Neuroscientific evidence for multisensory convergence and interaction

Emi Tanaka^{1,2}, Tetsuo Kida^{3,4}, Ryusuke Kakigi^{3,4} and Minoru Hoshiyama¹

¹Brain and Mind Research Center, Nagoya University, 1-1-20 Daiko-Minami, Higashi-ku, Nagoya, Aichi 461-8673, Japan

²Japan Society for the Promotion of Science, 5-3-1 Koji-Machi, Chiyoda, Tokyo 102-0083, Japan

³Department of Integrative Physiology, National Institute Physiological Sciences, 38 Nishigonaka, Myodaiji, Okazaki, Aichi 444-8585, Japan

⁴Department of Physiological Sciences, School of Life Sciences, Graduate University for Advanced Studies (SOKENDAI), Hayama, Miura, Kanagawa 240-0193, Japan

Received: June 28, 2017 / Accepted: July 31, 2017

Abstract Effective behaviors optimized for various situations are enabled by various sensory information. How does the brain deal simultaneously with information from different sensory systems? Investigation of multisensory processing began from neuroanatomical and neurophysiological studies at the level of single neurons. Accumulation of knowledge concerning neurons and progress in recording techniques of human brain activity have led to a massive expansion in recent neuroscientific studies using various combinations of brain activity recording and cognitive tasks. In this paper, we briefly review recent neuroscientific studies related to multisensory convergence and interaction, focusing on electrophysiological and imaging studies in humans.

Keywords : multisensory integration, human neuroimaging, brain, change detection, attention

Introduction

In various daily situations, using information from different sensory systems comprehensively enables us to perform appropriate behaviors. If the brain is likened to a big company, each sensory system represents a department with a different role. In a productive big company, each department plays their role and communicates essential information interactively and appropriately. Good communication will lead to appropriate judgments by the manager and an effective and efficient company. In addition, key sections or people may overlap and work across different departments. Similarly, our brain efficiently processes simultaneous information from different sensory systems to effectively control behavior. Previous studies targeting brain functions in physical fitness and sports science have examined the brain and spinal activities as well as the relevant perceptual, cognitive, and motor aspects at the behavior level, mainly in individual sensory modalities, and have assessed these activities separately in different sensory modalities. However, our daily life, including sports activities, takes part in a multisensory environment. Thanks to dramatic progress in scientific techniques and accumulated findings during this decade, we can now non-invasively assess human brain activity and behavioral metrics in response to multisensory or unisensory stimulation. Thus, a paradigm shift from uni-

sensory to multisensory experiments has taken place in the study of physical fitness and sports activity. Here we review neuroscientific research on multisensory processing including animal anatomy and physiology studies, and human neuroimaging.

Neuroscientific research on multisensory processing began to increase from the late 1970s with single neuron recording studies¹⁻⁴, which consolidated the principles of multisensory processing⁵⁻¹¹. Event-related potentials (ERPs), used to investigate human neural processing and cognitive mechanisms for over half a century, have been applied to the study of multisensory processing since the end of the 1990s. Due to the emergence and continuing development of functional neuroimaging techniques during this time, investigation of the relationship between multisensory neural activities and their behavioral, perceptual, and cognitive aspects in humans have continued to increase. In this paper we review human neuroscientific studies concerning multisensory processing, and discuss multisensory “convergence” and “interaction”, which play a role in the integration of multisensory information. These have been critical concepts since the early period of the study of multisensory integration in experimental animals, right up to the present, where neuroimaging studies of multisensory integration have dramatically advanced. We will describe what convergence and interaction are, the difference between them, the difference in each concept between animal physiology and human neuroimaging, pros and cons of the use of these two concepts, and

*Correspondence: tanaka.emi@i.mbox.nagoya-u.ac.jp

the new generation of scientific strategies in neuroimaging studies of multisensory integration.

Convergence in animal anatomical and physiological studies

Traditionally, the areas of the brain suggested to process multimodal inputs are the subcortical areas and higher multisensory cortical areas. This is largely based on data from non-human primates. At the cortical level in monkeys, there is anatomical evidence that multisensory sites include areas within the superior temporal sulcus (STS), intraparietal lobe (IPL), parieto-preoccipital cortex, posterior insula, and frontal regions such as the premotor, prefrontal, and anterior cingulate cortices^{2-4,12}.

Electrophysiological studies in primates have shown that these anatomically defined multisensory cortical regions have neurons responsive to stimulation of different sensory modalities⁶⁻¹¹. Also in the subcortical region, the superior colliculus (SC) is well known as a multisensory area that has multisensory neurons responsive to visual, auditory, and tactile stimuli^{5,6,8}. A functional link between multisensory integration in individual SC neurons and cross-modality attentive and orientation behaviors has been demonstrated using experiments in alert and anesthetized cats⁷. Thus, anatomical and electrophysiological studies have demonstrated single-neuronal multisensory convergence. Furthermore, a suppressive interaction exists in SC neurons¹³ and the higher-order somatosensory area IV¹⁴. Indeed, the response to the simultaneous presentation of different sensory stimuli is increased or decreased, compared to a linear sum of unisensory responses to separate presentations of the stimuli, providing evidence for multisensory integration. This non-linear interactive effect between stimuli from different sensory modalities has been extensively looked at in human neuroimaging studies and will be described later in more detail.

