

Acknowledgements.

This work was supported by grants from the BBSRC and MRC, UK, to the second author.

MULTISENSORY INTEGRATION

Caterina Cinel¹ and Glyn Humphreys²

¹Department of Psychology,
University of Essex, Wivenhoe, Colchester, C04 3SQ, UK

²Behavioural Brain Sciences Centre,
School of Psychology, University of Birmingham,
Birmingham, B15 2TT, UK

ABSTRACT

A fast, reliable and unambiguous perception of the environment requires the integration of all the information simultaneously acquired and processed by different perceptual systems. However, research has traditionally focused mostly on the processing of unimodal stimuli, which has effectively left unanswered the questions of how exactly information processed by different perceptual systems is integrated to give unified and coherent perceptual experiences. Only in recent years has the interest on multimodal perception increased and, with it, the number of experimental studies that investigate this issue.

We now know that the integration of multimodal information is something more than the simple adding together of information from the different senses. It has been shown, for example, that the perceived position from which a sound originates can be affected by the visual perception of an object, simultaneous presented (the ventriloquist effect). We also know that the perception of facial emotions can be affected by the hearing of vocal expressions of emotions, or that people can report felt textures as being seen or visual textures as being touched (when cross-modal illusory conjunctions are perceived).

In this chapter we will review evidence on cross-modal sensory integration, focusing in particular on spatial integration, spatial attention, object identification and feature binding.

1. INTRODUCTION

Our perception of the world is based on the integration of information gathered from different perceptual modalities: the objects we perceive can have shapes, colours, movements, but also sounds, odours and textures sensed through active touch. Despite this, research on perception and cognitive psychology has traditionally focused on unimodal processing, and it has only been in recent times that research on multimodal perception has increased to a significant level. Scientists are now forming a new view of human cognition where the senses interact and influence each other in extraordinary and sometimes unexpected ways. Multimodal perception is not simply the bringing together and summing of pieces of information acquired through different sensory channels, rather it is a much richer process in which input from one sensory modality can be modulated by contrasting information from a different modality. As a result, what we experience may be different from what our sensory systems have initially encoded.

2. VISUAL-AUDITORY SPATIAL INTERACTIONS

One of the most studied aspects of multimodal integration is spatial processing. This is often examined under conditions in which what is sensed by one modality is in conflict with what is sensed by a different modality. A typical example of spatial conflict between modalities is represented by the *ventriloquist effect* (Bermant and Welch, 1976; Bertelson, 1994; 1998; 1999).

In the ventriloquist effect the location of an auditory stimulus can be misjudged because of the presence of a visual stimulus at a different location. In a typical experimental setting, a flashing light and a brief sound are simultaneously presented in two different locations. Typically, the location of the sound source is reported closer to the light location than it actually was. The bias can also occur in the opposite direction – with the location of an auditory sound affecting visual localisation (though usually to a lesser degree; Bertelson and Radeau, 1981; Radeau and Bertelson, 1987). Furthermore, when, in a similar setting, subjects are asked to judge whether the visual and auditory stimuli have occurred in the same or at different locations, they often report that the stimuli occurred in a common location even when the locations differed (in which case *perceptual fusion* is said to occur; Bertelson and Radeau, 1981). The ventriloquist effect suggests that our brain “expects” two simultaneous events to be originating from the same source or object and, therefore, when they are not, there is an attempt to consider them not only temporally but also spatially “co-occurrent”.

Research on the ventriloquist effect has also focused on understanding the level of processing at which the effect originates. Is it at a strictly sensory level? Or does it originate at a later cognitive stage? To understand this, ventriloquism has been studied with many different kinds of stimuli and often under more realistic conditions with “everyday life” objects, for example, using speaking faces or hands playing an instrument as the visual objects and the corresponding spoken words or notes as auditory stimuli (Radeau and Bertelson, 1977). Settings like these suggest that ventriloquism might be caused by one’s expectations. However, the fact that ventriloquism can easily occur with unfamiliar, meaningless stimuli, or stimuli that are unrelated in the two modalities, shows that

interactions are not necessarily mediated by top-down biases and/or by attempts to establish coherence across sensory inputs (Bertelson and Aschersleben, 1998) (though these factors may contribute to performance too).

Recent experiments have also analysed the possible involvement of attentional orienting processes in the ventriloquist effect (Bertelson, Vroomen, de Gelder, and Driver, 2000; Vroomen, Bertelson, and de Gelder, 2001). Studies directly addressing cross-modal links in spatial attention have demonstrated that auditory attention can be shifted toward the location of a visual cue, and vice-versa (e.g. Spence and Driver, 1997; 1998a; see below). A shifting of attention to the visual stimulus could influence (if not cause) the ventriloquism effect, if participants tend to judge the location of a stimulus as being where they attend. To investigate this possibility, Bertelson and colleagues (Bertelson, Vroomen, de Gelder and Driver, 2000) proposed a variant of the ventriloquist paradigm. Observers had to localise a sound (which could be at the left, right or centre of a visual display) that was simultaneously presented with a visual ‘attractor’ stimulus, either on the left or on the right, which was a potential inducer of the ventriloquism effect. At the same time, observers had to detect the presence of a visual target presented either at the same place as the attractor or at the centre of the display, and, therefore, they had to focus attention on that location. It was found that the magnitude of the ventriloquist effect was not modulated by whether or not attention was focused (for target detection) on the same position as the visual attractor. This was taken as an evidence that the visual bias on audition is not caused by a shift of attention from the auditory to the visual stimulus, and it was claimed that “the ventriloquist phenomena reflect automatic interactions between sensory codes for location, arising at levels of representation that are pre-attentive, in the sense that they are uninfluenced by the current direction of deliberate attention of space” (Bertelson, et al., 2000).

To further test a possible involvement of attention, the ventriloquist effect has also been examined in a group of patients with unilateral visual neglect consequent to right hemisphere damage (Bertelson, Pavani, Làdavas, Vroomen and de Gelder, 2000). Patients with unilateral neglect tend to omit stimuli that are presented on the contralesional space (typically on the left, following right hemisphere damage in the patient). Neglect can be considered to be an attentional deficit, reflecting an abnormal competitive advantage in processing for the ipsilesional side (Heinke and Humphreys, 2003). Consistent with this, there is evidence that neglected stimuli are perceived and processed up to cognitive levels, but the patients seem incapable of attending to them (e.g., McGlinchey-Berroth, Milberg, Verfaellie, Alexander and Kilduff, 1993). If the ventriloquist effect was caused by attention being attracted by the visual distractor, then in patients with neglect there should be little or no ventriloquist effect when the visual distractor is presented on the neglected side. In contrast to this, neglect patients can show a ventriloquist effect when the face is on the left and the sound on the right, and indeed this shift can occur even when patients are not consciously aware of the contralesional, visual stimulus (Bertelson, 1999; Bertelson and Ashersleben, 1998; Bertelson, Vroomen, de Gelder and Driver, 2000; Vroomen, Bertelson and de Gelder, 2001). Such data suggest that attention does not play a direct role in the integration process. However other work, particularly on cross-modal illusory conjunctions (Cinèl, Humphreys and Poli, 2002), contravenes that view, as discuss below.

