

The Neural Correlates of Temporal Structure in Music

Daniel J. Levitin, PhD¹

In this article, the author reviews studies on the neuroanatomical underpinnings of *temporal structure in music*. Music's temporal structure has been studied using all the tools available to modern cognitive neuroscience, including genetic models, neuroimaging, mathematical models, and lesion studies. Regions of the brain that attend to and track ongoing temporal structure in music are bilateral, unlike those that subserve temporal structure in speech, which are predominantly left-lateralized. Musical structure processing appears to recruit networks in the prefrontal

cortex, inferior frontal cortex, superior temporal poles, and cerebellum (including the cerebellar vermis). When the music is emotionally meaningful, activity extends to the ventral tagmental area, the nucleus accumbens, and the hypothalamus. The author concludes with a discussion of what remains to be addressed in this rapidly growing field as well as the challenges to be faced.

Keywords: arts medicine; music medicine; neurologic music therapy; rhythmicity

This article begins with a review of what is meant by the phrase *temporal structure in music*. A brief recounting of some major findings on its neural correlates, chiefly as found in neuroimaging studies, follows. The article concludes with a discussion of what I believe still remains to be addressed in this rapidly growing field and, finally, what challenges still face us.

Structure in music is that which exists when one can differentiate an ordered sequence from a random sequence of musical events (Garner, 1974; Lerdaahl, 2001; Levitin & Menon, 2005; Patel, 2003; Pomerantz & Lockhead, 1991). One of the primary functions of the brain is to extract structure from the physical world. This is a difficult task, as the information that impinges on our sensory receptors (eardrums, retinas) is disorderly and chaotic, often a dense, undifferentiated stream of impulses (e.g., sound waves, photons) that our brains must organize into meaningful units. In listening to a symphony, for example, the sound waves that hit the eardrum

do not themselves differentiate an oboe from the air conditioner in the concert hall—this is the task of our brain, a process dubbed *auditory scene analysis* (Bregman, 1990); despite a great amount of effort on the part of computer scientists to instantiate such behavior in programs, no computer can do what all of us accomplish effortlessly.

Visual art is manifest over space. In contrast, the meaning of music (or similarly, poetry or speech) is manifest over time. Structure thus exists in music over time, in the relationship between successive notes of a melody or chords in a progression (the *horizontal* aspect of musical structure). Structure can be said to also exist in the simultaneously (or near-simultaneously) sounded notes that make up a chord (the *vertical* aspect of musical structure). Both the formation of chords and their sequence are described (or governed) by music theory.

Music's temporal structure has been studied using all the tools available to modern cognitive neuroscience, including genetic models (as in studies of Williams Syndrome; Levitin & Bellugi, 1998), neuroimaging (Janata et al., 2002; Knosche et al., 2005; Levitin & Menon, 2003), mathematical models (Krumhansl, 1990; Shepard, 1982), and lesion studies (Peretz & Hébert, 1995; Peretz & Zatorre, 2003). We have learned that regions of the

¹McGill University, Montreal, Canada

Address correspondence to: Daniel J. Levitin, Department of Psychology, McGill University, Montreal, QC H3A 1B1; e-mail: daniel.levitin@mcgill.ca.

Date received: February 22, 2009; accepted: May 01, 2009.

brain that attend to and track ongoing temporal structure in music are bilateral, unlike those that subservise temporal structure in speech, which are predominantly left-lateralized. Musical structure processing appears to recruit networks in the prefrontal cortex, inferior frontal cortex, superior temporal poles, and cerebellum (including the cerebellar vermis). When the music is emotionally meaningful, activity extends to the ventral tagmental area, the nucleus accumbens, and the hypothalamus (Menon & Levitin, 2005).

