Chapter 28: Auditory Processing by the Central Nervous System

Introduction

HEARING IS CRUCIAL FOR LOCALIZING and identifying sound; for humans, it is particularly important because of its role in the understanding and production of speech. The auditory system has several noteworthy features. Its subcortical pathway is longer than that of other sensory systems. Unlike the visual system, sounds can enter the auditory system from all directions, day and night, when we are asleep as well as when we are awake. The auditory system processes not only sounds emanating from outside the body (environmental sounds, sounds generated by others) but also self-generated sounds (vocalizations and chewing sounds). The location of sound stimuli in space is not conveyed by the spatial arrangement of sensory afferent neurons but is instead computed by the auditory system from representations of the physical cues.

Sounds Convey Multiple Types of Information to Hearing Animals

Hearing helps to alert animals to the presence of unseen dangers or opportunities and, in many species, also serves as a means for communication. Information about where sounds arise and what they mean must be extracted from the representations of the physical characteristics of sound at each of the ears. To understand how animals process sound, it is useful first to consider which cues are available.

Most vertebrates take advantage of having two ears for localizing sounds in the horizontal plane. Sound sources at different positions in that plane affect the two ears differentially: Sound arrives earlier and is more intense at the ear nearer the source (Figure 28–1A). Interaural time and intensity differences carry information about where sounds arise.

Figure 28–1

Cues for localizing sound sources in the horizontal plane.

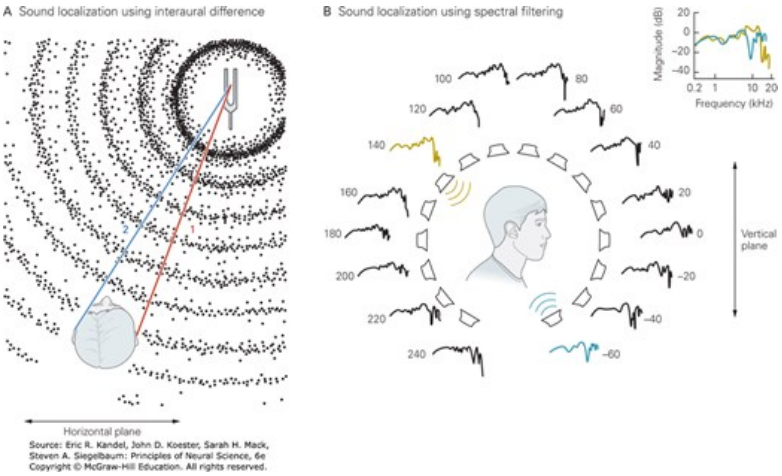
A. Interaural time and intensity differences are cues for localizing sound sources in the horizontal plane, or azimuth. A sound arising in the horizontal plane arrives differently at the two ears: Sounds arrive earlier and are louder at the ear nearer the source. A sound that arises directly in the front or back travels the same distance to the right and left ears and thus arrives at both ears simultaneously. Interaural time and intensity do not vary with the movement of sound sources in the vertical plane, so it is impossible to localize a pure sinusoidal tone in the vertical plane. In humans, the maximal interaural time difference is approximately 600 µs. High-frequency sounds, with short wavelengths, are deflected by the head, producing a sound shadow on the far side. (Adapted, with permission, from Geisler 1998.)

B. Mammals can localize broadband sounds in both the vertical and horizontal planes on the basis of spectral filtering. When a noise that has equal energy at all frequencies over the human hearing range (*white noise*) is presented through a speaker, the ear, head, and shoulders cancel energy at some frequencies and enhance others. The white noise that is emitted from the speaker has a flat power spectrum, but by the time the noise has reached the bottom of the ear canal, its spectrum is no longer flat.

In the figure, the sound energy at each frequency at the eardrum relative to that of the white noise is shown by the traces beside each speaker; these traces plot the relative sound magnitude in decibels against spectral frequency (*head-related transfer function*). The small plot in the upper right compares two head-related transfer functions: one for a noise that arises low and in front of a listener (**blue**) and one for a noise from behind the listener's head (**brown**). Head-related transfer functions have deep notches at frequencies greater than 8 kHz, whose frequencies vary depending on where the sounds arose. Sounds that lack energy at high frequencies and narrowband sounds are difficult to localize in the vertical plane. Since spectral filtering also varies in the horizontal plane, it provides the only location cue to animals that have lost hearing in one ear.

You can test the salience of these spectral cues with a simple experiment. Close your eyes as a friend jingles keys directly in front of you at various elevations. Compare your ability to localize sounds under normal conditions and when you distort the shape of both ears by pushing them with your

fingers from the back. (Data from D. Kistler and F. Wightman.)



The size of the head determines how interaural time delays are related to the location of sound sources; the neuronal circuitry determines the precision with which time delays are resolved. Because air pressure waves travel at roughly 340 m/s in air, the maximal interaural delay in humans is approximately 600 µs; in small birds, the greatest delay is only 35 µs. Humans can resolve the location of a sound source directly ahead to within approximately 1 degree, corresponding to an interaural time difference of 10 µs. Interaural time differences are particularly well conveyed by neurons that encode relatively low frequencies. These neurons can fire at the same position in every cycle of the sound and in this way encode the interaural time difference as an interaural phase difference. Sounds of high frequencies produce *sound shadows* or intensity differences between the two ears. For many mammals with small heads, high-frequency sounds provide the primary cue for localizing sound in the horizontal plane.

Mammals can localize sounds in the vertical plane and with a single ear using spectral filtering. High-frequency sounds, with wavelengths that are close to or smaller than the dimensions of the head, shoulders, and external ears, interact with those parts of the body to produce constructive and destructive interference, introducing broad spectral peaks and deep, narrow spectral notches whose frequency changes with the location of the sound (Figure 28–1B). High-frequency sounds from different origins are filtered differently because in mammals the shape of the external ear differs back-to-front as well as top-to-bottom. Animals learn to use these spectral cues to locate sound sources. If the shape of the ear is experimentally altered, even adult humans can learn to make use of a new pattern of spectral cues. If animals lose hearing in one ear, they lose interaural timing and intensity cues and must depend completely on spectral cues for localizing sounds.

How do we make sense of the complex and changing sounds that we hear? Most natural sounds contain energy over a wide range of frequencies and change rapidly with time. The information used to recognize sounds varies among animal species, and depends on listening conditions and experience. Human speech, for example, can be understood in the midst of noise, over electronic devices that distort sounds, and even through cochlear implants. One reason for its robustness is that speech contains redundant cues: The vocal apparatus produces sounds in which multiple parameters covary. At the same time, this makes the task of understanding how animals recognize patterns a complicated one. It is not clear which cues are used by animals under various conditions.

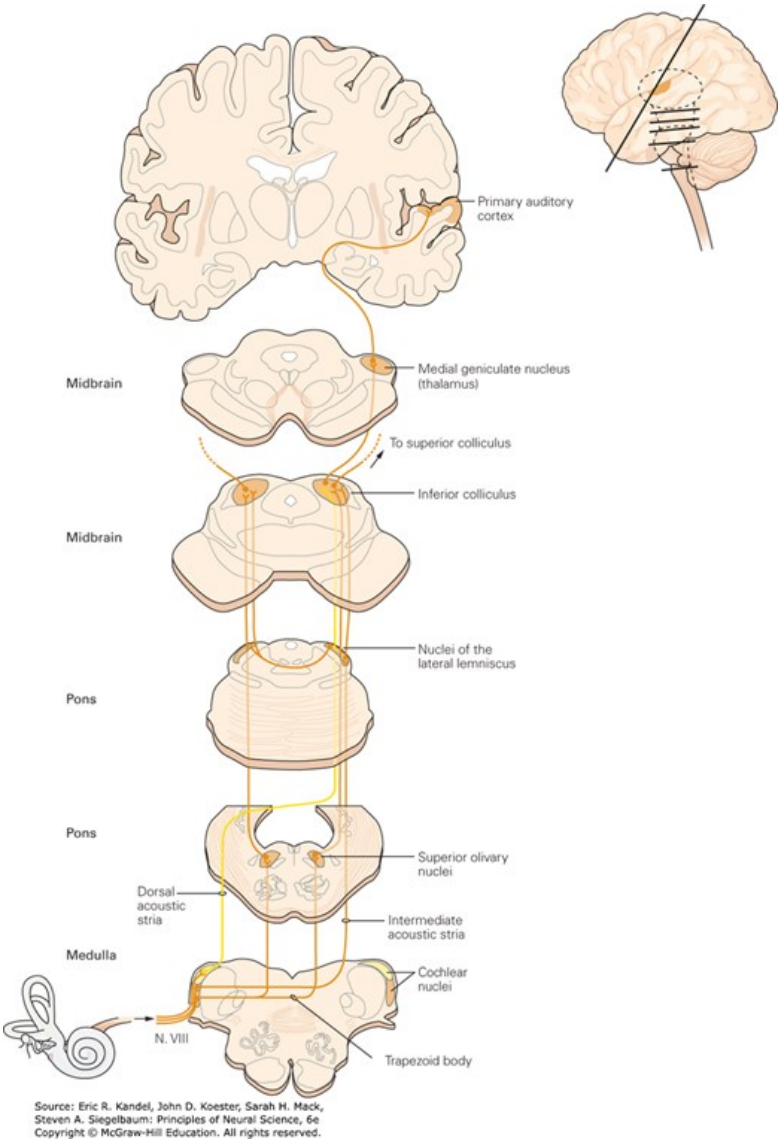
Music is a source of pleasure to human beings. Musical instruments and human voices produce sounds that have energy at the fundamental frequency that corresponds to its perceived pitch, as well as at multiples of that frequency, giving sounds a quality that allows us, for example, to distinguish a flute from a violin when their pitch is the same. Musical pitches are largely in the low-frequency range in which auditory nerve fibers fire in phase with sounds. In music, sounds are combined simultaneously to produce chords and successively to produce melodies. Euphonious, pleasant chords elicit regular, periodic firing in cochlear nerve fibers. In dissonant sounds, there is less regularity both in the sound itself and in the firing of auditory nerve fibers; the component frequencies are so close that they interfere with one another instead of periodically reinforcing one another.

The Neural Representation of Sound in Central Pathways Begins in the Cochlear Nuclei

The neural pathways that process acoustic information extend from the ear to the brain stem, through the midbrain and thalamus, to the cerebral cortex (Figure 28–2). Acoustic information is conveyed from cells in the cochlear ganglion (see Figure 26–17) to the cochlear nuclei in the brain stem. There information is received by several different types of neurons, most of which are arranged tonotopically.

Figure 28–2

The central auditory pathways extend from the brain stem through the midbrain and thalamus to the auditory cortex. The fibers in the cochlear nerve (cranial nerve VIII) terminate in the cochlear nuclei of the brain stem. The neurons of these nuclei project in several parallel pathways to the inferior colliculus. Their axons exit through the trapezoid body, intermediate acoustic stria, or dorsal acoustic stria. Some cells terminate directly in the inferior colliculus. Others contact cells in the superior olivary complex and in the nuclei of the lateral lemniscus, which in turn project to the inferior colliculus. Neurons of the inferior colliculus project to the superior colliculus and to the medial geniculate nucleus of the thalamus. Thalamic neurons project to the auditory cortex. The cochlear nuclei and the ventral nuclei of the lateral lemniscus are the only central auditory neurons that receive monaural input. (Adapted, with permission, from Brodal 1981.)



The axons of the different types of neurons take different routes to the brain stem and midbrain, where they terminate on separate targets. Some of the pathways from the cochlear nuclei to the contralateral inferior colliculus are direct; others involve one or two synaptic stages in brain stem auditory nuclei. From the bilateral inferior colliculi, acoustic information flows two ways: to the ipsilateral superior colliculus, where it participates in orienting the head and eyes in response to sounds, and to the ipsilateral thalamus, the relay to auditory areas of the cerebral cortex. The afferent auditory pathways from the periphery to higher brain regions include efferent feedback at many levels.

The Cochlear Nerve Delivers Acoustic Information in Parallel Pathways to the Tonotopically Organized Cochlear Nuclei

The afferent nerve fibers from cochlear ganglion cells are bundled in the cochlear or auditory component of the vestibulocochlear nerve (cranial nerve VIII) and terminate exclusively in the cochlear nuclei. The cochlear nerve in mammals contains two groups of fibers: a large number (95%) of myelinated fibers that receives input from inner hair cells and a small number (5%) of unmyelinated fibers that receive input from outer hair cells.

The larger, more numerous, myelinated fibers are much better understood than the unmyelinated fibers. Each type detects energy over a narrow range of frequencies; the tonotopic array of cochlear nerve fibers thus carries detailed information about how the frequency content of sounds varies from moment to moment. The unmyelinated fibers terminate both on the large neurons in the ventral cochlear nuclei and also on the small granule cells that surround the ventral cochlear nuclei. Because it is difficult to record from these tiny fibers, the information they convey to the brain is not well understood. The unmyelinated fibers integrate information from a relatively wide region of the cochlea but are not responsive to sound. It has been suggested that these fibers respond to cochlear damage and contribute to hyperacusis—pain after exposure to loud sounds that damages the cochlea.

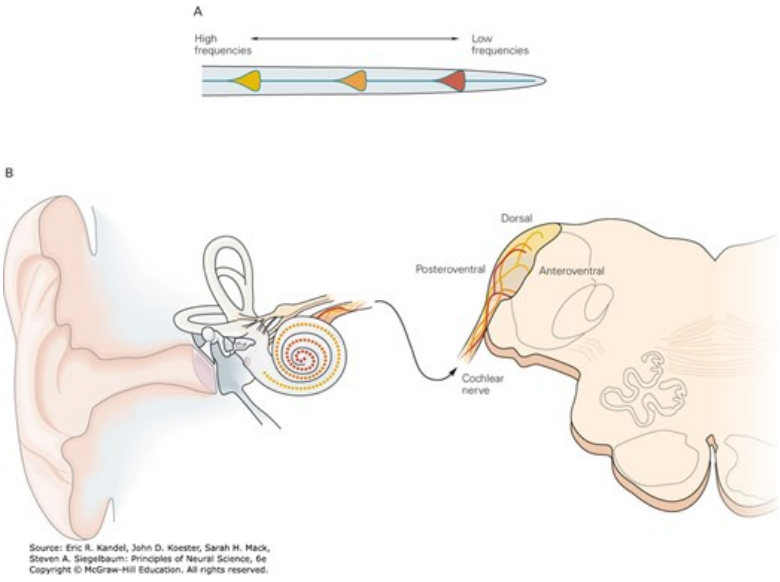
Two features of the cochlear nuclei are important. First, these nuclei are organized tonotopically. Fibers that carry information from the apical end of the cochlea, which detects low frequencies, terminate ventrally in the ventral and dorsal cochlear nuclei; those that carry information from the basal end of the cochlea, which detects high frequencies, terminate dorsally (Figure 28–3). Second, each cochlear nerve fiber innervates several different areas within the cochlear nuclei, contacting various types of neurons that have distinct projection patterns to higher auditory centers. As a result, the auditory pathway comprises at least four parallel ascending pathways that simultaneously extract different acoustic information from the signals carried by cochlear nerve fibers. Parallel circuits are a general feature of vertebrate sensory systems.

Figure 28–3

The dorsal and ventral cochlear nuclei.

A. Stimulation with three frequencies of sound vibrates the schematically uncoiled basilar membrane at three positions, exciting distinct populations of hair cells and their afferent nerve fibers.

B. Cochlear nerve fibers project in a tonotopic pattern to the cochlear nuclei. Those encoding the lowest frequencies (**red**) terminate most ventrally, whereas those encoding higher frequencies (**yellow**) terminate more dorsally. The cochlear nuclei include the ventral and dorsal nuclei. Each afferent fiber enters at the nerve root and splits into branches that run anteriorly (the ascending branch) and posteriorly (the descending branch). The ventral cochlear nucleus is thus divided functionally into anteroventral and posteroventral divisions.



