View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Abertay Research Portal

Event Segmentation and Biological Motion Perception in Watching Dance

(Abbreviated Title: Watching Dance)

Katie Noble, Donald Glowinski*, Helen Murphy, Corinne Jola°, Phil McAleer, Nikhil Darshane, Kedzie Penfield^, Sandhiya Kalyanasundaram[§], Antonio Camurri*, Frank E Pollick

University of Glasgow, Glasgow, G12 8QB UK

*University of Genoa, Genoa, 16145 Italy

°Neurospin, Gif-sur-Yvette Cedex 91191 France

^Queen Margaret University, Edinburgh, EH21 6UU UK

[§]Sangam Arts, San Jose USA

Address for correspondence:
Frank E Pollick
School of Psychology
University of Glasgow
Glasgow, G12 8QB
United Kingdom
Phone: +44 (0)141 330 3945
Fax: +44 (0)141 330 4606

Email: frank.pollick@glasgow.ac.uk

No conflict of interest

KEYWORDS:

Dance, biological motion, event segmentation, extrastriate body area (EBA), fMRI

Abstract

We used a combination of behavioral, computational vision and fMRI methods to examine human brain activity while viewing a 386 second video of a solo Bharatanatyam dance. A computational analysis provided us with a Motion Index (MI) quantifying the silhouette motion of the dancer throughout the dance. A behavioral analysis using 30 naïve observers provided us with the time points where observers were most likely to report event boundaries where one movement segment ended and another began. These behavioral and computational data were used to interpret the brain activity of a different set of 11 naïve observers who viewed the dance video while brain activity was measured using fMRI. Results showed that the Motion Index related to brain activity in a single cluster in the right Inferior Temporal Gyrus (ITG) in the vicinity of the Extrastriate Body Area (EBA). Perception of event boundaries in the video were related to the BA44 region of right Inferior Frontal Gyrus as well as extensive clusters of bilateral activity in the Inferior Occipital Gyrus which extended in the right hemisphere towards the posterior Superior Temporal Sulcus (pSTS).

INTRODUCTION

Most studies into the mechanisms of human movement perception have relied extensively on short videos of isolated body parts or representations of the whole body by points of lights (Grosbras, Beaton, & Eickhoff, 2012). Exploiting the simplicity of these displays, where precise control of motion can be obtained, advances have been made in revealing the neural mechanisms of action observation (Blake & Shiffrar, 2007). However, important questions remain about how viewing extended activity drives brain response and the processes involved in perceptually organizing action sequences. This is particularly important as evidence comparing brain activity during naturalistic viewing to brain activity when viewing short epochs has revealed more brain activity for naturalistic viewing (Bartels & Zeki, 2004). One possible reason advanced for this increased activity has been that the human brain is optimized for natural viewing (Bartels & Zeki, 2005). To address these issues we used fMRI to examine neural mechanisms involved while viewing a naturalistic video of an extended dance performance. Analysis related brain activity to a novel measure of the motion of the silhouette of the dancer (Glowinski, Camurri, Chiorri, Mazzarino & Volpe, 2009) and probed brain activity at times when the dance was perceived to have one event end and another begin (Zacks & Tversky, 2001). These measures focus respectively on the motion of the boundary that separates the dancer from the background and the temporal discontinuities that separate one part of a movement from another.

Dance provides a stimulus that has increasingly been found to be appropriate to study how human movement is processed (Blasing, et al., 2012; Christensen & Calvo-Merino, 2013; Calvo-Merino, Grezes, Glaser, Passingham, & Haggard, 2006; Cross, Hamilton, & Grafton, 2006; Reynolds, Jola, & Pollick, 2011). Dance also has structured patterns of movements that provides both a complex pattern of continuous motion as well as a set of discrete segments (deLahunta, Barnard, Nimmo-Smith, Potts, & Ramponi, 2006). We chose

the classical Indian dance style of Bharatanatyam dance since it has a strong narrative that is rich in gesture and affective movement (Hejmadi, Davidson, & Rozin, 2000), but the stories depicted and cultural significance are unfamiliar to western observers (Jola, Abedian-Amiri, Kuppuswamy, Pollick, & Grosbras, 2012; Jola & Grosbras, 2013). A lack of familiarity does not, however, necessarily imply a complete lack of understanding movement. For example, non-signers show competence in understanding the structure of a narrative given by sign language without grasping the meaning (Fenlon, Denmark, Campbell & Woll, 2007).

The cognitive capacity to perceive a continuous stream of activity in distinct parts (particularly beginning and end) is termed *event structure perception* (Zacks & Tversky, 2001), and it has been proposed that this segmentation into events is automatic and has specialized neural mechanisms that detect event boundaries by tracking social and physical changes. Evidence for neural mechanisms that detect event boundaries stems from Zacks, et al.'s (2001) study that revealed activity in right prefrontal cortex in the vicinity of the Frontal Eye Fields (FEF) as well as motion sensitive area hMT+. Further evidence for activity in hMT+ at the time of event boundaries was seen in several subsequent studies along with activity in the posterior temporal sulcus (pSTS) (Speer, Swallow, & Zacks, 2003; Zacks, Swallow, Vettel, & McAvoy, 2006). These areas of pSTS and hMT+ have been implicated in biological motion processing (Blake & Shiffrar, 2007; Grosbras, et al., 2012) along with the extrastriate body area (EBA), located adjacent and overlapping to hMT+ (Downing & Peelen, 2011; Peelen, Wiggett, & Downing, 2006; Ferri, Kolster, Jastorff, & Orban, 2013).

