

Crossmodal processing

Charles Spence · Daniel Senkowski ·
Brigitte Röder

Published online: 19 August 2009
© Springer-Verlag 2009

Introduction

Why does the voice of the ventriloquist appear to come from the lips of his/her dummy? Why does food lose its taste when your nose is blocked? Do the blind hear better or the deaf have enhanced visual sensitivity? These are all questions of interest to researchers in the area of crossmodal information processing and multisensory integration. While the human (and animal) senses have traditionally been studied in isolation, researchers have, in recent years, come to the realization that they cannot hope to understand perception until the question of how information from the different senses is combined has been satisfactorily addressed (see Calvert et al. 2004; Dodd and Campbell 1987; Lewkowicz and Lickliter 1994; Spence and Driver 2004).

Multisensory information processing is an ubiquitous part of our daily lives; As such, understanding the rules of crossmodal perception and multisensory integration offers the promise of many practical interventions in a variety of real-life settings, from the design of multisensory warning signals to capture a distracted driver's attention more effectively (see Ferris and Sarter 2008; Ho and Spence

2008) to the reduction in the unhealthy ingredients (such as salt, sugar, fat and carbonic acid) in the food we eat using multisensory illusions (Spence 2002), through to the development of more effective crossmodal compensatory training strategies for those attempting to deal with sensory loss (e.g., Merabet et al. 2005; Röder and Rösler 2004; Rouger et al. 2007).

This special issue of *Experimental Brain Research* comprises a collection of original research papers and review articles highlighting some of the most exciting basic research currently being undertaken in the area of crossmodal information processing and multisensory integration. However, before proceeding, it is, perhaps, worth saying a few words with regards to terminology. Over the years, scientists in this area have used a wide range of different terms to try and capture the phenomena that they have discovered and/or researched on. So, for example, in the literature we find terms such as polymodal, metamodal, multimodal, intermodal, multisensory and crossmodal (to name but a few)! In this special issue, we have tried to encourage the authors to restrict themselves primarily to just the latter two terms.

The term *crossmodal* is normally used to refer to situations in which the presentation of a stimulus in one sensory modality can be shown to exert an influence on our perception of, or ability to respond to, the stimuli presented in another sensory modality: As, for example, when the presentation of a spatially nonpredictive auditory cue results in a shift of a participant's spatial attention that leads to a short-lasting facilitation in their ability to perceive/respond to stimuli presented from the same side in another sensory modality (say vision; see McDonald et al. 2000; Spence and Driver 1997).

Multisensory integration, by contrast, was the term originally coined by neurophysiologists to describe the

C. Spence (✉)
Crossmodal Research Laboratory, Department of Experimental
Psychology, Oxford University, OX1 3UD Oxford, UK
e-mail: charles.spence@psy.ox.ac.uk

D. Senkowski
Department of Neurophysiology and Pathophysiology,
University Medical Center Hamburg-Eppendorf, 20246
Hamburg, Germany

B. Röder
Biological Psychology and Neuropsychology, University of
Hamburg, 20146 Hamburg, Germany

interactions they observed at the cellular level when stimuli were presented in different sensory modalities to anaesthetized animals (see Stein and Meredith 1993, for a very influential early overview). While the term was originally used to describe activity at the cellular level, researchers have, over the last couple of decades, increasingly been attempting to demonstrate similar effects at a behavioural level in both awake animals (e.g., Morgan et al. 2008; Populin and Yin 2002; Stein et al. 1989) and subsequently in awake people (Stein et al. 1996; Wallace et al. 2004; though see also Odgaard et al. 2003).