The Ventral Cochlear Nucleus Extracts Temporal and Spectral Information About Sounds

The principal cells of the unlayered ventral cochlear nucleus sharpen temporal and spectral information and convey it to higher centers of the auditory pathway. Three types of neurons are intermingled and form separate pathways through the brain stem (Figure 28–4).

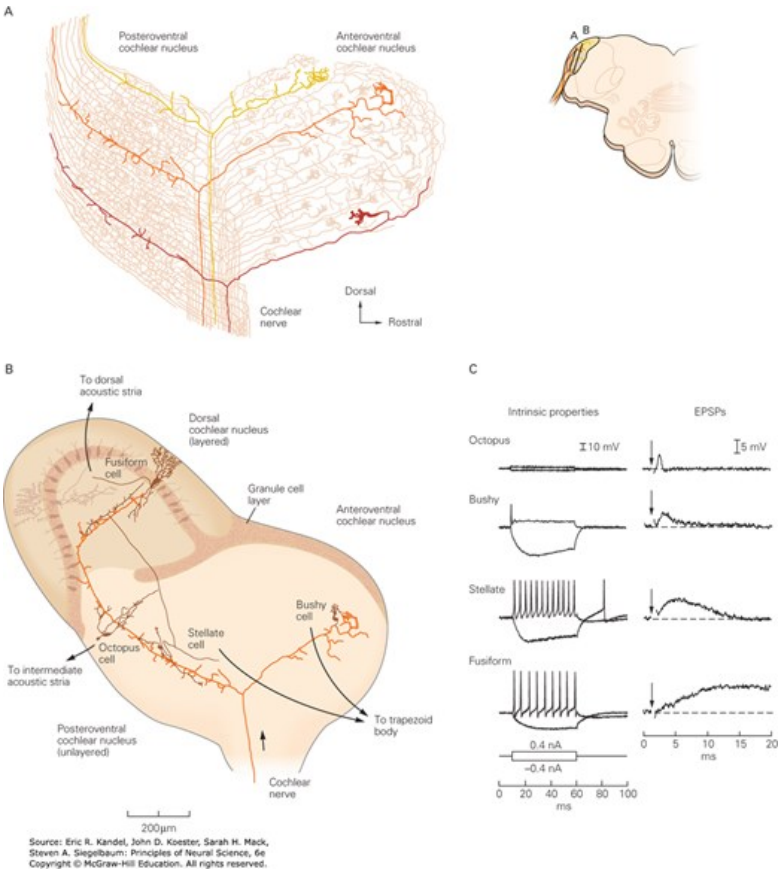
Figure 28–4

Different types of cells in the cochlear nuclei extract distinct types of acoustic information from cochlear nerve fibers.

A. The differing sizes and shapes of terminals along the length of each cochlear nerve fiber in the ventral cochlear nucleus of a newborn dog reflect differences in their postsynaptic targets. The large end bulbs form synapses on bushy cells; smaller boutons contact stellate and octopus cells. The nerve fibers shown here are color-coded as in Figure 28–3: the **yellow** fiber encodes the highest frequencies and the **red** fiber the lowest. (Adapted, with permission, from Cajal 1909.)

B. A layer of mouse granule cells (**light brown**) separates the unlayered ventral cochlear nucleus (**pink**) from the layered dorsal nucleus (**tan and light brown**). In the dorsal cochlear nucleus, the cell bodies of fusiform and granule cells are intermingled in a region between the outermost molecular layer and the deep layer. Cochlear nerve fibers, color-coded for frequency as in Figure 28–3, terminate in both nuclei but with different patterns of convergence on the principal cells. Each bushy, stellate, and fusiform cell receives input from a few auditory nerve fibers and is sharply tuned, whereas individual octopus cells are contacted by many auditory nerve fibers and are broadly tuned.

C. Differences in the intrinsic electrical properties of the principal cells of mouse cochlear nuclei are reflected in the patterns of voltage change in the cells. When steadily depolarized, stellate and fusiform cells fire repetitive action potentials, whereas repetitive firing in bushy and octopus cells is prevented by low-voltage-activated conductances. The low input resistance of bushy and octopus cells in the depolarized voltage range makes depolarizing voltage changes rapid but also small; the rise and fall of voltage changes in stellate and fusiform cells is slower. Synaptic potentials, too, are different. The brief synaptic potentials in bushy and octopus cells require larger synaptic currents but encode the timing of auditory nerve inputs more faithfully than do the longer-lasting synaptic potentials in stellate or fusiform cells. (Reproduced, with permission, from N. Golding.)



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e

Bushy cells project bilaterally to the superior olivary complex. This pathway has two parts. One courses through the medial superior olive and compares the time of arrival of sounds at the two ears; the other travels through the medial nucleus of the trapezoid body and the lateral superior olive and compares interaural intensity. Large spherical bushy cells sense low frequencies and project bilaterally to the medial superior olive, forming a circuit that detects interaural time delay and contributes to the localization of low-frequency sounds in the horizontal plane. The small spherical bushy

cells and globular bushy cells sense higher frequencies. Small spherical bushy cells excite the lateral superior olive ipsilaterally. The globular bushy cells, through calyceal endings, excite neurons in the contralateral medial nucleus of the trapezoid body that in turn inhibit principal cells of the lateral superior olive. Neurons in the lateral superior olive integrate the ipsilateral excitation and contralateral inhibition to measure interaural intensity and to localize sources of high-frequency sounds in the horizontal plane (see Figure 28–6).

Stellate cells terminate widely. They excite neurons in the ipsilateral dorsal cochlear nucleus, the medial olivocochlear efferent neurons in the ventral nucleus of the trapezoid body, the periolivary nuclei in the vicinity of the ipsilateral lateral superior olive, and the contralateral ventral nucleus of the lateral lemniscus, inferior colliculus, and thalamus. The tonotopic array of stellate cells encodes the spectrum of sounds.

Octopus cells excite targets in the contralateral paraolivary nucleus and terminate in large excitatory calyceal endings on neurons of the ventral nucleus of the lateral lemniscus, which in turn provide sharply timed glycinergic inhibition to the inferior colliculus. Octopus cells detect onsets of sounds that allow animals to detect brief gaps. They mark the spectral components that come from one source that necessarily start together.

The differences in the integrative tasks performed by these pathways through the ventral cochlear nucleus are reflected in cell morphology. The shapes of their dendrites reflect the way they collect information from cochlear nerve fibers. The dendrites of the sharply tuned bushy and stellate cells receive input from relatively few cochlear nerve fibers, whereas those of the broadly tuned octopus cells, in contrast, lie perpendicular to the path of cochlear nerve fibers, poised to receive input from many cochlear nerve fibers. Many of the inputs to bushy cells are from unusually large terminals that envelop the bushy cell bodies, meeting their need for large synaptic currents. The need for large synaptic currents in octopus cells is met by summing inputs from large numbers of small terminals.

The biophysical properties of neurons determine how synaptic currents are converted to voltage changes and over how long a time synaptic inputs are integrated. Octopus and bushy cells in the ventral cochlear nucleus are able to respond with exceptionally rapid and precisely timed synaptic potentials. These neurons have a prominent, low-voltage-activated K^+ conductance that confers a low input resistance and rapid responsiveness and prevents repetitive firing (Figure 28–4C). The large synaptic currents that are required to trigger action potentials in these leaky cells are delivered through rapidly gated, high-conductance, AMPA-type (α-amino-3-hydroxy-5-methylisoxazole-4-propionate) glutamate receptors at many synaptic release sites. In contrast, stellate cells, in which even relatively small depolarizing currents produce large protracted voltage changes, generate slower excitatory postsynaptic potentials (EPSPs) in response to synaptic currents, and *N*-methyl-D-aspartate (NMDA)-type glutamate receptors enhance those responses.

The Dorsal Cochlear Nucleus Integrates Acoustic With Somatosensory Information in Making Use of Spectral Cues for Localizing Sounds

Among vertebrates, only mammals have dorsal cochlear nuclei. The dorsal cochlear nucleus receives input from two systems of neurons that project to different layers (Figure 28–4A,B). Its principal cells, fusiform cells, integrate those two systems of inputs and convey the result directly to the contralateral inferior colliculus.

The outermost molecular layer is the terminus of a system of parallel fibers, the unmyelinated axons of granule cells that are scattered in and around the cochlear nuclei. This system transmits somatosensory, vestibular, and auditory information from widespread regions of the brain to the molecular layer.

The deep layer receives acoustic information. Not only cochlear nerve fibers but also stellate cells of the ventral cochlear nucleus terminate in the deep layer. Acoustic inputs are tonotopically organized in isofrequency laminae that run at right angles to parallel fibers.

Fusiform cells, the principal cells of the dorsal cochlear nucleus, integrate the two systems of inputs. Parallel fibers in the molecular layer excite fusiform cells through spines on apical dendrites in the molecular layer. Parallel fibers also terminate on spines of dendrites of cartwheel cells, interneurons that bear a strong resemblance to cerebellar Purkinje cells, which in turn inhibit fusiform cells. Cochlear nerve fibers and stellate cells in the ventral cochlear nucleus excite fusiform cells and inhibitory interneurons via synapses on the smooth basal dendrites in the deep layer.

Recent experiments suggest that the circuits of the dorsal cochlear nucleus distinguish between unpredictable and predictable sounds. An animal's own chewing or licking sounds, for example, are predictable and canceled through these circuits. The changes in spectral cues that arise when animals move their heads or ears or shoulders, changing the angle of incidence of sounds to the ears, are unpredictable, especially when an external sound source is moving. Somatosensory and vestibular information about the position of the head and ears, as well as descending information from higher

levels of the nervous system about the animal's own movements, pass through the molecular layer to modulate acoustic information that arrives in the deep layer.

The Superior Olivary Complex in Mammals Contains Separate Circuits for Detecting Interaural Time and Intensity Differences

In many vertebrates, including mammals and birds, neurons in the superior olivary complex compare the activity of cells in the bilateral cochlear nuclei to locate sound sources. Separate circuits detect interaural time and intensity differences and project to the inferior colliculi.

The Medial Superior Olive Generates a Map of Interaural Time Differences

Differences in arrival times at the ears are not represented at the cochlea. Instead, they are first represented in the medial superior olive where a map of interaural phase is created by a comparison of the timing of action potentials in the responses to sounds from the two ears. Sounds arrive at the near ear before they arrive at the far ear, with interaural time differences being directly related to the location of sound sources in the horizontal plane (Figure 28–5A).

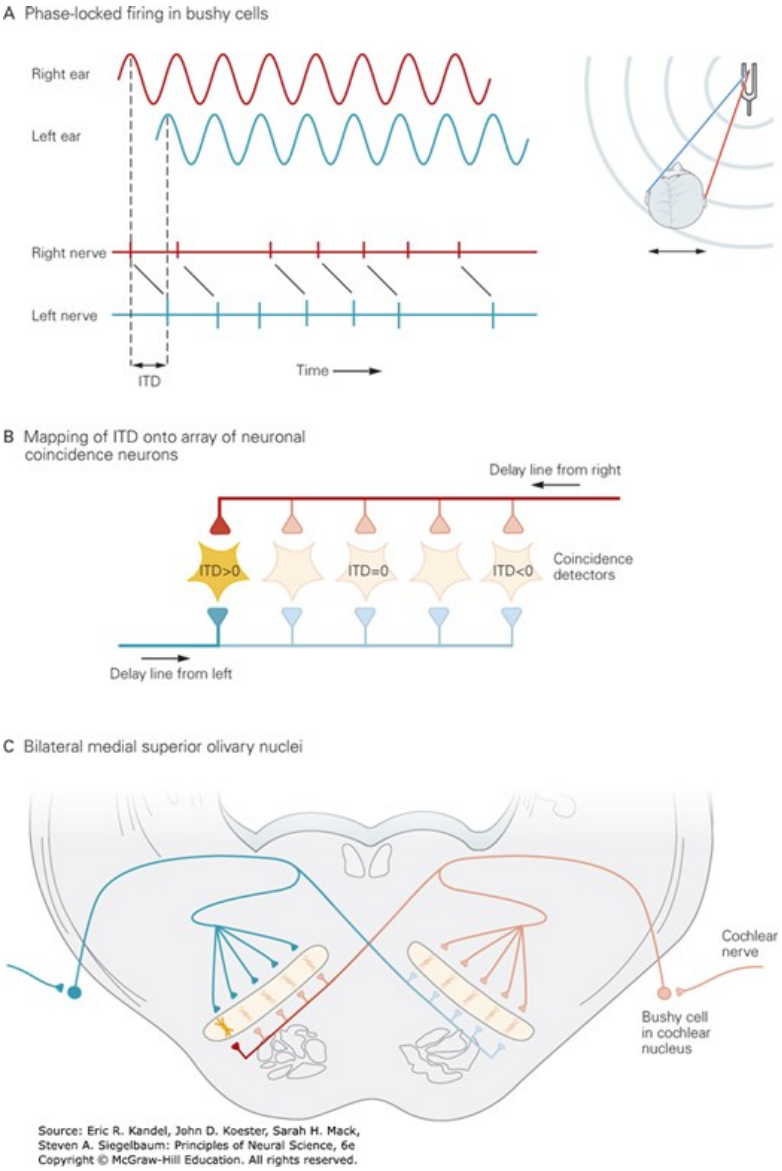
Figure 28–5

Interaural differences in the arrival of a sound help localize sound in the horizontal plane.

A. When a sound such as a pure tone arises from the right, the right ear detects the sound earlier than the left ear. The difference in the time of arrival at the two ears is the interaural time delay (**ITD**). Cochlear nerve fibers and their bushy cell targets fire in phase with pressure changes. Although individual bushy cells may fail to fire at some cycles, a set of cells will encode the timing of a low-frequency sound and its frequency with every cycle. Comparison of the onset of action potentials in the bushy cells at the two sides reveals the ITDs (**slanted black lines**).

B. Interaural time differences can be measured by an array of neurons whose inputs from the two ears are delay lines as proposed by Lloyd Jeffress (1948). Action potentials propagate to reach the nearest terminals before they reach the farthest ones; thus, in the delay line from the right, terminals will generate synaptic potentials sequentially from right to left, and in the delay line from the left, terminals will generate synaptic potentials sequentially from left to right. Suppose that such postsynaptic neurons are coincidence detectors, firing only when they receive excitatory postsynaptic potentials (EPSPs) simultaneously from the right and left. Sounds that arise at the midline reach the right and left ears simultaneously with no interaural time disparity (ITD = 0). The neuron in the middle of the array that receives input from equally long axons from the two sides will thus receive simultaneous EPSPs from the two sides. When sounds come from the right, signals from the right ear arrive at the central nervous system earlier than those from the left ear (ITD >0). Sound from the right generates synchronous EPSPs in the (**yellow**) neuron because the earlier arrival of sound from the right (**red**) is compensated by a longer conduction delay relative to that from the left (**blue**). Likewise, when sound arises from the left, the ITD <0 and conduction delays from the left (**blue**) compensate for the early arrival at the left. Such a neuronal circuit produces a map of interaural time disparities in the coincidence detectors; as sounds move from the right to left, they activate coincidence detectors sequentially from left to right. Such an arrangement of delay lines has been found in the nucleus laminaris of the barn owl, the homolog of the mammalian medial superior olivary nucleus.

