PERGAMON

Neuropsychologia 41 (2003) 1847–1862

NEUROPSYCHOLOGIA

www.elsevier.com/locate/neuropsychologia

Multisensory contributions to the perception of motion

Salvador Soto-Faraco[a,*], Alan Kingstone[b], Charles Spence[c]

[a] Departament de Psicologia Bàsica, Universitat de Barcelona, Pg. Vall d'Hebrón, 171, 08035 Barcelona, Spain
[b] University of British Columbia, Vancouver, Canada
[c] Department of Experimental Psychology, University of Oxford, Oxford, UK

Received 9 May 2003; accepted 20 June 2003

Abstract

The ability to process motion is crucial for coherent perception and action. While the majority of studies have focused on the unimodal factors that influence motion perception (see, for example, the other chapters in this Special Issue), some researchers have also investigated the extent to which information presented in one sensory modality can affect the perception of motion for stimuli presented in another modality. Although early studies often gave rise to mixed results, the development of increasingly sophisticated psychophysical paradigms are now enabling researchers to determine the spatiotemporal constraints on multisensory interactions in the perception of motion. Recent findings indicate that these interactions stand over-and-above the multisensory interactions documented previously for static stimuli, such as the oft-cited 'ventriloquism' effect. Neuroimaging and neuropsychological studies are also beginning to elucidate the network of neural structures responsible for the processing of motion information in the different sensory modalities, an important first step that will ultimately lead to the determination of the neural substrates underlying these multisensory contributions to motion perception.

Keywords: Motion perception; Somatosensation; Audition; Vision; Multisensory

1. Introduction

Our ability to extract information regarding movement is fundamental both for deriving and maintaining useful representations of the environment, and also for the planning and execution of action. As for many other perceptual domains, information regarding the movement of stimuli (i.e. their direction, speed, etc.) is often available via several sensory modalities simultaneously (think, for example, of tracking a partially occluded animal moving through the undergrowth). While vision may provide the most salient information with regard to stimulus motion, audition and somatosensation can also provide important cues, particularly when stimuli are occluded, or else move outside the current field of view (such as when objects move behind the head).

Despite its obvious adaptive importance, and after more than a century of intermittent research on this topic (see Ryan, 1940; Urbantschitsch, 1902; Zietz & Werner, 1927, for early work), our knowledge concerning how motion signals from different sensory modalities are integrated is still fairly limited. Nevertheless, there has been a rapid growth of interest in this area over the last few years, with researchers moving from the phenomenological approaches that characterized many early studies (e.g. Zapparoli & Reatto, 1969) to the development of increasingly sophisticated psychophysical paradigms that are currently enabling researchers to investigate the spatio-temporal constraints on multisensory contributions to the perception of dynamic stimuli.[1] Our understanding of these constraints on human perception has also been complemented by recent cognitive neuroscience studies that have begun to reveal the networks of neural structures involved in the integration of motion information across the senses.

In the present review, we start by evaluating the behavioural evidence regarding multisensory contributions to the perception of motion, as well as outlining some of the key constraints affecting such integration. We then proceed

---

* Corresponding author. Tel.: +34-93-312-5158; fax: +34-93-402-1363.
E-mail address: ssoto@psico.psi.ub.es (S. Soto-Faraco).

[1] Despite early claims for the perception of 'multisensory motion' based on phenomenological observation (e.g. Zapparoli & Reatto, 1969), the majority of contemporary researchers have argued that there is no such thing as 'multisensory motion' per se (i.e. intermodal apparent motion; e.g. Allen & Kolers, 1981; see also Hirsh & Sherrick, 1961), but instead there are only multisensory contributions to the unimodal perception of motion. That is, they believe that the phenomenal perception of motion is always tied to (or experienced within) a single sensory modality, although such unimodal perceptual experience can be modulated by the occurrence of stimuli (either moving or stationary) in other sensory modalities (cf. Driver & Spence, 2000).

0028-3932/$ – see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/S0028-3932(03)00185-4

to highlight recent neuroimaging and neurophysiological data regarding the neural substrates of motion processing that may underlie some of these multisensory perceptual interactions.[2]

2. Multisensory contributions to the perception of motion: behavioral evidence

For many years, researchers have investigated the question of whether the presentation of stimuli in one modality (either moving or stationary) can affect the perception of motion of stimuli presented in another modality (see Soto-Faraco & Kingstone, in press, for a detailed review). Many of these behavioral studies have examined whether the presentation of a static stimulus in one modality affects various aspects of motion processing in another modality, such as its trajectory (e.g. Hall & Earle, 1954; Hall, Earle, & Crookes, 1952; Lakatos, 1995; Sekuler, Sekuler, & Lau, 1997; Spelke, Born, & Chu, 1983; Watanabe & Shimojo, 2001a,b), its speed (e.g. Manabe & Riquimaroux, 2000), or the threshold for the perception of apparent motion (e.g. Allen & Kolers, 1981; Gilbert, 1939; Maass, 1938; Ohmura, 1987; Staal & Donderi, 1983; Zietz & Werner, 1927).

Other researchers have looked at the influence of the presentation of a moving stimulus in one modality upon the perception of a stationary stimulus presented in another modality (e.g. Ehrenstein & Reinhardt-Rutland, 1996; Mateeff, Hohnsbein, & Noack, 1985; Thurlow & Kerr, 1970). For example, Mateeff et al. (Experiment 1) reported that while participants were tracking a moving visual target with their eyes, a stationary sound presented for about 1 s appeared to move in the same direction as the visual stimulus. However, apparently contradictory results were reported by Ehrenstein and Reinhardt-Rutland. In their study, participants were placed at the centre of a visually-textured drum that was rotated around them for 2.5 min in either a clockwise or anticlockwise direction. An auditory stimulus, presented for 2–6 s from a loudspeaker cone placed behind the drum shortly after it had stopped rotating, was mislocalized in the opposite direction to that of the visual adapting stimulus. However, no impression of auditory movement was reported by the participants in this study. It is possible that the apparent auditory motion reported in Mateeff et al.'s study may simply reflect an artefact of the use of a forced choice response paradigm. That is, even if participants did not experience any auditory motion, they were still forced to respond that the sound moved either to the left or to the right on every trial. Meanwhile, the discrepancy between the direction of displacement of the auditory stimuli reported in these two studies may reflect the differing extents to which the movement of the visual stimuli/environment elicited sensations of self- as opposed to object-motion (see Arnoult, 1952; Gemelli, 1951; Thurlow & Kerr, 1970).

Researchers have also shown that the presentation of two stationary stimuli in different modalities can sometimes elicit a sensation of movement within one of the modalities (e.g. Hikosaka, Miyauchi, Takeichi, & Shimojo, 1996; Shimojo, Miyauchi, & Hikosaka, 1997; see also Maass, 1938). For example, Shimojo et al. reported that a horizontal line presented on a screen can actually appear to shoot from left-to-right (or right-to-left) if an auditory or tactile stimulus is presented on the left (or right, respectively) just before the line itself is presented.[3] However, these studies of the line-motion illusion, as well as the other studies discussed so far, are only marginally informative with regard to the question of how multisensory motion cues are integrated. This is because they do not specifically address the perceptual consequences of processing concurrent motion signals in different sensory modalities.

There are, however, a number of studies that have more directly addressed the effects of movement from one modality on the perceived movement of stimuli in another modality. Early studies by Zapparoli and Reatto (1969) and Anstis (1973; see p. 338) suggested that the direction of visual motion can have a profound effect on the perceived direction of auditory motion. However, given that these early studies used a phenomenological approach, their descriptive results should be treated, at best, as suggestive. Indeed, a subsequent empirical investigation by Allen and Kolers (1981) using psychophysical methods failed to support these early descriptive results. They found that both moving and static visual distractors increased the likelihood of people experiencing auditory apparent motion with respect to an auditory alone condition. In fact, both the static and the moving visual stimuli had a very similar crossmodal influence on the perception of auditory apparent motion stimuli, regardless of the relative direction (i.e. same versus opposite) of auditory and visual motion. Similarly, Staal and Donderi (1983) reported that the sequential presentation of a pair of 1000 Hz tones, one presented to either ear over headphones, reduced the threshold for the perception of apparent motion of a pair of visual stimuli (presented via LEDs situated directly in front of the observers). This

---

[2] Due to space constraints, this review will be restricted to evaluating multisensory contributions to the perception of movement of external stimuli. It should, however, be noted that there is an equally voluminous (though largely independent) literature investigating multisensory contributions to the perception of self-motion (typically focusing more on the integration of kinaesthetic, proprioceptive, gravitational, and visual cues; e.g. Dichgans & Brandt, 1978; Jeka, Oie, & Kiemel, 2000; Maioli & Poppele, 1991). There is also a growing body of research looking at the consequences of self-movement, or action, on the perception of movement of external stimuli (e.g. Arnoult, 1952; Gemelli, 1951; Jackson, 1953; Thurlow & Kerr, 1970; Wohlschläger, 2000; Wohlschläger & Wohlschläger, 1998; see also Neff, 1936, for a review of early studies).

[3] Although the 'line-motion' illusion, has been considered by some researchers to reflect an effect of attention in one modality on the perception of motion in another modality (e.g. Shimojo et al., 1997; see also Cavanagh, 1992), other researchers have argued for a non-attentional (low-level perceptual) account for this effect instead (e.g. see Downing & Treisman, 1997).

reduction in threshold took place both when the auditory and visual stimuli moved in the same direction and when they moved in opposite directions (although the reduction was numerically larger for congruent movement, it is unclear from Staal and Donderi's study whether the difference between these conditions reached statistical significance or not).

Mateeff et al. (1985) also conducted two further experiments in their study first described earlier in this article. In one experiment, participants judged the direction of motion of a briefly-presented sound moving to the left or right while they simultaneously tracked a visual stimulus moving in the same or opposite direction. The motion of the auditory stimulus was adjusted according to a psychophysical staircase procedure to determine the actual velocity and direction necessary for the sound to be perceived as stationary (computed as the point where participants made left and right responses with equal probability). The results showed that auditory stimuli needed to be presented in the contrary direction, and at a velocity of 25–50% of that of the visual target, in order to appear as being 'static'. This implies that the perception of auditory motion tended to be captured by the direction of visual motion. In a final experiment, Mateeff et al. also found that while the greatest visual capture effects occurred when participants overtly tracked the visual stimulus, a robust, albeit somewhat reduced, visual capture effect was still reported when participants maintained central fixation throughout the trial.[4] These latter results therefore represent some of the strongest evidence for an interaction in the processing of auditory and visual motion, prior to the recent rekindling of interest in this topic (see below).