Determining the exact processes that occur in the frontal regions has been difficult, and the findings are ambiguous. The frontal lobes are involved in mediating access to the mental lexicon in the case of language, although the musical equivalent has not yet been discovered. The prefrontal cortex also mediates access to the hippocampus and access to long-term and working memory for music. Brodmann Areas 44, 46, and 47 (bilaterally) have been broadly implicated in operations relating to musical syntax, tonal expectancies, and semantic retrieval (Koelsch et al., 2002; Koelsch et al., 2004; Koelsch, Maess, & Friederici, 2000; Patel, Gibson, Ratner, Besson, & Holcomb, 1998). It has been argued that the experimental paradigms that gave rise to these findings all have a property in common that might best be thought of as *temporal structure* and, further, that the involvement of Brodmann 47 may reflect its role in temporal perceptual organization; that is, it may instantiate a processing center for building up temporal expectations and determining (in the Gestalt sense) “what goes with what” (Levitin & Menon, 2005).

In spite of the great progress over the past decade, there remain many gaps in our understanding of musical structure and the brain. We still do not understand how music can cause such strong emotional reactions, and this is clearly related to its structure, not simply the overall spectrum or timbres of sounds—a randomized or “scrambled” sequence of notes is not able to elicit the same reactions as an ordered one (Lerdahl & Jackendoff, 1983; Levitin & Menon, 2005).

The relation between music and language appears to be a complex one (Besson & Friederici, 2005), and this presents yet another challenge for the coming decade. The emerging evidence from neuroimaging and lesion studies is that music recruits many of the same neural pathways as language does and yet also recruits distinct music-specific circuits.

The increased spatial resolution afforded by the newer, high-power fMRI machines (5 teslas and more) may help to substantially improve our understanding of processing specificity within subregions of what had previously been regarded as functionally homogeneous tissue.

Time, tempo, and structure in music appear to involve a neural network that links the cerebellum, parietal, frontal, and temporal lobes (Levitin, 2006), but the directionality and control sequence is poorly understood. As of this writing, a debate exists about the primacy of the basal ganglia versus the cerebellum in keeping the internal clock in humans that is crucial to understanding temporal events (Bhattacharjee, 2006). The historical view has been that the cerebellum contains such a time-keeper through a system of hierarchically arranged oscillators; increasing evidence suggests that this role is played by the basal ganglia. In approaching this debate, it may help to consider that the human cerebellum comprises both neo- and paleocerebellar regions, reflecting the phylogenetic evolution of this structure. The function that the cerebellum performs in modern humans may well be different than that in lower mammals. Moreover, it is possible that a basal ganglia–cerebellum circuit allows the cerebellum to control or mediate basal ganglia function when the cerebellum is intact and that when damaged these functions become decoupled. Again, better spatial resolution in fMRI along with new and more advanced analysis tools, including functional data analysis (Ramsay & Silverman, 1997, 2002; Vines, Nuzzo, & Levitin, 2005) and connectivity analysis (Bokde, Tagamets, Friedman, & Horwitz, 2001; Friston, 2005; Greicius, Krasnow, Reiss, & Menon, 2003; Menon & Levitin, 2005), will be important tools in the coming decade to properly address these issues. Transcranial Direct Current Stimulation may also be an important adjunct to current methods (Vines, Nair, & Schlaug, 2006; Vines, Schneider, & Schlaug, 2006), allowing experimental control over distinct neural subregions in real time.

Recent work has shown that visual input influences the perception of auditory temporal structure (Vines, Krumhansl, Wanderley, & Levitin, 2006; Vitouch, Sovdat, & Höller, 2006), and this is an area that is rapidly growing, evidenced by the formation of a new journal, *Gesture*, the aim of which is to better understand the role of paralinguistic and paramusical gestures that are perceived visually as an adjunct to interpreting the acoustical stream.

