A

Auditory Processing and Perception

Koedi S. Lawley[1] and Jonathan F. Prather[1,2]
[1]Neuroscience Program, Department of Zoology and Physiology, University of Wyoming, Laramie, WY, USA
[2]University of Wyoming, Laramie, WY, USA

Animals use a variety of complex signaling mechanisms to convey information to conspecifics and heterospecifics alike. Among those modalities, acoustic signaling is a widespread behavior across vertebrate groups. Acoustic signals can carry information over long distances, and precise changes in frequency and temporal sequencing can encode high-resolution information in a very efficient way. These signals generally contain information about traits of the individual producing them, and they can divulge everything from the individual's location in space, to its identity, physiological state, and even complex semantic content as in the case of human speech. Together, these features make acoustic signals and the associated auditory processing and perception an excellent means of communication.

Behavioral Value of Hearing and Acoustic Signaling

Acoustic signals move at great speeds through the environment and can be detected at long distances from their source. Therefore, auditory perception affords an excellent means of detecting and localizing other individuals in the environment. That ability can confer an advantage in noticing the movements of possible prey or detecting the approach of a nearby predator. Even beyond those unintended signals associated with movement, acoustic signals can also be produced intentionally in the form of vocalizations such as alarm calls. Because sound radiates from its source in all directions, these calls can inform the behavioral responses of many individuals, including members of the individual's own species and eavesdropping members of other species. For example, Madagascar spiny-tailed iguanas (*Opluris cuvieri cuvieri*) eavesdrop on the alarm calls that Madagascar paradise flycatchers (*Terpsiphone mutata*) produce in response to detecting a nearby predator (Ito and Mori 2010). Those two species share a mutual threat, and iguanas become hyper-vigilant in response to flycatcher calls (Ito and Mori 2010). Thus, by eavesdropping on the acoustic signals produced

by species that are similarly wary of specific nearby predators, it is thought that this helps iguanas evade predation.

Another key aspect of acoustic signaling and auditory processing is the recognition of individual identity. Identifying individuals is critical for distinguishing a mate, kin, or a threat from others within a group, and recognition of individual participants is necessary for almost all social behaviors. For example, wolves (*Canis lupus*) are capable of recognizing individuals based on the acoustic characteristics of their howls (Palacios et al. 2015). Similarly, female northern fur seals (*Callorhinus ursinus*) learn the calls produced by their pups, and a female will approach speakers that play the calls of their pups but ignore speakers playing the calls of other pups (Insley 2000). The elements of the nervous system that underlie that form of auditory processing incorporate not only sensory areas but also areas involved in memory, as a female seal's ability to use auditory processing to recognize her offspring persists even after several years of separation (Insley 2000).

Along with revealing an individual's identity, acoustic signals may also serve as an indicator of the individual's physiological state. For example, the vocalizations of rock hyraxes (*Procavia capensis*) contain cues about the physiological state and social rank of the individual (Koren and Geffen 2009). These examples are just a few of the almost innumerable cases in which acoustic signals and auditory perception play fundamental roles in the lives of many different types of organisms.

Hearing in Birds and Mammals: Structures and Mechanics

Upon first inspection, a striking difference between mammals and birds is the presence of pinna in mammals, such as the prominent ears on rabbits and many dogs, but the absence of any such structure in birds. Those external structures in mammals are thought to confer an advantage because they help to funnel sound into the ear canal, making even very faint sources of sound easier to detect (Yost 2006). Beyond that external

difference, however, the physical structures involved in detecting and processing sound are broadly similar between mammals and birds. For example, both groups have ear canals that are open to the surrounding air and that are diametrically opposed on the sides of their heads. This arrangement enables organisms to compare the timing with which a signal hits one ear first and then the other ear some brief time later. This interaural comparison is an integral component of how animals localize the source of that sound, reminiscent of how widely spaced eyes facilitate perception of visual depth. Despite the absence of pinna, some birds use the shape of their heads and specialized feathers to help funnel sounds into the ear canal in a way that is especially beneficial in detecting the location of the sound source (Wagner et al. 2013). Thus, similarities emerge even in the case of the most obvious anatomical difference between mammals and birds if we investigate the system more carefully.