Controversy currently surrounds the most appropriate interpretation of certain of these multisensory integration effects (see Holmes 2008, 2009; Pouget 2006; Stein, Stanford, Ramachandran, Perrault and Rowland, this volume). Furthermore, there is also some discussion regarding the specific criteria that should be met across different techniques/paradigms in order to designate that a particular finding or result either does (or does not) constitute an example of superadditivity (e.g., see Beauchamp 2005; Laurienti et al. 2005). That said, the term ‘multisensory integration’ is currently generally used to describe those situations in which multiple typically near-simultaneous stimuli presented in different sensory modalities are bound together. Today, there is a large body of research using behavioural, electrophysiological and neuroimaging techniques, as well as data from patients and mathematical modelling approaches, describing the principles of multisensory integration in both animals and humans (see Driver and Noesselt 2008; Stein and Stanford 2008; see also Goebel and van Atteveldt; Stein, Stanford, Ramachandran, Perrault and Rowland, this volume, for reviews). While much of the progress in recent years has come from studying the behaviour of adult organisms, there is increasing interest in the development of multisensory integration in early infancy and childhood (e.g., Bremner et al. 2008; Gori et al. 2008; Neil et al. 2006; Tremblay et al. 2007; Wallace and Stein 2007), building on the early comparative approaches captured in Lewkowicz and Lickliter’s (1994) seminal edited volume.

In certain cases, it has even been unclear whether the behavioural effect that one is dealing with reflects a crossmodal effect or a multisensory integration effect. This, for example, is currently the case for crossmodal exogenous attentional cuing studies in which a cue stimulus is presented in one modality only slightly (say 50 or 100 ms) before the target stimulus in another modality (e.g., see Macaluso et al. 2001; McDonald et al. 2001; Spence et al. 2004). Should any behavioural facilitation of participants’ performance in response to the visual target seen in such studies be accounted for in terms of an exogenous shift of spatial attention towards the location of the auditory cue, or to the multisensory integration of the

cue and target stimulus into a single multisensory stimulus (cf. Schneider and Bavelier 2003)? At present, we simply do not know.

The special issue on crossmodal processing

This special issue grew out of the 9th Annual Meeting of the International Multisensory Research Forum (IMRF) held in Hamburg, Germany, on the 16–19 July 2008. Each year, this meeting brings together an ever-increasing number of researchers from all corners of the globe, united by their interest in the question of how the brain (of humans and other species) binds the information transduced by the senses. When the first meeting was held in Oxford in 1999, there were around 125 delegates; In Hamburg, last summer, there were nearly 300! Although many of the contributors to this special issue were themselves present in Hamburg, we operated an open submission policy. This meant that anyone working in the field of multisensory research was allowed to submit their work for possible inclusion in this special issue. We were delighted by the number of submissions that we received and we are very grateful to the many reviewers who offered both their time and expertise during the review process. We are also particularly happy to be able to include a number of papers dealing with multisensory interactions involving the vestibular system, given that this important area of research has not always been as well represented at previous IMRF meetings as one might have hoped. We have grouped the resulting 26 articles (2 review articles and 24 research articles) that finally made it into the pages of this special issue into seven sections.

The first section provides a short update on the *principle mechanisms* of multisensory processing. We start with an excellent, timely and challenging review article by Barry Stein and his colleagues highlighting recent developments and progress in the field of multisensory research. In part, the authors evaluate some of the potential criticisms associated with the notion of superadditivity that were raised by Nick Holmes in his much talked about presentation at the IMRF meeting in Hamburg last year (Holmes 2008; see also Holmes 2009). In their article, Royal and his colleagues use single-neuron recordings in cats to examine the role of the spatio-temporal receptive fields (RFs) of neurons in multisensory processing in the cortex. Interestingly, multisensory responses were characterized by two distinct temporal phases of enhanced integration reflecting shorter response latencies and longer discharge durations. In the third article, Shall and her colleagues study spectro-temporal activity in the human electroencephalogram (EEG) during the processing of audio-visual temporal congruency. They provide evidence that waveform locking may

constitute an important mechanism for multisensory processing.

Given the growing interest in using *functional magnetic resonance imaging (fMRI)* to understand the neural correlates of multisensory processes, recent advances in this field are covered in the second section. In their invited review article, Goebel and van Attefeldt summarize some of the latest developments in multisensory fMRI. The section also includes research articles by Noa and Amedi and by Stevenson and his colleagues: Noa and Amedi used the fMRI repetition-suppression approach to examine multisensory interactions in the visuotactile representation of objects, while Stevenson et al. used fMRI to explore the brain correlates of crossmodal interactions between streams of audio, visual and haptic stimuli.

