Perceiving Biological Motion: Dissociating Visible Speech from Walking

Andrea Santi[1], Philip Servos[1,2], Eric Vatikiotis-Bateson[3], Takaaki Kuratate[3], and Kevin Munhall[3,4]

Abstract

■ Neuropsychological research suggests that the neural system underlying visible speech on the basis of kinematics is distinct from the system underlying visible speech of static images of the face and identifying whole-body actions from kinematics alone. Functional magnetic resonance imaging was used to identify the neural systems underlying point-light visible speech, as well as perception of a walking/jumping point-light body, to determine if they are independent. Although both point-light stimuli produced overlapping activation in the right middle occipital gyrus encompassing area KO and the right inferior temporal gyrus, they also activated distinct areas. Perception of walking biological motion activated a medial occipital area along the lingual gyrus close to the cuneus border, and the ventromedial frontal cortex, neither of which was activated by visible speech biological motion. In contrast, perception of visible speech biological motion activated right V5 and a network of motor-related areas (Broca's area, PM, M1, and supplementary motor area (SMA)), none of which were activated by walking biological motion. Many of the areas activated by seeing visible speech biological motion are similar to those activated while speech-reading from an actual face, with the exception of M1 and medial SMA. The motor-related areas found to be active during point-light visible speech are consistent with recent work characterizing the human ''mirror'' system (Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). ■

INTRODUCTION

Visual speech perception is functionally adaptive because it provides an alternative means to perceive speech under conditions in which the auditory modality is compromised (e.g., auditory impairment or noise in the environment). Hearing humans are natural speechreaders, to some degree. Seeing a speaker's face (especially the jaws, lips, and tongue) assists speech comprehension in good hearing conditions, when semantic content is difficult or when the speaker has an unfamiliar accent (Reisberg, McLean, & Goldfield, 1987), as well as under noisy hearing conditions (Sumby & Pollack, 1954). Furthermore, the McGurk effect shows that when visual speech is dubbed with different auditory speech, people report hearing an integrated percept of the two modalities' inputs (McGurk & MacDonald, 1976).

In a natural environment, there are two integrated pieces of visual information available for visual speech perception—static pictorial cues and kinematics. It is possible to study the kinematics of lipreading in the absence of pictorial cues, using point-light stimuli. Point-light stimuli were developed by Johansson (1973) to study biological motion perception of the whole body. These stimuli were created by placing point-lights on the joints of an actor who was filmed walking under dim-light conditions. Although each static frame looks like a random dot arrangement, observers immediately recognize the animated frames as a walking person. Consistent with studies using a fully illuminated face, a point-light talking mouth has been shown to assist the perception of speech in noise (Rosenblum, Johnson, & Saldana, 1996) and to interfere with audio speech perception when the auditory and visual streams are incongruent (Rosenblum & Saldana, 1996).

Insights into the neural system underlying visible speech kinematics have been gained through neuropsychological research. L. M., an akinetopsic patient with damage to the occipito-temporal cortex (likely including V5), is impaired on many low-level motion tasks and is also impaired at speechreading numbers mouthed under point-light conditions. Additional work suggests that L. M. is capable of deriving speech from still-mouth images (e.g., identifying vowels from images of faces) (Campbell, Zihl, Massaro, Munhall, & Cohen, 1997). The visual form agnosics HJA (damage to V1, V2, and V4) and DF (extrastriate damage) demonstrate a reversal of L. M.'s speechreading abilities and deficits. That is, both HJA and DF can speechread an actual moving face, but cannot identify speech sounds from still faces (Munhall,

[1]Wilfrid Laurier University, Waterloo, Canada, [2]CIHR Group on Action and Perception, [3]ATR Human Information Science Laboratories, Kyoto, Japan, [4]Queen's University, Kingston, Canada

Servos, Santi, & Goodale, 2002; Campbell, 1992). These findings imply that seeing speech movement relies, to some extent, on a separate pathway from seeing speech form. Interestingly, despite L. M.'s limited low-level motion perception and her deficits on higher-level kinematic speech perception, she shows preserved recognition of other point-light actions (McLeod, Dittrich, Driver, Perrett, & Zihl, 1996). Collectively, these findings indicate that seeing the kinematics of speech activates a neural system that is, to some degree, independent from those involved in seeing static speech forms and those involved in perception of nonspeech biological motion.

Although no studies have investigated the specific brain areas involved in speechreading from kinematics alone, a few neuroimaging studies have looked at areas of the brain that are active while speechreading a fully illuminated face. The areas of interest that these studies tend to report are the superior temporal gyrus (STG) and the inferior frontal gyrus (Campbell et al., 2001; Calvert et al., 1997).

Given the neuropsychological evidence for separate neural substrates underlying reading speech from form and reading speech from kinematics, the purpose of the present study was to determine what neural substrates are involved in processing visible speech kinematics. Moreover, we wished to determine whether the perception of speech biological motion recruits distinct neural substrates from the perception of other types of biological motion (e.g., walking).

The neuropsychological dissociation between the perception of nonbiological and biological motion in patient L. M. suggests that the neural system underlying biological motion perception may be separate from the nonbiological motion perception pathway involving area V5.

Human neuroimaging evidence is consistent with the idea that biological motion perception is processed in a pathway that is not critically dependent on area V5. There is mounting evidence that some portion of the superior temporal sulcus (STS) region (i.e., the STS and adjacent cortex along its straight segment on the surfaces of the STG and middle temporal gyrus (MTG), as well as the angular gyrus; Allison, Puce, & McCarthy, 2000) is involved in biological motion perception, although some other areas outside of the STS region have also been implicated (see Servos, Osu, Santi, & Kawato, 2002; Grezes et al., 2001; Vaina, Solomon, Chowdhury, Sinha, & Belliveau, 2001; Grossman et al., 2000; Bonda, Petrides, Ostry, & Evans, 1996; Howard et al., 1996). To determine whether speech and whole-body biological motion perception rely on different neural systems, we directly compared areas activated by both types of point-light stimuli in the same individuals.

