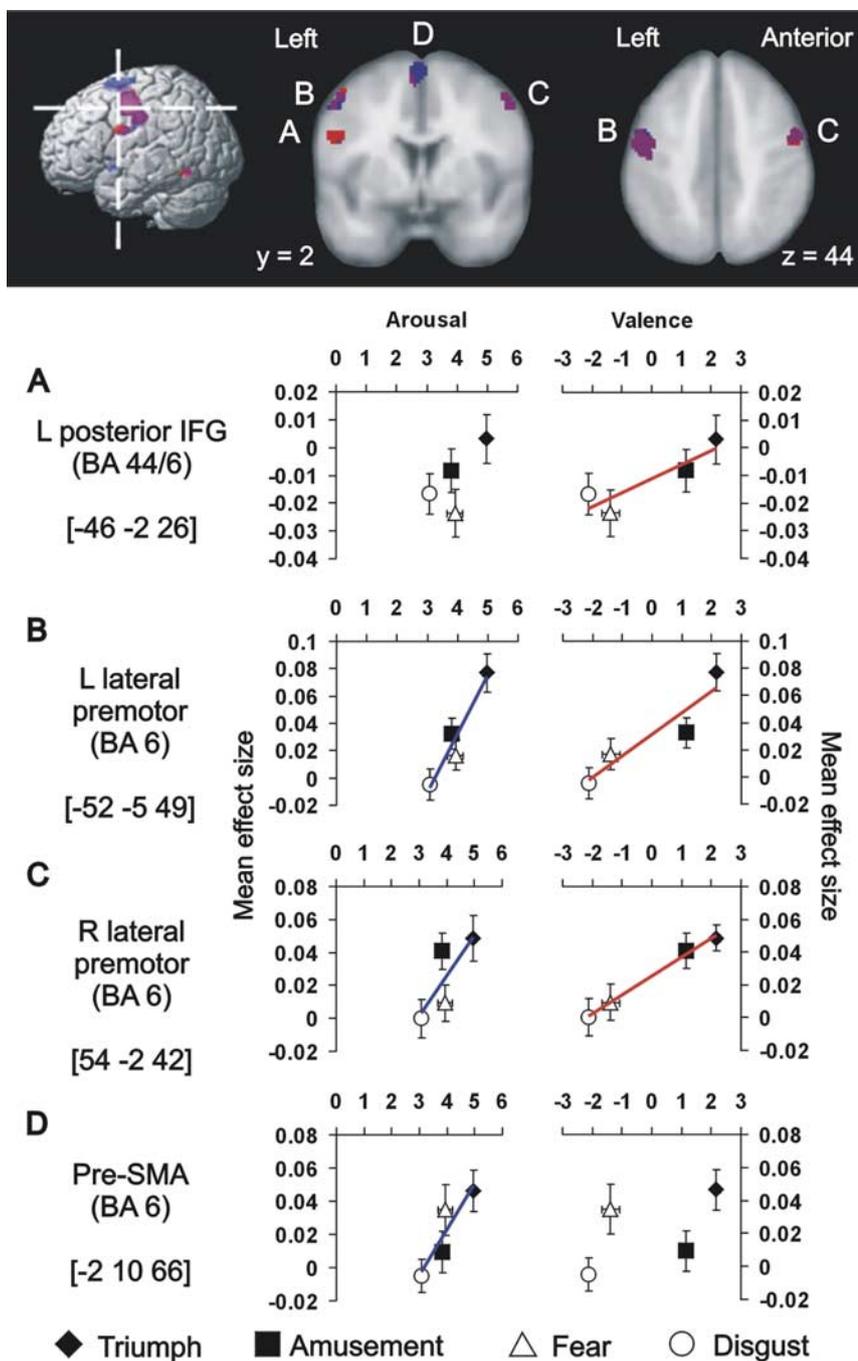
Figure 5. Auditory–perceptual responses to emotional vocalizations within premotor cortex correlate positively with emotional valence and arousal. The top shows left posterior IFG and pre-SMA (coronal section) and left and right lateral premotor (axial section) clusters from the statistical parametric maps presented in the bottom row of Figure 4, displayed on coronal and axial slices from a canonical averaged brain image in MNI stereotactic space. Statistical thresholds are the same as in Figure 4. *A–D* demonstrate the relationship between hemodynamic responses (mean effect sizes, arbitrary units) in each of these premotor regions (for details, see Fig. 3 legend) and measures of mean emotional valence and arousal for triumph (black diamonds), amusement (black squares), fear (white triangles), and disgust (white circles) vocal stimuli. Significant correlations between hemodynamic responses and arousal ratings are indicated by blue lines of best fit; significant correlations between hemodynamic responses and valence ratings are indicated by red lines of best fit. No lines of best fit are shown for nonsignificant correlations. Brodmann area (BA) locations of each peak are shown in brackets. Coordinates (in millimeters) give the location of the peaks in MNI space. Arousal and valence scales are the same as in Figure 1. Error bars indicate ±SEM. L, Left; R, right.

