Beyond Words: How Humans Communicate Through Sound

Nina Kraus[1,2,3,4] and Jessica Slater[1,2]

[1]Auditory Neuroscience Laboratory, Departments of [2]Communication Sciences,
[3]Neurobiology and Physiology, [4]Otolaryngology, Northwestern University, Evanston,
Illinois 60208; www.brainvolts.northwestern.edu; email: nkraus@northwestern.edu

Keywords

speech, music, rhythm, learning, neural plasticity, auditory processing

Abstract

Every day we communicate using complex linguistic and musical systems, yet
these modern systems are the product of a much more ancient relationship
with sound. When we speak, we communicate not only with the words we
choose, but also with the patterns of sound we create and the movements
that create them. From the natural rhythms of speech, to the precise timing
characteristics of a consonant, these patterns guide our daily communication.
By examining the principles of information processing that are common to
speech and music, we peel back the layers to reveal the biological foundations
of human communication through sound. Further, we consider how the
brain's response to sound is shaped by experience, such as musical expertise,
and implications for the treatment of communication disorders.

Contents

INTRODUCTION

The human relationship with sound is much deeper and more ancient than our relationship with words. The intoxicating sophistication and precision of modern languages can blind us to more fundamental aspects of auditory processing that underlie everyday communication. Yet a great deal of communicative power lies beneath the surface of words, and our modern languages and musical systems reflect how humans evolved in a world of sound (see **Figure 1**). Sound is created by physical movement, from the crunch of leaves underfoot, to the vibrations of vocal cords and
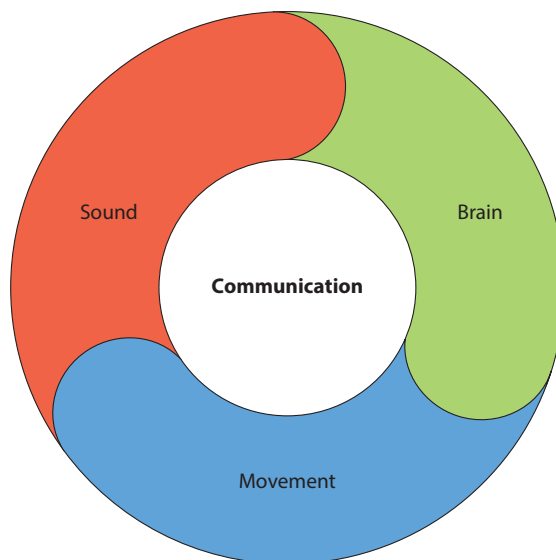
Figure 1

Our modern communication systems are the product of our ancient relationship with sound, rooted in the physical world.
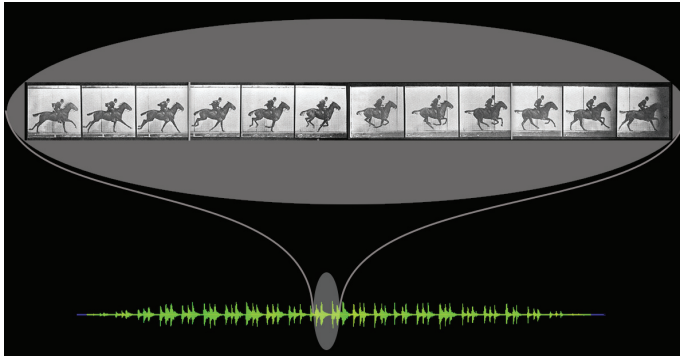
Figure 2

Sound provides an inventory of motion: The repeating movements of a galloping horse create repeating patterns in the resulting sound wave. Original images of a horse galloping by Eadweard Muybridge. Figure created by Adam Trefonides.

violin strings. The ability to make sense of sound helps us to construct accurate representations of our world, based on what we know about how sounds are produced. Just as a movie sound engineer will use sound effects to inform us about what is happening off camera (e.g., the wail of a siren or the door-slam of a departing lover), the sounds we hear as we navigate our daily lives create a dynamic map of objects around us and how they are moving in space. Sound extends our perceptual reach beyond the boundaries of our visual field. Though we can close our eyes or shift our gaze, our ears remain open, providing input from all directions. Auditory information therefore plays an important role in guiding our other senses, such as when we hear a sound and turn our head to look. Sound provides a kind of connective tissue, ensuring the continuity of our experience over time and choreographing the focus of our senses.

Sounds emanating from a single source share common characteristics and exhibit distinctive patterns that link them to their physical origin: The pitch and timbre of a sound are dependent on the resonant characteristics of an object, so a sequence of sounds with similar pitch and timbre implies a common source. If sounds are produced regularly in time, this may suggest an object in motion (see **Figure 2**). Leaves crunching with regularity and increasing volume may signal the approach of an animate leaf-crunching entity, such as a predator. The combination of loudness and rate of leaf crunching tells us about the size of the potential predator, how fast it is approaching, and, ultimately, which way to run.

Our brains are therefore performing constant computations to determine the underlying physicality of the sounds we hear, combining inputs from both ears and seeking out patterns that might inform us of approaching danger. When we communicate through sound, we are not only producing and perceiving acoustic signals; we are also exchanging detailed inventories of motion. Although the nuanced functions of modern language may seem far removed from a literal map of objects in space, knowing how a sound is produced can facilitate perception; for example, the mechanical properties of the human vocal apparatus place inherent constraints on the sound sequences likely to be generated in the course of a human utterance. Our motor planning and production systems possess implicit knowledge about which patterns of speech are most likely to be produced, based on the motor sequences required to produce them. Numerous studies have demonstrated that brain regions involved in speech production are activated when listening to speech (Watkins et al. 2003, Wilson et al. 2004), supporting the idea that "Speech is processed both as a sound and as an action" (Scott & Johnsrude 2003, p. 105). Given the close ties between speech and physical gesture, it is unsurprising that the brain integrates both sound and movement

to understand a communicative act. Further, processing speech as an action may help in situations where there is not a clear one-to-one mapping between sound and meaning. For example, a given phoneme can give rise to very different acoustic forms depending on the sounds that come before or after it and the characteristic articulation patterns of the individual talker. Putting the sound into an articulatory context may facilitate correct categorization of the phoneme despite acoustic ambiguity. Lack of physical context may also explain why the development of effective speech recognition technology has presented such a challenge. As the philosopher Wittgenstein (1953) stated, "If a lion could talk, we could not understand him," emphasizing that meaningful communication requires some degree of common experience between the communicators in order to understand the origin of the signal, how it was produced, and therefore what it means.

Our natural environments are complex, and there may be many inputs competing for our attention. Patterns help guide attention and streamline processing by grouping elements together into coherent objects. For example, timing cues play an important role in grouping sounds into an auditory object (Andreou et al. 2011, Shamma et al. 2011). Timing patterns must be integrated across multiple timescales, from the microsecond timing that helps us distinguish the crunch of leaves from the snap of a branch, to the slower rhythm of footsteps. Auditory information is also integrated with input from other modalities such as vision (Musacchia et al. 2008, Schutz & Lipscomb 2007). For example, the well-known McGurk effect is a demonstration of how conflicting visual information influences the perception of a sound: A video presenting repeated utterances of the syllable [ba] is dubbed on to the lip movements for [ga], resulting in normal-hearing adults perceiving a hybrid percept, the syllable [da] (McGurk & MacDonald 1976). Perception therefore involves not only identifying the presence of some kind of regularity or structure, but also integration of inputs across modalities and weighing between alternatives such that the "best fit" candidate wins out—in the case of the McGurk effect, the resulting percept is a compromise between conflicting alternatives. Both the immediate context of a sound and a perceiver's accumulated experience with sound over the course of a lifetime can shape how each new sound is interpreted.

In summary, our ancient relationship with sound is grounded in the physical world. The way that our brains evolved to make sense of sound is driven by how its physicality affects us and how our own physicality produces sound. As a result, our brains are constantly searching for patterns, particularly the kinds of patterns that tell us something about the physical world. The patterns our brains seek out determine how we group acoustic features together to form meaningful objects and streams, and these patterns are integral to the complex communication systems we use every day. Importantly, we are not disembodied listeners: Production and perception are entangled, and this entanglement is evident in the neural circuitry that supports our perception of sound (Kraus et al. 2015).

In the remaining sections of this review, we focus on how communication is guided by patterns, considering both statistical patterns and temporal patterns (see **Figure 3**). We then examine underlying biological mechanisms and how the interaction between production and perception is reflected in the brain. We pay particular attention to how these processing mechanisms relate to everyday communication skills such as reading and how their biological foundations are shaped by experience, including musical expertise.