C. Mammals use delay lines only in the nucleus contralateral to a sound source to form a map of interaural time differences. The bitufted neurons of the medial superior olivary nucleus form a sheet that is contacted on its lateral face by bushy cells from the ipsilateral cochlear nucleus and on the medial face by bushy cells from the contralateral cochlear nucleus. (Although it is depicted here schematically in a coronal section of the brain stem, the encoding of interaural disparities is in a sheet of neurons that also has a rostrocaudal dimension.) On the ipsilateral side, the branches of the bushy cell axon are of equal length and thus initiate synaptic currents in their targets in the medial superior olive simultaneously. On the contralateral side, the branches deliver synaptic currents sequentially first to the regions closest to the midline, and then to progressively more lateral regions. Neurons of the medial superior olive detect synchronous excitation from the two ears only when sounds arise from the contralateral half of space. When sounds arise from the right side, their early arrival at the right ear is compensated by progressively longer conduction delays to activate neurons more and more toward the lateral end of the left medial superior olive (the **yellow cell** is activated by a sound from the far right, as in part **B**). When sounds arise from the front and there is no interaural time difference, neurons in the anterior end of the medial superior olive are activated synchronously from both sides. Each medial superior olive forms a map of where sounds arise in the contralateral hemifield. (Adapted, with permission, from Yin 2002.)



Cochlear nerve fibers tuned to frequencies below 4 kHz and their bushy cell targets encode sounds by firing in phase with the pressure waves. This property is known as *phase-locking.* Although individual neurons may fail to fire at some cycles, some set of neurons fires with every cycle. In so doing, these neurons carry information about the timing of inputs with every cycle of the sound. Sounds arriving from one side evoke phase-locked firing that is consistently earlier at the near ear than at the far ear, resulting in consistent interaural phase differences (Figure 28–5A).

In 1948, Lloyd Jeffress suggested that an array of detectors of coincident inputs from the two ears, transmitted through *delay lines* comprised of axons with systematically differing lengths, could form a map of interaural time differences and thus a map of the location of sound sources (Figure 28–5B). In such a circuit, conduction delays compensate for the earlier arrival at the near ear. Interaural time delays increase systematically as sounds move from the midline to the side, resulting in coincident firing further toward the edge of the neuronal array.

Such neuronal maps have been found in the barn owl in the homolog of the medial superior olivary nucleus. Mammals and chickens use a variant of this input arrangement. The principal neurons of the medial superior olive form a sheet of one or a few cells' thickness on each side of the midline. Each neuron has two tufts of dendrites, one extending to the lateral face of the sheet, and the other projecting to the medial face of the sheet (Figure 28–5C). The dendrites at the lateral face are contacted by the axons of large spherical bushy cells from the ipsilateral cochlear nucleus, whereas the dendrites at the medial face are contacted by large spherical bushy cells of matching best frequency from the contralateral cochlear nucleus. The axons of bushy cells terminate in the contralateral medial superior olive with delay lines just as Jeffress had suggested, but the branches that terminate in the ipsilateral medial superior olive are of equal length (see Figure 28–5C).

The conduction delays are such that each medial superior olive receives coincident excitatory inputs from the two ears only when sounds come from the contralateral half of space. As sound sources move from the midline to the most lateral point on the contralateral side of the head, the earlier arrival of sounds at the contralateral ear needs to be compensated by successively longer delay lines. This results in inputs from the two ears coinciding at successively more posterior and lateral regions of the medial superior olive. Inhibition superimposed on these excitatory inputs plays a significant role in sharpening the map of interaural phase.

In encoding interaural phase, individual neurons in the medial superior olive provide ambiguous information about interaural time differences. Phase ambiguities are resolved when sounds have energy at multiple frequencies, as natural sounds almost always do. The sheet of neurons of the medial superior olive forms a representation of interaural phase along the rostrocaudal and lateromedial dimensions. The array of bushy cell inputs also imposes a tonotopic organization in the dorsoventral dimension. Sounds that contain energy at multiple frequencies evoke maximal coincident firing in a single dorsoventral column of neurons that localizes sound sources unambiguously. The beauty of using interaural phase to encode interaural time disparities is that the brain receives information about interaural time differences not just at the beginning and end of the sound but with every cycle of an ongoing sound.

Principal cells of the medial superior olive also receive sharply timed inhibition driven by sounds from both the ipsilateral and contralateral sides through the lateral and medial nuclei of the trapezoid body, respectively. Remarkably, the inhibition through pathways from both sides precedes the arrival of excitation and sharpens the summation of excitation even though inhibition is mediated through a pathway that has an additional synapse. The great conduction speed through the disynaptic pathway through the medial nucleus of the trapezoid body is made possible by the large axons of globular bushy cells and the large calyceal terminals of Held that activate neurons in the medial nucleus of the trapezoid body with short and consistently timed delays. The pathway that brings ipsilateral inhibition through the lateral nucleus of the trapezoid body is less well understood.

Each medial superior olive thus forms a map of the location of sound sources in the contralateral hemifield. The striking difference between this spatial representation of stimuli and those in other sensory systems is that it is not the result of the spatial arrangement of inputs, like retinotopic or somatosensory maps, but is inferred by the brain from computations made in the afferent pathways.

The Lateral Superior Olive Detects Interaural Intensity Differences

Sounds with wavelengths that are similar to or smaller than the head are deflected by the head, causing the intensity at the near ear to be greater than that at the far ear. In humans, interaural intensities can differ in sounds that have frequencies greater than about 2 kHz. Interaural intensity differences produced by such *head shadowing* are detected by a neuronal circuit that includes the medial nucleus of the trapezoid body and the lateral superior olive.

Although the lateral superior olive does not form a map of the location of sounds in the horizontal plane, it performs the first of several integrative steps that use interaural intensity differences to localize sounds. Neurons in this nucleus balance ipsilateral excitation with contralateral inhibition. Excitation comes from small spherical bushy cells and stellate cells in the ipsilateral ventral cochlear nucleus. Inhibition comes from a disynaptic pathway that includes globular bushy cells in the contralateral ventral cochlear nucleus and principal neurons of the ipsilateral medial nucleus of the trapezoid body (Figure 28–6A). Sounds that arise ipsilaterally generate relatively strong excitation and relatively weak inhibition, whereas those that arise contralaterally generate stronger inhibition than excitation. Neurons in the lateral superior olive are activated more strongly by sounds from the ipsilateral than from the contralateral hemifield. The firing of lateral superior olivary neurons is a function of the location of the sound source and thus carries information about where sounds arise in the horizontal plane (Figure 28–6B).

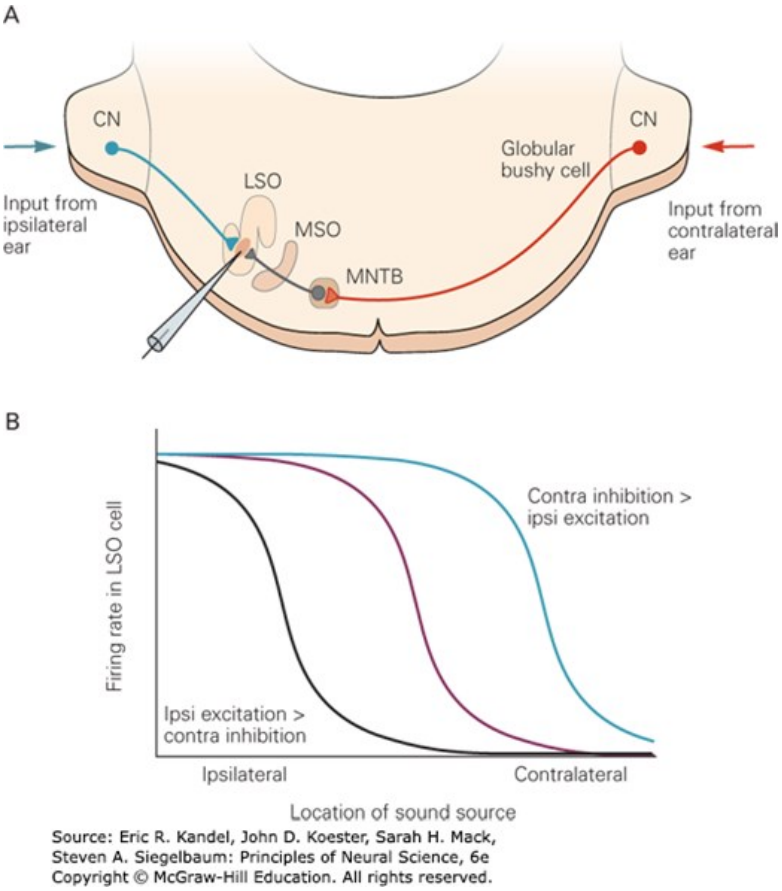
Figure 28–6

Interaural differences in the intensity of a sound also help localize sound in the horizontal plane.

A. Principal cells of the lateral superior olivary nucleus (**LSO**) receive excitatory input from the ipsilateral cochlear nucleus (**CN**) and inhibitory input from the contralateral cochlear nucleus. A coronal section through the brain stem of a cat illustrates the anatomical connections. Small spherical bushy cells and stellate cells in the ipsilateral ventral cochlear nucleus provide direct excitation. Globular bushy cells in the contralateral ventral cochlear nucleus project across the midline and excite neurons in the medial nucleus of the trapezoid body (**MNTB**) via large terminals, the calyces of Held. Cells of the medial nucleus of the trapezoid body inhibit neurons in the lateral superior olive as well as in the medial superior olive (**MSO**). For neurons of the lateral superior olive to compare intensities of the same sound, the timing of the ipsilateral excitatory input must be matched with the

timing of the contralateral inhibitory input. To this end, globular bushy cells have particularly large axons that terminate in a calyx of Held in the medial nucleus of the trapezoid body where synaptic transmission is strong and thus the synaptic delay is short and invariant in its timing.

B. The firing of neurons in the lateral superior olive reflects a balance of ipsilateral excitation and contralateral inhibition. When sounds arise from the ipsilateral side, excitation is relatively stronger and inhibition is relatively weaker than when sounds arise from the contralateral side. The transition between the dominance of excitation and inhibition reflects the location of the sound source.



In order to balance excitation and inhibition stimulated by one sound, the ipsilateral excitation and contralateral inhibition must arrive at neurons in the lateral superior olive at the same time. Thus, excitation that arises monosynaptically from the ipsilateral ventral cochlear nucleus must arrive at the same time as inhibition that arises disynaptically from the contralateral ventral cochlear nucleus. The inhibition comes from the medial nucleus of the trapezoid body whose inputs through large axons of globular bushy cells and large calyces of Held produce synaptic responses with short and consistently timed delays. The axons of small spherical bushy cells and stellate cells that carry ipsilateral excitation conduct more slowly than those of globular bushy cells.

The terminals of the globular bushy cells, the calyces of Held, engulf the cell bodies of trapezoid-body neurons so dramatically that they caught the attention of early anatomists and modern biophysicists. A single somatic terminal releases neurotransmitter at numerous release sites and generates large synaptic currents. The reliability of pre- and postsynaptic recordings at this synapse makes the site ideal for detailed studies of the mechanisms of synaptic transmission (Chapter 15).

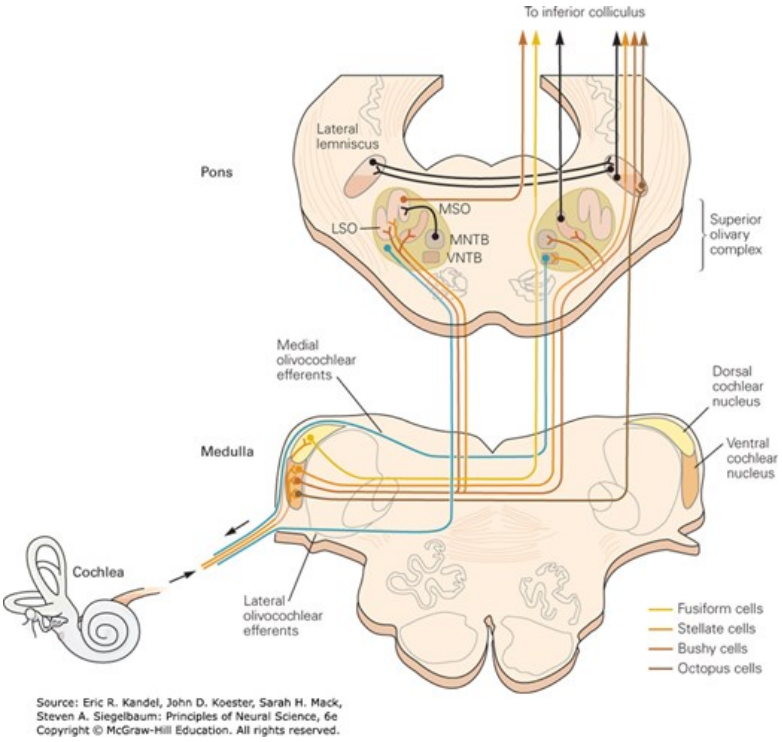
The Superior Olivary Complex Provides Feedback to the Cochlea

Although sensory systems are largely afferent, bringing sensory information to the brain, recent studies have led to an appreciation of the importance of efferent signaling at many levels of the auditory system. Olivocochlear neurons form a feedback loop from the superior olivary complex to hair cells in the cochlea. Their cell bodies lie around the major dense clusters of cell bodies in the olivary nuclei. In mammals, two groups of olivocochlear neurons have been functionally distinguished. The medial olivocochlear neurons' axons terminate on the outer hair cells bilaterally; the lateral olivocochlear neuron axons terminate ipsilaterally on the afferent fibers associated with inner hair cells.

Most medial olivocochlear neurons, with cell bodies that lie ventral and medial within the olivary complex, send their axons to the contralateral cochlea (Figure 28–7), but many also project to the ipsilateral cochlea. These cholinergic neurons act on hair cells through a special class of nicotinic acetylcholine receptor-channels formed from α9 and α10 subunits. The influx of Ca^{2+} through these channels leads to the opening of K^+ channels that hyperpolarize outer hair cells. These neurons thus mediate tuned negative feedback and are binaural, being driven predominantly but not exclusively by stellate cells of the contralateral ventral cochlear nucleus. Activity in these efferent fibers reduces the sensitivity of the cochlea and protects it from damage by loud sounds. Collateral branches of medial olivocochlear neurons terminate on stellate cells in the cochlear nucleus, acting on conventional nicotinic and muscarinic acetylcholine receptors, forming an excitatory feedback loop.

Figure 28–7

Major components of the ascending and descending auditory pathways. The auditory pathway is bilaterally symmetrical; the major connections among the nuclei that form the early auditory pathway are shown. The ascending pathway begins in the cochlea and progresses through several parallel pathways through the cochlear nuclei: the cochlear nuclei, the superior olivary nuclei, and the ventral and dorsal nuclei of the lateral lemniscus. These signals converge in the inferior colliculus, which projects to the medial geniculate body of the thalamus and thence to the cerebral cortex (see Figure 28–2). Some of the connections are through excitatory pathways (**colored lines**) and others through inhibitory pathways (**black lines**). These same nuclei are also interconnected through descending pathways (**blue lines**) and bilaterally through commissural projections. (**LSO**, lateral superior olivary nucleus; **MNTB**, medial nucleus of the trapezoid body; **MSO**, medial superior olive; **VNTB**, ventral nucleus of the trapezoid body).



Lateral olivocochlear neurons, with cell bodies that lie in and around the lateral superior olive, send their axons exclusively to the ipsilateral cochlea, where they terminate on the afferent fibers from inner hair cells. Charles Liberman and his colleagues demonstrated that these efferents balance the excitability of cochlear nerve fibers at the two ears.

Ventral and Dorsal Nuclei of the Lateral Lemniscus Shape Responses in the Inferior Colliculus With Inhibition

Fibers from the cochlear and superior olivary nuclei run in a band, or lemniscus, along the lateral edge of the brain as they ascend from the brainstem to the inferior colliculus. Along this band of fibers are groups of neurons that form the dorsal and ventral nuclei of the lateral lemniscus. Neurons in the ventral nuclei of the lateral lemniscus receive input from all major groups of principal cells of the ventral cochlear nuclei and respond predominantly to monaural input, driven by the contralateral ear, while neurons in the dorsal nucleus receive input from the lateral and medial

superior olivary nuclei and respond to inputs from both ears. Neurons in both subdivisions are inhibitory and project to the inferior colliculus. Their roles are intriguing but not fully understood.

