RESEARCH ARTICLE

Audio-visual facilitation of the mu rhythm

**Lucy M. McGarry · Frank A. Russo ·
Matt D. Schalles · Jaime A. Pineda**

Abstract Previous studies demonstrate that perception of action presented audio-visually facilitates greater mirror neuron system (MNS) activity in humans (Kaplan and Iacoboni in Cogn Process 8(2):103–113, 2007) and non-human primates (Keysers et al. in Exp Brain Res 153(4):628–636, 2003) than perception of action presented unimodally. In the current study, we examined whether audio-visual facilitation of the MNS can be indexed using electroencephalography (EEG) measurement of the mu rhythm. The mu rhythm is an EEG oscillation with peaks at 10 and 20 Hz that is suppressed during the execution and perception of action and is speculated to reflect activity in the premotor and inferior parietal cortices as a result of MNS activation (Pineda in Behav Brain Funct 4(1):47, 2008). Participants observed experimental stimuli unimodally (visual-alone or audio-alone) or bimodally during randomized presentations of two hands ripping a sheet of paper, and a control video depicting a box moving up and down. Audio-visual perception of action stimuli led to greater event-related desynchrony (ERD) of the 8–13 Hz mu rhythm compared to unimodal perception of the same stimuli over the C3 electrode, as well as in a left central cluster when data were examined in source space. These results are consistent with

Kaplan and Iacoboni's (in Cogn Process 8(2):103–113, 2007), findings that indicate audio-visual facilitation of the MNS; our left central cluster was localized approximately 13.89 mm away from the ventral premotor cluster identified in their fMRI study, suggesting that these clusters originate from similar sources. Consistency of results in electrode space and component space support the use of ICA as a valid source localization tool.

Keywords Mu rhythm · Mirror neuron system · Multimodal facilitation · Independent components analysis

Introduction

The auditory components of action play an important role in action perception and understanding. For instance, the sound of approaching footsteps precedes and signals the arrival of an individual, and hearing articulatory gestures is central to perceiving speech. Recent research demonstrates that auditory perception of human actions elicits activation of the human mirror neuron system (MNS) similar to that during the visual perception of action (Gazzola et al. 2006). In addition, audio-visual perception of action elicits enhanced MNS activation as compared to unimodal perception as measured by fMRI (Kaplan and Iacoboni 2007). In the current study, we sought to replicate findings that MNS activation is facilitated during the audio-visual perception of action, using measurement of the mu rhythm, an electroencephalography (EEG) oscillation with peaks at 10 and 20 Hz that is desynchronized during the execution and perception of action.

Mirror neurons were first recorded in area F5 in the monkey premotor cortex (di Pellegrino et al. 1992; Gallese et al. 1996; Rizzolatti et al. 1996a). Single-cell recordings in

L. M. McGarry (✉) · F. A. Russo
Department of Psychology, Ryerson University,
350 Victoria St., Toronto, ON M5B 2K3, Canada
e-mail: lmcgarry@psych.ryerson.ca

M. D. Schalles · J. A. Pineda
Department of Cognitive Science, University of California,
San Diego, San Diego, La Jolla, CA 92037-0515, USA

J. A. Pineda
Group in Neurosciences, University of California, San Diego,
San Diego, La Jolla, CA 92037-0515, USA

monkeys suggest that action execution stimulates firing of the same neurons as perception of another individual (monkey or human) performing the same action, or a different action with the same end goal (di Pellegrino et al. 1992; Gallese et al. 1996; Rizzolatti et al. 1996a). Evidence for mirror neurons in humans has been demonstrated using EEG, TMS, neuroimaging, and single-cell recordings in a combination of parietal, prefrontal, and premotor areas (Cohen-Seat et al. 1954; Pineda 2005, 2008; Fadiga et al. 1995; Maeda et al. 2002; Buccino et al. 2001; Grafton et al. 1996; Iacoboni et al. 1999, 2001; Rizzolatti et al. 1996b; Mukamel et al. 2010; for a review see Rizzolatti and Craighero 2004), demonstrating that MNS areas are activated during execution and perception of goal-directed action (Fadiga et al. 2005; Maeda et al. 2002), as well as towards different actions with similar intentions (Rizzolatti et al. 2001). More recently, Mukamel et al. (2010) used single-cell recordings in medial frontal and temporal cortices in humans and demonstrated single cells that fired towards the execution and perception of action (Mukamel et al. 2010).

In terms of anatomical connectivity, research in the primate brain demonstrates connections between discrete superior temporal regions and two distinct regions of the frontal lobes, namely the caudal dorsolateral prefrontal cortex and rostral and ventral prefrontal areas (Romanski et al. 1999). These frontal areas overlap with area F5, which is thought to be a homolog of Broca's area (Romanski et al. 1999). In humans, similar pathways connecting posterior superior temporal gyrus and dorsolateral prefrontal cortex are reported by Frey et al. (2008) using diffusion fibre tractography.

The perception of audio-visual (AV) action elicits supra-additive responses in areas comprising the MNS in humans, such as the STS and ventral premotor cortex (Kaplan and Iacoboni 2007; Keysers et al. 2003). In Kaplan and Iacoboni's (2007) fMRI study, videos were presented to participants depicting two hands ripping a sheet of paper, or a control video of a box moving up and down, accompanied by beeping sounds. For the experimental stimuli only, AV facilitation of action observation was observed in the ventral premotor cortex, on the border of areas 44, 6, 3a, and 3b, at Talaraich coordinates (−64, 0, 18). In the current study, we sought to replicate these findings and extend them to the realm of EEG mu suppression measurement.

There is a rich literature on AV integration and its contributions to perceptual sensitivity that may be relevant to questions about AV facilitation in the MNS. This literature demonstrates that multimodal processing results in enhanced perception of unimodal information (Driver 1996), earlier neural responses to multimodal information (Welch et al. 1986; Giard and Peronnet 1999), and in some cases supra-additive neural responding to multimodal

versus unimodal information (Stein and Meredith 1993). Stein and Meredith (1993) demonstrated in single-cell recordings of the superior colliculus that spatially and temporally congruent multimodal cues enhance responding in a multiplicative way, as compared to responses to either mode alone. When these responses are spatially separated, the response is suppressed. This suggests that multimodal relationships between stimuli affect stimulus processing at the unimodal level. It appears that this information then feeds back to influence how the unimodal stimuli are perceived. Although there are differences in functionality and connectivity between neurons in the superior colliculus and those in the MNS, it is possible that they respond in a similar manner to multimodal cues. In the current study, we expected that AV action stimuli might elicit earlier and more powerful responses compared to audio or visual unimodal action stimuli in the MNS, perhaps even in an additive or multiplicative way.