Basic evidence for multisensory integration in humans

In humans, population-level methods are commonly used for non-invasively recording neuronal activity and related signals, because it is impossible to non-invasively record single-neuronal activity in healthy subjects (although one patient study reported multisensory properties of single neurons in the medial temporal lobe). These techniques include functional magnetic resonance imaging (fMRI), positron emission tomography (PET), electroencephalography (EEG), and magnetoencephalography (MEG). Using these neuroimaging methods, voxel-wise or regional multisensory convergence has been extensively examined during the last decade. First, this type of multisensory convergence has been evidenced by the finding that sensory stimuli from different sensory modalities activate a common brain region or voxel. For example, in their fMRI study, Downar et al. found that

both unimodal and multimodal activation were elicited by an abrupt change of a unimodal stimulus during continuous multisensory stimulation. The multimodal responses occurred in the temporo-parietal junction (TPJ), middle temporal gyrus (MTG), inferior frontal gyrus (IFG), and insula¹⁵. They suggested that the multimodal activations in response to unimodal changes during continuous multimodal stimulation reflect the detection of changes in the sensory environment. Their experimental procedure enabled examination of the response to a common factor related to *change* elicited by different unimodal changes¹⁵.

The fMRI signal is a hemodynamic signal and therefore lacks temporal information, whereas EEG and MEG non-invasively record neuronal activity with high temporal resolution in the order of milliseconds. Concerning temporal aspects of multisensory convergence in humans, we previously recorded EEG responses to four different stimuli; visual, auditory, tactile, and pain. We found multisensory responses in the anterior cingulate cortex (ACC) and hippocampus, which are suggested to be regions important for high-level multisensory convergence, to all visual, auditory, tactile and pain stimuli. These responses were preceded at the same time interval by unisensory responses in modality-specific areas, including the middle occipital gyrus (MOG), superior temporal gyrus (STG), and second somatosensory cortex (S2), after each sensory stimulation (Fig. 1)¹⁶. An important finding of this study was the multisensory and unisensory responses to manipulation of the interstimulus interval, which was an experimental factor common to all sensory modalities tested here (vision, audition, touch, and pain). Another MEG study from our laboratory also showed the temporal dynamics of *change*-driven cortical responses in multisensory areas, such as the TPJ and middle and inferior frontal gyri (MFG, IFG), regardless of the sensory modalities where the change occurred. These multisensory activations were preceded by unisensory activations, both of which, in general, had peaks within 300 ms after the changes (Fig. 2). This result indicates that neural processes responsive to unimodal changes in the multisensory environment are distributed at different timings in these cortical areas¹⁷. In addition, another MEG study from our laboratory found similar cue-invariant activation of the TPJ, common to the onset, offset, and change in stimulation using faces, which was preceded by unisensory activation in the MOG and fusiform gyrus (Fig. 3)¹⁸. Considering this result and our previous multisensory MEG study¹⁷, this TPJ response appears to be related to higher-order processing irrespective of changes to the mean luminance and the sensory modality, and is evidence of regional multisensory convergence. A recent combined study of fMRI and DTI tractography in humans identified white matter tracts connecting the regions of modality-specific functional activation with the angular gyrus, a subregion of the TPJ¹⁹. This study provides a human anatomofunctional basis for multisensory con-

Fig. 1 Among several modality-specific activities, only one or two main sources are included in this figure to clearly show the similar sequence of activations among all sensory modalities. Multi-dipole analysis using brain electric source analysis (BESA) was applied to separate temporally overlapping multiple sources. A, temporal profiles of source strength; B, locations of the sources in the cingulate cortex and hippocampal region overlaid on slices of the standard brain; C, schematic presentation of the common sequential activations through 'early' sensory areas, 'late' sensory areas, the ACC and the hippocampal region; D, time course of each cortical activity in the 0.1 Hz interstimulus interval condition of individual data. Waveforms of each subject (gray lines) and their average (black line) are superimposed. Note the similar temporal profile of each cortical activity between the grand-averaged data (shown in A) and individual data. Rt, right; Lt, left; STG, superior temporal gyrus; ACC, anterior cingulate cortex; Hip, hippocampus (Tanaka et al., BMC Neurosci. 2008).

Fig. 2 Grand-averaged waveforms of each source activity across subjects for visual, auditory and tactile changes. Multi-dipole analysis using brain electric source analysis (BESA) was applied to separate temporally overlapping multiple sources. The waveforms of TPJ, MFG and MTG are displayed for each modality. S1, primary somatosensory area; S2, secondary somatosensory area; TPJ, temporo-parietal junction; MFG, middle frontal gyrus; MTG, middle temporal gyrus (Tanaka et al., BMC Neurosci. 2009).

vergence, consistent with our MEG findings in the TPJ. Thus, these fMRI, EEG, and MEG studies indicate that the TPJ, MTG, MFG, IFG, ACC, insula, hippocampus, and other regions respond to inputs from various sensory modalities. Consequently, these areas may be the loci of multisensory integration.

