

Published in final edited form as:

*Biol Cybern.* 2012 December ; 106(0): 617–625. doi:10.1007/s00422-012-0502-x.

## **Integrating Information from Different Senses in the Auditory Cortex**

**Andrew J. King** and **Kerry M.M. Walker**

Department of Physiology, Anatomy and Genetics, University of Oxford, Parks Road, Oxford OX1 3PT, United Kingdom

### **Abstract**

Multisensory integration was once thought to be the domain of brain areas high in the cortical hierarchy, with early sensory cortical fields devoted to unisensory processing of inputs from their given set of sensory receptors. More recently, a wealth of evidence documenting visual and somatosensory responses in auditory cortex, even as early as the primary fields, has changed this view of cortical processing. These multisensory inputs may serve to enhance responses to sounds that are accompanied by other sensory cues, effectively making them easier to hear, but may also act more selectively to shape the receptive field properties of auditory cortical neurons to the location or identity of these events. We discuss the new, converging evidence that multiplexing of neural signals may play a key role in informatively encoding and integrating signals in auditory cortex across multiple sensory modalities. We highlight some of the many open research questions that exist about the neural mechanisms that give rise to multisensory integration in auditory cortex, which should be addressed in future experimental and theoretical studies.

### **Keywords**

Multisensory; Auditory Cortex; Perception; Multiplex; Information; Neural code; Visual; Somatosensory

## **1 Introduction**

We live in a multisensory world in which the presence of most objects and events is frequently registered by more than one of our senses. The capacity to combine information across different sensory modalities can enhance perception, particularly if the inputs provided by individual modalities are weak, unreliable or ambiguous. The convergence and integration of multisensory cues within the brain have been shown to improve a wide range of functions from stimulus detection and localization to object recognition (Stein 2012). This has been demonstrated for many different combinations of sensory inputs and in many different species, emphasizing the fact that multisensory processing is a ubiquitous feature of the nervous system.

If perception and behavior are to benefit from combining information across different sensory modalities, it is essential that those inputs are integrated in an appropriate fashion. They may originate from the same or from different sources, so how does the brain combine the sensory cues associated with one source and separate them from those belonging to other sources? For example, visual cues provided by movements of the lips, face and tongue can improve speech intelligibility in the presence of other sounds (Sumbly and Pollack 1954). But if two people are talking at the same time, how does the listener know which voice belongs to which face? Indeed, watching lip movements that correspond to one syllable whilst listening to a different syllable often changes the speech sound that is heard – the well known McGurk effect – illustrating the potent way in which visual cues can modulate auditory perception (McGurk and MacDonald 1976).

Psychophysical studies in humans have examined the importance of physical factors such as spatial and temporal proximity in determining whether multiple sensory inputs are bound together or not (e.g. Sekuler et al. 1997; Slutsky and Recanzone 2001; Soto-Faraco et al. 2002; Zampini et al. 2005; van Wassenhove et al. 2007). In addition, the semantic relationship between the visual and auditory components of more naturalistic stimuli has been shown to influence multisensory integration (Doehrmann and Naumer 2008; Chen and Spence 2010). Consequently, these factors also need to be considered when investigating how multiple sensory cues combine to influence the sensitivity and stimulus selectivity of neurons in the brain.

In searching for the neural underpinnings of multisensory perceptual processing, most studies have, until recently, focussed on a relatively small number of brain regions, including the superior colliculus in the midbrain and higher-level association areas of the cortex in the frontal, parietal and temporal lobes (King 2004; Alais et al. 2010). This is because it was believed that the pooling and integration of information across the senses takes place only after extensive processing within modality-specific pathways that are assigned to each of the senses. After all, our sensory systems are responsible for transducing different forms of energy, and can give rise to perceptions, such as color or pitch, that are not obviously shared by other modalities.

It is now clear, however, that multisensory convergence is much more widespread and occurs at an earlier level in each sensory system, including the primary visual, auditory and somatosensory cortex, than previously thought (Ghazanfar and Schroeder 2006). This raises additional questions concerning the way inputs from different sensory modalities interact in the brain. First, can inputs from other sensory modalities drive neurons in early sensory cortex or do they just modulate the responses of neurons to the primary sensory modality? Second, do the different sensory inputs to a given neuron provide complementary information about, for example, stimulus location or identity? Finally, how do inputs from other modalities influence the activity of neurons in these brain regions without altering the modality specificity of the percepts to which they give rise. In this article, we will consider these issues in the context of visual and somatosensory inputs to the auditory cortex.

## 2. Multisensory convergence in the auditory cortex

Evidence for visual or somatosensory influences on the auditory cortex has been obtained in a number of species using either electrophysiological recordings (Wallace et al. 2004; Ghazanfar et al. 2005; Bizley et al. 2007; Lakatos et al. 2007; Kayser et al. 2008; Besle et al. 2009) or functional imaging (Pekkola et al. 2005; Schürmann et al. 2006; Kayser et al. 2007). Although these effects have been observed in core or primary auditory cortex, the strongest multisensory influences are found in surrounding belt and parabelt areas. Visual and somatosensory influences on auditory cortical responses could originate subcortically, as illustrated, for example, by the presence of multisensory neurons in parts of the thalamus that are known to project there (Benedek et al. 1997). It seems likely, however, that the multisensory interactions are mostly cortical in origin, reflecting inputs to the auditory cortex from high-level association areas, such as the superior temporal sulcus (Seltzer and Pandya 1994; Ghazanfar et al. 2008), or directly from visual and somatosensory cortical areas (Bizley et al. 2007; Budinger and Scheich 2009; Campi et al. 2010; Falchier et al. 2010).

Most of these studies have found that visual or somatosensory stimuli do not, by themselves, evoke responses in the auditory cortex. Instead, multisensory stimulation alters cortical responses to sound, suggesting that these non-auditory inputs are subthreshold and modulatory in nature. This is consistent with both anatomical (Budinger and Scheich 2009; Falchier et al. 2010) and physiological (Lakatos et al. 2007) evidence demonstrating that they terminate primarily outside layer IV, the main thalamorecipient zone of the auditory cortex. In some cases, however, spiking responses to simple visual stimuli, such as light flashes, have been observed in auditory cortex (Wallace et al. 2004; Brosch et al. 2005; Bizley et al. 2007; Kayser et al. 2008). The incidence of these neurons, particularly in the primary areas, is low and little is so far known about their physiological properties or whether they serve a different role from visual or somatosensory inputs that either increase or decrease neuronal responses to the dominant auditory modality. Interestingly, Bizley et al. (2007) found no difference in sound frequency tuning between unisensory auditory neurons and neurons that also responded to visual stimulation, suggesting that responsiveness to other modalities may not alter the more basic sound processing properties of these neurons.

