

Moving Multisensory Research Along: Motion Perception Across Sensory Modalities

Salvador Soto-Faraco,¹ Charles Spence,² Donna Lloyd,³ and Alan Kingstone⁴

¹*Departament de Psicologia Bàsica, Universitat de Barcelona, Barcelona, Spain;* ²*Department of Experimental Psychology, University of Oxford, Oxford, England;* ³*Magnetic Resonance and Image Analysis Research Centre, University of Liverpool, Liverpool, England;* and ⁴*Department of Psychology, University of British Columbia, Vancouver, British Columbia, Canada*

ABSTRACT—*The past few years have seen a rapid growth of interest regarding how information from the different senses is combined. Historically, the majority of the research on this topic has focused on interactions in the perception of stationary stimuli, but given that the majority of stimuli in the world move, an important question concerns the extent to which principles derived from stationary stimuli also apply to moving stimuli. A key finding emerging from recent work with moving stimuli is that our perception of stimulus movement in one modality is frequently, and unavoidably, modulated by the concurrent movement of stimuli in other sensory modalities. Visual motion has a particularly strong influence on the perception of auditory and tactile motion. These behavioral results are now being complemented by the results of neuroimaging studies that have pointed out the existence of both modality-specific motion-processing areas and areas involved in processing motion in more than one sense. The challenge for the future will be to develop novel experimental paradigms that can integrate behavioral and neuroscientific approaches in order to refine our understanding of multisensory contributions to the perception of movement.*

KEYWORDS—*perception; motion; multisensory integration; ventriloquist illusion; neuroimaging*

The phenomenon of ventriloquism, by which a skillful performer creates the illusion that his or her vocalizations come from the mouth of a puppet, provides a classic example of how different senses can influence one another. In this case, the brain is fooled by the conspicuous sight of the puppet, and tends to realign the perceived location of the sound source to match what is seen. Simplified versions of this illusion have been used by scientists for more than a century to study how the senses interact. However, in stark contrast with the type

of situation that we often encounter in naturalistic environments (where many of the stimuli move across space), traditional research has mostly considered interactions in the perception of stationary stimuli (e.g., one flash and one beep presented from different locations). This clearly limits our understanding of how the principles of multisensory integration may contribute to perceptual processes in real-world environments.

Over the years, cognitive scientists have occasionally documented profound interactions among the different sensory modalities in the perception of motion (see Soto-Faraco, Kingstone, & Spence, 2003, for a recent review). Early studies provided phenomenological evidence (verbal descriptions based on subjective experience) that motion perception in one sense could be influenced by the presentation of a stimulus in another sense. Later studies using more objective measurements of perception provided somewhat contradictory evidence, often failing to reproduce these initial results. Crucially, researchers at the time had no real insight into the causes of these discrepancies, and so a coherent picture regarding multisensory contributions to motion perception was not achieved.

The past few years have seen a rapid resurgence of interest in this area, driven in part by the general upturn of research related to multisensory processes (e.g., Stein & Meredith, 1993). Recent studies have highlighted a variety of important constraints on the processing of motion information across sensory modalities, and these constraints can account for some of the previous inconsistencies evident in the literature. Additionally, researchers are beginning to uncover the neural substrates underlying the multisensory integration of motion cues. This brief review highlights current understanding of the multisensory contributions to the perception of motion, at both the behavioral and neural levels.

CHARACTERIZATION OF MULTISENSORY INTEGRATION OF MOTION CUES

Spatiotemporal Constraints

Multisensory contributions to motion perception are most pronounced when moving stimuli in different sensory modalities occur at the same

Address correspondence to Salvador Soto-Faraco, Departament de Psicologia Bàsica, Universitat de Barcelona, Pg. Vall d'Hebrón 171, 08035 Barcelona, Spain; e-mail: ssoto@psico.psi.ub.es.

spatial locations and at approximately the same time, and fall off as the moving stimuli in different modalities are presented out of synchrony or from disparate spatial locations (e.g., Soto-Faraco, Lyons, Gazzaniga, Spence, & Kingstone, 2002). For example, in one recent study, people judged the direction of an auditory apparent motion stream¹ that moved from left to right or right to left, while at the same time attempting to ignore a visual apparent motion stream that moved in the same direction as the sound or in the opposite direction. When the two streams were presented from the same set of locations (i.e., the motion paths overlapped) and at the same time, the consequences of multisensory interactions were evident. In particular, people judged the direction of the sound correctly almost all the time if the sound and light moved in the same direction, but their performance fell to chance levels if the sound and light moved in opposite directions. Moreover, in these directionally conflicting trials, people often reported incorrectly that the sound moved in the same direction as the light.