Very early multisensory integration

As mentioned above, it has traditionally been believed that inputs from different sensory modalities are separately processed at the early stage for each sensory modality, and then converge at several stages, including the primary sensory cortices and association or heteromodal multisensory regions, via higher-order sensory cortices. However, extensive studies in both animals and humans during the last decade have revealed that multisensory convergence and interaction start in the primary or near the primary cortical regions at a very early stage after sensory input. Schroeder and colleagues have found neuronal activities of a multisensory nature in and around the auditory cortex, which has been considered to be a primary

modality-specific region. They recorded and analyzed multiunit activities and current source density to examine the laminar profile and timing of activities^{11,20-22}. They also found that rhythmic activities in the auditory cortex are modulated by tactile stimuli, suggesting it plays a modulatory role rather than a driving role²³. It has even been suggested that neocortical operations are essentially multisensory²⁴. Based on a number of anatomical findings, a recent review paper suggested different levels of multisensory convergence along sensory pathways from the early sensory cortices to sensory-specific association cortices and multisensory association cortices, culminating in maximally integrated regions such as the medial temporal lobe cortices and both lateral and medial prefrontal cortices²⁵. The authors also suggested different levels of multisensory divergence in the opposite direction to information flow. Other researchers have considered several possible neural mechanisms for multisensory interaction in sensory-specific cortices^{24,26,27}. The first is feedback effects from multisensory convergence zones in higher association regions to sensory specific regions. The second is feedback effects from multisensory conver-

Fig. 3 The time course of each source activity for all subjects to onset, offset and change of visual stimuli. Multi-dipole analysis using brain electric source analysis (BESA) was applied to separate temporally overlapping multiple sources. Gray and black lines indicate the waveforms of each subject and their grand average, respectively. LG, lingual gyrus (Tanaka et al., NeuroImage 2009).

gence zones at the border of two sensory-specific cortices that have recently been found. The third is direct cortico-cortical projections between low-level modality-specific regions including primary sensory-specific regions. The final mechanism is that different sensory-specific signals are integrated at subcortical levels, including in sensory-specific thalamic nuclei, and then fed forward to the sensory-specific cortices.

Electromagnetic measures of multisensory interaction

In animals and humans, brain regions that receive sensory inputs from different modalities have been revealed, as discussed earlier. However, it is still unclear whether these regions really contribute to multisensory integration and how? An answer to these questions may come from examination of changes in neuronal activities and relevant behaviors when these regions receive inputs simultaneously from different sensory modalities compared to when they receive separate inputs from individual modalities. The additive model is a popular method used to verify the interactive effect of multisensory inputs, and is

based on the comparison between the activation evoked by a multimodal input and the sum of the activation evoked by unimodal inputs from different senses. If there are differences between these activities, then it indicates the existence of non-linear interactive processing between the different signals from different sensory modalities at or before the stage reflected by this activation. EEG or MEG activations are appropriate for the use of this additive model because a scalp-recorded potential (or magnetic field) caused by two sources within the brain is the algebraic sum of potentials (or magnetic fields) resulting from each source. When a sum of unisensory responses subtracted from a multisensory response is a positive value, the interaction is called a “supra-additive” interaction. Inversely, when the difference is a negative value, it is called a “sub-additive” interaction. As described below, most evidence of multisensory interactive effects has been provided by using the additive model.

Giard and Peronnet first reported the audio-visual interactive effect at 40-90 ms and 155-200 ms in the visual cortical area, 90-110 ms in the auditory area, and 140-165 ms in the right frontal cortical area after stimulus onset

in a behavioral and electrophysiological experiment²⁸⁾. While several studies have replicated these interactive effects^{29,30)}, some studies have reported a long-lasting slow potential preceding the eliciting stimuli in the early interactive effect. If subjects perform a task (e.g., button-pressing the presentation of the eliciting stimulus), a slow potential might precede the eliciting stimulus. When using the additive model, the responses to two unisensory stimuli are summed and then compared with the response to a multisensory stimulus. Consequently, the slow potentials preceding unisensory stimuli are summed, but not for the multisensory stimulus. Thus, the waveform may be distorted by the duplication of stimulus-preceding slow potential only for the sum of the unisensory responses. In fact, the early interactive effect disappeared when the slow potential was removed by high-pass filtering, and thus the authors suggested that the early interactive effect is a virtual effect caused by the slow potential preceding the eliciting stimuli^{31,32)}. In contrast, the audio-somatosensory early interactive effect was observed even in a passive condition, where subjects had no task to perform and thus there was little slow potential preceding the eliciting stimuli³³⁾. Also studies have reported that interactive effects starting at ~100 ms after stimulus onset were still present even when the virtual effect of the stimulus-preceding slow potential was removed^{31,32,34)}. Subsequent studies have consistently found interactive effects starting at ~100 ms poststimulus onset using motion stimuli in vision and audition^{35,36)}. Thus, the discrepancy of the early interactive effect indicates a need for consideration of the experimental paradigm and analysis parameters when the additive model is adopted to examine multisensory interaction in EEG.