Activity in the MNS is hypothesized to be reflected in the mu rhythm, which is measured through scalp electroencephalography (EEG), and becomes desynchronized during the perception and execution of action (Cohen-Seat et al. 1954; Gastaut and Bert 1954; Cochin et al. 1999; Pineda 2005, 2008). Mu suppression is thought to reflect mirror neuron activity, specifically downstream regulation of the motor cortex via parietal and prefrontal mirroring areas (Pineda 2005, 2008). Measured over central electrodes, with frequency peaks at around 10 and 20 Hz, these rhythms are suppressed in healthy individuals during visual or auditory observation of action, as well as performance of action (Pineda 2008). The alpha component (8–13 Hz) of mu is thought to reflect MNS activation in the postcentral somatosensory cortex, with the beta band (15–25 Hz) reflecting slightly anterior motor activity (Pfurtscheller et al. 1997, 2000; Hari et al. 1998; Pineda 2008). Increased BOLD signal in frontal and parietal areas has been correlated with the suppression of the alpha EEG response (Laufs et al. 2003a, b), as well as marginally with the beta response (Laufs et al. 2003b). These negative correlations occurred in areas such as the frontal and parietal cortex, areas containing mirror neurons (Rizzolatti and Craighero 2004). Because these frequency bands overlap with mu, this supports the premise that the mu response is linked to activity in these regions. Consistent with this, Keuken et al. (2011) recently showed that using transcranial magnetic stimulation to disrupt activity in the inferior frontal gyrus directly affects the modulation of mu rhythms over sensorimotor cortex. These findings are consistent with the hypothesis that mu rhythm suppression during action observation reflects downstream modulation of activity in motor neurons in the primary motor cortex by mirror neurons in the inferior frontal gyrus that are involved in action planning (Pineda 2005).

Similar to other measurements of activation in the MNS (Rizzolatti and Craighero 2004), the mu rhythm is modulated during perception of human movement but not object movement (Cochin et al. 1998). Cochin et al. (1998) found that viewing human movements led to the suppression of the mu rhythm in alpha and beta frequency ranges, whereas viewing object movements did not. Also, like other measurements of the MNS (Rizzolatti and Craighero 2004), the mu rhythm is more sensitive to movements such as a hand operating on an object, rather than non-object-oriented human movement (Muthukumaraswamy et al. 2004).

In the current study, we examined whether AV action compared to unimodal action would facilitate event-related desynchronization (ERD; Pfurtscheller and Aranibar 1979; Pfurtscheller 1992; Neuper et al. 2006) of the mu rhythm. ERD corresponds to mu suppression and reflects a decreasing post-stimulus change in power in the mu band. Subjects viewed action and non-action audio-only (A), visual-only (V), and audio-visual (AV) congruent stimuli while their EEG mu rhythms were measured. We expected that mu suppression would be greatest during the perception of AV action stimuli. These results would suggest that the MNS responds selectively to action-related AV stimuli and that this facilitation can be captured through measurement of mu ERD. These findings will allow us to conduct future studies concerning AV facilitation of the MNS in healthy adults as well as patient groups using EEG technology, which is a non-invasive, temporally sensitive, and cost-effective measurement tool.

Methods

Participants

Thirty-four undergraduate students (22 women) from the University of California, San Diego, with a mean age of 21.0 (SD = ±3.2) and an average of 14.6 years of education (SD = ±2.0) participated in the study for course credit. Exclusion criteria included the presence of psychiatric or neurological disorders. One participant was excluded from analyses for psychiatric reasons. The study was reviewed and approved by the University of California, San Diego Human Research Protections Program.

Experimental design

All participants were involved in the same one-hour experiment. The independent variable was the stimulus type: Action A, V, or AV, control A, V, or AV, or self-performed action. The dependent variable of interest was the amount of mu suppression during perception of multimodal and unimodal action stimuli. Mu suppression during action percep-

tion conditions was also compared with control conditions, as well as with mu suppression during action execution.

Participants were seated in a comfortable chair 60 cm from the computer screen with speakers located on either side of the screen. Each video was repeated 20 times in random order. Perception trials were preceded by a 5-s black fixation cross on a white screen, while action execution trials were preceded by a 5-s green cross to signal that they should pick up a sheet of paper, followed by an 8-s red cross during the action execution trial to signal that they should begin ripping.

Apparatus

EEG Recording

EEG was recorded using a Neuroscan Synamps system and an electro-cap. Nineteen electrodes were placed, following the International 10–20 electrode placement method, at: FP1, FP2, F3, Fz, F4, F7, F8, C3, Cz, C4, P3, Pz, P4, T3, T4, T5, T6, O1, O2. One electrooculogram (EOG) electrode was placed below the eye, and reference electrodes were placed behind each ear (mastoids). Impedance was set to below 10 kΩ. All electrodes were amplified by a factor of $1,000\times$ and sampled at 500 Hz. Online bandpass filtering was set at 0.3–100 Hz (half-amplitude, 3 dB/octave roll-off). The experiment was conducted in a sound-attenuated chamber.

Stimuli

Stimuli consisting of short-duration videos were presented using NeuroBehavioral Presentation software run on a separate PC computer located outside the chamber that was connected to a monitor within the chamber. Sound was presented via speakers located on either side of the computer screen, at a consistent volume. Participants were asked to observe a series of 8-s videos in which 2 hands were depicted ripping a sheet of paper 6 times, with a 5-s ITI. The stimuli were identical to those used in Kaplan and Iacoboni's (2007) fMRI study and provided courtesy of the authors. Videos were either presented audio-visually (AV action condition), visual-only (V action condition), or audio-only (A action condition). Control stimuli consisted of a blinking square accompanied by a pure tone (261.7 Hz) that ranged in duration between 200 and 500 ms, with onset of beeps aligned with the onset of rips. Control stimuli were presented in audio-visual (AV control), visual (V control), and audio (A control) conditions. An additional condition was included in which participants were asked to pick up a sheet of paper resting on the arm of the chair they were sitting in during the intertrial interval (ITI), and then rip the sheet of paper 6 times during the following 8-s trial.

Data analysis

Using the EEGLab Matlab toolbox (DeLorme and Makeig 2004), channels were located using Montreal Neurological Institute (MNI) coordinates, and data were referenced to the average. Data were bandpass-filtered (3–30 Hz) and organized into trials. Data were initially examined in electrode space. Additionally, in order to isolate mu-specific components, a source localization procedure was performed on the data. It was found that multiple brain sources were contributing to activity at the central electrodes. Because of this, data were also examined using independent components analysis (ICA; DeLorme and Makeig 2004), isolating brain components instead of electrodes. Component decomposition was performed on the entire epoched data set using a second-order blind identification (SOBI) algorithm, with the number of correlation matrices set to 50. SOBI is a blind-source separation (BSS) technique that exploits the second-order statistics of the measurements to compute an estimate of the mixing matrix (full details can be found in Belouchrani et al. 1997; also see Moore et al. 2012). Dipoles were then fitted to each component, and trials containing amplitudes greater than 75 or less than −75 mV were rejected. An average of 17% of trials was deleted overall.

For the action condition, an average of 36% of trials had to be rejected due to excessive artefact, probably due to movement artefact from the gross movements involved in the procedure, compared to only 14.5% in the remaining conditions. The remaining trials elicited event-related spectral perturbations (ERSPs) that appeared largely synchronized, which is opposite to the trend that is normally observed in the mu band while participants are engaged in action. However, it is important to note that procedures employed in previous studies have involved much smaller movements (e.g. finger tapping) than was required in the current study. When the data were examined in electrode space, looking at differences between conditions over central electrodes C3 and C4, patterns appeared comparable to those observed in component space.