When sounds are produced in the environment and detected by the auditory system, acoustic energy propagating as vibrations of molecules in the air enters the outer ear and causes movement of the structures that transduce that energy into perception of sound (Yost 2006). Through the action of pinna or specialized feathers, acoustic energy is funneled into the canal of the outer ear, where it is eventually captured and transduced into neural activity by the remaining structures of the middle and inner ear.

In the middle ear, membranes and small bones serve as a conduit for the transmission of sound from movement of air into movement of the fluid in the inner ear (Yost 2006). Through those intricate mechanisms, the movement of air molecules that is inherent to acoustic energy is converted into motion of fluid in downstream structures where those motions are eventually transduced into activity of auditory neurons (Yost 2006). This transduction of acoustic energy and the associated physical motion into activity of neurons in the inner ear is broadly conserved across vertebrate groups (Manley 2017), but the structures and mechanisms that are employed to achieve sound transduction vary.

In both mammals and birds, acoustic energy in the canal of the outer ear impacts a thin membrane called the tympanic membrane (commonly called the "ear-drum"), and the tympanic membrane induces motion of a set of tiny bones that are collectively called the ossicles (Tucker 2017). These small ossicles within the middle ear are connected to form a chain for transmission of movement from the tympanic membrane to the opening to the inner ear known as the oval window (Yost 2006). At this level of resolution, additional differences begin to emerge between this system in mammals and the set of corresponding structures in birds, particularly in the number and relative orientation of the ossicles. For example, the ossicles consist of three bones – the malleus (hammer), incus (anvil), and stapes (stirrup) – in mammals and one bone – the columella – in birds (Tucker 2017; Yost 2006). Despite those structural differences, each of those systems achieves the same function of transferring movement and pressure from the outer ear, through the middle ear, and into the cochlea in the inner ear. For both groups, that pressure results in movement of the fluid and membranes within the inner ear. One of those membranes, the basilar membrane of the cochlea, contains sensory hair cells that are very sensitive to movement of the surrounding fluid. When those movements occur, they induce changes in the voltage of the hair cells, and those changes in voltage can be detected as electrical impulses, called action potentials, in neurons in the auditory nerve (Yost 2006). In both birds and mammals, these hair cells are the first step in routing auditory information into the central nervous system.

This comparison between mammals and birds makes it clear that nature has solved the challenge of auditory processing in several different ways. Behavioral studies performed using those groups and others have revealed that auditory information serves similar purposes across many species (e.g., localizing sound sources, permitting communication and social interactions). Many of the studies that have revealed what we know about these processes have been performed in birds, and for that reason many of the examples described in the remainder of this text focus on lessons learned from studies of birds.

Auditory Processing of Sound Location

The ability to resolve the location of a sound source is quite beneficial for detecting features of the ambient environment. This process is driven by comparing the timing of sounds arriving at each ear, with the sound produced by a specific source arriving first at the ear that is closer to the source and later at the ear that is farther from the source (Fig. 1). Mammals and birds have auditory systems that are excellent at localizing sounds by computing with millisecond accuracy the differences in time that it takes sound to reach each ear (Ashida and Carr 2011). The seminal studies that revealed this neural mapping of auditory space were performed using barn owls (*Tyto furcata pratimcola*) (Knudsen and Konishi 1978). Those owls are nocturnal predators, and they are exceptionally good at locating prey using only acoustic signals even in complete darkness. In order to locate their prey, such as a scurrying mouse, they rely on hearing and specifically on a mechanism called "coincidence detection" (Fig. 1). In this mechanism, action potentials associated with sounds arriving at the left and right ears are relayed to a specialized set of neurons in the brainstem (Wagner et al. 2013). Each of these specialized neurons receives input from each ear. Activity from either ear alone is not sufficient to activate those specialized brainstem neurons, but activity from each ear simultaneously is sufficient to activate those cells (Wagner et al. 2013). This requirement of activity from multiple inputs at the same time is the origin of the name "coincidence detectors."

