Chapter 6
Multisensory Processing in the Auditory Cortex

Andrew J. King, Amy Hammond-Kenny, and Fernando R. Nodal

Abstract The capacity of the brain to combine and integrate information provided by the different sensory systems has a profound impact on perception and behavior. This is especially the case for audition, with many studies demonstrating that the ability of listeners to detect, discriminate, or localize sounds can be altered in the presence of other sensory cues. For example, the availability of congruent visual stimuli can make it easier to localize sounds or to understand speech, benefits that are most apparent when auditory signals are weak or degraded by the presence of background noise. Multisensory convergence has been demonstrated at most levels of the auditory pathway, from the cochlear nucleus to the auditory cortex. This is particularly the case in extralemniscal nuclei from the midbrain upward but has also been observed in the tonotopically organized lemniscal or core projections. In addition to inheriting multisensory signals from subcortical levels, the auditory cortex receives visual and somatosensory inputs from other cortical areas. Although nonauditory stimuli can evoke spiking activity in auditory cortex, they typically modulate auditory responses. These interactions appear to provide contextual cues that signal the presence of an upcoming sound, but they can also increase the information conveyed by cortical neurons about the location or identity of sounds and may even recalibrate cortical responses when the information provided by different sensory modalities is conflicting. Identifying the neural circuitry responsible for the behavioral consequences of multisensory integration remains an area of intense investigation.

Keywords Attention · Auditory pathway · Cross-modal plasticity · Dorsal cochlear nucleus · Eye movement · Inferior colliculus · Somatosensory · Sound localization · Spectral cue · Speech · Superior colliculus · Thalamus · Ventriloquism illusion · Visual · Vocalization

A. J. King (✉) · A. Hammond-Kenny · F. R. Nodal
Department of Physiology, Anatomy and Genetics, University of Oxford, Oxford, UK
e-mail: andrew.king@dpag.ox.ac.uk; amy.hammond-kenny@merton.ox.ac.uk; fernando.nodal@dpag.ox.ac.uk

6.1 Introduction

Creating a unified sensory percept requires the integration of information from different sensory modalities. This process is traditionally viewed as occurring in two distinct phases in the brain. First, unisensory signals are processed by dedicated neural pathways, which are assumed to be largely independent and hierarchically organized. Second, once modality-specific computations have been performed, sensory information is combined and integrated in certain higher order association areas that implement different aspects of multisensory perception. In the cortex, classical multisensory areas have been described in the frontal, parietal, and temporal lobes, where their functions are thought to range from linking multiple sensory signals with the execution of particular motor actions to the merging of communication signals provided by the eyes and ears (reviewed by Cappe et al. 2012).

That sensory pathways are organized in this fashion stems from the different forms of energy (light, sound) that need to be detected. This necessitates the use of specialized transduction mechanisms for converting each form of energy into neural activity and imposes constraints on the associated neural circuits in order to overcome the differences between each sensory modality, such as the lack of spatial information at the cochlea or the differing temporal dynamics of visual and auditory processing. Furthermore, some of our perceptions, for example, the color of a flower or the pitch of someone's voice, do not have obvious equivalents in other sensory modalities. Nevertheless, it is often the case that we can identify or locate an object, such as a familiar person speaking, by using more than one of our senses. Although this cross-modal redundancy is extremely useful for perceptual stability should one set of cues disappear, such as when that person stops speaking or walks outside our field of view, sensory processing most commonly occurs in a multisensory context and the simultaneous availability of information across different modalities can have profound effects on perception and behavior.

A good example of this is provided by speech perception. If we want to understand the basis for this vital ability, it is necessary to consider not only how the brain responds to auditory information but also the motor aspects of speech production and, consequently, the associated visual articulation cues. Orofacial movements during speech production provide temporally correlated cues (Fig. 6.1; Chandrasekaran et al. 2009) that, when combined with acoustic signals, improve the detection and comprehension of speech, particularly if those signals are degraded by the presence of background sounds (Sumby and Pollack 1954; also see Grant and Bernstein, Chap. 3). The tendency to merge auditory-visual speech cues is further illustrated by the well-known McGurk illusion (McGurk and MacDonald 1976): pairing a voice articulating one syllable with a face articulating a different syllable can result in the perception of a novel token that represents a fusion of those syllables.

This work clearly indicates the capacity of the brain to integrate the informational content of auditory-visual speech. If the signals available in each modality are first processed independently and only subsequently combined at a specialized integration stage, one might expect the neural basis for the influence of vision on

Fig. 6.1 Visual and auditory statistics of human speech. (**A**) *Top*, example facial gestures at different frames from a video of a speaker uttering a sentence, with the *red ellipses* below each frame representing the area of the mouth opening. *Bottom:* graph shows the estimated area for each mouth contour in pixels squared as a function of time in seconds. *Numbers* refer to corresponding frames in the video. *Arrows* point to specific frames in the time series depicting different size of mouth opening. (**B**) Variation in the area of the mouth opening (*black*) and the broadband auditory envelope (*orange*) for a single sentence from a single subject as a function of time in seconds. (**C**) Heat map illustrating the robust coherence between the mouth opening and auditory signal as a function of both spectral frequency band and temporal modulation frequency for 20 subjects. *Dashed–line rectangle*, region of maximal coherence between the visual and auditory signals. Adapted from Chandrasekaran et al. (2009), with permission
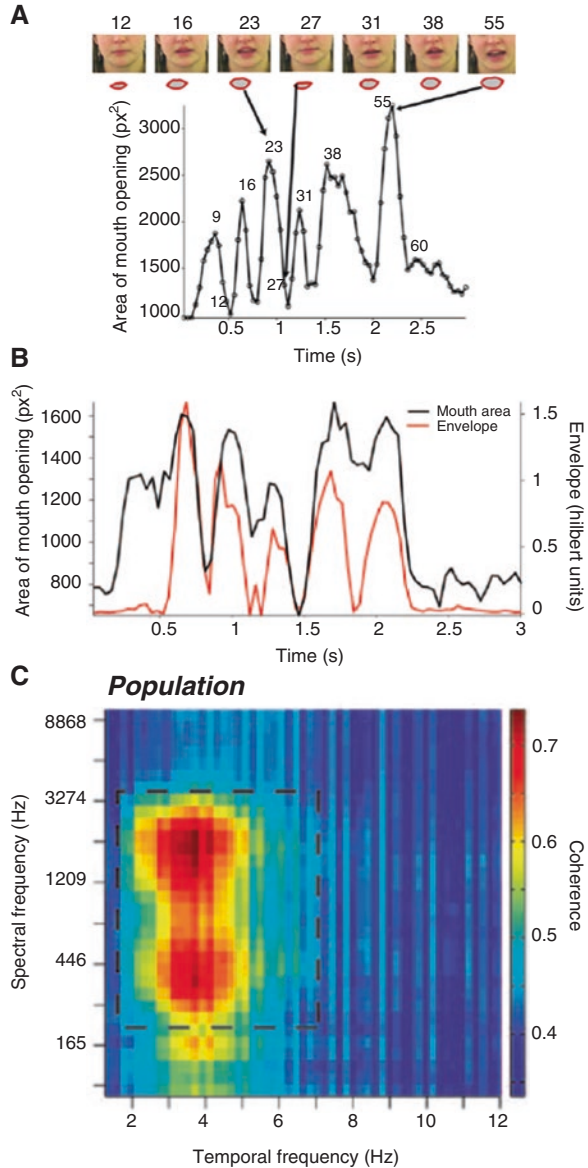
auditory speech intelligibility to reside in higher order speech-related areas such as the superior temporal sulcus (STS; see Beauchamp, Chap. 8). Although this is undoubtedly the case (Sekiyama et al. 2003; McGettigan et al. 2012), there is growing evidence that auditory and visual speech signals also interact as early as the primary auditory cortex (Schroeder et al. 2008; Okada et al. 2013). Furthermore, both cortical and subcortical auditory brain regions have been implicated in the various cross-modal effects that have been described for other dimensions of auditory

perception. Indeed, it is a general property of sensory systems that the availability of congruent multisensory cues can result in faster responses as well as improvements in the ability to detect, discriminate, or localize stimuli (Murray and Wallace 2012). It is therefore important to consider where and how those interactions take place as well as the nature of the information provided by the "nondominant" modality if we are to understand the impact of vision and other sensory modalities on auditory processing and perception.

This chapter considers these issues in the context of the auditory pathway as a whole but with a focus on visual and somatosensory influences on the auditory cortex and the implications of these effects for its primary role in hearing. Although similar questions can be asked about the functional significance of multisensory influences on processing in the visual or somatosensory cortex, the auditory cortex has been at the vanguard of research in this area. Consequently, these studies have the potential not only to improve our understanding of the computations performed by auditory cortical neurons but also to reveal general principles of how multisensory interactions influence perception and behavior.

6.2   Multisensory Versus Auditory Brain Areas

Conceptually, it is difficult to classify a given brain area as unisensory if stimuli belonging to different sensory modalities can influence the activity of the neurons found there. However, multisensory influences take different forms, ranging from a change in action potential firing in response to more than one type of sensory stimulus to cross-modal modulation of the spiking responses to one modality even if the other modality cues are by themselves ineffective in driving the neurons (Fig. 6.2). In the case of the auditory cortex, there is considerable evidence for modulatory effects of nonauditory inputs on responses to sound. These interactions have been found to be particularly prevalent in functional imaging experiments, which also show that visual cues alone can activate certain parts of the auditory cortex in humans (Calvert et al. 1997; Pekkola et al. 2005) and nonhuman primates (Kayser et al. 2007). Similar results have been obtained using electrophysiological measurements, with local field potential recordings demonstrating widespread effects of visual or somatosensory stimuli on sound-evoked responses in both primary and secondary areas of the auditory cortex (Ghazanfar et al. 2005; Kayser et al. 2008).