Components were manually inspected for each participant and any components contributing eye movement artefact or other muscular artefact were deleted. Components were clustered using the k-means procedure, and outliers were excluded based on three standard deviations. Eight clusters were formed.

For each cluster, ERSPs were compared across the alpha range of mu (8–13 Hz) as well as the beta mu component (17–23 Hz), for the time range containing the first two rips (0–2,500 ms). This narrower time range, and narrower beta frequency range (as compared to 15–25 Hz), was chosen to enhance power in the face of multiple comparisons across each time point and frequency

point in the ERSP comparison. It was also expected that earlier rips would elicit greater activity in mirroring areas before habituation occurred. ERSPs depict deviations in spectral power relative to baseline, specifically ERDs (post-stimulus decreasing change in power) and ERSs (post-stimulus increasing change in power), and allow for comparison at multiple time points and frequency points within a time/frequency band (described in Moore et al. 2012). The purpose of examining data at multiple points across a time and frequency window is to capture differences between conditions that may occur at one point in time or frequency but not another. If time and frequency are collapsed, this may cancel out some important differences between conditions, due to oscillating synchronization and desynchronization of activity. All ERSP results were corrected using the Benjamini and Hochberg false discovery rate correction (Benjamini and Hochberg 1995) with an alpha value of .05.

For each cluster, a one-way analysis of variance (ANOVA) was conducted with condition (A action, V action, AV action, A control, V control, AV control, action execution) as the within-subjects factor. Following this, pairwise comparisons were conducted on each action trial versus its respective control (i.e. A action vs. A control, V action vs. V control, and so on), as well as for unimodal versus multimodal action conditions, and unimodal versus multimodal control conditions.

Results

Corresponding significant multimodal facilitation was found towards perception of audio-visual action, but not audio-visual control stimuli, in the left central C3 electrode, as well as in a left central ICA cluster.

Electrode space analyses

Audio-visual facilitation of action was examined at each electrode site, for the 8–13 and 17–23 Hz ranges. Event-related spectral perturbations (ERSPs) were compared. The one-way ANOVA was significant for a main effect of condition as depicted in Fig. 1, across both frequency ranges. Pairwise comparisons demonstrated that there was audio-visual facilitation of mu ERD versus audio-alone as well as visual-alone at C3 and C4 electrode sites for 8–13 Hz, but not 17–23 Hz (see Fig. 2). Both A and V action stimuli generated greater ERD than control at the C3 location only, while AV action stimuli generated greater ERD than control at C3 and C4 locations, indicating that action facilitation of mu occurred at the C4 site when audio and visual presentation of action occurred together.
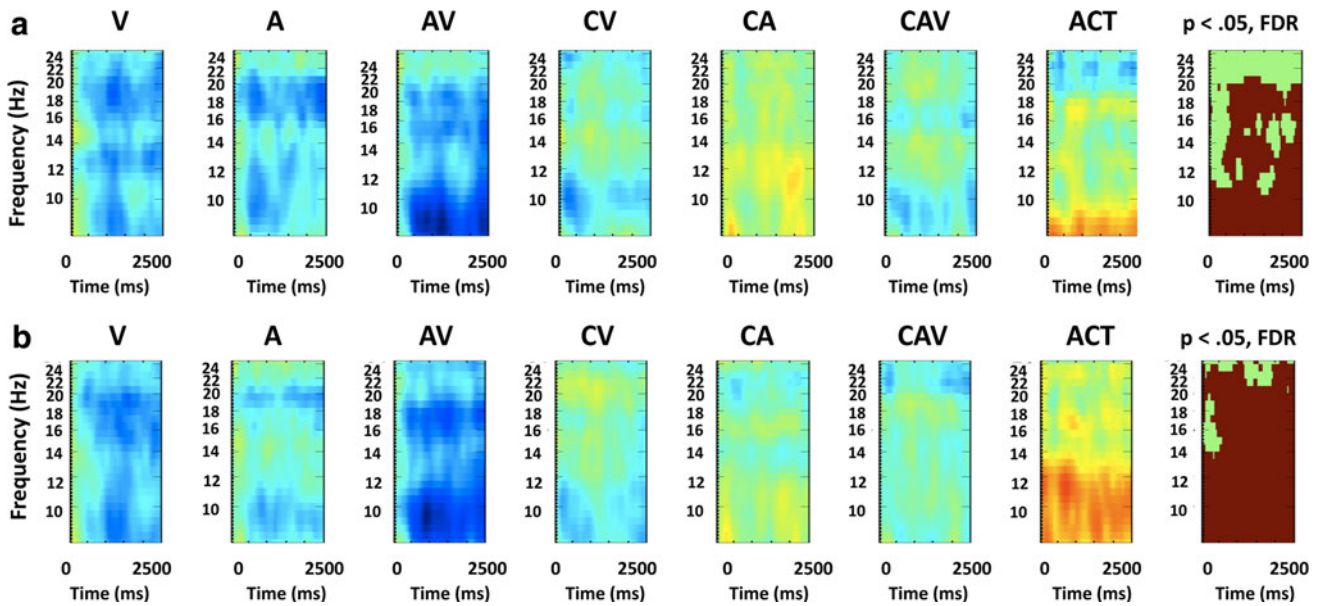
Fig. 1 ANOVAs for all conditions, 8–25 Hz, for central electrodes (**a** C3; **b** C4). On the ERSP graphs, warmer colours indicate synchronization and cooler colours indicate desynchronization. The *y*-axis indicates frequency and the *x*-axis indicates time. Graph on far *right*: areas that are significant for a main effect of condition are *maroon* in colour, with a Benjimini Hochberg FDR correction for multiple comparisons at an alpha level of .05. V, visual condition; A, auditory condition; AV, audio-visual action condition; CV, control visual condition; CA, control auditory condition; CAV, control audio-visual condition; ACT, action condition. Each ANOVA was significant for a main effect of condition (colour figure online)
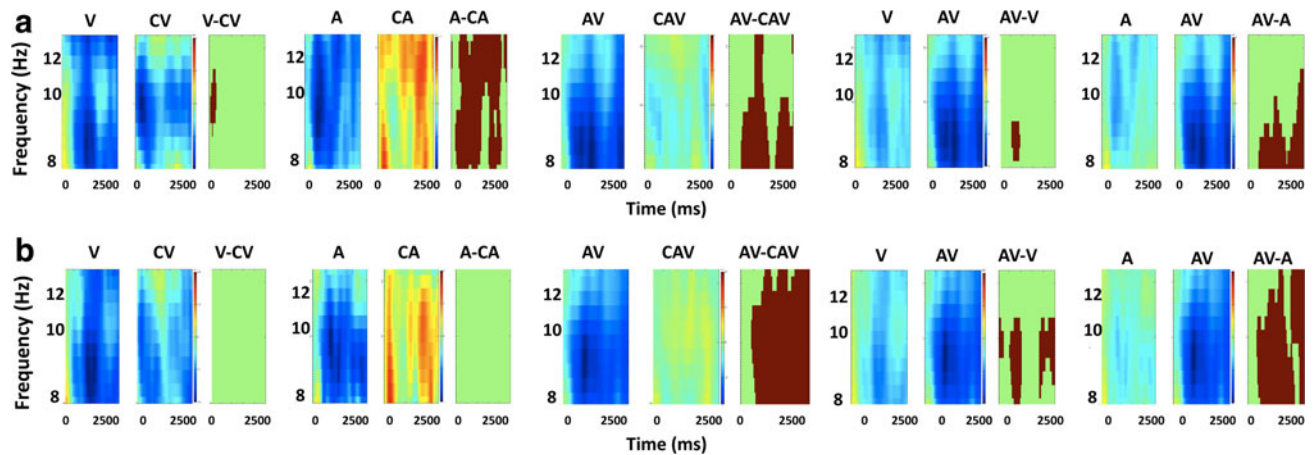
Fig. 2 Pairwise comparisons, 8–13 Hz, for central electrodes (**a** C3; **b** C4). Unimodal and bimodal conditions were compared to their respective control conditions, as well as to each other. On the ERSP graphs, *warmer colours* indicate synchronization and *cooler colours* indicate desynchronization. The *y*-axis indicates frequency and the *x*-axis indicates time. Graph to the *right* of each pair of ERSP graphs: areas that are significant for a main effect of condition are *maroon* in colour, with a Benjimini Hochberg 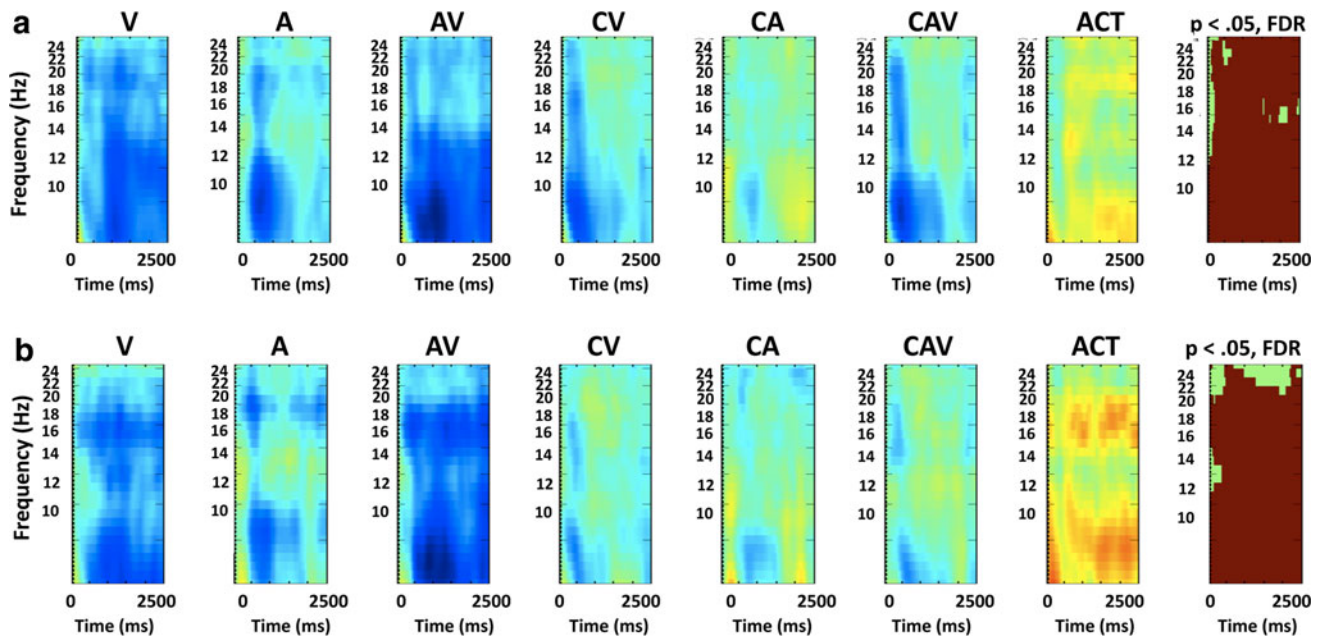FDR correction for multiple comparisons at an alpha level of .05. V, visual condition; A, auditory condition; AV, audio-visual action condition; CV, control visual condition; CA, control auditory condition; CAV, control audio-visual condition. For the C3 electrode, there was audio-visual facilitation in the 8–13 Hz range, and all action conditions generated greater ERD than control conditions. For C4, there was audio-visual facilitation of action, and only the audio-visual action condition generated greater mu ERD than control conditions (colour figure online)

