Musical Rhythm, Linguistic Rhythm, and Human Evolution

—

Aniruddh D. Patel
The Neurosciences Institute

THERE IS NOW A VIGOROUS debate over the evolutionary status of music. Some scholars argue that humans have been shaped by evolution to be musical, while others maintain that musical abilities have not been a target of natural selection but reflect an alternative use of more adaptive cognitive skills. One way to address this debate is to break music cognition into its underlying components and determine whether any of these are innate, specific to music, and unique to humans. Taking this approach, Justus and Hutsler (2005) and McDermott and Hauser (2005) suggest that musical pitch perception can be explained without invoking natural selection for music. However, they leave the issue of musical rhythm largely unexplored. This comment extends their conceptual approach to musical rhythm and suggests how issues of innateness, domain specificity, and human specificity might be addressed.

Key words: musical rhythm, linguistic rhythm, synchronization, basal ganglia, evolution

RECENT YEARS HAVE SEEN renewed interest in the idea that human minds have been shaped by natural selection for music (e.g., Mithen, 2005; Wallin, Merker, & Brown, 2000), an idea first proposed by Darwin (1871). Enthusiasm for the idea has spread rapidly, and there are a growing number of hypotheses about the possible adaptive roles of music in human evolution (see Fitch, in press, for one review). Some thinkers remain skeptical, however, regarding music as an enjoyable mental technology built from preexisting cognitive skills (e.g., Pinker, 1997). These skeptics echo the sentiment of William James, who said that love of music is "a mere incidental peculiarity of the nervous system, with no teleological significance" (cited in Langer, 1942, p. 210).

How can this debate be resolved? One approach, advocated in two recent articles in this journal (Justus & Hutsler, 2005; McDermott & Hauser, 2005), is to determine whether there are fundamental aspects of music cognition which are innate and which cannot be explained as byproducts or secondary uses of more clearly adaptive cognitive abilities such as auditory scene analysis or language. Demonstrating the existence of such aspects would favor adaptationist arguments. Without this demonstration, there is no reason to reject the null hypothesis that human minds have not been specifically shaped by natural selection for music.

This is a useful approach because it links evolutionary studies of music to empirical research, specifically on issues of the innateness, domain-specificity, and human-specificity of musical abilities. Reviewing what is known about these issues, the authors conclude that at present there is no compelling reason to reject the null hypothesis mentioned above. Both articles, however, focus almost exclusively on the perception of musical pitch. The purpose of this comment is to consider how the approach advocated by these papers could be applied to musical rhythm.

Is Musical Rhythm an Offshoot of Linguistic Rhythm?

From an evolutionary perspective, an obvious question about musical rhythm is its relationship to speech rhythm, since music and language both feature rich rhythmic organization (Jackendoff & Lerdahl, in press; Patel & Daniele, 2003). One area of overlap concerns perceptual grouping, the mental clustering of events into units (e.g., phrases) at different hierarchical levels. Grouping in music and speech shows many similarities. Music and language mark group boundaries in similar ways using pitch movements and durational lengthening, and cross-domain sensitivity to these grouping cues starts early in life (Jusczyk & Krumhansl, 1993). There is also evidence from neuropsychology and neuroimaging that the perception of grouping in the two domains uses similar brain substrates (Knösche et al., 2005; Patel, Peretz, Tramo, & Labrecque, 1998). Thus grouping in music may well be an offshoot of prosodic grouping abilities.

Turning from grouping to meter, the story is quite different. In every culture there is some form of music with a regular beat, a periodic pulse that affords temporal coordination between performers and elicits a synchronized motor response from listeners (Nettl, 2000). Although early theories of speech rhythm proposed an underlying isochronous pulse based on stresses or syllables (Abercrombie, 1967; Pike, 1945), empirical data have not supported this idea (Bertinetto, 1989; Dauer, 1983), and contemporary studies of speech rhythm have largely abandoned the isochrony issue (e.g., Grabe & Low, 2002; Ramus, Nespor, & Mehler, 1999).