3. SPATIAL INTERACTIONS BETWEEN TOUCH AND OTHER MODALITIES

Cross-modal interactions do not only take place between vision and audition, but also between other sensory modalities. For instance, similarly to the ventriloquist effect, sounds can be mis-localised towards the location of a felt stimulus (Caclin, Soto-Faraco, Kingstone and Spence, 2002), and again the effect can arise even when the felt stimuli are not attended (Cinel, Humphreys and Poli, 2002).

Visual-tactile interactions during the processing of spatial information have also been described. Pavani, Spence and Driver (2000; Spence, Pavani, and Driver, 1998) investigated how people rely both on vision and touch when they have to localise a tactile stimulus. In their experiments, they asked the subjects to judge whether a touch on their hand was on the index finger - “above” - or on the thumb - “below”. They found that the speed of response was influenced by a simultaneous visual distractor presented close to the stimulated hand, which could be either consistent or inconsistent with the response (Spence, Pavani, and Driver, 1998). The interference, however, was reduced when the visual distractor appeared further away from the hand. This suggests that there is a strong bias to integrate visual-tactile information when the stimuli are close to each other. Interference then occurs when that information is incongruent. These results are supported by neurophysiological studies (Graziano, Gregory, and Gross, 1994; Graziano and Gross, 1996) in which bimodal neurons have been detected in the premotor cortex and in the putamen of the macaque; these neurons have receptive fields (RFs) on the hand and corresponding visual RFs in the space immediately adjacent the tactile fields. The RFs of the visual neurons move (remaining centered on the hand) when the arm moves, but not when the gaze moves. This means that those bimodal cells are maximally sensitive to stimuli that are seen close to the hand. When there is spatial incongruence between what is seen and what is touched, the perceptual system is forced to disambiguate the information. This results in delayed response times. In contrast, visual stimuli presented at some distance from the hand do not fall inside the RFs of the bimodal cells, facilitating modality-specific selectivity. Interference from incongruent information in an irrelevant modality does not occur in this case.

A bizarre effect of visual- tactile spatial integration happens when artificial hands are used. Pavani and colleagues (Pavani et al., 2000) carried out experiments similar to the one described above, the only difference being that the subjects looked at two artificial hands exactly superimposed on the locations where their real hands would fall on top of a table, though their real hands were placed in equivalent positions under the table. A visual distractor was then presented on an artificial hand, but at some distance from the participant’s own hands, which simultaneously received tactile stimulation. The presence of the visual stimulus near the artificial hands produced visual-tactile interference, and subjects also reported that they felt the touch on the artificial hands (for a similar effect see also Botvinik and Cohen, 1998). However, the effect of the artificial hands disappeared when these hands were no longer aligned with the real hands. This suggests that proprioceptive perception (about the positions of the hands) is integrated with both visual and tactile information derived from the environment. When the seen position of a artificial hand is incompatible with what is perceived by the proprioceptive system, the hand is not “considered” as part of the body any more and, consequently, visual stimuli presented in the near vicinity do not interfere any more

with tactile spatial perception. The view of artificial hands (aligned with the real hands) can even improve touch sensation in patients where tactile detection is impaired (Rorden, Heutink, Greenfield, and Robertson, 1999). From this we may conclude that peripersonal space is coded in a multimodal manner, with inputs from vision, touch and proprioception. This multimodal representation might have evolved in order to optimise behaviours such as reaching and grasping objects close to the body, with multiple inputs from different modalities being used to compensate for any noise that might exist in one modality. The action of moving the hands to an object and grasping of it needs to be guided by the perceived distance of the object from the hand (e.g., from vision), the position of the hand and arm relative to the object (from vision and proprioception) and the touch of the object itself (for neuropsychological studies concerning visual-tactile and proprioceptive interactions, see di Pellegrino and Làdavas, 1997; Làdavas, di Pellegrino, Farnè, and Zeloni, 1998; Làdavas, Farnè, Zeloni, and di Pellegrino, 2000; Maravita, Spence, Clarke, and Driver, 2000).

A wide range of studies have demonstrated the existence of many areas in the brain in which multisensory integration can take place. This is indicated by a convergence of input coming from different unimodal areas, and the presence of neurons that can be activated by stimuli from two or more sensory modalities. For example, Stein and his colleagues (Kadunce, Vaughan, Wallace and Stein, 2001; Stein, 1998; Stein and Meredith, 1993; Stein and Wallace, 1996; Wallace and Stein, 1994) have documented the presence of multisensory neurons in the superior colliculus (SC), a midbrain structure which is involved in overt attention and orienting behaviour, and in the anterior ectosylvian sulcus, an association area in parietal cortex which seems to modulate SC multisensory neuron activity. In both structures the multisensory neurons respond to two or three modalities (vision, audition and touch). They have multiple receptive fields (RFs), each one responding to a different sensory modality, and overlapping one another. When two stimuli are presented in two different modalities and in similar positions, an enhancement of activity is registered; in contrast, neuronal activity is depressed when the two stimuli are presented in different positions. In the SC the multimodal neurons have small receptive fields while in the anterior ectosylvian sulcus the RFs are much larger, facilitating integration on a wider scale. Human studies, using fMRI, have also provided evidence for multimodal as well as unimodal representations of space. For example, Macaluso and Driver (2001) stimulated participants tactually on their left or right hand, or visually at locations close to the hands. They found posterior occipital regions responded in a spatially specific manner to visual stimuli, and the postcentral gyrus (part of somatosensory cortex) responded in a spatially specific manner to tactile stimuli. In contrast, the anterior part of the intraparietal sulcus responded in a spatially specific manner irrespective of the modality of the input. This suggests that this region of associative cortex can provide information about the positions of objects in space, independent of the sense by which the input was delivered.

4. CROSS-MODAL ATTENTION

In addition to space being represented in a multi- as well as a uni-modal manner, there is evidence for attentional processes operating across as well as within specific modalities. Research here has focused on processes such as orienting of attention, that are known to

influence the processing of unimodal – particularly visual – stimuli. The orienting of attention in the visual modality has been often analysed using a paradigm developed by Posner (e.g. Posner, Snyder, and Davidson, 1980). In a typical task, subjects have to fixate the centre of a display, where a cue and a target are presented on either side of the screen. The subjects' task is to give a response as soon as possible after the presentation of the target. Typically subjects respond more rapidly when cue and target are presented on the same side of the display – on valid cue trials – than when cue and target are presented on opposite sides – on invalid cue trials. The usual explanation for this is that the cue causes attention to move and then engage at the location where the cue was presented. If the target appears at the same position, attention is already focused on that location and responses to the target will be relatively fast. However, if the target is presented on the opposite side of space, then spatial attention will have to disengage, move and then re-engage at the new position, and this will result in a delayed response (Posner, Walker, Friedrich and Rafal, 1984).