Through an elegant means of comparing the latency of activity from the left ear versus the latency of activity from the right ear (Fig. 1), each neuron in the brainstem population becomes an indicator of a discrete interaural time difference. Activity from each ear propagates centrally from the ear to these brainstem neurons, and in nearly all cases, the activity from one ear affects a given brainstem neuron at a different time than activity from the other ear. But at one point in that convergence from each ear onto these brainstem neurons, activity from each ear reaches a given cell at the same time (open circles in Fig. 1b). In
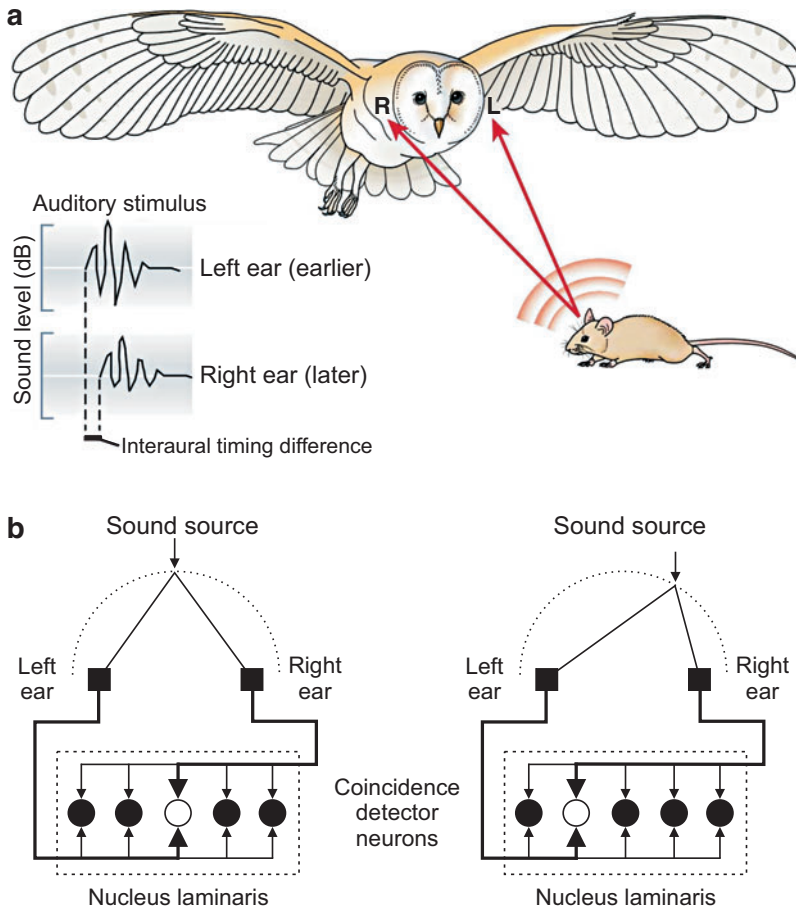
Auditory Processing and Perception, Fig. 1 The location of a sound source in the surrounding environment gives rise to activity at a specific location within a specialized neural network. (**a**) Acoustic energy generated by the mouse's movements or vocalizations reaches the owl's left (L) and right (R) ears at different times and with different intensities depending on the mouse's location in space. In this case, the left ear receives an earlier and more intense stimulus compared to the later and less intense signal received by the right ear. The difference in time of arrival between each ear is the interaural timing difference.

(Adapted from Fig. 1 in Knudsen 2002). (**b**) Coincidence detector neurons in the auditory brainstem (nucleus laminaris) receive input from each ear with different latency. Input from either ear alone is insufficient to activate those cells (filled circles), but those neurons respond robustly (open circles) when input from both ears arrives at the same time. Thus, each coincidence detector neuron becomes an indicator of the position of a sound source in the surrounding environment. (Adapted from Fig. 1 in Salomon et al. 2012)

that case, coincident input has occurred and that cell is activated (Fig. 1). Neurons that perform this comparison of arrival times in each ear reside in the nucleus laminaris (NL) in the brainstem of birds and in the homologous medial superior olive (MSO) in the brainstem of mammals (Ashida and Carr 2011). Thus, the activity of a given coincidence detector neuron indicates the presence of a sound source at the corresponding point in the surrounding space, enabling the organism to identify the position of the sound source in the horizontal plane. Barn owls have the additional impressive trait of asymmetric ear openings, with one ear biased toward detecting sounds that come from sources below the owl's head and the other ear biased toward detecting sounds from sources above the bird's head (Wagner et al. 2013). In this way, barn owls are

able to detect the position of a sound source along both the horizontal and vertical axes, revealing the position of that source in high resolution and all as a result of auditory input even in the absence of visual input.