This *congruency effect* (worse performance when motion signals move in conflicting, rather than congruent, directions) suggests that the perceived direction of auditory motion can be profoundly affected by the direction of visual motion. However, these effects were greatly reduced when the motion signals did not share common paths or were not presented at the same time. Thus, spatial and temporal coincidence play a critical role in multisensory integration. Such findings make sense when one considers that in the real world, a given object will often generate correlated multisensory motion information that originates from the same location at the same time. The same dependence on spatial and temporal synchronization has been reported to govern the response of single neurons to multisensory stimuli in experiments with animals (Stein & Meredith, 1993), and has also been shown, in humans, to modulate perception of the ventriloquist illusion (in which the perceived location of a static sound is attracted toward a flash of light presented simultaneously; Bertelson & de Gelder, in press).

The Nature of Multisensory Interactions in Motion Perception

Recent research has revealed at least three important characteristics of the perceptual interactions between motion signals in different sensory modalities (Soto-Faraco, Spence, & Kingstone, in press). First, one sense “dominates,” rather than interfering with, the perception of motion cues in the other sense. This phenomenon is referred to as *capture*. For example, as noted already, when the directions of visual and auditory motion conflict, people often perceive the auditory stimulus as moving in the same direction as the visual stimulus (other forms of capture can involve perception of speed or trajectory). Second, these multisensory effects involve motion features (e.g., direction of motion), and stand over and above effects found with static features (e.g., spatial location, such as in the classic ventriloquist illusion). Finally, recent findings suggest that these interactions have a perceptual basis, occurring at early processing stages that involve the way in which motion stimuli are experienced. In the case of the example just mentioned, for instance, a person genuinely experiences the sound as traveling in the same direction as the moving light, even if the two stimuli are physically presented in conflicting directions. This perceptual component, however, is not the only potential source of multisensory effects in motion perception. Several authors have

shown that response-related and higher-order cognitive factors can play an important role in multisensory interactions. For example, a person may be merely biased to respond that a sound moved in the same direction of a light without necessarily experiencing that the sound traveled in the same direction as the light (Meyer & Wuerger, 2001; but see Wohlschläger, 2000).

An Unbalanced Sensory Relationship

Visual motion signals typically dominate over motion signals in the other modalities. So, for example, although an observer will often hear a sound, or feel a tactile stimulus, as moving in the direction of a concurrent visual motion stimulus (even if it actually moves in the opposite direction), the direction in which a visual stimulus appears to move is rarely affected by motion cues from another sensory modality (see Fig. 1 for results from a study we conducted recently; also see Soto-Faraco et al., 2003). Such visual dominance has been reported frequently in research on stationary stimuli (such as in the *ventriloquism* effect; see Bertelson & de Gelder, in press).

The fact that visual motion signals dominate when placed in conflict with either auditory or tactile motion signals fits well with the *modality-appropriateness account* of sensory dominance (see Welch & Warren, 1986). According to this account, our perception of any particular stimulus attribute will normally be determined by information from the sense that provides the most appropriate, or accurate, information concerning that attribute. Given that vision normally provides more accurate information about the location and movement of stimuli than any of our other senses (such as audition, touch, taste, or smell), it is vision that would be expected to dominate in conflict situations (though sound can alter ambiguous visual motion; Sekuler, Sekuler, & Lau, 1997). Ernst and Banks (2002) recently outlined a mathematical formalization of this modality-appropriateness hypothesis. Their model is based on an optimal integration mechanism

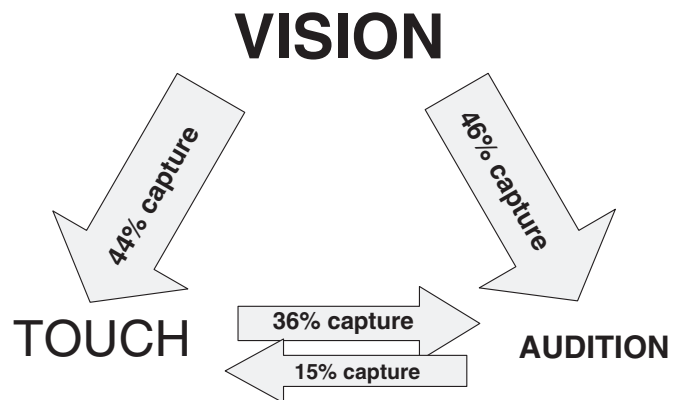


Fig. 1. Schematic representation of the results of a study (described in Soto-Faraco, Kingstone, & Spence, 2003) in which every possible pairing of target and distractor modality (across audition, vision, and touch) was evaluated by asking participants to judge the direction of a moving stimulus in the target sensory modality while attempting to ignore a moving stimulus in the distractor modality. The number within each arrow represents the magnitude of the congruency effect (the proportion of correct responses under the congruent condition minus the proportion of correct responses under the conflicting condition) obtained when the distractor was in the modality at the arrow’s origin, and the target was in the modality at the arrow’s end.