Since understanding the meaning of sounds is not greatly compromised by the loss of one ear, it would make sense that the largely monaural functions of the ventral nuclei of the lateral lemniscus involve the processing of the meaning of sounds. Furthermore, mammals vary in the information they extract from their acoustic environments, which may account for differences between species in the structure and function of the ventral nuclei of the lateral lemniscus.

A border that is more distinct in some mammalian species than in others separates the ventral and intermediate nuclei and the subdivisions of the ventral nucleus of the lateral lemniscus. Neurons differ in their shapes, biophysical properties, and pattern of convergence of cochlear nuclear inputs. One group of glycinergic neurons is innervated by large calyceal terminals from octopus cells. These could generate inhibitory temporal reference signals in the inferior colliculus. Some broadly tuned neurons fire almost exclusively at the onset of tones with sharply timed action potentials but convey periodicity in complex sounds, raising the question of whether these neurons might have a role in encoding pitch in music and speech. Others respond by firing as long as a tone is present; these neurons track the fluctuations in intensity or the envelopes of sounds, a feature that is useful for understanding the meaning of sounds including speech. Tuning curves of the neurons are variable, with many being broad or W-shaped.

Neurons in the dorsal nucleus are predominantly binaural, receiving input from the ipsilateral medial superior olive and from the lateral superior olive, primarily from the contralateral side. These neurons are GABAergic, targeting the inferior colliculi on both sides and also targeting the contralateral dorsal nucleus of the lateral lemniscus. Excitation in neurons of the dorsal nucleus is amplified by NMDA-type glutamate receptors so that the inhibition they generate in their targets outlasts sound stimuli for tens of milliseconds and thus has been termed persistent inhibition. To localize sounds accurately, animals must ignore the reflections of sounds from surrounding surfaces that arrive after the initial direct wave front. Psychophysical experiments have shown that mammals suppress all but the first-arriving sound, a phenomenon termed the *precedence effect*. It has been proposed that persistent inhibition in the inferior colliculus from the dorsal nucleus of the lateral lemniscus serves to suppress spurious localization cues such as echoes and thus that it contributes to the precedence effect.

Afferent Auditory Pathways Converge in the Inferior Colliculus

The inferior colliculus occupies a central position in the auditory pathway of all vertebrate animals because all auditory pathways ascending through the brain stem converge there (Figure 28–7). The most important sources of excitation are stellate cells from the contralateral ventral cochlear nucleus, fusiform cells from the contralateral dorsal cochlear nucleus, principal cells of the ipsilateral medial superior olive and of the contralateral lateral superior olive, principal cells of ipsilateral and contralateral dorsal nuclei of the lateral lemniscus, commissural connections from the contralateral inferior colliculus, and pyramidal cells in layer V of the auditory cortex. Important sources of inhibition include the nuclei of the lateral lemniscus, the ipsilateral lateral superior olive, the superior paraolivary nucleus, and the contralateral inferior colliculus.

The inferior colliculus of mammals is subdivided into the central nucleus, dorsal cortex, and external cortex. The central nucleus is tonotopically organized. Low frequencies are represented dorsolaterally and high frequencies ventromedially in laminae that have similar best frequencies. Fine mapping has shown that the tonotopic organization is discontinuous; the separation between best frequencies corresponds to psychophysically measured critical bands of approximately one-third octave. Although the central nucleus is organized tonotopically, the spectral range of inputs to these neurons is broader than at earlier stages in the auditory pathway. Inhibition can be broad and narrows the responses of excitatory neurons. Furthermore, tuning can be modulated by descending inputs from the cortex.

Many neurons in the central nucleus carry information about the location of sound sources. The majority of these cells are sensitive to interaural time and intensity differences, essential cues for localizing sounds in the horizontal plane. Neurons are also sensitive to spectral cues that localize sounds in the vertical plane. Physiological correlates of the precedence effect have been measured in the inferior colliculus, where inhibition suppresses simulated reflections of sounds.

The inferior colliculus is not only a convergence point but also a branch point for ascending or outflow pathways. Neurons of the central nucleus project to the external cortex of the inferior colliculus and also to the thalamus and the nucleus of the brachium of the inferior colliculus, both of which then project to the superior colliculus (or the optic tectum in birds).

Sound Location Information From the Inferior Colliculus Creates a Spatial Map of Sound in the Superior

Colliculus

The inferior colliculus is not only a convergence point but also a branch point for ascending or outflow pathways. Central nucleus neurons project to the thalamus and also to the external cortex of the inferior colliculus and the nucleus of the brachium of the inferior colliculus, both of which then project to the superior colliculus (or the optic tectum in birds).

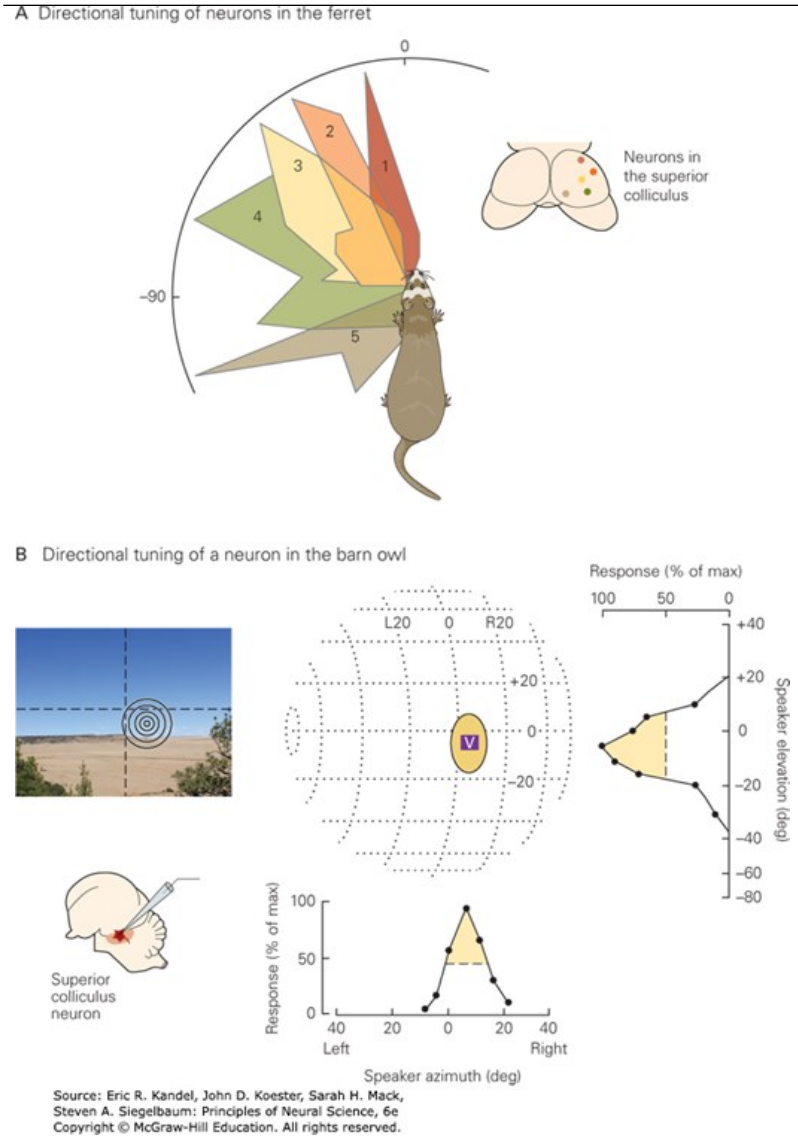
The superior colliculus is critical for reflexive orienting movements of the head and eyes to acoustic and visual cues in space. By the time the binaural sound cues and the monaural spectral cues that underlie mammalian sound localization reach the superior colliculus, they have been merged to create a spatial map of sound in which neurons are unambiguously tuned to specific sound directions. This convergence is critical since binaural differences in level and timing alone cannot unambiguously code for a single position in space. The spectral cues that provide information about vertical location must be taken into account, as different locations in the vertical plane can give rise to identical interaural differences in time or intensity. Such unambiguous spatial mapping occurs both in birds and in some mammals (Figure 28–8). In ferrets and guinea pigs, it occurs in the external cortex and the nucleus of the brachium of the inferior colliculus.

Figure 28–8

A spatial map of sound is formed in the superior colliculus.

A. Neurons in the ferret's superior colliculus are directionally tuned to sound in the horizontal plane. The illustration shows the firing rate profiles of collicular neurons 1 through 5 as a function of where the sounds are located, plotted in polar coordinates centered on the head. The drawing on the right shows the location of the recorded neurons in the colliculus. Note that neuron 1 responds best to sounds in front of the animal, whereas neurons that are located progressively more caudally in the colliculus gradually shift their responses to sounds that originate farther contralaterally. (Adapted, with permission, from King 1999.)

B. The normalized responses of a neuron in the superior colliculus of a barn owl to noise bursts presented at various locations along the horizon are plotted below (*bottom right*). The **yellow areas** in these tuning curves indicate where responses exceed 50% of the maximum. The sensitivity of the neuron to a particular location along the horizon or a particular elevation (*top right*) creates a discrete best auditory area in space for this neuron (*top middle*), shown as the colored ellipse on a plot of spatial locations with respect to a point straight in front of the owl. The neuron also responds to visual cues from the same area (the box labeled **V**). The photo illustrates the neuron's best area in space with respect to the position of the head (the intersection of the vertical and horizontal dotted lines indicates where the owl's head is pointing). The recording site for this neuron is also shown. (Adapted, with permission, from Cohen and Knudsen 1999. Copyright © 1999 Elsevier Science.)



Within the superior colliculus, the auditory map is aligned with maps of visual space and the body surface. Unlike the visual and somatosensory spatial maps, the auditory spatial map does not reflect the peripheral receptor surface; instead, it is computed from a combination of cues that identify the specific position of a sound source in space.

Auditory, visual, and somatosensory neurons in the superior colliculus all converge on output pathways in the same structure that controls orienting movements of the eyes, head, and external ears. The motor circuits of the superior colliculus are mapped with respect to motor targets in space and are aligned with the sensory maps. Such sensory-motor correspondence facilitates the sensory guiding of movements.

The Inferior Colliculus Transmits Auditory Information to the Cerebral Cortex

Auditory information ascends from the inferior colliculus to the medial geniculate body of the thalamus and from there to the auditory cortex. The pathways from the inferior colliculus include a lemniscal or core pathway and extralemniscal or belt pathways. Descending projections from the auditory cortex to the medial geniculate body are prominent both anatomically and functionally.

Stimulus Selectivity Progressively Increases Along the Ascending Pathway

A marked feature of auditory neurons at structures along the ascending pathway is their progressively increased stimulus selectivity. An auditory nerve fiber is primarily selective to one stimulus dimension, the frequency of a pure tone. The stimulus selectivity of neurons in the central auditory system

may be multidimensional, such as frequency, spectral bandwidth, sound intensity, modulation frequency, and spatial location. In this multidimensional acoustic space, neurons become more selective at successive auditory areas along the ascending pathway.

Many neurons in the auditory cortex (especially those in upper cortical layers) are highly selective to acoustic stimuli, such that the preferred (nearly optimal) stimulus of a neuron occupies only a small region of its receptive field in the multidimensional acoustic space. The region of the preferred stimulus becomes increasingly smaller at structures along the path to the auditory cortex (Figure 28–9A). Pure tones and broadband noises are two extreme cases of a wide range of acoustic stimuli that could preferentially drive auditory cortex neurons. The majority of neurons in the auditory cortex are preferentially driven by stimuli with greater spectral and temporal complexity than pure tones and broadband noises.

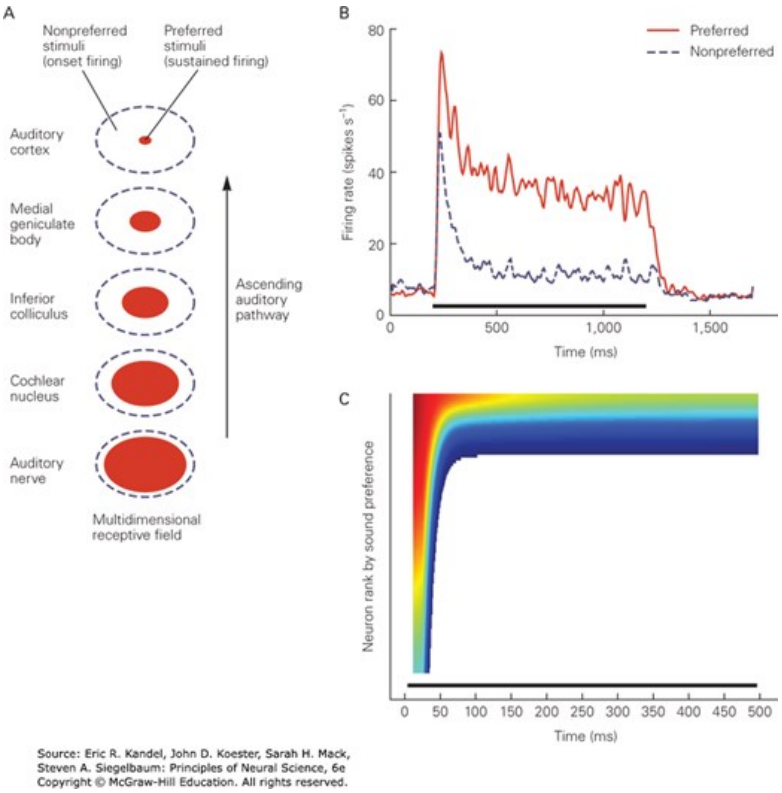
Figure 28–9

Stimulus selectivity increases along the ascending auditory pathway.

A. Stimulus selectivity and the relationship between sustained and onset firings along the ascending auditory pathway. Each open ellipse represents the multidimensional receptive field (RF) of a neuron illustrated on a two-dimensional plane. The filled ellipse represents the "sustained firing region" (corresponding to preferred stimuli) of a neuron's RF. The rest of the area within the RF is the "onset firing region" (corresponding to nonpreferred stimuli). A neuron exhibits sustained or onset firing depending on which region of the RF is stimulated. The neuron does not fire if stimuli fall outside the RF. (Adapted, with permission, from Wang 2018.)

B. Population-averaged firing rate in response to each neuron's preferred and nonpreferred stimuli from primary auditory cortex (A1). Extracellular recordings were made in awake marmoset monkeys. **Thick bar** = stimulus duration. (Adapted, with permission, from Wang et al. 2005. Copyright © 2005 Springer Nature.)

C. Distribution of activity among A1 neurons in response to a sound burst. On the *y*-axis, all A1 neurons are ranked according to their preference for a particular stimulus. The **blue-to-red color gradient** represents increasing firing rate. The neuron with the highest firing rate is located at the top end of the *y*-axis. **Black bar** = stimulus duration. Most neurons show a brief phasic response to the onset of the sound, but only those particularly tuned to the sound maintain their response until the end of the sound. (Adapted, with permission, from Middlebrooks 2005. Copyright © 2005 Springer Nature.)



The increased stimulus selectivity is also accompanied by changes in a neuron's firing pattern. When neurons are driven by their preferred stimuli, they respond not only with higher firing rates but also with sustained firing throughout the stimulus duration (Figure 28–9B). The receptive field of a cortical neuron contains a "sustained firing region" (corresponding to preferred stimuli) within a larger "onset firing region" (corresponding to nonpreferred stimuli). This explains why it is common for experimenters to observe onset (phasic) responses in auditory cortex when a continuous sound is played.