MEG is also a non-invasive method to record cortical activity from the surface of the scalp and accurately measure its temporal dynamics, which has the advantage of spatial resolution over EEG. Sams and colleagues reported that MEG responses in the supra-temporal auditory cortex were affected by visual information³⁷⁾. In an early study, Lam and colleagues reported the effect of continuous auditory stimulation on responses in the ipsilateral SII to median nerve stimulation³⁸⁾. Using the additive model, Raij et al. reported the convergence and interactions of brain activities related to the audio-visual integration mechanisms for letters³⁹⁾. Letters of the alphabet have phonemic and graphic aspects, which is useful to investigate audio-visual integration. They reported convergence in the lateral middle temporal areas around 225 ms, and suggested that this neural activity reflects free access multisensory processing, which occurs regardless of whether the letter is recognized or not. Their main result was a suppressive interaction in the frontal cortex, right temporo-occipito-parietal junction (280-345 ms), and STS (380-540 ms). Also, these interactions were stronger for letters than for controls. They suggested that left STS indicated audio-visual encoding and transformation of single letters,

activation of the right temporo-occipito-parietal junction reflected a perceptual classification function, and activation of the frontal cortex reflected a general attentional function. Using the same strategy, several MEG studies have identified neural correlates of audiotactile interaction at 130-220 ms poststimulus onset⁴⁰⁾ and 75-85 and 105-130 ms poststimulus onset⁴¹⁾. Furthermore, Raij and colleagues reported that audiovisual interaction started at 85 ms in the left auditory cortex, 80 ms in the right auditory cortex, and 74 ms in the visual cortex, i.e., 3-21 ms after inputs from the two sensory modalities converged in the sensory cortex⁴²⁾. Another MEG study of audiovisual interaction from another laboratory used a different strategy to show that the posterior parietal region responded greatest at 100 ms after multisensory stimulus onset, whereas activities in the superior temporal and posterior cingulate regions increased between 200 and 400 ms, and activities in the parahippocampal, dorsomedial, and orbitofrontal cortices increased between 100 and 400 ms poststimulus onset according to different task instructions⁴³⁾.

In sum, EEG and MEG interaction studies have consistently found that middle- and long-latency interactive effects at 100 ms poststimulus onset are generally observed for several sensory modalities, several tasks, both recording techniques (MEG and EEG), and their combination, whereas the early effect within 100 ms remains a matter of debate regarding the virtual effect of the stimulus-preceding slow potential that has been reported in EEG studies. Most of the interactive effects were suppressive or sub-additive.

Oscillations in multisensory integration

The merit of EEG and MEG is their high temporal resolution. This advantage also enables examinations at a wide range of frequencies from near-DC to more than 100 Hz by transforming time-dimension data to frequency-dimension data. Thus, effective methods of analysis include evoked response analysis, as described above, and frequency analysis, which is used to examine periodic oscillatory activities decomposed from a MEG/EEG waveform. Oscillatory activities in different frequencies are associated with different neural mechanisms and functions, and thus provide essential information about multisensory integration. However, the additive method described above cannot be simply applied to the examination of oscillations⁴⁴⁾. Time-frequency analysis employed to investigate oscillatory activities is generally performed using short-time Fourier or wavelet transformations. However, these are non-linear transformations, and thus the linear additive method causes an undesirable result. Therefore, several strategies are taken to examine the involvement of oscillatory activities in multisensory integration; e.g., the use of bootstrapping as a statistical analysis of the effect of simultaneous presentation from different sensory modalities⁴⁴⁾ and examination of multisensory integration

using non-simultaneous presentations (i.e., no need to use the additive model)⁴⁵. Additionally, recent studies have highlighted the effect of prestimulus oscillatory activities on the McGurk effect⁴⁶, the sound-induced flash illusion⁴⁷, and multisensory bistable integration⁴⁸. Another line of oscillation studies have emphasized the involvement of different oscillatory activities in bottom-up and top-down prediction errors generated in multisensory processing in terms of the prediction coding theory⁴⁹. Thus, oscillatory activities are associated with multisensory integration, but more relevant studies are warranted.