¹In apparent motion, two discrete events are presented in quick succession at different locations and produce the illusion of a single object moving from the location of the first event to the location of the second.

that evaluates the precision of each modality independently so that the more precise one modality is with respect to the other, the stronger its contribution to the final percept will be. It will be interesting to see how this model applies to the specific domain of multisensory integration of motion signals.

NEURAL UNDERPINNINGS OF MULTISENSORY MOTION INTEGRATION

Given such extensive interactions between the senses in the perception of motion, it is interesting to speculate on their neural bases. According to the traditional view of multisensory integration, information regarding motion in each sense is initially processed independently in modality-specific (or unimodal) brain areas (see Fig. 2). It is only at later stages of processing that information from different modalities converges in higher-order association areas. In accordance with this idea, visual motion processing has repeatedly been shown to involve visual area V5/MT+. Moreover, lesions in this part of the brain appear to impair just visual motion processing, while leaving auditory and tactile motion processing intact (Zihl, von Cramon, Mai, & Schmid, 1991). Similarly, researchers have also demonstrated the selective involvement of certain areas (such as the planum temporale, the inferior and superior parietal cortices, and the right insula) in the processing of auditory motion (Pavani, Macaluso, Warren, Driver, & Griffiths, 2002), whereas the primary and secondary somatosensory areas (SI and SII), located in the postcentral gyrus, play a major role in the perception of tactile motion (e.g., Hagen et al., 2002).

Recent studies have demonstrated that, in addition to these putatively modality-specific motion-processing areas, there are a number of brain areas that appear to be responsive to motion signals in more

than one sensory modality (e.g., Bremmer et al., 2001). Using functional magnetic resonance imaging (fMRI), these studies have shown that areas of the intraparietal sulcus, as well as the precentral gyrus (ventral premotor cortex, or PMv), can be activated by auditory, visual, or tactile motion signals (see Fig. 2). These findings converge with the results of animal studies using single-cell recording techniques (i.e., measuring the electrical activity in a single neuron), as such studies have found neurons that are responsive to both visual and somatosensory motion. Moreover, these neurons are found in brain regions that are homologous to the regions in the human brain that neuroimaging studies have indicated are responsive to multisensory motion signals (e.g., Duhamel, Colby, & Goldberg, 1998). It should be noted, though, that these recent fMRI studies failed to control for certain variables, such as non-motion-related activation (i.e., the baselines used were very different across the different modalities) and attentional factors (i.e., moving stimuli might simply capture attention more than stationary stimuli, and this might account for the differences in neural activity reported). However, when taken together, the evidence increasingly supports the idea that a network of brain areas is critically involved in processing motion information from more than one sensory modality.

Perhaps more controversially, Hagen et al. (2002) recently reported that tactile motion cues associated with the movement of a brush down the forearm resulted in the activation of V5/MT+. This suggests that what have traditionally been considered to be unimodal visual motion areas are involved in processing tactile motion as well (participants in this study had their eyes closed). At present, results showing recruitment of visual areas in a supposedly tactile task are open to an alternative interpretation in terms of visual imagery (e.g., activation in V5/MT+ may be evoked by visually imagining movement down the