Studies in the third section focus on the relationships between crossmodal processing and *perception*. The first article by Kawabe investigate crossmodal audio-visual interactions using a variant of the two-flash illusion (Shams et al. 2000; Watkins et al. 2007). In the second article, Chen and Yeh report on studies investigating the effects of the presentation of auditory stimuli on the phenomenon of repetition blindness in vision (see Kanwisher 1987), wherein observers fail to perceive the second occurrence of a repeated item in a rapid visual presentation stream. The next three articles examine the temporal aspects of multisensory perception: Barnett-Cowan and Harris provide some of the first evidence regarding multisensory timing involving vestibular stimulation. Meanwhile, Boenke and his colleagues investigate some of the factors, such as variations in stimulus duration, that influence the perception of temporal order for pairs of auditory and visual stimuli. In a very thorough series of empirical studies, Fujisaki and Nishida systematically compare synchrony perception for auditory, visual and tactile streams of stimuli, highlighting an intriguing peculiarity in adult humans with regard to the perception of the synchrony of auditory and tactile stimulus streams. Finally, in this section, Banissy and his colleagues describe the phenomenon of mirror-touch synaesthesia in which ‘the touch to another person induces a subjective tactile sensation on the synaesthete’s own body’ (see also Banissy and Ward 2007). Research in this area is increasingly starting to blur the boundary between synesthesia and ‘normal’ perception (see also Sagiv and Ward 2005; Parise and Spence 2009).

The fourth section of this special issue focuses on the *role of attention* in crossmodal processing. Using a cued forced-choice task, Hugenschmidt and her colleagues demonstrate that the effect of crossmodal selective attention in two multisensorially cued forced-choice tasks is preserved in older adults, and is comparable with the effects obtained in younger persons (see also Laurienti et al. 2006; Poliakoff et al. 2006, for earlier studies in this

area). In the second study in this section, Berger and Bühlhoff investigate the extent to which attending to one stimulus, while ignoring another, influences the integration of visual and inertial (vestibular, somatosensory and proprioceptive) stimuli. Juravle and Deubel examine the impact of action preparation on tactile attention using a novel crossmodal priming paradigm. They show that tactile perception is facilitated at the location of a goal-directed movement (saccade), as well as at the location of the effector of the movement (for simple finger lifting movements). Their results therefore provide evidence for a coupling between tactile attention and motor preparation (see also Gallace et al. *in press*). Finally, in a study using event-related potentials (ERPs), Durk Talsma and his colleagues demonstrate that attention to one or the other sensory stream (auditory or visual) can influence the processing of temporally asynchronous audiovisual stimuli. Together, the studies reported in this section therefore provide exciting new evidence concerning the extent to which attention modulates crossmodal information processing (see, Spence and Driver 2004).

The next section targets the role of *learning and memory* on crossmodal processing. Lacey and colleagues’ study shows that the learning of view-independence in visuo-haptic object representations results in the construction of bisensory, view-independent object representations, rather than of intermediate, unisensory, view-independent representations. In the next contribution, Petrini and her colleagues investigate the multisensory (audiovisual) processing of drumming actions in both jazz drummers and novices (see also Arrighi et al. 2006; Schutz and Kubovy *in press*). Their results suggest that through musical practice we learn to ignore variations in stimulus characteristics that would otherwise likely affect multisensory integration. It is worth noting in passing that the many articles on multisensory timing in this special issue capture a currently popular theme for multisensory researchers (especially for those presenting at the Hamburg meeting, where many papers and posters were on this topic). It would appear that there is something of a shift in the field from researchers focusing their attentions on the problems associated with *spatial* alignment and spatial representation across different coordinate frames (Röder et al. 2004; Spence and Driver 2004; Yamamoto and Kitazawa 2001) through to problems associated with multisensory *temporal* alignment (e.g., see King 2005; Spence and Squire 2003; Navarra et al. 2009).

By investigating the discrimination of motion for visual stimuli that had been paired with a relevant sound during training, Beer and Watanabe have been able to demonstrate that crossmodal learning is limited to those visual field locations that happened to overlap with the source of the sound; thus indicating that sounds can guide visual

plasticity at early stages of sensory information processing. In the final article in this section, Daniel Senkowski and his colleagues report enhanced gamma-band responses (activity > 30 Hz) for semantically matched as compared to non-matched audio-visual object pairings; thus providing evidence for the view that the dynamic coupling of neural populations may be a crucial mechanism underlying crossmodal semantic processing (see also Senkowski et al. 2008).