PATTERNS AND PREDICTION

Statistical Learning

One way the human brain makes sense of incoming sounds is by keeping track of statistical patterns and making predictions based on those patterns. For example, exposure to English quickly reveals

Precise neural encoding
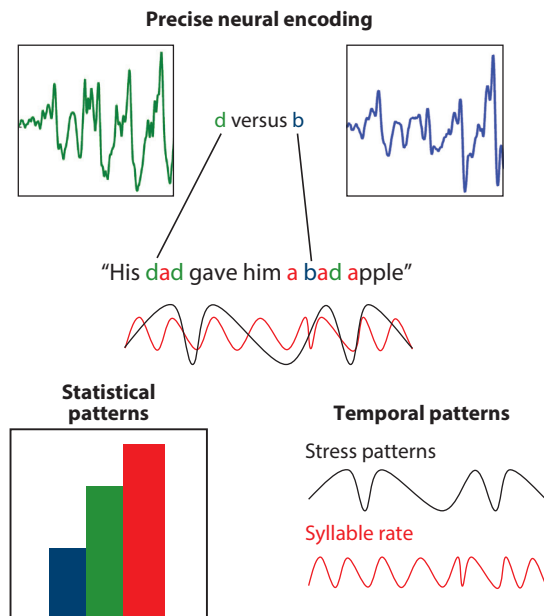
Figure 3

During speech perception, the brain is processing patterns across multiple timescales. Precise neural encoding of fine-timing features helps a listener differentiate one consonant from another; statistical characteristics and temporal patterns guide perception and help form predictions about what is coming next.

that the sound "th" is very likely to be followed by "a" but very unlikely to be followed by "p." By aggregating the probabilities of sound relationships over time, statistical learning enables a listener to discern meaningful structure in sound (Saffran et al. 1996). Predictions are constantly updated and corrected as further information comes in, and the timeframe over which statistical patterns are accrued can range from a short period of exposure (several minutes in an experimental setting) to a lifetime of experience, for example with a native language. Statistical learning has been observed even in infants and is considered a key mechanism underlying the natural acquisition of language (Saffran et al. 1996).

Sensitivity to environmental statistics is encoded throughout the auditory pathway and is not specific to humans: There is evidence that neurons adjust their firing rates in response to the sound-level statistics of their environment in individual auditory cortex neurons of rats (Ulanovsky et al. 2004), and from recordings in the auditory midbrain of guinea pigs (Dean et al. 2005, 2008). Statistics-based adaptation of firing rates has also been demonstrated in the rat thalamus (Antunes et al. 2010) and inferior colliculus (Malmierca et al. 2009), and even in the auditory nerve of anaesthetized cats (Wen et al. 2009). In the case of sound level, this statistical adaptation functions as a gain control, allowing the organisms to maximize their dynamic sensitivity within the parameters of the environment they are in. In a similar way, neurons can adjust their firing to the more complex statistical patterns of a language or musical system. Neonates less than 2 days old were able to pick up on statistical patterns in an artificial language to which they were exposed for one hour during sleep (Teinonen et al. 2009), reinforcing that humans seem to be innately wired with complex pattern detection mechanisms that do not even require active attention. An experiment assessing statistical learning abilities in cotton-top tamarins revealed that these nonhuman primates are also capable of probability-based pattern detection (Hauser

et al. 2001). However, a later study determined that human infants were better able to learn the complex grammatical structures that are characteristic of human language (Saffran et al. 2008), whereas the cotton-top tamarins were only able to learn simple grammars, suggesting that species may differ in the degree of sophistication of these statistical learning mechanisms.

Statistical learning is not limited to the auditory domain, with infants also demonstrating sensitivity to statistical patterns in sequences of visual stimuli (Kirkham et al. 2002). It is interesting to note that the basal ganglia, which play an important role in the generation of movement patterns such as walking, are also involved in pattern learning (Kotz et al. 2009), including the acquisition of linguistic and musical grammar (Conway & Pisoni 2008, Ettlinger et al. 2011, Ullman 2001). Whereas the dorsal pathway of the basal ganglia is associated with sensorimotor planning, the ventral pathway plays an important role in reward-based behavior. Although the extent of overlap between the various functions of the basal ganglia remains an active area of research, there is increasing evidence that neural circuitry previously associated with motor planning may in fact serve a broader function in pattern detection and prediction, for cognitive as well as motor action (Graybiel 2005, Schubotz 2007). The reward-based circuitry also introduces an important chemical ingredient for learning, in the form of the neurotransmitter dopamine. Further insight into probability-based learning is provided by examples in which an explicit reward is provided for a particular task: As the brain starts to build predictions based on experience, reward pathways are activated not only by a reward itself, but also in anticipation of likely reward (Knutson et al. 2001), with the degree of activation reflecting the statistical probabilities derived from prior experience (Morris et al. 2006). Dopamine therefore provides a chemical mechanism by which probability can be tracked within the brain and facilitate pattern-based prediction of future events.

Statistical learning mechanisms can themselves be shaped by experience. For example, there is evidence that musical expertise results in improved implicit learning abilities for both musical and linguistic structures (Ettlinger et al. 2011, François & Schön 2011, Skoe et al. 2013b), and musicians were better able to pick up statistical patterns in a novel Morse code–based language (Shook et al. 2013). Short-term musical training has also been shown to improve the ability to parse a speech stream into words (François et al. 2013). Receptivity to statistical learning seems to shift over the course of development; for example, while infants are able to differentiate metrical categories from another musical culture, adults only respond to those present in their own culture (Hannon & Trehub 2005). This suggests that as experience accumulates over a lifetime, the nervous system may maximize its efficiency in processing the most relevant distinctions at the cost of being able to adapt to new ones.

In summary, it has been suggested that the ability of neuronal populations to adapt their firing to statistical properties reflects a fundamental organizing principle of the nervous system, allowing an organism to function in dynamic environments (Kvale & Schreiner 2004, Tallal & Gaab 2006). The nervous system is able to encode patterns derived from the statistical features of incoming information and make predictions based on these patterns that guide both perception and action.

Temporal Patterns

In addition to aggregating probabilities over a period of time, the nervous system is also highly sensitive to how events are structured in time. Put simply, while statistical patterns allow prediction of what is likely to come next, temporal patterns indicate when something is likely to occur. As we know from the experience of listening to unfamiliar languages, separating a continuous stream of speech into meaningful units is no simple task. In addition to the statistical patterns discussed above, temporal patterns can also help a listener to discern the meaningful structure of speech (Cunillera et al. 2006, Cutler & Butterfield 1992, Nakatani & Schaffer 1978). In fact, there is

evidence that stress patterns in speech outweigh statistical cues for determining word boundaries when conflicting cues are pitted against each other (Johnson & Jusczyk 2001).

The ability to anticipate based on temporal patterns brings significant processing advantages. For example, the dynamic attending theory purports that the brain may modulate attention over time, such that attention is maximized when important events are most likely to occur (Large & Jones 1999). In other words, the brain synchronizes its activity with the temporal structures present in its environment and increases processing efficiency by allocating resources when they are most likely to be needed.

The effectiveness of this strategy has been demonstrated experimentally; for example, once a temporal pattern is established (e.g., via a repeating tone), linguistic discrimination judgments are made more rapidly if the target sound fits into the established temporal structure rather than deviating from it (Cason & Schön 2012, Quene & Port 2005). Similarly, timing regularities in a speech stream may result in a listener developing temporal expectancies, and these expectancies can influence how subsequent sounds are processed based on whether those sounds align with the expected timing or not (Pitt & Samuel 1990, Quene & Port 2005, Roncaglia-Denissen et al. 2013, Schmidt-Kassow & Kotz 2008). Further, a study by Morillon and colleagues (2014) revealed that moving a finger silently to a reference beat improved the separation of on-beat auditory target tones from distractors, suggesting that the allocation of attention over time was locked to the rhythmic motor activity.

Sensitivity to durational patterns is particularly important for understanding speech under degraded listening conditions (Slater & Kraus 2015, Smith et al. 1989), and violations of expectation can also influence processing. For example, prolonging the duration of less predictable words can help a listener recognize them in a novel context (Turk & Shattuck-Hufnagel 2014). It is important to note that patterns therefore provide a framework that can modulate processing in two ways, either by emphasizing the importance of elements that are consistent with the pattern or by drawing attention to elements that do not fit the pattern. Both sides of the coin are reflected in neural processing, with certain brain responses signaling deviation from expectation (for example, the mismatch negativity, which is enhanced in response to a deviant stimulus; Näätänen 1995), and others reflecting conformity to a pattern (for example, the enhanced subcortical response to speech sounds in a regular versus unpredictable context; Parbery-Clark et al. 2011). Our nuanced relationship with patterns—especially the tension between conformity and deviation—is integral to how we communicate, and it has been proposed that much of the emotional power of music stems from the creation (and violation) of temporal expectations within a predictable framework (Huron 2006).

Links with reading ability. The previous section emphasized that temporal patterns play an important role in guiding auditory perception. Although it is not always intuitive to think of reading as an auditory skill, typically developing children derive their first internal representations of linguistic meaning by parsing the sounds of speech. Effective parsing mechanisms are therefore critical to the development of accurate phonemic representations, which can later be mapped to the orthographic representations of written language.