RESULTS

The two types of biological motion produced overlapping activation in both the right middle occipital gyrus and the right inferior temporal gyrus. Additionally, both stimuli activated portions of the ventral occipito-temporal cortex and the STS region. Specifically, walking biological motion perception activated the fusiform gyrus bilaterally with a right hemisphere dominance and the right STS, while visible speech biological motion perception activated the right lingual gyrus and the STG bilaterally with a left hemisphere dominance. The common areas of activation across the two experiments are shown in Figure 1. Table 1 summarizes the areas activated by the walking and visible speech point-light stimuli.

Perception of jumping/walking biological motion activated the right MTG and the lingual gyrus at the cuneus border bilaterally, the right inferior frontal gyrus, and the ventromedial frontal cortex (superior frontal gyrus, anterior orbital gyrus, and frontomarginal gyrus).

Perception of the visible speech point-light stimulus activated the cerebellum, the right MTG, the left supramarginal gyrus, the precentral gyrus (including M1) bilaterally with a right hemisphere dominance, the inferior frontal gyrus bilaterally (Broca's area (Brodmann's area (BA) 44/45) and its right-sided homolog), and the left superior frontal gyrus. The two areas of activation within the left superior frontal gyrus are within the supplementary motor area (SMA). Given that a line drawn through the anterior commissure is considered, the border between the anterior and posterior SMA (Passingham, 1995), the SMA cluster with Talairach coordinates (−2, 0, 57) is centered on the border and the cluster with Talairach coordinates (−6, −20, 56) is within posterior SMA. Figure 2 displays some of the motor-related cortical activation (Broca's area, SMA, and M1) specific to the perception of visible speech biological motion.

DISCUSSION

The perception of visible speech and walking biological motion involves relatively independent processing streams. Overlap in activation of biological motion perception for these two forms was only found in the middle occipital gyrus ($z = -8$ to $+6$), encompassing area KO, and the right inferior temporal gyrus ($z = -7$ to -9). Vaina et al. (2001) found a similar area of activation within area KO (28, −84, 0) during perception of biological motion. Given that area KO is involved in processing kinetic borders (Dupont et al., 1997), Vaina et al. suggested that kinetic borders, corresponding to the outline of the body, and in our case the lips as well, could have been generated by integrating the different directions in which the point-lights moved. Walking biological motion perception activated right KO only, whereas visible speech biological motion perception activated KO bilaterally but with a right hemisphere dominance. The common region activated within the right IT during the perception of both types of biological motion is likely due to its role in object perception