The sixth section comprises studies using *motion stimuli* to examine crossmodal processing. Investigating the question of whether biological motion processing mechanisms contribute to audio-visual binding, van der Zwan and his colleagues show that auditory walking sequences containing matching gender information result in a facilitation of visual judgments of the gender of point-light walkers. By assessing the mismatch negativity in ERPs, Stekelenburg and Vroomen provide evidence that auditory and visual motion signals are integrated during early sensory information processing. Zvyagintsev and his colleagues used the dipole modelling of brain activity measured using magnetoencephalography (MEG) recordings to examine which neural systems trigger the perception of motion for audio-visual stimuli. These authors highlight a modulation of MEG activity, localized to primary auditory cortex; thus supporting the notion that audio-visual motion signals are integrated at the earliest stages of sensory processing (though see Kayser and Logothetis 2007, for a critical discussion of this issue).

The final section of this special issue comprises studies relating to the question of how *eye position* and *saccades* influence information processing crossmodally. In the first article, Harrar and Harris highlight a systematic shift in the perceived localization of tactile stimuli towards the location where participants happen to be fixating visually (see also Ho and Spence 2007). In a similar vein, Klingehoefer and Bremmer report that the perceived location of brief noise bursts that are presented before, during and after visually guided saccades are biased by the eye movements; thus, suggesting that sensory signals are represented in some form of crossmodal spatial representation. van Wanrooij and his colleagues examine the dynamics of saccades towards visual targets, while participants were instructed to ignore auditory distractors that were presented with various spatial and temporal disparities. They provide convincing evidence that both spatial alignment and timing influence the impact of auditory distractors on saccadic eye movements.

Taken together, the 26 articles that comprise this special issue highlight the range of both techniques and paradigms currently being brought to bear on questions of multisensory integration and crossmodal information processing. The results of the research highlighted here confirm the

view that similar mechanisms of multisensory integration appear to operate across different pairings of sensory modalities (though see Fujisaki and Nishida, this volume, for one intriguing exception in the temporal domain). What is more, the various techniques at the disposal of cognitive neuroscientists (from fMRI, MEG and EEG to single-cell electrophysiology, and from mathematical modelling to psychophysics) are increasingly showing just how early in information processing these crossmodal and multisensory effects can occur. Indeed, recent results have led some to question whether the whole brain might not in some sense better be considered as being multisensory (see Foxe and Schroeder 2005; Ghazanfar and Schroeder 2006; see also Pascual-Leone and Hamilton 2001).

In closing, we would like to express our sincere appreciation to the companies and institutions that supported the 9th Annual Meeting of the IMRF in Hamburg (listed in alphabetical order): Brain Products GmbH; CINACS graduate school; DIAEGO; Easycap GmbH; EUCognition; MES; Symrise; Unilever; and The University Hamburg.

References

- Arrighi R, Alais D, Burr D (2006) Perceptual synchrony of audiovisual streams for natural and artificial motion sequences. *J Vis* 6:260–268
- Banissy MJ, Ward J (2007) Mirror-touch synesthesia is linked with empathy. *Nat Neurosci* 10:815–816
- Beauchamp MS (2005) Statistical criteria in FMRI studies of multisensory integration. *Neuroinformatics* 3:93–113
- Bremmer AJ, Mareschal D, Lloyd-Fox S, Spence C (2008) Spatial localization of touch in the first year of life: early influence of a visual code and the development of remapping across changes in limb position. *J Exp Psychol Gen* 137:149–162
- Calvert GA, Spence C, Stein BE (eds) (2004) *The handbook of multisensory processes*. MIT Press, Cambridge
- Dodd B, Campbell R (eds) (1987) *Hearing by eye: the psychology of lip-reading*. LEA, Hillsdale
- Driver J, Noesselt T (2008) Multisensory interplay reveals crossmodal influences on ‘sensory-specific’ brain regions, neural responses, and judgments. *Neuron* 57:11–23
- Ferris TK, Sarter NB (2008) Cross-modal links among vision, audition, and touch in complex environments. *Hum Factors* 50:17–26
- Foxe JJ, Schroeder CE (2005) The case for feedforward multisensory convergence during early cortical processing. *Neuroreport* 16:419–423
- Gallace A, Zeeden S, Röder B, & Spence C (in press) Lost in the move? Secondary task performance impairs the detection of tactile change on the body surface. *Conscious Cogn*
- Ghazanfar AA, Schroeder CE (2006) Is neocortex essentially multisensory? *Trends Cogn Sci* 10:278–285
- Gori M, Del Viva M, Sandini G, Burr DC (2008) Young children do not integrate visual and haptic information. *Curr Biol* 18:694–698
- Ho C, Spence C (2007) Head orientation biases tactile localization. *Brain Res* 1144C:136–141