Buchtel and Butter carried out several experiments on endogenous¹ attentional orienting (Buchtel and Butter, 1987; Butter, Buchtel and Santucci, 1989). They found similar results to those obtained for visual orienting when the cue and target were presented in different modalities. For example, an auditory cue presented on one side decreased the speed of response to a visual stimulus presented on the opposite side, and, similarly, the speed of response to a tactile stimulus was influenced by a visual cue (Butter *et al.*, 1989). However, they did not find effects of visual cues on auditory targets. Using patients with parietal lobe damage, Farah, Wong, Monheit and Morrow (1989) found similar results, but in this case an auditory cue to the ipsilesional side disrupted the detection of a visual target on the contralesional side. These data are compatible with the idea of a supra-modal attentional network, located within the parietal lobe, that uses cues to develop spatial expectations irrespective of the modality of the cue. With auditory targets, however, responses may not be based on a location-coded representation, so spatial cueing may be less effective.

The idea that the basic mechanisms of spatial attention are similar independent of whether one is dealing with unimodal or multimodal stimuli has been confirmed by other research conducted systematically over the last few years. Spence and colleagues (Driver and Spence, 1998; Rorden and Driver, 1999; Spence and Driver, 1997; Spence, Nicholls, Gillespie, and Driver, 1998) have used paradigms similar to those of Buchtel and Butter. In one set of experiments (Spence *et al.*, 1998) normal subjects had to perform speeded discrimination tasks on stimuli preceded by spatially uninformative peripheral cues presented in a different modality. When cue and target stimuli were presented on opposite sides, the discrimination tasks were performed more slowly than when cues and targets were presented on the same side. Spence *et al.* found that auditory and visual cues affect the orienting of attention in touch (affecting tactile discriminations), that tactile cues can affect the orienting of visual attention, but (like Butter *et al.*, 1989) that auditory orienting was not affected by visual cues. From these studies it emerges that the mechanisms determining the orienting of attention can operate independently of the modalities in which stimuli are presented. Spence and colleagues suggest the existence of what they describe as “cross-modal links” in spatial attention, rather than opting for a view that there is a single, supramodal system.

¹ Endogenous attention is deliberately directed, on the basis of expectations or task requirements. On the contrary, exogenous attention is captured by a sudden stimulus, independently of our intentions.

Independently of whether there is a general attention system for all perceptual modalities or different modality-specific systems, the different modalities certainly share some resources. More evidence of this comes from studies of *inhibition of return* and the *attentional blink* (Arnell and Jolicoeur, 1999; Lloyd, Bolanowsky, Howard, and McGlone, 1999). Orienting of attention to a location is facilitated when a stimulus is immediately preceded by a cue on the same side of space. However, if the interval between the cue and the stimulus is increased (e.g., to 300 ms or longer), an inhibition of attentional orienting to the cue location is observed. In this case, RTs are slowed to targets at cued relative to uncued locations: “inhibition of return” (IOR; Posner and Cohen, 1984). IOR is observed typically when visual orienting of attention is tested (for a review see Klein, 2000), but it has also been described in the tactile (Lloyd *et al.*, 1999) and auditory (Mondor, Breau and Milliken, 1998) modalities. IOR has also been observed in multimodal stimulus presentation conditions (Lloyd *et al.*, 1999; Spence and Driver 1998a, 1998b). Spence and colleagues (Spence, Lloyd, McGlone, Nicholls and Driver, 2000), for example, presented a random sequence of visual, auditory and tactile targets, each preceded by a cue, (the interval between cue and target presentation varied between 950 and 2250 ms). They found that speeded detections of the target were slower when cue and target were presented on the same side of space, regardless of the modality of the cue. They concluded that IOR is “truly supramodal” and argued that it may involve brain areas (e.g., the superior colliculus and parietal cortex; see Bartolomeo, Chokron and Sieroff, 1999; Leipsen and Pollman, 2002; Sapir, Soroker, Berger and Henik, 1999; Vivas, Humphreys and Fuentes, 2003) in which neurons respond to visual, tactile and auditory inputs.

Temporal limitations on selection are demonstrated by the attentional blink (AB) phenomenon. The AB was first observed within vision (Raymond, Shapiro, and Arnell, 1992; Shapiro, Raymond and Arnell, 1994). It occurs when participants are required to respond to two targets presented within a rapid sequence of stimuli: the perception of the second target is impaired if the interval between the two targets is less than about 450 ms. This is an attentional rather than a perceptual limitation, since the second target can be identified relatively easily when the first target has to be ignored rather than identified. One proposal is that the AB reflects the time taken to consolidate a representation of the first target in short-term memory. During the time taken to consolidate the first target, the second target is not consolidated and can thus be lost for report. Arnell and Jolicoeur (1999) found that the AB phenomenon can occur under cross-modal presentation conditions, for example when a visual target follows an auditory target. This may reflect a central bottleneck, in which stimuli cannot be consolidated in short-term memory in parallel across different modalities (Arnell and Jolicoeur, 1999; Jolicoeur, 1999. Note, however, that Duncan, Martens and Ward, (1997) failed to find cross-modal AB using visual and auditory stimulus sequences)².

² It is worth noting that in some of their experiments Arnell and Jolicoeur used a “psychological refractory period” (PRP) paradigm, in which reaction times (RTs) are used as a measure of interference between two tasks (therefore, speeded responses are required). In the AB paradigm (Raymond *et al.*, 1992; Shapiro *et al.*, 1994) accuracy is measured instead of RTs (therefore, responses are unspeeded). This difference has led to different interpretations for the two phenomena. The most widely accepted interpretation for PRP (e.g., Pashler, 1994) is the dual-task interference is caused by interference on response selection operations. In contrast, AB theories have focused more on early stimulus encoding process (Raymond *et al.*, 1992; Shapiro *et al.*, 1994). This might suggest a different locus of interference for AB within vision and the cross-modal AB described by Arnell and Jolicoeur (1999).

Further evidence of cross-modal interactions in attention comes from neuropsychological data on brain damage patients. Studies on patients with brain lesions have shown that the parietal lobe plays an important role in the normal functioning of spatial attention (e.g. Driver, 1998; Friedrich *et al.*, 1998). Attentional deficits resulting from parietal damage have been observed most widely in visual perception, but they have also been reported in auditory and tactile perception. Mattingley, Driver, Beschin and Robertson (1997) examined three patients with right hemisphere damage who showed unimodal extinction for visual and tactile stimuli. That is, the patients could report a single stimulus presented on the contralesional side, but failed to detect the same item when it occurred simultaneously with another stimulus on the ipsilesional side. Interestingly, these patients also showed cross-modal extinction: a visual stimulus presented on the contralesional side was extinguished by a tactile stimulus presented on the ipsilesional side, and vice versa (a tactile stimulus was extinguished by a visual stimulus). Mattingley *et al.* concluded that the stimuli compete for attracting attention not only when they are within the same modality (Ward, Goodrich and Driver, 1994) but also when they are from different modalities (see also Duncan, 1996).