fMRI activation studies of multisensory integration in humans

In their early fMRI study, Calvert and colleagues examined the activation of the bilateral auditory cortex (Brodmann areas 41, 42, and 22) during silent lip-reading⁵⁰. They found that linguistic facial movements activated bilateral auditory cortices, but meaningless speech-like lip movements did not. Thus, they suggested that seen speech influences the perception of heard speech at a pre-lexical stage. Subsequently, they reported fMRI activation in the left STS (Brodmann areas 22/21) showing supra- and sub-additive response properties to congruent and incongruent audio-visual inputs, respectively⁵¹. Their suggestion was that these activation changes reflect the integration of modality-specific inputs and upstream multimodal areas such as the middle temporal gyrus (MTG). Calvert and colleagues suggest several criteria that provide evidence for multisensory integration in human fMRI studies; (1) responsive to stimuli from two different sensory modalities, (2) respond more to the multimodal stimulus than to either of the unisensory stimuli, (3) responsive to a multimodal stimulus, but not to either unisensory stimulus, and (4) respond more to a congruent multisensory stimulus compared to an incongruent stimulus. They also employed an additive model, which has been used in human and animal electrophysiological studies, as explained earlier, to provide evidence for multisensory integration^{51,52}.

Beauchamp and colleagues showed the importance of the posterior STS and MTG in integrating auditory and visual information about complex objects⁵³. Their experimental procedures were devised to extract activations related to multimodal processing from activations related to the subject's task, like movement. In this experiment, they showed activation mapping on individual brain surfaces. In another study, high resolution fMRI showed the patchy distribution of inputs in the human superior temporal sulcus⁵⁴. Although we could not describe all of the relevant fMRI studies here, a large number of studies have demonstrated the involvement of the pSTS, STG, TPJ, IFG, and surrounding regions, as well as modality-specific cortical regions, in multisensory convergence and interaction.

Multisensory single-neuronal convergence vs. regional convergence

As we have described above, a number of fMRI studies have reported multisensory regional or voxel-wise convergence in modality-specific, multisensory, and maximally integrative regions, but regional convergence revealed by human neuroimaging does not necessarily prove single-neuronal convergence in animal physiology. A single multisensory area, such as the pSTS, and even a single voxel in a multisensory region, have unisensory neurons as well as multisensory neurons, and the fMRI signal from these includes the composite activity from large neuronal ensembles. Therefore, the criteria above, used as evidence for multisensory integration in fMRI studies, do not prove single-neuron convergence. This is crucial in considering a translation from single-neuron level physiology to population-level physiology, and thus was extensively assessed using a modelling study of SC neurons and fMRI BOLD signals by Stein's group⁵⁵, which is one of the most successful research groups of multisensory integration in the SC. This modeling study showed no evidence for supra-additive activation in the modeled fMRI BOLD signal, and rather showed a tendency for sub-additive activation. The situation regarding the relationship between the BOLD signal and single-neuron action potentials in multisensory convergence is more complex, given that the BOLD signal better reflects the local field potential (LFP) than action potentials, i.e., synaptic inputs rather than outputs. Consequently, single-neuron studies and population-level studies may provide complementary information in the investigation of the neural mechanisms underlying multisensory integration. In this regard, MEG/EEG has a unique property that enables one to assess changes in neuronal activities at different frequencies, as described earlier. Thus, various directions of studies are undoubtedly necessary for further progress in the scientific study of multisensory integration. Integrative use of different techniques might reveal different actions of synaptic inputs and outputs in multisensory integration.

Decoding of multisensory object representations

As seen above, early fMRI studies of audiovisual integration have used various strategies, such as the additive model and comparison between congruent vs incongruent audiovisual stimuli, and consistently identified the involvement of the pSTS, STG, and IFG in multisensory regional convergence. In contrast, a series of recent studies have used machine-learning algorithms in multivoxel pattern analysis of fMRI. Using this decoding technique, researchers have successfully demonstrated that sound-implicating silent movies were reliably distinguished or decoded based on activity in the primary auditory cortex⁵⁶, that touch-implicating movies were decoded based on activity in the primary somatosensory cortex⁵⁷, and that the

pSTS contained modality-invariant object representations across vision and audition⁵⁸⁾. These studies imply that the primary sensory cortices can use information that should be primarily processed in the other primary cortical regions.

Brain stimulation study

Another line of studies of multisensory integration is not based on the recording aspect of multisensory processing, but on artificial stimulation techniques such as transcranial magnetic stimulation (TMS) and transcranial direct current stimulation (tDCS). Convento and colleagues reported that the cross-modal influences on phosphine perception elicited by TMS over the occipital cortex were facilitated by auditory and/or tactile inputs. They also added tDCS on occipital, temporal, and parietal areas to the experiment described above and confirmed the influence of excitability in other sensory cortical areas⁵⁹⁾. These perceptual enhancements induced by facilitation of another sensory area are evidence of integration of multisensory inputs.

Summary

We have reviewed neuroscientific non-human primate anatomy and electrophysiology, as well as human neuroimaging, studies of multisensory integration. Multisensory convergence and interaction play critical roles in multisensory integration, but convergence in human neuroimaging studies and primate single-neuron recordings must be cautiously interpreted and might be associated with different neural substrates underlying multisensory integration. The new generation of stimulation and analysis techniques, such as oscillation, connectivity, network, decoding, and brain stimulation techniques, will likely increase our understanding of the neural mechanisms underlying multisensory interplay in complex environments.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this article.