Multisensory convergence in the auditory cortex appears to be more limited, however, when the spiking responses of individual neurons or small clusters of neurons are considered. This may be because cortical local field potentials primarily reflect summed synaptic currents and their accompanying return currents and therefore capture the input activity of the neurons (Einevoll et al. 2013). Nevertheless, multisensory influences on the spiking behavior of auditory cortical neurons again range from a change in firing rate when otherwise ineffective stimuli are paired with a sound to responses evoked directly by visual or somatosensory
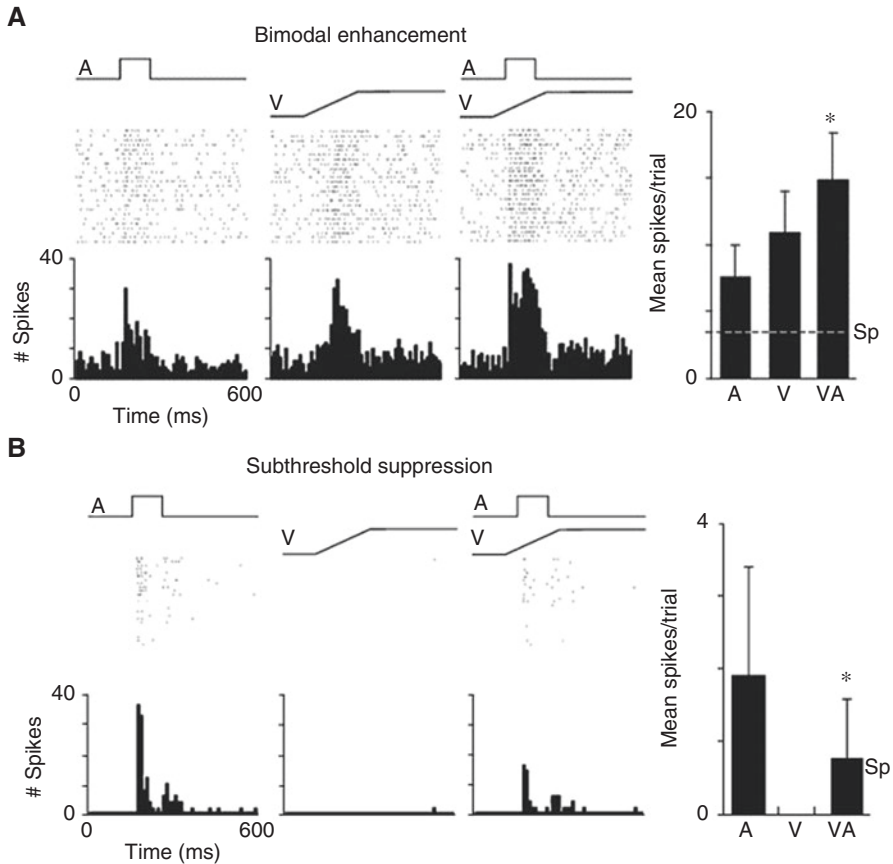
Fig. 6.2 Multisensory responses of neurons recorded in the auditory field of the anterior ectosylvian sulcus (FAES) of a cat to auditory, visual, and combined auditory-visual stimulation. (**A**) Example of a neuron that gave a suprathreshold spiking response to both auditory (*square wave A; top left*) and visual (*ramp V; top center*) stimuli presented alone and that generated a significantly enhanced response when the same stimuli were combined (*square wave and ramp together AV; top right*). (**B**) different FAES neuron that was activated by the auditory (*top left*) but not the visual stimulus (*top center*); in this case, presenting the two stimuli together led to a suppression of the auditory response (*top right*). In both (**A**) and (**B**), responses are shown in the form of raster plots (where each dot represents a spike with the response to multiple stimulus presentations arranged vertically; *center*), the corresponding peristimulus time histograms (*bottom*), and bar charts of the mean ± SD evoked activity for each stimulus type (*right*). *$P < 0.05$, paired *t*-test. *Sp*, spontaneous activity level. Adapted from Meredith and Allman (2009), with permission

stimuli (Fu et al. 2003; Bizley et al. 2007). This apparent continuum of multisensory properties could reflect differences in the way sensory inputs converge on neurons either in the cortex itself (Fig. 6.3; Clemo et al. 2012) or at an earlier level in the processing hierarchy.

It is unclear what functions spiking responses to nonauditory stimuli in auditory cortex might serve, unless they convey signals that can be processed as if they were
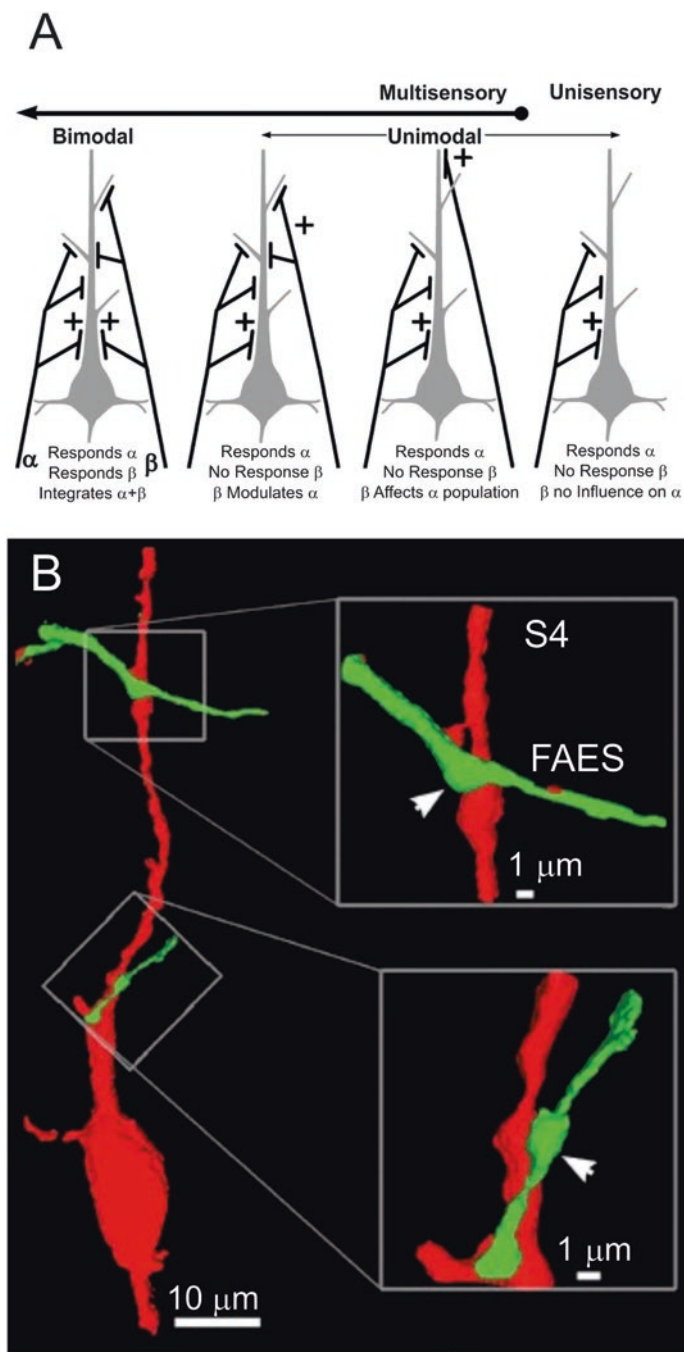
Fig. 6.3 Putative patterns of synaptic connectivity underlying the range of multisensory inter-
actions observed in the brain. (**A**) Neurons (*gray*) are depicted receiving afferent inputs (*black*)
from either one (*far right*) or two sensory modalities (α *and* β*; three left cases*). The simplified

auditory in origin. Indeed, it is possible that they are simply a consequence of synaptic inputs rising above threshold. In ferrets (*Mustela putorius*), the incidence of these nonauditory spiking responses increases between primary and high-level auditory cortical areas (Bizley et al. 2007), which likely reflects the greater density of projections to the latter from extralemniscal thalamic nuclei (Winer and Lee 2007) and from other sensory cortices (Bizley et al. 2007). Consequently, the relative proportion of neurons that receive subthreshold, modulatory inputs versus suprathreshold inputs that are capable of driving spiking activity is likely to be indicative of a progression from areas with a unisensory primary function to those more involved in merging independent inputs from different sensory modalities.

Another aspect to consider is the expected neural output of multisensory integration and to what extent it might vary in different parts of the brain. Electrophysiological recordings from multisensory neurons in the superior colliculus (SC) have led to the identification of several key principles by which different sensory inputs interact to govern the spiking activity of these neurons (King and Palmer 1985; Wallace et al. 1998). The SC is characterized by the presence of topographically aligned visual, auditory, and somatosensory maps. In such an organizational structure, the different modality signals arising from a particular location, and therefore potentially from the same source, can be represented by the same neurons. The strongest enhancement of the unisensory responses of SC neurons has been shown to occur when the component stimuli are weakly effective in eliciting a response and when those stimuli occur at approximately the same time and originate from the same region of space. By contrast, pairing strongly effective unisensory stimuli typically produces little or no enhancement as do multisensory signals that are widely separated in time or space. Indeed, this can result in a reduction of the firing rate elicited by unisensory stimulation. That these principles operate clearly makes sense because the relative timing and location of sensory signals are important factors in determining whether they belong to the same object and should therefore be bound together or to different objects.

Similar principles of multisensory integration have been observed in cortical neurons (Stein and Wallace 1996) and for various behavioral tasks, including the sensory-guided orienting responses with which the SC is likely to be involved (Stein et al. 1988; Bell et al. 2005). However, attempts to apply them to population and more indirect measures of neural activity, such as functional magnetic resonance imaging (fMRI), have turned out to be less straightforward (Stevenson et al. 2014). Moreover, it is an oversimplification to assume that improved perceptual abilities necessarily result from increased neural activity. Although functions in which the

Fig. 6.3 (continued) convergence patterns vary among the multisensory neurons so that although modality α evokes a spiking response in each case, modality β can result in a continuum of effects from producing a spiking response to modulating the response to α at the level of either the neuron or the neuronal population. (**B**) *Left*, confocal images of a neuron in the cat higher level somatosensory cortical area S4 (*red*) contacted by axons that originated in auditory FAES (*green*). *Right*, each axodendritic point of contact is enlarged to show the putative bouton swelling (*arrow*). Adapted from Clemo et al. (2012), with permission

auditory cortex plays a pivotal role, including speech perception and sound localization, can be enhanced by the availability of other sensory cues, cortical neurons frequently exhibit cross-modal suppression. Thus, the response to the primary auditory stimulus is often reduced when combined with visual or somatosensory stimuli (Bizley et al. 2007; Meredith and Allman 2009). Furthermore, some studies have stressed the contribution of changes in the timing rather than the magnitude of the responses in the presence of multisensory stimuli (Chandrasekaran et al. 2013).

Other studies in which single neuron recordings were made from the auditory cortex have also highlighted the importance of quantifying multisensory interactions in ways that go beyond simple changes in the number of action potentials evoked. Application of information theoretic analyses to the spike discharge patterns recorded from neurons in the auditory cortex has revealed that visual cues can enhance the reliability of neural responses and hence the amount of information transmitted even if the overall firing rate either does not change or is suppressed (Bizley et al. 2007; Kayser et al. 2010). This finding is consistent with earlier work demonstrating that the location and identity of sounds can be encoded by the temporal firing pattern of auditory cortical neurons (Furukawa and Middlebrooks 2002; Nelken et al. 2005). Moreover, as first highlighted by the principle of inverse effectiveness in the superior colliculus (Meredith and Stein 1986), it is important to take response magnitude into account when characterizing the effects of multisensory stimulation on neuronal firing patterns. Indeed, Kayser et al. (2010) showed that multisensory stimulation can have opposite effects on the magnitude and reliability of cortical responses according to how strong the responses are to sound alone, with these opposing modes of multisensory integration potentially having different functions. Thus, enhanced responses to weakly effective stimuli are likely to facilitate the detection of near-threshold events, whereas suppressed but more reliable responses may be particularly relevant for sound discrimination at stimulus levels that are more typical of everyday listening.

Population measures of auditory cortical activity in human (Luo et al. 2010; Thorne et al. 2011) and nonhuman primates (Lakatos et al. 2007; Kayser et al. 2008) also indicate that nonauditory inputs can modulate the phase of low-frequency oscillatory activity in the auditory cortex. This is thought to alter the excitability of the cortex, increasing the amplitude of responses evoked by temporally correlated auditory inputs and thereby providing another way in which visual or other sensory inputs can modulate neuronal responses to accompanying sounds without necessarily evoking spiking activity. Indeed, it has even been proposed that phase resetting may represent one of the "canonical" operating principles used by the brain to integrate different types of information (van Atteveldt et al. 2014; see Keil and Senkowski, Chap. 10).

Together, these findings stress the importance of investigating multisensory interactions at multiple levels, from the activity of individual neurons to more population-based signals, including local field potentials (LFPs), EEG, MEG, and fMRI, and of employing the appropriate metrics in each case to quantify the magnitude and nature of integration. This is particularly important for making sense of how nonauditory inputs influence the auditory cortex without altering its fundamental role in hearing.

6.3 Nonauditory Inputs at Different Levels of the Auditory Pathway

In considering the potential role of multisensory interactions in the auditory cortex, it is essential to examine the origin of nonauditory inputs as well as their entry point into the auditory pathway. This can provide insight into the type of stimulus-related information those inputs convey and the extent to which the signals provided have already been processed and integrated by the time they reach the cortex.