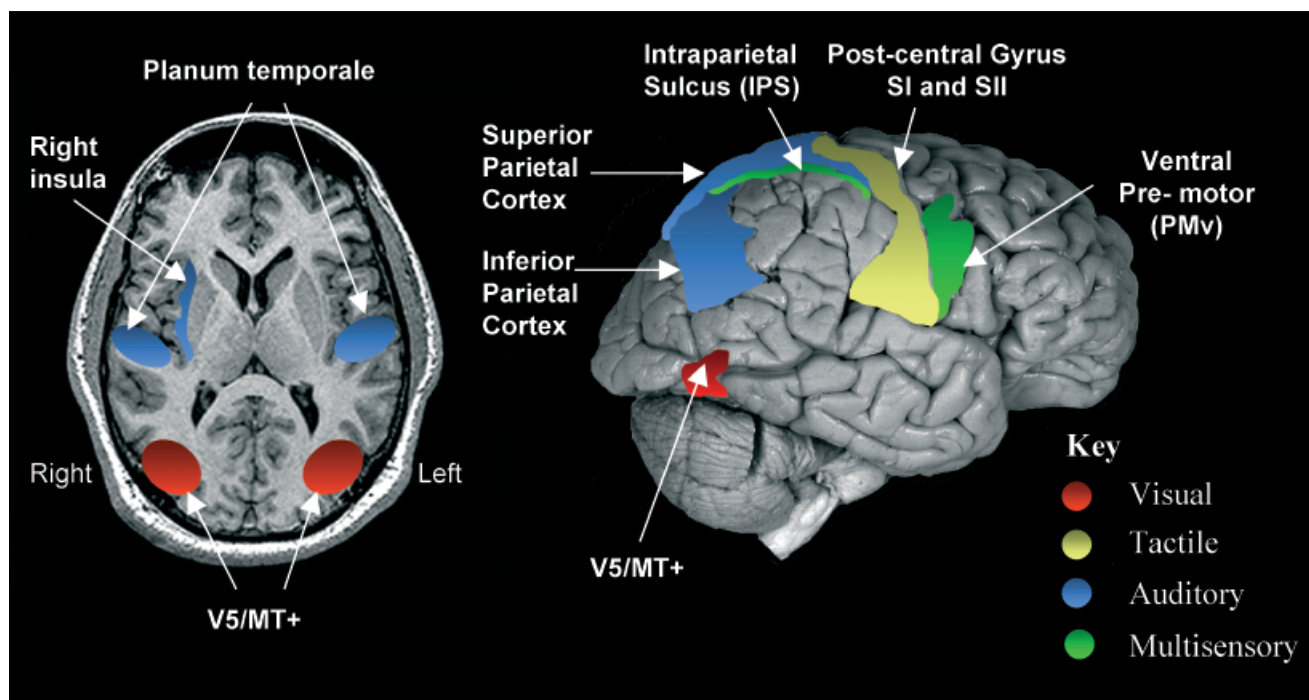


Fig. 2. An axial slice (a crosswise cut as seen from the bottom) magnetic resonance image of the brain (left illustration) and a right lateral view of the whole brain (right illustration) highlighting known neural substrates involved in unimodal visual, tactile, and auditory motion processing, as well as putative multisensory convergence sites for motion processing. SI and SII are the primary and secondary somatosensory areas, respectively.

arm, rather than being specifically associated with the processing of tactile motion). However, future studies, perhaps employing the technique of transcranial magnetic stimulation (TMS), which can be used to temporarily impair processing in V5/MT+, might help to rule out such a possibility. Indeed, the extensive reciprocal connections between the intraparietal sulcus multisensory area and V5/MT+ may reflect the crucial role of feedback from multisensory motion-processing sites to areas traditionally viewed as unimodal motion-processing areas. The existence of such feedback connections to unimodal motion areas may also help to explain the apparent conflict between the unimodal nature of our subjective experience of motion (presumably attributable to activation in putatively modality-specific brain structures) and the growing behavioral evidence for multisensory influences on early stages of motion perception (modulating the unimodal percept).

CURRENT DIRECTIONS AND UNRESOLVED QUESTIONS

The clear behavioral evidence for multisensory interactions and the existence of common neural substrates for multimodal motion processing suggest a number of important areas for future research. First, one issue that has plagued behavioral research is the distinction between cross-modal interactions at the perceptual level and those at a decisional level (i.e., caused by motivational or strategic factors that lead observers to respond on the basis of the erroneous sensory modality). Although many researchers have attempted to isolate purely perceptual effects, the relatively modest results of these studies, combined with recent findings indicating that perceptual and decisional processes can be intimately linked (Wohlschläger, 2000), mean that it may prove harder to disentangle their influence than some researchers have thought (e.g., Bertelson & de Gelder, in press).

Second, it seems that the way in which perceptual processes group incoming information within a single modality (principles of unimodal perceptual grouping) may also exert an important modulatory influence on the integration of multisensory motion cues, even under spatial and temporal coincidence, a situation that is optimal for integration. This not only could help explain some of the inconsistencies in previous research, but also could provide an interesting avenue for research in its own right. Finally, very different methods have been used in behavioral and neuroimaging studies. For example, most of the single-cell neurophysiological studies, in which polysensory motion areas have been identified, have presented tactile and visual motion stimuli at or near the face, whereas behavioral studies have traditionally used motion events that are at a distance from the observer. It will be important in future studies to use more similar paradigms in these different areas of research.