Functional imaging experiments and local field potential (LFP) recordings have consistently shown that visual or somatosensory stimuli typically enhance responses to acoustic stimuli and that these effects are widespread within the auditory cortex (reviewed by Kayser et al. 2009a). It is therefore tempting to conclude that these inputs provide fairly nonspecific contextual information about sound sources, which results in a general improvement in sound detection. In order to understand how multisensory stimulation affects auditory processing, however, it is necessary to investigate these interactions at the level of individual neurons. Neurons exhibiting a significant change in their firing rate in response to multisensory stimulation appear to be more sparsely distributed than expected on the basis of the imaging studies. Moreover, they frequently show crossmodal suppression when different modality stimuli are paired (Bizley et al. 2007; Kayser et al. 2008; Meredith and Allman 2009).

This apparent discrepancy reflects the fact that the suprathreshold firing of neurons is not necessarily well captured by imaging signals and LFPs, which may be related more to the inputs to the cortex (Logothetis et al. 2001; Kayser et al. 2009a). Nevertheless, while the great majority of single and multi-unit electrophysiological studies have focussed on changes in firing rate as a measure of multisensory integration, a reduced rate does not necessarily equate to a less informative response. Indeed, it has been shown that the temporal pattern of action potential firing in an auditory cortical neuron can provide significantly more information than the same neuron's firing rate about both the identity and location of sounds (Nelken et al. 2005). Similar information theoretic analyses have now been applied to the spike discharge patterns produced by neurons in the auditory cortex in response to multisensory stimulation, revealing that visual cues can enhance information processing even when there is no overall change in firing rate (Bizley et al. 2007; Kayser et al. 2010).

### 3. Specificity of multisensory interactions in the auditory cortex

There is considerable evidence that the auditory cortex plays a critical role in the localization (Heffner and Heffner 1990; Jenkins and Masterton 1982; Clarke et al. 2000; Nodal et al. 2010) and discrimination of sounds (Heffner and Heffner 1986; Harrington et al. 2001; Clarke et al. 2000). We might therefore expect crossmodal interactions to alter the capacity of these neurons to signal the location or identity of a sound source.

Evidence for this is provided by electrophysiological recordings in ferrets showing that presenting spatially coincident visual and auditory cues tends to increase the amount of information conveyed by the neurons about stimulus location (Bizley and King 2008; Fig. 1). In principle, this multisensory information enhancement could underlie the improved localization accuracy that is observed when visual and auditory stimuli are presented at the same location in space, although recordings would need to be made in behaving animals in order to confirm this. Activity within the human auditory cortex has been shown to vary on a trial-by-trial basis when subjects experience the ventriloquist illusion - where spatially disparate visual cues "capture" the perceived location of a sound source - implying a more direct link between cortical activity and perception (Bonath et al. 2007).

Visual signals can also sharpen the selectivity of neurons in the monkey auditory cortex to conspecific vocalizations (Ghazanfar et al. 2008). This effect was observed when the visual stimulus comprised a movie of a monkey vocalizing, but not when this image was replaced by a disc that was flashed on and off to approximate the movements of the animal's mouth. Similarly, Kayser et al. (2010) demonstrated that matching pairs of naturalistic sounds and movies evoke responses in auditory cortex that carry more stimulus-related information than those produced by the sounds alone, whereas this increase in transmitted information is reduced if the visual stimulus does not match the sound. Although there are so far relatively few studies of this sort, they do provide compelling evidence that the benefits of multisensory integration are determined by the correspondence between the different stimuli, including their semantic congruency, which could potentially provide the basis for a number of multisensory perceptual phenomena.

The simplest way of binding multisensory inputs that relate to a given stimulus attribute, such as the location or identity of the object of interest, is for the relevant regions of, say, visual and auditory cortex to be connected. This assumes, of course, that a division of labor exists, with different regions of the cortex specialized for processing different stimulus attributes. Evidence for this within the auditory cortex has been provided by a number of cortical deactivation (Lomber and Malhotra 2008), human neurological (Clarke et al. 2000), and brain imaging (Warren and Griffiths 2003; Ahveninen et al. 2006) studies. Moreover, Lomber et al. (2010) demonstrated that crossmodal reorganization takes place in congenitally deaf cats, with specific visual functions localized to different regions of the auditory cortex, and that activity in these regions accounts for the enhanced visual abilities of the animals.

Regional differences within the auditory cortex have also been described in the way that visual and auditory cues interact to determine the responses of the neurons found there. Thus, Ghazanfar et al. (2005) found that in order to generate response enhancement, a greater proportion of recording sites in the belt areas required the use of a real monkey face, whereas non-selective modulation of auditory cortical responses was more common in the core areas. Similarly, Bizley and King (2008) found that the amount of spatial information available in the neural responses varied across the ferret auditory cortex. They observed that spatial sensitivity to paired visual-auditory stimuli was highest in the anterior ectosylvian gyrus, supporting the notion that there is some functional segregation across the auditory cortex, with, in this species, the anterior fields more involved in spatial processing. This is consistent with the anatomical pattern of connections between visual and auditory cortex in the ferret, which also suggest that cross-modal inputs are arranged according to whether they carry information about spatial or non-spatial features (Bizley et al. 2007; Fig. 2).