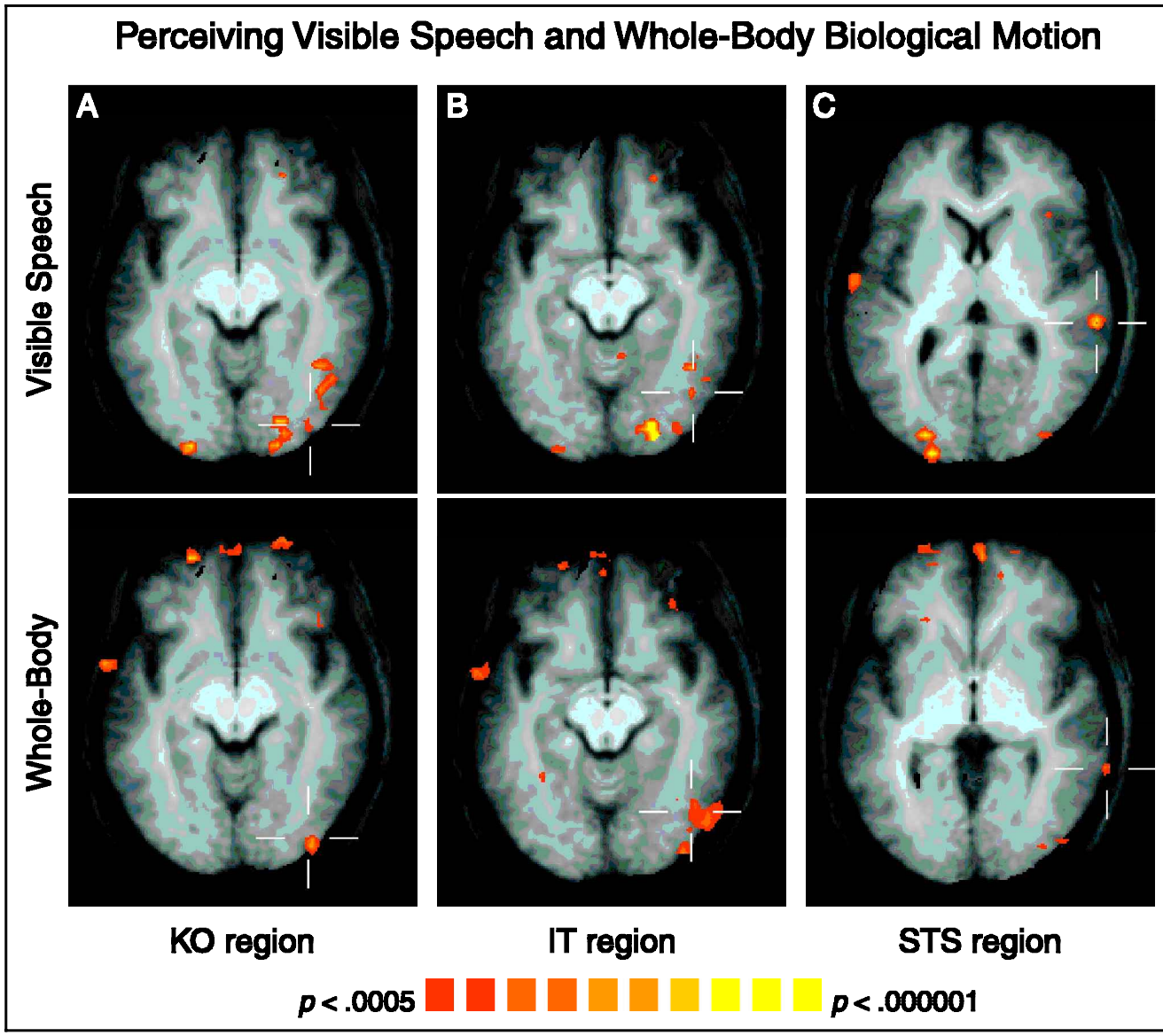

Figure 1. Common regions of activation in both biological motion perception experiments. The results of group analyses were overlaid on an averaged Talairached brain (the left hemisphere is on the left side of all images). The upper images show activation during speech-specific biological motion perception and the lower images show activation during whole-body biological motion perception (all axial views). (A) Common activation in the KO region. (B) Common activation within the IT region. (C) Common activation in the general STS region. Note that speech-specific biological motion perception activated a more anterior and superior portion of the STS region relative to whole-body biological motion perception.

(Ishai, Ungerleider, Martin, Schouten, & Haxby, 1999; Kanwisher, McDermott, & Chun, 1997). The two types of biological motion also produced separate regions of activation within IT—walking biological motion activation extended inferiorly and anteriorly into the fusiform gyrus, whereas visible speech biological motion activation extended more posteriorly into the lingual gyrus. Furthermore, walking biological motion produced bilateral fusiform/lingual gyrus activation that was right hemisphere dominant, whereas visible speech biological motion produced fusiform/lingual gyrus activation in the right hemisphere only. The general bias towards right hemisphere activation may be related to the greater role played by the right hemisphere in processing socially meaningful stimuli such as faces and biological motion (Perry et al., 2001; Borod, Haywood, & Koff, 1997).

Consistent with previous studies involving actual faces, visible speech biological motion perception produced bilateral activation on the upper bank of the STS in the STG with greater left activation (Campbell et al., 2001; Calvert et al., 1997), whereas perception of walking biological motion produced activation on the right lower bank of the STS in the MTG (Grossman et al., 2000; Bonda et al., 1996). Visible speech biological motion perception led to activation within a more superior and anterior STS region than walking biological motion perception. These findings suggest that linguistically meaningful biological motion is processed more

Table 1. Regions of Activation for Each Experiment Meeting a Threshold of $p < 5 \times 10^{-4}$ and a cluster size of 90 mm^3

	Talairach Coordinates				
Side Region	*x*	*y*	*z*	*Cluster Size (1 mm^3 Voxels)*	*BA*
Visible speech					
L middle occipital gyrus (KO)	−24	−91	5	1307	17/18
R lingual/middle occipital gyrus (KO)	23	−83	−2	1462	17/18
R inferior/middle temporal gyrus	41	−58	−2	1054	19/37
L superior temporal gyrus	−61	−8	13	445	22/42
R superior temporal gyrus	56	−29	9	176	22/42
L supramarginal gyrus	−48	−26	25	112	40
R precentral gyrus (M1)	57	−5	36	955	4/6
L precentral gyrus (M1)	−54	−5	40	458	4/6
L precentral gyrus (PM)	−35	−13	58	93	6
L inferior frontal gyrus (Broca's area)	−56	3	23	263	6/44
L inferior frontal gyrus (Broca's area)	−45	22	16	241	44/45
L inferior frontal gyrus (Broca's area)	−51	9	30	186	9/44
R inferior frontal gyrus	46	8	23	115	9/44
R inferior frontal gyrus	33	21	12	99	13
L superior frontal gyrus (SMA)	−6	−20	56	147	6
L superior frontal gyrus (SMA)	−2	0	57	1123	6
L cerebellum, culmen	−22	−52	−18	114	–
R cerebellum, culmen	5	−46	−9	94	–
Walk/jump					
R middle occipital gyrus (KO)	37	−81	0	642	17/18
L lingual gyrus	−9	−57	14	273	17/18
R lingual gyrus	6	−66	11	100	17
L fusiform gyrus	−32	−54	−13	980	19/37
L fusiform gyrus	−37	−45	−20	178	37
R inferior temporal/fusiform gyrus	40	−54	−13	3419	19/20/37
R middle temporal gyrus	36	−6	−27	169	21
R superior temporal sulcus	63	−44	2	92	21/22
R inferior frontal gyrus	39	25	−1	141	47
R superior frontal gyrus	2	62	4	282	10
L anterior orbital gyrus	−22	56	−2	361	10/11
L frontomarginal gyrus	−24	63	4	178	10
L frontomarginal gyrus	−3	59	−5	272	10/11
R frontomarginal gyrus	21	63	−1	397	10

Talairach coordinates report the center of gravity of the activation cluster. L = left, R = right, KO = kinetic occipital region, SMA = supplementary motor area, M1 = primary motor area, and PM = premotor cortex.