The perception of movement and the integration of multisensory cues are crucial for our adaptive interaction with the environment. It is, therefore, important that we develop our understanding of the constraints and mechanisms underlying multisensory contributions to motion perception, in order to derive a more comprehensive theory of perception with relevance for real-world situations.

Recommended Reading

Bremner, F., Schlack, A., Shah, N.J., Kubischik, M., Hoffman K.-P., Zilles, K., & Fink, G.R. (2001). (See References)

Hagen, M.C., Franzen, O., McGlone, F., Essick, G., Dancer, C., & Pardo, J.V. (2002). (See References)

Kitagawa, N., & Ichihara, S. (2002). Hearing visual motion in depth. *Nature*, *416*, 172–174.

Soto-Faraco, S., & Kingstone, A. (in press). Multisensory integration of dynamic information. In G. Calvert, C. Spence, & B. Stein (Eds.), *The handbook of multisensory processes*. Cambridge, MA: MIT Press.

Soto-Faraco, S., Kingstone, A., & Spence, C. (2003). (See References)

Acknowledgments—The authors would like to thank the Killam Trust for funding S.S.-F., the University of Oxford McDonnell-Pew Centre for Cognitive Neuroscience for funding a Network Grant to S.S.-F. and C.S., and the Human Frontier Science Program, National Science and Engineering Research Council of Canada, and Michael Smith Foundation for Health Research for funding A.K.

REFERENCES

Bertelson, P., & de Gelder, B. (in press). The psychology of multimodal perception. In C. Spence & J. Driver (Eds.), *Crossmodal space and crossmodal attention*. Oxford, England: Oxford University Press.

Bremner, F., Schlack, A., Shah, N.J., Kubischik, M., Hoffman, K.-P., Zilles, K., & Fink, G.R. (2001). Polymodal motion processing in posterior parietal and premotor cortex: A human fMRI study strongly implies equivalencies between humans and monkeys. *Neuron*, *29*, 287–296.

Duhamel, J.R., Colby, C.L., & Goldberg, M.E. (1998). Ventral intra-parietal area of the macaque: Congruent visual and somatic response properties. *Journal of Neurophysiology*, *79*, 126–136.

Ernst, M.O., & Banks, M.S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*, 429–433.

Hagen, M.C., Franzen, O., McGlone, F., Essick, G., Dancer, C., & Pardo, J.V. (2002). Tactile motion activates the human MT/V5 complex. *European Journal of Neuroscience*, *16*, 957–964.

Meyer, G.F., & Wuerger, S.M. (2001). Cross-modal integration of auditory and visual motion signals. *NeuroReport*, *12*, 2557–2560.

Pavani, F., Macaluso, E., Warren, J.D., Driver, J., & Griffiths, T.D. (2002). A common cortical substrate activated by horizontal and vertical sound movement in the human brain. *Current Biology*, *12*, 1584–1590.

Sekuler, R., Sekuler, A.B., & Lau, R. (1997). Sound alters visual motion perception. *Nature*, *385*, 308.

Soto-Faraco, S., Kingstone, A., & Spence, C. (2003). Multisensory contributions to the perception of motion. *Neuropsychologia*, *41*, 1847–1862.

Soto-Faraco, S., Lyons, J., Gazzaniga, M., Spence, C., & Kingstone, A. (2002). The ventriloquist in motion: Illusory capture of dynamic information across sensory modalities. *Cognitive Brain Research*, *14*, 139–146.

Soto-Faraco, S., Spence, C., & Kingstone, A. (in press). Cross-modal dynamic capture: Congruence effects in the perception of motion across sensory modalities. *Journal of Experimental Psychology: Human Perception and Performance*.

Stein, B.E., & Meredith, M.A. (1993). *The merging of the senses*. Cambridge, MA: MIT Press.

Welch, R.B., & Warren, D.H. (1986). Intersensory interactions. In K.R. Boff, L. Kaufman, & J.P. Thomas (Eds.), *Handbook of perception and performance: Vol. 1. Sensory processes and perception* (pp. 25–1–25–36). New York: John Wiley and Sons.

Wohlschläger, A. (2000). Visual motion priming by invisible actions. *Vision Research*, *40*, 925–930.

Zihl, J., von Cramon, D., Mai, N., & Schmid, C. (1991). Disturbance of movement vision after bilateral posterior brain damage: Further evidence and follow up observations. *Brain*, *114*, 2235–2252.