Activity in these regions of temporal cortex have also been shown to correlate between observers when watching Bharatanatyam dance (Jola, McAleer, Grosbras, Love, Morison & Pollick, 2013). This study used intersubject correlation (Hasson, Nir, Levy, Fuhrmann & Malach, 2004) to compare brain activity during uni- (audio and visual) and multisensory (audiovisual) presentation of dance and music and showed a general

enhancement of correlation for audiovisual stimuli. In the current research we more closely examine the data from the vision-only condition of Bharatanatyam dance of this previous study. We do this to explore both how action segmentation and motion perception could have contributed to the previously reported intersubject correlations of brain activity as well as whether any possible neural correlates of action segmentation and motion perception might have been missed by intersubject correlation analysis.

MATERIALS AND METHODS

Dance Stimulus and Calculation of Motion Index

Dance Video. The stimulus material was a standard PAL format recording of a classical Indian Bharatanatyam dance performed by a single female performer with traditional costume and performed in a dance studio against a plain grey background without any scenery. The Bharatanatyam dance was chosen in order to reduce the likelihood that the participants had any prior familiarity with the type of dance chosen for stimuli. The duration of the dance movie was 386 seconds. Details of the dance performance and the accompanying music and vocals are given in Jola et al. (2013). Although none of our participants heard any music or vocals during the scan it is still possible that the metrical structure of the music might have influenced the dance (Toiviainen, Luck & Thompson, 2013), thus we will summarize these aspects of the performance. The Carnatic music had a four-beat rhythm cycle that had duration of 8 seconds, with each cycle containing one line of the poem being sung in Tamil by a solo voice. In the original version (dance performed to music), the singing begins approximately 90 seconds into the movie and continued to the conclusion of the dance. Notably, the choreography of the dancer is fundamentally to the lyrics of the story being told rather than to the music.

Calculation of Motion Index. In order to obtain the Motion Index (MI) of the Indian dancer a sequence of image processing steps was required. These were all performed using the software tool EyesWeb (Camurri, De Poli, Leman, & Volpe, 2005; Camurri, Trocca, & Volpe, 2002; Volpe & Camurri, 2011). The Motion Index was computed as the area (i.e., number of pixels) of a Silhouette Motion Image, an image carrying information about the difference in area between individual silhouettes separated by 5 frames (167msec). Following Glowinski et al. (2009), a preliminary step consisted of motion segmentation, which aims at detecting the moving regions corresponding to the dancer (the silhouette). Motion segmentation included background subtraction, which consists of differencing between the current image and a reference background image. Other techniques to aid motion segmentation included techniques such as temporal differencing (frame differencing) and color extraction to handle inconsistencies due to several sources of noise including 1) noise in the background image; 2) lighting artifacts due to shadows and variable ambient illumination; and 3) overlapping of parts of the body ("internal motion", i.e., motion of parts of the body, e.g. arms, overlapped in the silhouette), and 4) variability in the conversion from a color image to grayscale for the background subtraction process. The result of this process was a single scalar value for each frame of video that reflected the total amount of silhouette motion in that frame. Since we were not interested in the rapid variations in MI, values were averaged over two seconds since this temporal granularity was both an effective histogram bin width in analysis of participant event segmentation judgments in the behavioral study and also matched the 2-second temporal resolution of the functional imaging data. The MI for the entire dance, after z-transformation is shown in Figure 1a.

Behavioral experiment to find event boundaries

To investigate where participants perceived the event boundaries in the dance performance we performed a behavioral Event Segmentation experiment. The participants for this experiment were different from those used in the fMRI scans. The results of this experiment informed, on average, when the boundaries occurred and were used in the fMRI experiment to examine brain activity around the time of the boundary. We chose to use group results to define event boundaries for fMRI data analysis since pilot data suggested individual responses did not reveal stable boundaries, and Zacks et al. (2006) suggest that group-level boundaries are preferable to use in this case.

Participants. Thirty participants of either sex, aged between 18 and 27, were recruited from the University of Glasgow Participant Pool. Participants were offered either course credit or a small cash payment for their participation. All were considered novices in their experience of dance in that they were not trained dance professionals, nor frequent viewers of professional dance.

Procedure. Participants were run individually in a quiet and dimly lit booth. Displays were shown without sound on a computer monitor under the control of custom software written using Presentation V11 (Neurobehavioral Systems, CA). Participants were instructed that the experimental task was to segment a dance movie into *events*: each of which can be defined as "a segment of time at a given location that is perceived to have a beginning and an end" (Zacks & Tversky, 2001). They were told they would watch a practice dance movie once and an experimental dance movie twice. The practice dance movie was presented prior to the experimental sequence in order to allow participants to familiarize themselves with the segmentation task before starting the experiment proper. Participants were instructed to press the 's' key on the computer keyboard whenever they determined an event boundary. When participants were comfortable that they fully understood the task they proceeded to the next phase. For the experimental movie participants watched the movie

twice, the first time the instruction was to ‘watch only’; the second time the instruction was to ‘segment’ by pressing the ‘s’ key.

fMRI Experiment

The brain imaging study involved participants viewing the Bharatanatyam dance movie while being scanned. Just as in the behavioral event segmentation experiment, the movie was observed without any sound. Participants were instructed simply to watch the movie and enjoy the performance, and a video link of the eye enabled us to monitor that they appeared alert for the entire scanning session. The fMRI experiment was performed on a separate set of individuals than the behavioral study. The participants were the same as used by the Jola et al. (2013) study; in fact the data analysed in this paper were from the vision-only condition reported in that paper. Analyses examined which brain areas covaried in activity with the MI and what brain activity could be related to the perception of event boundaries.