The recent resurgence of interest in multisensory phenomena has resulted in the development of a variety of novel psychophysical paradigms to investigate the nature of multisensory contributions to motion perception (e.g. see Kitagawa & Ichihara, 2002; Kitajima & Yamashita, 1999; Meyer & Wuerger, 2001; Shams, Allman & Shimojo, 2001; Soto-Faraco, Spence, & Kingstone, 2003; Vroomen & de Gelder, in press; Wuerger, Hofbauer & Meyer, 2002a). These studies have raised a number of important issues regarding the multisensory integration of motion information, such as the spatiotemporal modulation of such effects, the level of information processing at which such multisensory interactions take place, the hierarchy of sensory dominance for motion signals presented in different modalities, and the role of attention in both determining sensory dominance and modulating multisensory integration effects. These issues are discussed in more detail in the sections that follow.

2.1. Spatiotemporal modulation of multisensory contributions to motion perception

An extensive body of behavioural and neuroscience research has demonstrated that the most powerful multisensory integration effects occur when stimuli from different sensory modalities are presented from the same position at approximately the same time, and fall off as the spatial and/or temporal separation between the stimuli increases (e.g. see King & Palmer, 1985; Stein & Meredith, 1993). Importantly, however, when the early studies on multisensory integration of motion are seen in this light, it becomes unclear whether the null results reported in certain of the early studies on crossmodal motion perception (e.g. regarding the effects of directional congruency, Allen & Kolers, 1981; Staal & Donderi, 1983) reflect a genuine lack of interaction in motion signals between different sensory modalities, or simply a failure to integrate stimuli that happened to have been presented from very different spatial locations.[5] In the real world, stimuli coming from the same location are presumably more likely to originate from the same object and hence to provide consistent information concerning its movement. Given such an assumption of common origin, one might expect the majority of multisensory integration and sensory dominance effects to take place when there is a coincidence of the spatiotemporal characteristics of the various dynamic stimuli. Support for the importance of spatial factors in determining the results of previous studies comes from research showing that the magnitude of any multisensory contributions to motion perception decreases as the spatial distance between stimuli presented to different modalities increases (e.g. Hall & Earle, 1954; Meyer & Wuerger, 2001; Soto-Faraco, Lyons, Gazzaniga, Spence, & Kingstone, 2002; Wuerger et al., 2003). Similarly, several researchers have also demonstrated the importance of temporal synchrony for maximizing any multisensory contributions to motion perception, with multisensory contributions to motion perception falling off as the temporal asynchrony between the component signals increases (e.g. Ohmura, 1987; Soto-Faraco et al., 2002, in press).

2.2. Level of processing

Despite the divergence in methods used and results reported, some earlier studies suggested a robust effect of motion signals in one modality on the perception of motion signals in another modality (e.g. Anstis, 1973; Mateeff et al., 1985; Zapparoli & Reatto, 1969). However, it is important to try and distinguish the level(s) of processing at which such multisensory interactions take place. More

---

[4] This is an important result because it demonstrates the potential role of eye movements in modulating visual capture effects. This eye movement confound may also compromise the interpretation of many other studies, where no attempt was made to ensure, or monitor, constant fixation.

[5] Note also that other studies in which significant effects were found, but where different modality stimuli were presented from different locations or along different trajectories, may also have underestimated the true magnitude of any effects they reported (e.g. Ehrenstein & Reinhardt-Rutland, 1996; Gilbert, 1939; Ohmura, 1987; Mateeff et al., 1985).

specifically, a point of contention in the literature pertains to the distinction between multisensory integration processes that occur during the perception of stimuli, and other influences more related to decisional and/or response processes (i.e. processes that are not specific to motion perception per se, perhaps taking place after the properties of each individual motion signal has been computed). In simple terms, the former would affect the way in which the stimuli are actually experienced, while the latter would modify the way in which the responses to the stimuli are selected or executed, without necessarily implying changes in one's phenomenal experience. As with many research areas, it has proved difficult to determine the relative contribution of perceptual versus post-perceptual (or cognitive) factors to the behavioral effects reported in the laboratory (see, for example, Bertelson, 1998; Bertelson & Aschersleben, 1998; Choe, Welch, Guilford, & Juola, 1975; Soto-Faraco, Spence & Kingstone, 2003; Welch, 1999; Wohlschläger, 2000, on this issue).

Meyer and Wuerger (2001) attempted to address this topic by measuring their participants' ability to discriminate the direction of visual motion of random dot kinematograms as a function of the directional congruency of an irrelevant sound source (presented from pairs of loudspeaker cones situated either behind, or to one or the other side of the visual display). Their results were best fitted by a model assuming that the visual and auditory motion signals were perceived independently, and that the congruency effect of auditory motion (found when the visual signal was ambiguous) could be accounted for entirely by decisional biases (see also Wuerger et al., 2002a, for similar results). Meyer and Wuerger's results therefore highlight the important role that post-perceptual factors can have in the interpretation of crossmodal interactions in the domain of motion processing (see also Bertelson, 1998; Choe et al., 1975; Welch, 1999). The critical question remains whether, besides the generic non-perceptual processes highlighted by Meyer and Wuerger, there are also genuinely perceptual interactions that are specific to the processing of motion information.

One source of evidence that suggests potential interactions at early levels of sensory information processing comes from studies of adaptation after-effects (Kitagawa & Ichihara, 2002; Vroomen & de Gelder, in press). For example, Kitagawa and Ichiara studied crossmodal after-effects in the perception of motion in depth. They reported that a sound usually perceived as fixed in intensity was judged as if it was decreasing in intensity (i.e. as if it was receding) after adaptation to a looming square, or as increasing in intensity (i.e. as if it was approaching) after adaptation to a receding square. Although these results are consistent with an interpretation of multisensory interactions occurring at an early level of stimulus processing, the data are not conclusive on this point. Indeed, the potential contribution of post-perceptual processes such as response compatibility effects between the potential responses elicited by the target (increasing versus decreasing) and the direction of the adaptor stimulus (looming versus receding) cannot be ruled out.[6] As pointed out by several authors, these post-perceptual influences do not only take place for immediate sensory interactions, but can also influence after-effects as well (e.g. Bertelson & Aschersleben, 1998; Choe et al., 1975; Welch, 1999).

In another recent study, Vroomen and de Gelder (in press) addressed the contribution of the perceptual (versus post-perceptual) levels of processing using an auditory motion after-effect (the contingent auditory motion after-effect). In the critical conditions, the moving auditory adaptors were accompanied by either congruent or incongruent visual motion stimuli that participants had to track with their eyes. The results showed that the auditory after-effect was contingent upon the direction of the visual adaptor and not the auditory adaptor. Although Vroomen and de Gelder used a procedure in which the response alternatives were not orthogonal to the adapting stimuli (and are therefore also prone to post-perceptual response compatibility effects), the fact that static auditory adaptors accompanied by moving visual adaptors did not produce the after-effect supports a perceptual interpretation of their findings.

Some findings using online measures (i.e. immediate perceptual responses as opposed to after-effects) also suggest that the integration of motion signals can occur at specifically perceptual stages of processing. For example, Soto-Faraco et al. (in press) demonstrated that people could judge the direction of an otherwise unambiguous auditory apparent motion stream (moving either leftward or rightward) almost perfectly when a directionally congruent visual apparent motion stream was presented concurrently (or when no visual motion was present at all). In contrast, participants responded incorrectly on as many as half of all the trials when the visual motion occurred in the direction opposite to that of the apparent auditory motion (see Fig. 1). This crossmodal dynamic capture effect remained even when confidence ratings were taken after each response, and only those trials in which participants were highly confident about their judgements were included (Soto-Faraco et al., in press; see Fig. 1d). This finding therefore suggests that visually incongruent motion did not simply reduce the response certainty for participants (which would have increased the potential for cognitive factors to influence their decisions), but instead resulted in the sounds appearing to move in the same direction as the lights (i.e. visual capture of the perceived direction of auditory motion).

In a more recent study, Soto-Faraco et al. (submitted for publication) provided another demonstration of the genuinely perceptual basis of the crossmodal dynamic capture effect. Soto-Faraco et al. used psychophysical staircases

---

[6] Even when there is no potential congruency relationship between responses and irrelevant information, subtler cognitive influences may also influence performance (cf. Hubbard, 1995). Indeed, the mere awareness that some trials in an experiment contain conflicting information, while others do not, could form the basis for criterion shifts by observers (Bertelson & Aschersleben, 1998; Welch, 1999).

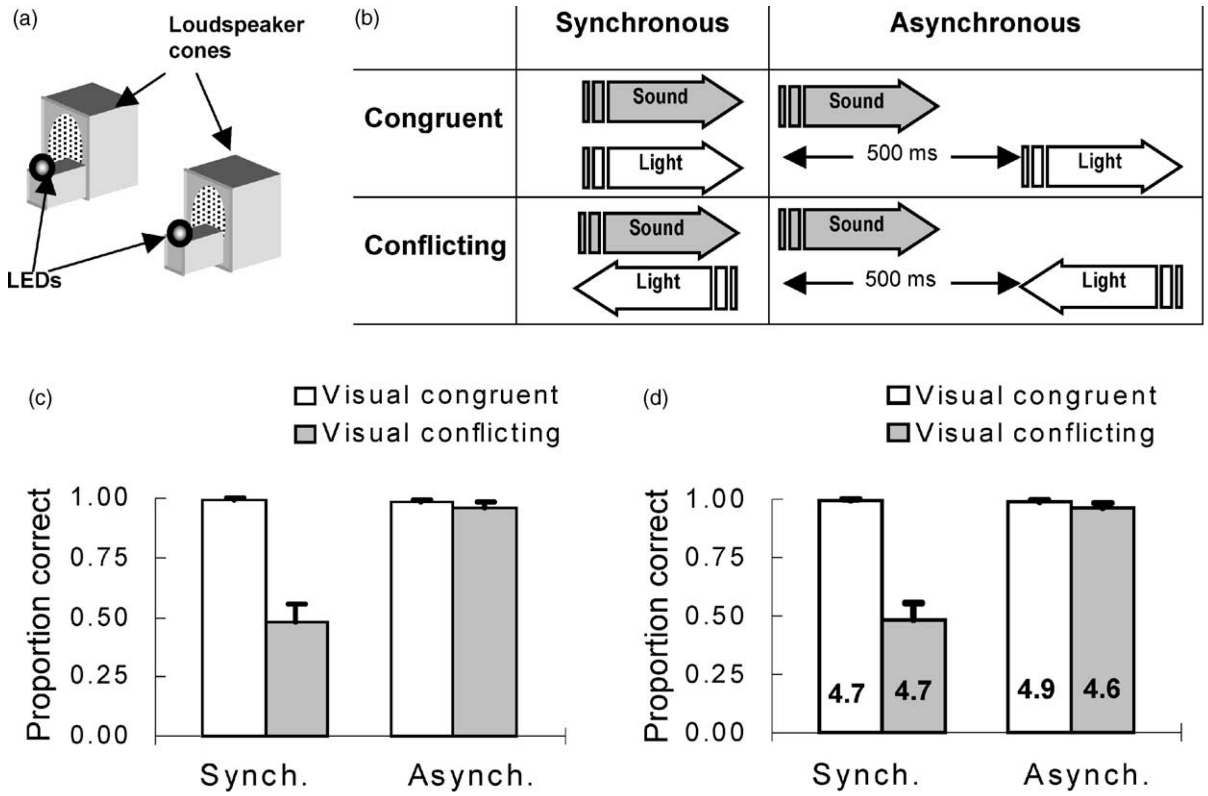

Fig. 1. (a) A schematic outline of the experimental set-up used in Soto-Faraco et al. (2002, in press). (b) A summary chart of the type of trials included in the experiments. These four types of trial (as well as the complementary ones where auditory motion was from right-to-left) were presented repeatedly, and in a random order. (c) Results from the auditory motion direction discrimination task (proportion correct + S.E.) as a function of the relative direction and synchrony of the visual motion distractor. (d) Data from a further experiment (Soto-Faraco et al., in press) in which participants rated their confidence (on a scale 1–5, where 1 was "guessing" and 5 was "positive") after each response. Only high confidence trials were included in this analysis. The number at the base of each bar is the average confidence rating for that condition (error bars once again represent + S.E.).

