Why musical hierarchies?

Courtney B. Hilton*, Rie Asano*, Cedric Boeckx
***** Co-first authors

To appear in *Behavioral and Brain Sciences*, as a commentary on *Origins of music in credible signaling,* by Mehr, S., Krasnow, M., Bryant, G., & Hagen, E. (https://psyarxiv.com/nrqb3/)

Abstract: Credible signaling may have provided a selection pressure for producing and discriminating increasingly elaborate proto-musical signals. But why evolve them to have hierarchical structure? We argue that the hierarchicality of tonality and meter is a byproduct of domain-general mechanisms evolved for reasons other than credible signaling.

The target article by Mehr and colleagues provides a welcome critique of prevailing evolutionary theories of music while also advancing their own credible signaling proposal. We find many aspects of this promising. However, while adaptations for rhythm and melody seem plausible, we take issue with the claim that credible signaling resulted in a "grammar-like, combinatorially generative interface" based on the "hierarchical organization of meter and tonality".

Coalition signaling provides plausible reasons to evolve the capacity to produce and discriminate rhythmically coordinated displays. This is supported in the cited data on birds (Hall & Magrath, 2007; Tobias et al., 2016) and primates (Geissmann, 2000). But it does not, as far as we can see, provide reasons to evolve hierarchical means of doing so. And indeed, these data only show evidence of rhythmic coordination in terms of temporal precision or synchronization and provide no evidence for or against hierarchy.

Likewise, parent-infant signaling provides evolutionary reasons for melodic signals and infant sensitivity to them. Comparison to animal contact calls (Bouchet et al., 2013; Leighton, 2017) and data on genomic imprinting disorders in humans (Mehr et al., 2017) supports these claims. But here too: why hierarchies? For the purpose of signaling attention to an infant, or for contact calls more generally, hierarchical organization poses no obvious advantage. There is also limited evidence for contact calls being hierarchically organized. Moreover, while some brain areas show differential responses to tonal structure from birth (Perani et al., 2010), behavioral sensitivity only begins to manifest at around 4 years of age before continuing to develop into the teenage years (Brandt et al., 2012; Corrigall & Trainor, 2014).

Taken together, while the hierarchical properties of meter and tonality are a design feature of the musical capacity, their presence is not so clearly motivated by credible signaling.

Hierarchies, however, are not unique to music. They are found in other cognitive domains such as language (Chomsky, 1957), vision (Bill et al., 2020), metacognition (Frith, 2012), and action planning (Miller et al., 1960). In non-human primates, they are found in social learning (Byrne & Russon, 1998) and tool use (Byrne et al., 2013; Greenfield, 1991). Musical hierarchicality may therefore be better conceived as using generic mechanisms evolved for reasons other than as a specific adaptation for credible signaling.

In our previous work, we have argued that the hierarchicality of both musical and linguistic structure derives from mechanisms originally evolved for action planning (Asano & Boeckx, 2015; see also: Jackendoff, 2009; Fitch & Martins, 2014). The inspiration for much of this thinking was Karl Lashley's (1951) prescient insight that complex actions generally, and those for music and language specifically, control their sequential manifestation through hierarchical plans. Doing so, he argued, was necessary for flexibility and robustness, especially for more complex and abstractly motivated actions in which the limitations of control by linear associative chaining are laid bare.

The primary neurocognitive mechanism underlying this capacity is hierarchical cognitive control and comprises a combination of executive functions (maintenance, selection, inhibition). Maintenance is subserved by prefrontal areas (together with their parietal connections) and selection and inhibition by the basal ganglia. The orchestration of these functional areas through a number of distinct cortico-basal ganglia-thalamacortical circuits enables complex and flexible behavior (Badre & Nee, 2018). Consistent with Lashley's insight, these neural circuits are not only implicated in action planning but also for processing musical and linguistic hierarchies (Asano et al., n.d.; Fitch & Martins, 2014; Jeon et al., 2014; Slevc & Okada, 2015).

Functional explanations of behavior are essential for understanding biological evolution. But based on these alone, the determination of how they are translated into mechanisms is too underconstrained. Are new mechanisms evolved de novo? Or are existing ones tweaked and put to new use? And then how may these and other proximate mechanisms in turn constrain the space of ultimate reasons that guides selection in a reciprocal cycle (Laland et al., 2011)? As Tinbergen (1963) suggested, the biological study of behavior (and cognitive systems, in the current paper) should give equal attention to each of four questions: mechanism, ontogeny, phylogeny, and function. Each provides unique constraints whose combined consilience is the basis for robust theory.

One notable "so what?" of all this for the target article is that adaptations for credible signaling may also have implications for language. According to our proposal, the structural complexity of both music and language partly derives from generic hierarchical cognitive control mechanisms that interface with auditory and motor systems. Compared to non-human primates, humans have substantially greater white-matter connectivity both within the hierarchical control circuits and through the dorsal auditory pathway that links motor, auditory, and parietal areas with the prefrontal cortex (Barrett et al., 2020; Rilling et al., 2008). Adaptations for producing and perceiving rhythmically coordinated audio-motor displays and for fine-scale vocal control of pitch conceivably include an expansion of this shared connectome (Merchant & Honing, 2014; Patel & Iversen, 2014), thus entangling the evolution of both domains. This would also be consistent with claims about music-to-language transfer effects more generally in ontogeny (Patel, 2011; Zatorre, 2013).

To conclude, the credible signaling proposal of Mehr and colleagues is commendable. But we suggest that it can be further improved by considering interactions of proximate and ultimate causes, and specifically how this may clarify the origins of musical hierarchies.

References

Asano, R., & Boeckx, C. (2015). Syntax in language and music: What is the right level of comparison? *Frontiers in Psychology*, *6*.