Participants. Twelve participants (6 male, mean age = 22.3, SD = 2.8) were recruited from the University of Glasgow Participant Pool. One female participant coughed during scanning and was excluded due to excessive head movement. All participants reported being right-handed and to have no history of neurological conditions. They were paid £30 for their participation. Ethics approval for the study was obtained from the Ethics Committee of the Faculty of Information and Mathematical Sciences, University of Glasgow.

fMRI data acquisition. Participants were scanned using a 3T Siemens Tim Trio MRI scanner (Erlangen, Germany). Anatomical data of the whole brain was acquired using a 3D magnetization prepared rapid acquisition gradient echo (MP-RAGE) T1-weighted sequence

(192 slices; 1mm cube isovoxel; Sagittal Slice; TR = 2.52; 256×256 image resolution). Functional scans obtained 32 Sagittal slices over nearly the entire brain volume (Echo Planar 2D imaging PACE; MoCo on; TR = 2000ms; TE = 30MS; 32 Sagittal Slices; 3mm cube isovoxel; 70×70 image resolution). The 386 seconds of the movie were contained in 193 volumes and were proceeded and ended by 5 volumes (10 seconds) of blank screen. The scanning session included the scans described here as well as others that comprised a separate experiment.

fMRI data analysis. Analysis was performed using BrainVoyager QX V2.3 (Brain Innovation B. V., Maastricht, Netherlands).

Pre-processing of data: Functional image data pre-processing included performing a slice scan time correction using sinc interpolation based on information about the TR (2000 ms) and the order of slice scanning (ascending, interleaved). In addition, 3-D motion correction (6df) was performed to detect and correct for small head movements by spatial alignment of all volumes of a subject to the first volume by rigid body transformations. The estimated translation and rotation parameters were inspected and never exceeded 3 mm or 3° (except for the previously mentioned participant who was removed due to coughing). A linear trend removal and temporal high-pass filtering with a cutoff of 0.0063 Hz were then applied. The functional MR images were smoothed using a Gaussian filter with full-width at half-maximum (FWHM) equal to 6 mm.

The anatomical data (DICOM format) of each subject were transformed into anterior commissure-posterior commissure (AC-PC) and Talairach standard space using a single transformation to avoid quality loss due to successive data sampling. The anatomical data for a participant were then aligned to the functional data that were similarly transformed into Talairach standard space.

Analysis relating Motion Index (MI) to brain activity: In order to test whether modulations in the Motion Index related to changes in brain activity reflected in the BOLD signal we introduced a parametric coding of the Motion Index into our general linear model. To achieve this, the z-transformed MI (Figure 1a) was convolved with a modeled hemodynamic response function to provide a predictor for activity at each of the 193 timepoints during the video. This model was applied to the whole brain and a criterion of $p< 0.001$, uncorrected was used to determine regions of significant activation. This was corrected for multiple comparisons using cluster-size thresholding (Forman, et al., 1995; Goebel, Esposito, & Formisano, 2006) and voxels surviving a criterion of $p<0.05$ are reported.

Analysis relating event boundaries to brain activity: The times of the 30 event boundaries derived from the behavioral experiment were used to probe the brain data (Zacks, et al., 2001). A random effects GLM, applied to the whole brain, was used to estimate the BOLD response over time courses of 22 second duration that included 4 seconds before an event boundary and 18 seconds after. This analysis identified brain regions that showed a reliable transient change in activity associated with the event boundaries. Data were analyzed across participants using a general linear model with the 12 levels of time at 2 second intervals serving as stick predictors over the 30 event boundaries revealed by the behavioral experiment. We consider only the changes where the BOLD signal moved positive relative to baseline. A criterion of $p<0.001$, uncorrected was used to determine regions of significant activation. This was corrected for multiple comparisons using cluster-size thresholding (Forman, et al., 1995; Goebel, et al., 2006) and voxels surviving a criterion of $p<0.05$ are reported. A region of interest analysis was also performed on the cluster found in the MI analysis using the same design applied to the whole brain.