The ability to make use of rhythmic cues when perceiving speech has been linked to reading skills (for review, see Tierney & Kraus 2013b). A recent study by Woodruff Carr et al. (2014) compared preschoolers who could synchronize to a beat with those who could not: Those who could synchronize had better prereading skills, such as the ability to separate words into their individual sounds, than those who were unable to synchronize. The good synchronizers also had more precise neural encoding of the temporal modulations in speech (Woodruff Carr et al. 2014), suggesting that one of the fundamental mechanisms common to language skills and

synchronization ability may be the precision with which the nervous system can encode temporal features. This is discussed further in the next section.

Impaired rhythmic abilities have been linked with language disorders such as dyslexia (Corriveau & Goswami 2009, Overy et al. 2003, Thomson & Goswami 2008), and a study assessing metric perception showed that performance strongly predicted reading ability and phonologic awareness across a population of normal and dyslexic subjects (Huss et al. 2010). These findings have led to the development of rhythm-based interventions to address reading difficulties such as dyslexia, with some success (Bhide et al. 2013; Overy 2000, 2003). However, the various studies demonstrating links between rhythm-related skills and language skills have assessed a wide variety of rhythmic skills, and the rhythm-based interventions have each trained different skills. A clearer understanding of the connections between specific aspects of rhythmic processing and language skills could lead to more accurate diagnosis and better targeted treatment approaches for language problems that are associated with rhythmic deficits. It is also important to note that links between rhythm and language skills across a population do not necessarily mean that training the former can improve the latter in an individual. Further longitudinal studies are needed to better understand the impact of training and to determine whether training that specifically targets certain rhythmic skills is indeed more effective than developing these skills within the context of broader musical training or using other auditory training approaches.

Distinct components of rhythm processing. A great deal remains to be learned about how rhythmic subskills relate with one another, and researchers have proposed various ways of categorizing them. Many rhythm-related activities involve the synchronization of movement to sound, such as tapping to a simple metronome or to the beat of a piece of music. Tapping accurately to a metronome involves perceiving the regular pattern, anticipating the next beat, and coordinating movement accordingly, as well as updating the motor plan based on any discrepancy between the produced movement and the target sound. Tasks involving the production of movement may yield different outcomes than those involving perceptual judgment alone, because humans have been shown to adapt their movements to timing perturbations that are below the threshold of perception (Repp 2000); in other words, individuals may adjust their tapping to a timing perturbation of 15 ms yet report that they did not perceive any shift. Some rhythmic tasks also rely more heavily on memory than others; for example, discriminating between complex rhythmic sequences involves maintaining one sequence in memory to compare with the next.

Tapping to the beat of music involves an additional layer of processing, since a listener must first extract the underlying pulse of the music and then synchronize with that pulse. Moving to the beat of music is something that comes naturally to most people, even young children, but beat and metrical structure are ultimately perceptual constructs. For example, the same piece of music could induce a different metrical percept in different listeners or under different conditions, and metrical structure can be perceived even when individual accented beats are absent (for example, in a syncopated melody) (Iversen et al. 2009). Tasks also differ in terms of complexity and predictability: Musical examples can be extremely complex yet still arranged around a roughly isochronous and predictable beat, whereas speech does not generally adhere to an isochronous framework. It is important to note, however, that live musical performance does contain significant timing fluctuations and more complex temporal patterns, closer to the free-flowing rhythms of speech (Palmer 1997, Repp 1992), and these subtle timing cues contribute greatly to the expressive quality of music (Ashley 2002). Despite the tempo fluctuations of live performances, listeners have no difficulty perceiving the underlying pulse of music; in fact, it has been suggested that beat perception may be helped by this natural timing variability (Rankin et al. 2009). This ability is thought to rely upon the entrainment of neural oscillators to the beat, and this self-sustaining

"internal beat" can be tolerant to minor variations and syncopations that may be present in the actual acoustic signal (Large & Snyder 2009, Nozaradan 2014, Nozaradan et al. 2012). Although the temporal patterns of natural speech seem less predictable than the beat of music, it is perhaps easier than one might expect to speak in unison with another speaker, even when the content is unfamiliar (Cummins 2013). This may reflect the close integration of auditory and motor systems, which may facilitate prediction based on articulatory and phrasing cues. Further explication of these nonperiodic aspects of timing and synchronization is a promising line of research that may provide important insight into the links between music and speech (Cummins 2013, Patel 2010).

As discussed previously, the tension between conformity and deviation is an important component of how the brain processes sound, and this tension is especially relevant to rhythm and temporal processing. Whereas the extraction of beat and metrical structure requires tolerance of minor timing variations, other elements of musical communication—such as expressive timing or coordination between players in an ensemble—require direct responses to those same minor variations. Experimentally, participants may be directed to perform specific tasks that isolate different aspects of rhythmic processing, such as beat-based tasks in which timing information is perceived in relation to an underlying beat (for example, detecting whether a target sound is on or off the beat), and duration-based tasks, which involve the perception of absolute timing information (such as determining whether two tones are of the same length) (Merchant et al. 2015). However, in more natural settings the brain must constantly find its own balance between latching onto stability and structure versus detecting and responding to deviations. The resulting interaction between the brain's internal time keeping and its ability to respond to incoming sounds is reflected in the neural circuitry underlying rhythm perception. The so-called motor regions that were discussed above, in the context of their role in pattern detection and perception, are actively involved in rhythm perception (Zatorre et al. 2007). These areas include the basal ganglia, frontal cortex, cerebellum, and midbrain, and because these areas are so highly interconnected, attempts to attach specific functions to distinct areas have met with limited success. However, it is thought that the basal ganglia play a particular role in generating internal representations of beat and metrical structure, whereas the cerebellum is important for coordinating precise motor movements and in tracking durations and timing in complex sound sequences (Grahn 2012, Grube et al. 2010, Merchant et al. 2015). These interconnected systems work together, integrating the big-picture patterns with the fine details, striving for the optimal balance between stability and flexibility.

The important role of motor areas in rhythmic processing is further emphasized by studies showing that movement influences metrical perception in both infants and adults. In the initial study with infants (Phillips-Silver & Trainor 2005), an experimenter bounced the infants on different beats of a rhythm pattern during a training period. The rhythm pattern had ambiguous metrical structure, and the training period specifically emphasized one of the two possible metrical structures. How the babies were bounced during the training phase influenced their subsequent listening preferences when they listened to two auditory versions of the rhythm patterns: The infants preferred to listen to the auditory version with intensity accents that matched the beats on which they had been bounced. There was no such effect when babies observed bouncing but were not bounced themselves, confirming that the effect was due to movement rather than visual input (Phillips-Silver & Trainor 2005). A follow-up study with adults (Phillips-Silver & Trainor 2007) involved participants bouncing to the beat of an auditory rhythm pattern by bending their knees, either emphasizing a waltz- or march-like metrical structure. After this training period, participants listened to auditory versions of the same sequences with increased intensities on certain beats, to match either the waltz or march form. As with the infant study, the bouncing pattern during the training period influenced how the participants perceived the subsequent auditory patterns,

with participants reporting greater similarity when the auditory patterns matched their bouncing patterns than when they did not (Phillips-Silver & Trainor 2007).

Research from our laboratory has focused on identifying distinct areas of rhythmic ability that are linked with reading ability. Interestingly, this work has revealed that rhythmic skills may in fact be broken down into distinct rhythmic "intelligences" that do not necessarily pattern together, and the broad connections observed between rhythm skills and reading ability may in fact reflect multiple underlying mechanisms. In particular, the ability to remember and reproduce rhythmic sequences is not necessarily linked to the ability to synchronize accurately to an auditory stimulus and adjust to fine timing perturbations, yet both abilities track with reading skills (Tierney & Kraus 2015). This distinction emphasizes that both music and language involve meaningful information at different timescales, and that effective communication involves integration across these timescales, with both fine temporal precision and sensitivity to rhythmic patterns playing a role.

COMMUNICATION THROUGH SOUND: BIOLOGICAL FOUNDATIONS

Precise Timing in the Auditory System

The auditory system, capable of much finer temporal resolution than other sensory systems, is specialized for timing (Griffiths et al. 2001). Highly efficient neural coding strategies have evolved across species to handle the microsecond timing sensitivity necessary for tasks such as sound localization (Mauk & Buonomano 2004), and in humans, precise encoding of temporal features is especially important for speech perception because the most meaningful parts of the signal are carried by the fastest-changing components, namely the consonants. The ability to encode these subtle timing characteristics is critical to developing accurate phonologic representations and therefore essential for the development of strong reading skills, with more precise subcortical timing linked to better reading abilities (Banai et al. 2009, Hornickel et al. 2009); reading difficulties are associated with less consistent timing in neural responses to speech (Hornickel & Kraus 2013). The ability to synchronize accurately with a metronome involves timing precision on the order of milliseconds (Madison & Merker 2004, Repp 2000, Thaut & Kenyon 2003), which in turn relies heavily on the precision and consistency of neural timing (Tierney & Kraus 2013a, 2014). Research with adolescents reveals that the ability to synchronize with a beat is associated with more consistent subcortical neural timing in the auditory system in response to speech sounds (Tierney & Kraus 2013a) as well as better cognitive and linguistic skills (Tierney & Kraus 2013c) (see **Figure 4**). These outcomes have led to the hypothesis that it may be the precise neural timing involved in entrainment to a beat that at least in part underpins the converging evidence that reading skills are strengthened by musical training, specifically by improving phonological awareness (reviewed in Tierney & Kraus 2014).