What was widely believed for visual attention – that attention is a limited resource and that, therefore, there is a competition for attracting it – seems now to apply for cross-modal perception: attentional resources appear to be shared between modalities. However, the extent of any sharing is not yet clear. Early discoveries of cross-modal attentional interference were interpreted in terms of a supramodal attentional system, which controls attention orienting independently of the modality in which the stimuli are presented (Farah *et al.*, 1989). Others have proposed the existence of amodal systems that mediate the interactions of unimodal systems. For instance, Arnell and Jolicouer (Arnell and Jolicouer, 1999; Jolicouer, 1999) claim that AB effects are produced by dual-task interference at a late stage of stimulus processing that requires central (supramodal) mechanisms. These mechanisms are capacity limited and only one operation can be performed optimally at any given time. Hence performance can be impaired even when two targets are presented in different modalities. In contrast, Spence, Driver and colleagues (e.g. Spence and Driver, 1998a), from their studies on attention orienting, have concluded that there are cross-modal links between unimodal attentional systems. They further suggest that cross-modal integration occurs on the basis of feedback projections from multimodal to unimodal areas (Driver and Spence, 2000; Macaluso and Driver, 2004). This view is supported by studies on cross-modal spatial attention such as that of Macaluso, Frith and Driver (2000a). They used event-related fMRI to show that tactile stimulation concurrent with visual stimulation, on the same side of space, can enhance the activity in the contralateral visual cortex (lingual gyrus, a supposedly unimodal area) above that achieved with visual stimulation alone. From a further analysis of the effective connectivity between brain areas, the authors suggested that the somatosensory cortex influences the unimodal visual cortex via back-propagations from multimodal parietal areas, allowing the visual cortex to be informed of tactile stimuli received simultaneously with visual stimuli. According to the authors, this is how cross-modal links in spatial attention are manifest. Further support to the idea of cross-modal links between unimodal areas has been provided by event-related potential (ERP) studies conducted by Eimer and colleagues (for a review see Eimer and Driver, 2001) where it is shown that directing attention to one side of space to detect target stimuli in one (relevant) modality can modulate early ERP components produced by the stimuli presented not only in that modality but also in other (irrelevant) modalities. For example, in a series of experiments participants were presented with stimuli in

two different modalities, vision and audition, one of which was task relevant, while the other was not (Eimer and Schroger, 1998; Eimer, van Velzen and Driver, 2004). The task was to detect targets presented in the relevant modality at a given location. When the relevant modality was vision and visual stimuli were presented at the attended location, there was an enhancement of the N1 components of ERPs in occipital areas, compared to when visual stimuli were presented at irrelevant locations (see also Mangun and Hillyard, 1991). However, enhancement of the N1 components at occipital areas were found even when audition was the relevant modality, while visual stimuli had to be ignored, thus demonstrating the existence of cross-modal attentional links between vision and audition. Comparable effects were found for early auditory components when audition was respectively the relevant and irrelevant modality (Eimer and Schroger, 1998. For similar studies where vision and touch were examined see Eimer and Driver, 2000).

5. CROSS-MODAL IDENTIFICATION

The majority of the work on multimodal integration has focused on spatial features of stimuli, particularly on their spatial coding and on how the spatial locations of stimuli affect attentional orienting (see above). However, there are also studies investigating multimodal integration in stimulus identification. What emerges here is that the senses “collaborate” with each other to improve, disambiguate or speed-up stimulus identification. Again, in experimental contexts, the collaboration between senses becomes more evident when what is experienced from one sense is in contrast with what felt from another. In such circumstances, there can be either more reliance on one sense than on the others - known as *perceptual dominance* - or a kind of trade-off between the information provided by one sense and that provided by another.

A classic case of perceptual dominance was reported by Gibson in 1933. In his experiments subjects touched straight edges while wearing prism spectacles, which made the straight edges appear curved. Although visual and kinaesthetic information was in conflict, the subjects did not experience that conflict and reported that the edges were curved. Vision dominated over kinaesthesia³. Later, vision was shown to dominate in many other different situations, and not only over kinaesthesia or touch (Klein and Posner, 1974; Posner, Nissen and Klein, 1976; Rock and Victor, 1964), but also over audition (Colavita, 1974). Posner and colleagues (Posner *et al.*, 1976) suggested that visual dominance is related to a weak alerting capacity of visual stimuli, compared to other modalities. In response to that, subjects are biased in their attention toward the visual modality, with a consequent reduction in attention to other modalities. However, vision is not always the dominant modality. In fact, there are circumstances in which touch is more reliable than vision (Lederman, Thorne and Jones, 1986). Interestingly, Heller (1982) found that, when people are asked to make texture judgments about abrasive surfaces, touch is better than vision for strict texture judgements, however, vision helps to control hand movement for haptic exploration. As a result, co-

³ The term kinaesthesia is used to indicate skin sensation and perception of limb movements, through receptors in the muscles, tendons and joints. In some cases, it is interchanged with the term proprioception. The term touch normally indicates the sensing, through cutaneous receptors, of spatial and surface properties of objects such as shape and texture.

operation between touch and vision can be better than single-modality exploration. This suggests that whether or not a sensory modality dominates over another depends on the nature of the task to be performed. For example, vision might be better for the processing of spatial features (e.g., size, shape and distance), while audition might give more precise recognition of temporal sequences. This might be the reason why, in the ventriloquism effect, a visual stimulus influences the perceived location of a sound to a much greater extent than the effect that a sound can have on the judgements about the location of a visual stimulus. On the other hand, touch appears to be better than the other senses for the discrimination of surface texture. When required to judge object size using both the visual and the tactile modalities, people seem to rely much more on vision than on touch (e.g. Rock and Victor, 1964). However, in tasks requiring texture judgements (e.g. smoothness, roughness), people can rely equally on one or the other sense (Lederman and Abbott, 1981; Heller, 1982; Jones and O'Neil, 1985), or touch can even dominate over vision (Klatzky, Lederman and Matula, 1993; Lederman, Thorne and Jones, 1986). In those cases perceptual dominance depends on the particular feature that has to be evaluated. This has been elegantly demonstrated by Klatzky, Lederman and Matula (1993) in a series of experiments where subjects had to compare objects with respect to either a material feature (which could be roughness, hardness, weight or temperature) or a geometric feature (either size or shape complexity). It was observed that tactile exploration was more frequent and was initiated faster than visual exploration when material judgments had to be made. However, when geometric properties had to be evaluated, touch was used less often than for material judgments, and it was used primarily for bringing the objects to better view, rather than for exploration.

The collaboration of our senses appears to be constant and fundamental for human cognition, even in cases in which we are apparently relying on one sense only. It might not be of surprise that both touch and vision contribute to the identification of surface textures, but how about sounds? Guest and colleagues (Guest, Catmur, Lloyd and Spence, 2002; see also Jousmäki and Hari, 1998) have shown that, perhaps surprisingly, auditory perception influences the identification of features that are apparently in the dominion of touch. They found that the tactile perception of roughness, was biased by an incongruent touch-related sound (the higher frequencies of the sound produced by touching the texture were either reduced or amplified in realtime). What was even more striking was that sound modulated also the perception of the wetness of the hands.