triumph, are typically encountered in group situations characterized by mutual and interactive expressions of emotion. Autistic children demonstrate reduced activation in posterior IFG during observation and imitation of emotional facial expressions that

correlates with measures of social dysfunction (Dapretto et al., 2006). Our findings suggest that vocal communications conveying positive emotions automatically activate motor representations encoded in posterior IFG, corresponding to a repertoire of orofacial gestures potentially appropriate to the emotional content of the perceived vocal stimulus. This process of auditory–motor interaction may be supported by the primate dorsal auditory pathway (Scott and Johnsrude, 2003; Hickok and Poeppel, 2004; Warren et al., 2005), which includes projections from posterior temporal auditory association cortex to posterior inferior frontal cortex (Deacon, 1992).

Listening to emotionally arousing vocal stimuli was associated with activation in pre-SMA. Depth electrode recordings from human pre-SMA have demonstrated mirror-like responses when viewing emotional faces (Krolak-Salmon et al., 2006). Pre-SMA corresponds to area F6 in nonhuman primates, which receives projections from prefrontal and cingulate cortex (Luppino et al., 1993), and has been implicated in the gating and overall control of visuomotor transformations on the basis of external contingencies and motivations (Rizzolatti and Luppino, 2001). Highly arousing emotional vocalizations therefore engage a region involved in higher-order aspects of complex motor control. Convergent valence- and arousal-related perceptual responses within the somatotopically arranged left and right lateral premotor and motor cortices (Buccino et al., 2001; Alkadhi et al., 2002) were maximal in the face motor area (Buccino et al., 2001; Carr et al., 2003; Leslie et al., 2004) but also extended into more ventral regions involved in the motor control of articulation (Murphy et al., 1997; Blank et al., 2002; Wilson et al., 2004). Lateral premotor activation was found to extend posteriorly into primary motor regions in both hemispheres. Activation of primary motor cortex is typically associated with overt movement, but our EMG study clearly demonstrated that listening to affective vocal stimuli does not elicit overt facial movements or vocalizations. In fact, our results are in keeping with previous fMRI and transcranial magnetic stimulation studies of motor responses during speech perception (Watkins et al., 2003; Wilson et al., 2004) and observation of facial expressions (Carr et al., 2003; Leslie et al., 2004), which have demonstrated that perception of orofacial actions alone is sufficient to increase activity in primary motor cortex.

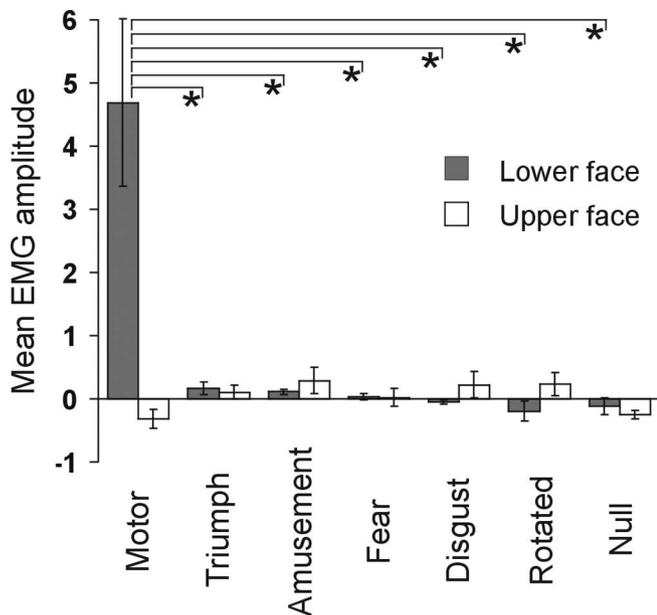Because of temporal constraints on fMRI data acquisition, we

Figure 6. EMG responses during listening to emotional vocalizations. Mean EMG amplitude values (microvolt units) were measured as the average change in the root mean square of recorded responses in the 1000–1499 ms time window poststimulus onset compared with prestimulus activity. EMG responses were measured over right-sided lower (zygomaticus major) and upper (corrugator supercilii) facial muscles are displayed for the four categories of emotional vocalization, the rotated baseline stimuli, the voluntary facial movement condition, and the null (resting) condition. Asterisks indicate the significance values associated with paired t test comparisons between the motor condition and each of the other experimental conditions (*$p < 0.03$). Error bars indicate $\pm$SEM.