Independent components analyses

Clusters

Eight clusters were formed. Clusters originating from central sources with frequency peaks in the mu range were selected for further examination. In addition, one occipital cluster generated audio-visual facilitation of suppression in the 8–13 Hz range, so this cluster was examined further to ensure that it was representative of alpha and not mu. For each of these clusters, the one-way ANOVA was significant for a main effect of condition at time and frequency points spanning 8–25 Hz (see Fig. 3 for ANOVA results for the two centrally generated clusters). A left central-parietal

Fig. 3 ANOVAs for each condition for the left central-parietal cluster (**a**), and mid central cluster (**b**), 8–25 Hz. On the ERSP graphs, *warmer colours* indicate synchronization and *cooler colours* indicate desynchronization. *V,* visual condition; *A,* auditory condition; *AV,* audio-visual action condition; *CV,* control visual condition; *CA,* control auditory condition; *CAV,* control audio-visual condition;

ACT, action condition. The *y*-axis indicates frequency and the *x*-axis indicates time. Graph on far *right*: areas that are significant for a main effect of condition are maroon in colour, with a Benjimini Hochberg FDR correction for multiple comparisons at an alpha level of .05. All ANOVAs were significant for a main effect of condition (colour figure online)

cluster generated multimodal action facilitation of mu ERD in the 8–13 Hz range, and a mid central cluster generated multimodal facilitation of mu ERD versus audio but not compared to visual action stimuli, in the 8–13 and 17–23 Hz ranges. A right central cluster that was also formed in the clustering analysis yielded no significant differences between ERSPs.