to assess the threshold SOA for the accurate perception of motion direction under unimodal and bimodal conditions. In unimodal displays, participants were presented with one apparent motion stream (either visual or auditory) and their task was to discriminate its direction (left or right). The SOA of each type of apparent motion stream (i.e. the interval between the onset of the first and the second sound, or the first and second light) was adjusted according to several interleaved staircases (ranging between 1000 ms SOA right-to-left streams and 1000 ms left-to-right streams) with variable step size. The staircases were iterated until a point of perceptual uncertainty was found (i.e. the average SOA where response reversals occurred). In bimodal displays, participants were presented with two apparent motion streams simultaneously (one in each modality) and they had to judge if the two motion streams moved in the same or different directions. The SOA of the apparent motion streams was again adjusted according to several interleaved staircases (ranging between 1000 ms SOA streams moving in the same direction and 1000 ms streams moving in opposite directions) until a point of perceptual uncertainty was reached (i.e. the SOA at which response reversals occurred, indicating that participants were unsure about if the two streams moved in the same direction or in opposite directions).

The data revealed that participants perceived the direction of motion in unimodal displays (visual or auditory) very accurately, even when the SOA was at its minimum value of 75 ms (see Fig. 2). However, in bimodal displays they could not discriminate whether two simultaneously presented apparent motion streams (one visual and the other auditory) moved in the same or different directions unless the SOA was greater than 300 ms. Moreover, for SOAs below 300 ms, participants systematically classified different-direction displays as same-direction displays (thus replicating the dynamic 'capture' result from Soto-Faraco et al. (2002, in press) discussed above). These data indicate that (1) it is not possible to evaluate the visual and auditory motion components of the bimodal displays independently at SOAs where information about the direction of motion is readily available under unimodal conditions; and (2) that the two motion signals appear to move in the same direction when they are, in fact, moving in different directions (indicating the potential dominance of one modality over the other). It is important to note that this psychophysical method controls for two important post-perceptual confounds that are often present in crossmodal studies in this area. First, in the bimodal condition there was no potential conflict between the available responses (same versus different), and the direction

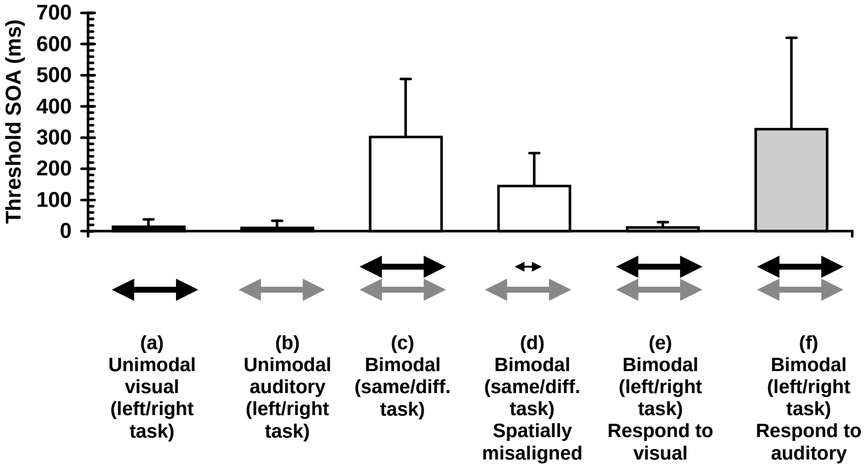

Fig. 2. Results in the staircase study reported by Soto-Faraco et al. (submitted for publication). Bars show the average SOA threshold (+S.D.) of perceptual uncertainty in each condition (described below). This threshold was assessed using the average SOA value of the last 12 reversals of each staircase for each participant. The higher the threshold the less accurate the performance. In the unimodal conditions, (a) and (b), the bars represent the average SOA value at which the left/right judgments were less certain. These staircases started with 1000 ms SOA streams (moving from left-to-right or from right-to-left) and the SOA was adjusted for each staircase in successive trials according to the participant's responses. The results show that this uncertainty SOA was close to 0 ms, indicating that unimodal perception was very accurate. In bimodal displays with the same/different task, (c) and (d), the bars represent the SOA at which participants were more uncertain in discriminating displays containing two simultaneous apparent motion streams moving in the same direction from displays containing two simultaneous apparent motion streams moving in opposite directions. The results show that when the two inputs were spatially aligned discrimination was worse than in the unimodal conditions and worse than when the two inputs were misaligned (see text). In bimodal displays with the left/right discrimination task, participants judged the direction of the visual (e) or the auditory apparent motion stream (f) from a display containing two apparent motion streams, one in each modality. When participants responded to the visual stream, their performance was as accurate as in the unimodal visual displays. However, when the participants responded to the auditory stream, the threshold for the left/right motion discrimination was as high as in the bimodal displays using the same/different task (c).

in which motion occurred in each modality (leftward versus rightward). That is, the response dimension was orthogonal to the directional information present in either modality (note that this contrasts with previous psychophysical studies; e.g. Kitagawa & Ichihara, 2002; Mateeff et al., 1985, Experiments 2 and 3; Vroomen & de Gelder, in press; see also Spence, Shore, & Klein, 2001b, on this issue). Second, the thresholds found reflected the point of perceptual uncertainty, and therefore, by definition the point at which participants became unaware as to whether the direction of motion of the two modalities was in conflict or not (see Bertelson & Aschersleben, 1998; Caclin, Soto-Faraco, Kingstone, & Spence, 2002, for similar approaches). This is important because without the knowledge about conflict, any potential cognitive bias cannot operate differentially for conflicting versus congruent motion conditions.

Two further controls were, however, necessary to ensure that these results reflected the consequences of multisensory integration. In one, we ruled out the potential account that performance in the bimodal condition was worse than in the unimodal condition simply because the necessity of having to attend to two streams instead of just one may have increased the attentional load sufficiently to impair performance (Spence, Nicholls, & Driver, 2001a). We repeated the staircase methodology, but now misaligned the visual and auditory inputs—a manipulation that should weaken multisensory integration (Stein & Meredith, 1993) relative to the previous study, while matching for any changes in attentional load. In this experiment, the discrimination thresholds decreased significantly. That is, discrimination of same-different direction displays was easier for misaligned inputs than for aligned inputs although the attentional requirements of the task were the same in both cases. In the other control experiment, we addressed the potential issue that bimodal displays were responded to less accurately than unimodal displays because a different type of discrimination response was used (same/different versus left/right discrimination, respectively). We ran the bimodal staircases again, but now used a left/right discrimination task instead of a same-different task. We recreated the divided attention conditions of the original bimodal experiment by leaving participants uncertain as to which modality they would have to respond to until after the displays had been presented (using an audiovisual post-stimulus response cue). In this experiment, we found that visual directional discrimination was as accurate as it had been in the unimodal condition. In contrast the thresholds for auditory directional discrimination were nearly identical to the bimodal thresholds obtained in the original same-different task. This result indicates that performance in the bimodal conditions was not due to the type of output task (or response) used. It also reveals an interesting asymmetry between auditory and visual motion cues that will be discussed in the next section.

To summarize, the question about the level of processing at which crossmodal effects occur has important implications for theories that seek to describe the mechanisms subserving multisensory integration. However, the issue is both complex and controversial. With the recent development of increasingly refined psychophysical tools, the picture that is emerging is one of different processing levels that are intimately linked, and both perceptual and post-perceptual stages of processing seem to play an important role. Indeed, while the perceptual/decisional distinction may be an intuitively plausible one, even defining these terms in a formal manner has proven to be a difficult task for researchers in the field.

2.3. Asymmetries

A typical feature of many classes of crossmodal interactions is that they often show modality dominance. For example, in the ventriloquist illusion, visual stimuli alter the perceived location of sounds, but the reverse effect is much weaker, if found at all (see Bertelson & de Gelder, in press, for a review). The same type of asymmetry (i.e. visual dominance) has also been found in many other domains as well (e.g. shape/size in touch-vision; Rock & Harris, 1967), whereas the reverse asymmetry (i.e. sound altering visual perception) has often been reported for tasks involving some form of temporal discrimination (e.g. Morein-Zamir, Soto-Faraco & Kingstone, 2003; Repp & Penel, 2002; Shams, Kamitani, & Shimojo, 2000; Welch, DuttonHurt, & Warren, 1986). In the domain of motion perception, the results of various studies suggest a dominance of visual motion cues over both auditory (Kitagawa & Ichihara, 2002; Soto-Faraco et al., in press, submitted for publication; though see also Manabe & Riquimaroux, 2000) and somatosensory motion cues (Soto-Faraco et al., 2000). For example, Kitagawa and Ichiara found that visual adaptors produced auditory motion after-effects, but that auditory adaptors did not produce visual motion after-effects (note that a similar asymmetry was also highlighted in early studies involving phenomenological reports; e.g. Anstis, 1973; Zapparoli & Reatto, 1969; see also Allen & Kolers, 1981; Ohmura, 1987).