Vocal Learning: The Auditory-Motor Connection?

In humans, sound plays a uniquely important role in the coordination of fine motor control, as is reflected in the ability to synchronize to much faster auditory sequences than visual (Repp 2003) and less variability when tapping to an auditory signal than to a visual stimulus (Chen et al. 2002, Hove et al. 2012, Kolers & Brewster 1985, Patel et al. 2005). Interestingly, this auditory advantage does not exist in other species, such as macaque monkeys (Zarco et al. 2009), despite their ability to accurately perceive timing intervals. In other words, synchronization skills require not only
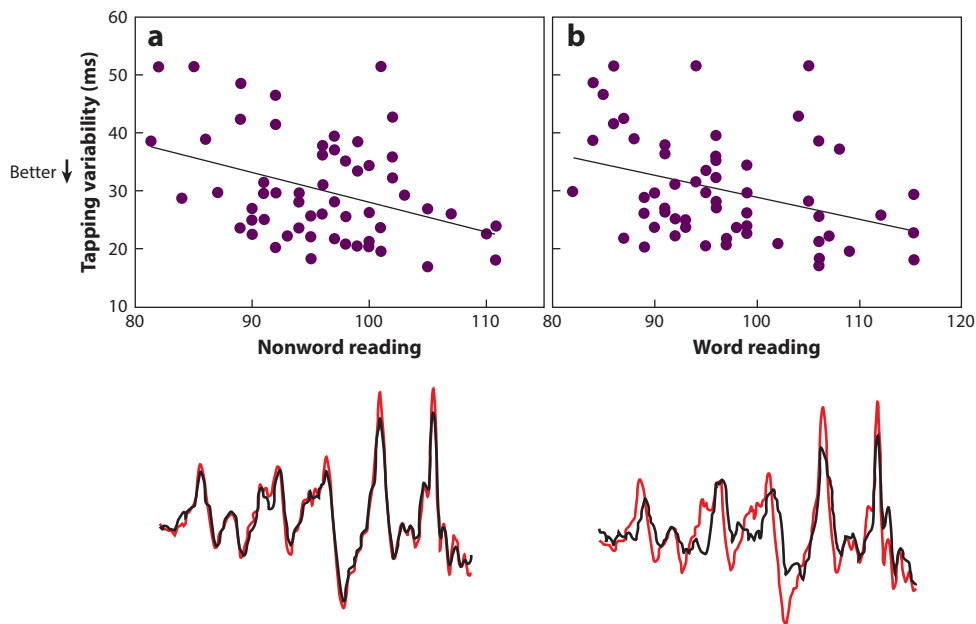
Figure 4

Synchronization to a metronome requires precise temporal encoding in the auditory system, which is also critical to the development of reading skills. (*Top*) Children who tap less variably to a beat have better reading skills. (*Bottom*) Less variable tappers (*a*) have less trial-to-trial variability in their neural response to the speech syllable /da/ than more variable tappers (*b*). Red and black waveforms represent two averages from the same recording session. Figure adapted with permission from Tierney & Kraus (2013a,b).

the precise encoding of sound, but also particular connections between the auditory and motor systems to allow that temporal precision to guide movement. In fact, the ability to synchronize to a regular beat is surprisingly rare in the animal kingdom (Patel et al. 2009). Nonhuman primates, such as monkeys, can perceive durations very accurately, yet extensive efforts to train monkeys to synchronize to an auditory stimulus have been notably unsuccessful (Merchant & Honing 2013, Zarco et al. 2009). Integration between auditory and motor systems is especially strong in vocal learning species, such as songbirds, which are not only able to mimic sounds in their environment but are also able to learn the underlying rules (e.g., grammar) of a sound system and create new sounds based on those rules. Coactivation of comparable motor and auditory brain regions has been observed in both humans and songbirds during vocal learning tasks (Brown et al. 2004), and it has been proposed that the auditory-motor feedback loops that evolved to support vocal learning may also underlie the ability to synchronize to a beat. Individual cases, such as the report of beat-keeping abilities in a sea lion (Cook et al. 2013), leave this a matter of continuing research, since sea lions have not previously been identified as a vocal learning species; however, other pinnipeds such as seals are known to be vocal learners. It may therefore be the vocal learning status of the sea lion that needs to be updated rather than the hypothesized connection between vocal learning and the ability to keep a beat (Patel & Iversen 2014). Furthermore, the common ground between vocal learning and beat synchronization may lie not only in connections between auditory and motor regions, but also more specifically in the motor areas involved in learning and generating patterns (e.g., the basal ganglia), since it is the ability to generate novel utterances based on inherent rules that distinguishes vocal learners from other species capable of vocalizations and vocal mimicry.

This is supported by converging evidence for a specific genetic factor, the *FOXP2* transcription factor gene, that is expressed in the basal ganglia circuitry of vocal learning species (for a review, see Fisher & Scharff 2009).

Neural Synchrony and Multiple Timescales

The brain must integrate information across multiple timescales at once, from the fine timing distinctions between consonant sounds to the longer-scale patterns of syllable rate and stress patterns. Integrating these streams of information into a coherent percept relies heavily on working memory, which determines the capacity for retaining sequences in memory and the scope of pattern matching that can be attempted. Neural oscillations may play an important role in coordinating activity across brain regions and across timescales because they allow the brain to "entrain" to various aspects of temporal structure (Giraud & Poeppel 2012). The cerebral hemispheres show specialization for fast versus slow timing characteristics, based on lateralization of oscillatory activity at different frequency bands (Belin et al. 1998, Poeppel 2003). Better reading abilities have been linked to stronger lateralization of relevant timing rates within the speech signal, such as right-hemispheric specialization for the slower syllable rate of speech (Abrams et al. 2009, Goswami 2011). The left-hemispheric specialization that supports rapid acoustic processing is linked to the precision of subcortical encoding of speech sounds, suggesting that accurate timing in the auditory brainstem is a fundamental ingredient in processing the fast acoustic properties of speech that are critical to the development of reading skills (Abrams et al. 2006, Tierney & Kraus 2013b). Neuronal entrainment is therefore happening simultaneously on multiple timescales, mirroring the multiple timescales of information that unfold in a spoken or musical phrase. Nested rates of brain activity can concurrently synchronize to different stimulus features, and increased coherence between different frequencies of neural oscillations (for example, theta and gamma bands) has been observed during short-term memory tasks (Schack et al. 2002) as well as during a task involving adaptation of speech movement based on auditory feedback (Sengupta & Nasir 2015). Coordination of pattern extraction and prediction across these timescales is therefore important not only for the efficient allocation of neural resources but also for learning. When this coordination breaks down, there can be behavioral consequences; for example, abnormal oscillatory patterns have been associated with reading impairments, such as dyslexia (Abrams et al. 2009, Goswami 2011, Heim et al. 2011, Nagarajan et al. 1999), and may contribute to the impaired multisensory integration that is common in dyslexics (Facoetti et al. 2010, Goswami 2011).

In summary, temporal patterns play an important role in guiding perception and are reflected in the rhythms of neural activity. Links between rhythmic skills and language-related skills, including reading, indicate some degree of common processing between these domains. Although there is still much to be understood about how various aspects of rhythmic processing relate with one another, there is evidence that at least two distinct areas of rhythmic skill are related with reading ability: first, the ability to distinguish fine timing characteristics, and second, the ability to remember and reproduce rhythmic sequences. A recurring theme throughout this review is that our experience of sound is rooted in the physical world, that sound is rooted in movement, and that our motor systems play an essential role in our perception of the inherent structure of sound. In the next section, we consider how these mechanisms can be shaped by experience.

THE IMPRINT OF EXPERIENCE

As we have already emphasized, our perception of communication sounds is influenced not only by acoustic features and immediate context, but also by our accumulated experience with sound

over the course of a lifetime. Each new speech utterance presents a new processing challenge: Much of our processing of sound happens automatically, without conscious attention, and this automatic processing is finely tuned by experience.

Automatic Sound Processing: Our Lab's Approach

The auditory system is a distributed but integrated system, capable of dynamic modulation of signal processing as well as long-term tuning of sensory function with experience. In our view, cognitive and sensory processing should not be viewed as distinct; rather, they are melded together through experience because what we have paid attention to in the past will shape our automatic response to new sounds in the present (Kraus & White-Schwoch 2015).