A musical beat typically occurs in the context of a meter, a hierarchical organization of beats in which some beats are perceived as stronger than others. Interestingly, speech also has a "metrical" hierarchy based on stress or prominence (Selkirk, 1984; Terken & Hermes, 2000), suggesting that a tendency to organize rhythmic sequences in terms of hierarchical prominence patterns may originate in language. Crucially, however, the "beats" of speech (stressed syllables) do not mark out a regular pulse. This difference has important cognitive consequences. In particular, the use of a perceptually isochronous pulse in music engages periodic temporal expectancies that play a basic role in music cognition (Jones 1976; Jones & Boltz, 1989), but which appear to play little or no role in ordinary speech perception (cf. Pitt & Samuel, 1990). Humans are able to extract periodicities from complex auditory stimuli, and can focus their expectancies on periodicities at different hierarchical levels in music (Drake, Baruch, & Jones, 2000). These periodic expectancies are the basis of motor synchronization to the beat on the part of listeners, as shown by the fact that listeners typically tap or move slightly *ahead* of the actual beat, indicating that synchronization is based on structured temporal anticipation.

Beat perception and synchronization (or BPS) is an aspect of rhythm which appears to be unique to music. This aspect of musical rhythm cannot be explained as a byproduct of linguistic rhythm, and thus merits attention in evolutionary studies. Indeed, hypotheses about the adaptive value of BPS have been offered in evolutionary theorizing about human music (e.g., Merker, 2000). Taking the approach to music evolution advocated by Justus & Hutsler (2005) and McDermott & Hauser (2005), the key questions about BPS concern its innateness, its domain specificity, and its human specificity. The remainder of this essay discusses strategies for addressing these issues, with the main focus being on human specificity.

Beat Perception and Synchronization: Innateness

Young infants do not synchronize their movements to a musical beat (Longhi, 2003). However, this is not evidence against innateness: Young infants also do not speak, even though there are good reasons to believe that humans are biologically predisposed to acquire speech. Thus one way to address the innateness of BPS is via developmental studies, in order to explore whether the brain seems specifically prepared to acquire this ability. At present we lack basic developmental information BPS, including the earliest age at which children reliably synchronize to a beat, and what percent of (musically untrained) children and adults attain this ability. Synchronization to a beat is attractive for behavioral study because it requires only gross motor skills (e.g., clapping, tapping, or bobbing up and down), yet has received relatively little developmental research (see Drake et al., 2000, and McAuley, Jones, Holub, Johnston, & Miller, 2006, for two relevant studies). This is an area where more developmental work is warranted, examining how innate predispositions and experience interact to produce BPS (cf. Hannon & Trehub, 2005; Phillips-Silver & Trainor, 2005; Repp, 2005).

Beat Perception and Synchronization: Domain-Specificity

One way to study the domain-specificity of BPS is to determine if brain damage which disrupts it also disrupts other nonmusical cognitive abilities. The neuropsychological literature contains descriptions of individuals with musical rhythmic disturbance after brain damage, or "acquired arrhythmia" (e.g., Di Pietro, Laganaro, Leemann, & Schnider, 2003; Fries & Swihart, 1990; Liégeois-Chauvel, Peretz, Babaï, Laguitton, & Chauvel, 1998; Mavlov, 1980; Peretz, 1990; Schuppert, Münte, Wieringa, & Altenmüller, 2000; Wilson, Pressing, & Wales, 2002). Two notable findings from this literature are that rhythmic abilities can be selectively disrupted, leaving pitch processing skills relatively intact, and that there are dissociations between rhythmic tasks requiring simple discrimination of temporal patterns and those requiring the evaluation or production of periodic patterns. However, no neuropsychological studies to date have examined relations between deficits in BPS and in other basic cognitive skills. If such relations can be found, this would suggest that BPS is based on abilities recruited from other brain functions.

Beat Perception and Synchronization: Human-Specificity

McDermott and Hauser (2005) argue that nonhuman animals (henceforth, animals) do not naturally produce music. Hence if an animal can acquire an ability which is part of human music, then this would suggest that the ability is not part of an adaptation for music. Applying this reasoning to BPS, the question arises whether animals can learn to synchronize to a musical beat. (Note that BPS differs in important ways from the synchronized displays of certain animals such as frogs, crickets, fireflies, etc. See Gerhardt & Huber, 2002, ch.8, and Patel, Iversen, Chen, & Repp, 2005, for further discussion).