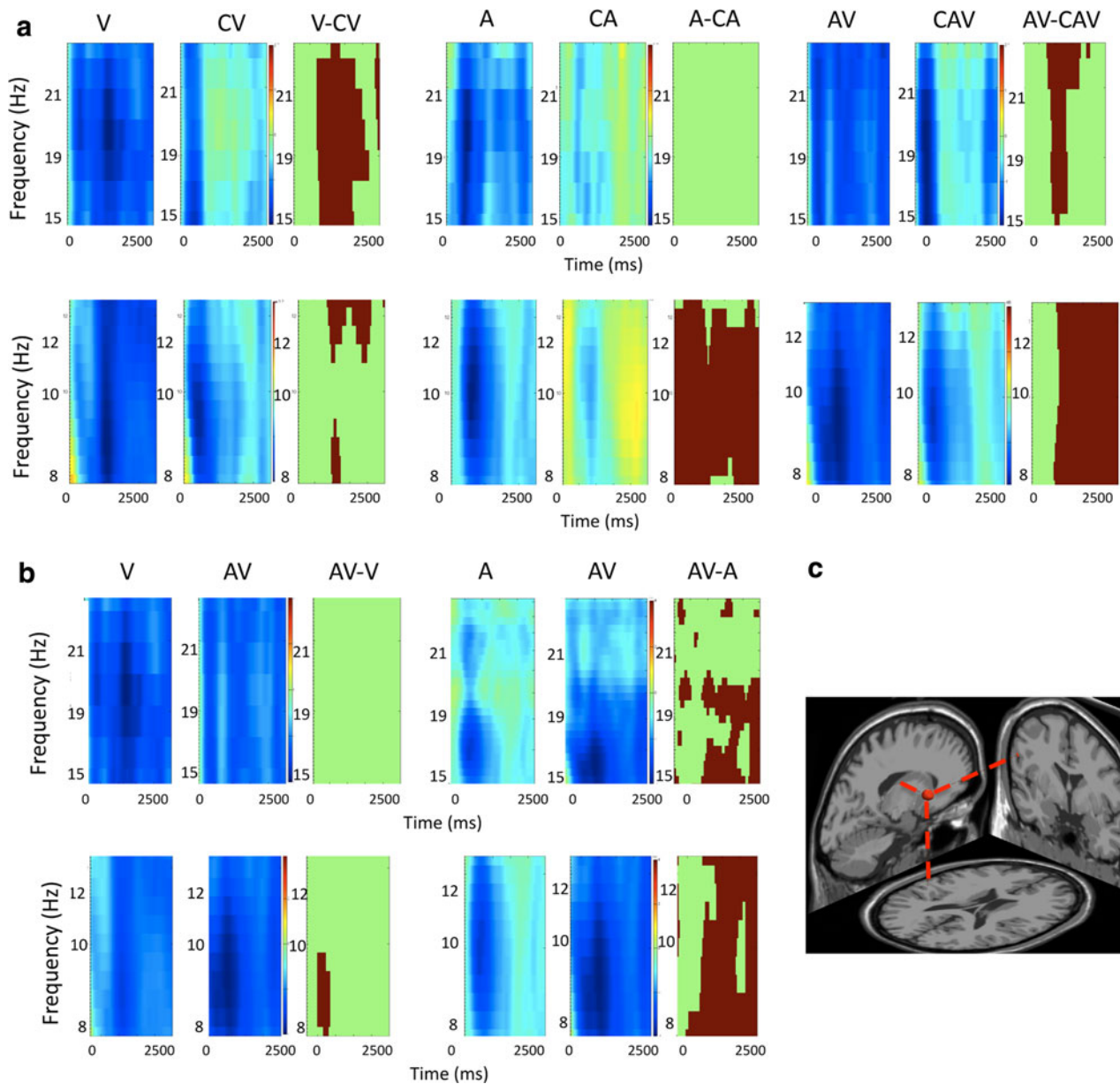*Left central-parietal cluster*

The left central-parietal cluster with mean Talairach coordinates at (−57, −12, 18) (see Fig. 4) generated greater mu ERD in the 8–13 Hz range but not the 17–23 Hz range during perception of AV action versus V action or A action, demonstrating bimodal facilitation of the mu rhythm. In addition, this cluster generated greater mu ERD for V action as compared to V control, A action versus A control, and AV action versus AV control stimuli in the 8–13 Hz range, as well as V action compared to V control and AV action versus AV control in the 17–23 Hz range. Please see Fig. 4 for visual depictions of these comparisons and their significance. A pairwise comparison of ERSPs for multimodal versus unimodal control conditions yielded significant differences between A and AV, but not between V and AV conditions.

Middle central cluster

A middle central cluster with mean Talairach coordinates at (−9, −2, 13) (see Fig. 5) generated greater mu ERD in the 8–13 Hz range and 17–23 Hz range during the perception of AV action as compared to A but not V action stimuli. In addition, greater mu ERD was found in the 8–13 and 17–23 Hz ranges for V action as compared to V control, A action compared to A control, and AV action compared to AV control stimuli. See Fig. 5 for visual depictions of these comparisons and their significance. There was no multimodal facilitation of mu ERD for control stimuli.

Left occipital cluster

A left occipital cluster with mean Talairach coordinates at (−44, −62, 8) was also examined. AV action stimuli elicited greater suppression in the 8–13 Hz alpha range than A or V action stimuli, but AV control stimuli did not elicit multimodal facilitation of alpha suppression compared to control stimuli. ERSPs generated from this cluster showed no differentiation between action and control conditions. There were also no differences across these conditions in a right occipital cluster that was also formed in the clustering analysis.

Fig. 4 a Pairwise comparisons between action and control conditions; left central-parietal cluster. ERSP comparisons are depicted for each action condition versus its respective control condition. On the ERSP graphs, *warmer colours* indicate synchronization and *cooler colours* indicate desynchronization. The *y*-axis indicates frequency and the *x*-axis indicates time. Graph to the right of each pair of ERSP graphs: areas that are significant for a main effect of condition are *maroon* in colour, with a Benjimini Hochberg FDR correction for multiple comparisons at an alpha level of .05. *V,* visual condition; *A,* auditory condition; *AV,* audio-visual action condition; *CV,* control visual condition; *CA,* control auditory condition; *CAV,* control audio-visual condition.

Each of the action conditions elicited greater desynchronization along the 8–13 and 17–23 Hz mu range than their respective control conditions, except for audio which generated greater desynchronization for action versus control at the 8–13 Hz but not the 17–23 Hz range. **b** Pairwise comparisons between multimodal and unimodal conditions; left central-parietal cluster. ERSP comparisons are depicted for the AV action condition versus A and V action conditions. Audio-visual facilitation was found versus both audio and visual action conditions in the 8–13 Hz range, and versus the audio condition in the 17–23 Hz range. **c** Mean dipole location for left central-parietal cluster. Talairach coordinates: −57, −12, 18 (colour figure online)