However, the appropriate interpretation of such asymmetries is by no means obvious, because they can reflect a variety of qualitatively distinct causes. On the one hand, they might reflect the underlying organization of the perceptual systems (e.g. visual motion input receives a higher weighting in the integration process than the motion cues presented to other sensory modalities). On the other hand, these asymmetries could just as well be caused by the particular properties of the stimuli used (e.g. perhaps the visual inputs used in the majority of studies were simply more intense, or elicited a better ‘quality’ of motion). In an experiment addressing the asymmetry issue, Soto-Faraco et al. (2003) found that the congruency effects of visual motion direction on the perception of auditory motion direction (discussed above) did not occur in the reverse direction. In their study, auditory apparent motion did not influence the perception of visual motion direction even when using an ISI where the apparent motion threshold of the two modalities was equivalent, suggesting that visual dominance in motion may reflect intrinsic processing characteristics rather than a particular imbalance between the two inputs.[7]

2.4. The role of attention and perceptual grouping

One important aspect in the interpretation of perceptual asymmetries is the relative distribution of *attention* across different sensory modalities. Many studies have shown that the degree to which motion information is processed depends on the allocation of attention both within a particular sensory modality (e.g. Alais & Blake, 1999; Beer & Röder, 2002; Chaudhuri, 1990; see Raymond, 2000, for a review), as well as between sensory modalities (e.g. Berman & Colby, 2002; Spence, Chan & Simm, 2000; though see Rees & Lavie, 2001). Following on from the previous section, several researchers have postulated that visual input may often dominate over input from other modalities because of a pervasive bias to attend preferentially toward the visual modality, rather than to any of the other sensory modalities (e.g. Posner, Nissen, & Klein, 1976; Spence et al., 2001b). Consistent with such a view are older studies that have reported sensory dominance patterns that are to some degree malleable by specific attentional manipulations (e.g. Canon, 1970, 1971; Warren & Schmitt, 1978). However, the majority of more recent studies have failed to demonstrate any role of attention on multisensory integration in the static ventriloquism effect (e.g. Bertelson, Vroomen, de Gelder, & Driver, 2000; Driver, 1996; Spence & Driver, 2000; Vroomen, Bertelson, & de Gelder, 2001; see Bertelson & de Gelder, in press, for a review).

Given this mixed pattern of results in previous studies, we recently studied the effects of attention on the crossmodal dynamic capture phenomenon reported earlier. In a first experiment, we used simultaneous auditory and visual apparent motion displays to evaluate the effects of dividing attention between the two modalities (participants were uncertain about which modality they would have to report on). When compared with the usual focused attention conditions (where participants knew in advance which modality constituted the target modality), any modulation in dynamic capture attributable to the direction of attention was found to be only marginal (see Fig. 3a and b).

---

[7] Even when the two modalities are matched in one or more respects, it is difficult to ensure a perfect equilibrium that renders a clear interpretation in terms of information processing differences, rather than simply in terms of stimulus properties per se. However, this reflects the more general problem of assessing modality dominance present in crossmodal literature as a whole, given the difficulties associated with trying to match stimulus intensity across different sensory modalities (see Soto-Faraco & Kingstone, in press; Spence et al., 2001b). One potential way to circumvent this problem would be to conduct parametric studies using different saliency values across the two modalities involved. Along these lines, Soto-Faraco et al. (2002) found that large variations in sound intensity had only a mild effect on the degree of crossmodal dynamic capture.

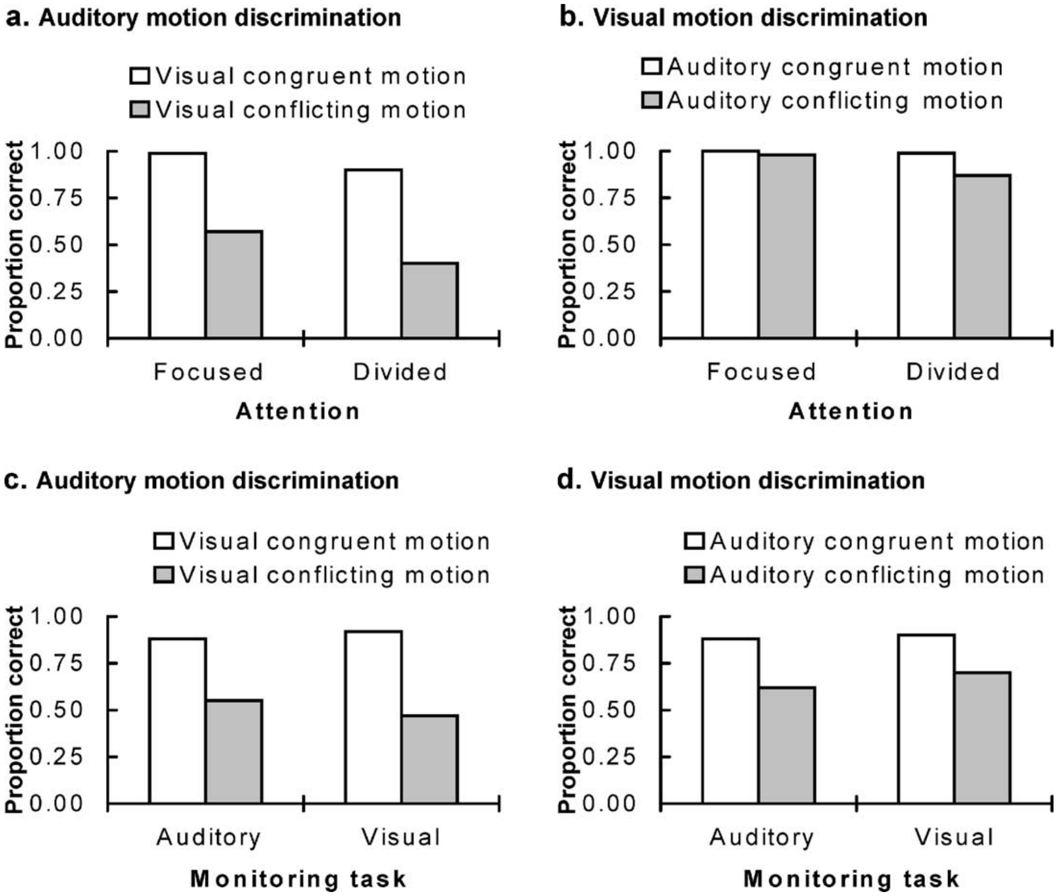

Fig. 3. Results from two additional experiments reported in this review. (a) and (b) show the results from an experiment in which the simultaneous apparent motion streams were presented under two different attention conditions: focused (attend to the sound motion or attend to the visual motion) or divided (attend to both motion signals simultaneously). Accuracy in discriminating the direction of auditory motion as a function of congruency of visual distractors and attention condition is shown in (a), and the accuracy of discriminating the direction of visual motion is shown as a function of congruency of auditory distractors and attention in (b). (c–d) Represent the data from a second experiment where the same motion discrimination task was performed in combination with a monitoring task at fixation (searching for pre-defined targets in a rapidly-presented visual or auditory stream). The modality of the target motion display in the directional discrimination task and the stream containing the target in the monitoring task were varied systematically. (c) Shows auditory motion discrimination performance as a function of visual congruency and the sensory modality of the monitored stream. (d) Shows visual motion discrimination performance as a function of auditory congruency and the modality of the monitored stream.

In a subsequent experiment, we addressed whether this marginal effect of attention on dynamic capture would reach significance if the participant's task was made more attentionally-demanding (cf. Spence et al., 2001a). The participants in this new experiment were required to perform a switch between a sustained attention task and a direction of motion discrimination task. The primary task was to monitor constantly one of two rapidly presented streams of stimuli (one visual and the other auditory), both located centrally at fixation, in search for pre-specified targets. The secondary task was to respond to the direction of the visual or the auditory component of a display consisting of two simultaneous apparent motion streams (one presented in each modality) that could be directionally congruent or conflicting. The motion discrimination trials were identical to those used in previous experiments and interspersed at unpredictable times, during the sustained monitoring task. All motion trials contained a task-switch (from the target monitoring to the motion direction task), and half of them also contained a modality switch (i.e. from monitoring the auditory stream and responding to visual motion or vice versa).

In the auditory motion direction task, the congruency effect (as a function of visual motion direction, see Fig. 3c) was again significant, replicating our previous results. In the visual motion task, for the first time, we observed a small (22%) but significant congruency effect (as a function of auditory motion direction, see Fig. 3d). No effects of modality-switching were detected over and above the task-switch present in all conditions (e.g. directional discrimination of visual motion was affected equally by auditory motion no matter whether switching from a visual or from an auditory monitoring task). These results therefore suggest that under conditions of high attentional load, some modulation of visual motion processing as a function of irrelevant auditory motion is possible. This finding appears to conflict with recent studies where variations in the distribution of attention failed to modulate the ventriloquist illusion (Driver, 1996; Spence & Driver, 2000; see Bertelson & de Gelder, in press, for a review). One potential explanation for the discrepancy may lie in the type of attention manipulated. Whereas these recent studies have manipulated the spatial distribution of attention, the present experiment, as well as older studies showing significant attentional modulation on ventriloquism, manipulated non-spatial aspects of attention (i.e. the distribution of attention to one or the other sensory modality or task instead; Spence et al., 2001a).

Another potentially important factor in the study of multisensory integration of motion signals relates to perceptual grouping. As highlighted above, temporal and spatial constraints are important factors for multisensory integration in general, and for motion integration in particular. Related to these factors, parsing inputs into one perceptual group or into segregated streams is another determinant of multisensory integration. Although some researchers have investigated this question for the case of static stimuli (see Vroomen & de Gelder, 2000), the influence of perceptual grouping on the integration of motion signals has received little consideration (though see Watanabe & Shimojo, 2001b). In a series of studies currently under way in this laboratory (Sanabria, Soto-Faraco, Chan, & Spence, submitted for publication), we have observed that the crossmodal dynamic capture of perceived auditory motion by simultaneous visual motion is significantly reduced when the distractor visual motion is embedded within a more spatially extended visual stream of stimuli moving from one side to the other (see also Churchland, Ramachandran, & Sejnowski, 1994). Note that this factor may constitute a potential explanation for Meyer and Wuerger's (2001) failure to observe any perceptual effects in their study. For the particular type of stimuli used by Meyer and Wuerger (visual random dot kinematograms and auditory white noise), may not have provided sufficient cues to encourage perceptual grouping, and consequently, the two motion signals may have been treated as independent streams (see Wuerger, Hofbauer, & Meyer, 2002b; Wuerger et al., 2003, for more recent evidence supporting this account).

2.5. The role of dynamic information

Another potential issue to address when dealing with the behavioural effects of multisensory integration of motion signals is the dissociation between static versus dynamic interactions. Allen and Kolers' (1981) results (discussed above) implied that the effects of visual motion stimuli on an auditory apparent motion task were no different from the effects of static stimuli (see also Staal & Donderi, 1983, for similar results). It therefore becomes an important question to assess the degree to which any crossmodal interactions in the perception of motion results can be accounted for by processes already known to occur (and well-documented) for the case of static stimuli. Some researchers have used paradigms where motion streams do not contain abrupt onsets or offsets at the start or end points, therefore minimizing the presence of any static cues (e.g. Meyer & Wuerger, 2001). Meanwhile, other investigations have addressed the question of static versus dynamic integration processes directly (Soto-Faraco et al., 2002, 2003, in press). For example, Soto-Faraco et al. (2002) used the auditory discrimination of motion direction task with either congruent or conflicting visual distractor motion, and manipulated the SOA between the two elements of each apparent motion stream (the first pairing of sound/light and the second pairing of sound/light). The normal crossmodal congruency effect broke down when the SOA was beyond the threshold where participants typically experienced apparent motion. Because at larger SOAs all the elements of the displays were still present, and the only change was that a motion illusion was no longer perceived, these results suggest that specifically dynamic features played a crucial role in the crossmodal interactions observed.