The auditory midbrain is an information-processing hub in which cortical and sensory inputs converge (Bajo et al. 2009, Nelken 2008, Suga & Ma 2003). Direct connections exist between the auditory midbrain and brain regions important for motor control and coordination, including the cerebellum, basal ganglia, and premotor cortex (Casseday et al. 2002). Investigation of the evoked auditory brainstem response to complex sounds (cABR, of which the inferior colliculus is a primary generator) has proved a fruitful approach in exploring experience-based plasticity (Skoe & Kraus 2010). The cABR preserves stimulus characteristics with great fidelity. By comparing stimulus and response characteristics in both time and frequency domains, it is possible to assess the neural representation of fundamental characteristics such as pitch, timing, and timbre. The cABR reflects experience with sound (Bidelman et al. 2011, Kraus & Chandrasekaran 2010, Krishnan et al. 2005, Tzounopoulos & Kraus 2009) and therefore provides an objective biologic snapshot of sound processing in an individual (Kraus 2011). This approach has been used to develop early biomarkers of language development (White-Schwoch et al. 2015) as well as to explore how auditory processing is shaped by expertise, such as musical experience, and impairment, such as dyslexia and autism (reviewed in Kraus & Nicol 2014).

The Musician Signature

Experience does not shape auditory processing with a simple volume knob effect; rather, it selectively enhances relevant components and attenuates irrelevant inputs (Kraus & White-Schwoch 2015). What is relevant to any given individual will determine the focus of his or her experience and, over time, this accumulated experience results in distinct "neural signatures" of expertise (Kraus & Nicol 2014). For example, the brain response of musicians can be tuned in to the specific timbre of the instrument they play, such that the familiar timbre produces a larger response than does the sound of another instrument (Strait et al. 2012). The style of musical playing can also affect automatic processing of sound, as shown by jazz musicians demonstrating greater sensitivity to subtle acoustic variations in their preattentive brain responses in comparison with musicians of other genres (Vuust et al. 2012). However, some of the effects of experience are more general and can transfer to other domains. Converging evidence suggests that the ability to parse a complex auditory scene can be strengthened by musical practice, and there are numerous examples of musical expertise conferring advantages for speech processing (Patel 2011). For example, the neural representation of fine timing characteristics is more precise in musicians than in nonmusicians: Musicians show greater neural differentiation of contrastive consonant sounds than do nonmusicians across the life span (Parbery-Clark et al. 2012c, Strait et al. 2013), and this has been demonstrated longitudinally in elementary school children following two years of music training (Kraus et al. 2014; reviewed in Kraus & Strait 2015). Furthermore, musicians' subcortical encoding of sound is influenced by statistical predictability (Parbery-Clark et al. 2011, Skoe et al.

2013b), suggesting that musical experience may strengthen the ability of the nervous system to benefit from underlying patterns in sound in both musical and linguistic contexts.

Recent work indicates that just a few years of musical training in childhood can influence the neural encoding of sound in adulthood, years after the training has ceased (Skoe & Kraus 2012, White-Schwoch et al. 2013), emphasizing that past experience with sound can shape automatic sound processing in the present, even many years later. It is proposed that the observed enhancements in the neural encoding of sound may be driven in a top-down manner by strengthened cognitive function, such as enhanced working memory and attention, which help to tune sensory function based on prior experience and specific task demands (Ahissar et al. 2009, Conway et al. 2009, Kraus & Chandrasekaran 2010, Kraus et al. 2012, Kraus & White-Schwoch 2015, Nelken & Ulanovsky 2007). This results in more effective targeting of neural resources and better synchronization of underlying neural activity, thereby promoting more consistent and robust neural responses (Krizman et al. 2012, Parbery-Clark et al. 2012b, Skoe & Kraus 2013, Tierney et al. 2015). For example, when listeners are presented with two concurrent speech streams and asked to attend to one and ignore the other, their neural responses to the attended stream are more consistent from trial to trial than the responses to ignored speech, demonstrating that attention can help to coordinate firing across populations of neurons (Strait & Kraus 2011). Neural processing advantages in musicians are linked to behavioral advantages in everyday communication tasks, such as understanding speech in a noisy background. Musicians across the life span demonstrate superior speech-in-noise perception and auditory working memory compared with nonmusician peers (for a review, see Strait & Kraus 2013, Swaminathan et al. 2015, Zendel et al. 2015), and improved speech-in-noise perception was observed longitudinally in elementary school children following two years of music class (Slater et al. 2015). Recent work reveals that the ability to perceive speech in noise is correlated with performance on a rhythm perception task (Slater & Kraus 2015), emphasizing that temporal patterns are especially important for following speech under difficult listening conditions.

Clinical and Educational Implications

Understanding the fundamental mechanisms by which humans process communication sounds not only is interesting from a theoretical perspective but also has great utility in clinical and educational settings. Many of the same aspects of neural processing that are strengthened in auditory experts, such as musicians, are found to be deficient in populations with language-related disorders (for a review, see Tierney & Kraus 2013b), suggesting that training and remediation that appropriately target underlying mechanisms could be effective in treating individuals with language-based deficits. Furthermore, the biological benefits of musical expertise may counteract some of the natural declines in neural processing associated with aging (Parbery-Clark et al. 2009, 2012a) as well as the negative impact on neural function that may result from living in poverty (Skoe et al. 2013a). Longitudinal studies have demonstrated positive effects of school-based music training on the neural processing of speech (Kraus & Strait 2015) and have revealed that neural processing of speech can be strengthened by short-term training in children with language-based learning problems (for a review, see Kraus & Hornickel 2012). The potential role of music-based interventions for the treatment of language disorders is a continuing area of research.

CONCLUSIONS

In this review we have emphasized that the sophisticated systems of modern communication are rooted in our more ancient relationship with sound. This relationship is grounded in the physical world and is reflected in the patterns that our brains seek out as we try to make sense of the sounds

we hear. These patterns also provide a framework for communicating with others, and there is close integration between our ability to produce as well as perceive communication sounds. The same neural networks involved in generating the movements to produce sound are also intimately involved in the perception and prediction of underlying patterns. Therefore, communication is not simply the transfer of sound signals from one person to another but rather is an interaction between physical entities. Evidence indicates that when two individuals synchronize their movements, this increases affiliation between them and promotes bonding (Cirelli et al. 2012, Hove & Risen 2009, Launay et al. 2013). The inherent rhythms of music and speech facilitate interpersonal synchrony by allowing us to predict what is coming next and align our movements with others (Phillips-Silver et al. 2010). From the dyadic exchange between mother and child (Censullo et al. 1985) to conversational turn-taking (Auer et al. 1999) and improvisational jazz (Berkowitz 2010), these patterns in time not only streamline information processing, they also help us to connect.

An important aspect of synchrony is that it is a natural signature of emotion, since the potent neural chemistry of emotion can trigger activation across multiple brain regions within a very short space of time (Scherer 2013) and may stem from the basic signaling mechanisms underlying fight or flight behavior. Interpersonal synchrony may build upon this foundation such that synchronized activity with another person implies shared emotion, and it has been suggested that synchrony may foster social bonds between individuals by blurring the perceived boundary between self and other (Tarr et al. 2014). In a sense, this blurring of boundaries between self and other is what communication is all about: Words may scratch the surface, but sound can move us beyond words.

FUTURE ISSUES

1. Understanding how different timescales of information are integrated in the brain will be an important area of future research.

2. Identifying distinct components of rhythmic processing will be important for understanding connections between music and language skills.

3. Longitudinal studies will be important to assess whether rhythm-based training can improve language skills within an individual.

4. Identifying biological markers of communication deficits will allow for more effective treatment and early intervention.