were unable to incorporate an additional motor condition involving a negative facial expression, such as frowning. Thus, our delineation of cortical regions involved in the generation of facial expressions was based solely on a single positive facial expression. It is doubtful, however, that the spatial resolution of fMRI would have been sufficient to demonstrate significant topographical differences in activation for different facial expressions in motor and premotor cortical regions. Moreover, the EMG study failed to demonstrate any significant increase in brow muscle activity during perception of negative-emotion vocalizations. Therefore, we would argue that the absence of a motor condition involving a negative facial expression does not compromise the validity of our results.

Together, our findings suggest that listeners' motor responses to emotional vocalizations involved more than direct imitative activation of representations of facial or vocal expressions. The recruitment of mirror regions during action perception has been attributed not only to unconscious imitation and action recognition (Gallese et al., 1996; Rizzolatti and Arbib, 1998; Rizzolatti and Luppino, 2001; Kohler et al., 2002; Ferrari et al., 2003; Keysers et al., 2003) but also to more complex functions such as understanding the intention or goal of a perceived action (Ferrari et al., 2005; Iacoboni et al., 2005) or the preparation of nonimitative motor responses (Leslie et al., 2004). Based on our findings, we speculate that listening to vocal expressions of positive or highly arousing emotions activates representations of vocal and facial gestures appropriate to the emotion being communicated. The mirroring of social cues, a process not limited to imitation, is strongly associated with positive valence; for example, mirroring of body posture, gestures, and intonation is linked to enhanced establishment of rapport (Chartrand and Bargh, 1999). The greater propensity for positive-valence communications to auto-

matically activate motor representations may be a crucial component in the formation of empathic responses.

We are all familiar with the experience of responding to laughter or cheering with an involuntary smile or laugh. On the basis of this study, we argue that this impulse to respond to affective vocal communications with appropriate orofacial gestures is mediated by the automatic activation of orofacial motor cortical fields. We suggest that the enhanced motor response to perception of positive emotions provides a mechanism for mirroring the positive emotional states of others during primate social interaction. Mirroring behavior improves ease of social interaction (Chartrand and Bargh, 1999): motor facilitation in response to vocal communication of positive emotions may therefore provide a fundamental mechanism for establishing cohesive bonds between individuals in primate social groups. Given the importance of individual bonds and group cohesion for survival in many social species, such a mechanism may not be restricted to primates.

References

Adolphs R (2003) Cognitive neuroscience of human social behaviour. Nat Rev Neurosci 4:165–178.

Alkadhi H, Crelier GR, Boendermaker SH, Golay X, Hepp-Reymond MC, Kollias SS (2002) Reproducibility of primary motor cortex somatotopy under controlled conditions. Am J Neuroradiol 23:1524–1532.

Blank SC, Scott SK, Murphy KM, Warburton E, Wise RJS (2002) Speech production: Wernicke, Broca and beyond. Brain 125:1829–1838.

Blesser B (1972) Speech perception under conditions of spectral transformation. I. Phonetic characteristics. J Speech Hear Res 15:5–41.

Brett M, Anton JL, Valabregue R, Poline JB (2003) Region of interest analysis using an SPM toolbox [abstract]. NeuroImage 16:497.

Buccino G, Binkofski F, Fink GR, Fadiga L, Fogassi L, Gallese V, Seitz RJ, Zilles K, Rizzolatti G, Freund HJ (2001) Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. Eur J Neurosci 13:400–404.

Carr L, Iacoboni M, Dubeau MC, Mazziotta JC, Lenzi GL (2003) Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. Proc Natl Acad Sci USA 100:5497–5502.

Chartrand TL, Bargh JA (1999) The chameleon effect: the perception-behavior link and social interaction. J Pers Soc Psychol 76:893–910.

Dapretto M, Davies MS, Pfeifer JH, Scott AA, Sigman M, Bookheimer SY, Iacoboni M (2006) Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. Nat Neurosci 9:28–30.

Deacon TW (1992) Cortical connections of the inferior arcuate sulcus cortex in the macaque brain. Brain Res 573:8–26.

Dimberg U, Thunberg M, Elmehed K (2000) Unconscious facial reactions to emotional facial expressions. Psychol Sci 11:86–89.

Ekman P (1992) An argument for basic emotions. Cogn Emot 6:169–200.

