

Multisensory integration: from fundamental principles to translational research

Georg F. Meyer · Uta Noppeney

Published online: 29 July 2011
© Springer-Verlag 2011

We perceive our natural environment via multiple senses. How does our mind and brain integrate these diverse sensory inputs into a coherent and unified percept? This challenging and exciting question has been the focus of a growing multidisciplinary community of researchers that meet regularly at the annual meeting of the International Multisensory Research Forum (IMRF). The IMRF meeting brings together researchers that investigate multisensory integration at multiple levels of description ranging from neurophysiology to behaviour, with research interests from theory to applications. Traditionally, multisensory research has focused on the characterization of the fundamental principles and constraints that govern multisensory integration. Research has then moved forward to address questions of how multisensory integration emerges during development and may be perturbed in cases of disease or ageing. In the long-term, multisensory research will have direct impact in translational studies investigating the benefits of a multisensory environment for patients that are impaired when presented with information in one sensory modality alone. Obviously, this myriad of research topics can only be addressed by combining findings from a multitude of methods including psychophysics, neurophysiology and non-invasive structural and functional imaging in humans. Further, since its inception multisensory research has constantly gained impetus from

computational models. Computational models contribute substantially to the progress made in multisensory research by providing a deeper understanding of the current empirical findings and conversely making predictions that guide future research. Most prominently, the normative Bayesian perspective continues to inspire inquiries into the optimality of multisensory integration across various species.

This special issue on multisensory processing has resulted from the IMRF meeting held at Liverpool University in 2010. In accordance with previous procedures, the call for papers was not restricted to meeting attendees but was open to the entire multisensory community. As has been the tradition since the first special IMRF issue, we received a large number of high quality submissions leading to strong competition. Many excellent submissions had to be rejected or transferred to other special issues because of space limitations. Nevertheless, we hope that the collection of manuscripts included in this special issue will provide a rich source of reference for the wider multisensory community.

Given the multidisciplinary of the IMRF community, the submitted manuscripts cover a range of the topics that have briefly been highlighted above. For coarse reference, we have grouped the manuscripts into five sections:

Principles and fundamental constraints of multisensory integration

A fundamental aim in multisensory research is to delineate the constraints that enable integration of sensory signals. Sensory inputs should be combined when they come from a common source, but segregated when they come from different sources. Spatiotemporal and semantic congruence

G. F. Meyer (✉)
Department of Experimental Psychology, Liverpool University,
Liverpool L69 7ZA, UK
e-mail: georg@liv.ac.uk

U. Noppeney
Max Planck Institute for Biological Cybernetics,
Spemannstr. 41, 72076 Tuebingen, Germany
e-mail: uta.noppeney@tuebingen.mpg.de

have long been identified as important factors determining the emergence of multisensory integration. Fiebelkorn et al. have investigated the role of spatial concordance for auditory facilitation of visual-target detection. They demonstrate that auditory facilitation of visual-target detection persists regardless of retinal eccentricity and despite wide audiovisual misalignments (Fiebelkorn et al. 2011, this issue). These results suggest that multisensory facilitation may at least in part be mediated through a spatially non-specific mechanism.

Rather than focusing on multisensory constraints, Alsius and Soto-Faraco investigated processing constraints that are already present in individual sensory modalities per se and may in turn shape multisensory binding. Using visual and auditory search tasks for speaking faces, they demonstrate that search efficiency declines with the number of signals for faces, but not for auditory streams. The authors argue that a key difference between visual and auditory modalities is that spatial information is obligatory for visual tasks, while auditory matching can be based on temporal information. (Alsius and Soto Faraco 2011, this issue).

A key question is obviously: what are the benefits of multisensory integration? Within a maximum likelihood estimation framework, it is beneficial to integrate information about a particular property from multiple senses in order to increase estimation efficiency. In other words, multisensory integration helps to reduce an estimator's variance. Indeed, many previous psychophysics studies have demonstrated that humans are often near optimal when integrating information from multiple senses (Ernst and Banks 2002). Mendonça and colleagues (2011, this issue) have expanded this body of research by demonstrating that humans are near optimal when integrating the velocity of audiovisual congruent biological motion stimuli.