It is a remarkable fact that despite decades of research in psychology and neuroscience in which animals have been trained to do elaborate tasks, there is not a single report of an animal being trained to tap, peck, or move in synchrony with an auditory beat. One might object that such a behavior is unnatural for an animal, but this misses the point. Monkeys, for example, are often trained to do highly ecologically unnatural tasks in neuroscience experiments (such as tracing ellipses in the air) for the purpose of research on neural mechanisms of perception or motor control. Thus the relevant question is whether an animal *could* learn BPS. If so, this would indicate that natural selection for music is not necessary to account for BPS.

A question which immediately arises is which animals one should study. Chimps and bonobos may seem the obvious choice. Among the great apes they are the most closely related to humans. They are also highly intelligent, as evidenced by research with language-trained apes such as Kanzi (Savage-Rumbaugh, Shanker, & Taylor, 1998). Furthermore, chimps and bonobos produce short bouts of rhythmic 'drumming' with their hands or feet as part of display or play behavior (Arcadi, Robert, & Boesch, 1998; Fitch, in press; Kugler & Savage-Rumbaugh, 2002), meaning that they can voluntarily produce rhythmic movements on a time scale appropriate for BPS.

Despite these facts, there are reasons to question whether apes (and nonhuman primates in general) are capable of BPS. These reasons pertain to the brain circuits that are involved in beat perception and motor control. Perceptual research on humans using fMRI indicates that rhythms that do (vs. do not) have a regular beat are associated with increased activity in the basal ganglia (Grahn, 2004). This deep-brain structure is known to be an essential part of the distributed circuit (involving the cerebral cortex, basal ganglia, and thalamus) involved in interval timing, that is, in gauging temporal intervals in the time range relevant to musical beat perception (Matell & Meck, 2000). Importantly, the basal ganglia are also involved in motor control and sequencing (cf. Janata & Grafton, 2003), meaning that a brain structure involved in perceptually "keeping the beat" is also involved in the coordination of patterned movement.

If BPS simply required that a common brain structure be involved in interval timing and motor control, then one would expect that chimps (and many other animals) would be capable of BPS. This is because the basal ganglia subserve interval timing and motor control functions across a wide range of species, including primates and rodents (Buhusi & Meck, 2005). However, I suspect BPS requires more than just a common brain structure that handles both of these functions. This is because BPS involves a special relationship between *auditory* temporal intervals and patterned movement, as evidenced by the fact that visual rhythms poorly induce BPS in humans (Patel et al., 2005). Yet the interval timing abilities of the basal ganglia are amodal, applying equally well to intervals defined by auditory vs. visual events. This suggests that some additional force in human evolution modified the basal ganglia in a way that affords a tight coupling between auditory input and motor output.

One plausible candidate for this evolutionary force is vocal learning. Vocal learning involves learning to produce vocal signals based on auditory experience and sensory feedback. This ability seems commonplace to us, since every child exhibits it as part of learning to speak. An evolutionary perspective, however, reveals that vocal learning is an uncommon trait, having arisen in only a few groups of animals (including songbirds, parrots, cetaceans, and some pinnipeds; cf. Fitch, in press; Merker, 2005). Notably, humans are unique among primates in exhibiting complex vocal learning (Egnor & Hauser, 2004).

Vocal learning requires a tight coupling between auditory input and motor output in order to match vocal production to a desired model. This online integration of the auditory and motor system places special demands on the nervous system. Neurobiological research on birds indicates that vocal learning is associated with modifications to the basal ganglia, which play a key role in mediating a link between auditory input and motor output during learning (Doupe, Perkel, Reiner, & Stern, 2005). Since there are many anatomical parallels between basal ganglia anatomy in birds and mammals, it seems plausible to suggest that human

basal ganglia have also been modified by natural selection for vocal learning (cf. Jarvis, 2004). The resulting tight coupling between auditory input and motor output may be a necessary foundation for BPS.

In the spirit of comparative research advocated by McDermott and Hauser, the foregoing observations can be condensed into a specific and testable hypothesis, namely that having the neural circuitry for complex vocal learning is a necessary prerequisite for the ability to synchronize with an auditory beat. This "vocal learning and rhythmic synchronization hypothesis" predicts that attempts to teach nonhuman primates to synchronize to a beat will not be successful. Furthermore, it suggests that if primates do fail at BPS it would be premature to conclude that BPS is unique to humans.

Determining whether nonhuman vocal learners can acquire BPS will be an essential part of probing the human-specificity of musical abilities.