Discussion

Our results indicate that mu ERD, an index of MNS activity, is enhanced by the multimodal presentation of action associated with the left central-parietal area. This cluster generated more desynchronized activity in the 8–13 Hz component of the mu band during perception of audio-visual action compared to either visual or audio action, and each action condition generated greater ERD than control conditions. Similar results were obtained under examination of the
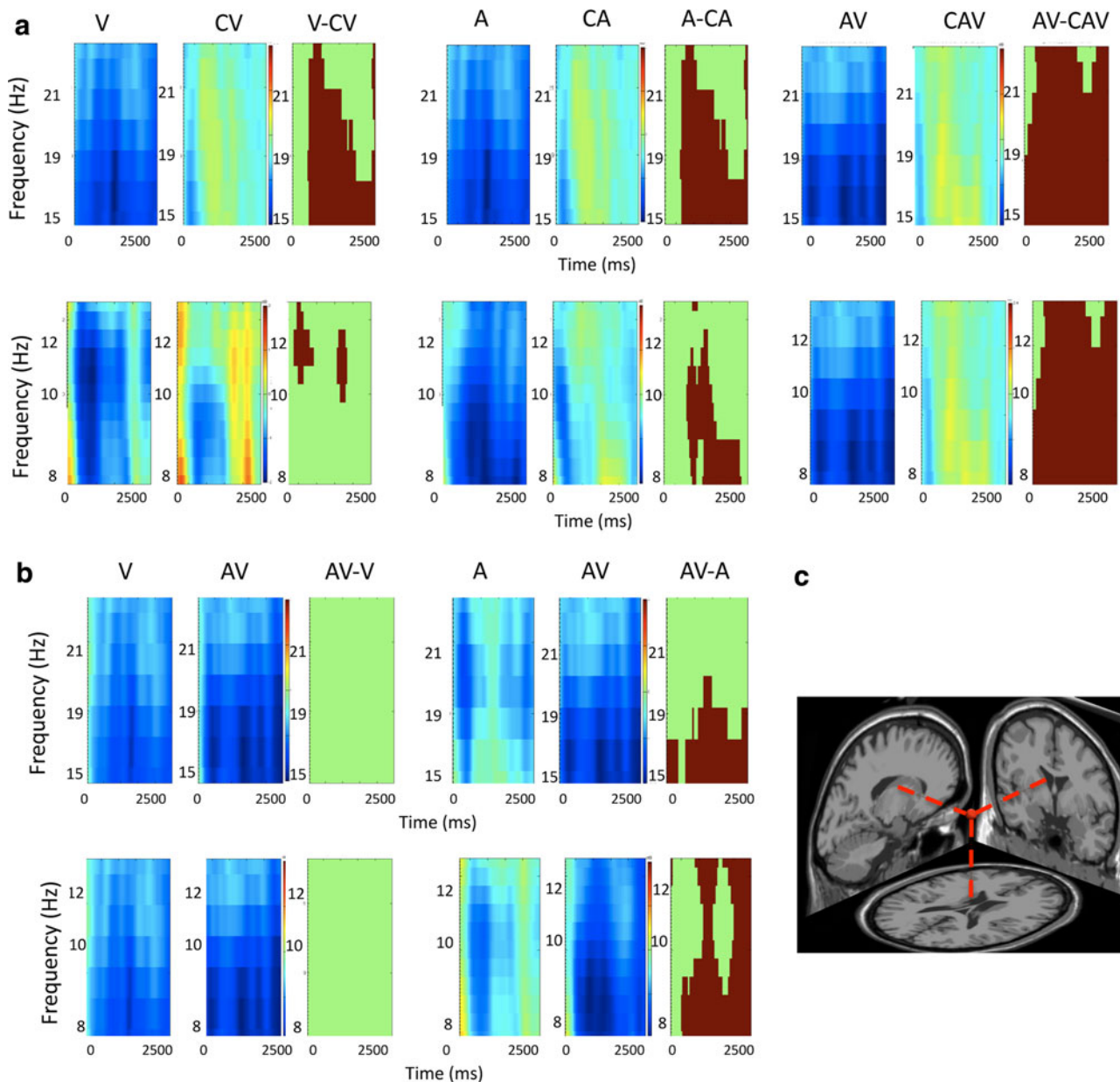
Fig. 5 a Pairwise comparisons between action and control conditions; mid central cluster. ERSP comparisons are depicted for each action condition versus its respective control condition. On the ERSP graphs, *warmer colours* indicate synchronization and *cooler colours* indicate desynchronization. The *y*-axis indicates frequency and the *x*-axis indicates time. Graph to the right of each pair of ERSP graphs: areas that are significant for a main effect of condition are *maroon* in colour, with a Benjimini Hochberg FDR correction for multiple comparisons at an alpha level of .05. *V,* visual condition; *A,* auditory condition; *AV,* audio-visual action condition; *CV,* control visual condition; *CA,* control auditory condition; *CAV,* control audio-visual condition. Each of the action conditions elicited greater desynchronization along the 8–13 Hz mu range than their respective control conditions. **b** Pairwise comparisons between multimodal and unimodal conditions; mid central cluster. ERSP comparisons are depicted for the AV action condition versus the A condition at 8–13 and 17–23 Hz. The AV action condition elicited greater desynchronization along the 8–13 and 17–23 Hz mu ranges than the A condition. There were no significant differences between AV and V conditions. **c** Mean dipole location for mid central cluster. Talairach coordinates: −9, −2, 13 (colour figure online)

C3 electrode in electrode space, an electrode that is also located in the left central area, where mu suppression towards action stimuli is commonly found. The mean of this cluster is strikingly close to the ventral premotor area of the MNS, identified in Kaplan and Iacoboni's (2007) fMRI study

as generating facilitated activation during the perception of multimodal action. The ventral premotor cortex identified in Kaplan and Iacoboni's (2007) study is approximately 13.89 mm away from the mean of this cluster, suggesting that these clusters could originate from the same source.

The control audio-visual condition also generated more ERD than the audio but not the visual condition in this cluster (and not in electrode space for C3). This may indicate that the audio-visual facilitation that occurs in the ventral premotor cortex is due to an additive combination of audio-visual facilitation and action perception, rather than audio-visual facilitation being necessarily action related.

A middle central cluster generated greater mu ERD in the alpha and beta mu bands during perception of visual, audio, and audio-visual action relative to control conditions, and also generated greater ERD of audio-visual versus audio action but not visual action stimuli. Rather than multimodal facilitation, it is likely that the greater contribution of visual action to mu ERD in this cluster was responsible for the enhanced ERD in both visual and audio-visual conditions. There was no enhanced ERD of mu for control AV versus A or V stimuli, supporting the idea that enhanced ERD occurred primarily for action stimuli in this cluster.

An occipital cluster generated greater alpha ERD during perception of audio-visual versus visual but not auditory stimuli. However, this cluster did not differentiate between action and non-action conditions. This suggests that this cluster was generating occipital alpha activity. Changes in the occipital alpha band towards multimodal stimuli are likely to be related to changes in visual attention elicited by AV stimulation (Sauseng et al. 2005), whereas mu is specifically desynchronized during the perception and execution of biological motion. Occipital alpha did not show similar facilitation of AV versus A or V control stimuli. Perhaps this is because the control stimuli were not as visually complex as the experimental stimuli.

Interestingly, while audio action stimuli generated greater ERD of the mu band than audio control stimuli in both the left central and middle central clusters, visual action stimuli appeared to generate greater ERD than audio in both instances, though this was not significant at any point in the time window after the FDR correction for multiple comparisons. This could mean that while audio action information contributes to mu desynchronization, desynchronization of the mu rhythm relies more on visual information than audio information. There may also be stimuli such as speech or singing, where the auditory signal is more informative about movement intention than the visual signal. However, there are other possible reasons why visual stimuli elicited greater mu suppression than audio in the current study. Visual stimuli were present onscreen for the entirety of each 8-s video. However, audio information was only available while the ripping occurred and went away while the hands in the video prepared to create an additional rip. The pervasiveness of the visual stimuli could explain the enhanced ERD. In addition, the presence of visual stimuli always preceded sound during audio-visual

trials. Follow-up studies will equate the timing of visual and audio stimulus presentation, and experiment with other types of stimuli, in order to account for these potential factors.