The basis for *vision* influencing music may derive from the observation that music cannot exist without movement and an evolutionary history during which sound and sight were coupled. Because sound is transmitted via vibrating molecules, some physical action is required to set those molecules in motion in the first place (Levitin, McAdams, & Adams, 2002; Levitin & Tirovolas, 2009). Even when lying perfectly still, listeners show activation in those regions of the brain that are involved in motor action planning, including the cerebellum, basal ganglia, and cortical motor areas (Levitin, 2006; Levitin & Menon, 2003). Tapping in synchrony to the pulse of a musical sequence (by humans) engages the presupplementary motor area, the supplemental motor area, the dorsal premotor cortex, the dorsolateral prefrontal cortex, the inferior parietal lobule, and lobule VI of the cerebellum, as measured by the blood oxygenation level dependent (BOLD) response (Chen, Penhune & Zatorre, 2008).

Additional connections have been noted between auditory and kinesthetic senses (Phillips-Silver & Trainor, 2005, 2007, 2008). In one study, adult and infant participants bounced to a duple (march) or triple (waltz) meter while listening to an unaccented, ambiguous rhythm. The kinesthetically-felt meter biased the perceptual representation and subsequent recall of the sequences. In effect, the movement itself created the (cross-modal) accented beat. This interactive process was found to be mediated by the vestibular system: Although full-body movement appears to be the most effective in engendering the movement–sound interaction, head movement alone is capable of producing it, although body movement alone is not (Phillips-Silver & Trainor, 2008). Additional data implicate the dorsal premotor cortex (dPMC) in rhythmic synchronization. Participants tapped to rhythmic sequences of varying levels of metric difficulty; greater difficulty was correlated with increased dPMC activation (Chen et al., 2008).

The way in which auditory meaning unfolds over time is described by rules that we call *syntax* in language and *harmony* in music. Linguists differentiate different aspects of the speech stream through the separate study of phonetics, phonology, morphology, syntax, semantics, and pragmatics. Attempts to find strictly analogous units in the study of music have been unsuccessful (Lerdahl & Jackendoff, 1983), yet many scientific researchers still refer to the syntactic structure of music as a way of drawing out the parallels between music and language (Gjerdingen, 1992;

Koelsch et al., 2004; Koelsch et al., 2000; Levitin & Menon, 2003; Maess, Koelsch, Gunter, & Friederici, 2001; Patel, 2003). In language, the syntax of an utterance—the relation between words—can be studied on its own; as soon as we introduce a dictionary we are at the level of semantics, and as soon as we introduce a listener we are at the level of pragmatics. Syntax and semantics in music appear to be far less separable than they are in speech. No one has yet identified the equivalent to *parts of speech*, or *actors* and *agents* in music. Even attempts to distinguish “legal” or “well-formed” musical phrases from “illegal” or “malformed” phrases necessarily involve not just the structure of the musical phrase with respect to tonality and harmony but also its meaning and interpretation. An emerging and compelling idea is that the proper level at which to analyze temporal structure in music may in fact be the pragmatic level (Ashley, 2006). This is because we can rarely divorce notions of consonance and dissonance, or tension and resolution, from a discussion of musical structure, and these concepts are cognitive, necessarily involving a listener.

Pitch structure in melodies has been shown to be characterized by the $1/f$ or $1/f^2$ distribution (Gilden, Thornton, & Mallon, 1995; Patel & Balaban, 2000; Voss & Clarke, 1975), whereby the power spectrum is proportional to the reciprocal of frequency. A major unsolved problem is the extent to which rhythmic sequences in music can be characterized by a $1/f$ distribution.

We stand at the beginning of a new age in understanding human behavior and the brain. The study of music can provide an important window into brain function because it is a complex faculty that involves virtually every cognitive process we have identified—attention, memory, categorization, motor action planning, prediction, emotion, and communication. The study of music in psychology and neuroscience also enjoys a distinguished history, dating back to Aristotle and Pythagoras, and including Wundt, Fechner, Helmholtz, and James. Although relegated to the fringe of experimental psychology for much of the 20th century, recent decades have seen exponential growth in research on the psychology of music (Levitin, 1999). The coming 10 to 20 years promise to solve many questions that have remained unanswered for centuries, and I believe that the temporal structure of music is at the core of the mystery of what makes music special to us and how it is that music can move us so emotionally.