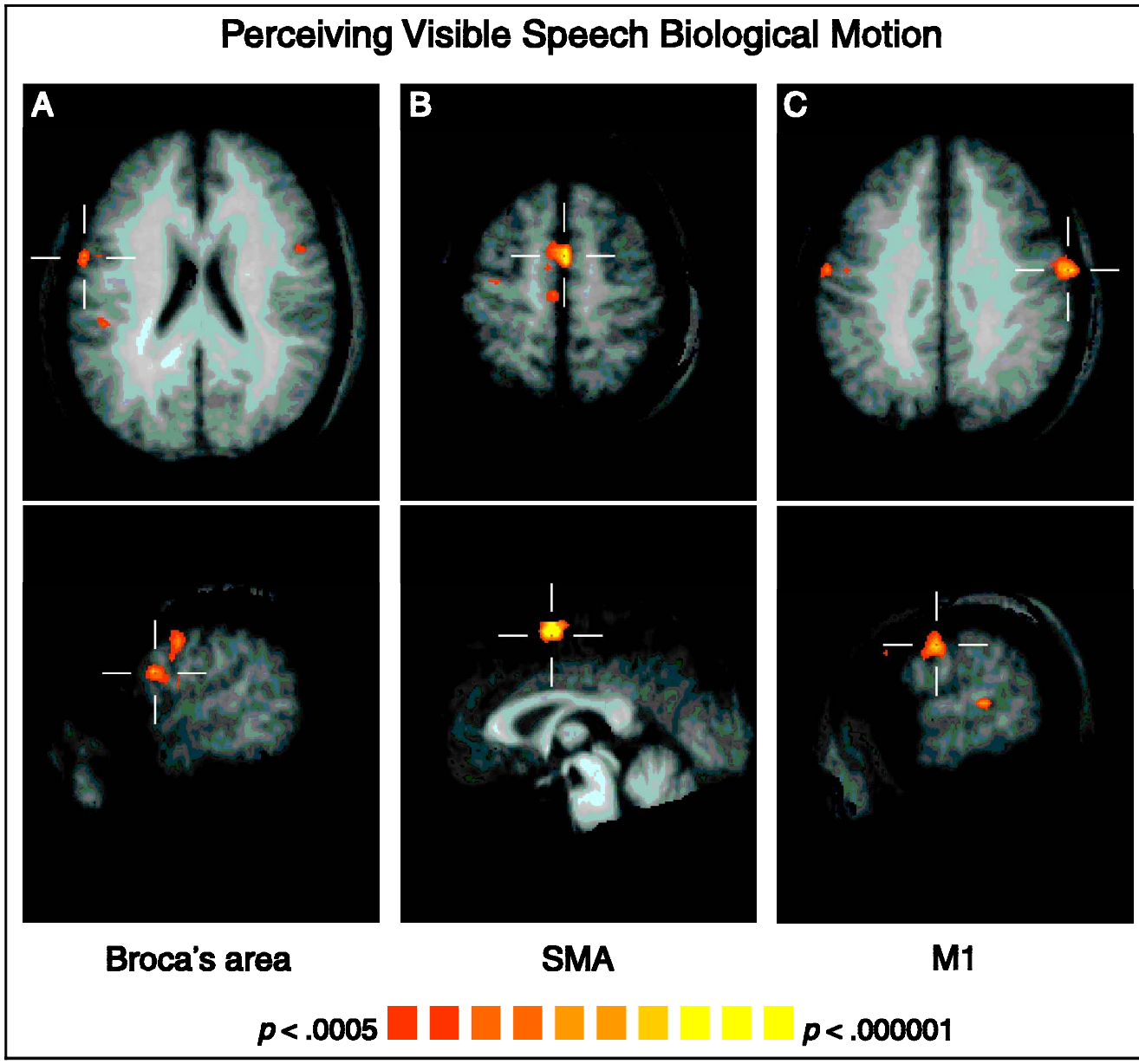

Figure 2. Motor-related areas activated during speech-specific biological motion perception. The results of group analysis were overlaid on an averaged Talairached brain (the left hemisphere is on the left side of the axial images; upper row: axial views, lower row: sagittal views). (A) Activation within Broca's area. (B) Activation within the medial SMA. (C) Activation within the right M1.

anteriorly in the STS region than nonlinguistically meaningful biological motion. These findings are also consistent with previous work showing that nonlinguistic mouth movements activate relatively posterior portions of the STS region (Puce, Allison, Bentin, Gore, & McCarthy, 1998).

The portions of the STG (BA 22/42) that were activated by seeing speech have previously been shown to be engaged in auditory word processing (Chee, O'Craven, Bergida, Rosen, & Savoy, 1999; Binder et al., 1994). In contrast, such activation within the auditory cortex was not observed during the perception of the walking biological motion stimuli. Our observation of auditory cortex activation during the perception of visible speech biological motion complements and extends those of Calvert et al. (1997) by showing that the kinematics alone of a visible speech stimulus are sufficient to activate the auditory cortex (but see Bernstein et al., 2002).

Studies of neuropsychological patients with either left or right hemisphere damage indicate that the left hemisphere is necessary for lipreading while the right hemisphere is not (Campbell, 1992). Divided visual field studies in neurologically intact subjects generally support this conclusion (Smeele, Massaro, Cohen, & Sittig, 1998; Campbell, De Gelder, & De Haan, 1996; Diesch, 1995; Baynes, Funnell, & Fowler, 1994; although see Campbell, 1986). This is consistent with our finding

that there was greater activation within the left STG. However, we also observed some activation within the right STG supporting functional magnetic resonance imaging studies reporting right STG activation when observers lipread (Campbell et al., 2001; Calvert et al., 1997; Calvert & Campbell, 2003).