There are cases in which, when information is incongruent, none of the senses seems to dominate, which results in a form of trade-off between modalities. An example of this is given by the McGurk effect (MacDonald and McGurk, 1978; McGurk and MacDonald, 1976). The McGurk effect occurs when subjects are looking at lip movements for one syllable whilst the sound produced is a different syllable. The sound experienced then is a third syllable. For example, when the lip movement is “ba” and the sound produced is “ga”, the syllable commonly experienced is “da”. This illusion has been originally described as a case of perceptual fusion. Consistent with this, stimulus identification is based equally on the two modalities, and neither dominates over the other (Easton and Basala, 1982, found that when words were presented instead of syllables the subjects relied almost completely on audition, while ignoring the visual information).

Another example of the mutual influence of sound and vision has also emerged from studies on the perception of emotions. De Gelder and Vroomen (2000) have shown that when asked to identify the emotion (happiness/sadness of different degrees) expressed by a face,

subjects are biased by the emotion expressed by a concurrent voice (also expressing happiness/sadness), even when subjects are asked to ignore the voice. The opposite was also found: emotion identification of a voice can be biased by the emotion expressed by a concurrent face. This suggests that face and voice expressions appear to be equally important for the perception of emotions. De Gelder and Vroomen attributed this phenomenon to perceptual rather than cognitive (or post-perceptual) mechanisms.

6. CROSS-MODAL FEATURE BINDING

Our perception of coherent objects, rather than collections of features, requires that the elements making up an object be bound together, so that the object can be properly encoded and identified. Feature binding has been widely studied in visual perception, but there is still little understanding of how feature binding is performed when the different features of an object are perceived by different sensory modalities. One way of decomposing the binding process, to analyse the underlying processes, is to study cases in which binding goes wrong – when so-called illusory conjunctions (ICs) of features occur. This happens when participants form an illusory percept by combining features that belong to different objects.

The first experimental study of ICs was conducted by Treisman (Treisman and Schmidt, 1982). She suggested that different visual features may be coded independently by early preattentive vision, with the features being represented in separate spatial maps. Subsequently, the features are bound by attention to their common location (Treisman, 1998; Treisman and Gelade, 1980). According to Treisman, ICs can occur when objects and their features are not under the focus of attention. ICs have often been demonstrated with features presented within the same modality (typically in the visual modality, but also in audition; e.g., Hall, Pastore, Acker and Huang, 2000). Cinel, Humphreys and Poli (2002) found that ICs could be formed not only within a single modality but also when stimuli are presented in different modalities. In the typical experiments, participants had to touch a textured bar (occluded from view) and, at the same time, they looked at a visual display where two shapes with textures were briefly presented. The participants had to identify both the orientation of the tactile bar (this task forced participants to sense the tactile texture without necessarily focusing attention on it) and the shape and texture of each visual stimulus. It was found that cross-modal IC errors could be made, in which touched textures were reported as having been seen (tactile→visual ICs). In a further experiment (where visual shapes, but not textures, had to be reported along with the texture and orientation of the touched bar) the reverse result was observed, i.e., visual textures were reported as being felt (visual→tactile ICs). Cross-modal visual→tactile ICs were as frequent as cross-modal tactile→visual ICs, under similar presentation conditions. The experiments of Cinel et al. (2002) suggest that multimodal integration of sensory information happens at very low levels of perceptual processing, so that mis-attributions of modalities can arise under conditions of inattention. There was, however, no clear evidence of whether cross-modal ICs were strictly perceptual illusions or whether they originate from a higher cognitive level. ICs might have been caused by memory failure. In this perspective, ICs would originate because, even though initially features are correctly perceived, the information about the correct binding might decay rapidly and, by the time a verbal response is given, participants might only remember what features they have

perceived, but forget what feature belonged to what object. Several subsequent experiments provided explicit tests of whether cross-modal ICs were due to memory failures or to impaired perceptual binding of stimuli to modalities. On no occasion was evidence found for the errors arising in memory. For example, in a variant of the experiments designed to minimise the load on perceptual report, participants were given a forced-choice task where they had just to say whether or not a target stimulus (defined by shape and texture) was presented. If ICs were an effect of memory, then they should have occurred less frequently than in full-report tasks. Against this, cross-modal ICs were made with a frequency comparable (if not higher) to that observed in full-report experiments. In other experiments cross-modal ICs were found to be independent of the memory load, providing more evidence for a perceptual account for cross-modal ICs. Apparently, under conditions of limited attention, a stimulus registered in one modality may create linked activation in another modality and sometimes can be mis-interpreted in terms of this secondary activity.

Another important aspect of feature binding concerns the role of spatial attention. ICs are relatively frequent for stimuli that are not under the focus of attention or in conditions of high attentional load (i.e., when attention is divided between a large number of items) (Tresiman and Schmidt, 1982). This holds for cross-modal as well as for unimodal ICs. In Cinel *et al.* (2002) there was no evidence for cross-modal ICs when attention was focused on the visual and tactile textures. This suggests that focusing attention on a feature prevented that feature from being bound to the incorrect modality, as well as to other features within the same modality. For example, when we attend to a stimulus we may facilitate top-down activation that helps to consolidate initial bottom-up activation from a stimulus, stabilising its representation in space, time and sensory quality.

Cinel and colleagues, in their studies of cross-modal feature binding, also reported data from a patient with left parietal damage. Damage to the parietal lobe is often associated with impairments of attention, particularly to the contralesional side (see Halligan, Fink, Marshall and Vallar 2003). The patient examined showed significant numbers of cross-modal ICs even when using relatively long stimulus exposure and these tended to be greater when visual stimuli fell on his contralesional side (in the right visual field). ICs occurred when there was concurrent presentation of the felt and seen textures, but not when their presentations were staggered, providing more evidence for a perceptual account of cross-modal ICs – binding across modalities being more likely when the stimuli have a common time code. Taken together the evidence is consistent with information converging preattentively for binding from different sensory modalities, and with binding process being modulated by the parietal lobe.

There is converging evidence suggesting that the parietal lobe is a potential locus for cross-modal integration as well as playing an important role in spatial attention (e.g. Andersen, Snyder, Bradley, and Xing, 1997; Banati, Goerres, Tjoa, Aggleton, and Grasby, 2000; Macaluso *et al.*, 2000a, 2000b). For example, in a tactile-visual matching task where participants had to make a cross-modal match between a tactile stimulus and some visual stimuli, Banati *et al.* (2000) found that the inferior parietal lobes were amongst the areas where an increased activity was detected. It is possible that the normal cross-modal integration process involves “tagging” information in order to co-register it with information from other modalities. Parietal damage may disrupt the ability to “read” such tags, so that inputs are mis-attributed to the wrong modality of origin.