Despite these compelling results, the evidence for functional segregation in studies that have recorded the responses of single neurons in auditory cortex is comparatively weak. These studies have instead emphasized that individual neurons throughout the auditory cortex are sensitive to a variety of perceptual features, including the pitch, identity and spatial location of sounds (Stecker and Middlebrooks 2003; Recanzone 2008; Bizley et al. 2009). Neurons can differ systematically across cortical fields in the extent to which they are sensitive to one particular feature of sound or another (Tian et al. 2001; Stecker et al. 2005; Bizley et al. 2009), but these responses are not strictly selective for any given feature. This raises an interesting problem regarding multisensory integration in auditory cortex: how can inputs from other sensory systems interact informatively with auditory responses in order to improve sound identification or localization? That is, how do we arrive at a cross-modal enhancement of any one sound feature? An alternative way in which neural responses might invariantly represent multiple perceptual features is to compartmentalize these representations in the spectral or time domain, rather than spatially across different regions of the auditory cortex. This concept is often referred to as “multiplexing”, and evidence for its use in sensory systems has accumulated in recent years (reviewed by Panzeri et al. 2010).

## 4. Multiplexing stimulus information

In the most commonly observed form of cortical multiplexing, precise spike timing and low frequency network oscillations together provide information about sensory stimuli that is additional to that conveyed by spike firing rate. In the visual cortex, the timing of spikes relative to the phase of LFP oscillations in the delta band (1-4 Hz) has been shown to carry information about the contents of naturalistic images (Montemurro et al. 2008). Similarly, Kayser et al. (2009b) found that the timing of spikes with respect to LFP phase in the theta (4-8 Hz) frequency range in auditory cortex was informative about the type of sound being presented to awake, passively listening monkeys. In both these studies, the information provided by the *phase-of-firing* multiplexed code was complementary to that of classic spike-rate codes, thereby enhancing the sensory acuity of the spike code through more global network activation (Fig. 3A).

Lakatos et al. (2007) showed that a slightly different form of phase-of-firing code may also have important implications for crossmodal modulation of auditory cortical responses. They found that while somatosensory stimulation alone was insufficient to evoke spikes in A1 neurons, it did reset the phase of the LFP oscillations in this region. The spiking response to proceeding auditory stimulation was then enhanced or suppressed according to whether the sound-evoked activity arrived in auditory cortex during LFP oscillatory peaks or troughs, respectively (Fig. 3B). We refer to this effect, which was observed for LFP phase in the gamma, theta and delta frequency bands, as cross-modal *phase-based-enhancement* of auditory spike rates. Along similar lines, Kayser et al. (2008) have shown that visual stimulation can reset the phase of low frequency LFPs in auditory cortex, thereby enhancing spiking responses to concurrent auditory stimulation. These oscillatory effects would lead to the prediction that stimulus onset asynchronies corresponding to the period of theta frequencies are optimal for multisensory integration, though this psychophysical result has not yet been demonstrated.

It remains a challenge for future studies to establish a causal behavioral link between LFP phase and multisensory integration. Nevertheless, the aforementioned studies have made the critical first step of illustrating a potential neural mechanism through which other sensory modalities might modulate the spike representations of sounds in auditory cortex. It has also now been shown that visual stimulation can reset low-frequency neural oscillations in human auditory cortex (Luo et al. 2010; Thorne et al. 2011). In keeping with the data from non-human primates (Lakatos et al. 2007), Thorne et al. (2011) found that visual phase-resetting in auditory cortex was confined to particular spectral bands (in this case, high theta to high alpha), and required that visual stimuli preceded auditory stimulation by 30-75 ms.

We have recently described another form of multiplexing in ferret auditory cortex, in which the spike rate of a neuron is modulated by the identity (i.e. spectral timbre) of an artificial vowel sound in one time window and the periodicity - the basis for the perception of pitch - of the vowel in a later time window (Walker et al. 2011; Fig. 4). This allows the neurons to carry mutually invariant information about these two perceptual features. In line with this neurophysiological finding, ferrets were found to detect changes in vowel identity faster than they detected changes in the periodicity of these sounds (Walker et al. 2011). It is

possible that this form of *time-division* multiplexing in auditory cortex is utilized by inputs from other sensory modalities either to selectively modulate auditory responses within a feature-relevant time window or to multiplex different sensory cues in those instances where each is capable of driving the neuron.

These possibilities have yet to be tested directly, but previous studies have shown that visual and auditory information is temporally segregated in auditory cortical neurons, just as it is in other brain areas, in a manner consistent with time-division multiplexing. In anesthetized mammals, the first spike latencies of primary auditory cortical responses to simple sounds are usually in the region of 20-50 ms (Recanzone et al. 2000; Bizley et al. 2007), whereas responses to light flashes have latencies ranging from 40 to >200 ms (Bizley et al. 2007; Bizley and King 2008) (Fig. 1). In awake animals, the first-spike latency of A1 responses to sounds are often <10 ms (Lakatos et al. 2005), but visual-evoked activity in primary visual cortex still occurs typically at least 10-20 ms later than this (Maunsell and Gibson 1992; Schroeder et al. 1998). Responses to visual stimulation are further delayed relative to auditory responses in non-primary auditory cortex (Bizley et al. 2007), and this increased response asynchrony may allow the later sustained responses to auditory stimuli observed in these regions (Bendor and Wang 2008; Walker et al. 2011) to remain temporally distinct from visual responses.

## 5. Possible neural mechanisms for signal multiplexing

A number of potential neurophysiological mechanisms have been proposed through which neurons might create or read out multiplexed neural signals. Phase-of-firing and phase-based-enhancement require a single neuron in brain area 'x' to be sensitive to low-frequency oscillations across the general population of neurons that provide input to 'x'. Blumhagen et al. (2011) have demonstrated that this form of information is extracted from mitral cells in the olfactory bulb that project to the zebrafish dorsal telencephalon via the long membrane time constants of the latter. Although the biophysical properties of dorsal telencephalon neurons act as low-pass filters on their inputs, attenuating the effects of synchronous inputs, mitral cell synchrony is still apparent in the in the timing of their action potentials (Blumhagen et al. 2011). Through a similar mechanism, peaks of low-frequency oscillatory inputs from other modalities may cause excitatory post-synaptic potentials in primary auditory cortical neurons that give rise to phase-based-enhancement of spike rates.