5. Motor regions of the brain may play a broader role in pattern-based behavior than previously thought.

6. Further investigation of dynamic synchrony, for example, two people speaking in unison, may help reveal predictive mechanisms important for speech.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

LITERATURE CITED

Abrams DA, Nicol T, Zecker S, Kraus N. 2009. Abnormal cortical processing of the syllable rate of speech in poor readers. *J. Neurosci.* 29:7686–93

Abrams DA, Nicol T, Zecker SG, Kraus N. 2006. Auditory brainstem timing predicts cerebral asymmetry for speech. *J. Neurosci.* 26:11131–37

Ahissar M, Nahum M, Nelken I, Hochstein S. 2009. Reverse hierarchies and sensory learning. *Philos. Trans. R. Soc. B* 364:285–99

Andreou L-V, Kashino M, Chait M. 2011. The role of temporal regularity in auditory segregation. *Hear. Res.* 280:228–35

Antunes FM, Nelken I, Covey E, Malmierca MS. 2010. Stimulus-specific adaptation in the auditory thalamus of the anesthetized rat. *PLOS ONE* 5:e14071

Ashley R. 2002. Do[n't] change a hair for me: the art of jazz rubato. *Music Percept.* 19:311–32

Auer P, Couper-Kuhlen E, Müller F. 1999. *Language in Time: The Rhythm and Tempo of Spoken Interaction*. Oxford, UK: Oxford Univ. Press

Bajo VM, Nodal FR, Moore DR, King AJ. 2009. The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nat. Neurosci.* 13:253–60

Banai K, Hornickel J, Skoe E, Nicol T, Zecker S, Kraus N. 2009. Reading and subcortical auditory function. *Cereb. Cortex* 19:2699–707

Belin P, Zilbovicius M, Crozier S, Thivard L, Fontaine A, et al. 1998. Lateralization of speech and auditory temporal processing. *J. Cogn. Neurosci.* 10:536–40

Berkowitz AL. 2010. *The Improvising Mind: Cognition and Creativity in the Musical Moment*. New York: Oxford Univ. Press

Bhide A, Power A, Goswami U. 2013. A rhythmic musical intervention for poor readers: a comparison of efficacy with a letter-based intervention. *Mind Brain Educ.* 7:113–23

Bidelman GM, Gandour JT, Krishnan A. 2011. Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *J. Cogn. Neurosci.* 23:425–34

Brown S, Martinez MJ, Hodges DA, Fox PT, Parsons LM. 2004. The song system of the human brain. *Cogn. Brain Res.* 20:363–75

Cason N, Schön D. 2012. Rhythmic priming enhances the phonological processing of speech. *Neuropsychologia* 50:2652–58

Casseday JH, Fremouw T, Covey E. 2002. The inferior colliculus: a hub for the central auditory system. In *Integrative Functions in the Mammalian Auditory Pathway*, ed. D Oertel, AN Popper, RR Fay, pp. 238–318. New York: Springer

Censullo M, Lester B, Hoffman J. 1985. Rhythmic patterning in mother-newborn interaction. *Nurs. Res.* 34:342–46

Chen Y, Repp BH, Patel AD. 2002. Spectral decomposition of variability in synchronization and continuation tapping: comparisons between auditory and visual pacing and feedback conditions. *Hum. Mov. Sci.* 21:515–32

Cirelli LK, Wan SJ, Trainor LJ. 2014. Fourteen-month-old infants use interpersonal synchrony as a cue to direct helpfulness. *Philos. Trans. R. Soc. B* 369:20130400

Conway CM, Pisoni DB. 2008. Neurocognitive basis of implicit learning of sequential structure and its relation to language processing. *Ann. N. Y. Acad. Sci.* 1145:113–31

Conway CM, Pisoni DB, Kronenberger WG. 2009. The importance of sound for cognitive sequencing abilities: the auditory scaffolding hypothesis. *Curr. Dir. Psychol. Sci.* 18:275–79

Cook P, Rouse A, Wilson M, Reichmuth C. 2013. A California sea lion (*Zalophus californianus*) can keep the beat: motor entrainment to rhythmic auditory stimuli in a non vocal mimic. *J. Comp. Psychol.* 127:412–27

Corriveau KH, Goswami U. 2009. Rhythmic motor entrainment in children with speech and language impairments: tapping to the beat. *Cortex* 45:119–30

Cummins F. 2013. Joint speech: the missing link between speech and music? *Percepta: J. Musical Cogn.* 1:17–32

Cunillera T, Toro JM, Sebastián-Gallés N, Rodríguez-Fornells A. 2006. The effects of stress and statistical cues on continuous speech segmentation: an event-related brain potential study. *Brain Res.* 1123:168–78

Cutler A, Butterfield S. 1992. Rhythmic cues to speech segmentation: evidence from juncture misperception. *J. Mem. Lang.* 31:218–36

Dean I, Harper NS, McAlpine D. 2005. Neural population coding of sound level adapts to stimulus statistics. *Nat. Neurosci.* 8:1684–89

Dean I, Robinson BL, Harper NS, McAlpine D. 2008. Rapid neural adaptation to sound level statistics. *J. Neurosci.* 28:6430–38

Ettlinger M, Margulis EH, Wong PC. 2011. Implicit memory in music and language. *Front. Psychol.* 2:211

Facoetti A, Trussardi AN, Ruffino M, Lorusso ML, Cattaneo C, et al. 2010. Multisensory spatial attention deficits are predictive of phonological decoding skills in developmental dyslexia. *J. Cogn. Neurosci.* 22:1011–25

Fisher SE, Scharff C. 2009. *FOXP2* as a molecular window into speech and language. *Trends Genet.* 25:166–77

François C, Chobert J, Besson M, Schön D. 2013. Music training for the development of speech segmentation. *Cereb. Cortex* 23:2038–43

François C, Schön D. 2011. Musical expertise boosts implicit learning of both musical and linguistic structures. *Cereb. Cortex* 21:2357–65

Giraud A-L, Poeppel D. 2012. Cortical oscillations and speech processing: emerging computational principles and operations. *Nat. Neurosci.* 15:511–17

Goswami U. 2011. A temporal sampling framework for developmental dyslexia. *Trends Cogn. Sci.* 15:3–10

Grahn JA. 2012. Neural mechanisms of rhythm perception: current findings and future perspectives. *Top. Cogn. Sci.* 4:585–606

Graybiel AM. 2005. The basal ganglia: learning new tricks and loving it. *Curr. Opin. Neurobiol.* 15:638–44

Griffiths TD, Uppenkamp S, Johnsrude I, Josephs O, Patterson RD. 2001. Encoding of the temporal regularity of sound in the human brainstem. *Nat. Neurosci.* 4:633–37

Grube M, Lee K-H, Griffiths TD, Barker AT, Woodruff PW. 2010. Frontiers: Transcranial magnetic theta-burst stimulation of the human cerebellum distinguishes absolute, duration-based from relative, beat-based perception of subsecond time intervals. *Front. Psychol.* 1:171

Hannon EE, Trehub SE. 2005. Metrical categories in infancy and adulthood. *Psychol. Sci.* 16:48–55

Hauser MD, Newport EL, Aslin RN. 2001. Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins. *Cognition* 78:B53–64

Heim S, Friedman JT, Keil A, Benasich AA. 2011. Reduced sensory oscillatory activity during rapid auditory processing as a correlate of language-learning impairment. *J. Neurolinguist.* 24:538–55

Hornickel J, Kraus N. 2013. Unstable representation of sound: a biological marker of dyslexia. *J. Neurosci.* 33:3500–4

Hornickel J, Skoe E, Nicol T, Zecker S, Kraus N. 2009. Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception. *PNAS* 106:13022–27

Hove MJ, Iversen JR, Zhang A, Repp BH. 2012. Synchronization with competing visual and auditory rhythms: Bouncing ball meets metronome. *Psychol. Res.* 77:388–98

Hove MJ, Risen JL. 2009. It's all in the timing: Interpersonal synchrony increases affiliation. *Soc. Cogn.* 27:949–60

Huron DB. 2006. *Sweet Anticipation: Music and the Psychology of Expectation*. Cambridge, MA: MIT Press

Huss M, Verney JP, Fosker T, Mead N, Goswami U. 2010. Music, rhythm, rise time perception and developmental dyslexia: Perception of musical meter predicts reading and phonology. *Cortex* 47:674–89

Iversen JR, Repp BH, Patel AD. 2009. Top-down control of rhythm perception modulates early auditory responses. *Ann. N. Y. Acad. Sci.* 1169:58–73

Johnson EK, Jusczyk PW. 2001. Word segmentation by 8-month-olds: when speech cues count more than statistics. *J. Mem. Lang.* 44:548–67

Kirkham NZ, Slemmer JA, Johnson SP. 2002. Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition* 83:B35–42

Knutson B, Adams CM, Fong GW, Hommer D. 2001. Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J. Neurosci.* 21:RC159

Kolers PA, Brewster JM. 1985. Rhythms and responses. *J. Exp. Psychol.: Hum. Percept. Perform.* 11:150–67

Kotz SA, Schwartze M, Schmidt-Kassow M. 2009. Non-motor basal ganglia functions: a review and proposal for a model of sensory predictability in auditory language perception. *Cortex* 45:982–90

Kraus N. 2011. Listening in on the listening brain. *Phys. Today* 64:40–45

Kraus N, Chandrasekaran B. 2010. Music training for the development of auditory skills. *Nat. Rev. Neurosci.* 11:599–605

Kraus N, Hornickel J. 2012. Meaningful engagement with sound for strengthening communication skills. In *Auditory Processing Disorders: Assessment, Management and Treatment*, ed. D Geffner, D Ross-Swain, pp. 693–717. San Diego: Plural

Kraus N, Nicol T. 2014. The cognitive auditory system: the role of learning in shaping the biology of the auditory system. In *Perspectives on Auditory Research*, ed. AN Popper, RR Fay, pp. 299–319. New York: Springer Sci.+Bus. Media