Auditory Processing of Social Information

In addition to serving as a beacon of the location of a sound source, many acoustic signals such as speech and songs performed by birds also convey important social information. For example, song plays many key roles in songbird social behavior, including revealing the singer's identity, helping the singer establish and maintain a territory, and playing a central role in courtship and mate choice (Catchpole and Slater 2008). In most of the species that have been studied, especially those that reside in North America, male birds sing but females do not (Catchpole and Slater 2008). That ability of male birds to sing is associated with a network of specialized brain sites that are large and obvious in the male brain but are atrophied to much smaller sizes or are absent in the female brain (Mooney et al. 2008). Intriguingly, male songbirds learn their songs in a developmental progression that is strikingly parallel to the way that humans learn the sounds that we use in speech (Doupe and Kuhl 1999). In that process, the learner imitates the sounds that he hears performed by other members of his species and relies on auditory feedback to refine his imitations of those vocalizations, and many investigations have focused on determining how and where those complex auditory signals are processed in the songbird brain.

Electrophysiological recordings and pathway tracing studies from many research groups have revealed that activity originating in hair cells of the inner ear is propagated through a network of auditory neurons in the brainstem and thalamus and eventually reaches a forebrain area called Field L, which is the avian analog of the mammalian primary auditory cortex (Fig. 2) (reviewed in Prather 2013). From Field L, activity is passed along to secondary cortical areas including the

caudal mesopallium (CM) and the caudomedial nidopallium (NCM) (Fig. 2) (Vates et al. 1996). As we ascend through that circuit, neurons become more selectively responsive to specific characteristics of those behaviorally relevant sounds (i.e., songs and calls). For example, Field L is broadly responsive to both song and a wide variety of other sounds including pure tones and white noise (Lewicki and Arthur 1996). In contrast, neurons in CM and NCM are more selectively responsive to specific aspects of auditory stimuli such as the novel or familiar status of the stimulus (Gentner and Margoliash 2003; Mello et al. 1995) and the identity of the associated individual that produced that vocalization (Menardy et al. 2012). Furthermore, immediate early gene expression within NCM is also positively correlated with the strength of song learning (Terpstra et al. 2004), suggesting that these forebrain areas may contribute to not only auditory processing but also formation and recall of memories regarding these behaviorally relevant auditory experiences.

Additional studies have also revealed highly selective auditory sites that are interconnected with CM and NCM. One of those sites is a forebrain nucleus called HVC (abbreviation used as a proper name, Fig. 2), and the auditory responses of HVC neurons are among the most selective sensory responses ever described. For example, HVC neurons respond to not only song but also more specifically to the bird's own song as opposed to songs or calls produced by members of the bird's own species (reviewed in Mooney et al. 2008). Furthermore, HVC neurons can even represent different songs in the bird's repertoire (Prather et al. 2008), revealing that those neurons are not simply representing classes of vocalizations (e.g., self-generated songs vs. songs performed by others). Instead, they are representing specific sounds used in vocal communication. In support of that idea, HVC neurons that respond to individual songs in the bird's own repertoire are responsive to specific elements of those songs (e.g., one specific transition between two notes), and those cells respond to playback of that element regardless of whether it was performed by the bird where that HVC cell
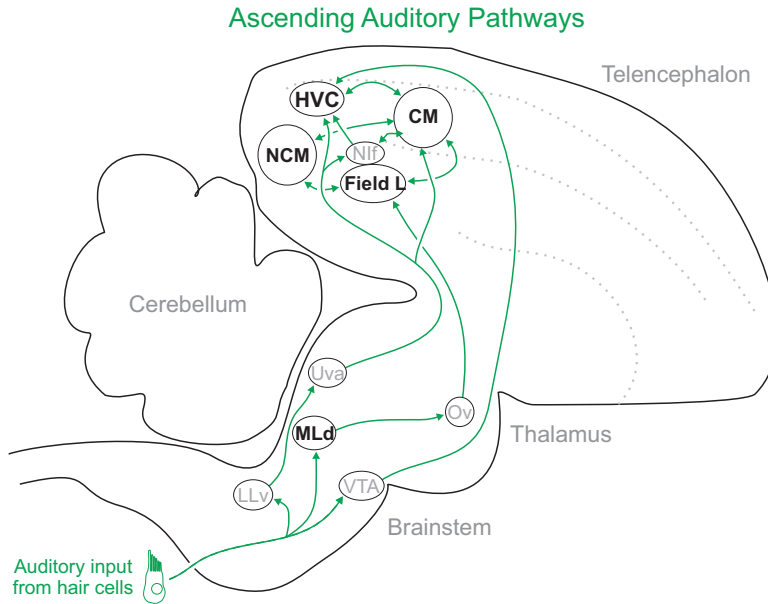
Auditory Processing and Perception, Fig. 2 Auditory input is processed through an ascending neural network. In this depiction of a parasagittal section of the songbird brain, auditory input is transduced by hair cells in the cochlea and passed through the brainstem, thalamus, and eventually to auditory processing centers in the telencephalon. Areas highlighted in this text are in black font, and additional areas are shown in gray. Not all connections are shown. Arrows indicate the directionality of each connection, and gray dotted lines indicate lamina that serve as landmarks that are helpful in identifying locations in the brain. Some structures, such as HVC, are sexually dimorphic (large in males and smaller or absent in females). Other structures, such as MLd, Field L, CM, and NCM, are robustly present in both sexes. Abbreviations are: *CM* caudal mesopallium, *Field L* auditory thalamorecipient neurons in the forebrain, *HVC* abbreviation used as a proper name, sometimes referred to as the high vocal center, *LLv* lateral lemniscus, ventral nucleus; *MLd* dorsal lateral nucleus of the mesencephalon, *NCM* caudal medial nidopallium, *NIf* nucleus interfacialis of the nidopallium, *Ov* ovoidalis, *Uva* nucleus uvaeformis, *VTA* ventral tegmental area. (Adapted from Fig. 2 in Prather 2013)