Time-division multiplexing requires: a) a neural signal representing the onset of the stimulus event; b) a mechanism through which spike rates may be integrated over a particular time period with respect to that event; and c) downstream neurons that can decode the temporal firing patterns of multiplexed signals in order to integrate the features of a single perceptual object. Again, each of these requirements can be realized through previously proposed properties of neural circuits. The stimulus event may be signalled by the stimulus-triggered low-frequency oscillations described by Lakatos et al. (2007) and Kayser et al. (2008). Alternatively, this physiological timing signal could be provided by low-threshold, broadly tuned neurons, with post-stimulus response time profiles similar to those of onset cells found in the cochlear nucleus (Blackburn and Sachs 1989) and auditory thalamus (Kvasnak et al. 2000). Spike rates of A1 neurons may then be integrated within defined time windows by

downstream neurons through a balance of inhibitory and excitatory inputs. For example, integration over early, brief time periods has been shown to arise in rat hippocampal neurons as a result of delayed feed-forward inhibition following initial excitation (Pouille and Scanziani 2001). In the same manner, early and transient inhibition summed with excitatory inputs should result in sensitivity to late, but not early, spikes from the excitatory neurons (Fig. 5). Mathematical models have demonstrated that, in principle, single integrate-and-fire neurons are capable of learning to decode the type of complex temporal spiking patterns that would be provided by multiplexed inputs (Gutig and Sompolinsky 2006). While many of these mechanisms remain to be investigated in auditory cortex, the evidence to date suggests that they are feasible and utilized in other neural circuits.

## 6. Implications of multisensory integration dynamics for cortical processing

A final challenge that has yet to be addressed within the above physiological models is the temporal dynamics of multisensory integration. A large body of psychophysical evidence has shown that acoustic and visual events are integrated over a temporal window of several hundred milliseconds, with a bias towards asynchronies in which the visual stimulus leads the auditory stimulus (reviewed by Vatakis and Spence 2010). However, the precise temporal integration window varies depending on factors such as the type of stimulus (e.g. speech vs non-speech), the medium in which it is presented, and the distance of the perceived object from the subject (King 2005), and can also change in response to behavioural training (Powers et al. 2009). In the real world, it makes good sense that auditory and visual cues should be integrated over different time windows in a context-dependent manner, as the differences in the speed of light and sound produce larger auditory lags for distant objects (e.g. several seconds for a lightning strike) than proximate ones (e.g. tens of milliseconds for typical conversational speech).

In contrast to the temporal variability identified in the psychophysical literature, initial electrophysiological studies of multiplexing in passively listening (Lakatos et al. 2007; Kayser et al. 2008, 2009b) and anesthetized (Walker et al. 2011) animals have identified relatively stable time windows for encoding and integrating information in the auditory cortex. However, recent imaging data in humans (Powers et al. 2012) suggests that these physiological integration windows may become more dynamic when tested across a range of stimulus types or when measured in behaving animals. Neuromodulators such as acetylcholine, noradrenaline and serotonin can modulate the excitability of auditory cortical neurons (Foote et al. 1975; Hurley and Hall 2011) and are released during behavioral tasks (Stark and Scheich 1997; Butt et al. 2009). Acetylcholine release from the basal forebrain has been shown to alter the reliability of spike firing and synchrony among neurons in the visual cortex (Goard and Dan 2009), which will presumably influence the way cortical neurons multiplex sensory signals. The temporal parameters of multisensory integration may therefore be adaptable through the mechanisms of attention and arousal.



## 7. Concluding remarks

It is now widely accepted that multisensory convergence occurs to a surprisingly large degree in early sensory areas, including core and belt regions of the auditory cortex, which were previously thought to be unisensory in nature. Nevertheless, the auditory cortex is still primarily an auditory structure and a challenge for future work will be to identify how visual and somatosensory inputs influence the activity of its neurons in a behaviorally relevant manner without compromising their role in hearing. Recent studies of the physiological basis of multisensory integration have demonstrated the power of information theoretic approaches in determining whether neurons become more reliable at detecting or encoding a particular sound in the presence of other sensory cues. Investigations of various forms of multiplexing in neural responses show that we need to consider different features of the neural code, rather than relying solely on firing rate, in order to understand fully how multisensory cues are represented in the brain.

The present review has focussed on visual and somatosensory influences on auditory cortical encoding of sounds, but auditory influences on non-auditory cortex may likewise enhance the processing of other sensory modalities (Meredith et al. 2009). The extent to which these forms of multisensory integration utilize common neural mechanisms is unknown and largely unexplored. Therefore, while the first hints of the neural basis of multisensory integration are beginning to be uncovered, there remains a wealth of questions about these circuits that may be addressed in future physiological, behavioral and theoretical studies.

## Acknowledgements

Our research is supported by the Wellcome Trust through a Principal Research Fellowship to A. J. King (WT076508AIA).