At the subcortical level, nonauditory inputs have been identified in most of the main relay stations of the ascending auditory pathway. The complexity of this network, which includes more levels of subcortical processing than in other sensory modalities, makes it challenging to determine the role of these inputs. Furthermore, we currently have a poor understanding of the extent to which the multisensory interactions in the auditory cortex are inherited from the thalamus and therefore reflect bottom-up processing or arise from the convergence of inputs from other cortical areas.

Before discussing where nonauditory influences are found along the auditory pathway, it is first important to consider briefly the route by which acoustic information passes from the cochlea to the cortex (Fig. 6.4). Auditory nerve fibers transmit information from the cochlea to the cochlear nucleus (CN) in the brainstem, which is the first relay for the ascending auditory pathway. The CN comprises three subdivisions (anteroventral, posteroventral, and dorsal), within which are found several different neuron types that differ in their anatomical location, morphology, cellular physiology, synaptic inputs, and spectrotemporal response properties. The output from the CN takes the form of multiple, parallel ascending pathways with different targets. One of these is the superior olivary complex, where sensitivity to binaural localization cues emerges. The various ascending tracts then innervate the nuclei of the lateral lemniscus and all converge in the inferior colliculus (IC) in the midbrain, which therefore provides a relay for the outputs from each of the brainstem auditory centers. The IC comprises a central nucleus, which is surrounded by a dorsal cortex (DCIC), a lateral (or external) cortex (LCIC) and a rostral cortex, which can be distinguished by their connections and response properties. The IC in turn delivers much of the auditory input to the SC, which, as we have seen, is a major site for the integration of multisensory spatial information, and also projects to the medial geniculate body (MGB) in the thalamus, which serves as the gateway to the auditory cortex.

Classically, the ascending auditory pathway is thought to comprise a core or lemniscal projection, which is characterized by a precise tonotopic organization at each level from the CN to the primary auditory cortical fields. In addition, the extralemniscal projection includes parts of the IC, the MGB, and a belt of the auditory cortex surrounding the core tonotopic fields (Fig. 6.4). The defining features of neurons in extralemniscal areas are that they tend to show broader frequency tuning than those in the lemniscal projection and their tonotopic organization is less well defined or even nonexistent. Furthermore, they often receive inputs from other sensory modalities.
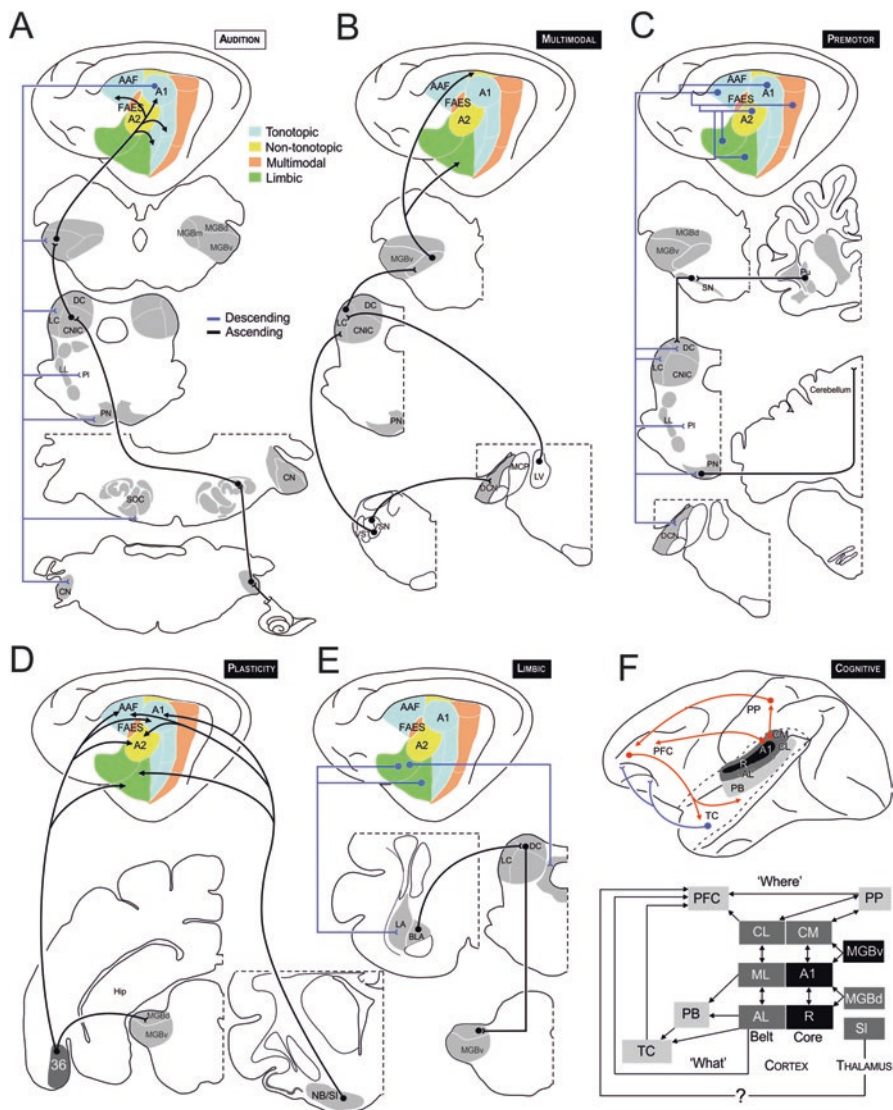
Fig. 6.4 Ascending and descending pathways of the cat auditory cortex with some of the main ascending (*black*) and descending (*blue*) connections shown according to their putative functional role and/or nature of the information transmitted. (**A**) Principal connections of the tonotopic lemniscal pathway. (**B**) Ascending connections in the extralemniscal pathway, highlighting auditory brain areas that receive projections from other sensory systems. (**C**) Descending cortical projections to premotor brain areas that participate in vocalization production and other motor functions. (**D**) Ascending connections associated with plasticity in the auditory cortex because of their cholinergic nature. (**E**) Descending cortical connections to the limbic system that are thought to contribute to emotional responses. (**F**) Putative cognitive streams involved in sound identification and localization (What and Where, respectively) described in macaques on the basis of the connectivity between the auditory cortex and prefrontal cortex. *A1* primary auditory cortex; *A2*, secondary

   Although the functional significance of multisensory convergence within the subcortical auditory pathway is often unclear, there are instances where information from other sensory modalities makes an important contribution to the "unisensory" role of the neurons in question. Perhaps the best example is to be found in the dorsal CN (DCN). The complex excitatory and inhibitory interactions displayed by type IV neurons in the DCN of the cat (*Felis catus*) allow these neurons to signal the presence of spectral notches that are generated by the directional filtering properties of the external ear (Yu and Young 2000). Together with the finding that lesions of the pathway by which DCN projection neurons reach the IC result in impaired head orienting responses to broadband sounds (May 2000), this points to a likely role for this nucleus in sound localization. But cats are able to move their ears, shifting the locations at which spectral notches occur relative to the head. Consequently, information about the ongoing position of the pinnae is required to maintain accurate sound localization. This appears to be provided by muscle proprioceptors located in and around the pinna of the external ear, with the DCN combining monaural acoustical cues to sound source direction with somatosensory inputs about the orientation of the pinna (Kanold and Young 2001). More recent work in rats (*Rattus norvegicus*) indicates that multisensory computations in the DCN may also help distinguish moving sound sources from the apparent movement produced by motion of the head, suggesting that the integration of auditory and vestibular inputs helps to create a surprisingly sophisticated representation of spatial information at this early stage of auditory processing (Wigderson et al. 2016).

   There is also evidence to suggest that somatosensory projections to the DCN are involved not only in sound localization but also in suppressing the effects of self-generated noises on the central auditory system, such as those produced by vocalizing and masticating (Shore and Zhou 2006). In support of this adaptive filter function, paired stimulation of the auditory and trigeminal nerves shows that neurons in the DCN are capable of multisensory integration and, more importantly, that the majority of multisensory interactions elicited are suppressive (Shore 2005; Koehler and Shore 2013). Interestingly, a related effect has also been described in tinnitus sufferers, whereby some individuals are able to modulate the loudness of their tinnitus by activating the trigeminal system using orofacial stimulation (Pinchoff et al. 1998). It has been suggested that the tinnitus percept arises, at least in part, from increased spontaneous activity in the DCN (Kaltenbach 2007).

Fig. 6.4 (continued) auditory cortex; *AAF*, anterior auditory field; *AL*, anterolateral area of the belt auditory cortex; *BLA*, basolateral nucleus of the amygdala; *CM*, caudomedial area of the belt auditory cortex; *CN*, cochlear nucleus; *CNIC*, central nucleus of the inferior colliculus (IC); *CL*, caudolateral area of the belt auditory cortex; *DC*, dorsal cortex of the IC; *DCN*, dorsal CN; *Hip*, hippocampus; *MGBv*, -*d*, and -*m*, medial geniculate body (ventral, dorsal, and medial divisions, respectively); *LA*, lateral nucleus of the amygdala; *LC*, lateral cortex of the IC; *LL*, lateral lemniscus; *NB/SI*, nucleus basalis/substantia innominata; *PB*, parabelt auditory cortex; *PFC*, prefrontal cortex; *PP*, posterior parietal cortex; *Pl*, paralemniscal area; *PN*, pontine nuclei; *Pu*, putamen; *R*, auditory cortical area; *SI*, suprageniculate nucleus, lateral part; *SN*, substantia nigra; *SOC*, superior olivary complex; *TC*, temporal cortex; *36*, cortical area 36. Adapted from Winer and Lee (2007), with permission

Therefore, it is conceivable that in addition to suppressing neural responses to self-generated sounds, somatosensory inputs to the DCN may reduce abnormal activity associated with phantom sounds, highlighting the therapeutic potential of harnessing somatosensory inputs to the DCN to alleviate tinnitus.

Beyond the DCN, responses to somatosensory stimulation have also been described in the auditory midbrain and, particularly, in the LCIC. This activity again likely reflects the influence of inputs from multiple sources, which include the dorsal column nuclei, the spinal trigeminal (Sp5) nucleus, and the somatosensory cortex (Shore and Zhou 2006; Lesicko et al. 2016). But whereas somatosensory inputs to the DCN originate principally from the pinnae, in accordance with their presumed role in the processing of spectral localization cues, those to the IC suggest a diffuse input from the entire body (Aitkin et al. 1981). Thus, somatosensory responses in the IC are not just inherited from the DCN and may serve to suppress the effects of self-generated noises regardless of their spatial origin.

The first responses to visual stimulation in the auditory pathway appear to be found in the midbrain, and recordings in behaving monkeys (*Macaca mulatta*) have reported that the prevalence of visual influences on IC neurons may be surprisingly high (Porter et al. 2007). This is supported by the presence of sparse inputs from the retina to the DCIC (Morin and Studholme 2014) and from the visual cortex to various subdivisions of the IC (Cooper and Young 1976; Gao et al. 2015). However, the primary source of visual input to the auditory midbrain, and potentially therefore to other parts of the auditory pathway, appears be the SC. Indeed, in ferrets, the nucleus of the brachium of the IC (Doubell et al. 2000) and the LCIC (Stitt et al. 2015) have reciprocal connections with the SC. This provides a source of retinotopically organized input into different regions of the IC, which may play a role in coordinating and updating the alignment of maps of visual and auditory space in the SC (Doubell et al. 2000; Stitt et al. 2015). Potentially related to this is the finding that auditory responses in the monkey IC are modulated by changes in gaze direction (Groh et al. 2001; Zwiers et al. 2004). If accurate gaze shifts are to be made to auditory targets, it is essential that eye position signals are incorporated in the brain's representation of auditory space (see also Willett, Groh, and Maddox, Chap. 5). This is well-known to be the case in the SC (Jay and Sparks 1984; Hartline et al. 1995), and these findings indicate that this process most likely begins in the IC. Beyond a role in spatial processing, nonauditory inputs to the IC could contribute to other aspects of multisensory behavior. A single case study of a human patient with a unilateral IC lesion reported a weaker McGurk effect for audiovisual speech stimuli in the contralesional hemifield (Champoux et al. 2006), although it is unclear whether this reflects multisensory processing in the IC itself.