Future studies should also explore whether the MNS relies more on auditory information when visual stimuli are degraded. The inverse effectiveness rule of multisensory integration suggests that the likelihood of multisensory integration increases as the ambiguity of its unimodal constituents increases (Meredith and Stein 1986). There is considerable behavioural evidence in support of this rule (e.g. Alais and Burr 2004; Walden et al. 1977), so it is reasonable to expect that when information from the visual modality is degraded, participants may rely more heavily on audio information, and there might be more audio-visual integration observed.

Previous research has demonstrated that multimodal stimulus attributes are perceived earlier and influence perception of unimodal stimulus attributes. In Driver's (1996) study, it was shown that multimodal cues are processed before unimodal cues are fully processed. In this case, it led viewers to perceive the spatial location multimodally before they would have processed the spatial location of the auditory source, and this influenced their perception of the auditory spatial location. The colour of a food item has been shown to affect the gustatory experience of flavour intensity (DuBose et al. 1980), the magnitude of a singer's head movement has been shown to influence the auditory experience of sung interval size (Thompson et al. 2010), and the sound of parchment on the skin influences the tactile experience of texture (Jousmäki and Hari 1998). It has been suggested that the time course of cross-modal perception occurs earlier than that of unimodal perception, serving to influence processing of unimodal stimulus properties (Welch et al. 1986). This is also observed for the experience of orthogonal multimodal cues. Welch et al. (1986) demonstrated that auditory beeps influenced perception of the rate of unrelated visual flickers. Giard and Peronnet (1999) found an early event-related potential (ERP) signal whose pattern of responding suggests that an auditory cue, even if it is unrelated to the visual cue, can affect visual processing at an early stage. In terms of processing in the MNS, we might expect that AV processing will also occur earlier than A or V alone. Future studies should examine this by equating onset time of audio and visual stimuli to be able to compare latency of mu suppression in unimodal and bimodal conditions.

Our findings complement recent research, which has also suggested the presence of AV neurons in the MNS that generate supra-additive responding during perception of AV congruent stimuli, as opposed to A-, V-, or AV-incongruent stimuli (Barraclough et al. 2005; Keysers et al. 2003; Kaplan and Iacoboni 2007). This facilitation occurs

for action-related stimuli only (Barraclough et al. 2005; Kaplan and Iacoboni 2007). AV facilitation has previously been measured using single-cell recordings in monkeys (Barraclough et al. 2005, Keysers et al. 2003; Kohler et al. 2002), as well as fMRI studies in humans (Kaplan and Iacoboni 2007). AV facilitation occurs only when audio-visual stimuli represent spatially and temporally congruent, goal-directed action. This implies that mirror neurons not only respond selectively to sets of actions serving a single intention, but that they serve to single out matching goal-directed actions from multiple modalities (Keysers et al. 2003). Research suggests that AV integration of action and non-action stimuli occurs first in the STS (Barraclough et al. 2005; Calvert et al. 2000; Keysers et al. 2003; Skipper et al. 2005), and the current study supports findings that AV action stimuli are further processed in the left ventral premotor cortex (Keysers et al. 2003; Kaplan and Iacoboni 2007).

Keysers et al. (2003) used single-cell recordings in monkeys to measure MNS activity during the perception of AV action. They found AV mirror neurons in the ventral premotor cortex of the monkey. These neurons fire whether action is performed, heard, or seen. Of 33 single cells that they measured in this area, 22 showed both visual and auditory selectivity, and 8 of these 22 neurons demonstrated additive or supra-additive responding to AV action, as compared to A and V action alone. This evidence supports findings that that AV integration continues to occur within the mirror neuron system as well as in the STS (Kaplan and Iacoboni 2007).

Future studies should examine whether incongruent audio-visual action stimuli generate a comparable degree of multimodal facilitation as compared to congruent action stimuli. Keysers et al. (2003) found that individual audio-visual neurons in monkeys discriminate in firing towards congruent versus incongruent action stimuli. It is possible that measurement in humans using fMRI or EEG could be more difficult to detect differences in brain responsivity to congruent versus incongruent audio-visual actions, since these devices measure summed activity of individual neurons. However, it is possible that overall activity will be greater during congruent stimuli than incongruent, since it is likely that a greater overall number of neurons will desynchronize during congruent than incongruent audio-visual action perception.

Conclusions

In the current study, we have identified a left central-parietal cluster, possibly associated with the ventral premotor cortex, that generates greater ERD in the 8–13 Hz frequency range during perception of audio-visual action as compared to audio action, visual action, and non-action stimuli. Similar results were found when data were examined over the C3 electrode in channel space, bolstering the results and also bolstering accuracy of the ICA method. These results are consistent with findings in fMRI literature in humans (Kaplan and Iacoboni 2007) and single-cell recordings in monkeys (Keysers et al. 2003) that indicate supra-additive responding of MNS activity the ventral premotor cortex during multimodal action perception. In addition, these findings demonstrate that this multimodal action facilitation can be detected using EEG, allowing for time-sensitive and cost-effective data collection. Future studies should examine multimodal facilitation of mu oscillations in autistic populations, in order to learn whether multimodal integration impairments (Le Bel, Pineda and Sharma 2009) are related to dysfunctional AV facilitation of the MNS and to aid the development of MNS-related training and therapy for autism.

References

Alais D, Burr D (2004) The ventriloquist effect results from near-optimal bimodal integration. Curr Biol 14(3):257–262

Barraclough NE, Xiao D, Baker CI, Oram MW, Perrett DI (2005) Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. J Cogn Neurosci 17(3):377–391

Belouchrani A, Abed-Meraim K, Cardoso J (1997) A blind source separation technique using second-order statistics. IEEE Trans Signal Process 45(2):434–444

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. J R Stat Soc Ser B (Methodological) 57(1):289–300

Buccino G, Binkofski F, Fink GR, Fadiga L, Fogassi L, Gallese V et al (2001) Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. Euro J Neurosci 13:400–404

Calvert GA, Campbell R, Brammer MJ (2000) Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. Curr Biol 10(11):649–657

Cochin S, Barthelemy C, Lejeune B, Roux S, Martineau J (1998) Perception of motion and qEEG activity in human adults. Electroencephalogr Clin Neurophysiol 107(4):287–295

Cochin S, Barthelemy C, Roux S, Martineau J (1999) Observation and execution of movement: similarities demonstrated by quantified electroencephalography. Euro J Neurosci 11(5):1839–1842

Cohen-Seat G, Gastaut H, Faure J, Heuyer G (1954) Etudes expérimentales de l'activité nerveuse pendant la projection cinématographique. Rev Int Filmologie 5:7–64

Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 134:9–21

di Pellegrino G, Fadiga L, Fogassi L, Gallese V, Rizzolatti G (1992) Understanding motor events: a neurophysiological study. Exp Brain Res 91(1):176–180

Driver J (1996) Enhancement of selective listening by illusory mislocation of speech sounds due to lip-reading. Nature 381(2):66–68

DuBose CN, Cardello AV, Maller O (1980) Effects of colorants and flavorants on identification perceived flavor intensity and hedonic quality of fruit flavored beverages and cake. J Food Sci 45(5):1393–1399

Fadiga L, Fogassi L, Pavesi G, Rizzolatti G (1995) Motor facilitation during action observation: a magnetic stimulation study. J Neurophysiol 73(6):2608–2611