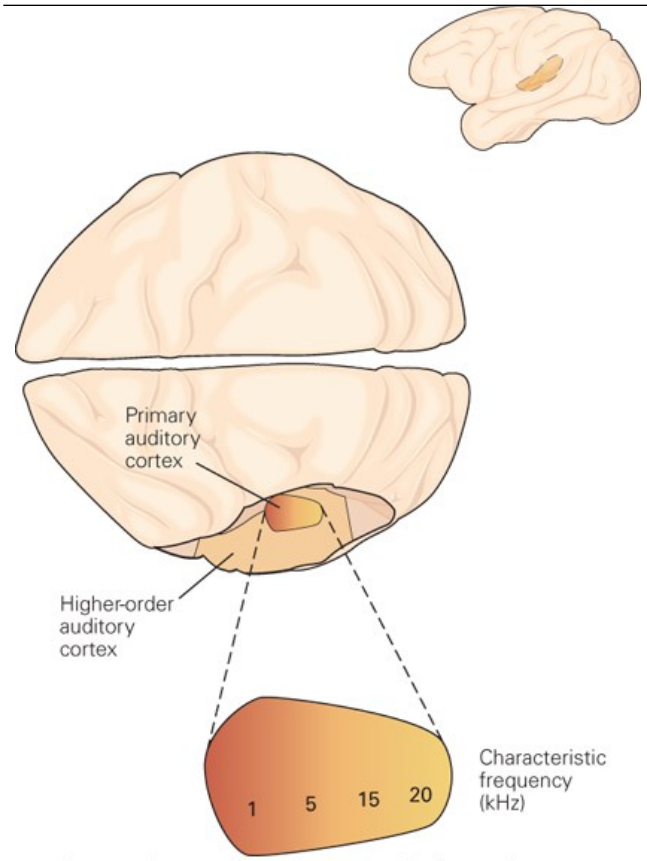
The discovery of how sustained firing is evoked in the auditory cortex is important because it provides a direct link between neural firing and the perception of a continuous acoustic event. Such sustained firing by auditory cortex neurons has been observed only in awake animals. In contrast, an auditory nerve fiber typically shows sustains firing in response to a wide range of acoustic signals as long as the spectral energy of the stimulus falls within the neuron's receptive field, under either anesthetized or awake conditions. When David Hubel and his colleagues ventured into the auditory cortex more than half a century ago, they were puzzled by how difficult it was to drive neurons in the auditory cortex of awake cats. Now we know it was because they were probably recording from highly selective neurons and using nonpreferred stimuli. The availability of digital technology since then has made it possible to create and test a large battery of acoustic stimuli in search of the preferred stimulus of a highly selective neuron in auditory cortex. The overall picture elucidated by experimenters is that when a sound is heard, the auditory cortex first responds with transient discharges (encoding the onset of a sound) across a relatively large population of neurons. As the time passes, the activation becomes restricted to a smaller population of neurons that are preferentially driven by the sound (Figure 28–9C), which results in a selective representation of the sound within the neuronal population and over time. Because each neuron has its own preferred stimulus that differs from preferred stimuli of other neurons, neurons in the auditory cortex collectively cover the entire acoustic space with their sustained firing regions. Therefore, any particular sound can evoke sustained firing throughout its duration in a particular population of neurons in the auditory cortex. In other words, the region of auditory cortex activated by acoustic stimulation in whole-brain imaging (eg, functional magnetic resonance imaging [fMRI], positron emission tomography [PET]) comprises neurons that are preferentially driven by the acoustic stimulus.

The Auditory Cortex Maps Numerous Aspects of Sound

The auditory cortex includes multiple distinct functional areas on the dorsal surface of the temporal lobe. The most prominent projection is from the ventral division of the medial geniculate nucleus to the primary auditory cortex (A1, or Brodmann's area 41). As in the subcortical structures, the neurons in this cytoarchitectonically distinct region are arranged tonotopically. In monkeys, neurons tuned to low frequencies are found at the rostral end of A1, while those responsive to high frequencies are in the caudal region (Figure 28–10). Thus, like the visual and somatosensory cortices, A1 contains a map reflecting the sensory periphery.

Figure 28–10

The auditory cortex of primates has multiple primary and secondary areas. The expanded figure of the primary auditory cortex shows its tonotopic organization. The primary areas are surrounded by higher-order areas (see Figure 28–11).



Source: Eric R. Kandel, John D. Koester, Sarah H. Mack, Steven A. Siegelbaum: Principles of Neural Science, 6e

Because the cochlea encodes discrete frequencies at different points along the basilar membrane, however, a one-dimensional frequency map from the periphery is spread across the two-dimensional surface of the cortex, with a smooth frequency gradient in one direction and isofrequency contours along the other direction. In many species, subregions of the auditory cortex that represent biologically significant frequencies are larger than others because of extensive inputs, similar to the large area in the primary visual cortex devoted to inputs from the fovea.

In addition to frequency, other features of auditory stimuli are mapped in the primary auditory cortex, although the overall organization is less clear and precise than for vision. Auditory neurons in A1 are excited either by input from both ears (EE neurons), with the contralateral input usually stronger than the ipsilateral contribution, or by a unilateral input (EI). The EI neurons are inhibited by stimulation of the opposite ear.

Certain neurons in A1 also seem to be organized according to bandwidth, that is, according to their responsiveness to a narrow or broad range of frequencies. Neurons near the center of the isofrequency contours are tuned more narrowly to bandwidth or frequency than those located away from the center. Distinct subregions of A1 form clusters of cells with narrow or broadband tuning within individual isofrequency contours. Within intracortical circuits, neurons receive input primarily from neurons with similar bandwidths and characteristic frequencies. This modular organization of bandwidth selectivity may allow redundant processing of incoming signals through neuronal filters of varying bandwidths as well as center frequencies, which could be useful for the analysis of spectrally complex sounds such as species-specific vocalizations, including speech.

Several other parameters are represented in A1. These include neuronal response latency, loudness, modulation of loudness, and the rate and direction of frequency modulation. Although it remains to be seen how these various maps intersect, this array of parameters clearly endows each neuron and each location in A1 with the ability to represent many independent variables of sound and thus allows for a great diversity of neuronal selectivity.

As is true for visual and somatosensory areas of the cortex, sensory representation in A1 can change in response to alterations in input pathways. After peripheral hearing loss, tonotopic mapping in A1 can be altered so that neurons that were previously responsive to sounds within the lost range of hearing will begin to respond to adjacent frequencies. The work of Michael Merzenich and others has shown that behavioral training of adult animals can also result in large-scale reorganization of the auditory cortex, so that the most behaviorally relevant frequencies—those specifically associated

with attention or reinforcement—come to be overrepresented.

The auditory areas of young animals are particularly plastic. In rodents, the frequency organization of A1 emerges gradually during development from an early, crude frequency map. Raising animals in acoustic environments in which they are exposed to repeated tone pulses of a particular frequency results in a persistent expansion of cortical areas devoted to that frequency, accompanied by a general deterioration and broadening of the tonotopic map. This result not only suggests that the development of A1 is experience-dependent but also raises the possibility that early exposure to abnormal sound environments can create long-term disruptions of high-level sensory processing. A greater understanding of how this happens and whether it is also true for human fetuses and infants may provide insights into the origin and remediation of disorders in which central auditory processing is impaired, such as many forms of dyslexia. Moreover, the ability to induce plasticity in the auditory cortex of adults by engaging attention or reward raises new hopes for brain repair even in adulthood.

The primary auditory area of mammals is surrounded by multiple distinct regions, some of which are tonotopic. Adjacent tonotopic fields have mirror-image tonotopy: The direction of tonotopy reverses at the boundary between fields. In monkeys, as many as 7 to 10 secondary (belt) areas surround the three or four primary or primary-like (core) areas (see Figure 28–11). The secondary areas receive input from the core areas of the auditory cortex and, in some cases, from thalamic nuclei. Electrophysiological and imaging studies have confirmed that A1 in humans lies on Heschl's gyrus, in the temporal lobe, medial to the Sylvian fissure. In addition, recent fMRI studies have revealed that in humans, just as in monkeys, pure tones activate primarily core areas, whereas the neurons of belt areas prefer complex sounds such as narrowband noise bursts.

A Second Sound-Localization Pathway From the Inferior Colliculus Involves the Cerebral Cortex in Gaze Control

Many neurons in the auditory cortex have broad spatial tuning, but neurons with narrow spatial tuning are also found when studied in awake animals. In monkeys, auditory cortex neurons are tuned to both frontal space and rear space (outside the coverage of vision), as well as the space above and below the horizontal plane. In contrast to the auditory midbrain, however, there is yet no evidence for a spatially organized map of sound in any of the cortical areas sensitive to sound location.

The sound-localization pathways in the cortex originate in the central nucleus of the inferior colliculus and ascend through the auditory thalamus and the primary and secondary cortical areas, eventually reaching the frontal eye fields involved in gaze control. Eye or head movements can be elicited by stimulating the frontal eye fields, which connect directly to brain stem tegmentum premotor nuclei that mediate gaze changes as well as to the superior colliculus. But why should there be this second sound-localization pathway connected to gaze control circuitry when the midbrain pathway from location-sensitive neurons in the inferior colliculus to the superior colliculus to gaze control circuitry directly controls orientation movements of the head, eyes, and ears?

Behavioral experiments shed light on this question. Although lesions of A1 can result in profound sound-localization deficits, no deficiency is seen when the task is simply to indicate the side of the sound source by pushing a lever. The deficit becomes apparent only when the animal must approach the location of a brief sound source; that is, when the task is the more complex one of forming an image of the source, remembering it, and moving toward it.

Experiments in barn owls have produced particularly compelling evidence. The ability of owls to orient to sounds in space is unaffected by inactivation of the avian equivalent of the frontal eye fields. Similarly, when the midbrain sound localization pathway is disrupted by pharmacological inactivation of the superior colliculus, the probability of an accurate head turn is decreased, but animals still respond correctly more than half of the time. In contrast, when both structures are inactivated, animals are completely unable to orient accurately to acoustic stimuli on the contralateral side. Thus, cortical and subcortical sound-localization pathways have parallel access to gaze control centers, perhaps providing some redundancy. Moreover, when only the frontal eye fields are inactivated, birds lose their ability to orient their gaze toward a target that has been extinguished and must be remembered, just as is seen with mammalian A1 lesions. Thus, in both mammals and birds, cortical pathways are required for more complex sound-localization tasks.

This appears to be a general difference between cortical and subcortical pathways. Subcortical circuits are important for rapid and reliable performance of behaviors that are critical to survival. Cortical circuitry allows for working memory, complex recognition tasks, and selection of stimuli and evaluation of their significance, resulting in slower but more differentiated performance. Examples of this also exist in auditory pathways not involved in localization. Conditioned fear responses to simple auditory stimuli are mediated by direct rapid pathways from the auditory thalamus to the amygdala; they can still be elicited after cortical inactivation. However, fear responses that require more complex discrimination of auditory stimuli require pathways through the cortex and are accordingly slower but more specific.

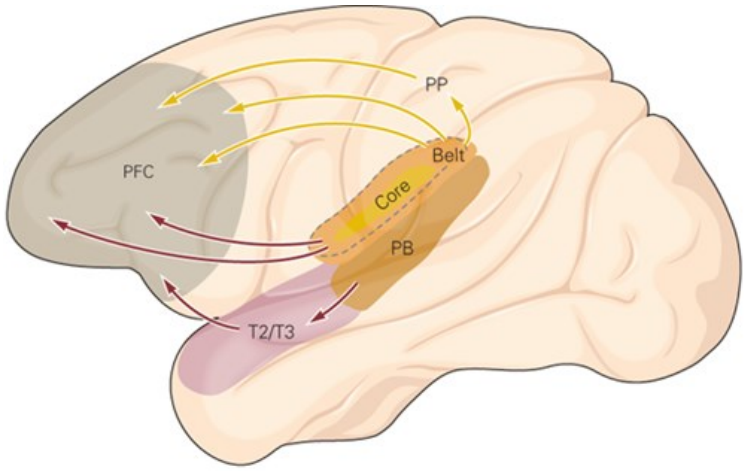
Auditory Circuits in the Cerebral Cortex Are Segregated Into Separate Processing Streams

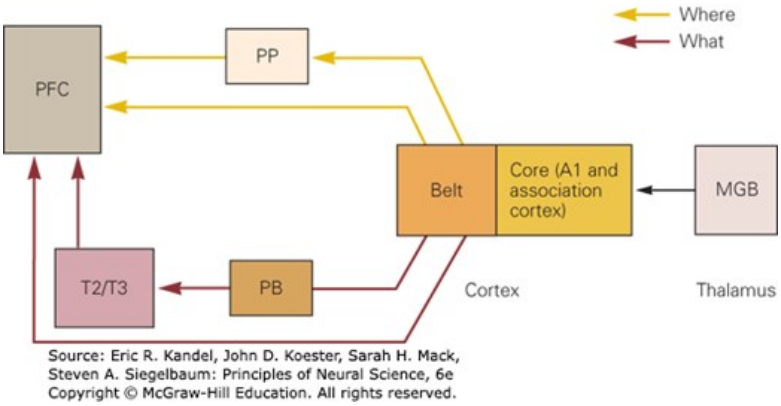
In the visual system, the output from the primary visual cortex is segregated into separate dorsal and ventral streams concerned respectively with object location in space and object identification. A similar division of labor is thought to exist in the somatosensory cortex, and recent evidence suggests that the auditory cortex also follows this plan.

Anatomical tracing studies of the three most accessible belt areas in monkeys show that the more rostral and ventral areas connect primarily to the more rostral and ventral areas of the temporal lobe, whereas the more caudal area projects to the dorsal and caudal temporal lobe. In addition, these belt areas and their temporal lobe targets both project to largely different areas of the frontal lobes (Figure 28–11).

Figure 28–11

The "what" and "where" streams in the auditory cortical system of primates. The ventral "what" stream and dorsal "where" stream originate in different parts of primary and belt cortex and ultimately project to distinct regions of prefrontal cortex through independent paths. (**MGB**, medial geniculate body of the thalamus; **PB**, parabelt cortex; **PFC**, prefrontal cortex; **PP**, posterior parietal cortex; **T2**/**T3**, areas of temporal cortex.) (Adapted, with permission, from Rauschecker and Tian 2000. Copyright 2000 National Academy of Sciences; adapted from Romanski and Averbeck 2009.)





The frontal areas receiving anterior auditory projections are generally implicated in nonspatial functions, whereas those that are targets of posterior auditory areas are implicated in spatial processing. Electrophysiological and imaging studies provide support for this. Caudal and parietal areas are more active when a stimulus must be localized or moves, and ventral areas are more active during identification of the same stimulus or analysis of its pitch. Thus anterior-ventral pathways may identify auditory objects by analyzing spectral and temporal characteristics of sounds, whereas the more dorsal-posterior pathways may specialize in sound-source location, detection of sound-source motion, and spatial segregation of sources.

Although the idea that all sensory areas of the cerebral cortex initially segregate object identification and location is attractive, it is likely an oversimplification. It is clear that the medial-belt areas of the auditory cortex project to both dorsal and ventral frontal cortices, and neurons with broad spatial responsiveness are distributed throughout caudal and anterior areas. Nonetheless, although the details may differ between systems, the basic concept holds that sensory systems deconstruct stimuli into features and analyze each type in discrete pathways.

The Cerebral Cortex Modulates Sensory Processing in Subcortical Auditory Areas

An intriguing feature of all mammalian cortical areas, and one shared by the auditory system, is the massive projection from the cortex back to lower areas. There are almost 10 times as many corticofugal fibers entering the sensory thalamus as there are axons projecting from the thalamus to the cortex. Projections from the auditory cortex also innervate the inferior colliculus, olivocochlear neurons, some basal ganglionic structures, and even the dorsal cochlear nucleus.