RESULTS

Behavioral Results. Each participant had a set of recorded times for their key presses at perceived event boundaries while watching the movie. These times of event boundaries for each participant were combined into a histogram using bin durations of two-seconds (Zacks, 2004) and results are shown in Figure 1b. A threshold of one standard deviation of the average button press frequency was used to define event boundaries and is shown as the dashed line. In the situation where a bin defining a boundary consecutively followed another the more significant was chosen, if they were equal the first boundary was used and the second discarded. This resulted in a total of 30 boundaries and an average segment duration of 12.8 s (SD = 3.5). Examination of the activity at these boundaries revealed a variety of actions including kneeling down, standing up, head and arm movements as well as stationary postures. Examination of the histogram revealed that even the highest peak had only one-third of participants marking it as a boundary suggesting limited agreement among the participants. To evaluate whether the number of boundaries was reasonable we had a Bharatanatyam dancer and a movement analyst evaluate the dance, and at a coarse level of analysis they found 32 and 34 event boundaries respectively. While there was still variability in the placement of the boundaries, these results support that the number of boundaries found is appropriate.

fMRI Results. The whole brain analysis examining the effect of MI as a modulation parameter is shown in Figure 2a. This analysis revealed a single cluster with peak activation in the right Inferior Temporal Gyrus (rITG) at Talairach coordinates 47, -71, 0 ($F=4.879$, $p= 0.000672$, cluster size= 248 mm^3, Brodmann Area 37). The whole brain analysis examining the effect of event boundaries is shown in Figure 2b-c. This analysis returned three clusters (Table 1) with peak voxels located in the right Inferior Occipital Gyrus (IOG), the left Middle

Occipital Gyrus (lMOG) and the right Inferior Frontal Gyrus (rIFG). Comparison of Figures 2a and 2b reveal that the voxels involved with event boundaries encompass those related to the MI. A final region of interest analysis examined whether there was a significant effect of event segmentation within the area found in rITG defined by the analysis of MI covariation with brain activity. Results, revealed a significant effect (df =10, t= 3.638, p = 0.00455) appearing as an increase in activity that peaked 2 seconds after the time of the event boundary.

DISCUSSION

We explored the perception of human movement over an extended movement sequence. Our results showed that activation at the right Inferior Temporal Gyrus was related to a novel Motion Index that quantified motion as the global silhouette change of a dancer. In addition, times at event boundaries of the dance were related to activation in the right Inferior Frontal Gyrus and bilaterally in the occipitotemporal cortex. The region sensitive to the Motion Index in right hemisphere was incorporated in a larger right occipitotemporal activation associated with event segmentation. Relating these results to Jola et al. (2013) indicate that the extended correlation of brain activity among observers reported in right temporal cortex could be driven by mechanisms of event segmentation and motion processing of the silhouette. These results, using a continuous performance rather than short discrete actions, advance our knowledge of the brain mechanisms behind visual representations of human motions and their transformation into more elaborate action representations.

The activation associated with the Motion Index found at Talairach coordinates (47, -71, 0) is quite similar to the coordinates of (44, -71, 3) revealed as a peak of correlated brain activity between observers while viewing the same dance (Jola et al., 2013). This right hemisphere region was also identified by a recent meta-analysis to have the greatest

probability of being found in experiments on body movements (Grosbras, et al., 2012). In the left hemisphere this region (-45, -64, 1) has been reported to be a primary node in the processing of biological motion, demonstrating functional connectivity with occipital and parietal regions (McKay, Simmons, McAleer, Marjoram, Piggot & Pollick, 2012). This right hemisphere activation is also quite similar to coordinates (48, -70, 0) previously reported for localization of the EBA (Downing, Peelen, Wiggett, & Tew, 2006). It is known that the EBA is sensitive to photorealistic depictions of the whole body but previous evidence suggests that this representation is not dynamic (Downing & Peelen, 2011) and although the EBA is near to motion area hMT+, experiments using multivoxel pattern analysis reported distinct areas of motion and body shape sensitivity between these areas (Peelen, et al., 2006). However, it has been suggested that the response to biological motion within EBA reflects an integration of form and motion (Jastorff & Orban, 2009) and evidence from Thompson and Baccus (2012) suggest independent neural populations within EBA for the processing form and motion. While it is possibly not surprising that the EBA would be implicated by a motion index that represents silhouette changes of a moving body, the activity we report is consistent with a joint calculation between shape sensitive neurons to define a body contour and motion sensitive neurons to calculate the global changes of this contour. Such cooperative mechanisms have previously been suggested (Downing & Peelen, 2011) and could be crucial for functions that use motion to segregate form from background such as how motion can break camouflage.

As most studies of human perception have been based on point light displays little is known about the perception of silhouettes of the body in motion. It is known theoretically that a static silhouette contains rich information about form (Koenderink, 1990) and that time-varying deformations of silhouettes of even simple shapes can provide a vivid impression of three-dimensional form (Pollick, 1994). The computational recovery of form

from changing sihouettes involves resolving an ambiguity between motion and form to account for how each is simultaneously contributing to changes in the image (Giblin, Pollick, Rycroft, 1994). While such ambiguity is fundamental to many visual processes (Koenderink & Van Doorn, 1997) it is acute in the case of silhouettes where local deformation of the silhouette depends upon motion and the curvature of the surface in complex ways. Our calculation of a Motion Index does not decompose changes of the silhouette into different factors of form and motion (Tomasi & Kanade, 1992; Cadieu & Olshausen, 2012). Thus, the brain activity related to our motion index might be due to motion features, form features or the process of factorizing the image motion into form and motion features. The pattern of activations associated with our Motion Index also has parallels with the model proposed by Lange and Lappe (2006). In Stage 1 of this model body postures are obtained, for example through the activity of snapshot neurons that represent body postures by the activity of neurons with large receptive fields that encompass the entire body (Giese & Poggio, 2003). In Stage 2 of processing the difference between postures is calculated and this differencing is what our Motion Index calculation achieves through the difference between consecutive silhouettes. In this way our reported activations are consistent with Stage 2 processing.