References

- 1) Meredith MA and Stein BE. 1983. Interactions among converging sensory inputs in the superior colliculus. *Science* 221: 389-391. doi: 10.1126/science.6867718.
- 2) Jones EG and Powell TP. 1970. An anatomical study of converging sensory pathways within the cerebral cortex of the monkey. *Brain* 93: 793-820. doi: 10.1093/brain/93.4.793.
- 3) Chavis DA and Pandya DN. 1976. Further observations on corticofrontal connections in the rhesus monkey. *Brain Res* 117: 369-386. doi: 10.1016/0006-8993(76)90089-5.
- 4) Seltzer B and Pandya DN. 1978. Afferent cortical connec-

- tions and architectonics of the superior temporal sulcus and surrounding cortex in the rhesus monkey. *Brain Res* 149: 1-24. doi: 10.1016/0006-8993(78)90584-X.
- 5) Meredith MA and Stein BE. 1986. Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *J Neurophysiol* 56: 640-662.
- 6) Wallace MT, Meredith MA and Stein BE. 1993. Converging influences from visual, auditory, and somatosensory cortices onto output neurons of the superior colliculus. *J Neurophysiol* 69: 1797-1809.
- 7) Wallace MT and Stein BE. 1994. Cross-modal synthesis in the midbrain depends on input from cortex. *J Neurophysiol* 71: 429-432.
- 8) Wallace MT, Wilkinson LK and Stein BE. 1996. Representation and integration of multiple sensory inputs in primate superior colliculus. *J Neurophysiol* 76: 1246-1266.
- 9) Hikosaka K, Iwai E, Saito H and Tanaka K. 1988. Polysensory properties of neurons in the anterior bank of the caudal superior temporal sulcus of the macaque monkey. *J Neurophysiol* 60: 1615-1637.
- 10) Mistlin AJ and Perrett DI. 1990. Visual and somatosensory processing in the macaque temporal cortex: the role of 'expectation'. *Exp Brain Res* 82: 437-450. doi: 10.1007/BF00231263.
- 11) Schroeder CE and Foxe JJ. 2002. The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. *Brain Res Cogn Brain Res* 14: 187-198. doi: 10.1016/S0926-6410(02)00073-3.
- 12) Mesulam MM and Mufson EJ. 1982. Insula of the old world monkey. I. Architectonics in the insulo-orbito-temporal component of the paralimbic brain. *J Comp Neurol* 212: 1-22. doi: 10.1002/cne.902120102.
- 13) Kadunce DC, Vaughan JW, Wallace MT, Benedek G and Stein BE. 1997. Mechanisms of within- and cross-modality suppression in the superior colliculus. *J Neurophysiol* 78: 2834-2847.
- 14) Dehner LR, Keniston LP, Clemo HR and Meredith MA. 2004. Cross-modal circuitry between auditory and somatosensory areas of the cat anterior ectosylvian sulcal cortex: a 'new' inhibitory form of multisensory convergence. *Cereb Cortex* 14: 387-403. doi: 10.1093/cercor/bhg135.
- 15) Downar J, Crawley AP, Mikulis DJ and Davis KD. 2000. A multimodal cortical network for the detection of changes in the sensory environment. *Nat Neurosci* 3: 277-283.
- 16) Tanaka E, Inui K, Kida T, Miyazaki T, Takeshima Y and Kakigi R. 2008. A transition from unimodal to multimodal activations in four sensory modalities in humans: an electrophysiological study. *BMC Neurosci* 9: 116. doi: 10.1186/1471-2202-9-116.
- 17) Tanaka E, Kida T, Inui K and Kakigi R. 2009. Change-driven cortical activation in multisensory environments: an MEG study. *Neuroimage* 48: 464-474. doi: 10.1016/j.neuroimage.2009.06.037.
- 18) Tanaka E, Inui K, Kida T and Kakigi R. 2009. Common cortical responses evoked by appearance, disappearance and change of the human face. *BMC Neurosci* 10: 38. doi: 10.1186/1471-2202-10-38.
- 19) Bonner MF, Peelle JE, Cook PA and Grossman M. 2013. Heteromodal conceptual processing in the angular gyrus. *Neuroimage* 71: 175-186. doi: 10.1016/j.neuroimage.2013.01.006.
- 20) Schroeder CE, Smiley J, Fu KG, McGinnis T, O'Connell MN