Author Note

Address correspondence to: Aniruddh D. Patel, The Neurosciences Institute, 10640 John Jay Hopkins Dr., San Diego, CA 92121. E-MAIL *apatel@nsi.edu*

References

ABERCROMBIE, D. (1967). *Elements of general phonetics.* Chicago: Aldine Publishing Company.

ARCADI, A. C., ROBERT, D., & BOESCH, C. (1998). Buttress drumming by wild chimpanzees: Temporal patterning, phrase integration into loud calls, and preliminary evidence for individual distinctiveness. *Primates, 39,* 503–516.

BERTINETTO, P. (1989). Reflections on the dichotomy "stress" vs. "syllable-timing." *Revue de Phonétique Appliquée,* 91–93, 99–130.

BUHUSI, C. V., & MECK, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews, Neuroscience, 6,* 755–765.

DARWIN, C. (1871). *The descent of man, and selection in relation to sex.* London: John Murray.

DAUER, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics, 11,* 51–62.

DI PIETRO, M., LAGANARO, M., LEEMANN, B., & SCHNIDER, A. (2003). Amusia: Selective rhythm processing following left temporoparietal lesion in a professional musician with conduction aphasia. *Brain and Language, 87,* 152–153.

DOUPE, A. J., PERKEL, D. J., REINER, A., & STERN, E. A. (2005). Birdbrains could teach basal ganglia research a new song. *Trends in Neurosciences, 28,* 353–363.

DRAKE, C., BARUCH, C., & JONES, M. R. (2000). The development of rhythmic attending in auditory sequences: Attunement, reference period, focal attending. *Cognition, 77,* 251–288.

EGNOR, S. E. R., & HAUSER, M. D. (2004). A paradox in the evolution of primate vocal learning. *Trends in Neurosciences, 27,* 649–654.

FITCH, W. T. (in press). The biology and evolution of music: A comparative perspective. *Cognition.*

FRIES, W., & SWIHART, A. A. (1990). Disturbance of rhythm sense following right hemisphere damage. *Neuropsychologia, 28,* 1317–1323.

GEISSMANN, T. (2000). Gibbon song and human music from an evolutionary perspective. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 103–123). Cambridge, MA: MIT Press.

GERHARDT, B., & HUBER, F. (2002). *Acoustic communication in insects and anurans.* Chicago: University of Chicago Press.

GRABE, E., & LOW, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. In C. Gussenhoven & N. Warner (Eds.), *Laboratory phonology 7* (pp. 515–546). Berlin: Mouton de Gruyter.

GRAHN, J. (2004). *Behavioural and functional imaging studies of rhythm processing.* Unpublished doctoral dissertation, University of Cambridge, UK.

HANNON, E. E., & TREHUB, S. E. (2005). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences of the USA, 102,* 12639–12643.

JACKENDOFF, R., & LERDAHL, F. (in press). The capacity for music: What is it, and what's special about it? *Cognition.*

JANATA, P., & GRAFTON, S. T. (2003). Swinging in the brain: Shared neural substrates for behaviors related to sequencing and music. *Nature Neuroscience, 6,* 682–687.

JARVIS, E. D. (2004). Learned birdsong and the neurobiology of human language. *Annals of the New York Academy of Sciences, 1016,* 749–777.

JONES, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review, 83,* 323–355.

JONES M. R., & BOLTZ, M. (1989). Dynamic attending and responses to time. *Psychological Review, 96,* 459–491.

JUSCZYK, P., & KRUMHANSL, C. (1993). Pitch and rhythmic patterns affecting infants' sensitivity to musical phrase structure. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 627–640.

JUSTUS, T., & HUTSLER, J. J. (2005). Fundamental issues in the evolutionary psychology of music: Assessing innateness and domain-specificity. *Music Perception, 23,* 1–27.

KNÖSCHE, T. R., NEUHAUS, C., HAUEISEN, J., ALTER, K., MAESS, B., WITTE, O. W., et al. (2005). The perception of phrase structure in music. *Human Brain Mapping, 24,* 259–273.

KUGLER, K., & SAVAGE-RUMBAUGH, S. (2002, June). Rhythmic drumming by Kanzi an adult male bonobo (*Pan paniscus*) at the language research center. Paper presented at the 25th meeting of the American Society of Primatologists, Oklahoma University, OK.