- Ho C, Spence C (2008) The multisensory driver: implications for ergonomic car interface design. Ashgate Publishing, Aldershot
- Holmes NP (2008) The seemingly inviolable principle of inverse effectiveness: in search of a null hypothesis. Paper presented at the 9th annual meeting of the international multisensory research form, Hamburg, Germany, July 16–19th
- Holmes NP (2009) The principle of inverse effectiveness in multisensory integration: some statistical considerations. *Brain Topogr* 21:168–176
- Kanwisher NG (1987) Repetition blindness: type recognition without token individuation. *Cognition* 27:117–143
- Kayser C, Logothetis NK (2007) Do early sensory cortices integrated cross-modal information? *Brain Struct Funct* 1173:102–109
- King AJ (2005) Multisensory integration: strategies for synchronization. *Curr Biol* 15:R339–R341
- Laurienti PJ, Perrault TJ, Stanford TR, Wallace MT, Stein BE (2005) On the use of superadditivity as a metric for characterizing multisensory integration in functional neuroimaging studies. *Exp Brain Res* 166:289–297
- Laurienti PJ, Burdette JH, Maldjian JA, Wallace MT (2006) Enhanced multisensory integration in older adults. *Neurobiol Aging* 27:1155–1163
- Lewkowicz DJ, Lickliter R (eds) (1994) The development of intersensory perception: comparative perspectives. LEA, Hillsdale
- Macaluso E, Frith CD, Driver J (2001) A reply to J. J. McDonald, W. A. Teder-Sälejärvi, & L. M. Ward, Multisensory integration and crossmodal attention effects in the human brain. *Science* 292:1791
- McDonald JJ, Teder-Sälejärvi WA, Hillyard SA (2000) Involuntary orienting to sound improves visual perception. *Nature* 407:906–908
- McDonald JJ, Teder-Sälejärvi WA, Ward LM (2001) Multisensory integration and crossmodal attention effects in the human brain. *Science* 292:1791
- Merabet LB, Rizzo JF, Amedi A, Somers DC, Pascual-Leone A (2005) Opinion: what blindness can tell us about seeing again: merging neuroplasticity and neuroprostheses. *Nat Rev Neurosci* 6:71–77
- Morgan ML, DeAngelis GC, Angelaki DE (2008) Multisensory integration in macaque visual cortex depends on cue reliability. *Neuron* 59:662–673
- Navarra J, Hartcher-O'Brien J, Piazza E, Spence C (2009) Adaptation to audiovisual asynchrony modulates the speeded detection of sound. *Proc Natl Acad Sci USA* 106:9169–9173
- Neil PA, Chee-Ruiter C, Scheier C, Lewkowicz DJ, Shimojo S (2006) Development of multisensory spatial integration and perception in humans. *Dev Psychol* 9:454–464
- Odgaard EC, Arieh Y, Marks LE (2003) Cross-modal enhancement of perceived brightness: sensory interaction versus response bias. *Percept Psychophys* 65:123–132
- Parise C, Spence C (2009) 'When birds of a feather flock together': synesthetic correspondences modulate audiovisual integration in non-synesthetes. *PLoS ONE* 4(5):e5664. doi:10.1371/journal.pone.0005664
- Pascual-Leone A, Hamilton R (2001) The metamodal organization of the brain. In: Casanova C, Ptito M (eds) *Progress in brain research*, vol 134. Elsevier, Amsterdam, pp 427–445
- Poliakoff E, Ashworth S, Lowe C, Spence C (2006) Vision and touch in ageing: crossmodal selective attention and visuotactile spatial interactions. *Neuropsychologia* 44:507–517
- Populin LC, Yin TCT (2002) Bimodal interactions in the superior colliculus of the behaving cat. *J Neurosci* 22:2826–2834