Kraus N, Slater J, Thompson EC, Hornickel J, Strait DL, et al. 2014. Music enrichment programs improve the neural encoding of speech in at-risk children. *J. Neurosci.* 34:11913–18

Kraus N, Strait DL. 2015. Emergence of biological markers of musicianship with school-based music instruction. *Ann. N. Y. Acad. Sci.* 1337:163–69

Kraus N, Strait DL, Parbery-Clark A. 2012. Cognitive factors shape brain networks for auditory skills: spotlight on auditory working memory. *Ann. N. Y. Acad. Sci.* 1252:100–7

Kraus N, White-Schwoch T. 2015. Unraveling the biology of auditory learning: a cognitive-sensorimotor-reward framework. *Trends Cogn. Sci.* In press

Krishnan A, Xu Y, Gandour J, Cariani P. 2005. Encoding of pitch in the human brainstem is sensitive to language experience. *Brain Res. Cogn. Brain Res.* 25:161–68

Krizman J, Marian V, Shook A, Skoe E, Kraus N. 2012. Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *PNAS* 109:7877–81

Kvale MN, Schreiner CE. 2004. Short-term adaptation of auditory receptive fields to dynamic stimuli. *J. Neurophysiol.* 91:604–12

Large EW, Jones MR. 1999. The dynamics of attending: how people track time-varying events. *Psychol. Rev.* 106:119–59

Large EW, Snyder JS. 2009. Pulse and meter as neural resonance. *Ann. N. Y. Acad. Sci.* 1169:46–57

Launay J, Dean RT, Bailes F. 2013. Synchronization can influence trust following virtual interaction. *Exp. Psychol.* 60:53–63

Madison G, Merker B. 2004. Human sensorimotor tracking of continuous subliminal deviations from isochrony. *Neurosci. Lett.* 370:69–73

Malmierca MS, Cristaudo S, Pérez-González D, Covey E. 2009. Stimulus-specific adaptation in the inferior colliculus of the anesthetized rat. *J. Neurosci.* 29:5483–93

Mauk MD, Buonomano DV. 2004. The neural basis of temporal processing. *Annu. Rev. Neurosci.* 27:307–40

McGurk H, MacDonald J. 1976. Hearing lips and seeing voices. *Nature* 264:746–48

Merchant H, Grahn J, Trainor L, Rohrmeier M, Fitch WT. 2015. Finding the beat: a neural perspective across humans and non-human primates. *Philos. Trans. R. Soc. Lond. B* 370:20140093

Merchant H, Honing H. 2013. Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. *Front. Neurosci.* 7:274

Morillon B, Schroeder CE, Wyart V. 2014. Motor contributions to the temporal precision of auditory attention. *Nat. Commun.* 5:5255

Morris G, Nevet A, Arkadir D, Vaadia E, Bergman H. 2006. Midbrain dopamine neurons encode decisions for future action. *Nat. Neurosci.* 9:1057–63

Musacchia G, Strait D, Kraus N. 2008. Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. *Hear. Res.* 241:34–42

Näätänen R. 1995. The mismatch negativity: a powerful tool for cognitive neuroscience. *Ear Hear.* 16:6–18

Nagarajan S, Mahncke H, Salz T, Tallal P, Roberts T, Merzenich MM. 1999. Cortical auditory signal processing in poor readers. *PNAS* 96:6483–88

Nakatani LH, Schaffer JA. 1978. Hearing "words" without words: prosodic cues for word perception. *J. Acoust. Soc. Am.* 63:234–45

Nelken I. 2008. Processing of complex sounds in the auditory system. *Curr. Opin. Neurobiol.* 18:413–17

Nelken I, Ulanovsky N. 2007. Mismatch negativity and stimulus-specific adaptation in animal models. *J. Psychophysiol.* 21:214–23

Nozaradan S. 2014. Exploring how musical rhythm entrains brain activity with electroencephalogram frequency-tagging. *Philos. Trans. R. Soc. Lond. B* 369:20130393

Nozaradan S, Peretz I, Mouraux A. 2012. Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. *J. Neurosci.* 32:17572–81

Overy K. 2000. Dyslexia, temporal processing and music: the potential of music as an early learning aid for dyslexic children. *Psychol. Music* 28:218–29

Overy K. 2003. Dyslexia and music. From timing deficits to musical intervention. *Ann. N.Y. Acad. Sci.* 999:497–505

Overy K, Nicolson RI, Fawcett AJ, Clarke EF. 2003. Dyslexia and music: measuring musical timing skills. *Dyslexia* 9:18–36

Palmer C. 1997. Music performance. *Annu. Rev. Psychol.* 48:115–38

Parbery-Clark A, Anderson S, Hittner E, Kraus N. 2012a. Musical experience offsets age-related delays in neural timing. *Neurobiol. Aging* 33:1483.e1–4

Parbery-Clark A, Anderson S, Hittner E, Kraus N. 2012b. Musical experience strengthens the neural representation of sounds important for communication in middle-aged adults. *Front. Aging Neurosci.* 4:30

Parbery-Clark A, Skoe E, Kraus N. 2009. Musical experience limits the degradative effects of background noise on the neural processing of sound. *J. Neurosci.* 29:14100–7

Parbery-Clark A, Strait D, Kraus N. 2011. Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians. *Neuropsychologia* 49:3338–45

Parbery-Clark A, Tierney A, Strait DL, Kraus N. 2012c. Musicians have fine-tuned neural distinction of speech syllables. *Neuroscience* 219:111–19

Patel AD. 2010. *Music, Language, and the Brain*. New York: Oxford Univ. Press

Patel AD. 2011. Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Front. Psychol.* 2:142

Patel AD, Iversen JR. 2014. The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. *Front. Syst. Neurosci.* 8:57

Patel AD, Iversen JR, Bregman MR, Schulz I. 2009. Studying synchronization to a musical beat in nonhuman animals. *Ann. N. Y. Acad. Sci.* 1169:459–69

Patel AD, Iversen JR, Chen Y, Repp BH. 2005. The influence of metricality and modality on synchronization with a beat. *Exp. Brain Res.* 163:226–38

Phillips-Silver J, Aktipis CA, Bryant GA. 2010. The ecology of entrainment: foundations of coordinated rhythmic movement. *Music Percept.* 28:3–14

Phillips-Silver J, Trainor LJ. 2005. Feeling the beat: Movement influences infant rhythm perception. *Science* 308:1430

Phillips-Silver J, Trainor LJ. 2007. Hearing what the body feels: auditory encoding of rhythmic movement. *Cognition* 105:533–46

Pitt MA, Samuel AG. 1990. The use of rhythm in attending to speech. *J. Exp. Psychol.: Hum. Percept. Perform.* 16:564–73

Poeppel D. 2003. The analysis of speech in different temporal integration windows: cerebral lateralization as "asymmetric sampling in time." *Speech Commun.* 41:245–55

Quene H, Port RF. 2005. Effects of timing regularity and metrical expectancy on spoken-word perception. *Phonetica* 62:1–13

Rankin SK, Large EW, Fink PW. 2009. Fractal tempo fluctuation and pulse prediction. *Music Percept.* 26:401–13

Repp BH. 1992. Diversity and commonality in music performance: an analysis of timing microstructure in Schumann's "Traumerei." *J. Acoust. Soc. Am.* 92:2546–68

Repp BH. 2000. Compensation for subliminal timing perturbations in perceptual-motor synchronization. *Psychol. Res.* 63:106–28

Repp BH. 2003. Rate limits in sensorimotor synchronization with auditory and visual sequences: the synchronization threshold and the benefits and costs of interval subdivision. *J. Mot. Behav.* 35:355–70

Roncaglia-Denissen MP, Schmidt-Kassow M, Kotz SA. 2013. Speech rhythm facilitates syntactic ambiguity resolution: ERP evidence. *PLOS ONE* 8:e56000

Saffran J, Hauser M, Seibel R, Kapfhamer J, Tsao F, Cushman F. 2008. Grammatical pattern learning by human infants and cotton-top tamarin monkeys. *Cognition* 107:479–500

Saffran JR, Aslin RN, Newport EL. 1996. Statistical learning by 8-month-old infants. *Science* 274:1926–28

Schack B, Vath N, Petsche H, Geissler H-G, Möller E. 2002. Phase-coupling of theta-gamma EEG rhythms during short-term memory processing. *Int. J. Psychophysiol.* 44:143–63

Scherer K. 2013. The evolutionary origin of multimodal synchronization in emotional expression. *J. Anthropol. Sci.* 91:1–16

Schmidt-Kassow M, Kotz SA. 2008. Entrainment of syntactic processing? ERP-responses to predictable time intervals during syntactic reanalysis. *Brain Res.* 1226:144–55

Schubotz RI. 2007. Prediction of external events with our motor system: towards a new framework. *Trends Cogn. Sci.* 11:211–18