Understanding this core issue in cognition may help us to better understand communication in the broadest sense and, in the process, to help us better understand ourselves as a species.

Acknowledgments

The author is grateful to the National Sciences and Engineering Research Council of Canada (NSERC) and to the GRAMMY Foundation for support during the preparation of this article.

Declaration of Conflicting Interests

The author has declared that there are no conflicts of interests in the authorship and publication of this contribution.

References

- Ashley, R. (2006, August). *Musical communication and musical comprehension: A perspective from pragmatics*. Paper presented at the 9th International Conference on Music Perception and Cognition, Bologna, Italy.
- Besson, M., & Friederici, A. D. (2005). Part II: Language and music—A comparison. Introduction. *Annals of the New York Academy Science*, 1060, 57-58.
- Bhattacharjee, Y. (2006). A timely debate about the brain. *Science*, 311(5761), 596-598.
- Bokde, A. L., Tagamets, M.-A., Friedman, R. B., & Horwitz, B. (2001). Functional interactions of the inferior frontal cortex during the processing of words and word-like stimuli. *Neuron*, 30(2), 609-617.
- Bregman, A. S. (1990). *Auditory scene analysis*. Cambridge, MA: MIT Press.
- Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Moving on time: Brain network for auditory-motor synchronization is modulated by rhythm complexity Band musical training. *Journal of Cognitive Neuroscience*, 20, 226-239.
- Friston, K. J. (2005). Models of brain function in neuroimaging. *Annual Review of Psychology*, 56, 57-87.
- Garner, W. R. (1974). *The processing of information and structure*. Potomac, MD: Erlbaum.
- Gilden, D. L., Thornton, T., & Mallon, M. W. (1995, March 24). 1/f noise in human cognition. *Science*, 267, 1837-1839.
- Gjerdingen, R. O. (1992). Learning syntactically significant temporal patterns of chords—A masking field embedded in an Art-3 architecture. *Neural Networks*, 5, 551-564.
- Greicius, M. D., Krasnow, B. D., Reiss, A. L., & Menon, V. (2003). Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences*, 100, 253-258.
- Janata, P., Birk, J. L., Horn, J. D. V., Leman, M., Tillmann, B., & Bharucha, J. J. (2002). The cortical topography of tonal structures underlying Western music. *Science*, 298, 2167-2170.
- Knosche, T. R., Neuhaus, C., Haueisen, J., Alter, K., Maess, B., Witte, O., et al. (2005). Perception of phrase structure in music. *Human Brain Mapping*, 24, 259-273.
- Koelsch, S., Gunter, T. C., Cramon, D. Y. v., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach speaks: A cortical “language-network” serves the processing of music. *NeuroImage*, 17, 956-966.
- Koelsch, S., Kasper, E., Sammler, d., Schulze, K., Gunter, T., & Friederici, A. D. (2004). Music, language and meaning: brain signatures of semantic processing. *Nature Neuroscience*, 7, 302-307.
- Koelsch, S., Maess, B., & Friederici, A. D. (2000). Musical syntax is processed in the area of Broca: An MEG study. *NeuroImage*, 11(5, Suppl. 1), S56.
- Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.
- Lerdahl, F. (2001). *Tonal pitch space*. New York: Oxford University Press.
- Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.
- Levitin, D. J. (1999). Review of D. Deutsch, *The Psychology of Music*, Second Edition [Book review]. *Music Perception*, 16(4), 495-506.
- Levitin, D. J. (2006). *This is your brain on music: The science of a human obsession*. New York: Dutton.
- Levitin, D. J., & Bellugi, U. (1998). Musical abilities in individuals with Williams Syndrome. *Music Perception*, 15, 357-389.
- Levitin, D. J., McAdams, S., & Adams, R. L. (2002). Control parameters for musical instruments: A foundation for new mappings of gesture to sound. *Organised Sound*, 7, 171-189.
- Levitin, D. J., & Menon, V. (2003). Musical structure is processed in “language” areas of the brain: A possible role for Brodmann Area 47 in temporal coherence. *NeuroImage*, 20, 2142-2152.
- Levitin, D. J., & Menon, V. (2005). The neural locus of temporal structure and expectancies in music: Evidence from functional neuroimaging at 3 Tesla. *Music Perception*, 22, 563-575.
- Levitin, D. J., & Tirovolas, A. K. (2009). Current advances in the cognitive neuroscience of music. *The Year in Cognitive Neuroscience 2009: Annals of the New York Academy of Sciences*, 1156, 211-231.
- Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca’s area: An MEG study. *Nature Neuroscience*, 4, 540-545.