resides or by another member of its species (Prather et al. 2008). Beyond simply responding to acoustic parameters of the song, an integrated approach combining behavioral observation and electrophysiological recording of the activity of individual HVC neurons in awake and freely behaving birds also revealed that some HVC neurons respond to not simply the acoustic properties of auditory stimuli, but rather the animal's perception of the identity of those stimuli (Prather et al. 2009) and the syntax in which those sounds are produced (Fujimoto et al. 2011). Thus, neural activity in the songbird brain is correlated with many specific aspects of auditory processing, affording researchers many opportunities to study the neural mechanisms that underlie auditory processing and perception.

Summary

The studies described here highlight the value of birds as a behaviorally rich and experimentally tractable context in which to continue to investigate how the brain enables auditory processing and perception. Regions of the mammalian cortex are also specialized for auditory processing, yet our understanding remains incomplete regarding how the structure of individual neurons and the circuits they compose results in auditory perception. Cortical areas that are specialized for perception and performance of the sounds we use in speech have been identified in the human brain, yet our understanding of the functional role of those sites also remains incomplete. For example, Wernicke's area in the human brain has been implicated in our perception of the sounds that

we use to communicate using language, and another region of the human brain called Broca's area has been implicated in the production of those sounds. Recent research has blurred the distinction between those brain sites, suggesting that both areas contribute to both performance and perception of the sounds that compose one of the primary ways in which we use auditory processing in our everyday lives (Flinker et al. 2015; Tremblay and Dick 2016). The obvious behavioral significance of those human brain areas highlights them as important areas for additional study, but the function of individual neurons and the circuits they compose remains challenging to detect and interpret in humans. By incorporating a blend of behavioral and neurobiological approaches, animal studies such as those performed using songbirds afford an opportunity to understand the relation between activity in specific neurons and the emergence of an organism's perception of auditory experiences.

Cross-References

▶ Auditory Signals
▶ Aves (Birds)
▶ Cognition
▶ Ethology
▶ Nervous System, The
▶ Owls
▶ Songbird Learning

References

Ashida, G., & Carr, C. E. (2011). Sound localization: Jeffress and beyond. *Current Opinion in Neurobiology, 21*, 745–751.

Catchpole, C. K., & Slater, P. J. B. (2008). *Birdsong: Biological themes and variations*. Cambridge: Cambridge University Press.

Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: Common themes and mechanisms. *Annual Review of Neuroscience, 22*, 567–631.

Flinker, A., Korzeniewska, A., Shestyuk, A. Y., Franaszczuk, P. J., Dronkers, N. F., Knight, R. T., & Crone, N. E. (2015). Redefining the role of Broca's area in speech. *Proceedings of the National Academy of Sciences of the United States of America, 112*, 2871–2875.