- Pouget A (2006) Neural basis of Bayes-optimal multisensory integration: theory and experiments. Paper presented at the 7th annual meeting of the international multisensory research forum. Trinity College, Dublin, 18–21 June
- Röder B, Rösler F (2004) Compensatory plasticity as a consequence of sensory loss. In: Calvert GA, Spence C, Stein BE (eds) *The handbook of multisensory processes*. MIT Press, Cambridge, pp 719–747
- Röder B, Rösler F, Spence C (2004) Early vision impairs tactile perception in the blind. *Curr Biol* 14:121–124
- Rouger J, Lagleyre S, Fraysse B, Deneve S, Deguine O, Barone P (2007) Evidence that cochlear-implanted deaf patients are better multisensory integrators. *Proc Natl Acad Sci USA* 104:7295–7300
- Sagiv N, Ward J (2005) Cross-modal interactions: lessons from synaesthesia. *Prog Brain Res* 155:263–275
- Schneider KA, Bavelier D (2003) Components of visual prior entry. *Cogn Psychol* 47:333–366
- Schutz M, Kubovy M (in press) Causality in audio-visual sensory integration. *J Exp Psychol: Human Percept Perform*
- Senkowski D, Schneider TR, Foxe JJ, Engel AK (2008) Crossmodal binding through neural coherence: implications for multisensory processing. *Trends Neurosci* 31:401–409
- Shams L, Kamitani Y, Shimojo S (2000) What you see is what you hear: sound induced visual flashing. *Nature* 408:788
- Spence C (2002) *The ICI report on the secret of the senses*. The Communication Group, London
- Spence C, Driver J (1997) Audiovisual links in exogenous covert spatial orienting. *Percept Psychophys* 59:1–22
- Spence C, Driver J (eds) (2004) *Crossmodal space and crossmodal attention*. Oxford University Press, Oxford
- Spence C, Squire SB (2003) Multisensory integration: maintaining the perception of synchrony. *Curr Biol* 13:R519–R521
- Spence C, McDonald J, Driver J (2004) Exogenous spatial cuing studies of human crossmodal attention and multisensory integration. In: Spence C, Driver J (eds) *Crossmodal space and crossmodal attention*. Oxford University Press, Oxford, pp 277–320
- Stein BE, Meredith MA (1993) *The merging of the senses*. MIT Press, Cambridge
- Stein BE, Stanford TR (2008) Multisensory integration: current issues from the perspective of the single neuron. *Nat Rev Neurosci* 9:255–267
- Stein BE, Meredith MA, Honeycutt WS, McDade L (1989) Behavioral indices of multisensory integration: orientation to visual cues is affected by auditory stimuli. *J Cogn Neurosci* 1:12–24
- Stein BE, London N, Wilkinson LK, Price DP (1996) Enhancement of perceived visual intensity by auditory stimuli: a psychophysical analysis. *J Cogn Neurosci* 8:497–506
- Tremblay C, Champoux F, Voss P, Bacon BA, Lepore F, Théoret H (2007) Speech and non-speech audio-visual illusions: a developmental study. *PLoS ONE* 8:e742
- Wallace MT, Stein BE (2007) Early experience determines how the senses will interact. *J Neurosci* 97:921–926
- Wallace MT, Roberson GE, Hairston WD, Stein BE, Vaughan JW, Schirillo JA (2004) Unifying multisensory signals across time and space. *Exp Brain Res* 158:252–258
- Watkins S, Shams L, Josephs O, Rees G (2007) Activity in human V1 follows multisensory perception. *NeuroImage* 37:572–578
- Yamamoto S, Kitazawa S (2001) Reversal of subjective temporal order due to arm crossing. *Nat Neurosci* 4:759–765