- Menon, V., & Levitin, D. J. (2005). The rewards of music listening: Response and physiological connectivity of the mesolimbic system. *NeuroImage*, 28, 175-184.
- Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, 6, 674-681.
- Patel, A. D., & Balaban, E. (2000). Temporal patterns of human cortical activity reflect tone sequence structure. *Nature*, 404, 80-84.
- Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related study. *Journal of Cognitive Neuroscience*, 10, 717-733.
- Peretz, I., & Hébert, S. (1995). Music processing after brain damage: The case of rhythm without melody. In Steinberg, R. (Ed.), *Music and the mind machine*. Berlin: Springer-Verlag.
- Peretz, I., & Zatorre, R. J. (Eds.). (2003). *The cognitive neuroscience of music*. New York: Oxford University Press.
- Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science*, 308, 1430.
- Phillips-Silver, J., & Trainor, L. J. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition*, 105, 533-546.
- Phillips-Silver, J., & Trainor, L. J. (2008). Vestibular influence on auditory metrical interpretation. *Brain and Cognition*, 67, 94-102.
- Pomerantz, J. R., & Lockhead, G. R. (1991). Perception of structure: An overview. In Lockhead, G. R., & Pomerantz, J. R. (Eds.), *The perception of structure* (pp. 1-22). Washington, DC: American Psychological Association.
- Ramsay, J. O., & Silverman, B. W. (1997). *Functional data analysis*. New York: Springer-Verlag.
- Ramsay, J. O., & Silverman, B. W. (2002). *Applied functional data analysis*. New York: Springer-Verlag.
- Shepard, R. N. (1982). Structural representations of musical pitch. In Deutsch, D. (Ed.), *Psychology of music* (pp. 343-390). San Diego: Academic Press.
- Vines, B. W., Krumhansl, C. L., Wanderley, M. M., & Levitin, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition*, 101, 80-113.
- Vines, B. W., Nair, D. G., & Schlaug, G. (2006). Contralateral and ipsilateral transcranial direct current stimulation. *NeuroReport*, 17, 671-674.
- Vines, B. W., Nuzzo, R. L., & Levitin, D. J. (2005). Analyzing musical dynamics: Quantifying musical emotion with differential calculus and functional data objects. *Music Perception*, 23, 139-154.
- Vines, B. W., Schnider, N. M., & Schlaug, G. (2006). Testing for causality with tDCS: pitch memory and the left supra-marginal gyrus. *NeuroReport*, 17, 1047-1050.
- Vitouch, O., Sovdat, S., & Höller, N. (2006, August 22-26). *Audio-vision: Visual input drives perceived music tempo*. Paper presented at the 9th International Conference on Music Perception and Cognition, Bologna, Italy.
- Voss, R. F., & Clarke, J. (1975). Music from 1/f noise. *Journal of the Acoustical Society of America*, 63.

Daniel J. Levitin, PhD, is James McGill Professor of Psychology and Neuroscience in the Department of Psychology at McGill University in Montreal, Quebec, Canada.