In addition to the common areas of activation produced by the two types of biological motion (i.e., KO, ventral pathway, and STS region), both stimuli activated separate cortical regions as well. Perception of walking biological motion activated a medial area in the lingual gyrus bilaterally, which was not observed during visible speech biological motion perception. This finding replicates previous work in our laboratory in which walking biological motion stimuli activated area VP (Servos et al., 2002). As Servos et al. (2002) suggest, the lingual gyrus and STS findings may parallel findings from the face perception literature. In face perception, there is a double dissociation between deriving identity (fusiform gyrus) and deriving socially salient information (STS). Servos et al. suggest that the lingual gyrus is associated with deriving "form-from-motion," whereas the STS is involved in interpreting the form-from-motion in a socially meaningful context. The present findings extend these hypotheses by suggesting that language-related biological motion does not have access to VP.

Activation of the ventromedial prefrontal cortex (BA 10 and 11) during walking biological motion perception is consistent with the notion that this region is involved in perceiving socially meaningful stimuli and processing whether they lead to rewarding or aversive consequences (for a review, see Adolphs, 1999). Furthermore, the orbito-frontal cortex (including BA 11) likely plays a role in a network of areas (amygdala, STS, and orbito-frontal cortex) important to perceiving socially meaningful stimuli through STS inputs (Allison et al., 2000). Previous studies have shown orbito-frontal cortical activation during perception of such socially meaningful stimuli as faces and gaze (Allison, Puce, Spencer, & McCarthy, 1999; Wicker, Michel, Henaff, & Decety, 1998). Moreover, the orbito-frontal cortex is thought to play a role in social reinforcement (Rolls, 2000).

Perception of visible speech biological motion activated the right MTG (area V5), whereas perception of walking biological motion did not. Campbell et al. (2001) also found greater activation within the MTG (BA 37) in the right hemisphere (58, −53, −2) during speechreading compared with counting gurns. Given that the local motion properties are identical in both our stimulus and control conditions, the V5 activation for visible speech is somewhat surprising. In contrast, the V5 activation observed by Campbell et al. might have been due to possible differences in the amplitudes and/or velocities of the speech mouth movements relative to the gurning movements. Nevertheless, both our study and that of Campbell et al. observed only right V5 activation, implying that right V5 is more active during the processing of speech movements relative to random movements. The reasons for this are unclear.

Perception of visible speech biological motion activated a network of motor-related areas (Broca's area, PM, SMA, and M1) that were not activated during walking biological motion perception. The only motor-related area uniquely activated by the walking biological motion stimuli was located within the inferior frontal gyrus (BA 47). A previous neuroimaging study also found a similar area within the inferior frontal gyrus (BA 45 and 47) active during whole-body biological motion perception (Vaina et al., 2001).

The extensive activation of motor areas during the perception of visible speech biological motion, in the current study, is possibly due to activation of a human mirror system that is sensitive to the kinematics of facial movements. Mirror neurons, which respond both when an action is observed and when the same action is executed, were first discovered in monkey F5—the monkey homolog of Broca's area (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996; di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992). The mirror neuron system is thought to be important for learning, imitating, and understanding the actions of others (Gallese et al., 1996) and more broadly for understanding other people's states of mind (Gallese & Goldman, 1998). In humans, the inferior frontal cortex (including Broca's area), M1, parietal cortex, STS, and somatosensory cortices (SI and SII) appear to be involved in the mirror system (Avikainen, Forss, & Hari, 2002; Nishitani & Hari, 2000; Iacoboni et al., 1999; Hari et al., 1998; Rizzolatti et al., 1996). For example, Buccino et al. (2001) found premotor cortex activation during the observation of another individual performing actions with various effectors (mouth, arm/hand, and foot). The premotor areas (BA 6, 44, and 45) that they found to be active during observation of mouth actions are in similar locations to those that we found to be active during lipreading. Interestingly, Buccino et al. found that right hemisphere activation was greater than left hemisphere activation during observation of mouth actions, which is also true of our activation within BA 4 and 6. Campbell et al. (2001) also observed activation in inferior frontal regions during a lipreading task (see also Calvert & Campbell, 2003), but unlike Buccino et al., in a nonlinguistic gurning condition, they did not observe this frontal activation. At this point, the extent to which linguistically and nonlinguistically meaningful actions share common neural substrates within the "mouth" mirror system is not entirely clear.

We did find an additional motor-related area that was active during the perception of visible speech biological motion-area SMA. This region might in principle also form part of the human mirror system. For example, the rostral posterior SMA appears to be recruited during imagining and execution of action, while the more caudal parts of the SMA are active during motor execution

(Stephan et al., 1995). More relevantly, the SMA has been shown to play a role in speech production. Electrical stimulation of the SMA can cause either speech production or speech arrest (Fried et al., 1991). Thus, it is possible that the SMA contains mirror neurons for both speech and nonspeech action.

An additional source of the inferior frontal gyrus (Broca's area) and SMA activation might lie in the role that they play in phonological and semantic processing (Burton, Noll, & Small, 2001; Chee et al., 1999; Fiez, 1997). Broca's area is known to play a role in phonological processing (Fiez, 1997), and some work implicates portions of the SMA in semantic processing (Chee et al., 1999). However, there is little overlap between the anterior portion of SMA ($y \geq +9$) that Chee et al. (1999) observed during semantic processing and the SMA activation observed in the present study ($y = +9$ to -10). Thus, semantic-specific processing probably does not account for much of the SMA activation that we observed in the present study. It should also be noted that subjects in the present study likely engaged in relatively little phonological and semantic processing given the relative difficulty that naive observers have in extracting words solely on the basis of face kinematics. Even with video images of the full face, silent speechreading performance is generally poor in the general population (Summerfield, 1992). For the point-light stimuli tested here, pilot subjects found it virtually impossible to identify words without an accompanying acoustic signal.

To sum up, the extensive activation of motor-related areas observed during the perception of visible speech biological motion, in the current study, is likely a consequence of the engagement of the human mirror system (most likely M1 and probably Broca's area and SMA) with the possible additional involvement of phonological processing in Broca's area and to a lesser extent semantic processing in SMA.