Insights into possible functions of this feedback have come from the bat's auditory system. Silencing of frequency-specific cortical areas leads to decreased responses in thalamus and inferior colliculus in the corresponding frequency-specific areas, whereas activation of cortical projections increases and sharpens the responses of some neurons. The auditory cortex can therefore actively adjust and improve auditory signal processing in subcortical structures. A variety of evidence suggests that cortical feedback also occurs in other mammals. This challenges the view of ascending sensory pathways as purely feedforward circuits and suggests that we should regard the thalamus and cortex as reciprocally and highly interconnected circuits in which the cortex exercises some top-down control of perception.

The Cerebral Cortex Forms Complex Sound Representations

The Auditory Cortex Uses Temporal and Rate Codes to Represent Time-Varying Sounds

An important function of the auditory system is to represent time-varying sounds across multiple time scales, from a few milliseconds to tens and hundreds of milliseconds or even longer. In the auditory nerve, firing patterns largely mirror the temporal structure of sounds, firing in phase with sounds to the limit of the phase-locking. The precision of this temporally based neural representation gradually decreases as information ascends toward the auditory cortex due to synaptic integration at the soma and dendrites.

The upper limit of the phase-locking to periodic sounds progressively decreases along the ascending auditory pathway from approximately 3,000 Hz in the auditory nerve to less than approximately 300 Hz in the medial geniculate body in the thalamus and less than 100 Hz in A1. The upper limit of the phase-locking in A1 is similar to that found in the primary visual and somatosensory areas of cortex. In the auditory cortex, the temporal firing pattern alone is inadequate to represent the entire range of time-varying sounds that are perceived by humans and animals.

Cortical neurons use an alternative method to represent time-varying sounds that change more rapidly than the upper limit of the phase-locking in A1. When an animal listens to a sequence of periodic clicks, two types of neural responses are observed in A1. One population of neurons displays phase-locked periodic firing in response to click trains with long intervals between clicks or slowly varying sounds, but not to click trains with short intervals between clicks or rapidly varying sounds (Figure 28–12A). The second population of neurons does not respond to click trains at long interclick intervals, but instead fires increasingly rapidly as the interclick interval becomes shorter (Figure 28–12B). These two populations of A1 neurons, referred to as *synchronized* and *nonsynchronized*, respectively, have complementary response properties. Neurons of the synchronized population *explicitly* represent slowly occurring sound events by synchronized neural firing (a temporal code), whereas neurons of the nonsynchronized population *implicitly* represent rapidly changing sound events by changes in average firing rates (a rate code).

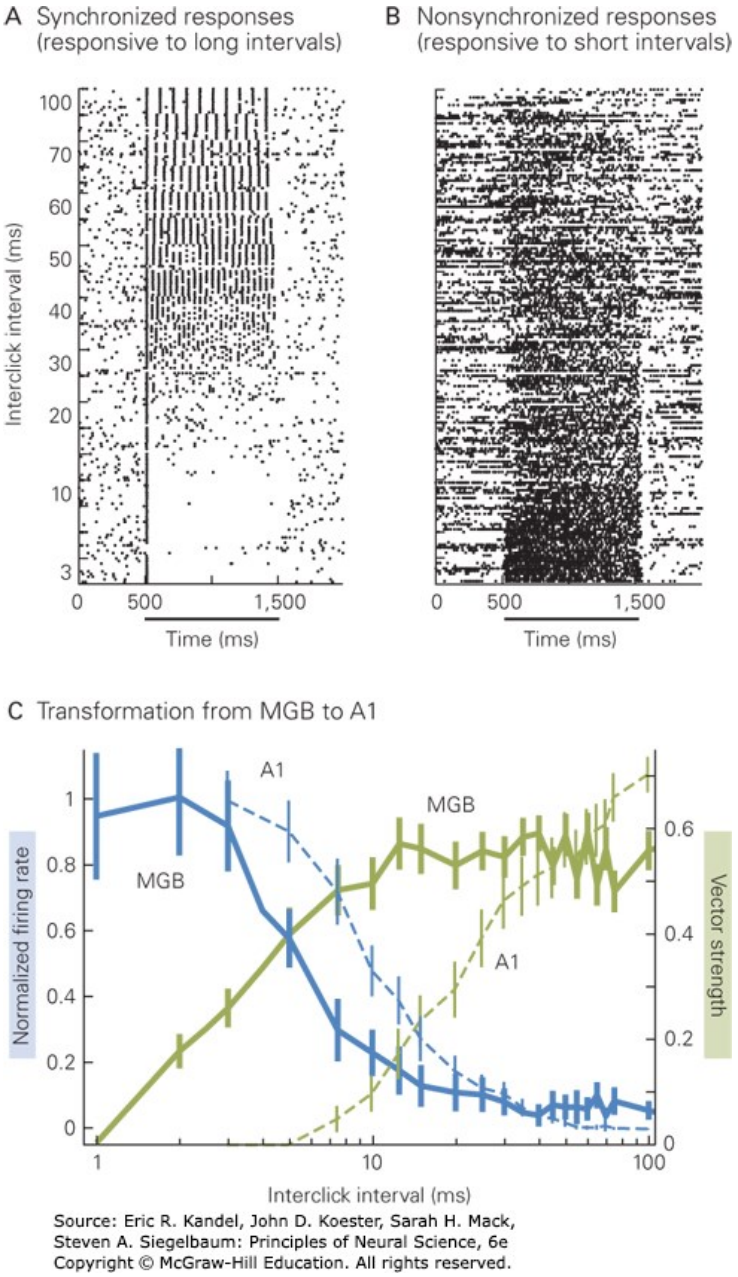
Figure 28–12

Temporal and rate coding of time-varying sounds.

A. Stimulus-synchronized responses of a neuron to periodic click trains recorded from A1 of an awake marmoset. The horizontal bar below the *x*-axis indicates the duration of the stimulus. (Adapted, with permission, from Lu, Liang, and Wang 2001. Copyright © 2001 Springer Nature.)

B. Nonsynchronized responses of a neuron to periodic click trains recorded from A1 in an awake marmoset. (Adapted, with permission, from Lu, Liang, and Wang 2001. Copyright © 2001 Springer Nature.)

C. Comparison of temporal response properties between primary auditory cortex (**A1**) and medial geniculate body of the thalamus (**MGB**). Stimulus-synchronized responses are quantified by vector strength, a measure of the strength of phase-locking. Nonsynchronized responses are quantified by the normalized firing rate (data curves identified as A1 rate and MGB rate). Error bars represent standard error of the mean. (Adapted, with permission, from Bartlett and Wang 2007.)



The nonsynchronized neurons have been observed in the auditory cortex of awake primates and rodents. In A1, neural representation changes from a temporal code to a rate code at the interclick interval of about 25 ms, corresponding to a repetition rate of approximately 40 Hz (Figure 28–12A,B). This is near the boundary where our perception of a periodic click train changes from being “discrete” to “continuous.”

The combination of temporal and rate codes to represent the whole range of time-varying sounds is the consequence of a progressive transformation beginning in the auditory nerve, where only a temporal code (phase-locking) is available. The progressive reduction in the upper limit of the phase-locking along the ascending auditory pathway is accompanied by the emergence of firing-rate-based representations. In the medial geniculate body of the thalamus, the intersection between temporal and rate codes is at a shorter interclick interval than in A1 (Figure 28–12C). This indicates that neurons in the medial geniculate body can phase-lock to more rapidly time-varying sounds than A1 neurons, but still utilize a rate code to represent

rapidly changing sounds beyond their phase-locking limit.

The prevalence of rate-coding neurons in A1 has important functional implications. It shows that a considerable transition from temporal to rate coding has taken place by the time auditory signals reach the auditory cortex. The importance of the nonsynchronized neural responses is that they represent transformed instead of preserved temporal information. It suggests that cortical processing of sound streams operates on a segment-by-segment basis rather than on a moment-by-moment basis, as found in the auditory nerve. This is necessary for complex integration because higher-level processing tasks require temporal integration over a time window. The reduction in A1 of the upper limit of phase-locking is a prerequisite for multisensory integration in the cerebral cortex. Auditory information is encoded at the periphery at a much faster temporal modulation rate than visual or tactile information, but phase-locking is similar across primary sensory areas of the cortex. The slowing of the phase-locking limit along the ascending auditory pathway and accompanying transition from a temporal code to a rate code are necessary for auditory information to be integrated in the cerebral cortex with information from other sensory modalities that are intrinsically slower.

Primates Have Specialized Cortical Neurons That Encode Pitch and Harmonics

Pitch perception is crucial for perceiving speech and music and for recognizing auditory objects in a complex acoustic environment. Pitch is the percept that allows harmonically structured periodic sounds to be perceived and ordered on a musical scale. Pitch carries crucial linguistic information in tonal languages such as Chinese and prosodic information in European languages. We use pitch to identify a particular voice from a noisy background in a cocktail party. When listening to an orchestra, we hear the melody of the soloist over the background of accompanying instruments.

An important phenomenon for understanding pitch is the perception of “missing fundamental,” also referred to as the residue pitch. When the harmonics of a fundamental frequency are played together, the pitch is perceived as the fundamental frequency even if the fundamental frequency is missing. For example, the harmonics of the fundamental frequency of 200 Hz are at 400, 600, 800 Hz, and so on. Playing the frequencies 400, 600, and 800 Hz together will generate a pitch perception of 200 Hz, even though a distinct frequency component of 200 Hz is not physically present in the sound. We encounter this phenomenon routinely when we listen to music over speakers that are too small to generate sounds at low frequencies.

Many combinations of frequencies can give rise to a common fundamental frequency or pitch, making it a particularly valuable auditory cue. This is especially useful when pitch conveys behaviorally important information, as in the case of human speech or animal vocalizations. Sounds propagated through the environment can become spectrally degraded, losing high or low frequencies. While such spectral filtering distorts spectral information, the perception of the missing fundamental is robust despite the loss of some harmonic components.

The ability to perceive pitch is not unique to humans; birds, cats, and monkeys can also pick out pitch. Monkeys are capable of spectral pitch discrimination, melody recognition, and octave generalization, each of which requires the perception of pitch. Marmoset monkeys (*Callithrix jacchus*), a highly vocal New World primate species whose hearing range is similar to that of humans, exhibit human-like pitch perception. Marmosets are able to discriminate a missing fundamental in harmonic sounds with a precision as small as one semitone for the periodicity above 440 Hz.

Given that both humans and some animals experience a pitch that generalizes across a variety of sounds with the same periodicity (including harmonic sounds with a missing fundamental), it is reasonable to expect that some neurons extract pitch from complex sounds. Xiaoqin Wang and his colleagues discovered a decade ago that a small region in the auditory cortex of marmoset monkeys contains “pitch-selective neurons.” These neurons are tuned to pure tones with a best frequency and respond to harmonic complexes with a fundamental frequency near its best frequency even when the harmonics lay outside the neuron’s excitatory-frequency response area (Figure 28–13A).

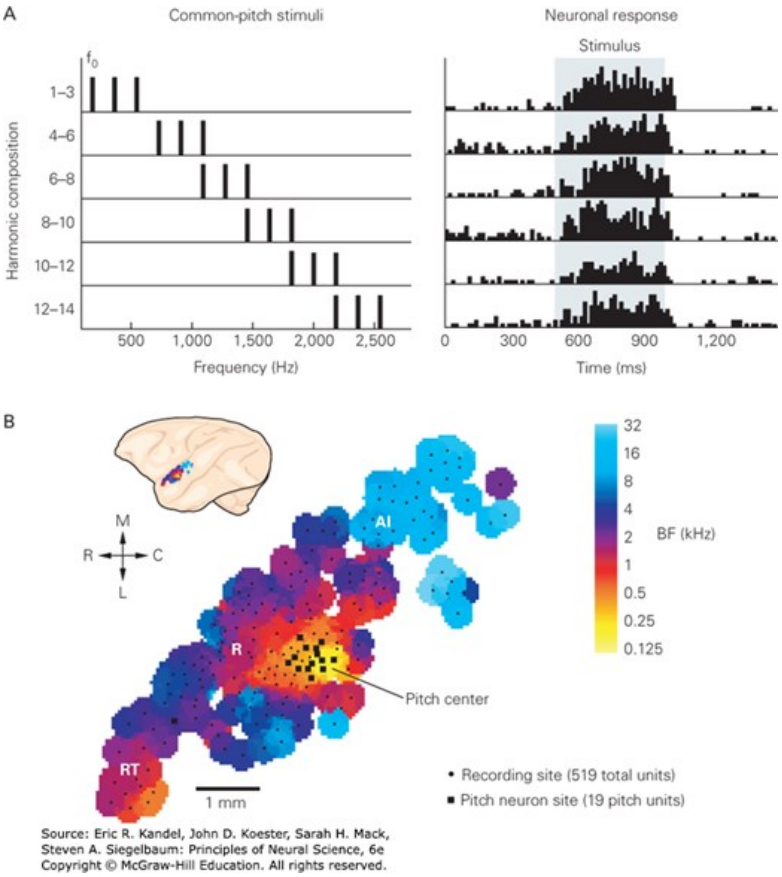
Figure 28–13

Pitch is encoded by specialized neurons in primate auditory cortex.

A. An example of a pitch-selective neuron recorded from marmoset auditory cortex. *Left*: Frequency spectra of a series of harmonic stimuli that share the same fundamental frequency ($\mathbf{f_0}$). *Right*: Peristimulus time histogram of the neuron’s response to the stimuli (stimulus duration indicated by the shaded region). (Adapted, with permission, from Bendor and Wang 2005. Copyright © 2005 Springer Nature.)

B. Anatomical organization of the marmoset auditory cortex and the location of a pitch center. **Top:** Side view of the marmoset brain. **Bottom:** Tonotopic map of the left auditory cortex characterized in one marmoset. Pitch-selective neurons (**black squares**) are clustered near the low-frequency border between A1 and area **R** (rostral auditory cortex). Frequency reversals indicate the borders between A1/R and **R/RT** (rostrotemporal

auditory cortex). (**BF**: best frequency.) (Adapted from Bendor and Wang 2005. Copyright © 2005 Springer Nature.)



A pitch-selective neuron responds to pitch-evoking sounds (eg, harmonic sounds, click trains) when the pitch is near the neuron's preferred best frequency. Pitch-selective neurons increase their firing rates as the behavioral salience of pitch increases and prefer sounds with periodicity over aperiodic sounds. It is important to note that the pitch-selective neurons in marmoset monkeys, which extract and code for pitch embedded in harmonic sounds (a highly nonlinear computation), are distinctly different from neurons in subcortical areas or A1 that merely "reflect" information on pitch in their firing patterns.

The region containing the pitch-selective neurons in marmoset monkeys is confined to the low-frequency border of A1, the rostral auditory cortex (area R), and lateral belt areas (Figure 28–13B). Human imaging studies have identified a restricted region at the lateral end of Heschl's gyrus anterolateral to A1 that extracts pitch of harmonic complex sounds and is sensitive to changes in pitch salience. The location of this region mirrors the location of the pitch center in marmoset monkeys (Figure 28–13B).

The core regions of auditory cortex in marmosets also contain a class of harmonic template neurons that respond weakly or not at all to pure tones or two-tone combinations but respond strongly to particular combinations of multiple harmonics. The harmonic template neurons show stronger responses to harmonic sounds than inharmonic sounds and selectivity for particular harmonic structures. In contrast to the pitch-selective neurons that are localized within a small cortical region lateral to the low-frequency border between A1 and R and have best frequencies less than 1,000 Hz, the harmonic template neurons are distributed across A1 and R and have best frequencies ranging from approximately 1 kHz to approximately 32 kHz, a range that covers the entire hearing range of marmosets.