The activation associated with event segmentation revealed extensive bilateral regions in occipitotemporal cortex. These regions are similar to those reported previously in event segmentation (Speer, et al., 2003; Zacks, et al., 2001; Zacks, et al., 2006) and also consistent with the intersubject correlation results of Jola et al. (2013). The right hemisphere peak of activation for event segmentation was near to the peak found for the Motion Index, however the event segmentation cluster was much larger, overlapping the Motion Index cluster and extending towards superior regions of the temporal cortex. This finding suggests that while activity related to the Motion Index contributed to segmentation it was not the sole factor driving segmentation. This could explain why there was not a clear relationship between

Motion Index and the pattern of segmentation illustrated in Figure 1. Because our Motion Index reflected global motion of the entire body we can speculate that the additional activation with event segmentation related to more local (i.e. body part) aspects of the motion. Consistent with this, recent findings have revealed brain regions dorsal to our Motion Index cluster that have activity correlated to the hand speed of actors sitting at a table and performing everyday activities (McAleer, Pollick, Love, Crabbe & Zacks, in press). Activity in these dorsal regions can be related to results from monkey where it has been shown that the response of individual neurons in temporal cortex are tuned to individual movement segments (Vangeneugden, Pollick & Vogels, 2009).

The activation associated with event segmentation was also found in the right inferior frontal gyrus. However, unlike the occipitotemporal result, activation in BA44 has not been reported previously in event segmentation tasks although more dorsal regions of prefrontal cortex have been reported. This difference is significant since BA44 has been implicated in various aspects of action understanding (Rizzolatti & Fabbri-Destro, 2008). In particular, BA44 has been shown to be active in the segmentation of viewed action sequences at the level of movement chunks (Koechlin, Ody, & Kouneiher, 2003; Koechlin & Summerfield, 2007). In this theoretical framework the prefrontal cortex is arranged hierarchically for processing actions with anterior regions processing more abstract aspects of motor plans and posterior regions containing increasingly specific aspects of the motor plan. The activation of BA44 with segmentation of the dance into discrete events is thus consistent with the segmentation occurring at a coarse grain that separates the various chunks/phrases of the dance.

Our results point to a right hemisphere frontotemporal circuit that incorporates analysis of dynamic changes in body movement in temporal cortex with hierarchical action representations in prefrontal cortex. Although a lack of finding activations in left hemisphere

might be due to a lack of statistical power, the findings are consistent with several reports of right hemisphere dominance in the processing of actions (Downing & Peelen, 2011; Grosbras, et al., 2012). Further evidence of a right hemisphere circuit come from evidence of anatomic connectivity between temporal (BA37) and prefrontal (B44) cortices (Glasser & Rilling, 2008) that have been implicated in the processing of speech prosody (Meyer, Alter, Friederici, Lohmann, & von Cramon, 2002). Given the similarities between modulation of speech in prosody and the modulation of activity indicated by the Motion Index there are potential shared functions. In particular, prosody requires analysis over long durations (Hickok & Poeppel, 2004) and is effective at signaling boundaries. It has also been shown that a form of visual prosody can be obtained from head movements (Munhall, Jones, Callan, Kuratate, & Vatikiotis-Bateson, 2004).

Why this particular dance would visually contain prosodic information is an open question. As mentioned previously, it has been shown that metrical structure of music can influence movement (Toiviainen, et al., 2010), so that is a possibility. However, in this dance the emphasis of the choreography is to the lyrics rather than the music. The dance choreography includes highly codified facial and hand gestures, which express symbolic meaning contained in the lyrics and this suggests that visual prosody would arise from alignment of expressive dance movement to the prosody and meaning of the lyrics. This expressive aspect of movement was not present in previous studies of event segmentation and this could explain why a similar result was not reported in these previous studies. An observation to guide future research is that since observers had no familiarity with Tamil language, Carnatic music, mythological content or Bharatanatyam dance, that whatever the mechanisms involved in encoding this hypothesized visual prosody, their interpretation invokes a general mechanism that does not rely on culture-specific processes.

Acknowledgements

We would like to thank Karin Petrini, Lawrie McKay, Scott Love and members of the MARCS lab of the University of Western Sydney for their helpful suggestions. We thank the AHRC for support of the project 'Watching Dance: Kinesthetic Empathy'. DG and AC were supported by the EU FP7 ICT FET SIEMPRE Project No. 250026 (www.siempre.infomus.org). Finally, we thank Matthew Reason for providing the dance video and to Anna Kuppuswamy for her dance performance.

References

Bartels, A., & Zeki, S. (2004). The chronoarchitecture of the human brain—natural viewing conditions reveal a time-based anatomy of the brain. Neuroimage, *22*(1), 419–433.

Bartels, A., & Zeki, S. (2005). The chronoarchitecture of the cerebral cortex. Philos T Roy Soc B, *360*(1456), 733–750.

Blake, R., & Shiffrar, M. (2007). Perception of human motion. Annu Rev Psychol, *58*, 47-73.

Blasing, B., Calvo-Merino, B., Cross, E. S., Jola, C., Honisch, J., & Stevens, C. J. (2012). Neurocognitive control in dance perception and performance. Acta Psychol, *139*(2), 300-308.

Cadieu, C. F., & Olshausen, B. A. (2012). Learning intermediate-level representations of form and motion from natural movies. Neural Comput, *24*(4), 827–866.