Two important aspects of cross-modal feature binding emerging from the study on cross-modal ICs are the spatial and temporal distribution of the stimuli. Cross-modal ICs were more frequent for stimuli presented in the same hemifield than for stimuli presented across hemifields, and when the stimuli occurred at the same time (Cinél et al., 2002). It appears that our perceptual systems have evolved to be sensitive to temporal co-incidence, where stimuli that are concurrent in time and space tend to be perceived as belonging to the same object. Other evidence on temporal cues for cross-modal binding comes from studies on stimulus segregation. Vroomen and de Gelder (2000) found that discrimination of a visual target embedded in a sequence of visual distractors improved when the target was synchronous to a high tone sound embedded in a sequence of low tone sounds. An easy interpretation of this facilitation effect might be that the sound triggers attentional alerting, which, in turn, makes the visual target detection easier (see also Doyle and Snowden, 2001 for similar conclusions). However, Vroomen and de Gelder demonstrated that this was not the case. An alternative explanation might be that the synchronised presentation of sound and visual target leads to perceptual grouping and, as a consequence of that, the visual target bound to an auditory tone segregates more easily from the background of visual distractors. Surprisingly, this effect was not observed any more when the high tone was part of a simple melody, and therefore grouped with the other sounds. Melody grouping prevented the high tone sound from being bound to the visual target.

7. CONCLUSION

In this chapter we have described some of the recent research on multisensory integration. Even though there is now considerable evidence of cross-modal integration, understanding the mechanisms involved is still a challenge. One of the main problems is to discover the level of processing at which integration occurs. Based on the view that multisensory integration occurs naturally in the real world, we might presume that integration takes place at early stages of perceptual processing through, for example, connections converging from unimodal to multimodal cells (Stein and Meredith, 1993). In this case, one would experience only the result of integration, and cross-modal divergence could not be experienced at all. However, there could also be post-perceptual integration, which might be the result of an adjustment where, for example, top-down processes (from memory or semantic systems) are involved. We have described several studies where researchers have tried to give evidence favouring one interpretation rather than the other. In reality, the conceptual distinction between perceptual and post-perceptual origins of cognitive mechanisms is not always clear, and, therefore, finding evidence for one or the other of the two possible interpretations can be even more difficult...

7.1. Future Research

There remain numerous questions about how multisensory integration takes place. These include:

- While in some conditions it is easily understandable why a perceptual modality can dominate over another, it is not yet completely clear why this occurs in other conditions – such as in the ventriloquism effect, where vision affects audition but not vice versa, or in cross-modal attentional orienting, where attentional orienting is affected by an auditory cue, but not vice-versa. This might depend on basic differences between the sensory modalities on how stimuli are processed (e.g., location coding being fundamental to visual and tactile inputs, but not to audition).
- Cross-modal feature binding: exploring feature binding between vision, touch and audition. At present, studies have examined binding between vision and touch. These may represent two highly compatible forms of input, unlike other sensory modalities (such as audition), which might have quite different sensory properties. What factors govern the binding of audition to vision and touch, and vice versa?
- There is evidence that selection within a single modality is affected by grouping (e.g., Ward et al., 1994). Do similar processes take place across modalities (e.g., see Vroomen and de Gelder, 2000)?
- Understanding the neural substrates of multisensory integration. There are studies revealing the presence of multimodal neurons in some brain areas, with the neurons responding maximally to multisensory stimulation (e.g., Stein, Wallace, Stanford and Jiang, 2002). While these neurons are likely to be important for multisensory integration, they might not be the only way of bringing cross-modal information together. Another different way of multimodal integration can be through interactions between unimodal areas of the brain (see Malacuso and Driver, 2004). Another possibility is temporal synchronisation of neural activity (Singer and Gray, 1995). Further research is required to provide a more detailed account of whether each of these processes operate, and if so, how they co-operate to facilitate integration.

REFERENCES

- Andersen, R. A., Snyder, L. H., Bradley, D. C., and Xing, J. (1997). Multimodal representation of space in the posterior parietal cortex and its use in planning movement. *Annual Review in Neuroscience*, 20, 303-330.
- Arnell, K.M., and Jolicouer, P. (1999). The attentional blink across stimulus modalities: evidence for central processing limitations. *Journal of Experimental Psychology: Human Perception and Performance*, 25(3), 630-648.
- Banati, R.B., Goerres, G.W., Tjoa, C., Aggleton, J.P., and Grasby, P. (2000). The functional anatomy of visual -tactile integration in man: a study using positron emission tomography. *Neuropsychologia*, 38, 115-124.
- Bartolomeo, P., Chokron, S., and Sieroff, E. (1999). Facilitation instead of inhibition for repeated right-side events in left neglect. *Neuroreport*, 10, 3353-3357.
- Bermant, R.I., and Welch, R.B. (1976). The effect of degree of visual-auditory stimulus separation upon the spatial interaction of vision and audition. *Perceptual and Motor Skills*, 43, 487-493.

- Bertelson, P. (1994). The cognitive architecture behind auditory-visual interaction in scene analysis and speed identification. *Current Psychology of Cognition*, 13, 69-95.
- Bertelson, P. (1998). Starting from the ventriloquist: The perception of multimodal events. In M. Sabourin, F. I. M. Craik and M. Rober (Eds.), *Advances in Psychological Sciences, Vol. 1: Biological and Cognitive Aspects*. Hove, UK: Psychology Press.
- Bertelson, P. (1999). Ventriloquism: A case of crossmodal perceptual grouping. In G. Aschersleben, T. Bachmann, and J. Müsseler (Eds.), *Cognitive Contributions to the Perception of Spatial and Temporal Events*. Elsevier Science B. V.: Amsterdam.
- Bertelson, P., and Aschersleben, G. (1998). Automatic visual bias of perceived auditory location. *Psychonomic Bulletin and Review*, 5, 482-489.
- Bertelson, P., Pavani, F., Làdavas, E., Vroomen, J., and de Gelder, B. (2000). Ventriloquism in patients with unilateral visual neglect. *Neuropsychologia*, 38, 1634-1642.
- Bertelson, P., and Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Perception and Psychophysics*, 29, 578-587.
- Bertelson, P., Vroomen, J., de Gelder, B., and Driver, J. (2000). The ventriloquist effect does not depend on the direction of deliberate visual attention. *Perception and Psychophysics*, 62, 321-332.
- Botvinick, M., and Cohen, J. (1998). Rubber hands “feel” touch that eyes see. *Nature*, 391, 756.
- Buchtel, H.A., and Butter, C.M. (1988). Spatial attention shifts: Implication for the role of polysensory mechanisms. *Neuropsychologia*, 26, 499-509.
- Butter, C.M., Buchtel, H.A., and Santucci, R. (1989). Spatial attentional shifts: Further evidence for the role of polysensory mechanisms using visual and tactile stimuli. *Neuropsychologia*, 27, 1231-1240.
- Caclin, A., Soto-Faraco, S., Kingstone, A., and Spence, C. (2002). Tactile “capture” of audition. *Perception and Psychophysics*, 64:616-30.
- Cinèl, C., Humphreys, G.W., and Poli, R. (2002). Cross-modal illusory conjunctions between vision and touch. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 1243-66.
- Colavita, F.B. (1974). Human sensory dominance. *Perception and Psychophysics*, 16, 409-412.
- de Gelder, B., and Vroomen, J. (2000). The perception of emotions by ear and eye. *Cognition and Emotion*, 14, 289-311.
- di Pellegrino, G., and Làdavas, E. (1997). Seeing where your hands are. *Nature*, 388, 730.
- Doyle, M.C., and Snowden, R.J. (2001). The effects of sound upon visual target acquisition: single targets. *Perception*, 30, 795-810.
- Driver, J. (1998). Object-based and egocentric visual neglect. In N. Burgess and J. O’Keefe (Eds.), *The hippocampal and parietal foundations of spatial cognition*. Oxford: Oxford University Press.
- Driver, J., and Spence, C. (1998). Attention and the crossmodal construction of space. *Trends in Cognitive Science*, 2, 254-262.
- Driver, J., and Spence, C. (2000). Multisensory perception: Beyond modularity and convergence. *Current Biology*, 10, R731-R735.
- Duncan, J. (1996). Coordinated brain systems in selective perception and action. In T. Inui and J.L. McClelland (Eds.), *Attention and Performance XVI*. MIT Press, Cambridge, MA, pp. 311-331.