The thalamus is the final subcortical level in the auditory pathway at which multisensory processing occurs. In addition to inheriting nonauditory inputs via ascending projections from earlier stages in the pathway, the medial division of the MGB (MGBm) is innervated by the spinal cord, whereas the dorsal nucleus of the MGB (MGBd) and the suprageniculate nucleus receive inputs from the SC (Jones and Burton 1974; Katoh and Benedek 1995). An added complication when discussing multisensory processing in the thalamus is that we need to consider not only those

subdivisions comprising the auditory thalamus itself (e.g., the lemniscal ventral nucleus of the MGB (MGBv) and the extralemniscal MGBm and MGBd), but also those subdivisions regarded as higher order or multisensory, such as the pulvinar, which project to and receive inputs from auditory as well as other cortical areas (de la Mothe et al. 2006a; Scott et al. 2017). A detailed description of these projections is beyond the scope of this chapter (see Cappe et al. 2012 for a review). However, their existence is important to note, given that they provide a potential route for transferring information between different cortical areas, including those belonging to different sensory modalities (Rouiller and Welker 2000; Sherman 2016). Moreover, cortico-thalamo-cortical circuits can also involve the primary sensory thalamus. Thus, visual and whisker signals are combined in the ventral posterior medial region of the thalamus in mice (*Mus musculus*) (Allen et al. 2017), whereas activation of the primary somatosensory cortex in this species can alter the activity of neurons in the MGB (Lohse et al. 2017).

6.4   Origins of Visual and Somatosensory Inputs to the Auditory Cortex

The studies discussed so far show that multisensory information is incorporated at most stages along the ascending auditory pathway, with nonauditory inputs primarily, but not exclusively, targeting extralemniscal regions. Therefore, at the cortical level, it seems reasonable to expect that nonauditory influences will be most apparent in the cortical belt areas because of their extralemniscal inputs, and this has been confirmed by anatomical and physiological experiments in a range of species (e.g., Bizley et al. 2007). In addition to its subcortical origin, however, multisensory convergence in the auditory cortex has been shown to result from inputs from other sensory as well as higher level association cortical areas.

Anatomical tracing studies have identified direct corticocortical connections between different sensory areas in several species. As with subcortical levels of the auditory pathway, inputs from visual and somatosensory cortical areas are distributed primarily to noncore parts of the auditory cortex, such as the caudomedial belt areas in marmosets (*Callithrix jacchus*; de la Mothe et al. 2006b) and macaque monkeys (Falchier et al. 2010) or the fields on the anterior and posterior ectosylvian gyri in ferrets (Bizley et al. 2007). This is consistent with the greater incidence of multisensory neurons in those regions and with the often nonlemniscal nature of their auditory response properties. Nevertheless, the activity of neurons in the core auditory cortex can be modulated and sometimes even driven by nonauditory inputs. In the ferret, for example, around 20% of neurons in the core areas, the primary auditory cortex and the anterior auditory field, were shown to be sensitive to visual (Bizley et al. 2007) or tactile (Meredith and Allman 2015) stimulation. Although sparse projections from primary or secondary sensory areas were observed in these studies, the greatest proportion of retrograde labeling following tracer injections in

the core auditory cortex was found in visual area 20 (Bizley et al. 2007) and in the rostral suprasylvian sulcal somatosensory area (Meredith and Allman 2015). This would suggest that core auditory areas in ferrets are mainly innervated by higher order visual and somatosensory cortical areas. Direct connections between A1 and other primary and secondary sensory cortical areas have also been described in rodents (Fig. 6.5; Budinger et al. 2006; Stehberg et al. 2014). Similarly, in marmosets, the core auditory cortex is innervated by the secondary somatosensory cortex and by the STS (Cappe and Barone 2005), whereas other studies in primates suggest that nonauditory influences on A1 most likely originate from the thalamus as well as from multisensory association areas like the STS (Smiley and Falchier 2009).

Most of the anatomical studies have used retrograde tracer injections to reveal the origins of projections to the auditory cortex. Although this approach does not provide a clear picture of the extent and the laminar distribution of the terminal fields in the auditory cortex, it is possible to infer something about the nature of the inputs from the laminar origin of the projection neurons. Thus, feedforward corticocortical projections typically originate in the supragranular layers and terminate in granular layer IV, whereas feedback corticocortical projections are more likely to originate in the infragranular layers and to terminate in the supragranular and infragranular layers. After retrograde tracer injections into A1, labeled cells were found predominantly in the infragranular layers of the projecting cortices (Cappe and Barone 2005; Budinger et al. 2006). This suggests that the core auditory cortex receives mainly feedback projections from other cortical areas and is consistent with physiological evidence in monkeys that somatosensory inputs target the supragranular layers and have a modulatory influence on A1 activity (Lakatos et al. 2007). However, feedforward projections to the auditory cortex cannot be excluded because several studies have also reported retrogradely labeled cells in the supragranular layers of other cortical areas (Cappe and Barone 2005; Budinger et al. 2006).

The relative contributions of thalamocortical and corticocortical projections to multisensory processing in the auditory cortex are poorly understood. However, Budinger et al. (2006) estimated that approximately 60% of nonauditory inputs to gerbil A1 originate subcortically, with the remaining 40% arising from other sensory or multisensory cortical areas. It is therefore clear that a hierarchy of multisensory

Fig. 6.5 (continued) experiments. *Numbers next to the arrows* connecting the cortical areas represent the number of labeled cells found in the supragranular layers minus the number in the infragranular layers divided by the total of labeled cells; positive values indicate putative feedforward projections and negative values indicate putative feedback projections. Although the strongest connections to the primary sensory cortices come from their modality-specific thalamic nuclei, crossmodal inputs arise from other sensory cortices and the (extralemniscal) thalamus. *DLG*, dorsal lateral geniculate nucleus; *LD*, laterodorsal thalamic nucleus; *LP*, lateral posterior thalamic nucleus; *MGBmz*, MGB marginal zone; *Po*, posterior thalamic nuclear group; *S1*, primary somatosensory cortex; *Sg*, suprageniculate nucleus; *V1*, primary visual cortex; *VL*, ventrolateral thalamic nucleus; *VLG*, ventral lateral geniculate nucleus; *VPL*, ventral posterolateral thalamic nucleus; *VPM*, ventral posteromedial thalamic nucleus. Adapted from Henschke et al. (2015), with permission
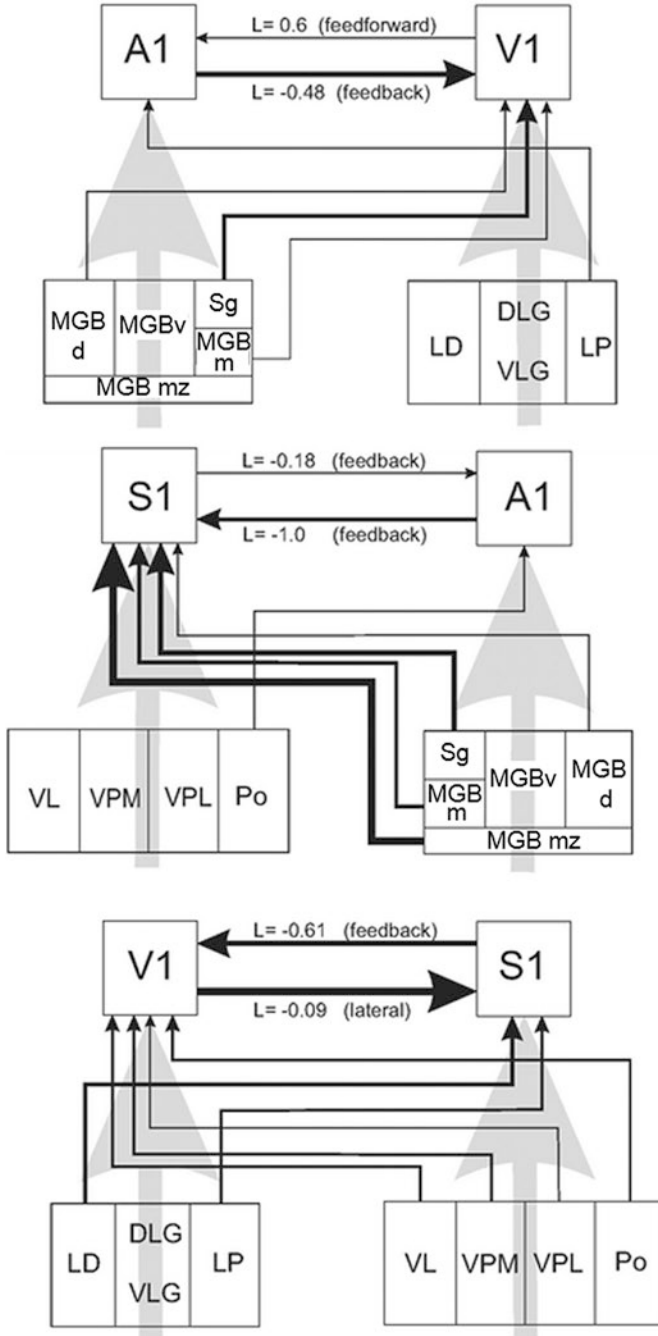
Fig. 6.5 Summary of the direct thalamocortical and corticocortical connections of the primary auditory, visual, and somatosensory cortices in the Mongolian gerbil (*Meriones unguiculatus*). *Thickness of the lines* indicates the strength of the connections as revealed by retrograde tracing

processing exists within the auditory pathway and that the auditory cortex in particular is likely to be involved in various functions that depend on the integration of information across different sensory modalities.

6.5 Functional Significance of Multisensory Interactions in the Auditory Cortex

Because there are so many subcortical and cortical sources of nonauditory inputs in the auditory pathway, it is challenging to pinpoint specific functions for the cross-modal influences that can be observed at the level of the cortex. Indeed, establishing a causal relationship between multisensory interactions at the neural and behavioral levels is particularly difficult because this field of research has yet to benefit to any great degree from the experimental approaches, such as optogenetics, that are now available for interrogating the functions of specific neural circuits (Olcese et al. 2013; Wasserman et al. 2015).

Nevertheless, insights into what those functions might be can be obtained by knowing the sources of input to particular auditory cortical areas and, of course, by measuring how the responses of the neurons change in the presence of stimuli belonging to other sensory modalities. In addition to amplifying the responses of auditory cortical neurons, particularly to relatively weak sounds, visual stimuli have been shown to induce more specific effects on the sensitivity and even the selectivity of these neurons. As discussed in Sect. 6.1, speech perception can be profoundly influenced by the talker's facial gestures, with studies in macaque monkeys demonstrating that neural responses to conspecific vocalizations are enhanced when accompanied by a video clip of an animal vocalizing but not when paired with a disk presented to mimic the opening of the mouth (Ghazanfar et al. 2005; Ghazanfar 2009). Similarly, by pairing complex naturalistic audiovisual stimuli, including videos and the accompanying sounds of conspecific animals, Kayser et al. (2010) found that the information gain in the auditory cortical responses was reduced when the auditory and visual cues were no longer matched in their dynamics or semantic content.