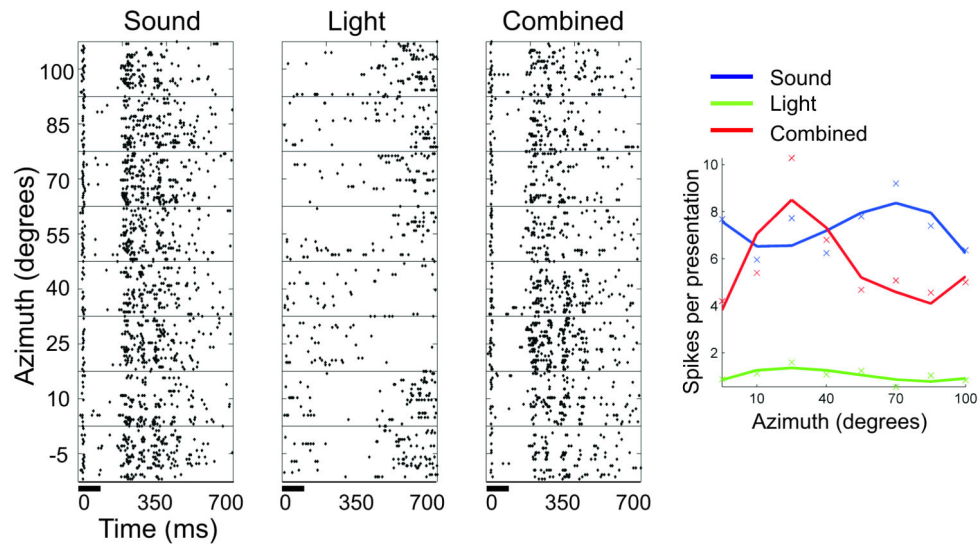
## References

- Ahveninen J, Jääskeläinen IP, Raji T, Bonmassar G, Devore S, Hämäläinen M, Levänen S, Lin FH, Sams M, Shinn-Cunningham BG, Witzel T, Belliveau JW. Task-modulated “what” and “where” pathways in human auditory cortex. *Proc Natl Acad Sci U S A*. 2006; 103:14608–14613. [PubMed: 16983092]
- Alais D, Newell FN, Mamassian P. Multisensory processing in review: from physiology to behaviour. *Seeing Perceiving*. 2010; 23:3–38. [PubMed: 20507725]
- Bendor D, Wang X. Neural response properties of primary, rostral, and rostrotemporal core fields in the auditory cortex of marmoset monkeys. *J Neurophysiol*. 2008; 100:888–906. [PubMed: 18525020]
- Benedek G, Perény J, Kovács G, Fischer-Szátmári L, Katoh YY. Visual, somatosensory, auditory and nociceptive modality properties in the feline suprageniculate nucleus. *Neuroscience*. 1997; 78:179–189. [PubMed: 9135099]
- Besle J, Bertrand O, Giard MH. Electrophysiological (EEG, sEEG, MEG) evidence for multiple audiovisual interactions in the human auditory cortex. *Hear Res*. 2009; 258:143–151. [PubMed: 19573583]
- Bizley JK, King AJ. Visual-auditory spatial processing in auditory cortical neurons. *Brain Res*. 2008; 1242:24–36. [PubMed: 18407249]

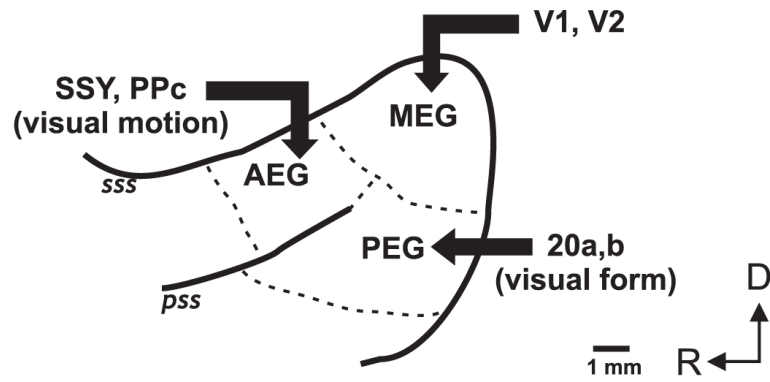
- Bizley JK, Nodal FR, Bajo VM, Nelken I, King AJ. Physiological and anatomical evidence for multisensory interactions in auditory cortex. *Cereb Cortex*. 2007; 17:2172–2189. [PubMed: 17135481]
- Bizley JK, Walker KMM, Silverman BW, King AJ, Schnupp JWH. Interdependent encoding of pitch, timbre, and spatial location in auditory cortex. *J Neurosci*. 2009; 29:2064–2075. [PubMed: 19228960]
- Blackburn CC, Sachs MB. Classification of unit types in the anteroventral cochlear nucleus: PST histograms and regularity analysis. *J Neurophysiol*. 1989; 62:1303–1329. [PubMed: 2600627]
- Blumhagen F, Zhu P, Shum J, Scharer YP, Yaksi E, Deisseroth K, Friedrich RW. Neuronal filtering of multiplexed odour representations. *Nature*. 2011; 479:493–498. [PubMed: 22080956]
- Bonath B, Noesselt T, Martinez A, Mishra J, Schwiecker K, Heinze HJ, Hillyard SA. Neural basis of the ventriloquist illusion. *Curr Biol*. 2007; 17:1697–1703. [PubMed: 17884498]
- Brosch M, Selezneva E, Scheich H. Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. *J Neurosci*. 2005; 25:6797–6806. [PubMed: 16033889]
- Budinger E, Scheich H. Anatomical connections suitable for the direct processing of neuronal information of different modalities via the rodent primary auditory cortex. *Hear Res*. 2009; 258:16–27. [PubMed: 19446016]
- Butt AE, Chavez CM, Flesher MM, Kinney-Hurd BL, Araujo GC, Miasnikov AA, Weinberger NM. Association learning-dependent increases in acetylcholine release in the rat auditory cortex during auditory classical conditioning. *Neurobiol Learn Mem*. 2009; 92:400–409. [PubMed: 19467339]
- Campi KL, Bales KL, Grunewald R, Krubitzer L. Connections of auditory and visual cortex in the prairie vole (*Microtus ochrogaster*): evidence for multisensory processing in primary sensory areas. *Cereb Cortex*. 2010; 20:89–108. [PubMed: 19395525]
- Chen YC, Spence C. When hearing the bark helps to identify the dog: semantically-congruent sounds modulate the identification of masked pictures. *Cognition*. 2010; 114:389–404. [PubMed: 19909945]
- Clarke S, Bellmann A, Meuli RA, Assal G, Steck AJ. Auditory agnosia and auditory spatial deficits following left hemispheric lesions: evidence for distinct processing pathways. *Neuropsychologia*. 2000; 38:797–807. [PubMed: 10689055]
- Doehrmann O, Naumer MJ. Semantics and the multisensory brain: how meaning modulates processes of audio-visual integration. *Brain Res*. 2008; 1242:136–150. [PubMed: 18479672]
- Falchier A, Schroeder CE, Hackett TA, Lakatos P, Nascimento-Silva S, Ulbert I, Karmos G, Smiley JF. Projection from visual areas V2 and prostriata to caudal auditory cortex in the monkey. *Cereb Cortex*. 2010; 20:1529–1538. [PubMed: 19875677]
- Foote SL, Freedman R, Oliver AP. Effects of putative neurotransmitters on neuronal activity in monkey auditory cortex. *Brain Res*. 1975; 86:229–242. [PubMed: 234774]
- Ghazanfar AA, Chandrasekaran C, Logothetis NK. Interactions between the superior temporal sulcus and auditory cortex mediate dynamic face/voice integration in rhesus monkeys. *J Neurosci*. 2008; 28:4457–4469. [PubMed: 18434524]
- Ghazanfar AA, Maier JX, Hoffman KL, Logothetis NK. Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *J Neurosci*. 2005; 25:5004–5012. [PubMed: 15901781]
- Ghazanfar AA, Schroeder CE. Is neocortex essentially multisensory? *Trends Cog Sci*. 2006; 10:278–285.
- Goard M, Dan Y. Basal forebrain activation enhances cortical coding of natural scenes. *Nat Neurosci*. 2009; 12:1444–1449. [PubMed: 19801988]
- Gutig R, Sompolinsky H. The tempotron: a neuron that learns spike timing-based decisions. *Nat Neurosci*. 2006; 9:420–428. [PubMed: 16474393]
- Harrington IA, Heffner RS, Heffner HE. An investigation of sensory deficits underlying the aphasia-like behaviour of macaques with auditory cortex lesions. *Neuroreport*. 2001; 12:1271–1221. [PubMed: 11338205]
- Heffner HE, Heffner RS. Effect of unilateral and bilateral auditory cortex lesions on the discrimination of vocalizations by Japanese macaques. *J Neurophysiol*. 1986; 56:683–701. [PubMed: 3783215]