- Duncan, J., Martens, S., and Ward, R. (1997). Restricted attentional capacity within but not between sensory modalities. *Nature*, 387, 808-810.
- Easton, R.D., and Basala, M. (1982). Perceptual dominance during lipreading. *Perception and Psychophysics*, 32, 562-570.
- Eimer, M., and Driver, J. (2000). An event-related brain potential study of cross-modal links in spatial attention between vision and touch. *Psychophysiology*, 37, 697-705.
- Eimer, M., and Driver, J. (2001). Crossmodal links in endogenous and exogenous spatial attention: Evidence from event-related brain potential studies. *Neuroscience and Biobehavioral Reviews*, 25, 497-511.
- Eimer, M., and Schöger, E. (1998). ERP effects of intermodal attention and cross-modal links in spatial attention. *Psychophysiology*, 35, 313-327.
- Eimer, M., van Velzen, J., and Driver, J. (2004). ERP evidence for cross-modal audiovisual effects of endogenous spatial attention within hemifields. *Journal of Cognitive Neuroscience*, 16, 272-288.
- Farah, M.J., Wong, A.B., Monheit, M.A., and Morrow, L.A. (1989). Parietal lobe mechanisms of spatial attention: modality specific or supramodal? *Neuropsychologia*, 27(4), 461-470.
- Friedrich, F.J., Egly, R., Rafal, R.D., and Beck, D. (1998). Spatial attention deficits in humans: A comparison of superior parietal and temporal-parietal junction lesions. *Neuropsychology*, 12, 193-207.
- Gibson, J.J. (1962). Observations on active touch. *Psychological Review*, 69, 477-491.
- Graziano, M.S.A., Gregory, S.Y., and Gross, C.G. (1994). Coding of visual space by premotor neurons. *Science*, 266, 1054-1057.
- Graziano, M.S.A., and Gross, C.G. (1996). Multiple pathways for processing visual space. In T. Inui and J.L. McClelland (Eds). *Attention and Performance XVI*. MIT Press, Cambridge, MA, pp. 181-207.
- Guest, S., Catmur, C., Lloyd, D., and Spence, C. (2002). Audiotactile interactions in roughness perception. *Experimental Brain Research*, 146, 161-171.
- Hall, M.D., Pastore, R.E., Acker, B.E., and Huang, W. (2000). Evidence for auditory feature integration with spatially distributed items. *Perception and Psychophysics*, 62, 1243-1257.
- Halligan, P.W., Fink, G.R., Marshall, J.C., and Vallar, G. (2003). Spatial cognition: evidence from visual neglect. *Trends in Cognitive Sciences*, 7, 125-133.
- Heinke, D. and Humphreys, G.W. (2003). Attention, spatial representation and visual neglect: Simulating emergent attention and spatial memory in the Selective Attention for Identification Model (SAIM). *Psychological Review*, 110, 29-87.
- Heller, M.A. (1982). Visual and tactual texture perception: Intersensory cooperation. *Perception and Psychophysics*, 31, 339-344.
- Jolicouer, P. (1999). Restricted attentional capacity between sensory modalities. *Psychonomic Bulletin and Review*, 6, 87-92.
- Jones, B., and O'Neil, S. (1985). Combining vision and touch in texture perception. *Perception and Psychophysics*, 37, 66-72.
- Jousmäki, V., and Hari, R. (1998). Parchment-skin illusion: Sound-biased touch. *Current Biology*, 8, R190.

- Kadunce, D.C., Vaughan, L.V., Wallace, M.T., and Stein, B.E. (2001). The influence of visual and auditory receptive field organization on multisensory integration in the superior colliculus. *Experimental Brain Research*, 139, 303-310.
- Klatzky, R.L., Lederman, S.J., and Matula, D.E. (1993). Haptic exploration in the presence of vision. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 726-743.
- Klein, R.M. (2000). Inhibition of return. *Trends in Cognitive Science*, 4, 138-147.
- Klein, R.M., and Posner, M.I. (1974). Attention to visual and kinesthetic component of skills. *Brain Research*, 71, 401-411.
- Làdavas, E., Farnè, A., Zeloni, G., and di Pellegrino, G. (2000). Seeing or not seeing where your hands are. *Experimental Brain Research*, 131, 458-467.
- Làdavas, E., di Pellegrino, G., Farnè, A., and Zeloni, G. (1998). Neuropsychological evidence of an integrated visuotactile representation of peripersonal space in humans. *Journal of Cognitive Neuroscience*, 10, 581-589.
- Lederman, S.J., and Abbott, S.G. (1981). Texture perception: Studies of intersensory organization using a discrepancy paradigm, and visual versus tactile psychophysics. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 902-915.
- Lederman, S.J., Thorne, G., Jones, B. (1986). Perception of texture by vision and touch: Multidimensionality and intersensory integration. *Journal of Experimental Psychology: Human Perception and Performance*, 12, 169-180.
- Lepsein, J., and Pollman, S. (2002) Covert reorienting and inhibition of return: An event-related fMRI study. *Journal of Cognitive Neuroscience*, 14, 127-144.
- Lloyd, D.M., Bolanowsky, S.J., Howard, L., and McGlone, F. (1999). Mechanisms of attention in touch. *Somatosensory and Motor Research*, 16(1), a 3-10.
- Macaluso, E., and Driver, J. (2001) Spatial attention and cross-modal interactions between vision and touch. *Neuropsychology*, 39, 1304-1316.
- Macaluso, E., and Driver, J. (2004) Functional imaging of crossmodal spatial representations and crossmodal spatial attention. In C. Spence and J. Driver (Eds.), *Crossmodal space and crossmodal attention*. (pp. 247-276). Oxford University Press: Oxford.
- Macaluso E., Frith C., and Driver J. (2000a). Modulation of human visual cortex by crossmodal spatial attention. *Science*, 289, 1206-1208.
- Macaluso E., Frith C., and Driver J. (2000b). Selective spatial attention in vision and touch: Unimodal and multimodal mechanisms revealed by PET. *Journal of Neurophysiology*, 83, 3062-3075.
- MacDonald, J., and McGurk, H. (1978). Visual influences on speech perception. *Perception and Psychophysics*, 24, 253-257.
- Mangun, G.R., and Hillyard, S.A. (1991). Modulations of sensory-evoked brain potentials indicate changes in perceptual processing during visual-spatial priming. *Journal of Experimental Psychology: Human Perception and Performance*, 17, 1057-1074.
- Maravita, A., Spence, C., Clarke, K., and Driver, J. (2000). Vision and touch through the looking glass in a case of crossmodal extinction. *NeuroReport*, 11, 3521-3526.
- Mattingley, J.B., Driver, J., Beschin, N., and Robertson I.J. (1997). Attentional competition between modalities: Extinction between touch and vision after right hemisphere damage. *Neuropsychologia*, 35(6), 867-880.