Leo et al. take a different approach and investigate whether an auditory looming or receding stimulus can improve subjects' performance in a visual discrimination task. Importantly, here, the auditory stimulus does not provide any relevant information about the property judged in the visual discrimination task. Nevertheless, they demonstrate that uninformative looming sounds improve subjects' visual discrimination performance, when auditory and visual inputs were spatially congruent (Leo et al. 2011, this issue).

Temporal congruency of visual and auditory signals is used as a marker in an EEG study by Proctor and Meyer, who investigated whether visual face processing is influenced by the temporal congruency of the audiovisual signals. Their ERP study reveals early effects of audiovisual congruency at around 135 ms. Yet, the later N170 component generally associated with configurational face

processing is unaffected by audiovisual temporal congruency pointing to an independence of this processing stage from auditory inputs (Proctor and Meyer 2011, this issue). In contrast, a role of synchrony and spatial concordance in face perception emerged in the study by Mazurega et al. focusing on the enfacement effect (Mazurega et al. 2011, this issue). Observers judged a stranger's face as more similar to their own in terms of physical and personality features, if visual and tactile stimuli were applied concurrently to the observed and observer's face in a spatiotemporally congruent fashion.

Space and body

A series of papers focused on interactions between space and body representations. One important question is how space is represented and processed in different sensory modalities. Maij et al. point to processing similarities across vision and touch by showing that movement-related mislocalization of haptic representations is analogous to the systematic errors made during saccades in visual perception (Maij et al. 2011, this issue). Yet, sensory modalities can represent space in different reference systems raising the question of how spatial representations interact and are transformed between different sensory modalities (e.g. eye vs. head-centred). Pritchett and Harris have examined the influence of head and eye position on the perceived location of touch. They demonstrate that subjects' perceived tactile position is shifted by eye and head position indicating influences of a visual reference frame on tactile location (Pritchett and Harris 2011, this issue).

Along similar lines, Van Barneveld et al. studied the influence of head roll on the perceived auditory location in head-centred and world-centred coordinates. When judging spatial location in a world-centred reference system, subjects made systematic errors in the direction of head roll suggesting an influence of the vestibular system on auditory spatial location (Van Barneveld et al. 2011, this issue). Schomaker and colleagues (2011, this issue) supplemented this body of research with a learning perspective. They examined how spatial mapping across sensory modalities can be changed using visual-vestibular motor recalibration. Subjects had to maintain an arm position in space while their whole body was rotated. Importantly, they received incorrect visual information about the rotation degree from three views: first-person, top view or mirror view. Their intriguing results demonstrate that the naturalistic first person view is most powerful for inducing visual-vestibular recalibration.

While recalibration emerges at relatively short time-scales, Jola et al. investigated the effect of long-term training (i.e. dancer's expertise) on the precision of

endpoint position judgments based on visual, proprioceptive or both sorts of information. They show that trained dancers show better integration of local proprioceptive and visual signals. Further, dancers rely more on proprioceptive signals than untrained participants when judging hand position (Jola et al. 2011, this issue).

Multisensory integration may not only vary across individuals because of training or expertise but also because of subjects' strategies. Lacey et al. showed that object and spatial imagery alter the dominance of texture and space in object recognition jointly for both vision and touch (Lacey et al. 2011, this issue).

Effects of personality traits, disease and ageing on multisensory integration

One exciting new avenue of research characterizes multisensory integration in populations with particular personality traits or diseases. Studying altered multisensory integration in diseases may help us to better understand the underlying mechanisms by which diseases affect higher cognitive functions. Further, multisensory integration performance may also be used as an additional diagnostic marker in early stages of diseases. Koizumi and coworkers investigated the effects of anxiety on the interpretation of emotion in the face–voice pairs (Koizumi et al. 2011, this issue). Consistent with previous results using unisensory emotional Stroop tasks (Gotlib and McCann 1984), subjects with high trait anxiety were more prone to interference from task-irrelevant negative emotional facial cues when interpreting the emotion of the task-relevant voices and vice versa.

Previous studies have already suggested that multisensory integration may be altered in individuals with autism spectrum disorders. In line with this, Saalasti et al. (2011, this issue) showed systematic differences in audiovisual speech perception between individuals with Asperger syndrome and a control group. When an auditory /apa/ was paired with a visual /aka/, individuals with Asperger syndrome were more likely to perceive /ata/ than the control subjects. These changes in audiovisual integration may contribute to and explain the difficulties of individuals with Asperger syndrome in face-to-face communication.