Asano, R., Boeckx, C., & Seifert, U. (n.d.). Hierarchical control as a shared neurocognitive mechanism for language and music. *(Submitted)*.

Badre, D., & Nee, D. E. (2018). Frontal Cortex and the Hierarchical Control of Behavior. *Trends in Cognitive Sciences*, *22*(2), 170–188.

Barrett, R. L. C., Dawson, M., Dyrby, T. B., Krug, K., Ptito, M., D'Arceuil, H., Croxson, P. L., Johnson, P. J., Howells, H., Forkel, S. J., Dell'Acqua, F., & Catani, M. (2020). Differences in Frontal Network Anatomy Across Primate Species. *The Journal of Neuroscience*, *40*(10), 2094–2107.

Bill, J., Pailian, H., Gershman, S. J., & Drugowitsch, J. (2020). Hierarchical structure is employed by humans during visual motion perception. *Proceedings of the National Academy of Sciences*, 9.

Bouchet, H., Blois-Heulin, C., & Lemasson, A. (2013). Social complexity parallels vocal complexity: A comparison of three non-human primate species. *Frontiers in Psychology*, *4*.

Brandt, A., Gebrian, M., & Slevc, L. R. (2012). Music and Early Language Acquisition. *Frontiers in Psychology*, *3*.

Byrne, R. W., & Russon, A. E. (1998). Learning by imitation: A hierarchical approach. *Behavioral and Brain Sciences*, *21*(5), 667–684.

Byrne, R. W., Sanz, C. M., & Morgan, D. B. (2013). Chimpanzees plan their tool use. In C. Sanz, J. Call, & C. Boesch (Eds.), *Tool Use in Animals* (pp. 48–64). Cambridge University Press.

Chomsky, N. (1957). *Syntactic Structures*. Walter de Gruyter.

Corrigall, K. A., & Trainor, L. J. (2014). Enculturation to musical pitch structure in young children: Evidence from behavioral and electrophysiological methods. *Developmental Science*, *17*(1), 142–158.

Fitch, W. T., & Martins, M. D. (2014). Hierarchical processing in music, language, and action: Lashley revisited. *Annals of the New York Academy of Sciences*, *1316*(1), 87–104.

Frith, C. D. (2012). The role of metacognition in human social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1599), 2213–2223.

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

Greenfield, P. (1991). Language, tools and the brain: The ontogeny and phylogeny of hierarchically organized sequential behavior. *Behavioral and Brain Sciences*, *14*(4), 570–571.

Hall, M. L., & Magrath, R. D. (2007). Temporal coordination signals coalition quality. *Current Biology*, *17*(11), R406–R407.

Jackendoff, R. (2009). Parallels and Nonparallels between Language and Music. *Music Perception*, *26*(3), 195–204.

Jeon, H.-A., Anwander, A., & Friederici, A. D. (2014). Functional Network Mirrored in the Prefrontal Cortex, Caudate Nucleus, and Thalamus: High-Resolution Functional Imaging and Structural Connectivity. *Journal of Neuroscience*, *34*(28), 9202–9212.

Laland, K. N., Sterelny, K., Odling-Smee, J., Hoppitt, W., & Uller, T. (2011). Cause and Effect in Biology Revisited: Is Mayr's Proximate-Ultimate Dichotomy Still Useful? *Science*, *334*(6062), 1512–1516.

Lashley, K. S. (1951). The Problem of Serial Order in Behavior. In *Cerebral Mechanisms in Behavior* (p. 36).

Leighton, G. M. (2017). Cooperative breeding influences the number and type of vocalizations in avian lineages. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1868), 20171508.

Mehr, S. A., Kotler, J., Howard, R. M., Haig, D., & Krasnow, M. M. (2017). Genomic Imprinting Is Implicated in the Psychology of Music. *Psychological Science*, 13.

Merchant, H., & Honing, H. (2014). Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. *Frontiers in Neuroscience*, *7*.

Miller, G., Galanter, E., & Pibram, K. (1960). *Plans and Structure of Behavior*. Henry Holt and Company.

Patel, A. D. (2011). Why would Musical Training Benefit the Neural Encoding of Speech? The OPERA Hypothesis. *Frontiers in Psychology*, *2*.

Patel, A. D., & Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: The Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in Systems Neuroscience*, *8*, 57.

Perani, D., Saccuman, M. C., Scifo, P., Spada, D., Andreolli, G., Rovelli, R., Baldoli, C., & Koelsch, S. (2010). Functional specializations for music processing in the human newborn brain. *Proceedings of the National Academy of Sciences*, *107*(10), 4758–4763.

Rilling, J. K., Glasser, M. F., Preuss, T. M., Ma, X., Zhao, T., Hu, X., & Behrens, T. E. J. (2008). The evolution of the arcuate fasciculus revealed with comparative DTI. *Nature Neuroscience*, *11*(4), 426–428.

Slevc, L. R., & Okada, B. M. (2015). Processing structure in language and music: A case for shared reliance on cognitive control. *Psychonomic Bulletin & Review*, *22*(3), 637–652.

Tinbergen, N. (1963). On aims and methods of Ethology. *Zeitschrift Für Tierpsychologie*, *20*, 410–433.

Tobias, J. A., Sheard, C., Seddon, N., Meade, A., Cotton, A. J., & Nakagawa, S. (2016). Territoriality, Social Bonds, and the Evolution of Communal Signaling in Birds. *Frontiers in Ecology and Evolution*, *4*.

Zatorre, R. J. (2013). Predispositions and Plasticity in Music and Speech Learning: Neural Correlates and Implications. *Science*, *342*(6158), 585–589.