Whereas in the periphery single auditory nerve fibers encode individual components of harmonic sounds, the properties of the harmonic template neurons reveal harmonically structured receptive fields for extracting harmonic patterns. The change in neural representation of harmonic sounds from auditory nerve fibers to the auditory cortex reflects a principle of neural coding in sensory systems. Neurons in sensory pathways transform the representation of physical features, such as the frequency of sounds in hearing or luminance of images in vision, into a representation of perceptual features, such as pitch in hearing or curvature in vision. Such features lead to the formation of auditory or visual percepts. The harmonic template neurons in the auditory cortex are key to processing sounds with harmonic structures such as animal vocalizations, human speech, and music.

Insectivorous Bats Have Cortical Areas Specialized for Behaviorally Relevant Features of Sound

Although it is generally assumed that upstream auditory areas perform increasingly specialized functions related to hearing, much less is known about the functions of serial relays in the auditory system compared to the visual system. In humans, one of the most important aspects of audition is its role in processing language, but we know relatively little about how speech sounds are analyzed by neural circuits. New techniques for imaging the human brain are gradually providing insights into the functional specialization of cortical areas associated with language (Chapter 55).

Evidence for specialized analysis of complex auditory signals in the cerebral cortex comes from studies of insectivorous bats. These animals find their prey almost entirely through *echolocation*, emitting ultrasonic pulses of sound that are reflected by flying insects. Bats analyze the timing and structure of the echoes to help locate and identify the targets, and discrete auditory areas are devoted to processing different aspects of the echoes.

Many bats, such as the mustached bat studied by Nobuo Suga and his collaborators, emit echolocating pulses with two components. An initial *constant-frequency* (CF) component consists of several harmonically related sounds. These harmonics are emitted stably for tens to hundreds of milliseconds, akin to human vowel sounds. The constant-frequency component is followed by a sound that decreases steeply in frequency, the *frequency-modulated* (FM) component, which resembles the rapidly changing frequency of human consonants (Figure 28–14A).

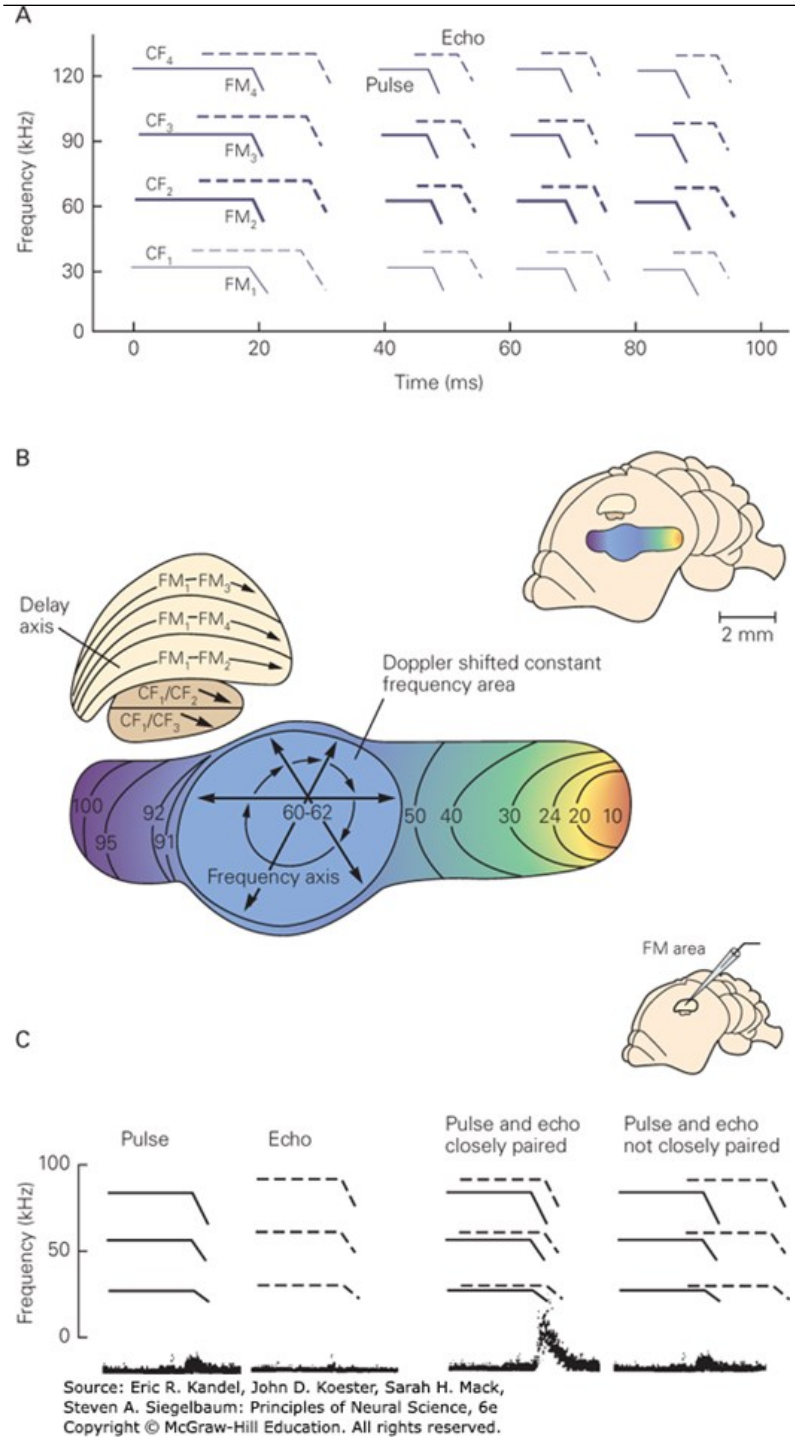
Figure 28–14

The auditory system of the bat has specialized areas for locating sounds.

A. A sonogram of an animal's calls (**solid lines**) and the resultant echoes (**dashed lines**) illustrates the two components of the call: the protracted, harmonically related constant-frequency (**CF**) signal and the briefer frequency-modulated (**FM**) signal. The duration of the calls declines as the animal approaches its target. (Adapted, with permission, from Suga 1984.)

B. A view of the cerebral hemisphere of the mustached bat shows three of the functional areas within the auditory cortex. The FM area is where the distance from the target is computed; the CF area is where the velocity of the target is computed; and the Doppler-shifted CF area is specialized for the identification of small fluttering objects. The expanded cortical representation of Doppler-shifted CF signals near the second harmonic of the call frequency (60–62 kHz) forms the acoustic "fovea." (Adapted, with permission, from Suga 1984.)

C. The FM-FM combination-sensitive neuron shown does not respond significantly to either pulses or echoes alone, but responds very strongly to a closely paired pulse-echo. However, the neuron is also sensitive to the time difference between the pulse and echo, as seen in the record on the right, where the neuron fails to respond to a pulse-echo combination that is not closely paired. (Adapted, with permission, from Suga et al. 1983.)



The FM sounds are used to determine the distance to the target. The bat measures the interval between the emitted sound and the returning echo, which corresponds to a particular distance, based on the relatively constant speed of sound. Neurons in the FM-FM area of auditory cortex (Figure 28–14B) respond preferentially to pulse-echo pairs separated by a specific delay. Moreover, these neurons respond better to particular combinations of sounds than to the individual sounds in isolation; such neurons are called *feature detectors* (Figure 28–14C). The FM-FM area contains an array of such detectors, with preferred delays systematically ranging from 0.4 to 18 ms, corresponding to target ranges of 7 to 280 cm (Figure 28–14B). These neurons are organized in columns, each of which is responsive to a particular combination of stimulus frequency and delay. In this way, the bat, like the barn owl in its inferior colliculus, is able to represent an acoustic feature that is not directly represented by sensory receptors.

The CF components of bat calls are used to determine both the speed of the target relative to the bat and the acoustic image of the target. When an

echolocating bat is flying toward an insect, the sounds reflected from the insect are Doppler-shifted to a higher frequency at the bat's ear, for the bat is moving toward the returning sound waves from the target, causing a relative speeding up of these waves at its ear. Similarly, a receding insect yields reflections of lowered frequency at the bat's ear. Neurons in the CF-CF area (Figure 28–14B) are sharply tuned to a combination of frequencies close to the emitted frequency or its harmonics. Each neuron responds best to a combination of a pulse of a particular fundamental frequency with an echo corresponding to the first or second harmonic of the pulse, Doppler-shifted to a specific extent. As in the FM-FM area, neurons do not respond to the pulse or echo alone, but rather to the combination of the two CF signals.

CF-CF neurons are arranged in columns, each encoding a particular combination of frequencies. These columns are arranged regularly along the cortical surface, with the fundamental frequency along one axis and the echo harmonics along a perpendicular axis. This dual-frequency coordinate system creates a map wherein a specific location corresponds to a particular Doppler shift and thus a particular target velocity, ranging systematically from –2 m/s to 9 m/s.

The CF components of returning echoes are also used for detailed frequency analysis of the acoustic image, presumably important in its identification. The Doppler-shifted constant-frequency area (DSCF) of the mustached bat is a dramatic expansion of the primary auditory cortex's representation of frequencies between 60 kHz and 62 kHz, corresponding well to the set of returning echoes from the major CF component of the bat's call (Figure 28–14B). Within the DSCF area, individual neurons are extremely sharply tuned to frequency, so that the tiny changes in frequency created by fluttering moth wings are easily detected.

Transient inactivation of some of these specialized cortical areas while the bat performs a discrimination task strikingly supports the importance of their functional specialization in behavior. Silencing of the DSCF selectively impairs fine frequency discrimination while leaving time perception intact. Conversely, inactivation of the FM-FM area impairs the bat's ability to detect small differences in the time of arrival of two echoes, while leaving frequency perception unchanged.

Investigation of this auditory system was greatly facilitated by knowledge of the stimuli relevant to bats. It remains to be seen whether these cortical areas are functionally or anatomically analogous to particular fields in cats, monkeys, and humans. Regardless, the choice of appropriate stimuli is likely to be as important in studying these other species as it has been in studies of bats.

The Auditory Cortex Is Involved in Processing Vocal Feedback During Speaking

Vocal communication involves both speaking and hearing, often taking place concurrently. When we speak, the sound of our voice is delivered not only to the intended listener but also back to our own ears. Such feedback to our auditory system during vocal production is conducted not only through the air but also through bone and can be loud as a result of the proximity of the mouth and the ears.

The auditory system must distinguish an auditory percept as being self-generated or externally generated. To monitor external sounds from the acoustic environment during speaking, self-generated sounds have to be masked. At the same time, the auditory system must also monitor our own voice in order to detect errors in vocal production. An accurate representation of one's own voice through vocal feedback is crucial to maintaining desired vocal production and to the learning of a new language. In humans and animals, perturbations of the vocal feedback can lead to alterations in vocal production, and interruptions or blockages of the vocal feedback can result in degradation in vocal learning.

The evidence for the involvement of the auditory cortex in processing vocal feedback comes from both human and animal studies. Responses in the auditory cortex of human subjects to their own voice while speaking are smaller than the responses to the playback of the same sounds. This reduction can be observed in electrocorticographical (ECoG) recordings (Figure 28–15A) or with a variety of imaging methods (eg, fMRI, PET, magnetoencephalography [MEG]).

Figure 28–15

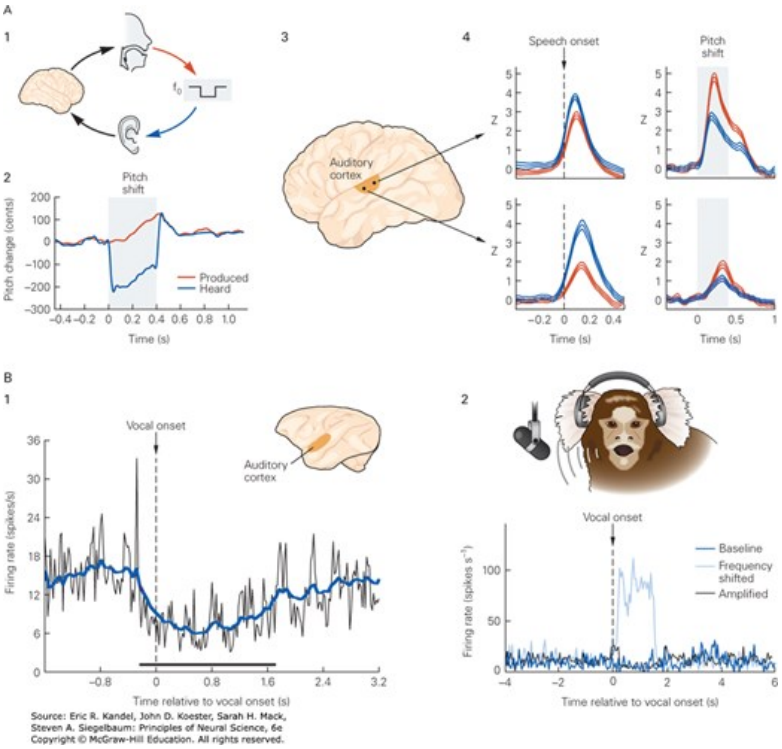
Vocal feedback processing in auditory cortex.

A. Examples of vocalization-induced suppression and sensitivity to pitch perturbation in human cerebral cortex. **1.** A subject's vocalizations (**red arrow**) went through a digital signal processor that shifted pitch and delivered the distorted auditory feedback (**blue arrow**) to the subject's earphones. **2.** Pitch track of an example trial shows the pitch recorded by the microphone (produced) and the pitch delivered to the earphones (heard). Shaded region indicates the time interval when the signal processor shifted the pitch by −200 cents (1 cent = 1/1200 octave). **3.** The locations of

electrodes that recorded from two sites in the auditory cortex on the surface of the superior temporal gyrus. **4.** The Z variable represents the power in the 50 to 150 Hz (high-γ) range of cortical activity, which has been shown to correlate well with neuronal spiking activity. It was extracted from the signals recorded at each electrode in the speaking (**red**) and listening (**blue**) conditions. Vertical lines in the left column of plots indicate vocalization onset, and shaded regions in the right column of plots indicate the onset and offset of perturbation. The response of a subject's auditory cortex to his or her own self-produced vocalization is generally smaller than the response seen when the subject passively listens to playback of the same vocalization (*left column).* The response of auditory cortex to the perturbation during active phonation (speaking) is enhanced (*right column*). (Adapted, with permission, from Houde and Chang, 2015.)

B. 1. Vocalization-induced suppression of neural activity in marmoset monkey auditory cortex. Population-averaged firing rate of all vocalization-suppressed responses are aligned by vocal onset (a "Phee" call). The **blue line** is a moving average (100 ms window) and shows that suppression begins prior to vocalization (indicated by **arrow**). The **thick bar** indicates the period over which suppression is continuously significant (P <0.05). (Adapted, with permission, from Eliades and Wang 2003.)

2. Neurons subject to vocalization-induced suppression are sensitive to vocal feedback perturbations. **Top:** Self-produced vocalizations with or without feedback alterations were delivered to the marmoset through a customized headphone. **Bottom:** This auditory cortical neuron was suppressed during normal vocalization (**dark blue**) but showed a large increase in firing rate when the auditory feedback of the vocalization was shifted in the frequency domain (**light blue**). Amplifying auditory feedback alone did not generate firing rate changes (**black**). (Adapted, with permission, from Eliades and Wang 2008; Crapse and Sommer 2008.)



Single-neuron recordings from the auditory cortex of vocalizing monkeys have shown that self-initiated vocalizations result in suppression of cortical responses to monkeys' own vocalizations, of external sounds heard during vocalization, and also spontaneous activity (Figure 28–15B). Because in many instances firing rates are suppressed to below spontaneous activity, the suppression is likely caused by inhibition. Neurons suppressed by self-initiated vocalizations show frequency and intensity tuning, as is typical of auditory cortical neurons, and respond to the playback of vocalizations.