These visual influences have been measured in different auditory cortical areas, including A1. The complexity of the visual information involved in interpreting articulation cues makes it unlikely that the auditory cortex receives this information directly from early visual cortices. Instead, simultaneous recordings in the auditory cortex and STS showed that spiking activity in the auditory cortex was coordinated with the oscillatory dynamics of the STS (Ghazanfar et al. 2008). Thus, in the case of communication signals, the integration of multisensory information in the auditory cortex likely depends, at least in part, on top-down inputs from this area of association cortex and probably also from other cortical areas that have been shown to be entrained by lip movements during speech (Park et al. 2016).

The other major area where the functional significance of cross-modal interactions in the auditory cortex is starting to become clear is sound localization. An intact auditory cortex is required for normal spatial hearing, and inactivation studies in

cats suggest that this reflects the contribution of A1 plus certain higher level auditory cortical fields, such as the posterior auditory field (PAF; Malhotra and Lomber 2007). Studies such as these have contributed to the notion that segregated cortical processing streams exist for different auditory functions (Fig. 6.4F). Although the extent to which a division of labor exists across the auditory cortex in the processing of different sound features remains controversial (Schnupp et al. 2011; Rauschecker 2018), these findings raise the possibility that the way nonauditory stimuli influence the processing of spatial and nonspatial sound properties may be area specific.

Studies in ferrets have provided some support for this hypothesis. As expected from the extensive subcortical processing that takes place in the auditory pathway, neurons across different auditory cortical fields encode both spatial and nonspatial sound features. However, neurons located in the auditory fields located on the posterior ectosylvian gyrus in this species are more sensitive to stimulus periodicity and timbre than to spatial location (Bizley et al. 2009). In keeping with a potentially greater role in stimulus identification, this region receives inputs from areas 20a and 20b, which have been implicated in visual form processing (Manger et al. 2004). Conversely, neurons that are most informative about the azimuthal location of auditory, visual or paired auditory-visual stimuli are found on the anterior ectosylvian gyrus (Bizley and King 2008), which is innervated by a region of extrastriate visual cortex that has been implicated in spatial processing (Fig. 6.6; Philipp et al. 2006; Bizley et al. 2007).

The interpretation of these results needs to be treated with some caution because relatively little research has so far been carried out on higher level visual or auditory cortical fields in ferrets, so a detailed understanding of the functions of these areas is not yet available. However, the cross-modal reorganization observed following deafness is consistent with the notion that visual inputs target auditory cortical areas with related functions. Relative to hearing animals, congenitally deaf cats exhibit superior visual localization in the peripheral field and lower movement detection thresholds (Lomber et al. 2010). Cooling PAF, one of the key auditory cortical fields involved in spatial hearing, produced a selective loss of this enhanced visual localization, whereas deactivating the dorsal zone of the auditory cortex raised visual motion detection thresholds to values typical of hearing animals (Fig. 6.7). These findings therefore suggest that cross-modal plasticity occurs in cortical regions that share functions with the nondeprived sensory modality.

In keeping with the effects of matching naturalistic auditory-visual stimuli in nonhuman primates, the presence of spatially congruent visual stimuli produced an overall gain in the spatial information available in the responses of ferret auditory cortical neurons (Fig. 6.8; Bizley and King 2008). However, these effects were found to vary from neuron to neuron, and the largest proportion of neurons that showed an increase in transmitted spatial information when visual and auditory stimuli were presented together was actually found in the posterior suprasylvian field, where sensitivity to sound periodicity and timbre is most pronounced.

Although these studies have shown that information coding in the auditory cortex can be enhanced by the availability of matching visual cues, relatively few
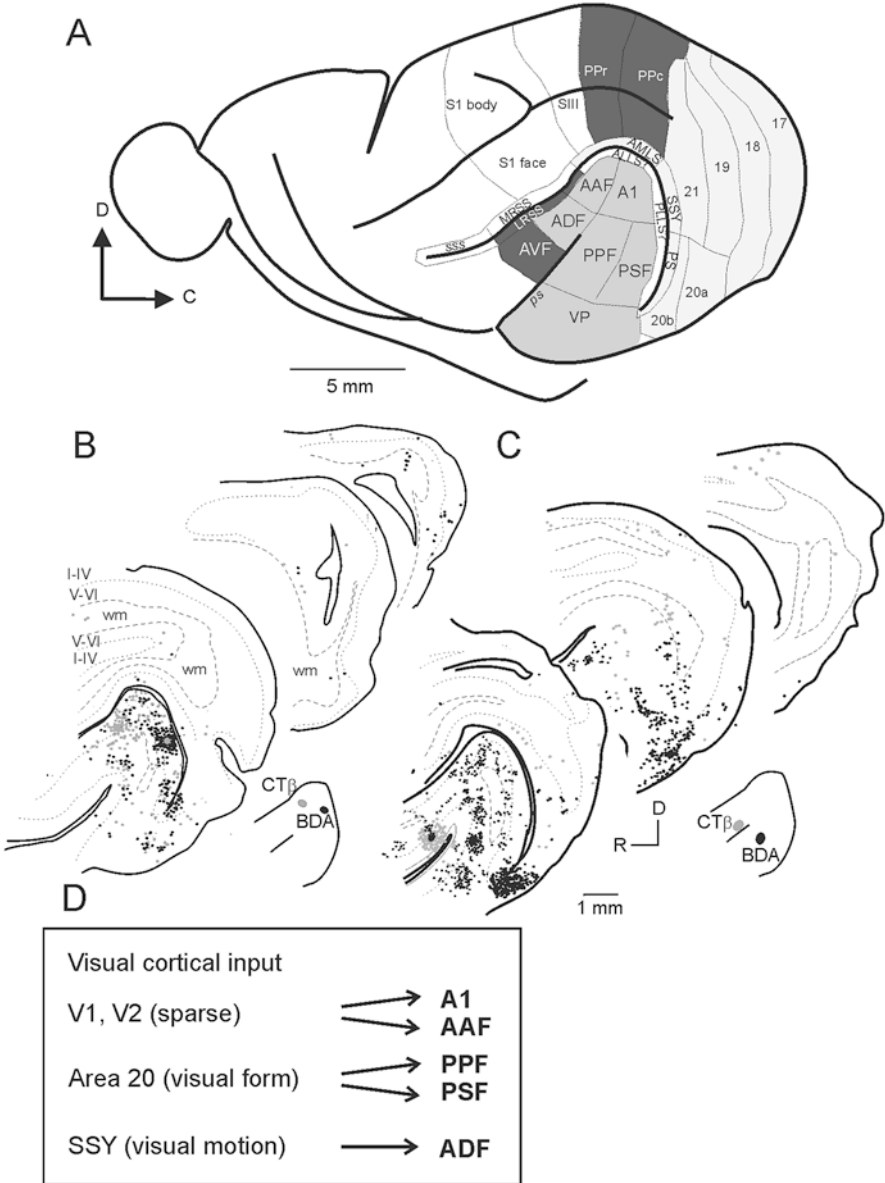
Fig. 6.6 Visual inputs to ferret auditory cortex. (**A**) Visual (areas 17-20, PS, SSY, AMLS), posterior parietal (PPr, PPc), somatosensory (S1, SIII, MRSS), and auditory (A1, AAF, PPF, PSF, and ADF) areas are shown. In addition, LRSS and AVF are multisensory regions, although many of the areas usually classified as modality specific also contain some multisensory neurons. (**B**) Location of neurons in the visual cortex that project to the auditory cortex. Tracer injections made into the core auditory cortex (A1: biotinylated dextran amine, *black*; AAF: cholera toxin subunit β, *gray*) result in retrograde labeling in the early visual areas. *Dotted lines*, limit between cortical layers IV and V; *dashed lines*, delimit the white matter. (**C**) Tracer injections made into belt the auditory cortex.

have measured neuronal activity while the animals carry out multisensory tasks (e.g., Brosch et al. 2005; Chandrasekaran et al. 2013). Consequently, the behavioral relevance of the cross-modal effects observed under anesthesia or in awake, nonbehaving animals remains speculative. Moreover, where auditory cortical recordings have been made in behaving animals, there are indications that task engagement can be accompanied by the emergence of responses to nonauditory stimuli (Brosch et al. 2005; Lakatos et al. 2009) and that the modulatory nature of these stimuli may differ. Thus, visible mouth movements improve the ability of monkeys to detect vocalizations, with this behavioral advantage accompanied by shorter latency responses by auditory cortical neurons rather than changes in the magnitude or variability of their firing rates (Fig. 6.9; Chandrasekaran et al. 2013). This again stresses the importance of considering both rate and temporal codes when investigating the impact of multisensory integration at the level of the auditory cortex.

Measuring cortical activity during behavior has provided other insights into the neural basis for cross-modal influences on auditory perception. Because visual information is normally more accurate and reliable in the spatial domain, it can provide a reference for calibrating the perception of auditory space. This is particularly the case during development when vision plays a key role in aligning the neural maps of space in the SC, as revealed by the changes produced in the auditory spatial receptive fields when the visual inputs are altered (reviewed in King 2009). This cross-modal plasticity compensates for growth-related changes and individual differences in the relative geometry of sense organs. But as illustrated by the ventriloquism illusion and related phenomena (Zwiers et al. 2003), vision can also be used in adulthood to alter the perceived location of sound sources to resolve short-term spatial conflicts between these modalities. The neural basis for the ventriloquism illusion is poorly understood, but event-related potential and fMRI measurements have revealed changes in the activity levels in the auditory cortex on trials in which participants experienced a shift in perceived sound location in the direction of a misaligned visual stimulus (Fig. 6.10; Bonath et al. 2007, 2014).

Fig. 6.6 (continued) *Gray*, retrograde labeling after an injection of CTβ into the anterior fields (on the borders of ADF and AVF); *black*, retrograde labeling resulting from a BDA injection into the posterior fields PPF and PSF. Note the difference in the extent and distribution of labeling after injections into the core and belt areas of auditory cortex. (**D**) Summary of sources of visual cortical input to different regions of auditory cortex, with their likely functions indicated. *ADF*, anterior dorsal field; *ALLS*, anterolateral lateral suprasylvian visual area; *AMLS*, anteromedial lateral suprasylvian visual area; *AVF*, anterior ventral field; *BDA*, biotinylated dextran amine; *C*, caudal; *CTβ*, cholera toxin subunit β; *D*, dorsal; *I-VI*, cortical layers; *LRSS*, lateral bank of the rostral suprasylvian sulcus; *MRSS*, medial bank of the rostral suprasylvian sulcus; *PLLS*, posterolateral lateral suprasylvian area; *PPF*, posterior pseudosylvian field; *PSF*, posterior suprasylvian field; *pss*, pseudosylvian sulcus; *PPc*, caudal posterior parietal cortex; *PPr*, rostral posterior parietal cortex; *PS*, posterior suprasylvian area; *R*, rostral; *S1*, primary somatosensory cortex; *SIII*, tertiary somatosensory cortex; *SSY*, suprasylvian cortex; *VP*, ventroposterior area; *wm*, white matter. Adapted from Bizley et al. (2007), with permission
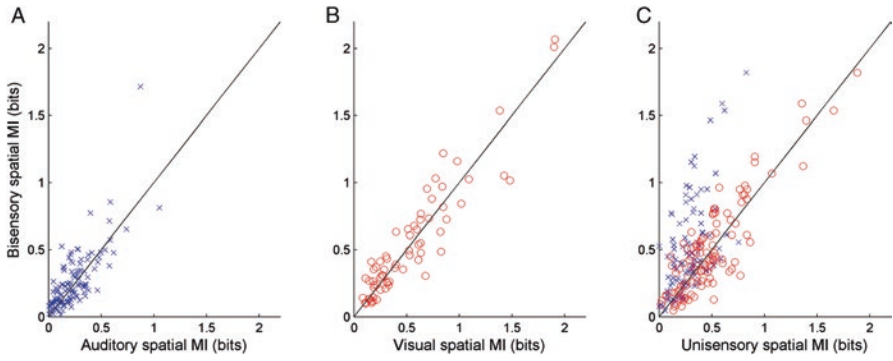
Fig. 6.7 Pairing auditory and visual stimulation produces an overall increase in the spatial information conveyed by ferret auditory cortex neurons that were driven by auditory (**A**), visual (**B**), or both auditory and visual (**C**) stimuli. Each symbol (*blue crosses*, auditory; *red circles*, visual) shows the estimated mutual information (MI) between the stimulus location and the spike trains evoked by unisensory stimulation (*x*-axis) and by combined visual-auditory stimulation (*y*-axis) for a different neuron. Higher values indicate that the responses conveyed more information about the location of the stimuli so points above the line mean that more information was transmitted in response to combined visual-auditory stimulation than in the unisensory condition. Reproduced from Bizley and King (2008), with permission