- Heffner HE, Heffner RS. Effect of bilateral auditory cortex lesions on sound localization in Japanese macaques. *J Neurophysiol.* 1990; 64:915–931. [PubMed: 2230934]
- Hurley LM, Hall IC. Context-dependent modulation of auditory processing by serotonin. *Hear Res.* 2011; 279:74–84. [PubMed: 21187135]
- Jenkins WM, Masterton RB. Sound localization: effects of unilateral lesions in central auditory system. *J Neurophysiol.* 1982; 47:987–1016. [PubMed: 7108581]
- Kayser C, Logothetis NK, Panzeri S. Visual enhancement of the information representation in auditory cortex. *Curr Biol.* 2010; 20:19–24. [PubMed: 20036538]
- Kayser C, Petkov CI, Augath M, Logothetis NK. Functional imaging reveals visual modulation of specific fields in auditory cortex. *J Neurosci.* 2007; 27:1824–1835. [PubMed: 17314280]
- Kayser C, Petkov CI, Logothetis NK. Visual modulation of neurons in auditory cortex. *Cereb Cortex.* 2008; 18:1560–1574. [PubMed: 18180245]
- Kayser C, Petkov CI, Logothetis NK. Multisensory interactions in primate auditory cortex: fMRI and electrophysiology. *Hear Res.* 2009a; 258:80–88. [PubMed: 19269312]
- Kayser C, Montemurro MA, Logothetis NK, Panzeri S. Spike-phase coding boosts and stabilizes information carried by spatial and temporal spike patterns. *Neuron.* 2009b; 61:597–608. [PubMed: 19249279]
- King AJ. The superior colliculus. *Curr Biol.* 2004; 14:R335–338. [PubMed: 15120083]
- King AJ. Multisensory integration: strategies for synchronization. *Curr Biol.* 2005; 15:R339–341. [PubMed: 15886092]
- Kvasnák E, Popelár J, Syka J. Discharge properties of neurons in subdivisions of the medial geniculate body of the guinea pig. *Physiol Res.* 2000; 49:369–378. [PubMed: 11043925]
- Lakatos P, Chen CM, O’Connell MN, Mills A, Schroeder CE. Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron.* 2007; 53:279–292. [PubMed: 17224408]
- Lakatos P, Pincze Z, Fu KM, Javitt DC, Karmos G, Schroeder CE. Timing of pure tone and noise-evoked responses in macaque auditory cortex. *Neuroreport.* 2005; 16:933–937. [PubMed: 15931064]
- Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A. Neurophysiological investigation of the basis of the fMRI signal. *Nature.* 2001; 412:150–157. [PubMed: 11449264]
- Lomber SG, Malhotra S. Double dissociation of ‘what’ and ‘where’ processing in auditory cortex. *Nat Neurosci.* 2008; 11:609–616. [PubMed: 18408717]
- Lomber SG, Meredith MA, Kral A. Cross-modal plasticity in specific auditory cortices underlies visual compensations in the deaf. *Nat Neurosci.* 2010; 13:1421–1427. [PubMed: 20935644]
- Luo H, Liu Z, Poeppel D. Auditory cortex tracks both auditory and visual stimulus dynamics using low-frequency neuronal phase modulation. *PLoS Biol.* 2010; 8:e1000445. [PubMed: 20711473]
- Maunsell JH, Gibson JR. Visual response latencies in striate cortex of the macaque monkey. *J Neurophysiol.* 1992; 68:1332–1344. [PubMed: 1432087]
- McGurk H, MacDonald J. Hearing lips and seeing voices. *Nature.* 1976; 264:746–748. [PubMed: 1012311]
- Meredith MA, Allman BL. Subthreshold multisensory processing in cat auditory cortex. *Neuroreport.* 2009; 20:126–131. [PubMed: 19057421]
- Meredith MA, Allman BL, Keniston LP, Clemo HR. Auditory influences on non-auditory cortices. *Hear Res.* 2009; 258:64–71. [PubMed: 19303926]
- Montemurro MA, Rasch MJ, Murayama Y, Logothetis NK, Panzeri S. Phase-of-firing coding of natural visual stimuli in primary visual cortex. *Curr Biol.* 2008; 18:375–80. [PubMed: 18328702]
- Nodal FR, Kacelnik O, Bajo VM, Bizley JK, Moore DR, King AJ. Lesions of the auditory cortex impair azimuthal sound localization and its recalibration in ferrets. *J Neurophysiol.* 2010; 103:1209–1225. [PubMed: 20032231]
- Panzeri S, Brunel N, Logothetis NK, Kayser C. Sensory neural codes using multiplexed temporal scales. *Trends Neurosci.* 2010; 33:111–120. [PubMed: 20045201]
- Pekkola J, Ojanen V, Autti T, Jääskeläinen IP, Möttönen R, Tarkiainen A, Sams M. Primary auditory cortex activation by visual speech: an fMRI study at 3 T. *Neuroreport.* 2005; 16:125–128. [PubMed: 15671860]