Calvo-Merino, B., Grezes, J., Glaser, D. E., Passingham, R. E., & Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. Curr Biol, *16*(19), 1905-1910.

Camurri, A., De Poli, G., Leman, M., & Volpe, G. (2005).“Toward Communicating Expressiveness and Affect in Multimodal Interactive Systems for Performing Art and Cultural Applications”, IEEE Multimedia, Vol.12, No.1, pp.43-53, IEEE Computer Society Press.

Camurri, A., Trocca, R., & Volpe, G. (2002). Interactive systems design: A KANSEI-based approach. In Proceedings of the 2002 Conference on New Interfaces for Musical Expression (pp. 155-162). limerick, Ireland: University of limerick, Department of Computer Science and Information Systems.

Christensen, J.F., Calvo-Merino, B. (2013). Dance as a subject for empirical aesthetics. Psychol Aesthet Crea, 7, 76-88.

Cross, E. S., Hamilton, A. F. D. C., & Grafton, S. T. (2006). Building a motor simulation de novo: Observation of dance by dancers. Neuroimage, *31*(3), 1257-1267.

deLahunta, S., Barnard, P., Nimmo-Smith, I., Potts, J., & Ramponi, C. (2006). Densities of agreement: Making visible some intangible properties of dance. Dance Theat J, *21*(3), 17-23.

Downing, P. E., & Peelen, M. V. (2011). The role of occipitotemporal body-selective regions in person perception. Cogn Neurosci, *2*(3-4), 186-203.

Downing, P. E., Peelen, M. V., Wiggett, A. J., & Tew, B. D. (2006). The role of the extrastriate body area in action perception. Soc Neurosci, *1*(1), 52-62.

Fenlon, J., Denmark, T., Campbell, R., & Woll, B. (2007). Seeing sentence boundaries. Sign Lang Linguist, *10*(2), 177-200.

Ferri, S., Kolster, H., Jastorff, J., & Orban, G. A. (2013). The overlap of the EBA and the MT/V5 cluster. Neuroimage, *66*(C), 412–425.

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved Assessment of Significant Activation in Functional Magnetic-Resonance-Imaging (Fmri) - Use of a Cluster-Size Threshold. Magn Reson Med, *33*(5), 636-647.

Giese, M. A., & Poggio, T. (2003). Neural mechanisms for the recognition of biological movements. Nat Rev Neurosci, *4*(3), 179-192.

Glasser, M. F., & Rilling, J. K. (2008). DTI Tractography of the Human Brain's Language Pathways. Cereb Cortex, *18*(11), 2471-2482.

Glowinski D., Camurri, A., Chiorri C., Mazzarino, B., Volpe, G. (2009). Validation of an algorithm for segmentation of full-body movement sequences by perception: a pilot experiment, in Sales Dias, M.; Gibet, S.; Wanderley, M.M.; Bastos, R. (Eds.),

Gesture-Based Human-Computer Interaction and Simulation, Lecture Notes in Computer Science, Volume 5085, p. 239-244, Springer-Verlag Berlin / Heidelberg, 2009.

Goebel, R., Esposito, F., & Formisano, E. (2006). Analysis of Functional Image Analysis Contest (FIAC) data with BrainVoyager QX: From single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. Hum Brain Mapp, *27*(5), 392-401.

Grosbras, M. H., Beaton, S., & Eickhoff, S. B. (2012). Brain regions involved in human movement perception: A quantitative voxel-based meta-analysis. Hum Brain Mapp, *33*(2), 431-454.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. Science, *303(5664)*, 1634-1640.

Hejmadi, A., Davidson, R. J., & Rozin, P. (2000). Exploring Hindu Indian emotion expressions: Evidence for accurate recognition by Americans and Indians. Psychol Sci, *11*(3), 183-187.

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. Cognition, *92*(1-2), 67-99.

Jastorff, J., & Orban, G. A. (2009). Human Functional Magnetic Resonance Imaging Reveals Separation and Integration of Shape and Motion Cues in Biological Motion Processing. J Neurosci, *29*(22), 7315-7329.

Jola, C., Abedian-Amiri, A., Kuppuswamy, A., Pollick, F. E., & Grosbras, M. H. (2012). Motor Simulation without Motor Expertise: Enhanced Corticospinal Excitability in Visually Experienced Dance Spectators. PLoS One, *7*(3).

Jola, C., McAleer, Ph., Grosbras, M.-H., Love, S.A., Morison, G., Pollick, F.E. (2013). Uni- and multisensory brain areas are synchronised across spectators when watching unedited dance. IPerception.

Jola, C. & Grosbras, M.-H. (2013). In the here and now: enhanced motor corticospinal excitability in novices when watching live compared to video recorded dance. Cogn Neurosci, *4*, 90-98.

Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. Science, *302*(5648), 1181-1185.

Koechlin, E., & Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. Trends Cogn Sci, *11*(6), 229-235.

Koenderink, J.J. (1990). Solid Shape. MIT Press.

Koenderink, J. J., & Van Doorn, A. J. (1997). The generic bilinear calibration-estimation problem. Int J Comput Vision, *23*(3), 217–234.

Lange, J., & Lappe, M. (2006). A model of biological motion perception from configural form cues. J Neurosci, *26*(11), 2894-2906.