The vocalization-induced suppression begins several hundred milliseconds prior to the onset of vocalization (Figure 28–15B), suggesting that these neurons receive modulatory signals originating in vocal production circuits. In humans, vocal production is carried out by cortical areas in the frontal lobe, from Broca's area to premotor and motor cortex. In humans and monkeys, axons from the premotor cortex to auditory regions of the superior temporal gyrus have been described, and presumably, they mediate the vocalization-induced suppression. This modulatory connection is not active when humans or monkeys simply listen to vocal sounds played to them.

Why do we suppress our auditory cortex when we speak? A simple answer is that this suppression helps reduce the masking effect of our own voice, which can be very loud. A more interesting answer is that this suppression results from a vocal feedback-monitoring network in auditory cortex. In humans, there is less or no suppression of auditory cortex if vocal feedback is experimentally altered through earphones, for example, when the pitch of the voice is shifted (Figure 28–15A). In marmoset monkeys, neurons suppressed by self-initiated vocalizations may become less suppressed or even excited when the animal hears its own frequency-shifted vocalizations (Figure 28–15C). This sensitivity to feedback perturbations suggests that neurons exhibiting vocalization-induced suppression are part of a network responsible for monitoring vocal feedback signals. The presence of vocal feedback-related neural activity in the auditory cortex of both humans and monkeys suggests that the auditory cortex combines both internal modulation and vocal feedback responses, rather than merely responding to sensory signals coming through the ears.

Not all neurons in the auditory cortex are suppressed by speaking or vocalizing. A smaller proportion (~30%) of neurons in marmoset A1 increase their responses during self-initiated vocalizations, consistent with their auditory response characteristics. In contrast to vocalization-induced suppression, vocalization-related excitation begins after the onset of vocalization and is likely the result of feedback through the ascending auditory pathway. The vocalization-related excitation may help maintain the sensitivity of the auditory cortex to the external acoustic environment during speaking or vocalizing.

Vocalization-induced suppression of auditory responses has been observed in several mammalian subcortical structures, including the brain stem and inferior colliculus. Such suppression begins a few milliseconds before or is synchronized with vocal production. In contrast, cortical suppression begins several hundred milliseconds before the vocal onset. It is possible that subcortical suppression of auditory responses during speaking or vocalizing is initiated by cortical commands.

Highlights

1. Sound impinging on two ears carries information that the brain uses to compute where sounds arise and what they mean. Sounds are characterized by the amount of energy at one or more frequencies. To determine where sounds arise in the horizontal plane, many mammals compute differences in the time of arrival at the two ears for sounds less than approximately 3,000 Hz. To determine where sounds arise in the vertical dimension and whether they arise from the front or the back, mammals use spectral filtering of sounds greater than approximately 6,000 Hz by the head, shoulders, and external ears.

2. Acoustic information is brought to the brain from the cochlea by auditory nerve fibers, each sharply tuned to a narrow range of frequencies and together representing the entire hearing range of the animal. Auditory nerve fibers terminate in the ventral and dorsal cochlear nuclei, distributing acoustic information to four major groups of principal cells that form parallel ascending pathways through the brain stem. The topographic organization of the auditory nerve inputs imparts a tonotopic organization to the ipsilateral cochlear nuclei that is preserved all along the auditory pathway, including auditory cortex.

3. A marked feature of auditory neurons at processing stations along the ascending pathway is their progressively increasing stimulus selectivity.

4. The ventral cochlear nucleus extracts three features of sounds: (a) The monaural pathways through octopus cells of the ventral cochlear nucleus, the superior paraolivary nucleus, and ventral nucleus of the lateral lemniscus detect coincident firing of auditory nerve fibers that is useful for detecting onsets and gaps in sounds. (b) Stellate cells detect and sharpen the encoding of spectral peaks and valleys and convey that spectral information to the dorsal cochlear nucleus, olivocochlear neurons in the ventral nucleus of the lateral lemniscus, ventral nucleus of the lateral lemniscus, inferior colliculus, and thalamus. Spectral information is used for understanding the meaning of sounds and for localizing their sources. (c) Bushy cells sharpen and convey information about the fine structure of sounds, which is used in the binaural pathways through the medial and lateral superior olivary nuclei to make the interaural comparisons of timing and intensity of sounds at the two ears, which are used to localize sound sources along the azimuth.

5. The dorsal cochlear nucleus integrates acoustic signals with somatosensory information in its principal cells. Somatosensory information helps distinguish the spectral cues generated by an animal's own movements, which are biologically uninteresting, from those that arise from the environment.

6. Auditory brainstem pathways converge in the inferior colliculus. The inferior colliculus feeds acoustic information through the medial geniculate body of the thalamus to auditory cortex.

7. A projection from the inferior colliculus carries information about the location of sounds to the superior colliculus, a part of the brain that controls reflexive orienting movements of the head and eyes.

8. Within auditory cortex, auditory neurons continue to become more selective to the stimuli to which they respond. Subregions of the auditory cortex represent different biologically significant features such as pitch of tones that form harmonic complexes. Auditory cortex also transforms rapidly varying features of sounds into firing-rate-based representations, while representing slowly varying sounds using spike timing.

9. Auditory circuits in the cerebral cortex are segregated into separate processing streams, with dorsal and ventral streams concerned respectively with sound location in space and sound identification.

10. The cerebral cortex modulates processing in subcortical auditory areas. Projections from the auditory cortex innervate the thalamus, inferior colliculus, olivocochlear neurons, some basal ganglionic structures, and even the dorsal cochlear nucleus.

11. Auditory cortex is involved in processing vocal feedback signals during speaking. Speaking induces suppression of neural activity in auditory cortex that begins several hundred milliseconds prior to the vocal onset. This suppression results from a vocal feedback-monitoring network that functions to guide vocal production and learning.

Donata Oertel

Xiaoqin Wang

Selected Reading

Bendor DA, Wang X. 2005. The neuronal representation of pitch in primate auditory cortex. Nature 436:1161–1165. [PubMed: 16121182]

Chase SM, Young ED. 2006. Spike-timing codes enhance the representation of multiple simultaneous sound-localization cues in the inferior colliculus. J Neurosci 26:3889–3898. [PubMed: 16611804]

Eliades SJ, Wang X. 2008. Neural substrates of vocalization feedback monitoring in primate auditory cortex. Nature 453:1102–1106. [PubMed: 18454135]

Gao E, Suga N. 2000. Experience-dependent plasticity in the auditory cortex and the inferior colliculus of bats: role of the corticofugal system. Proc Natl Acad Sci U S A 97:8081–8086. [PubMed: 10884432]

Hofman PM, Van Riswick JG, Van Opstal AJ. 1998. Relearning sound localization with new ears. Nat Neurosci 1:417–421. [PubMed: 10196533]

Joris PX, Smith PH, Yin TC. 1998. Coincidence detection in the auditory system: 50 years after Jeffress. Neuron 21:1235–1238. [PubMed: 9883717]

Joris PX, Yin TCT. 2007. A matter of time: internal delays in binaural processing. Trends Neurosci 30:70–78. [PubMed: 17188761]

Oertel D, Young ED. 2004. What's a cerebellar circuit doing in the auditory system? Trends Neurosci 27:104–110. [PubMed: 15102490]

Schneider DM, Mooney R. 2018. How movement modulates hearing. Annu Rev Neurosci 41:553–572. [PubMed: 29986164]

Schreiner CE, Read HL, Sutter ML. 2000. Modular organization of frequency integration in primary auditory cortex. Annu Rev Neurosci 23:501–529. [PubMed: 10845073]

Suga N. 1990. Cortical computational maps for auditory imaging. Neural Netw 3:3–21.

Wang X. 2018. Cortical coding of auditory features. Annu Rev Neurosci 41:527–552. [PubMed: 29986161]

Zhang LI, Bao S, Merzenich MM. 2001. Persistent and specific influences of early acoustic environments on primary auditory cortex. Nat Neurosci

4:1123–1130. [PubMed: 11687817]

References

Bartlett EL, Wang X. 2007. Neural representations of temporally-modulated signals in the auditory thalamus of awake primates. J Neurophysiol 97:1005–1017. [PubMed: 17050830]

Bendor DA, Wang X. 2006. Cortical representations of pitch in monkeys and humans. Curr Opin Neurobiol 16:391–399. [PubMed: 16842992]

Brodal A. 1981. *Neurological Anatomy in Relation to Clinical Medicine*. New York: Oxford Univ. Press.

Cajal SR. 1909. *Histologie du Systeme Nerveux de l'Homme et des Vertebres.* Paris: A. Maloine.

Cariani PA, Delgutte B. 1996. Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. J Neurophysiol 76:1698–1716. [PubMed: 8890286]

Cohen YE, Knudsen EI. 1999. Maps versus clusters: different representations of auditory space in the midbrain and forebrain. Trends Neurosci 22:128–135. [PubMed: 10199638]

Crapse TB, Sommer MA. 2008. Corollary discharge circuits in the primate brain. Curr Opin Neurobiol 18:552–557. [PubMed: 18848626]

Darrow KN, Maison SF, Liberman MC. 2006. Cochlear efferent feedback balances interaural sensitivity. Nat Neurosci 9:1474–1476. [PubMed: 17115038]

Eliades SJ, Wang X. 2003. Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. J Neurophysiol 89:2194–2207. [PubMed: 12612021]

Feng L, Wang X. 2017. Harmonic template neurons in primate auditory cortex underlying complex sound processing. Proc Natl Acad Sci U S A 114:E840–E848. [PubMed: 28096341]

Gao L, Kostlan K, Wang Y, Wang X. 2016. Distinct subthreshold mechanisms underlying rate-coding principles in primate auditory cortex. Neuron 91:905–919. [PubMed: 27478016]

Geisler CD. 1998. *From Sound to Synapse, Physiology of the Mammalian Ear.* New York: Oxford Univ. Press.

Houde JF, Chang EF 2015. The cortical computations underlying feedback control in vocal production. Curr Opin Neurobiol 33:174–181. [PubMed: 25989242]

Hubel DH, Henson CO, Rupert A, Galambos R. 1959. Attention units in the auditory cortex. Science 129:1279–1280. [PubMed: 13658956]

Jeffress LA. 1948. A place theory of sound localization. J Comp Physiol Psychol 41:35–39. [PubMed: 18904764]

Kanold PO, Young ED. 2001. Proprioceptive information from the pinna provides somatosensory input to cat dorsal cochlear nucleus. J Neurosci 21:7848–7858. [PubMed: 11567076]

King AJ. 1999. Sensory experience and the formation of a computational map of auditory space in the brain. BioEssays 21:900–911. [PubMed: 10517863]

King AJ, Bajo VM, Bizley JK, et al. 2007. Physiological and behavioral studies of spatial coding in the auditory cortex. Hear Res 229:106–115. [PubMed: 17314017]

Liberman MC. 1978. Auditory-nerve response from cats raised in a low-noise chamber. J Acoust Soc Am 63:442–455. [PubMed: 670542]

Lu T, Liang L, Wang X. 2001. Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. Nature Neurosci 4:1131–1138. [PubMed: 11593234]

Merzenich MM, Knight PL, Roth GL. 1975. Representation of cochlea within primary auditory cortex in the cat. J Neurophysiol 38:231–249. [PubMed: 1092814]

Mesgarani N, Cheung C, Johnson K, Chang EF. 2014. Phonetic feature encoding in human superior temporal gyrus. Science 343:1006–1010. [PubMed: 24482117]

Middlebrooks JC. 2005. Auditory cortex cheers the overture and listens through the finale. Nature Neurosci 8:851–852. [PubMed: 16136671]

Musicant AD, Chan JCK, Hind JE. 1990. Direction-dependent spectral properties of cat external ear: new data and cross-species comparisons. J Acoust Soc Am 87:757–781. [PubMed: 2307774]

Oertel D, Bal R, Gardner SM, Smith PH, Joris PX. 2000. Detection of synchrony in the activity of auditory nerve fibers by octopus cells of the mammalian cochlear nucleus. Proc Nat Acad Sci U S A 97:11773–11779.

Palmer AR, King AJ. 1982. The representation of auditory space in the mammalian superior colliculus. Nature 299:248–249. [PubMed: 7110344]

Penagos H, Melcher JR, Oxenham AJ. 2004. A neural representation of pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging. J Neurosci 24:6810–6815. [PubMed: 15282286]

Raman IM, Zhang S, Trussell LO. 1994. Pathway-specific variants of AMPA receptors and their contribution to neuronal signaling. J Neurosci 14:4998–5010. [PubMed: 7913958]

Rauschecker JP, Tian B. 2000. Mechanisms and streams for processing of “what” and “where” in auditory cortex. Proc Nat Acad Sci U S A 97:11800–11806.

Rauschecker JP, Tian B, Hauser M. 1995. Processing of complex sounds in the macaque nonprimary auditory cortex. Science 268:111–114. [PubMed: 7701330]

Recanzone GH, Schreiner CE, Merzenich MM. 1993. Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. J Neurosci 13:87–103. [PubMed: 8423485]

Remington ED, Wang X. 2019. Neural representations of the full spatial field in auditory cortex of awake marmoset (*Callithrix jacchus*). Cereb Cortex 29:1199–1216. [PubMed: 29420692]

Riquimaroux H, Gaioni SJ, Suga N. 1991. Cortical computational maps control auditory perception. Science 251: 565–568. [PubMed: 1990432]

Romanski LM, Averbeck BB. 2009. The primate cortical auditory system and neural representation of conspecific vocalizations. Annu Rev Neurosci 32:315–346. [PubMed: 19400713]

Sadagopan S, Wang X. 2009. Nonlinear spectrotemporal interactions underlying selectivity for complex sounds in auditory cortex. J Neurosci 29:11192–11202. [PubMed: 19741126]

Schreiner CE, Winer JA. 2007. Auditory cortex mapmaking: principles, projections, and plasticity. Neuron 56:356–365. [PubMed: 17964251]

Scott LL, Mathews PJ, Golding NL. 2005. Posthearing developmental refinement of temporal processing in principal neurons of the medial superior

olive. J Neurosci 25:7887–7895. [PubMed: 16135745]

Song X, Osmanski MS, Guo Y, Wang X. 2016. Complex pitch perception mechanisms are shared by humans and a New World monkey. Proc Natl Acad Sci U S A 113:781–786. [PubMed: 26712015]

Spirou GA, Young ED. 1991. Organization of dorsal cochlear nucleus type IV unit response maps and their relationship to activation by bandlimited noise. J Neurophysiol 66:1750–1768. [PubMed: 1765805]

Suga N, O'Neill WE, Kujirai K, Manabe T. 1983. Specificity of combination-sensitive neurons for processing of complex biosonar signals in auditory cortex of the mustached bat. J Neurophysiol 49:1573–626. [PubMed: 6875639]

Suga N. 1984. Neural mechanisms of complex-sound processing for echolocation. Trends Neurosci 7:20–27.

Tollin DJ, Yin TC. 2002. The coding of spatial location by single units in the lateral superior olive of the cat. II. The determinants of spatial receptive fields in azimuth. J Neurosci 22:1468–1479. [PubMed: 11850473]

Wang X, Lu T, Snider RK, Liang L. 2005. Sustained firing in auditory cortex evoked by preferred stimuli. Nature 435:341–346. [PubMed: 15902257]

Warr WB. 1992. Organization of olivocochlear efferent systems in mammals. In: DB Webster, AN Popper, RR Fay (eds). *The Mammalian Auditory Pathway: Neuroanatomy,* pp. 410–448. New York: Springer.

Winer JA, Saint Marie RL, Larue DT, Oliver DL. 1996. GABAergic feedforward projections from the inferior colliculus to the medial geniculate body. Proc Natl Acad Sci U S A 93:8005–8010. [PubMed: 8755593]

Yin TCT. 2002. Neural mechanisms of encoding binaural localization cues in the auditory brainstem. In: D Oertel, RR Fay, AN Popper (eds). *Integrative Functions in the Mammalian Auditory Pathway,* pp. 238–288. New York: Springer.