Visible speech and walking biological motion perception rely on networks that have cortical regions in common as well as regions that are relatively independent. Both forms of biological motion activated area KO, the ventral pathway, and the STS region. For both types of biological motion, shape processing appears to be largely taken on by regions within the right hemisphere. However, at the level of the STS, we see that the left hemisphere becomes dominant in processing speech-related biological motion, while the right STS maintains its dominant role in processing whole-body biological motion. VP and the ventromedial cortex were uniquely activated by walking biological motion perception, while extensive motor-related cortex and V5 were uniquely activated by the perception of visible speech biological motion. Although many of the areas activated by visible speech biological motion are similar to those activated during speechreading from an actual face, the current findings implicate a more extensive mirror system, including M1 and SMA, in lipreading speech kinematics. Future work will hopefully identify the entire cortical network involved in a speech mirror system and the degree to which this system is driven purely by the perception of facial kinematics as compared to the perception of the dynamics of an actual face speaking or the perception of static images of a speaking face.

METHODS

Subjects

Ten right-handed, native English-speaking subjects with normal or corrected-to-normal vision participated in the experiment (6 men and 4 women, 22–25 years old, $\bar{x}_{age} =$ 23 years). Prior to the experiment, an informed consent was obtained from each of the subjects. The study protocol was in accordance with the ethical guidelines of the Robarts Research Institute (London, ON).

Stimuli

The visible speech stimuli were created by positioning 17 infrared emitting diodes on a speaker's face below the eyes with an emphasis on the area around the mouth. The positions were recorded with an Optotrak optoelectronic imaging system (Northern Digital, Waterloo, ON). The 3-D movements of the markers were then registered to a head reference frame to eliminate head motion. The 3-D (x, y, z) data streams for each marker were used to create the point-light animation. The control for this stimulus was created by scrambling the markers' positions within the area of the head sphere and rotating their axes randomly. The stimuli were 320 × 240 pixels in size (11.6° × 8.7°).

The jumping/walking point-light stimuli were created in a fashion similar to that of the talking stimuli. Twenty-one light points were recorded from the major joints of the body (head, neck, torso, and both shoulders, elbows, wrists, hands, hips, knees, ankles, heels, and toes) using the Optotrak. The control stimuli for the jumping/walking person were created by applying autoregressive models to the stimuli using TIMSAC such that the dots' local motion properties (speeds and amplitudes) were preserved, whereas their linking structure was not (for details, see methods by Servos et al., 2002). The stimuli were 300 × 300 pixels in size (10.8° × 10.8°). The two actions (jumping and walking) alternated throughout the 36-sec stimulus half-cycle in order to maintain the participants' interest. The duration between alternations of the two actions were 2, 3, or 4 sec. The intervals randomly alternated, but in such a way that each action was presented for an equal duration of time during the stimulus half-cycle.

Procedure

Each participant partook in two 6-min experiments with experiment order counterbalanced across subjects. Both

experiments consisted of five 72-sec cycles. Either a talking point-light face or a jumping and walking point-light person alternated with moving control point-lights every 36 sec. The stimuli were presented in QuickTime and were projected by an SVGA projector onto a screen located approximately 60 cm from the subject's eyes. The subject viewed the stimuli through a mirror angled above his or her head within the headcoil. In both experiments, the subject's task was to fixate in the center of the display, while attending to the stimulus. In the jumping/walking experiment, subjects were asked to attend to when the displays changed from one type of body movement to the other. In the visible speech experiment, subjects were asked to do their best to lipread. The subjects were informed at the beginning of the study that the speech stimuli were made up of multiple sentences that each contained a noun, adjective, and verb; each having three possible alternatives (people/father/women, shopped/played/walked, badly/quickly/slowly). Average sentence duration was 2.24 sec. Subjects were shown examples of the point-light stimuli prior to the experiment.

Functional Magnetic Resonance Imaging

Image acquisition was performed with a 4-T whole-body imaging system (Varian, Palo Alto, CA; Siemens, Erlangen, Germany; Robarts imager located at the Robarts Research Institute in London, Canada) with a 90-cm-diameter horizontal bore and a whole-body 68-cm-diameter gradient set with a maximum strength of 40 mT/m and a slew rate of 120 mT/m/s. A transmit–receive cylindrical hybrid birdcage radio frequency coil (Barbieri, Gati, Rutt, & Menon, 2000) was used for transmission and detection of signal. A global shim was performed followed by anatomic imaging to localize the area of interest (entire brain and upper half of cerebellum; 23 contiguous 5-mm axial slices). Each functional volume was acquired using a segmented T2*-weighted gradient echo–echo planar imaging (EPI) pulse sequence with navigator echo correction. Each imaging run produced 120 continuous acquisitions of the brain volume (TE = 10 msec, TR = 750 msec, 4 shots, FA = 40°, 64 × 64 matrix size; FOV = 192 × 192 mm; volume acquisition time of 3 sec). High-resolution (256 × 256) 3-D T1-weighted anatomical volumes were acquired (TI = 600 msec; TR = 10 msec; TE = 5.5 msec; FA = 15°; FOV = 192 × 192 mm). The anatomical volume consisted of 128 contiguous structural images with a slice thickness of 1.25 mm.