Schutz M, Lipscomb S. 2007. Hearing gestures, seeing music: Vision influences perceived tone duration. *Perception* 36:888–97

Scott SK, Johnsrude IS. 2003. The neuroanatomical and functional organization of speech perception. *Trends Neurosci.* 26:100–7

Sengupta R, Nasir SM. 2015. Redistribution of neural phase coherence reflects establishment of feedforward map in speech motor adaptation. *J. Neurophysiol.* 113:2471–79

Shamma SA, Elhilali M, Micheyl C. 2011. Temporal coherence and attention in auditory scene analysis. *Trends Neurosci.* 34:114–23

Shook A, Marian V, Bartolotti J, Schroeder SR. 2013. Musical experience influences statistical learning of a novel language. *Am. J. Psychol.* 126:95–104

Skoe E, Kraus N. 2010. Auditory brain stem response to complex sounds: a tutorial. *Ear Hear.* 31:302–24

Skoe E, Kraus N. 2012. A little goes a long way: how the adult brain is shaped by musical training in childhood. *J. Neurosci.* 32:11507–10

Skoe E, Kraus N. 2013. Musical training heightens auditory brainstem function during sensitive periods in development. *Front. Psychol.* 4:622

Skoe E, Krizman J, Kraus N. 2013a. The impoverished brain: disparities in maternal education affect the neural response to sound. *J. Neurosci.* 33:17221–31

Skoe E, Krizman J, Spitzer E, Kraus N. 2013b. The auditory brainstem is a barometer of rapid auditory learning. *Neuroscience* 243:104–14

Slater J, Kraus N. 2015. The role of rhythm in perceiving speech in noise: a comparison of percussionists, vocalists and non-musicians. *Cogn. Process*. In press

Slater J, Skoe E, Strait DL, O'Connell S, Thompson E, Kraus N. 2015. Music training improves speech-in-noise perception: longitudinal evidence from a community-based music program. *Behav. Brain Res.* 291:244–52

Smith MR, Cutler A, Butterfield S, Nimmo-Smith I. 1989. The perception of rhythm and word boundaries in noise-masked speech. *J. Speech Hear. Res.* 32:912–20

Strait DL, Chan K, Ashley R, Kraus N. 2012. Specialization among the specialized: Auditory brainstem function is tuned in to timbre. *Cortex* 48:360–62

Strait DL, Kraus N. 2011. Can you hear me now? Musical training shapes functional brain networks for selective auditory attention and hearing speech in noise. *Front. Psychol.* 2:113

Strait DL, Kraus N. 2013. Biological impact of auditory expertise across the life span: musicians as a model of auditory learning. *Hear. Res.* 308:109–21

Strait DL, O'Connell S, Parbery-Clark A, Kraus N. 2013. Musicians' enhanced neural differentiation of speech sounds arises early in life: developmental evidence from ages 3 to 30. *Cereb. Cortex* 24:2512–21

Suga N, Ma X. 2003. Multiparametric corticofugal modulation and plasticity in the auditory system. *Nat. Rev. Neurosci.* 4:783–94

Swaminathan J, Mason CR, Streeter TM, Best V, Kidd G Jr., Patel AD. 2015. Musical training, individual differences and the cocktail party problem. *Sci. Rep.* 5:11628

Tallal P, Gaab N. 2006. Dynamic auditory processing, musical experience and language development. *Trends Neurosci.* 29:382–90

Tarr B, Launay J, Dunbar RI. 2014. Music and social bonding: "self-other" merging and neurohormonal mechanisms. *Front. Psychol.* 5:1096

Teinonen T, Fellman V, Näätänen R, Alku P, Huotilainen M. 2009. Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neurosci.* 10:21

Thaut MH, Kenyon GP. 2003. Rapid motor adaptations to subliminal frequency shifts during syncopated rhythmic sensorimotor synchronization. *Hum. Mov. Sci.* 22:321–38

Thomson JM, Goswami U. 2008. Rhythmic processing in children with developmental dyslexia: Auditory and motor rhythms link to reading and spelling. *J. Physiol.* 102:120–29

Tierney A, Kraus N. 2013a. The ability to move to a beat is linked to the consistency of neural responses to sound. *J. Neurosci.* 33:14981–88

Tierney A, Kraus N. 2013b. Music training for the development of reading skills. *Progress in Brain Research. Changing Brains: Applying Brain Plasticity to Advance and Recover Human Ability*, ed. MM Merzenich, M Nahum, TM Van Vleet, 207:209–21. Amsterdam: Elsevier

Tierney A, Kraus N. 2014. Auditory-motor entrainment and phonological skills: precise auditory timing hypothesis (PATH). *Front. Hum. Neurosci.* 8:949

Tierney AT, Kraus N. 2013c. The ability to tap to a beat relates to cognitive, linguistic, and perceptual skills. *Brain Lang.* 124:225–31

Tierney A, Kraus N. 2015. Evidence for multiple rhythmic skills. *PLOS ONE.* In press

Tierney AT, Krizman J, Kraus N. 2015. Music training alters the course of adolescent auditory development. *PNAS* 112:10062–67

Turk A, Shattuck-Hufnagel S. 2014. Timing in talking: What is it used for, and how is it controlled? *Philos. Trans. R. Soc. B* 369:20130395

Tzounopoulos T, Kraus N. 2009. Learning to encode timing: mechanisms of plasticity in the auditory brainstem. *Neuron* 62:463–69

Ulanovsky N, Las L, Farkas D, Nelken I. 2004. Multiple time scales of adaptation in auditory cortex neurons. *J. Neurosci.* 24:10440–53

Ullman MT. 2001. A neurocognitive perspective on language: the declarative/procedural model. *Nat. Rev. Neurosci.* 2:717–26

Vuust P, Brattico E, Seppänen M, Näätänen R, Tervaniemi M. 2012. Practiced musical style shapes auditory skills. *Ann. N. Y. Acad. Sci.* 1252:139–46

Watkins KE, Strafella AP, Paus T. 2003. Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia* 41:989–94

Wen B, Wang GI, Dean I, Delgutte B. 2009. Dynamic range adaptation to sound level statistics in the auditory nerve. *J. Neurosci.* 29:13797–808

White-Schwoch T, Woodruff Carr K, Anderson S, Strait DL, Kraus N. 2013. Older adults benefit from music training early in life: biological evidence for long-term training-driven plasticity. *J. Neurosci.* 33:17667–74

White-Schwoch T, Woodruff Carr K, Thompson EC, Anderson S, Nicol T, et al. 2015. Auditory processing in noise: a preschool biomarker for literacy. *PLOS Biol.* 13:e1002196

Wilson SM, Saygin AP, Sereno MI, Iacoboni M. 2004. Listening to speech activates motor areas involved in speech production. *Nat. Neurosci.* 7:701–2

Wittgenstein L. 1953. *Philosophical Investigations*. Oxford, UK: Blackwell

Woodruff Carr K, White-Schwoch T, Tierney AT, Strait DL, Kraus N. 2014. Beat synchronization predicts neural speech encoding and reading readiness in preschoolers. *PNAS* 111:14559–64

Zarco W, Merchant H, Prado L, Mendez JC. 2009. Subsecond timing in primates: comparison of interval production between human subjects and rhesus monkeys. *J. Neurophysiol.* 102:3191–202

Zatorre RJ, Chen JL, Penhune VB. 2007. When the brain plays music: auditory–motor interactions in music perception and production. *Nat. Rev. Neurosci.* 8:547–58

Zendel BR, Tremblay CD, Belleville S, Peretz I. 2015. The impact of musicianship on the cortical mechanisms related to separating speech from background noise. *J. Cogn. Neurosci.* 27:1044–59

ANNUAL REVIEWS
Connect With Our Experts
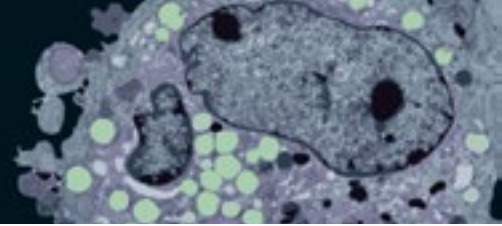
New From Annual Reviews:

Annual Review of Cancer Biology
cancerbio.annualreviews.org · Volume 1 · March 2017

ONLINE NOW!

Co-Editors: Tyler Jacks, *Massachusetts Institute of Technology*
Charles L. Sawyers, *Memorial Sloan Kettering Cancer Center*

The *Annual Review of Cancer Biology* reviews a range of subjects representing important and emerging areas in the field of cancer research. The *Annual Review of Cancer Biology* includes three broad themes: Cancer Cell Biology, Tumorigenesis and Cancer Progression, and Translational Cancer Science.

ANNUAL REVIEWS | CONNECT WITH OUR EXPERTS

650.493.4400/800.523.8635 (US/CAN)
www.annualreviews.org | service@annualreviews.org

Contents