2.6. Interim conclusions

Taken together, a number of recent psychophysical studies now demonstrate that multisensory contributions to motion processing can reflect both decisional/response-related components as well as genuinely perceptual interactions in the processing of motion signals. These studies highlight the importance of various perceptual grouping principles, such as spatial and temporal co-occurrence, as well as common onset and offset in determining the magnitude of any multisensory interactions taking place. Additionally, research has also revealed some evidence that attention may have a modulatory role in determining the magnitude of any multisensory contributions to motion processing, and that visual motion signals tend to dominate over motion signals available to other sensory modalities. Given this growing body of robust behavioural evidence demonstrating multisensory contributions to motion processing, we now turn to the brain correlates of motion processing in search of possible neural substrates for these effects.

3. Neural correlates of multisensory integration of motion

The brain is rich in both subcortical and cortical areas that demonstrate a sensitivity to various properties of moving stimuli such as directionality or velocity. We now know a great deal about the neural networks involved in visual motion processing, but our knowledge of the networks of brain areas involved in auditory and somatosensory motion processing is currently somewhat more limited. Moreover, the study of how motion information from different modalities interacts in the human brain is still in its infancy. After a brief review of motion processing networks identified for vision, audition and somatosensation, we will discuss recent studies concerned with the brain areas that may subserve multisensory interactions in the processing of motion information.

3.1. Visual motion

Motion selective properties have been described as early in the nervous system as the retina in some species (Barlow & Levick, 1965). With regard to the central nervous system, it is well known that most visual neurons in the cat superior colliculus (SC), a sub-cortical structure, are highly sensitive to motion in a particular direction (e.g. Rauschecker & Harris, 1989; Stein & Arigbede, 1972; Sterling & Wickelgren, 1969). However, the most extensive motion-related regions are to be found in the cortex, where several functionally distinct areas along the so-called dorsal pathway (Ungerleider & Mishkin, 1982) describe a network for the processing of visual motion. The main motion processing center is V5/MT, located in the occipito-temporal cortex, and surrounding areas (the MT+ complex), as revealed by both human (e.g. Watson et al., 1993; Zihl, von Cramon, & Mai, 1983; Zihl, von Cramon, Mai, & Schmid, 1991) and non-human primate studies (e.g. Ungerleider & Desimone, 1986). The V1/V2 complex and V3 are also considered to be part of this network, as they contain populations of neurons sensitive to certain aspects of motion, and relay information to V5/MT (e.g. Orban, Kennedy, & Bullier, 1986). Higher order association areas in the parietal cortex that receive projections from V5/MT as well as from V1/V2 are also involved in the motion processing network. Amongst them, the ventral intra-parietal area (VIP) within the intra parietal sulcus (IPS) is strongly interconnected with V5/MT (e.g. Maunsell & Van Essen, 1983), and has been shown to contain directionally-sensitive neurons in the primate brain (e.g. Colby, Duhamel & Goldberg, 1993; Duhamel, Colby & Goldberg, 1991, 1998). A homologous region in the human cortex often shows activation in functional magnetic resonance imaging (fMRI) studies investigating motion-related brain activity (e.g. Bremmer et al., 2001b). Finally, the prefrontal cortex is also involved in the dorsal pathway and possibly in the motion processing network as well. On the basis of anatomical data, one would expect the ventral premotor cortex (PMv) to be involved in some aspects of visual motion processing, as some areas within this region show strong interconnections with areas VIP and V5/MT (e.g. Luppino, Murata, Govoni & Matelli, 1999).

As we will see below, some of these higher order cortical areas are also implicated in the processing of motion in other sensory modalities, and therefore provide potential candidates sites for the multisensory integration of motion information.

3.2. Auditory and somatosensory motion

The main cues to the spatial location of sounds (in the horizontal plane) are interaural differences in time of arrival (phase) and intensity. Therefore, motion direction can be characterized by changes of these parameters over time. Neurons sensitive to such auditory motion cues have been found at a subcortical level in both the superior (Rauschecker & Harris, 1989) and inferior colliculi (Altman, 1968; Ingham et al., 2001) of lower mammals. At a cortical level, nonhuman primate studies have also revealed auditory motion selectivity in neurons of the primary auditory cortex (Ahissar, Ahissar, Bergman, & Vaadia, 1992). In accord with these animal data, human lesion studies have also highlighted the importance of subcortical structures (Griffiths et al., 1997a) as well as primary auditory cortex (Griffiths et al., 1997b) in auditory motion processing. However, the results of neuropsychological and neuroimaging studies underline the importance of additional areas beyond the primary auditory cortex for auditory motion perception. These include the planum temporale (Baumgart, Gaschler-Markefski, Woldorff, Heinze, & Scheich, 1999; Pavani, Macaluso, Warren, Driver, & Griffiths, 2002; Warren, Zielinski, Green, Rauschecker, & Griffiths, 2002), the right posterior parietal cortex and the right insula (Griffiths et al., 1996, 1998), as well as the superior parietal cortex and the premotor cortex (Pavani, Macaluso, Warren, Driver, & Griffiths, 2002). These studies are now therefore beginning to reveal a number of auditory motion-related areas that may represent part of a cortical network for auditory motion processing.

Regarding somatosensation, various animal studies using neurophysiological techniques have revealed direction selective neurons in primary (Costanzo & Gardner, 1980; Hyvarinen & Poranen, 1978) as well as secondary (Whitsel, Roppolo, & Werner, 1972) somatosensory areas (SI and SII, respectively). Higher-order areas such as Brodmann areas 5 (Sakata, Takaoka, Kawaraski, & Shibutani, 1973) and 7b (Graziano, Yap, & Gross, 1994; Leinonen & Nyman, 1979; Leinonen, Hyvarinen, Nyman, & Linnankoski, 1979; Robinson & Burton, 1980) in the superior parietal cortex have been shown to contain neurons responsive to the direction of motion across the skin. Further along the processing stream, area VIP in the posterior parietal cortex receives input originating from the primary somatosensory cortex as well as from area 7b, and contains neurons sensitive to tactile motion (Duhamel et al., 1998). fMRI studies have also suggested the presence of tactile motion-related activation in the human IPS for tactile stimuli on the face (e.g. Bremmer, Schlack, Duhamel, Graf, & Fink, 2001a). Meanwhile, a recent PET study by Hagen et al. (2002) revealed activation in SI, bilaterally in SII, and in the inferior parietal lobule in response to somatosensory motion along the forearm.

3.3. Multisensory convergence of motion information in animal studies

Several studies have addressed potential regions where the processing of motion information from more than one sensory modality may converge. The results of these studies are now beginning to reveal a number of potential candidate regions that may form part of a multisensory integration network for motion processing, and help to rule out other areas.

At a subcortical level, one potential candidate area is the superior colliculus (SC), because it is a multisensory convergence site containing neurons that show motion sensitivity properties. However, to date, evidence for a major role of the SC in multisensory integration of motion information is surprisingly weak. Indeed, animal studies show that a high proportion of the visual neurons in the SC have directional selectivity (e.g. Stein & Arigbede, 1972; Sterling & Wickelgren, 1969), but despite initial reports of auditory motion selective neurons (Wickelgren, 1971) it seems that, in general, directional selectivity in auditory and somatosensory neurons of the SC is rare (e.g. Clemo & Stein, 1987; Rauschecker & Harris, 1989). Moreover, bimodal neurons that show clear motion selectivity in two modalities have as yet to be reported in this structure.

Area VIP, in the posterior parietal cortex, is another multisensory convergence site that, as shown by the studies cited above, receives projections from visual motion areas (MT+ complex) as well as from somatosensory areas related to motion processing (SI, SII, 5 and 7b). VIP has also been shown to contain neurons responsive to spatial properties of auditory stimuli (Schlack et al., 2000), although its specific role in auditory motion is less clear. Indeed, several studies with primates have found directionally-sensitive neurons for stimuli in different modalities in VIP, indicating its potential role in multisensory integration of motion information (Bremmer et al., 2001b; Colby et al., 1993; Duhamel et al., 1998). For example, Duhamel et al. studied the direction sensitivity of bimodal neurons (visual-somatosensory) in area VIP of the monkey and found that the preferred direction of these neurons often coincided across both sensory modalities. As their receptive fields are mostly circumscribed to the face and surrounding areas in close extrapersonal space, Duhamel et al. hypothesized that one of the roles that VIP may accomplish is to provide a common frame of reference for somatosensory and optical flow information during navigation (see also Bremmer et al., 2001b).

The superior temporal polysensory area (STP), a multisensory converge site located around the superior temporal sulcus, also receives projections from certain motion-related areas. However, it is unlikely that STP plays any specific role in motion integration, as this area does not appear to show motion-related activity, at least for simple stimuli (Bremmer et al., 2001a,b; Bruce, Desimone & Gross, 1981; Desimone & Gross, 1979; though see also Section 3.4).

Finally, another region that receives strong projections from motion-related areas in various modalities is the ventral premotor cortex (PMv). In particular, PMv contains bimodal tactile-visual neurons with similar functional properties to those found in VIP, and seems to have a role in processing objects moving in close extrapersonal space near the face (e.g. Graziano et al., 1994; Graziano, Hu, & Gross, 1997). In recent reports, some portions of the PMv of the monkey (polysensory zone, or PZ; Graziano & Ghandi, 2000; Graziano, Gross, Taylor, & Moore, in press) have been found to contain neurons that encode specific properties of motion in vision and touch (i.e. in most of these neurons, directional selectivity coincides in the two sensory modalities). In contrast with bimodal neurons in VIP, however, PZ neurons often encode space in head-centered coordinates (independent of eye orientation). As PZ is strongly connected with the spinal cord and the primary motor areas, Graziano et al. (in press) have recently hypothesized that this area may be the final stage in a multisensory integration network specialized with defensive behavior to threatening stimuli near (or approaching) to the body.

3.4. Multisensory convergence of motion information in humans

Human neuropsychological lesion studies have reported motion processing deficits circumscribed to a single modality (vision or audition), but these motion-specific deficits have not been found to encompass more than one modality. For example, Zihl et al. (1983) reported a patient with a bilateral lesion in the lateral portions of the temporo-occipital cortex, who failed to perceive visual motion (a syndrome called cerebral akinetopsia) but had spared auditory and somatosensory motion perception abilities (Zihl et al., 1991). In the auditory modality, Griffiths et al. (1997b) reported a patient with damage to the right temporal cortex and the right insula who failed to perceive auditory motion, while retaining normal visual motion perception. To date, we are unaware of any neuropsychological evidence supporting the existence of specific failures to integrate motion information across different sensory modalities.