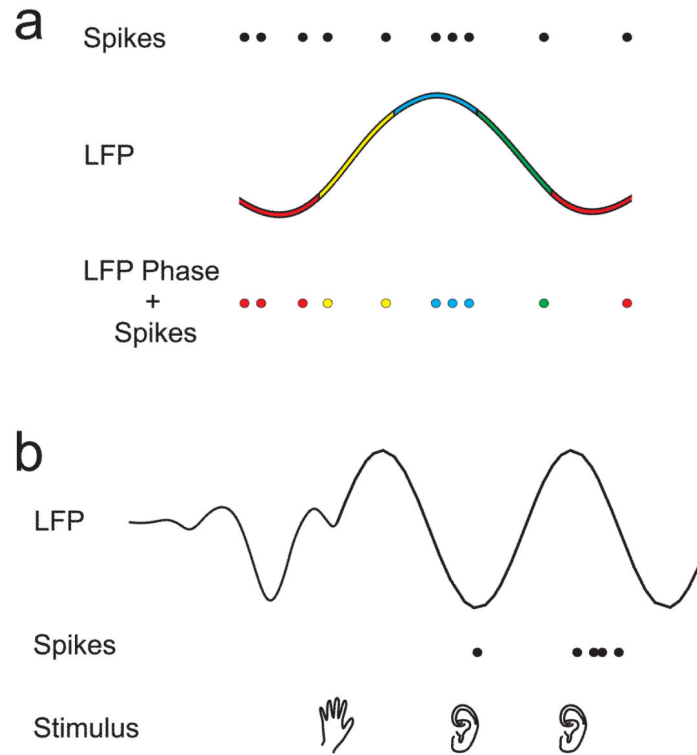
- Pouille F, Scanziani M. Enforcement of temporal fidelity in pyramidal cells by somatic feed-forward inhibition. *Science*. 2001; 293:1159–1163. [PubMed: 11498596]
- Powers AR 3rd, Hevey MA, Wallace MT. Neural correlates of multisensory perceptual learning. *J Neurosci*. 2012; 32:6263–6274. [PubMed: 22553032]
- Powers AR 3rd, Hillock AR, Wallace MT. Perceptual training narrows the temporal window of multisensory binding. *J Neurosci*. 2009; 29:12265–12274. [PubMed: 19793985]
- Recanzone GH, Guard DC, Phan ML. Frequency and intensity response properties of single neurons in the auditory cortex of the behaving macaque monkey. *J Neurophysiol*. 2000; 83:2315–2331. [PubMed: 10758136]
- Recanzone GH. Representation of con-specific vocalizations in the core and belt areas of the auditory cortex in the alert macaque monkey. *J Neurosci*. 2008; 28:13184–13193. [PubMed: 19052209]
- Schroeder CE, Mehta AD, Givre SJ. A spatiotemporal profile of visual system activation revealed by current source density analysis in the awake macaque. *Cereb Cortex*. 1998; 8:575–592. [PubMed: 9823479]
- Schürmann M, Caetano G, Hlushchuk Y, Jousmäki V, Hari R. Touch activates human auditory cortex. *Neuroimage*. 2006; 30:1325–1331. [PubMed: 16488157]
- Sekuler R, Sekuler AB, Lau R. Sound alters visual motion perception. *Nature*. 1997; 385:308. [PubMed: 9002513]
- Seltzer B, Pandya DN. Parietal, temporal, and occipital projections to cortex of the superior temporal sulcus in the rhesus monkey: a retrograde tracer study. *J Comp Neurol*. 1994; 343:445–463. [PubMed: 8027452]
- Slutsky DA, Recanzone GH. Temporal and spatial dependency of the ventriloquism effect. *Neuroreport*. 2001; 12:7–10. [PubMed: 11201094]
- Soto-Faraco S, Lyons J, Gazzaniga M, Spence C, Kingstone A. The ventriloquist in motion: illusory capture of dynamic information across sensory modalities. *Brain Res Cogn Brain Res*. 2002; 14:139–146. [PubMed: 12063137]
- Stark H, Scheich H. Dopaminergic and serotonergic neurotransmission systems are differentially involved in auditory cortex learning: a long-term microdialysis study of metabolites. *J Neurochem*. 1997; 68:691–697. [PubMed: 9003057]
- Stecker GC, Harrington IA, Macpherson EA, Middlebrooks JC. Spatial sensitivity in the dorsal zone (area DZ) of cat auditory cortex. *J Neurophysiol*. 2005; 94:1267–1280. [PubMed: 15857970]
- Stecker GC, Middlebrooks JC. Distributed coding of sound locations in the auditory cortex. *Biol Cybern*. 2003; 89:341–349. [PubMed: 14669014]
- Stein, BE., editor. *The New Handbook of Multisensory Processing*. MIT Press; Cambridge, MA: 2012.
- Sumby WH, Pollack I. Visual contribution to speech intelligibility in noise. *J Acoust Soc Am*. 1954; 26:212–215.
- Thorne JD, De Vos M, Viola FC, Debener S. Cross-modal phase reset predicts auditory task performance in humans. *J Neurosci*. 2011; 31:3853–3861. [PubMed: 21389240]
- Tian B, Reser D, Durham A, Kustov A, Rauschecker JP. Functional specialization in rhesus monkey auditory cortex. *Science*. 2001; 292:290–293. [PubMed: 11303104]
- van Wassenhove V, Grant KW, Poeppel D. Temporal window of integration in auditory-visual speech perception. *Neuropsychologia*. 2007; 45:598–607. [PubMed: 16530232]
- Vatakis, A.; Spence, C. Audiovisual temporal integration for complex speech, object-action, animal call, and musical stimuli. In: Naumer, MJ.; Kaiser, J., editors. *Multisensory Object Perception in the Primate Brain*. Springer-Verlag; Berlin Heidelberg: 2010.
- Walker KM, Bizley JK, King AJ, Schnupp JW. Multiplexed and robust representations of sound features in auditory cortex. *J Neurosci*. 2011; 31:14565–14576. [PubMed: 21994373]
- Wallace MT, Ramachandran R, Stein BE. A revised view of sensory cortical parcellation. *Proc Natl Acad Sci USA*. 2004; 101:2167–2172. [PubMed: 14766982]
- Warren JD, Griffiths TD. Distinct mechanisms for processing spatial sequences and pitch sequences in the human auditory brain. *J Neurosci*. 2003; 23:5799–5804. [PubMed: 12843284]
- Zampini M, Guest S, Shore DI, Spence C. Audio-visual simultaneity judgments. *Percept Psychophys*. 2005; 67:531–544. [PubMed: 16119399]