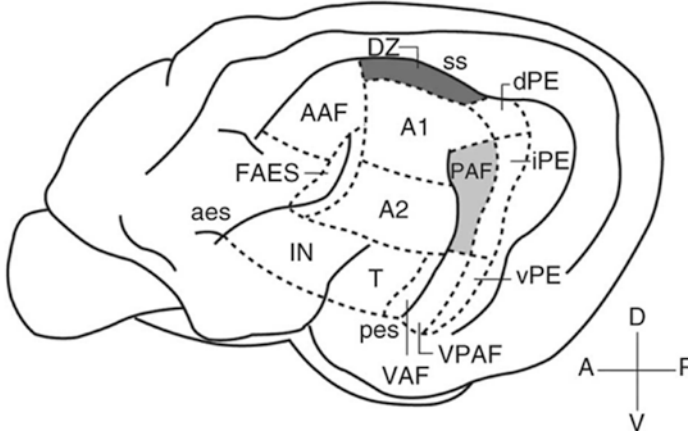| Task	PAF deactivation	DZ deactivation
Visual localization in the peripheral field	Deficit	No deficit
Movement detection	No deficit	Deficit

Fig. 6.8 *Top*, double dissociation of visual functions in the auditory cortex of the congenitally deaf cat. Bilateral deactivation of the PAF, but not the DZ, resulted in the loss of enhanced visual localization in the far periphery. On the other hand, bilateral deactivation of the DZ, but not the PAF, resulted in higher movement detection thresholds. *Bottom*, lateral view of the cat cerebrum highlighting the locations of the PAF and DZ. *A*, anterior; *aes*, anterior ectosylvian sulcus; *dPE*, dorsal posterior ectosylvian area; *DZ*, dorsal zone of auditory cortex; *IN*, insular region; *iPE*, intermediate posterior ectosylvian area; *P*, posterior; *PAF*, posterior auditory field; *pes*, posterior ectosylvian sulcus; *ss*, suprasylvian sulcus; *T*, temporal region; *V*, ventral; *VAF*, ventral auditory field; *VPAF*, ventral posterior auditory field; *vPE*, ventral posterior ectosylvian area. Reproduced from Lomber et al. (2010), with permission
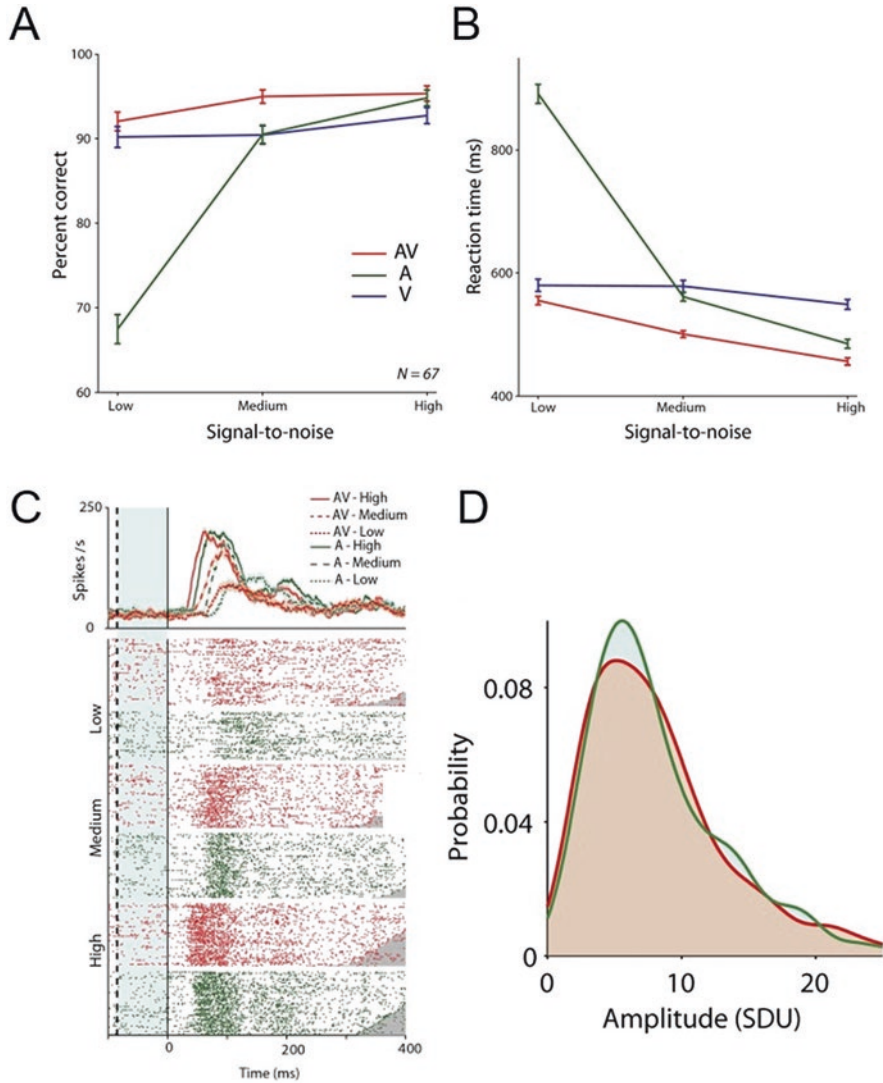
Fig. 6.9 Auditory cortical correlate of the ability of monkeys to detect auditory-visual vocalizations. Accuracy (**A**) and reaction time (**B**) for three different signal-to-noise levels for monkeys trained to detect auditory-visual vocalizations and their component auditory or visual stimulus are shown. Note the superior performance when both modality cues are available. Values are means ± SE. (**C**) Peristimulus time histogram (*top*) and rasters (*bottom*) showing the responses to auditory (A), visual (V), and auditory-visual stimulation (AV) at the three signal-to-noise levels. *Solid line,* onset of the auditory stimulus; *dashed line*, onset of the visual stimulus; *blue shading*, time period when only visual input was present. The auditory cortex responds faster with the addition of mouth motion. (**D**) Probability density of peak magnitudes for the spiking responses in the AV (*red*) and A (*green*) conditions. The *x*-axis depicts the change in normalized response magnitude in standard deviation units (SDU); the *y*-axis depicts the probability of observing that response magnitude. No systematic changes in the magnitude or variability of the firing rate were observed with the addition of mouth motion. Adapted from Chandrasekaran et al. (2013), with permission
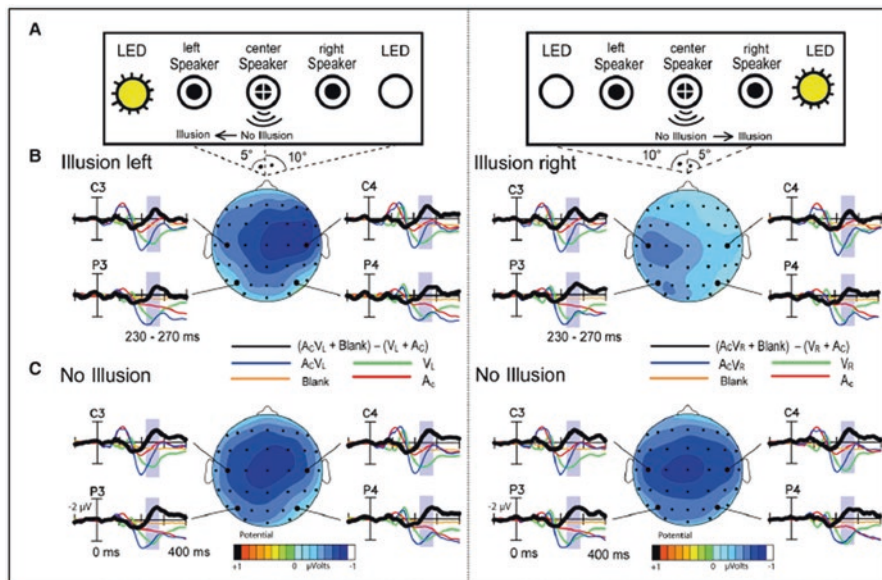
Fig. 6.10 Auditory cortical correlates of the ventriloquism illusion. (**A**) Tones were presented from left, center, or right loudspeakers, either alone or in combination with flashes from a LED on the right or left side. *Left*, stimulus combination of central tone (A_C) plus left flash (V_L); *right*, A_C plus right flash (V_R) combination. (**B**) Grand averaged event-related potential (ERP) waveforms to auditory (*red*), visual (*green*), blank (*orange*), and auditory-visual (*blue*) stimuli together with the multisensory difference waves ([AV + blank] − [A + V]; *thick black*) recorded from central (C3, C4) and parietal (P3, P4) electrodes on trials where the ventriloquist illusion was present (i.e., subjects perceived the sound as coming from the speaker on the same side as the flash). Topographical voltage maps are of the N260 component measured as mean amplitude over 230–270 ms (*shaded areas*) in the multisensory difference waves. Note larger amplitude contralateral to the side of the flash and perceived sound. (**C**) Grand average ERPs and topographical voltage distributions of the N260 component on trials where the ventriloquist illusion was absent (i.e., subjects correctly reported the sound location to be at the center). Note bilaterally symmetrical voltage distributions of N260. Reproduced from Bonath et al. (2007), with permission

The growing evidence that the auditory cortex may provide a substrate for visual influences on spatial hearing raises an important question. In the SC, each of the sensory representations is topographically organized and together they form overlapping maps of space. A shift in the visual map is therefore readily translated into a corresponding adjustment in the representation of auditory space by systematically retuning the neurons to a new set of spatial cue values, as illustrated by recordings from the optic tectum, the homologous structure to the SC, in barn owls (Knudsen 2002). In the mammalian cortex, however, there is no map of auditory space, and it is currently thought that the sound source azimuth is likely to be encoded by a comparison of activity between neurons with heterogeneous spatial sensitivity within each hemisphere (Stecker et al. 2005; Keating et al. 2015). Although it remains unclear how visual inputs, whether they originate subcortically

or from other parts of the cortex, might "map" onto this arrangement, the finding that the ventriloquism illusion is associated with a change in the balance of activity between the left and right auditory cortical areas (Bonath et al. 2014) raises testable hypotheses.

Maintaining concordant multisensory representations of space in the brain requires continuous recalibration because the spatial information provided by each modality is, at least initially, encoded using different reference frames (see Willett, Groh, and Maddox, Chap. 5). Thus, visual signals are encoded using eye-centered retinal coordinates, whereas auditory signals are head centered because the location of a sound source is derived from interaural time and level differences in conjunction with the spectral localization cues generated by the head and each external ear. An important strategy used by the brain to cope with this is to incorporate information about current gaze direction into the brain's representation of auditory space. As stated in Sect. 6.3, this process begins in the midbrain and is widespread in the monkey auditory cortex (Werner-Reiss et al. 2003), with at least some of the effects of eye position likely to arise from feedback from the parietal or frontal cortex (Fu et al. 2004). The importance of oculomotor information has also been demonstrated behaviorally by the surprising finding that looking toward a sound while keeping the head still significantly enhances the discrimination of both interaural level and time differences (Maddox et al. 2014).