Ekman P (2003) Emotions revealed: recognizing faces and feelings to improve communication and emotional life. London: Weidenfeld and Nicolson.

Fadiga L, Craighero L, Buccino G, Rizzolatti G (2002) Speech listening specifically modulates the excitability of tongue muscles: a TMS study. Eur J Neurosci 15:399–402.

Ferrari PF, Gallese V, Rizzolatti G, Fogassi L (2003) Mirror neuron responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. Eur J Neurosci 17:1703–1714.

Ferrari PF, Rozzi S, Fogassi L (2005) Mirror neurons responding to observation of actions made with tools in monkey ventral premotor cortex. J Cogn Neurosci 17:212–226.

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

Genovese CR, Lazar NA, Nichols T (2002) Thresholding of statistical maps in functional neuroimaging using the false discovery rate. NeuroImage 15:870–878.

Geyer S, Matelli M, Luppino G, Zilles K (2000) Functional neuroanatomy of the primate isocortex. Anat Embryol (Berl) 202:443–474.

Hall DA, Haggard MP, Akeroyd MA, Palmer AR, Summerfield AQ, Elliott MR, Gurney EM, Bowtell RW (1999) "Sparse" temporal sampling in auditory fMRI. Hum Brain Mapp 7:213–223.

Hauk O, Johnsrude I, Pulvermuller F (2004) Somatotopic representation of action words in human motor and premotor cortex. Neuron 41:301–307.

Hennenlotter A, Schroeder U, Erhard P, Catrop F, Haslinger B, Stoecker D, Lange KW, Ceballos-Baumannn AO (2005) A common neural basis for receptive and expressive communication of pleasant facial affect. Neuro-Image 26:581–591.

Hickok G, Poeppel D (2004) Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. Cognition 92:67–99.

Hietanen JK, Surakka V, Linnankoski I (1998) Facial electromyographic responses to vocal affect expressions. Psychophysiology 35:530–536.

Iacoboni M, Molnar-Szakacs I, Gallese V, Buccino G, Mazziotta JC, Rizzolatti G (2005) Grasping the intentions of others with one's own mirror neuron system. PLoS Biol 3:e79.

Keysers C, Kohler E, Umiltà MA, Nanetti L, Fogassi L, Gallese V (2003) Audiovisual mirror neurons and action recognition. Exp Brain Res 153:628–636.

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

Krolak-Salmon P, Hénaff MA, Vighetto A, Bauchet F, Bertrand O, Mauguière F, Isnard J (2006) Experiencing and detecting happiness in humans: the role of the supplementary motor area. Ann Neurol 59:196–199.

Leslie KR, Johnson-Frey SH, Grafton ST (2004) Functional imaging of face and hand imitation: towards a motor theory of empathy. NeuroImage 21:601–607.

Luppino G, Matelli M, Camarda R, Rizzolatti G (1993) Corticocortical connections of area F3 (SMA-proper) and area F6 (pre-SMA) in the macaque monkey. J Comp Neurol 338:114–140.

Murphy K, Corfield DR, Guz A, Fink GR, Wise RJ, Harrison J, Adams L (1997) Cerebral areas associated with motor control of speech in humans. J Appl Physiol 83:1438–1447.

Preston SD, de Waal FB (2002) Empathy: its ultimate and proximate bases. Behav Brain Sci 25:1–20.

Rizzolatti G, Arbib MA (1998) Language within our grasp. Trends Neurosci 21:188–194.

Rizzolatti G, Luppino G (2001) The cortical motor system. Neuron 31:889–901.

Russell JA (1980) A circumplex model of affect. J Pers Soc Psychol 39:1161–1178.

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

Scott SK, Young AW, Calder AJ, Hellawell DJ, Aggleton JP, Johnson M (1997) Impaired auditory recognition of fear and anger following bilateral amygdala lesions. Nature 385:254–257.

Scott SK, Blank SC, Rosen S, Wise RJS (2000) Identification of a pathway for intelligible speech in the left temporal lobe. Brain 123:2400–2406.

Warren JE, Wise RJ, Warren JD (2005) Sounds do-able: auditory-motor transformations and the posterior temporal plane. Trends Neurosci 28:636–643.

Watkins K, Paus T (2004) Modulation of motor excitability during speech perception: the role of Broca's area. J Cogn Neurosci 16:978–987.

Watkins KE, Strafella AP, Paus T (2003) Seeing and hearing speech excites the motor system involved in speech production. Neuropsychologia 41:989–994.

Wilson SM, Saygin AP, Sereno MI, Iacoboni M (2004) Listening to speech activates motor areas involved in speech production. Nat Neurosci 7:701–702.