McAleer, P., Pollick, F.E., Love, S.A., Crabbe, F. & Zacks, J. (in press). The role of kinematics in cortical regions for continuous human motion perception. Cogn Affect Behav Ne,

McKay, L.S., Simmons, D.R., McAleer, P., Marjoram, D., Piggot, J. & Pollick, F.E. (2012). Do distinct atypical cortical networks process biological motion information in adults with Autism Spectrum Disorders? Neuroimage, *59(2)*, 1524-1533.

Meyer, M., Alter, K., Friederici, A. D., Lohmann, G., & von Cramon, D. Y. (2002). FMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. Hum Brain Mapp, *17*(2), 73-88.

Munhall, K. G., Jones, J. A., Callan, D. E., Kuratate, T., & Vatikiotis-Bateson, E. (2004). Visual prosody and speech intelligibility - Head movement improves auditory speech perception. Psychol Sci, *15*(2), 133-137.

Peelen, M. V., Wiggett, A. J., & Downing, P. E. (2006). Patterns of fMRI activity dissociate overlapping functional brain areas that respond to biological motion. Neuron, *49*(6), 815-822.

Reynolds, D., Jola, C., & Pollick, F. E. (2011). Dance Research Electronic-Introduction Dance and Neuroscience-New Partnerships. Dance Res, *29*(2), 260-269.

Rizzolatti, G., & Fabbri-Destro, M. (2008). The mirror system and its role in social cognition. Curr Opin Neurobiol, *18*(2), 179-184.

Speer, N. K., Swallow, K. M., & Zacks, J. M. (2003). Activation of human motion processing areas during event perception. Cogn Affect Behav Ne, *3*(4), 335-345.

Thompson, J. C., & Baccus, W. (2012). Form and motion make independent contributions to the response to biological motion in occipitotemporal cortex. Neuroimage, *59*(1), 625–634.

Toiviainen, P., Luck, G. & Thompson, M. (2010). Embodied meter: hierarchical eigenmodes in music-induced movement. Music Percept, *28(1)*, 59-70.

Tomasi, C., & Kanade, T. (1992). Shape and motion from image streams under orthography: a factorization method. Int J Comput Vision, *9*(2), 137–154.

Vangeneugden, J., Pollick, F., & Vogels, R. (2009). Functional Differentiation of Macaque Visual Temporal Cortical Neurons Using a Parametric Action Space. Cereb Cortex, *19(3)*, 593-611.

Volpe, G. & Camurri A. (2011) A system for embodied social active listening to sound and music content. ACM J Comput Cult Herit, *4*(1), pp.2-23.

Zacks, J. M. (2004). Using movement and intentions to understand simple events. Cognitive Sci, *28*(6), 979-1008.

Zacks, J. M., Braver, T. S., Sheridan, M. A., Donaldson, D. I., Snyder, A. Z., Ollinger, J. M., et al. (2001). Human brain activity time-locked to perceptual event boundaries. Nat Neurosci, *4*(6), 651-655.

Zacks, J. M., Swallow, K. M., Vettel, J. M., & McAvoy, M. P. (2006). Visual motion and the neural correlates of event perception. Brain Res, *1076*, 150-162.

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and cognition. Psychol Bull, *127*(1), 3-21.

Figure Legends

Figure 1. A) The z-score transformation of the Motion Index of the dancer for the time course of the entire dance. Each plot point represents the average motion index averaged over two seconds. Note, that the extreme negative z-scores correspond to the dancer at the start and end when the Motion Index was near zero. B) A histogram of button press frequencies for the 30 participants asked to identify segment boundaries. The bin width is two seconds and the dashed line shows the threshold of 1 standard deviation used as a criterion to define a boundary. Comparison of panels A and B reveals that a fraction of the event boundaries are aligned with peaks or troughs of the Motion Index.

Figure 2. The results of the brain imaging experiment. Panel A shows the inferior temporal gyrus brain area that was found to covary with the motion index of the dancer. Panel B shows bilateral brain areas in the inferior occipital gyrus, which were activated at the times of event boundaries. Panel C shows the inferior frontal gyrus brain area that was also activated at the times of event boundaries.

Table 1.

Brain activations related to event segmentation. The location, size and statistic are given for each brain area showing a main effect of time relative to the event boundaries.

Area	Hemisphere	Talairach Coordinates	BA	Size (Voxels)	Statistic t	p
Inferior Occipital Gyrus	Right	44, -71, -6	19	17419	6.4068	0.000279
Middle Occipital Gyrus	Left	-43, -77, -9	18	6686	5.7892	0.000327
Inferior Frontal Gyrus	Right	56, 16, 18	44	307	5.4786	0.000413