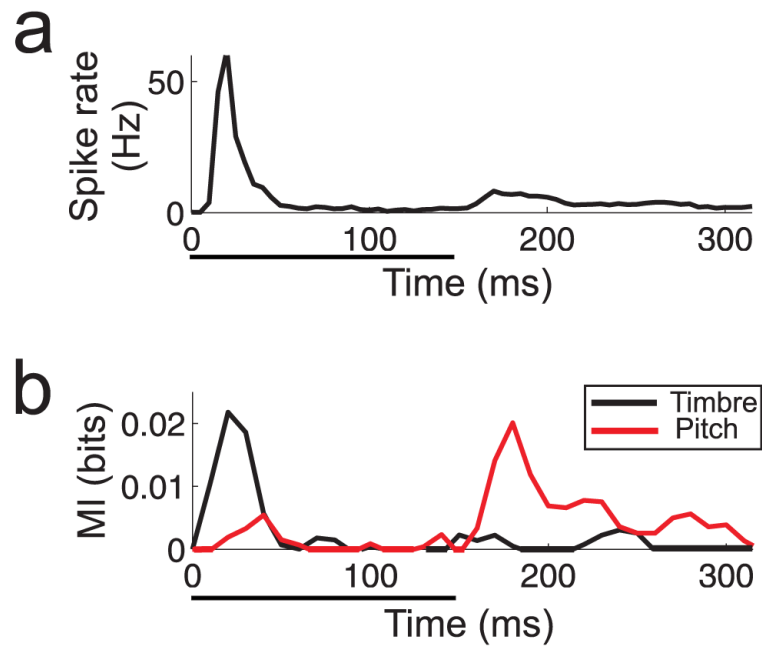
**Fig. 1.** Multisensory spatial receptive fields of a ferret auditory cortical neuron. The raster plots in the 3 left panels show the action potential responses of the neuron to 100 ms bursts of broadband noise (*far left plot*), a white light-emitting diode (*center raster plot*), or both the sound and light together (*right raster plot*). In each raster plot, the timing of stimuli are indicated by a horizontal bar at the bottom of the plot, and responses to stimuli are plotted across a range of stimulus locations in the horizontal plane (y-axis). The spike rate functions of the neuron for auditory (blue line), visual (green line), and audiovisual (red line) stimulation are summarized in the plot to the far right. In this example, pairing spatially coincident visual and auditory stimuli resulted in a more sharply tuned spatial response profile that carried significantly more information about stimulus location. Adapted from Bizley and King (2008).



**Fig. 2.** Schematic of auditory cortex in the ferret, showing organization of visual inputs. Auditory cortical areas are located on the anterior, middle and posterior ectosylvian gyri (AEG, MEG and PEG). Other abbreviations: PPc, caudal posterior parietal cortex; SSY, suprasylvian cortex; V1, primary visual cortex; V2, secondary visual cortex; sss, suprasylvian sulcus; pss, pseudosylvian sulcus; D, dorsal; R, rostral.



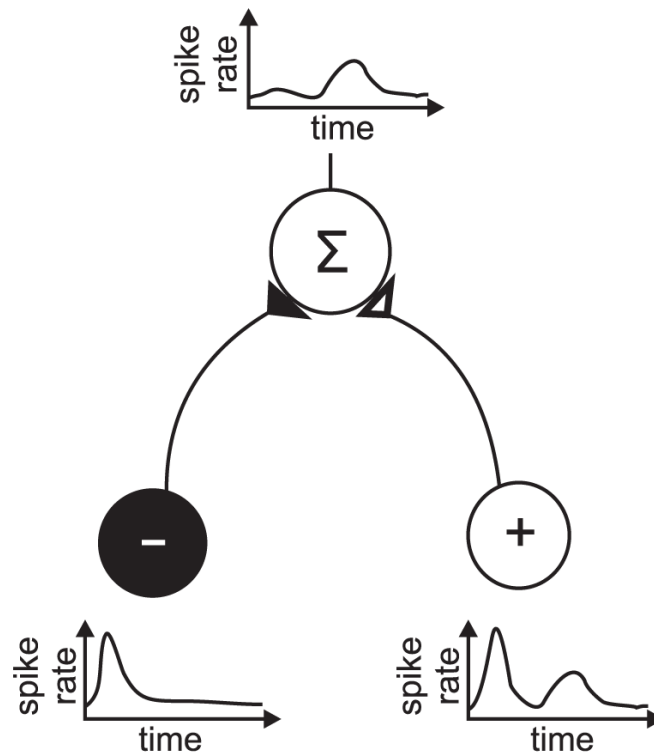
**Fig. 3.** Cartoons of 2 forms of spectral multiplexing observed in auditory cortex. **a.** Phase-of-firing multiplexing described by Kayser et al. (2009b). LFPs have been categorized into four phases, represented in color. The pattern of action potentials (top row) or local field potential (LFP) phase (middle row) alone do not provide as much information about auditory stimuli as the timing of action potentials with respect to the LFP phase in which they occur (bottom row). **b.** Phase-based-enhancement of spiking responses described by Lakatos et al. (2007). A somatosensory stimulus (hand, bottom row) resets the phase of the LFP (top row) in auditory cortex, but does not result in significant spiking activity (middle row). Spiking responses to subsequent auditory stimulation (ears, bottom row) are suppressed during troughs and enhanced during peaks of the ongoing LFP.



**Fig. 4.**

Time-division multiplexing within a single auditory cortex neuron described by Walker et al. (2011). **a.** Post-stimulus time histogram, showing the mean firing rate of the neuron across all presentations of an artificial vowel sound presented over 4 locations in horizontal space, 4 values of the fundamental frequency (“pitch”), and 4 spectral identities (“timbre”). Sound presentation time is indicated by the black horizontal bar below the plot. **b.** Mutual Information carried by the neuron for either the timbre (black line) or pitch (red line) of the sound throughout its response. Note that these values peak at different time points.





**Fig. 5.** Schematic of excitatory and inhibitory summation to achieve delayed temporal integration of excitatory signals. The target neuron (above, " $\Sigma$ ") receives input from an early inhibitory interneuron (black, "-") and both an early and late excitatory input (white, "+"), the result of which is integration of excitatory signals in the later time window only.