Multisensory integration may not only be used as a diagnostic marker, but also help individuals that show impairments when exposed to unisensory stimuli alone. This is demonstrated in the study by Elliott and colleagues investigating the effect of ageing on multisensory integration for the control of movement timing (Elliott et al. 2011, this issue). They showed that older adults' performance when synchronizing actions to auditory and tactile temporal cues was degraded to a greater extent than for young

adults in both unimodal and bimodal conditions. Yet, since the elderly subjects benefitted similarly from integrating multisensory temporal cues, multisensory information may be used to mitigate behavioural deficits observed under unisensory conditions in action coordination to complex timing cues.

Collectively, these novel lines of research have the potential to take multisensory research from basic to translational research. Multisensory integration may then have a real impact on diagnostics, therapy and engineering developments.

Methodological approaches to characterize the neural basis of multisensory integration

Three contributions applied and further developed functional (fMRI & M/EEG) and structural imaging techniques to characterize the neural basis of multisensory integration.

Beer et al. (2011, this issue) used diffusion tensor imaging to show direct white matter connections between auditory and visual cortex that may serve as one pathway mediating audiovisual integration. Naumer et al. (2011, this issue) applied data-driven group spatial independent component analysis (ICA) to a functional magnetic resonance imaging (fMRI) data set acquired during a passive audiovisual (AV) object observation experiment and a recognition task. The independent component maps proved useful for identification of candidate regions of interest that can be characterized further by more specific statistical analyses.

Keitel and coworkers (2011, this issue) demonstrate that intermodal attention modulates the auditory and visual steady state responses using amplitude-modulated multi-speech babble and a stream of nonsense letter sets. Their results suggest that frequency tagging of responses in complex and more natural environments is a potentially interesting paradigm to study of intermodal attention and multisensory integration.

Computational models of multisensory integration

As often stated, empirical data are blind without appropriate theoretical frameworks. Two theoretical contributions complemented the excellent experimental work in psychophysics and neuroimaging described so far. Lim et al. (2011, this issue) used a computational network of spiking neurons to provide insight into the role of connectional parameters in shaping the multisensory properties of neurons in convergence areas. Changes in extrinsic, intrinsic and local inhibitory contacts influenced the proportion of neurons that generate integrated multisensory responses as well as the magnitude of integration. Cuppini

and colleagues (2011, this issue) simulated the changes that occur during neurodevelopment in multisensory integration and receptive field sizes of neurons by training a network with modality-specific and multisensory signals. Adjusting the synaptic weights according to Hebbian learning rules of potentiation and depression induced a sharpening of the neuron's receptive field size and crossmodal response properties as observed during maturation.

Coda

This special issue clearly demonstrates the broad scope of the research currently performed in multisensory integration. Multisensory research continues to thrive through synergistic interactions between computational models, basic research and translational studies. The recent surge of translational research demonstrates the potential impact that multisensory research could exert when combined with clinical work or robotics. Conversely, current experimental research pursues two complementary perspectives and research goals: The first stream employs simple well-controlled stimuli to characterize fundamental principles of multisensory integration. The second stream investigates how multisensory integration emerges in an ecologically valid context and hence uses complex, naturalistic stimuli such as biological motion or speech signals. These less-controlled environmentally valid studies will benefit from recent developments in data-driven analysis approaches.

We believe that progress in multisensory integration evolves from exactly those tensions and from communication across different research fields and approaches.

We look forward to exciting new developments and results that will be presented at the next IMRF meeting that will take place from the 17th to 20th October 2011 in Fukuoka, Japan.

Before closing, we would like to thank all the reviewers who provided critical and constructive comments on the submissions within a short timeframe, and all authors who submitted papers to this edition. Special thanks are due to Patrick Haggard who acted as supervising editor and the team at Springer for making the editing of this issue possible.