Table 1

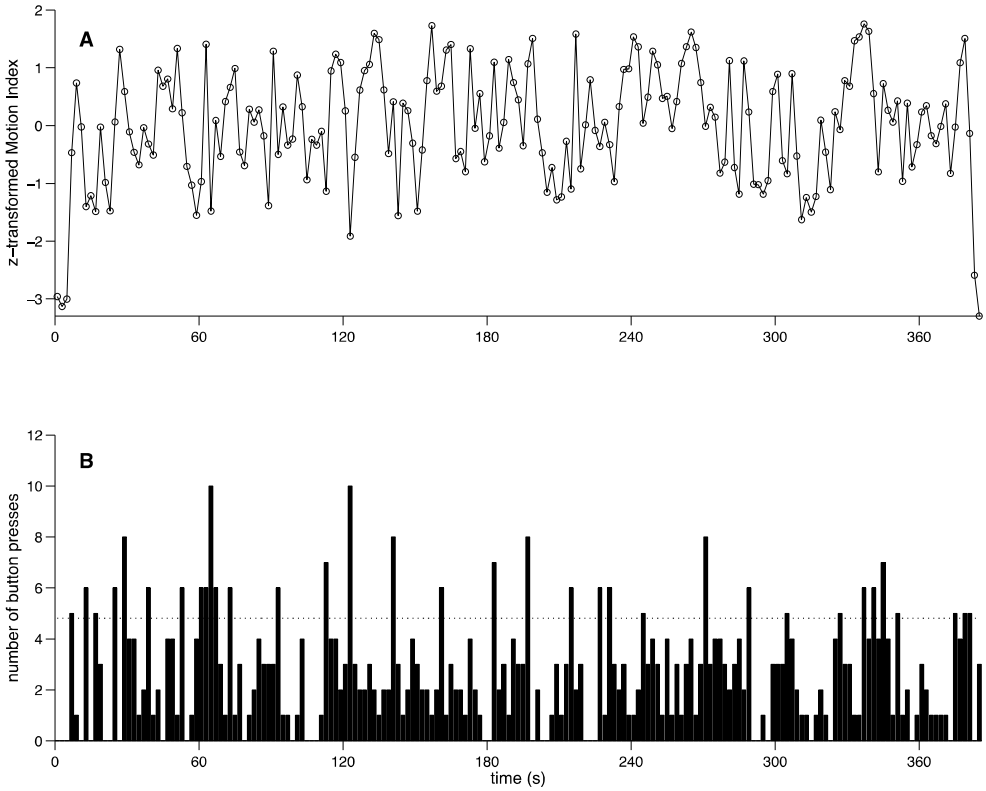
A
z−transformed Motion Index
2
1
0
−1
−2
−3
0
60
120
180
240
300
360
B
number of button presses
12
10
8
6
4
2
0
0
60
120
180
240
300
360
time (s)

Figure 1

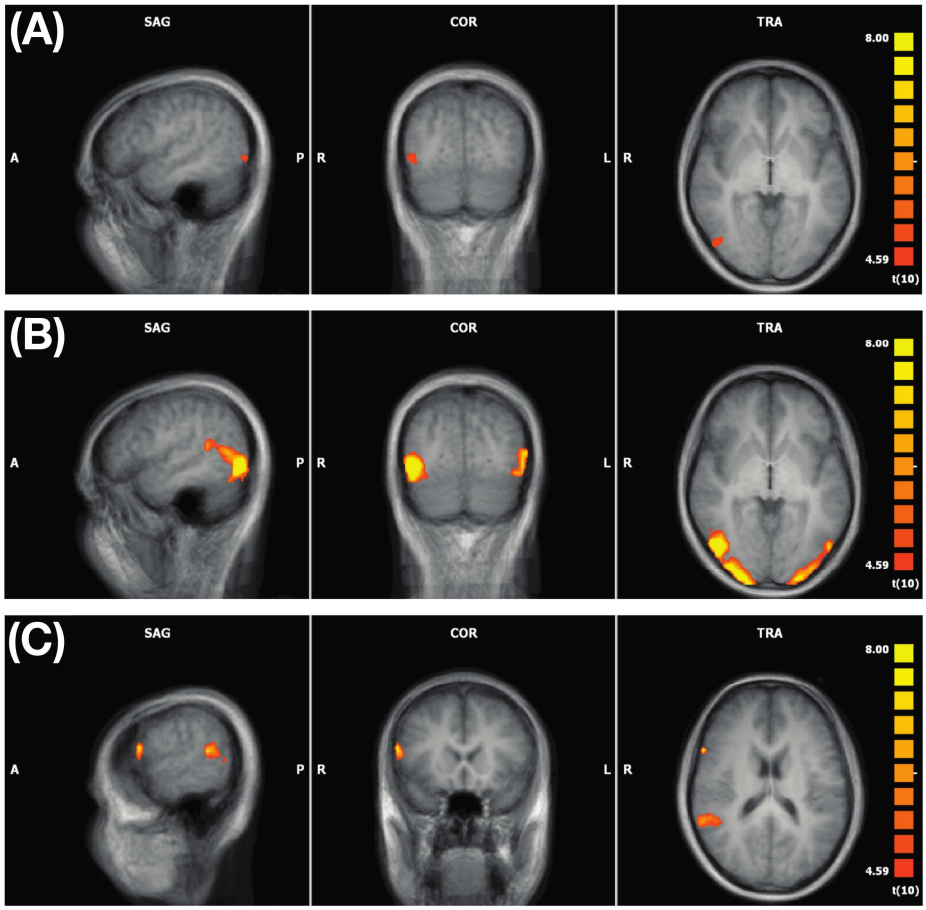
(A)
SAG
COR
TRA
A
P
R
L
R
8.00
4.59
t(10)
(B)
SAG
COR
TRA
A
P
R
L
R
8.00
4.59
t(10)
(C)
SAG
COR
TRA
A
P
R
L
R
8.00
4.59
t(10)

Figure 2