Neuroimaging studies have, however, provided some interesting insights into potential multisensory integration of motion information in the human brain (Bremmer et al., 2001b; Hagen et al., 2002; Lewis, Beauchamp & DeYoe, 2000). Lewis et al. used fMRI to compare the brain regions that were active during the presentation of visual motion and during the presentation of auditory motion. Sites of activation common to motion in these two modalities were found in the lateral parietal cortex (possibly IPS), lateral frontal cortex (pre-central sulcus), the anterior midline (potentially the anterior cingulate) and the anterior part of the insula. Some of these co-activated regions are consistent with potential multisensory integration sites found in animal studies (i.e. IPS and lateral frontal cortex). However, the method used in Lewis et al.'s study raises some concerns regarding the dissociation between the activation due to motion-specific processes, and activation related to commonalities between the visual and auditory tasks not specifically related to motion. Indeed, Lewis et al. found a very similar pattern of neural activation resulting from a control task regarding auditory pitch discrimination of static sounds, suggesting that most of the activity found when using motion stimuli may not have been specific to motion processing per se.

In another study, Bremmer et al. (2001b) used fMRI to study the pattern of brain activity related to auditory, somatosensory, and visual motion. One of the regions showing activity in response to the three types of motion was the posterior parietal cortex bilaterally, with the local maxima lying in the depth of the IPS, suggesting the potential involvement of the human VIP area. The other region of co-activation common to the three modalities was the ventral premotor cortex (PMv), an area known to contain polymodal neurons that display similarity in function with those found in VIP (e.g. Graziano, Yap, & Gross, 1994; see also Graziano et al., in press). Finally, Bremmer et al.'s study also revealed co-activation in the lateral inferior postcentral cortex. The authors hypothesized that this activation corresponded to SII or surrounding somatosensory association cortex. However, Bremmer et al.'s data needs to be interpreted with care, as the type of motion and the baselines used in each sensory modality were significantly different from each other. Whereas visual motion consisted of random dot displays moving across the fronto-parallel plane and was compared with static dot displays, tactile motion consisted or airflow across the forehead and was compared to the absence of stimulation. The auditory stimulus consisted of binaural beats travelling across the ears and was compared with a no-stimulation baseline (other than the scanner noise). Additionally, there may have been a common element of expectation and attention across all stimulus modalities (see Shulman et al., 2002). Nevertheless, Bremmer et al.'s data do suggest that equivalent areas may be involved in polymodal motion processing in both monkey and human.

In another fMRI study, not specifically related to multisensory integration, Howard et al. (1996) investigated the pattern of brain activation during different types of unimodal visual motion. Among other results, Howard and colleagues reported activation in the superior temporal gyrus (STG), an auditory association area related to the perception of speech. These researchers raised the possibility that this activity could be related to the role that visual information has in speech perception (e.g. Calvert et al., 1997). Alternatively, the authors suggested that this activation in STG could correspond to the human analogue of STP (superior temporal polysensory area), where neurons responsive to visual and auditory stimuli have been identified. However, in the light of the previously discussed animal and human findings, the potential involvement of STP (and/or STG) in multisensory integration of motion cues must remain, at best, speculative. In fact, it would appear that any involvement in motion processing may be specific to the processing of animated visual speech stimuli (see Calvert, Campbell, & Brammer, 2000; Lewis et al., 2000, on this issue).

3.5. A network of multisensory convergence of motion information?

The evidence discussed so far in both animals and humans would appear to conform to the idea that motion information processing proceeds in a fairly feed-forward manner (see, for example, Lewis et al., 2000). That is, motion information progresses from unimodal processing areas (such as V5/MST for vision), to higher order polysensory areas (such as VIP). This strictly feed-forward view of information processing has recently been challenged on several fronts (see Driver & Spence, 2000; Pouget, Deneve & Duhamel, 2002). A number of findings now suggest that multisensory integration may be subserved, at least in part, by back-projections from these multisensory convergence sites to areas traditionally considered to be strictly unimodal (e.g. Calvert et al., 1997; Macaluso, Frith, & Driver, 2000). In support of this view, anatomical evidence reveals that most projections between different brain areas are strongly reciprocal, thus suggesting a strong role of re-entrant mechanisms (e.g. Zeki, 1993). With regard to motion processing, a recent study using positron emission tomography (PET) found activation of hMT (human analog of the monkey's MIT) when participants were presented with just tactile motion consisting of a brush being stroked down the arm (Hagen et al., 2002). At first pass, this finding would appear to reinforce the idea that the neural processes subserving multisensory integration, including the computation of dynamic properties, are highly interactive. Indeed, this idea seems to receive support from tracer studies in the monkey brain showing the existence of projections from IPS back to MT (Blatt, Andersen, & Stoner, 1990; Rosa, Soares, Fiorani, & Gattass, 1993). However, in Hagen et al.'s PET study with humans, one needs also to consider the potential role of mental imagery before any unequivocal interpretation can be put forward (i.e. the possibility that when patients were instructed to attend to the movement of the tactile stimulus down the arm while keeping their eyes closed, they may have used visual imagery, and that this might have been sufficient to activate MT/V5; e.g. see Cohen et al., 1996; Zangaladze, Epstein, Grafton, & Sathian, 1999). In a similar vein, Lewis et al.'s (2000) fMRI study discussed earlier revealed the suppression of activation of hMT in response to auditory motion; these mechanisms of suppression may also be related to attention (Berman & Colby, 2002).

4. Conclusions

Behavioral research into the nature of crossmodal influences on the perception of stimulus motion has a long, albeit intermittent, history. However, it is only in the last few years, thanks in part to the development of a range of new psychophysical paradigms, that a consistent picture has started to emerge. Recent studies have demonstrated that spatial and temporal factors play an important role in motion integration, just as for many other crossmodal effects. There is now convincing behavioral evidence to support the role of dynamic properties in the integration of motion signals, over and above the multisensory interactions that affect the perception of static stimuli. Attention has also been demonstrated to have a modest role in modulating multisensory contributions to motion perception. The pattern of sensory dominance that is starting to emerge in relation to the multisensory contributions to motion perception matches that reported previously for the static ventriloquist illusion; Namely, visual motion cues appear to dominate over auditory and somatosensory motion cues. However, further research is needed to confirm the source of these hierarchical effects in the behavioral data (i.e. are they stimulus-related, or do they reflect an underlying bias to attend preferentially to a given sensory modality). Ongoing research in this laboratory has also started to highlight perceptual grouping as an important (if somewhat neglected) factor determining the nature of any multisensory contributions to motion perception.

Several findings now suggest that the multisensory integration of motion signals across different sensory modalities can occur early in perceptual processing. In addition, there is also unequivocal support for the view that post-perceptual (i.e. decisional and response-related) factors may have played an important role in determining performance in many previous behavioral studies. However, the relative contributions of these various levels of processing is still a matter of some debate, partly because of the lack of formal definitions about perceptual and post-perceptual factors. One approach to operationalizing these terms comes from signal detection theory (e.g. Green & Swets, 1966). The presentation of a motion stimulus in one modality may affect responses to motion presented in the other modality only by introducing a *bias* (β) in the criterion (c) set to decide whether certain motion property is present in that modality. This could be contrasted with other situations in which the presentation of motion stimuli in one modality leads to a shift in *sensitivity* (d') for the perception of whether that certain motion property is present in another sensory modality. One might consider changes in sensitivity to reflect genuine perceptual effects, whereas the former biases might be more likely to reflect decisional and response bias factors instead. Implementing these concepts to the problem of levels of processing in multisensory integration of motion may provide useful insights in the future. Another promising way to approach this question is to turn to neuroimaging studies, where the controversy about perceptual versus post-perceptual factors can perhaps be re-framed in terms of the specific brain areas involved and the temporal succession of their activation during the processing of multisensory motion cues.

Indeed, recent neuroscience research is beginning to help researchers to elucidate the networks of cortical structures responsible for the processing of motion in the different sensory modalities. In particular, neuroimaging studies in humans and neurophysiological studies in animals have highlighted a number of brain regions that are activated in response to the movement of a stimulus in a specific modality versus those areas that are co-activated by motion information in two or more sensory modalities. These convergence sites provide plausible candidates as the neural substrates of the behavioural multisensory interactions in motion processing that have been reported, either directly, or via re-entrant processing. Indeed, given the purportedly unimodal phenomenology of motion perception, together with the clear evidence for multisensory contributions to this type of perceptual experience, one might speculate that interactions take place in higher-level areas with multisensory neurons, while unimodal motion areas are responsible for the unimodal phenomenology of motion perception (cf. Driver & Spence, 2000).

Nevertheless, the present review has identified an important gap between behavioural, neurophysiological and neuroimaging studies in terms of the types of stimulation used and the specific aspects of motion processing addressed (though see Bushara et al., 2003, for a recent exception). Clearly, future neuroimaging studies that contrast the effects of presenting motion in one versus two modalities, and also consider the effects of presenting motion in congruent versus incongruent directions in different sensory modalities, will be particularly informative in this regard. Conversely, it might also be fruitful for behavioural studies to began to address multisensory interactions in motion processing in close extrapersonal space, a situation that has shown to give rise to clear physiological correlates associated with the obvious behavioural relevance of such stimuli.

Acknowledgements

The authors would like to thank E. Macaluso, D. Lloyd, and A. Rodriguez-Fornells for helpful comments on an earlier draft of this manuscript, and also the University of Oxford McDonnell-Pew Centre for Cognitive Neuroscience for funding a Network Grant to S.S.-F. and C.S. Funding to AK was provided by Human Frontier Science Program, the Natural Sciences and Engineering Research Council of Canada, and the Michael Smith Foundation for Health Research. Correspondence regarding this article should be addressed to S.S.-F. (email: ssoto@psico.psi.ub.es) or to C.S. (email: charles.spence@psy.ox.ac.uk).

References

Ahissar, M., Ahissar, E., Bergman, H., & Vaadia, E. (1992). Encoding of sound-source location and movement: Activity of single neurons and interactions between adjacent neurons in the monkey auditory cortex. *Journal of Neurophysiology*, *67*, 203–215.

Alais, D., & Blake, R. (1999). Neural strength of visual attention gauged by motion adaptation. *Nature Neuroscience*, *2*, 1015–1018.

Allen, P. G., & Kolers, P. A. (1981). Sensory specificity of apparent motion. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 1318–1326.

Altman, J. A. (1968). Are there neurons detecting direction of sound source motion? *Experimental Neurology*, *22*, 13–25.

Anstis, S. M. (1973). Hearing with the hands. *Perception*, *2*, 337–341.

Arnoult, M. D. (1952). Localization of sound during rotation of the visual environment. *American Journal of Psychology*, *65*, 48–58.

Barlow, H. B., & Levick, W. R. (1965). Mechanism of directionally selective units in rabbit's retina. *Journal of Physiology*, *178*, 477–504.