Fadiga L, Craighero L, Olivier E (2005) Human motor cortex excitability during the perception of others' action. Curr Opin Neurobiol 15(2):213–218

Frey S, Campball J, Pike G, Petrides M (2008) Dissociating human language pathways with high angular resolution diffusion fiber tractography. J Neurosci 28(45):11435–11444

Gallese V, Fadiga L, Fogassi L, Rizzolatti G (1996) Action recognition in the premotor cortex. Brain 119(2):593

Gastaut HJ, Bert J (1954) EEG changes during cinematographic presentation; moving picture activation of the EEG. Electroencephalogr Clin Neurophysiol 6(3):433

Gazzola V, Aziz-Zadeh L, Keysers C (2006) Empathy and the somatotopic auditory mirror system in humans. Curr Biol 16(18):1824–1829

Giard M, Peronnet F (1999) Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. J Cogn Neurosci 11(5):473–490

Grafton ST, Arbib MA, Fadiga L, Rizzolatti G (1996) Localization of grasp representations in humans by PET: 2. Observation compared with imagination. Exp Brain Res 112:103–111

Hari R, Forss N, Avikainen S, Kirveskari E, Salenius S, Rizzolatti G (1998) Activation of human primary motor cortex during action observation: a neuromagnetic study. Proc Natl Acad Sci 95(25):15061

Iacoboni M, Woods RP, Brass M, Bekkering H, Mazziotta JC, Rizzolatti G (1999) Cortical mechanisms of human imitation. Science 286:2526–2528

Iacoboni M, Koski LM, Brass M, Bekkering H, Woods RP et al (2001) Reafferent copies of imitated actions in the right superior temporal cortex. Proc Natl Acad Sci USA 98:13995–13999

Jousmäki V, Hari R (1998) Parchment-skin illusion: sound-biased touch. Curr Biol 8(6):190

Kaplan JT, Iacoboni M (2007) Multimodal action representation in human left ventral premotor cortex. Cogn Process 8(2):103–113

Keuken MC, Hardie A, Dorn BT, Dev S, Paulus MP, Jonas KJ et al (2011) The role of the left inferior frontal gyrus in social perception: An rTMS study. Brain Res 1383(6):196–205

Keysers C, Kohler E, Umiltà MA, Nanetti L, Fogassi L, Gallese V (2003) Audiovisual mirror neurons and action recognition. Exp Brain Res 153(4):628–636

Kohler E, Keysers C, Umilta MA, Fogassi L, Gallese V, Rizzolatti G (2002) Hearing sounds, understanding actions: action representation in mirror neurons. Science 297(5582):846

Laufs H, Kleinschmidt A, Beyerle A, Eger E, Salek-Haddadi A, Preibisch C et al (2003a) EEG-correlated fMRI of human alpha activity. Neuroimage 19(4):1463–1476

Laufs H, Krakow K, Sterzer P, Eger E, Beyerle A, Salek-Haddadi A et al (2003b) Electroencephalographic signatures of attentional and cognitive default modes in spontaneous brain activity fluctuations at rest. Proc Natl Acad Sci USA 100(19):11053

Le Bel RM, Pineda JA, Sharma A (2009) Motor-auditory-visual integration: The role of the human mirror neuron system in communication and communication disorders. J Commun Disord 42(4):299–304

Maeda F, Kleiner-Fisman G, Pascual-Leone A (2002) Motor facilitation while observing hand actions: specificity of the effect and role of observer's orientation. J Neurophysiol 87:1329–1335

Meredith MA, Stein BE (1986) Visual auditory and somatosensory convergence on cells in superior colliculus results in multisensory integration. J Neurophysiol 56:640–662

Moore A, Gorodnitsky I, Pineda J (2012) EEG mu component responses to viewing emotional faces. Behav Brain Res 226(1):309–316

Mukamel R, Ekstrom AD, Kaplan J, Iacoboni M, Fried I (2010) Single-neuron responses in humans during execution and observation of actions. Curr Biol 20:750–756

Muthukumaraswamy SD, Johnson BW, McNair NA (2004) Mu rhythm modulation during observation of an object-directed grasp. Cogn Brain Res 19(2):195–201

Neuper C, Wortz M, Pfurtscheller G (2006) ERD/ERS patterns reflecting sensorimotor activation and deactivation. Prog Brain Res 159:211–222

Pfurtscheller G (1992) Event-related synchronization (ERS): an electrophysiological correlate of cortical areas at rest. Electroencephalogr Clin Neurophysiol 83(1):62–69

Pfurtscheller G, Aranibar A (1979) Evaluation of event-related decynchronization (ERD) preceding and following voluntary self-paced movement. Electroencephalogr Clin Neurophysiol 46:138–146

Pfurtscheller G, Neuper C, Andrew C, Edlinger G (1997) Foot and hand area mu rhythms. Int J Psychophysiol 26:121–135

Pfurtscheller G, Neuper C, Krausz G (2000) Functional dissociation of lower and upper frequency mu rhythms in relation to voluntary limb movement. Clin Neurophysiol 111(10):1873–1879

Pineda JA (2005) The functional significance of mu rhythms: translating "seeing" and "hearing" into "doing". Brain Res Rev 50(1):57–68

Pineda JA (2008) Sensorimotor cortex as a critical component of an'extended'mirror neuron system: does it solve the development correspondence and control problems in mirroring? Behav Brain Funct 4(1):47

Rizzolatti G, Craighero L (2004) The mirror-neuron system. Annu Rev Neurosci 27:169–192

Rizzolatti G, Fadiga L, Gallese V, Fogassi L (1996a) Premotor cortex and the recognition of motor actions. Cogn Brain Res 3(2):131–141

Rizzolatti G, Fadiga L, Matelli M, Bettinardi V, Paulesu E et al (1996b) Localization of grasp representation in humans by PET: 1. Observation versus execution. Exp Brain Res 111:246–252

Rizzolatti G, Fogassi L, Gallese V (2001) Neurophysiological mechanisms underlying the understanding and imitation of action. Nat Rev Neurosci 2(9):661–670

Romanski LM, Tian B, Fritz J, Mishkin M, Goldman-Rakic PS, Rauschecker JP (1999) Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. Nat Neurosci 2(12):1131

Sauseng P, Klimesch W, Stadler W, Schabus M, Doppelmayr M, Hanslmayr S et al (2005) A shift of visual spatial attention is selectively associated with human EEG alpha activity. Euro J Neurosci 22(11):2917–2926

Skipper JI, Nusbaum HC, Small SL (2005) Listening to talking faces: motor cortical activation during speech perception. Neuroimage 25(1):76–89

Stein BE, Meredith MA (1993) The merging of the senses. The MIT Press, MA, USA

Thompson WF, Russo RA, Livingstone S (2010) Facial expressions of pitch structure in music performance. Psychon Bull Rev 17:317–322

Walden BE, Prosek RA, Montgomery AA, Scherr CK, Jones CJ (1977) Effects of training on the visual recognition of consonants. J Speech Hear Res 20:130–145

Welch RB, DutionHurt LD, Warren DH (1986) Contributions of audition and vision to temporal rate perception attention perception. Psychophysics 39(4):294–300