- and Hackett TA. 2003. Anatomical mechanisms and functional implications of multisensory convergence in early cortical processing. *Int J Psychophysiol* 50: 5-17. doi: 10.1016/S0167-8760(03)00120-X.
- 21) Hackett TA, De La Mothe LA, Ulbert I, Karmos G, Smiley J and Schroeder CE. 2007. Multisensory convergence in auditory cortex, II. Thalamocortical connections of the caudal superior temporal plane. *J Comp Neurol* 502: 924-952. doi: 10.1002/cne.21326.
 - 22) Smiley JF, Hackett TA, Ulbert I, Karmos G, Lakatos P, Javitt DC and Schroeder CE. 2007. Multisensory convergence in auditory cortex, I. Cortical connections of the caudal superior temporal plane in macaque monkeys. *J Comp Neurol* 502: 894-923. doi: 10.1002/cne.21325.
 - 23) Lakatos P, Chen CM, O'Connell MN, Mills A and Schroeder CE. 2007. Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron* 53: 279-292. doi: 10.1016/j.neuron.2006.12.011.
 - 24) Ghazanfar AA and Schroeder CE. 2006. Is neocortex essentially multisensory? *Trends Cogn Sci* 10: 278-285. doi: 10.1016/j.tics.2006.04.008.
 - 25) Man K, Kaplan J, Damasio H and Damasio A. 2013. Neural convergence and divergence in the mammalian cerebral cortex: from experimental neuroanatomy to functional neuroimaging. *J Comp Neurol* 521: 4097-4111. doi: 10.1002/cne.23408.
 - 26) Driver J and Noesselt T. 2008. Multisensory interplay reveals crossmodal influences on 'sensory-specific' brain regions, neural responses, and judgments. *Neuron* 57: 11-23. doi: 10.1016/j.neuron.2007.12.013.
 - 27) Tyll S, Budinger E and Noesselt T. 2011. Thalamic influences on multisensory integration. *Commun Integr Biol* 4: 378-381. doi: 10.4161/cib.4.4.15222.
 - 28) Giard MH and Peronnet F. 1999. Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J Cogn Neurosci* 11: 473-490. doi: 10.1162/089892999563544.
 - 29) Fort A, Delpuech C, Pernier J and Giard MH. 2002. Dynamics of cortico-subcortical cross-modal operations involved in audio-visual object detection in humans. *Cereb Cortex* 12: 1031-1039. doi: 10.1093/cecor/12.10.1031.
 - 30) Molholm S, Ritter W, Murray MM, Javitt DC, Schroeder CE and Foxe JJ. 2002. Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Brain Res Cogn Brain Res* 14: 115-128. doi: 10.1016/S0926-6410(02)00066-6.
 - 31) Teder-Salejari WA, McDonald JJ, Di Russo F and Hillyard SA. 2002. An analysis of audio-visual crossmodal integration by means of event-related potential (ERP) recordings. *Brain Res Cogn Brain Res* 14: 106-114. doi: 10.1016/S0926-6410(02)00065-4.
 - 32) Talsma D and Woldorff MG. 2005. Selective attention and multisensory integration: multiple phases of effects on the evoked brain activity. *J Cogn Neurosci* 17: 1098-1114. doi: 10.1162/0898929054475172.
 - 33) Murray MM, Molholm S, Michel CM, Heslenfeld DJ, Ritter W, Javitt DC, Schroeder CE and Foxe JJ. 2005. Grabbing your ear: rapid auditory-somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cereb Cortex* 15: 963-974. doi: 10.1093/cercor/bhh197.
 - 34) Busse L, Roberts KC, Crist RE, Weissman DH and Woldorff MG. 2005. The spread of attention across modalities and space in a multisensory object. *Proc Natl Acad Sci U S A* 102: 18751-18756. doi: 10.1073/pnas.0507704102.
 - 35) Senkowski D, Saint-Amour D, Kelly SP and Foxe JJ. 2007. Multisensory processing of naturalistic objects in motion: a high-density electrical mapping and source estimation study. *Neuroimage* 36: 877-888. doi: 10.1016/j.neuroimage.2007.01.053.
 - 36) Liu B, Jin Z, Wang Z and Gong C. 2011. The influence of temporal asynchrony on multisensory integration in the processing of asynchronous audio-visual stimuli of real-world events: an event-related potential study. *Neuroscience* 176: 254-264. doi: 10.1016/j.neuroscience.2010.12.028.
 - 37) Sams M, Aulanko R, Hamalainen M, Hari R, Lounasmaa OV, Lu ST and Simola J. 1991. Seeing speech: visual information from lip movements modifies activity in the human auditory cortex. *Neurosci Lett* 127: 141-145. doi: 10.1016/0304-3940(91)90914-F.
 - 38) Lam K, Kakigi R, Kaneoke Y, Naka D, Maeda K and Suzuki H. 1999. Effects of visual and auditory stimulation on somatosensory evoked magnetic fields. *Clin Neurophysiol* 110: 295-304. doi: 10.1016/S0168-5597(98)00059-8.
 - 39) Raji T, Uutela K and Hari R. 2000. Audiovisual integration of letters in the human brain. *Neuron* 28: 617-625. doi: 10.1016/S0896-6273(00)00138-0.
 - 40) Lutkenhoner B, Lammertmann C, Simoes C and Hari R. 2002. Magnetoencephalographic correlates of audiotactile interaction. *Neuroimage* 15: 509-522. doi: 10.1006/nimg.2001.0991.
 - 41) Gobbels R, Schürmann M, Forss N, Juottonen K, Buchner H and Hari R. 2003. Activation of the human posterior parietal and temporoparietal cortices during audiotactile interaction. *Neuroimage* 20: 503-511. doi: 10.1016/S1053-8119(03)00312-4.
 - 42) Raji T, Ahveninen J, Lin FH, Witzel T, Jaaskelainen IP, Letham B, Israeli E, Sahyoun C, Vasios C, Stufflebeam S, Hamalainen M and Belliveau JW. 2010. Onset timing of cross-sensory activations and multisensory interactions in auditory and visual sensory cortices. *Eur J Neurosci* 31: 1772-1782. doi: 10.1111/j.1460-9568.2010.07213.x.
 - 43) Diaconescu AO, Alain C and McIntosh AR. 2011. The co-occurrence of multisensory facilitation and cross-modal conflict in the human brain. *J Neurophysiol* 106: 2896-2909. doi: 10.1152/jn.00303.2011.
 - 44) Senkowski D, Gomez-Ramirez M, Lakatos P, Wylie GR, Molholm S, Schroeder CE and Foxe JJ. 2007. Multisensory processing and oscillatory activity: analyzing non-linear electrophysiological measures in humans and simians. *Exp Brain Res* 177: 184-195. doi: 10.1007/s00221-006-0664-7.
 - 45) Ahveninen J, Huang S, Ahlfors SP, Hamalainen M, Rossi S, Sams M and Jaaskelainen IP. 2016. Interacting parallel pathways associate sounds with visual identity in auditory cortices. *Neuroimage* 124: 858-868. doi: 10.1016/j.neuroimage.2015.09.044.
 - 46) Keil J, Müller N, Ihssen N and Weisz N. 2012. On the variability of the McGurk effect: audiovisual integration depends on prestimulus brain states. *Cereb Cortex* 22: 221-231. doi: 10.1093/cercor/bhr125.
 - 47) Keil J, Müller N, Hartmann T and Weisz N. 2014. Prestimulus beta power and phase synchrony influence the sound-induced flash illusion. *Cereb Cortex* 24: 1278-1288. doi: 10.1093/cer-