References

- Alsius A, Soto-Faraco S (2011) Searching for audiovisual correspondence in multiple speaker scenarios (this issue). doi:10.1007/s00221-011-2624-0
- Beer AL, Plank T, Greenlee MW (2011) Diffusion tensor imaging shows white matter tracts between human auditory and visual cortex (this issue). doi:10.1007/s00221-011-2715-y
- Cuppini C, Stein BE, Rowland BA, Magosso E, Ursino M (2011) A computational study of multi-sensory maturation in the superior colliculus (SC) (this issue). doi:10.1007/s00221-011-2714-zA
- Elliott MT, Wing AM, Welchman AE (2011) Effect of ageing on multisensory integration for the control of movement timing (this issue). doi:10.1007/s00221-011-2740-x
- Ernst MM, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415:429–433
- Fiebelkorn IC, Foxe JJ, Butler JS, Molholm S (2011) Auditory facilitation of visual-target detection persists regardless of retinal eccentricity and despite wide audiovisual misalignments (this issue). doi:10.1007/s00221-011-2670-7
- Gotlib IH, McCann CD (1984) Construct accessibility and depression: an examination of cognitive and affective factors. *J Pers Soc Psychol* 47(2):427–439
- Jola C, Davis A, Haggard P (2011) Proprioceptive integration and body representation: insights into dancers' expertise (this issue). doi:10.1007/s00221-011-2743-7
- Keitel C, Schröger E, Saupe K, Müller MM (2011) Sustained selective intermodal attention modulates processing of language-like stimuli (this issue). doi:10.1007/s00221-011-2667-2
- Koizumi A, Tanaka A, Imai H, Hiramatsu S, Hiramoto E, Sato T, de Gelder B (2011) Effects of anxiety on the interpretation of emotion in the face–voice pairs (this issue). doi:10.1007/s00221-011-2668-1
- Lacey S, Lin JB, Sathian K (2011) Object and spatial imagery dimensions in visuo-haptic representations (this issue). doi:10.1007/s00221-011-2623-1
- Leo F, Romei V, Freeman E, Ladavas E, Driver J (2011) Looming sounds enhance orientation sensitivity for visual stimuli on the same side as such sounds (this issue). doi:10.1007/s00221-011-2742-8
- Lim HK, Keniston LP, Shin JH, Allman BL, Meredith MA, Cios KJ (2011) Connectional parameters determine multisensory processing in a spiking network model of multisensory convergence (this issue). doi:10.1007/s00221-011-2671-6
- Maij F, de Grave DDJ, Brenner E, Smeets JBJ (2011) Misjudging where you felt a light switch in a dark room (this issue). doi:10.1007/s00221-011-2680-5
- Mazzurega M, Pavani F, Paladino MP, Schubert TW (2011) Self-other bodily merging in the context of synchronous but arbitrary-related multisensory inputs (this issue). doi:10.1007/s00221-011-2744-6
- Mendonça C, Santos JA, López-Moliner J (2011) The benefit of multisensory integration with biological motion signals (this issue). doi:10.1007/s00221-011-2620-4
- Naumer MJ, van den Bosch JFF, Wibral M, Kohler A, Singer W, Kaiser J, van de Ven V, Muckli L (2011) Investigating human audio-visual object perception with a combination of hypothesis-generating and hypothesis-testing fMRI analysis tools (this issue). doi:10.1007/s00221-011-2669-0
- Pritchett LM, Harris LR (2011) Perceived touch location is coded using a gaze signal (this issue). doi:10.1007/s00221-011-2713-0
- Proctor BJ, Meyer GF (2011) Electrophysiological correlates of facial configuration and audio–visual congruency: evidence that face processing is a visual rather than a multisensory task (this issue). doi:10.1007/s00221-011-2724-x
- Saalasti S, Tiippana K, Kätsyri J, Sams M (2011) The effect of visual spatial attention on audiovisual speech perception in adults with Asperger syndrome (this issue). doi:10.1007/s00221-011-2751-7
- Schomaker J, Tesch J, Bühlhoff HH, Bresciani J-P (2011) It is all me: the effect of viewpoint on visual–vestibular recalibration (this issue). doi:10.1007/s00221-011-2723-y
- Van Barneveld DCPBM, Van Grootel TJ, Alberts B, Van Opstal AJ (2011) The effect of head roll on perceived auditory zenith (this issue). doi:10.1007/s00221-011-2741-9