Baumgart, F., Gaschler-Markefski, B., Woldorff, M. G., Heinze, H. J., & Scheich, H. (1999). A movement-sensitive area in auditory cortex. *Nature*, *400*, 724–726.

Berman, R. A., & Colby, C. L. (2002). Auditory and visual attention modulate motion processing in area MT+. *Cognitive Brain Research*, *14*, 64–74.

Bertelson, P. (1998). Starting from the ventriloquist: The perception of multimodal events (vol. 2). In M. Sabourin, C. Fergus, et al. (Eds.), *Advances in psychological science*: *Biological and cognitive aspects* (pp. 419–439). Hove, England: Psychology Press.

Bertelson, P., & Aschersleben, G. (1998). Automatic visual bias of perceived auditory location. *Psychonomic Bulletin* & *Review*, *5*, 482–489.

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

Bertelson, P., Vroomen, J., de Gelder, B., & Driver, J. (2000). The ventriloquist effect does not depend on the direction of deliberate visual attention. *Perception* & *Psychophysics*, *62*, 321–332.

Blatt, G. J., Andersen, R. A., & Stoner, G. R. (1990). Visual receptive field organization and cortico-cortical connections of the lateral intraparietal area (area LIP) in the macaque. *Journal of Computational Neurology*, *299*, 421–445.

Bremmer, F., Schlack, A., Duhamel, J. R., Graf, W., & Fink, G. R. (2001a). Space coding in primate parietal cortex. *Neuroimage*, *14*, S46–S51.

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

Bruce, C., Desimone, R., & Gross, C. G. (1981). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *Journal of Neurophysiology*, *46*, 369–384.

Caclin, A., Soto-Faraco, S., Kingstone, A., & Spence, C. (2002). Tactile 'capture' of audition. *Perception* & *Psychophysics*, *64*, 616–630.

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

Calvert, G. A., Campbell, R., & Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Current Biology*, *10*, 649–657.

Canon, L. K. (1970). Intermodality inconsistency of input and directed attention as determinants of the nature of adaptation. *Journal of Experimental Psychology*, *84*, 141–147.

Canon, L. K. (1971). Directed attention and maladaptive 'adaptation' to displacement of the visual field. *Journal of Experimental Psychology*, *88*, 403–408.

Cavanagh, P. (1992). Attention-based motion perception. *Science*, *257*, 1563–1565.

Chaudhuri, A. (1990). Modulation of the motion aftereffect by selective attention. *Nature*, *344*, 60–62.

Choe, C. S., Welch, R. B., Gilford, R. M., & Juola, J. F. (1975). The 'ventriloquist effect': Visual dominance or response bias? *Perception* & *Psychophysics*, *18*, 55–60.

Churchland, P. S., Ramachandran, V. S., Sejnowski, T. J. (1994). A critique of pure vision. In C. Koch & J. L. Davis (Eds.), *Large-scale neuronal theories of the brain* (pp. 23–60). Cambridge, MA: MIT Press.

Clemo, H. R., & Stein, B. E. (1987). Responses to direction of stimulus movement are different for somatosensory and visual cells in cat superior colliculus. *Brain Research*, *405*, 313–319.

Cohen, M. S., Kosslyn, S. M., Breiter, H. C., DiGirolamo, G. J., Thompson, W. L., Anderson, A. K., Brookheimer, S. Y., Rosen, B. R., & Belliveau, J. W. (1996). Changes in cortical activity during mental rotation. A mapping study using functional MRI. *Brain*, *119*, 89–100.

Colby, C. L., Duhamel, J. R., & Goldberg, M. E. (1993). Ventral intra-parietal area of the macaque: Anatomical location and visual response properties. *Journal of Neurophysiology*, *69*, 902–914.

Costanzo, R. M., & Gardner, E. P. (1980). A quantitative analysis of responses of direction-sensitive neurons in somatosensory cortex of awake monkeys. *Journal of Neurophysiology*, *43*, 1319–1341.

Desimone, R., & Gross, C. G. (1979). Visual areas in the temporal cortex of the macaque. *Brain Research*, *178*, 363–380.

Dichgans, J., & Brandt, T. (1978). Visual-vestibular interaction (vol. VIII). In R. Held, H. W. Leibowitz, & H. L. Teuber (Eds.), *Handbook of sensory physiology*: *Perception* (pp. 755–804). Berlin: Springer.

Downing, P. E., & Treisman, A. M. (1997). The line-motion illusion: Attention or impletion? *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 768–779.

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

Driver, J., & Spence, C. (2000). Multisensory perception: Beyond modularity and convergence. *Current Biology*, *10*, R731–R735.

Duhamel, J. R., Colby, C. L., & Goldberg, M. E. (1991). Congruent representations of visual and somatosensory space in single neurons of monkey ventral intra-parietal cortex (area VIP). In J. Palliard (Ed.), *Brain and space* (pp. 223–236). Oxford: Oxford University Press.

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.

Ehrenstein, W. H., & Reinhardt-Rutland, A. H. (1996). A cross-modal aftereffect: Auditory displacement following adaptation to visual motion. *Perceptual and Motor Skills*, *82*, 23–26.

Gemelli, A. (1951). The effect of illusory perception of movement on sound localization. In G. Ekman, G. Johansson, & C. I. Sandström (Eds.), *Essays in psychology dedicated to David Katz* (pp. 104–116). Uppsala: Almqvist and Wiksells Boktryckeri AB.

Gilbert, G. M. (1939). Dynamic psychophysics and the phi phenomenon. *Archives of Psychology*, *237*, 5–43.

Graziano, M. S. A., & Ghandi, S. (2000). Location of the polysensory zone in the precentral gyrus of anesthetized monkeys. *Experimental Brain Research*, *135*, 259–266.

Graziano, M. S. A., Gross, C. G., Taylor, C. S. R., & Moore, T. (in press). A system of multimodal areas in the primate brain. In Spence, C., & Driver, J. (Eds.), *Crossmodal space and crossmodal attention*. Oxford: Oxford University Press.

Graziano, M. S. A., Hu, X., & Gross, C. G. (1997). Visuo-spatial properties of ventral premotor cortex. *Journal of Neurophysiology*, *77*, 2268–2292.

Graziano, M. S. A., Yap, G. S., & Gross, C. G. (1994). Coding of visual space by premotor neurons. *Science*, *266*, 1054–1057.

Green, D. M. & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Griffiths, T. D., Bates, D., Rees, A., Witton, C., Gholkar, A., & Green, G. G. R. (1997a). Sound movement detection deficit due to a brainstem lesion. *Journal of Neurology, Neurosurgery and Psychiatry*, *62*, 522–526.

Griffiths, T. D., Rees, A., Witton, C., Cross, P. M., Shakir, R. A., & Green, G. G. R. (1997b). Spatial and temporal auditory processing deficits following right hemisphere infarction. *Brain*, *120*, 785–794.

Griffiths, T. D., Rees, A., Witton, C., Shakir, R. A., Henning, G. B., & Green, G. G. R. (1996). Evidence for a sound movement area in the human cerebral cortex. *Nature*, *383*, 425–427.

Griffiths, T. D., Rees, G., Rees, A., Green, G. G. R., Witton, C., Rowe, D., Buchel, C., Turner, R., & Frackowiak, R. S. J. (1998). Right parietal cortex is involved in the perception of sound movement in humans. *Nature Neuroscience*, *1*, 74–79.

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.

Hall, K. R. L., & Earle, A. E. (1954). A further study of the pendulum phenomenon. *Quarterly Journal of Experimental Psychology*, *6*, 112–124.

Hall, K. R. L., Earle, A. E., & Crookes, T. G. (1952). A pendulum phenomenon in the visual perception of apparent motion. *Quarterly Journal of Experimental Psychology*, *4*, 109–120.

Hikosaka, O., Miyauchi, S., Takeichi, H., & Shimojo, S. (1996). Multimodal spatial attention visualized by motion illusion. In T. Inui, & J. L. McClelland (Eds.), *Attention and performance XVI*: *Information integration in perception and communication* (pp. 237–261). Cambridge, MA: MIT Press.

Hirsh, I. J., & Sherrick Jr., C. E. (1961). Perceived order in different sense modalities. *Journal of Experimental Psychology*, *62*, 423–432.

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.

Hubbard, T. L. (1995). Environmental invariants in the representation of motion: Implied dynamics and representation of motion: Implied dynamics and representational momentum, gravity, friction, and centripetal force. *Psychonomic Bulletin* & *Review*, *2*, 322–338.

Hyvarinen, J., & Poranen, A. (1978). Movement-sensitive and direction and orientation-selective cutaneous receptive fields in the hand area of the post-central gyrus in monkeys. *Journal of Physiology*, *283*, 523–537.

Ingham, N. J., Hart, H. C., & McAlpine, D. (2001). Spatial receptive fields of inferior colliculus neurons to auditory apparent motion in free field. *Journal of Neurophysiology*, *85*, 23–33.

Jackson, C. V. (1953). Visual factors in auditory localization. *Quarterly Journal of Experimental Psychology*, *5*, 52–65.

Jeka, J., Oie, K. S., & Kiemel, T. (2000). Multisensory information for human postural control: Integrating touch and vision. *Experimental Brain Research*, *134*, 107–125.

King, A. J., & Palmer, A. R. (1985). Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Experimental Brain Research*, *60*, 492–500.

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

Kitajima, N., & Yamashita, Y. (1999). Dynamic capture of sound motion by light stimuli moving in three-dimensional space. *Perceptual and Motor Skills*, *89*, 1139–1158.

Lakatos, S. (1995). The influence of visual cues on the localisation of circular auditory motion. *Perception*, *24*, 457–465.

Leinonen, L., & Nyman, G. (1979). II. Functional properties of cells in anterolateral part of area 7 associative face area of awake monkeys. *Experimental Brain Research*, *34*, 321–333.

Leinonen, L., Hyvarinen, J., Nyman, G., & Linnankoski, I. (1979). I. Functional properties of neurons in the lateral part of associative area 7 in awake monkeys. *Experimental Brain Research*, *34*, 299–320.

Lewis, J. W., Beauchamp, M. S., & DeYoe, E. A. (2000). A comparison of visual and auditory motion processing in human cerebral cortex. *Cerebral Cortex*, *10*, 873–888.

Luppino, G., Murata, A., Govoni, P., & Matelli, M. (1999). Largely segregated parietofrontal connections linking rostral intraparietal cortex (areas AIP and VIP) and the ventral premotor cortex (areas F5 and F4). *Experimental Brain Research*, *128*, 181–187.

Maass, H. (1938). Über den Einfluss akustischer Rhythmen auf optische Bewegungsgestaltungen [About the influence of acoustic rhythms on visual motion]. (Sander, F. Ganzheit und Gestalt. Psychol. Untersuch. VIII). *Archiv für die Gesamte Psychologie*, *100*, 424–464.