- cor/bhs409.
- 48) Leonardelli E, Braun C, Weisz N, Lithari C, Occelli V and Zampini M. 2015. Prestimulus oscillatory alpha power and connectivity patterns predispose perceptual integration of an audio and a tactile stimulus. *Hum Brain Mapp* 36: 3486-3498. doi: 10.1002/hbm.22857.
 - 49) Arnal LH, Wyart V and Giraud AL. 2011. Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nat Neurosci* 14: 797-801. doi: 10.1038/nn.2810.
 - 50) Calvert GA, Bullmore ET, Brammer MJ, Campbell R, Williams SC, McGuire PK, Woodruff PW, Iversen SD and David AS. 1997. Activation of auditory cortex during silent lipreading. *Science* 276: 593-596. doi: 10.1126/science.276.5312.593.
 - 51) Calvert GA, Campbell R and Brammer MJ. 2000. Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Curr Biol* 10: 649-657. doi: 10.1016/S0960-9822(00)00513-3.
 - 52) Calvert GA, Hansen PC, Iversen SD and Brammer MJ. 2001. Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. *Neuroimage* 14: 427-438. doi: 10.1006/nimg.2001.0812.
 - 53) Beauchamp MS, Lee KE, Argall BD and Martin A. 2004. Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron* 41: 809-823. doi: 10.1016/S0896-6273(04)00070-4.
 - 54) Beauchamp MS, Argall BD, Bodurka J, Duyn JH and Martin A. 2004. Unraveling multisensory integration: patchy organization within human STS multisensory cortex. *Nat Neurosci* 7: 1190-1192. doi: 10.1038/nn1333.
 - 55) Laurienti PJ, Perrault TJ, Stanford TR, Wallace MT and 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. doi: 10.1007/s00221-005-2370-2.
 - 56) Meyer K, Kaplan JT, Essex R, Webber C, Damasio H and Damasio A. 2010. Predicting visual stimuli on the basis of activity in auditory cortices. *Nat Neurosci* 13: 667-668. doi: 10.1038/nn.2533.
 - 57) Meyer K, Kaplan JT, Essex R, Damasio H and Damasio A. 2011. Seeing touch is correlated with content-specific activity in primary somatosensory cortex. *Cereb Cortex* 21: 2113-2121. doi: 10.1093/cercor/bhq289.
 - 58) Man K, Kaplan JT, Damasio A and Meyer K. 2012. Sight and sound converge to form modality-invariant representations in temporoparietal cortex. *J Neurosci* 32: 16629-16636. doi: 10.1523/JNEUROSCI.2342-12.2012.
 - 59) Convento S, Vallar G, Galantini C and Bolognini N. 2013. Neuromodulation of early multisensory interactions in the visual cortex. *J Cogn Neurosci* 25: 685-696. doi: 10.1162/jocn_a_00347.