- McGlinchey-Berroth, R., Milberg, W.P., Verfaellie, M., Alexander, M. and Kilduff, P.T. (1993). Semantic processing in the neglected field: Evidence from a lexical decision task. *Cognitive Neuropsychology*, 10, 79-108.
- McGurk, H., and MacDonald, J. (1978). Hearing lips and seeing voices: A new illusion. *Nature*, 264, 746-748.
- Mondor, D.A., Breau, L.M., and Milliken, B. (1998). Inhibitory processes in auditory selective attention: Evidence of a location-based and a frequency-based inhibition of return. *Perception and Psychophysics*, 60, 296-302.
- Pavani, F., Spence, C., and Driver, J. (2000). Visual capture of touch: Out-of-the-body experiences with rubber gloves. *Psychological Science*, 11, 353-359.
- Posner, M.I., and Cohen, Y. (1984). Components of visual orienting. In H. Bouma, D. Bouwhuis (Eds.), *Attention and Performance X*. Erlbaum.
- Posner, M.I., Nissen, M.J., and Klein, R.M. (1976). Visual Dominance: An information-processing account of its origins and significance. *Psychological Review*, 83, 157-171.
- Posner, M. I., Snyder, C.R., and Davidson, B.J., (1980). Attention and the detection of signals. *Journal of Experimental Psychology*, 109, 160-174.
- Posner, M. I., Walker, J.A., Friedrich, F.J., and Rafal, R. (1984). Effects of parietal lobe injury on covert orienting of visual attention. *Journal of Neuroscience*, 4, 1863-74.
- Radeau, M., and Bertelson, P. (1977). Adaptation to auditory-visual discordance and ventriloquism in semirealistic situations. *Perception and Psychophysics*, 22, 137-146.
- Radeau, M., and Bertelson, P. (1987). Auditory-visual interaction and the timing of inputs: Thomas (1941) revisited. *Psychological Research*, 49, 17-22.
- Raymond, J.A., Shapiro, K.L., and Arnell, K.M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, 18, 849-860.
- Rock, I., and Victor, J. (1964). Vision and touch: An experimentally created conflict between the two senses. *Science*, 143, 594-596.
- Rorden, C., and Driver, J. (1999). Does auditory attention shifts in the direction of an upcoming saccade? *Neuropsychologia*, 37, 357-377.
- Rorden, C., Heutink, J., Greenfield, E., and Robertson, I.H. (1999). When a rubber hand "feels" what a real hand cannot. *NeuroReport*, 10, 135-138.
- Sapir, A, Soroker, N., Berger, A. and Henik, A. (1999). Inhibition of Return in spatial attention: Direct evidence for collicular generation. *Nature Neuroscience*, 2(12).
- Shapiro, K.L., Raymond, J.A., and Arnell, K.M. (1994). Attention to visual-pattern information produces the attentional blink in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 357-371.
- Singer, W., and Gray, C.M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience*, 18, 555-586.
- Spence, C., and Driver, J. (1997). Audiovisual links in exogenous covert spatial orienting. *Perception and Psychophysics*, 59, 1-22.
- Spence, C., and Driver, J. (1998a). Cross-modal links in exogenous covert spatial orienting between touch, audition and vision. *Perception and Psychophysics*, 60, 544-557.
- Spence, C., and Driver, J. (1998b). Inhibition of return following an auditory cue. The role of central reorienting of events. *Experimental Brain Research*, 118, 352-360.

- Spence, C., Lloyd, D., McGlone, F., Nicholls, M.E.R., and Driver, J. (2000). Inhibition of return is supramodal: a demonstration between all possible pairings of vision, touch, and audition. *Experimental Brain Research*, 134, 42-48.
- Spence, C., Nicholls, M.E.R., Gillespie, N., and Driver, J. (1998). Crossmodal links in exogenous covert spatial orienting between touch, audition, and vision. *Perception and Psychophysics*, 60, 544-557.
- Spence, C., Pavani, F., and Driver, J. (1998). What crossing the hands can reveal about visuo-tactile links in spatial attention. *Abstracts of the Psychonomic Society*, 3, 13.
- Stein, B.E. (1998). Neural mechanisms for synthesizing sensory information and producing adaptive behaviours. *Experimental Brain Research*, 123, 124-135.
- Stein, B.E., and Meredith, M. A. (1993). *The merging of the senses*. Cambridge, MA: MIT Press.
- Stein, B.E., and Wallace, M. T. (1996). Comparison of cross-modality integration in midbrain and cortex. *Progress in Brain Research*, 112, 289-299.
- Stein, B.E., Wallace, M.W., Stanford, T.R., and Jiang, W. (2002). Cortex governs multisensory integration in the midbrain. *Neuroscientist*, 8, 306-314.
- Treisman, A., and Schmidt, H. (1982). Illusory conjunction in the perception of objects. *Cognitive Psychology*, 14, 107-141.
- Treisman, A. (1998). Feature binding, perception and object perception. *Philosophical Transaction of the Royal Society*, 353, 1295-1306.
- Treisman, A., and Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, 12, 97-136.
- Treisman, A., and Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, 14, 107-141.
- Vivas, A.B., Humphreys, G.W. and Fuentes, L.J. (2003) Inhibitory processing following damage to the parietal lobe. *Neuropsychologia*, 41, 1531-1540.
- Vroomen, J., Bertelson, P., and de Gelder, B. (2001). The ventriloquist effect does not depend on the direction of automatic attention. *Perception and Psychophysics*, 63, 651-659.
- Vroomen, J., and 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.
- Wallace, M.T., and Stein, B.E. (1994). Cross-modal synthesis in the midbrain depends on input from association cortex. *Journal of Neurophysiology*, 71, 429-432.
- Ward, R., Goodrich, S., and Driver, J. (1994). Grouping reduces visual extinction: Neuropsychological evidence for weight-linkage in visual selection. *Visual Cognition*, 1, 101-129.