Data Analysis

Analyses were conducted with BrainVoyager 4.4. Both the anatomical and functional data were isovoxeled (1 mm^3) and transformed into Talairach and Tournoux (1988) stereotaxic coordinates. The Talairached brains of the 10 subjects were averaged to account for any variation in cortical anatomy across the subjects. The functional data were preprocessed to remove any linear trends and slow drifts. For each experiment, a model of the hemodynamic response was fit to a concatenation of each subject's data at each voxel to generate group average maps. Specifically, we performed a General Linear Model, Multiple Regression analysis that is represented by the following equation: $Y_z(t) = Y'_z(t) + e(t) = \beta_{confound} + \beta_1 X_1(t) + e(t)$, where $Y_z(t)$ is the measured standardized signal at time point t, $\beta_{confound}$ is an estimate of the standardized signal during baseline episodes, $X_1(t)$ is the value of the predictor variable at time point t, β_1 is the standardized regression coefficient, $Y'_z(t)$ is the best predicted standardized signal, and $e(t)$ is the unexplained data ($Y_z(t) - Y'_z(t)$). The measured signal ($Y_z(t)$) at each voxel is a concatenation of each subject's signal time course that have each undergone z-normalization (i.e., each subject's signal time course has $\bar{x} = 0$ and $SD = 1$) to account for differences in average signal strength and variation across subjects. The predictor variable (X_1) is the modeled hemodynamic response over the 6-min experiment (five on/off cycles) appended 10 times (number of subjects) (thus, 10 subjects × 120 time points = 1200 time points). Regions displaying activation had to both meet a threshold of $p < .0005$ (uncorrected for multiple comparisons) and be 90 mm^3 in size (equivalent to two functional voxels). The Talairach Daemon (http://ric.uthscsa.edu/projects/talairachdaemon.html) and Duvernoy (1991) brain atlas were conjointly used in identifying the anatomical location of the functional clusters of activation.

Acknowledgments

We would like to thank Rieko Osu for her help in the generation of the ''walking'' stimuli. This research was supported by NSERC (K. M. and P. S.), the Canada Research Chairs Program (P. S.), and by the CRL Keihanna Info-Communications Research Laboratories and the Telecommunications Advancement Organization of Japan (E. V.-B., T. K., and K. M.).

Reprint requests should be sent to Philip Servos, Department of Psychology, Wilfrid Laurier University, Waterloo, ON, Canada N2L 3C5, or via e-mail: pservos@wlu.ca.

The data reported in this experiment have been deposited in The fMRI Data Center (http://www.fmridc.org). The accession number is 2-2003-113JE.

REFERENCES

Adolphs, R. (1999). Social cognition and the human brain. *Trends in Cognitive Sciences, 3,* 469–479.

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences, 4,* 267–278.

Allison, T., Puce, A., Spencer, D. D., & McCarthy, G. (1999). Electrophysiological studies of human face perception: I. Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cerebral Cortex, 9,* 415–430.

Avikainen, S., Forss, N., & Hari, R. (2002). Modulated activation of the human S1 and S11 cortices during observation of hand actions. *Neuroimage, 15,* 640–646.

Barbieri, E. A., Gati, J. S., Rutt, B. K., & Menon, R. S. (2000). A transmit-only/receive-only (TORO) RF system for high-field MRI/MRS applications. *Magnetic Resonance in Medicine, 43,* 284–289.

Baynes, K., Funnell, M. G., & Fowler, C. A. (1994). Hemispheric contributions to the integration of visual and auditory information in speech perception. *Perception and Psychophysics, 55,* 633–641.

Bernstein, L. E., Auer, E. T., Moore, J. K., Ponton, C. W., Don, M., & Singh, M. (2002). Visual speech perception without primary auditory cortex activation. *NeuroReport, 13,* 311–315.

Binder, J. R., Rao, S. M., Hammeke, T. A., Yetkin, F. Z., Jesmanowicz, A., Bandettini, P. A., Wong, E. C., Estkowski, L. D., Goldstein, M. D., Haughton, V. M., & Hyde, J. S. (1994). Functional magnetic resonance imaging of human auditory cortex. *Annals in Neurology, 35,* 662–672.

Bonda, E., Petrides, M., Ostry, D., & Evans, A. (1996). Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *Journal of Neuroscience, 16,* 3737–3744.

Borod, J. C., Haywood, C. S., & Koff, E. (1997). Neuropsychological aspects of facial asymmetry during emotional expression: A review of the normal adult literature. *Neuropsychology Review, 7,* 41–60.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R. J., Zilles, K., Rizzolatti, G., & Freund, H.-J. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience, 13,* 400–404.

Burton, M. W., Noll, D. C., & Small, S. L. (2001). The anatomy of auditory word processing: Individual variability. *Brain and Language, 77,* 119–131.

Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C. R., McGuire, P. K., Woodruff, P. W. R., Iversen, S. D., & David, A. S. (1997). Activation of auditory cortex during silent lipreading. *Science, 276,* 593–596.

Calvert, G. A., & Campbell, R. (2003). Reading speech from still and moving faces: The neural substrates of visible speech. *Journal of Cognitive Neuroscience, 15,* 57–70.

Campbell, R. (1986). The lateralization of lip-read sounds: A first look. *Brain and Cognition, 5,* 1–21.

Campbell, R. (1992). The neuropsychology of lipreading. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences, 335,* 39–45.

Campbell, R., De Gelder, B., & De Haan, E. (1996). The lateralization of lip-reading: A second look. *Neuropsychologia, 34,* 1235–1240.

Campbell, R., MacSweeney, M., Surguladze, S., Calvert, G., McGuire, P., Suckling, J., Brammer, M. J., & David, A. S. (2001). Cortical substrates for the perception of face actions: An fMRI study of the specificity of activation for seen speech and for meaningless lower-face acts (gurning). *Cognitive Brain Research, 12,* 233–243.

Campbell, R., Zihl, J., Massaro, D., Munhall, K., & Cohen, M. M. (1997). Speechreading in the akinetopsic patient, L. M. *Brain, 120,* 1793–1803.