Macaluso, E., Frith, C., & Driver, J. (2000). Modulation of human visual cortex by crossmodal spatial attention. *Science*, *289*, 1206–1208.

Maioli, C., & Poppele, R. E. (1991). Parallel processing of multisensory information concerning self-motion. *Experimental Brain Research*, *87*, 119–125.

Manabe, K., & Riquimaroux, H. (2000). Sound controls velocity perception of visual apparent motion. *Journal of the Acoustical Society of Japan*, *21*, 171–174.

Mateeff, S., Hohnsbein, J., & Noack, T. (1985). Dynamic visual capture: Apparent auditory motion induced by a moving visual target. *Perception*, *14*, 721–727.

Maunsell, J. H. R., & Van Essen, D. C. (1983). The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *Journal of Neuroscience*, *3*, 2563–2580.

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

Morein-Zamir, S., Soto-Faraco, S., & Kingstone, A. (2003). Auditory capture of vision: Examining temporal ventriloquism. *Cognitive Brain Research*, *17*, 154–163.

Neff, W. S. (1936). A critical investigation of the visual apprehension of movement. *American Journal of Psychology*, *48*, 1–42.

Ohmura, H. (1987). Intersensory influences on the perception of apparent movement. *Japanese Psychological Research*, *29*, 1–9.

Orban, G. A., Kennedy, H., & Bullier, J. (1986). Velocity sensitivity and direction selectivity of neurons in areas V1 and V2 of the monkey: Influence of eccentricity. *Journal of Neurophysiology*, *56*, 462–480.

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.

Posner, M. I., Nissen, M. J., & Klein, R. M. (1976). Visual dominance: An information-processing account of its origins and significance. *Psychological Review*, *83*, 157–171.

Pouget, A., Deneve, S., & Duhamel, J. R. (2002). A computational perspective on the neural basis of multisensory spatial representations. *Nature Reviews Neuroscience*, *3*, 741–747.

Rauschecker, J. P., & Harris, L. R. (1989). Auditory and visual neurons in the cat's superior colliculus selective for the direction of apparent motion stimuli. *Brain Research*, *490*, 56–63.

Raymond, J. E. (2000). Attentional modulation of visual motion perception. *Trends in Cognitive Sciences*, *4*, 42–50.

Rees, G., & Lavie, N. (2001). What can functional imaging reveal about the role of attention in visual awareness? *Neuropsychologia*, *39*, 1343–1353.

Repp, B. H., & Penel, A. (2002). Auditory dominance in temporal processing: New evidence from synchronization with simultaneous visual and auditory sequences. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 1085–1099.

Robinson, C. J., & Burton, H. (1980). Somatic submodality distribution within the second somatosensory (SII), 7b, retroinsular, post-auditory, and granular insular cortical areas of M. fascicularis. *Journal of Computational Neurology*, *192*, 93–108.

Rock, I., & Harris, C. S. (1967). Vision and touch. *Scientific American*, *216*, 96–104.

Rosa, M. G., Soares, J. G., Fiorani, M., & Gattass, R. (1993). Cortical afferents of visual area MT in the Cebus monkey: Possible homologies between New and Old World monkeys. *Visual Neuroscience*, *10*, 827–855.

Ryan, T. A. (1940). Interrelations of the sensory systems in perception. *Psychological Bulletin*, *37*, 659–698.

Sakata, H., Takaoka, Y., Kawaraski, A., & Shibutani, H. (1973). Somatosensory properties of neurons in the superior parietal cortex (area 5) of the Rhesus monkey. *Brain Research*, *64*, 3877–3881.

Sanabria, D., Soto-Faraco, S., Chan, J., & Spence, C. (2003). *Intramodal perceptual grouping modulates multisensory integration: Evidence from the crossmodal dynamic capture task*. Submitted for publication.

Schlack, A., Sterbing, S., Hartung, H., Hoffmann, K. P. & Bremmer, F. (2000). Auditory responsiveness in the macaque ventral intraparietal area (VIP). *Society for Neuroscience Abstracts*, 26

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

Shams, L., Allman, J., & Shimojo, S. (2001). Illusory visual motion induced by sound. *Society for Neuroscience Abstracts*, 511.10.

Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature*, *408*, 788.

Shimojo, S., Miyauchi, S., & Hikosaka, O. (1997). Visual motion sensation yielded by non-visually driven attention. *Vision Research*, *12*, 1575–1580.

Shulman, G. L., Tansy, A. P., Kincade, M., Petersen, S. E., McAvoy, M. P., & Corbetta, M. (2002). Reactivation of networks involved in preparatory states. *Cerebral Cortex*, *12*, 590–600.

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. (2000). The role of movement and attention in modulating audiovisual and audiotactile 'ventriloquism' effects. *Abstracts of the Psychonomic Society*, *5*, 40.

Soto-Faraco, S., Spence, C., & Kingstone, A., (in press). Cross-modal dynamic capture. *Journal of Experimental Psychology: Human Perception and Performance*.

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 (2003). *Automatic visual capture of auditory motion*. Submitted for publication.

Spelke, E. S., Born, W. S., & Chu, F. (1983). Perception of moving, sounding objects by four-month-old infants. *Perception*, *12*, 719–732.

Spence, C., Chan, J. & Simm, A. (2000). *Evaluating the 'Perceptual Load' hypothesis in audition and crossmodally*. Poster presented at the 7th Annual Meeting of the Cognitive Neuroscience Society, San Francisco, April 10th.

Spence, C., & Driver, J. (2000). Attracting attention to the illusory location of a sound: Reflexive crossmodal orienting and ventriloquism. *Neuroreport*, *11*, 2057–2061.

Spence, C., Nicholls, M. E. R., & Driver, J. (2001a). The cost of expecting events in the wrong sensory modality. *Perception* & *Psychophysics*, *63*, 330–336.

Spence, C., Shore, D. I., & Klein, R. M. (2001b). Multisensory prior entry. *Journal of Experimental Psychology: General*, *130*, 799–832.

Staal, H. E., & Donderi, D. C. (1983). The effect of sound on visual apparent movement. *American Journal of Psychology*, *96*, 95–105.

Stein, B. E., & Arigbede, M. O. (1972). Unimodal and multimodal response properties of neurons in the cat's superior colliculus. *Experimental Neurology*, *36*, 179–196.

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

Sterling, P., & Wickelgren, B. G. (1969). Visual receptive fields in the superior colliculus of the cat. *Journal of Neurophysiology*, *32*, 1–15.

Thurlow, W. R., & Kerr, T. P. (1970). Effects of a moving visual environment on localization of sound. *American Journal of Psychology*, *83*, 112–118.

Ungerleider, L. G., & Desimone, R. (1986). Cortical connections of visual area MT in the macaque. *Journal of Computational Neurology*, *248*, 190–222.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of visual behaviour* (pp. 549–586). Cambridge, MA: MIT Press.

Urbantschitsch, V. (1902). Ueber die Beeinflussung subjektiver Gesichtsempfindungen [On the manipulation of subjective sensations at the face]. *Pflügers Archiv für die geschichte Physiologie*, *94*, 347–448.

Vroomen, J., Bertelson, P., & de Gelder, B. (2001). The ventriloquist effect does not depend on the direction of automatic visual attention. *Perception* & *Psychophysics*, *63*, 651–659.

Vroomen, J., & de Gelder, B. (in press). Visual motion influences the contingent auditory motion aftereffect. *Psychological Science*.

Vroomen, J., & de Gelder, B. (2000). Sound enhances visual perception: Cross-modal effects of auditory organization on vision. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 1583–1590.

Warren, D. H., & Schmitt, T. L. (1978). On the plasticity of visual-proprioceptive bias effect. *Journal of Experimental Psychology: Human Perception and Performance*, *4*, 302–310.

Warren, J. D., Zielinski, B. A., Green, G. G., Rauschecker, J. P., & Griffiths, T. D. (2002). Perception of sound-source motion by the human brain. *Neuron*, *34*, 139–148.

Watanabe, K., & Shimojo, S. (2001a). Postcoincidence trajectory duration affects motion event perception. *Perception* & *Psychophysics*, *63*, 16–28.

Watanabe, K., & Shimojo, S. (2001b). When sound affects vision: Effects of auditory grouping on visual motion perception. *Psychological Science*, *12*, 109–116.

Watson, J. D., Myers, R., Frackowiak, R. S., Hajnal, J. V., Woods, R. P., Mazziotta, J. C., Shipp, S., & Zeki, S. (1993). Area V5 of the human brain: evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cerebral Cortex*, *3*, 79–94.

Welch, R. B. (1999). Meaning, attention, and the "unity assuption" in the intersensory bias of spatial and temporal perceptions. In G. Aschersleben, T. Bachmann, & J. Müsseler (Eds.), *Cognitive contributions to the perception of spatial and temporal events* (pp. 371–387). Amsterdam: Elsevier.

Welch, R. B., DuttonHurt, L. D., & Warren, D. H. (1986). Contributions of audition and vision to temporal rate perception. *Perception* & *Psychophysics*, *39*, 294–300.

Whitsel, B. L., Roppolo, J. R., & Werner, G. (1972). Cortical information processing of stimulus motion on primate skin. *Journal of Neurophysiology*, *35*, 691–717.

Wickelgren, B. G. (1971). Superior colliculus: Some receptive field properties of bimodally responsive cells. *Science*, *173*, 69–71.

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

Wohlschläger, A., & Wohlschläger, A. (1998). Mental and manual rotation. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 397–412.

Wuerger, S. M., Hofbauer, M., & Meyer, G. F. (2002a). *The integration of auditory and visual motion signals*. Submitted for publication.

Wuerger, S. M., Hofbauer, M., & Meyer, G. F. (2002b). The integration of auditory and visual motion signals. *Journal of Vision*, *2*, 663a.

Wuerger, S. M., Roehrbein, Meyer, G. F., Hofbauer, M., Schill, K., & Zetzche, C. (2003). Auditory and visual motion signals have to be co-localized to be effectively integrated. Paper presented at the Vision Science Society Meeting, Sarasota.

Zangaladze, A., Epstein, C. M., Grafton, S. T., & Sathian, K. (1999). Involvement of visual cortex in tactile discrimination of orientation. *Nature*, *401*, 587–590.

Zapparoli, G. C., & Reatto, L. L. (1969). The apparent movement between visual and acoustic stimulus and the problem of intermodal relations. *Acta Psychologica*, *29*, 256–267.

Zeki, S. (1993). *A vision of the brain*. Oxford: Blackwell Scientific.

Zietz, K., & Werner, H. (1927). Über die dynamische Struktur der Bewegung. Werner's Studien über Strukturgesetze, VIII. [On the dynamic structure of movement. Werner's studies on the laws of structure, VIII]. *Zeitschrift für Psychologie*, *105*, 226–249.

Zihl, J., von Cramon, D., & Mai, N. (1983). Selective disturbance of movement vision after bilateral brain damage. *Brain*, *106*, 313–340.

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.