6.6   Concluding Remarks

It is increasingly clear that focusing exclusively on the responses of neurons to the acoustic properties of sound is insufficient to understand how activity in the central auditory pathway, and the cortex in particular, underpins perception and behavior. Because increasingly naturalistic conditions are being used to study auditory processing, more attention is being paid to the interplay between the senses. It is now known that multisensory interactions are a property of many neurons in the auditory pathway, just as they are for other sensory systems. These interactions most commonly take the form of a modulation of auditory activity, with other sensory inputs providing contextual cues that signal the presence of an upcoming sound, thereby making it easier to hear. Additionally, they may convey more specific information about the location or identity of multisensory objects and events and enhance or recalibrate the tuning properties of the auditory neurons without changing their primary role in hearing.

Although the application of more sophisticated analytical approaches has provided valuable insights into how multisensory signals are encoded by individual auditory neurons, there is currently little understanding of the way in which populations of neurons interact to represent those signals. Moreover, given that multisensory interactions are so widespread in the brain, it remains a daunting task to decipher the specific circuits that underlie a particular behavior. Indeed, it is becoming increasingly clear that multiple circuits exist for mediating the influence of one modality on

another, as shown by recent experiments in mice illustrating the different routes by which activity in the auditory cortex can suppress that in the visual cortex (Iurilli et al. 2012; Song et al. 2017). Identification of behaviorally relevant circuits is a necessary step toward an improved understanding of the cellular and synaptic mechanisms underlying multisensory interactions.

The effects of multisensory processing on perception are well documented in humans but understandably less so in other species. As more is learned about the brain regions and cell types that mediate multisensory interactions, it will be necessary to develop new behavioral paradigms to probe their role in merging different sensory stimuli and resolving conflicts between them. This will enable further assessment of the role of attention and task engagement in multisensory processing, which has so far been largely restricted to studies in primates, as well as investigation into the role of sensory experience in shaping the connections and response properties of neurons in the auditory cortex and elsewhere in the brain so that they integrate other sensory inputs that are commonly associated with sounds.

Compliance with Ethics Requirements Andrew J. King declares that he has no conflict of interest.

Amy Hammond-Kenny declares that she has no conflict of interest.

Fernando R. Nodal declares that he has no conflict of interest.

References

Aitkin, L. M., Kenyon, C. E., & Philpott, P. (1981). The representation of the auditory and somatosensory systems in the external nucleus of the cat inferior colliculus. *Journal of Comparative Neurology, 196*(1), 25–40.

Allen, A. E., Procyk, C. A., Brown, T. M., & Lucas, R. J. (2017). Convergence of visual and whisker responses in the primary somatosensory thalamus (ventral posterior medial region) of the mouse. *The Journal of Physiology, 595*(3), 865–881.

Bell, A. H., Meredith, M. A., Van Opstal, A. J., & Munoz, D. P. (2005). Crossmodal integration in the primate superior colliculus underlying the preparation and initiation of saccadic eye movements. *Journal of Neurophysiology, 93*(6), 3659–3673.

Bizley, J. K., & King, A. J. (2008). Visual-auditory spatial processing in auditory cortical neurons. *Brain Research, 1242*, 24–36.

Bizley, J. K., Nodal, F. R., Bajo, V. M., Nelken, I., & King, A. J. (2007). Physiological and anatomical evidence for multisensory interactions in auditory cortex. *Cerebral Cortex, 17*(9), 2172–2189.

Bizley, J. K., Walker, K. M., Silverman, B. W., King, A. J., & Schnupp, J. W. (2009). Interdependent encoding of pitch, timbre, and spatial location in auditory cortex. *The Journal of Neuroscience, 29*(7), 2064–2075.

Bonath, B., Noesselt, T., Martinez, A., Mishra, J., Schwiecker, K., Heinze, H. J., & Hillyard, S. A. (2007). Neural basis of the ventriloquist illusion. *Current Biology, 17*(19), 1697–1703.

Bonath, B., Noesselt, T., Krauel, K., Tyll, S., Tempelmann, C., & Hillyard, S. A. (2014). Audio-visual synchrony modulates the ventriloquist illusion and its neural/spatial representation in the auditory cortex. *NeuroImage, 98*, 425–434.

Brosch, M., Selezneva, E., & Scheich, H. (2005). Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. *The Journal of Neuroscience, 25*(29), 6797–6806.

Budinger, E., Heil, P., Hess, A., & Scheich, H. (2006). Multisensory processing via early cortical stages: Connections of the primary auditory cortical field with other sensory systems. *Neuroscience, 143*(4), 1065–1083.

Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C., McGuire, P. K., Woodruff, P. W., Iversen, S. D., & David, A. S. (1997). Activation of auditory cortex during silent lipreading. *Science, 276*(5312), 593–596.

Cappe, C., & Barone, P. (2005). Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. *European Journal of Neuroscience, 22*(11), 2886–2902.

Cappe, C., Rouiller, E. M., & Barone, P. (2012). Cortical and thalamic pathways for multisensory and sensorimotor interplay. In M. M. Murray & M. T. Wallace (Eds.), *The neural bases of multisensory processes* (pp. 15–30). Boca Raton, FL: CRC Press.

Champoux, F., Tremblay, C., Mercier, C., Lassonde, M., Lepore, F., Gagné, J. P., & Théoret, H. (2006). A role for the inferior colliculus in multisensory speech integration. *NeuroReport, 17*(15), 1607–1610.

Chandrasekaran, C., Trubanova, A., Stillittano, S., Caplier, A., & Ghazanfar, A. A. (2009). The natural statistics of audiovisual speech. *PLoS Computational Biology, 5*(7), e1000436.

Chandrasekaran, C., Lemus, L., & Ghazanfar, A. A. (2013). Dynamic faces speed up the onset of auditory cortical spiking responses during vocal detection. *Proceedings of the National Academy of Sciences of the United States of America, 110*(48), E4668–E4677.

Clemo, H. R., Keniston, L. P., & Meredith, M. A. (2012). Structural basis of multisensory processing convergence. In M. M. Murray & M. T. Wallace (Eds.), *The neural bases of multisensory processes* (pp. 3–14). Boca Raton, FL: CRC Press.

Cooper, M. H., & Young, P. A. (1976). Cortical projections to the inferior colliculus of the cat. *Experimental Neurology, 51*(2), 488–502.

de la Mothe, L. A., Blumell, S., Kajikawa, Y., & Hackett, T. A. (2006a). Thalamic connections of the auditory cortex in marmoset monkeys: Core and medial belt regions. *Journal of Comparative Neurology, 496*(1), 72–96.

de la Mothe, L. A., Blumell, S., Kajikawa, Y., & Hackett, T. A. (2006b). Cortical connections of the auditory cortex in marmoset monkeys: Core and medial belt regions. *Journal of Comparative Neurology, 496*(1), 27–71.

Doubell, T. P., Baron, J., Skaliora, I., & King, A. J. (2000). Topographical projection from the superior colliculus to the nucleus of the brachium of the inferior colliculus in the ferret: Convergence of visual and auditory information. *European Journal of Neuroscience, 12*(12), 4290–4308.

Einevoll, G. T., Kayser, C., Logothetis, N. K., & Panzeri, S. (2013). Modelling and analysis of local field potentials for studying the function of cortical circuits. *Nature Reviews Neuroscience, 14*(11), 770–785.

Falchier, A., Schroeder, C. E., Hackett, T. A., Lakatos, P., Nascimento-Silva, S., Ulbert, I., Karmos, G., & Smiley, J. F. (2010). Projection from visual areas V2 and prostriata to caudal auditory cortex in the monkey. *Cerebral Cortex, 20*(7), 1529–1538.

Fu, K.-M. G., Johnston, T. A., Shah, A. S., Arnold, L., Smiley, J., Hackett, T. A., Garraghty, P. E., & Schroeder, C. E. (2003). Auditory cortical neurons respond to somatosensory stimulation. *The Journal of Neuroscience, 23*(20), 7510–7515.

Fu, K.-M. G., Shah, A. S., O'Connell, M. N., McGinnis, T., Eckholdt, H., Lakatos, P., Smiley, J., & Schroeder, C. E. (2004). Timing and laminar profile of eye-position effects on auditory responses in primate auditory cortex. *Journal of Neurophysiology, 92*(6), 3522–3531.

Furukawa, S., & Middlebrooks, J. C. (2002). Cortical representation of auditory space: Information-bearing features of spike patterns. *Journal of Neurophysiology, 87*(4), 1749–1762.

Gao, P. P., Zhang, J. W., Fan, S. J., Sanes, D. H., & Wu, E. X. (2015). Auditory midbrain processing is differentially modulated by auditory and visual cortices: An auditory fMRI study. *NeuroImage, 123*, 22–32.

Ghazanfar, A. A. (2009). The multisensory roles for auditory cortex in primate vocal communication. *Hearing Research, 258*(1-2), 113–120.

Ghazanfar, A. A., Maier, J. X., Hoffman, K. L., & Logothetis, N. K. (2005). Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *The Journal of Neuroscience, 25*(20), 5004–5012.

Ghazanfar, A. A., Chandrasekaran, C., & Logothetis, N. K. (2008). Interactions between the superior temporal sulcus and auditory cortex mediate dynamic face/voice integration in rhesus monkeys. *The Journal of Neuroscience, 28*(17), 4457–4469.

Groh, J. M., Trause, A. S., Underhill, A. M., Clark, K. R., & Inati, S. (2001). Eye position influences auditory responses in primate inferior colliculus. *Neuron, 29*(2), 509–518.

Hartline, P. H., Vimal, R. L., King, A. J., Kurylo, D. D., & Northmore, D. P. (1995). Effects of eye position on auditory localization and neural representation of space in superior colliculus of cats. *Experimental Brain Research, 104*(3), 402–408.

Henschke, J. U., Noesselt, T., Scheich, H., & Budinger, E. (2015). Possible anatomical pathways for short-latency multisensory integration processes in primary sensory cortices. *Brain Structure and Function, 220*(2), 955–977.

Iurilli, G., Ghezzi, D., Olcese, U., Lassi, G., Nazzaro, C., Tonini, R., Tucci, V., Benfenati, F., & Medini, P. (2012). Sound-driven synaptic inhibition in primary visual cortex. *Neuron, 73*(4), 814–828.

Jay, M. F., & Sparks, D. L. (1984). Auditory receptive fields in primate superior colliculus shift with changes in eye position. *Nature, 309*(5966), 345–347.

Jones, E. G., & Burton, H. (1974). Cytoarchitecture and somatic sensory connectivity of thalamic nuclei other than the ventrobasal complex in the cat. *Journal of Comparative Neurology, 154*(4), 395–432.

Kaltenbach, J. A. (2007). The dorsal cochlear nucleus as a contributor to tinnitus: mechanisms underlying the induction of hyperactivity. *Progress in Brain Research, 166*, 89–106.

Kanold, P. O., & Young, E. D. (2001). Proprioceptive information from the pinna provides somatosensory input to cat dorsal cochlear nucleus. *The Journal of Neuroscience, 21*(19), 7848–7858.

Katoh, Y. Y., & Benedek, G. (1995). Organization of the colliculo-suprageniculate pathway in the cat: A wheat germ agglutinin-horseradish peroxidase study. *Journal of Comparative Neurology, 352*(3), 381–397.

Kayser, C., Petkov, C. I., Augath, M., & Logothetis, N. K. (2007). Functional imaging reveals visual modulation of specific fields in auditory cortex. *The Journal of Neuroscience, 27*(8), 1824–1835.