LANGER, S. (1942). *Philosophy in a new key: A study in the symbolism of reason, right, and art.* Cambridge, MA: Harvard University Press.

LIÉGEOIS-CHAUVEL, C., PERETZ, I., BABAÏ, M., LAGUITTON, V., & CHAUVEL, P. (1998). Contribution of different cortical areas in the temporal lobes to music processing. *Brain, 121,* 1853–1867.

LONGHI, E. (2003). *The temporal structure of mother-infant interactions in musical contexts.* Unpublished doctoral dissertation, University of Edinburgh.

MATELL, M. S., & MECK, W. H. (2000). Neuropsychological mechanisms of interval timing behaviour. *BioEssays, 22,* 94–103.

MAVLOV, L. (1980). Amusia due to rhythm agnosia in a musician with left hemisphere damage: A non-auditory supramodal defect. *Cortex, 16,* 331–338.

MCAULEY, J. D., JONES, M. R., HOLUB, S., JOHNSTON, H., & MILLER, N. S. (2006). The time of our lives: Lifespan development of timing and event tracking. *Journal of Experimental Psychology: General, 135,* 348–367.

MCDERMOTT, J., & HAUSER, M. D. (2005). The origins of music: Innateness, development, and evolution. *Music Perception, 23,* 29–59.

MERKER, B. (2000). Synchronous chorusing and human origins. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 315–327). Cambridge, MA: MIT Press.

MERKER, B. (2005). The conformal motive in birdsong, music, and language: An introduction. *Annals of the New York Academy of Sciences, 1060,* 17–28.

MITHEN, S. (2005). *The singing Neanderthals: The origins of music, language, mind and body.* London: Weidenfeld & Nicolson.

NETTL, B. (2000). An ethnomusicologist contemplates universals in musical sound and musical culture. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 463–472). Cambridge, MA: MIT Press.

PATEL, A. D., & DANIELE, J. R. (2003). An empirical comparison of rhythm in language and music. *Cognition, 87,* B35–B45.

PATEL, A. D., IVERSEN, J. R., CHEN, Y., & REPP, B. H. (2005). The influence of metricality and modality on synchronization with a beat. *Experimental Brain Research, 163,* 226–238.

PATEL, A. D., PERETZ, I., TRAMO, M., & LABRECQUE, R. (1998). Processing prosodic and musical patterns: A neuropsychological investigation. *Brain and Language, 61,* 123–144.

PERETZ, I. (1990). Processing of local and global musical information by unilateral brain-damaged patients. *Brain, 113,* 1185–1205.

PHILLIPS-SILVER, J., & TRAINOR, L. J. (2005). Feeling the beat in music: Movement influences rhythm perception in infants. *Science, 308,* 1430.

PIKE, K. N. (1945). *The intonation of American English.* Ann Arbor: University of Michigan Press.

PINKER, S. (1997). *How the mind works.* New York: Norton.

PITT, M. A., & SAMUEL, A. G. (1990). The use of rhythm in attending to speech. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 564–573.

RAMUS, F., NESPOR, M., & MEHLER, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition, 73,* 265–292.

REPP, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review, 12,* 969–992.

SAVAGE-RUMBAUGH, S., SHANKER, S. G., & TAYLOR, T. J. (1998). *Apes, language, and the human mind.* New York: Oxford University Press.

SCHUPPERT, M., MÜNTE, T. F., WIERINGA, B. M., & ALTENMÜLLER, E. (2000). Receptive amusia: Evidence for cross-hemispheric neural networks underlying music processing strategies. *Brain, 123,* 546–559.

SELKIRK, E. O. (1984). *Phonology and syntax: The relation between sound and structure.* Cambridge, MA: MIT Press.

TERKEN, J., & HERMES, D. (2000). The perception of prosodic prominence. In M. Horne (Ed.), *Prosody: Theory and experiment* (pp. 89–127). Dordrecht: Kluwer.

WALLIN, N. L., MERKER, B., & BROWN, S. (Eds.). (2000). *The origins of music.* Cambridge, MA: MIT Press.

WILSON, S. J., PRESSING, J. L., & WALES, R. J. (2002). Modelling rhythmic function in a musician post-stroke. *Neuropsychologia, 40,* 1494–1505.