Fujimoto, H., Hasegawa, T., & Watanabe, D. (2011). Neural coding of syntactic structure in learned vocalizations in the songbird. *The Journal of Neuroscience, 31*, 10023–10033.

Gentner, T. Q., & Margoliash, D. (2003). Neuronal populations and single cells representing learned auditory objects. *Nature, 424*, 669–674.

Insley, S. J. (2000). Long-term vocal recognition in the northern fur seal. *Nature, 406*, 404–405.

Ito, R., & Mori, A. (2010). Vigilance against predators induced by eavesdropping on heterospecific alarm calls in a non-vocal lizard *Oplurus cuvieri cuvieri* (Reptilia: Iguania). *Proceedings of the Biological Sciences, 277*, 1275–1280.

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

Knudsen, E. I., & Konishi, M. (1978). A neural map of auditory space in the owl. *Science, 200*, 795–797.

Koren, L., & Geffen, E. (2009). Complex call in male rock hyrax (*Procavia capensis*): A multi-information distributing channel. *Behavioral Ecology and Sociobiology, 63*, 581–590.

Lewicki, M. S., & Arthur, B. J. (1996). Hierarchical organization of auditory temporal context sensitivity. *The Journal of Neuroscience, 16*, 6987–6998.

Manley, G. A. (2017). Comparative auditory neuroscience: Understanding the evolution and function of ears. *Journal of the Association for Research in Otolaryngology, 18*, 1–24.

Mello, C., Nottebohm, F., & Clayton, D. (1995). Repeated exposure to one song leads to a rapid and persistent decline in an immediate-early genes response to that song in zebra finch telencephalon. *The Journal of Neuroscience, 15*, 6919–6925.

Menardy, F., Touiki, K., Dutrieux, G., Bozon, B., Vignal, C., Mathevon, N., & Del Negro, C. (2012). Social experience affects neuronal responses to male calls in adult female zebra finches. *The European Journal of Neuroscience, 35*, 1322–1336.

Mooney, R., Prather, J. F., & Roberts, T. (2008). Neurophysiology of birdsong learning. In H. Eichenbaum (Ed.), *Learning and memory: A comprehensive reference* (pp. 441–474). Oxford: Elsevier.

Palacios, V., Font, E., Marquez, R., & Carazo, P. (2015). Recognition of familiarity on the basis of howls: A playback experiment in a captive group of wolves. *Behaviour, 152*, 593–614.

Prather, J. F. (2013). Auditory signal processing in communication: Perception and performance of vocal sounds. *Hearing Research, 305*, 144–155.

Prather, J. F., Peters, S., Nowicki, S., & Mooney, R. (2008). Precise auditory-vocal mirroring in neurons for learned vocal communication. *Nature, 451*, 305–310.

Prather, J. F., Nowicki, S., Anderson, R. C., Peters, S., & Mooney, R. (2009). Neural correlates of categorical perception in learned vocal communication. *Nature Neuroscience, 12*, 221–228.

Salomon, R., Heinrich, E., & Joost, R. (2012). Modeling the nucleus laminaris of the barn owl: Achieving 20 ps resolution on a 85-MHz-clocked digital device. *Frontiers in Computational Neuroscience, 6*, 6.

Terpstra, N. J., Bolhuis, J. J., & den Boer-Visser, A. M. (2004). An analysis of the neural representation of birdsong memory. *The Journal of Neuroscience, 24*, 4971–4977.

Tremblay, P., & Dick, A. S. (2016). Broca and Wernicke are dead, or moving past the classic model of language neurobiology. *Brain and Language, 162*, 60–71.

Tucker, A. S. (2017). Major evolutionary transitions and innovations: The tympanic middle ear. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 372*, 20150483.

Vates, G. E., Broome, B. M., Mello, C. V., & Nottebohm, F. (1996). Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches. *The Journal of Comparative Neurology, 366*, 613–642.

Wagner, H., Kettler, L., Orlowski, J., & Tellers, P. (2013). Neuroethology of prey capture in the barn owl (*Tyto alba* L.) *Journal of Physiology, Paris, 107*, 51–61.

Yost, W. A. (2006). *Fundamentals of hearing: An introduction* (p. 338). Burlington: Academic.