Kayser, C., Petkov, C. I., & Logothetis, N. K. (2008). Visual modulation of neurons in auditory cortex. *Cerebral Cortex, 18*(7), 1560–1574.

Kayser, C., Logothetis, N. K., & Panzeri, S. (2010). Visual enhancement of the information representation in auditory cortex. *Current Biology, 20*(1), 19–24.

Keating, P., Dahmen, J. C., & King, A. J. (2015). Complementary adaptive processes contribute to the developmental plasticity of spatial hearing. *Nature Neuroscience, 18*(2), 185–187.

King, A. J. (2009). Visual influences on auditory spatial learning. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 364*(1515), 331–339.

King, A. J., & Palmer, A. R. (1985). Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Experimental Brain Research, 60*(3), 492–500.

Knudsen, E. I. (2002). Instructed learning in the auditory localization pathway of the barn owl. *Nature, 417*(6886), 322–328.

Koehler, S. D., & Shore, S. E. (2013). Stimulus-timing dependent multisensory plasticity in the guinea pig dorsal cochlear nucleus. *PLoS One, 8*(3), e59828.

Lakatos, P., Chen, C. M., O'Connell, M. N., Mills, A., & Schroeder, C. E. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron, 53*(2), 279–292.

Lakatos, P., O'Connell, M. N., Barczak, A., Mills, A., Javitt, D. C., & Schroeder, C. E. (2009). The leading sense: Supramodal control of neurophysiological context by attention. *Neuron, 64*(3), 419–430.

Lesicko, A. M., Hristova, T. S., Maigler, K. C., & Llano, D. A. (2016). Connectional modularity of top-down and bottom-up multimodal inputs to the lateral cortex of the mouse inferior colliculus. *The Journal of Neuroscience, 36*(43), 11037–11050.

Lohse, M., Bajo, V. M., & King, A. J. (2017). Types and distribution of multisensory interactions in auditory thalamus. *Association for Research in Otolaryngology Abstracts*, 280.

Lomber, S. G., Meredith, M. A., & Kral, A. (2010). Cross-modal plasticity in specific auditory cortices underlies visual compensations in the deaf. *Nature Neuroscience, 13*(11), 1421–1427.

Luo, H., Liu, Z., & Poeppel, D. (2010). Auditory cortex tracks both auditory and visual stimulus dynamics using low-frequency neuronal phase modulation. *PLoS Biology, 8*(8), e1000445.

Maddox, R. K., Pospisil, D. A., Stecker, G. C., & Lee, A. K. (2014). Directing eye gaze enhances auditory spatial cue discrimination. *Current Biology, 24*(7), 748–752.

Malhotra, S., & Lomber, S. G. (2007). Sound localization during homotopic and heterotopic bilateral cooling deactivation of primary and nonprimary auditory cortical areas in the cat. *Journal of Neurophysiology, 97*(1), 26–43.

Manger, P. R., Nakamura, H., Valentiniene, S., & Innocenti, G. M. (2004). Visual areas in the lateral temporal cortex of the ferret (*Mustela putorius*). *Cerebral Cortex, 14*(6), 676–689.

May, B. J. (2000). Role of the dorsal cochlear nucleus in the sound localization behavior of cats. *Hearing Research, 148*(1-2), 74–87.

McGettigan, C., Faulkner, A., Altarelli, I., Obleser, J., Baverstock, H., & Scott, S. K. (2012). Speech comprehension aided by multiple modalities: Behavioural and neural interactions. *Neuropsychologia, 50*(5), 762–776.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*(5588), 746–748.

Meredith, M. A., & Allman, B. L. (2009). Subthreshold multisensory processing in cat auditory cortex. *NeuroReport, 20*(2), 126–131.

Meredith, M. A., & Allman, B. L. (2015). Single-unit analysis of somatosensory processing in the core auditory cortex of hearing ferrets. *European Journal of Neuroscience, 41*(5), 686–698.

Meredith, M. A., & Stein, B. E. (1986). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology, 56*(3), 640–662.

Morin, L. P., & Studholme, K. M. (2014). Retinofugal projections in the mouse. *Journal of Comparative Neurology, 522*(16), 3733–3753.

Murray, M. M., & Wallace, M. T. (Eds.). (2012). *The neural bases of multisensory processes*. Boca Raton, FL: CRC Press.

Nelken, I., Chechik, G., Mrsic-Flogel, T. D., King, A. J., & Schnupp, J. W. H. (2005). Encoding stimulus information by spike numbers and mean response time in primary auditory cortex. *Journal of Computational Neuroscience, 19*(2), 199–221.

Okada, K., Venezia, J. H., Matchin, W., Saberi, K., & Hickok, G. (2013). An fMRI study of audio-visual speech perception reveals multisensory interactions in auditory cortex. *PLoS One, 8*(6), e68959.

Olcese, U., Iurilli, G., & Medini, P. (2013). Cellular and synaptic architecture of multisensory integration in the mouse neocortex. *Neuron, 79*(3), 579–593.

Park, H., Kayser, C., Thut, G., & Gross, J. (2016). Lip movements entrain the observers' low-frequency brain oscillations to facilitate speech intelligibility. *eLife, 5*, e14521.

Pekkola, J., Ojanen, V., Autti, T., Jääskeläinen, I. P., Möttönen, R., Tarkiainen, A., & Sams, M. (2005). Primary auditory cortex activation by visual speech: an fMRI study at 3 T. *NeuroReport, 16*(2), 125–128.

Philipp, R., Distler, C., & Hoffmann, K. P. (2006). A motion-sensitive area in ferret extrastriate visual cortex: an analysis in pigmented and albino animals. *Cerebral Cortex, 16*(6), 779–790.

Pinchoff, R. J., Burkard, R. F., Salvi, R. J., Coad, M. L., & Lockwood, A. H. (1998). Modulation of tinnitus by voluntary jaw movements. *American Journal of Otolaryngology, 19*(6), 785–789.

Porter, K. K., Metzger, R. R., & Groh, J. M. (2007). Visual- and saccade-related signals in the primate inferior colliculus. *Proceedings of the National Academy of Sciences of the United States of America, 104*(45), 17855–17860.

Rauschecker, J. P. (2018). Where, when, and how: Are they all sensorimotor? Towards a unified view of the dorsal pathway in vision and audition. *Cortex, 98*, 262–268.

Rouiller, E. M., & Welker, E. (2000). A comparative analysis of the morphology of corticothalamic projections in mammals. *Brain Research Bulletin, 53*(6), 727–741.

Schnupp, J., Nelken, I., & King, A. (2011). *Auditory neuroscience: Making sense of sound*. Cambridge, MA: MIT Press.

Schroeder, C. E., Lakatos, P., Kajikawa, Y., Partan, S., & Puce, A. (2008). Neuronal oscillations and visual amplification of speech. *Trends in Cognitive Sciences, 12*(3), 106–113.

Scott, B. H., Saleem, K. S., Kikuchi, Y., Fukushima, M., Mishkin, M., & Saunders, R. C. (2017). Thalamic connections of the core auditory cortex and rostral supratemporal plane in the macaque monkey. *Journal of Comparative Neurology, 525*(16), 3488–3513.

Sekiyama, K., Kanno, I., Miura, S., & Sugita, Y. (2003). Auditory-visual speech perception examined by fMRI and PET. *Neuroscience Research, 47*(3), 277–287.

Sherman, S. M. (2016). Thalamus plays a central role in ongoing cortical functioning. *Nature Neuroscience, 16*(4), 533–541.

Shore, S. E. (2005). Multisensory integration in the dorsal cochlear nucleus: Unit responses to acoustic and trigeminal ganglion stimulation. *European Journal of Neuroscience, 21*(12), 3334–3348.

Shore, S. E., & Zhou, J. (2006). Somatosensory influence on the cochlear nucleus and beyond. *Hearing Research, 216-217*, 90–99.

Smiley, J. F., & Falchier, A. (2009). Multisensory connections of monkey auditory cerebral cortex. *Hearing Research, 258*(1-2), 37–46.

Song, Y.-H., Kim, J.-H., Jeong, H.-W., Choi, I., Jeong, D., Kim, K., & Lee, S. H. (2017). A neural circuit for auditory dominance over visual perception. *Neuron, 93*(4), 940–954.

Stecker, G. C., Harrington, I. A., & Middlebrooks, J. C. (2005). Location coding by opponent neural populations in the auditory cortex. *PLoS Biology, 3*(3), e78.

Stehberg, J., Dang, P. T., & Frostig, R. D. (2014). Unimodal primary sensory cortices are directly connected by long-range horizontal projections in the rat sensory cortex. *Frontiers in Neuroanatomy, 8*, 93.

Stein, B. E., & Wallace, M. T. (1996). Comparisons of cross-modality integration in midbrain and cortex. *Progress in Brain Research, 112*, 289–299.

Stein, B. E., Huneycutt, W. S., & Meredith, M. A. (1988). Neurons and behavior: The same rules of multisensory integration apply. *Brain Research, 448*(2), 355–358.

Stevenson, R. A., Ghose, D., Fister, J. K., Sarko, D. K., Altieri, N. A., Nidiffer, A. R., Kurela, L. R., Siemann, J. K., James, T. W., & Wallace, M. T. (2014). Identifying and quantifying multisensory integration: A tutorial review. *Brain Topography, 27*(6), 707–730.

Stitt, I., Galindo-Leon, E., Pieper, F., Hollensteiner, K. J., Engler, G., & Engel, A. K. (2015). Auditory and visual interactions between the superior and inferior colliculi in the ferret. *European Journal of Neuroscience, 41*(10), 1311–1320.

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *The Journal of Acoustical Society of America, 26*(2), 212–215.

Thorne, J. D., De Vos, M., Viola, F. C., & Debener, S. (2011). Cross-modal phase reset predicts auditory task performance in humans. *The Journal of Neuroscience, 31*(10), 3853–3861.

Van Atteveldt, N., Murray, M. M., Thut, G., & Schroeder, C. E. (2014). Multisensory integration: Flexible use of general operations. *Neuron, 81*(6), 1240–1253.

Wallace, M. T., Meredith, M. A., & Stein, B. E. (1998). Multisensory integration in the superior colliculus of the alert cat. *Journal of Neurophysiology, 80*(2), 1006–1010.

Wasserman, S. M., Aptekar, J. W., Lu, P., Nguyen, J., Wang, A. L., Keles, M. F., Grygoruk, A., Krantz, D. E., Larsen, C., & Frye, M. A. (2015). Olfactory neuromodulation of motion vision circuitry in *Drosophila*. *Current Biology, 25*(4), 467–472.

Werner-Reiss, U., Kelly, K. A., Trause, A. S., Underhill, A. M., & Groh, J. M. (2003). Eye position affects activity in primary auditory cortex of primates. *Current Biology, 13*(7), 554–562.

Wigderson, E., Nelken, I., & Yarom, Y. (2016). Early multisensory integration of self and source motion in the auditory system. *Proceedings of the National Academy of Sciences of the United States of America, 113*(29), 8308–8313.

Winer, J. A., & Lee, C. C. (2007). The distributed auditory cortex. *Hearing Research, 229*(1-2), 3–13.

Yu, J. J., & Young, E. D. (2000). Linear and nonlinear pathways of spectral information transmission in the cochlear nucleus. *Proceedings of the National Academy of Sciences of the United States of America, 97*(22), 11780–11786.

Zwiers, M. P., Van Opstal, A. J., & Paige, G. D. (2003). Plasticity in human sound localization induced by compressed spatial vision. *Nature Neuroscience, 6*(2), 175–181.

Zwiers, M. P., Versnel, H., & Van Opstal, A. J. (2004). Involvement of monkey inferior colliculus in spatial hearing. *The Journal of Neuroscience, 24*(17), 4145–4156.