Chee, M. W. L., O'Craven, K. M., Bergida, R., Rosen, B. R., & Savoy, R. L. (1999). Auditory and visual word processing studied with fMRI. *Human Brain Mapping, 7,* 15–28.

Diesch, E. (1995). Left and right hemifield advantages of fusions and combinations in audiovisual speech perception. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 48A,* 320–333.

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research, 91,* 176–180.

Dupont, P., De Bruyn, B., Vandenberghe, R., Rosier, A., Michiels, J., Marchal, G., Mortelmans, L., & Orban, G. A. (1997). The kinetic occipital region in human visual cortex. *Cerebral Cortex, 7,* 283–292.

Duvernoy, H. M. (1991). *The human brain, surface, three-dimensional sectional anatomy and MRI.* Wien, Germany: Springer-Verlag.

Fiez, J. A. (1997). Phonology, semantics, and the role of the left inferior prefrontal cortex. *Human Brain Mapping, 5,* 79–83.

Fried, I., Katz, A., McCarthy, G., Sass, K. J., Williamson, P., Spencer, S. S., & Spencer, D. D. (1991). Functional organization of human supplementary motor cortex studied by electrical stimulation. *Journal of Neuroscience, 11,* 3656–3666.

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

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences, 2,* 493–501.

Grezes, J., Fonlupt, P., Bertenthal, B., Delon-Martin, C., Segebarth, C., & Decety, J. (2001). Does perception of biological motion rely on specific brain regions? *Neuroimage, 13,* 775–785.

Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., & Blake, R. (2000). Brain areas involved in perception of biological motion. *Journal of Cognitive Neuroscience, 12,* 711–720.

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. *Proceedings of the National Academy of Sciences, U.S.A., 95,* 15061–15065.

Howard, R. J., Brammer, M., Wright, I., Woodruff, P. W., Bullmore, E. T., & Zeki, S. (1996). A direct demonstration of functional specialization within motion-related visual and auditory cortex of the human brain. *Current Biology, 6,* 1015–1019.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science, 286,* 2526–2528.

Ishai, A., Ungerleider, L. G., Martin, A., Schouten, J. L., & Haxby, J. V. (1999). Distributed representation of objects in the human ventral visual pathway. *Proceedings of the National Academy of Sciences, U.S.A., 96,* 9379–9384.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics, 14,* 201–211.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17,* 4302–4311.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264,* 746–748.

McLeod, P., Dittrich, W., Driver, J., Perrett, D., & Zihl, J. (1996). Preserved and impaired detection of structure from motion by a 'motion-blind' patient. *Visual Cognition, 3,* 363–391.

Munhall, K. G., Servos, P., Santi, A., & Goodale, M. A. (2002). Dynamic visual speech perception in a patient with visual form agnosia. *NeuroReport, 13,* 1793–1796.

Nishitani, N., & Hari, R. (2000). Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences, U.S.A., 97,* 913–918.

Passingham, R. E. (1995). The status of the premotor areas: Evidence from PET scanning. In W. R. Ferrell & U. Proske

(Eds.), *Neural control of movement* (pp. 167–178). New York: Plenum.
Perry, R. J., Rosen, H. R., Kramer, J. H., Beer, J. S., Levenson, R. L., & Miller, B. L. (2001). Hemispheric dominance for emotions, empathy and social behaviour: Evidence from right and left handers with frontotemporal dementia. *Neurocase, 7,* 145–160.
Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *Journal of Neuroscience, 18,* 2188–2199.
Reisberg, D., McLean, J., & Goldfield, A. (1987) Easy to hear but hard to understand: A lip-reading advantage with intact auditory stimuli. In B. Dodd & R. Campbell (Eds.), *Hearing by eye: The psychology of lip-reading* (pp. 97–113). London: Erlbaum.
Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research, 3,* 131–141.
Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., & Fazio, F. (1996). Localization of grasp representations in humans by PET: 1. Observation versus execution. *Experimental Brain Research, 111,* 246–252.
Rolls, E. T. (2000). The orbitofrontal cortex and reward. *Cerebral Cortex, 10,* 284–294.
Rosenblum, L. D., Johnson, J. A., & Saldana, H. M. (1996). Point-light facial displays enhance comprehension of speech in noise. *Journal of Speech and Hearing Research, 39,* 1159–1170.
Rosenblum, L. D., & Saldana, H. M. (1996). An audiovisual test of kinematic primitives for visual speech perception. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 318–331.
Servos, P., Osu, R., Santi, A., & Kawato, M. (2002). The neural substrates of biological motion perception: An fMRI study. *Cerebral Cortex, 12,* 772–782.
Smeele, P. M., Massaro, D. W., Cohen, M. M., & Sittig, A. C. (1998). *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1232–1242.
Stephan, K. M., Fink, G. R., Passingham, R. E., Silbersweig, D., Ceballos-Baumann, A. O., Frith, C. D., & Frackowiak, R. S. J. (1995). Functional anatomy of the mental representation of upper extremity movements in healthy subjects. *Journal of Neurophysiology, 73,* 373–386.
Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America, 26,* 212–215.
Summerfield, Q. (1992). Lipreading and audio-visual speech perception. *Philosophical Transactions of the Royal Society of London, B, 335,* 71–77.
Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain.* New York: Thieme.
Vaina, L. M., Solomon, J., Chowdhury, S., Sinha, P., & Belliveau, J. W. (2001). Functional neuroanatomy of biological motion perception in humans. *Proceedings of the National Academy of Sciences, U.S.A., 98,* 11656–11661.
Wicker, B., Michel, F., Henaff, M., & Decety, J. (1998). Brain regions involved in the perception of gaze: A PET study. *Neuroimage, 8,* 221–227.