

## THREE QUESTIONS CONCERNING CONSONANCE PERCEPTION

---

PETER M. C. HARRISON

*Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany*

I DISCUSS THREE FUNDAMENTAL QUESTIONS underpinning the study of consonance: 1) What features cause a particular chord to be perceived as consonant? 2) How did humans evolve the ability to perceive these features? 3) Why did humans evolve to attribute particular aesthetic valences to these features (if they did at all)? The first question has been addressed by several recent articles, including Friedman, Kowalewski, Vuvan, and Neill (2021), with the common conclusion that consonance in Western listeners is driven by multiple features such as harmonicity, interference between partials, and familiarity. On this basis, it seems relatively straightforward to answer the second question: each of these consonance features seems to be grounded in fundamental aspects of human auditory perception, such as auditory scene analysis and auditory long-term memory. However, the third question is harder to resolve. I describe several potential answers, and argue that the present evidence is insufficient to distinguish between them, despite what has been claimed in the literature. I conclude by discussing what kinds of future studies might be able to shed light on this problem.

*Received: November 30, 2020, accepted December 10, 2020.*

**Key words:** consonance, evolution

---

MUSIC PSYCHOLOGISTS HAVE LONG BEEN interested in consonance, the sense in which certain combinations of tones seem to fit well together while other combinations seem to fit poorly. The psychological study of consonance is underpinned by three fundamental research questions: 1) What features cause a particular chord to be perceived as consonant? 2) How did humans evolve the ability to perceive these features? 3) Why did humans evolve to attribute particular aesthetic valences to these features (if they did at all)?

Several recent studies, including Friedman, Kowalewski, Vuvan, and Neill (2021), study the first question using regression modeling (Eerola & Lahdelma, 2020; Harrison & Pearce, 2020; Smit, Milne, Dean, & Weidemann, 2019). The common conclusion of these studies is that consonance in Western listeners is driven by multiple features (though see McDermott, Lehr, & Oxenham, 2010, for contradicting claims). All studies identify distinct contributions of harmonicity and interference between partials, and several additionally identify contributions of familiarity as operationalized in various ways (Eerola & Lahdelma, 2020; Harrison & Pearce, 2020; Smit et al., 2019).

The second question seems also to have a straightforward answer: each feature seems to be grounded in fundamental aspects of human auditory perception. Harmonicity is a key feature underpinning fusion processes in auditory scene analysis that is particularly useful for understanding vocalizations (e.g., McPherson et al., 2020). Interference between partials produces characteristic amplitude modulations (“beating”) that propagate through the auditory system (e.g., Vencovský, 2016). Familiarity is meanwhile a simple byproduct of auditory long-term memory.

The third question is less tractable. The literature suggests several speculative answers:

*Auditory scene analysis.* Humans evolved to like harmonicity and dislike interference because acoustic environments with these properties are easy to understand and hence safer for the organism (Huron, 2001).

*Vocal similarity.* Humans evolved to like harmonicity and to dislike interference between partials because this encouraged them to attend carefully to vocalizations, which carry important information about conspecifics (Bowling, Purves, & Gill, 2018).

*Social bonding.* Humans evolved to prefer sounds with high harmonicity and low interference because this preference encouraged collaborative music making, which in turn had adaptive functions of fostering social bonding (Savage et al., 2020).

These three positive hypotheses may be contrasted with the following null hypothesis:

*The null hypothesis (H0).* Humans never evolved any particular aesthetic predispositions towards harmonicity or interference. The Western preference for harmonicity and dislike of interference is simply due to cultural convention (McDermott, Schultz, Undurraga, & Godoy, 2016).<sup>1</sup>

New behavioral data must be evaluated in the light of all competing hypotheses. Now, Friedman et al. (2021) write that their results provide support for the vocal similarity hypothesis, because the vocal similarity hypothesis successfully predicts that humans will prefer chords that exhibit high harmonicity and low interference. However, this conclusion is problematic, as their results are also consistent with each of the other hypotheses as detailed above. Bowling et al. (2018) similarly make a problematic conclusion in favor of the vocal similarity hypothesis without properly considering the alternative hypotheses.

In particular, before making any conclusions about Q3, it is essential to establish that humans are naturally predisposed towards harmonicity and against interference, rather than these preferences being solely culturally learned. McDermott et al. (2016) present important negative evidence here, showing that the Tsimane' people do not share Western preferences for harmonicity or for the kinds of chords that Westerners perceive as consonant; likewise, the existence of beat diaphony traditions (Ambrazevičius, 2017; Florian, 1981; Vassilakis, 2005; Vyčiniënė, 2002) suggests that listeners in some musical cultures can develop inverse preference patterns to Western listeners. However, a complete solution to this question requires large-scale studies examining preferences cross-culturally.

Large-scale cross-species studies could meanwhile deliver compelling support for one of the three positive hypotheses. If harmonicity liking and interference disliking were common to most animals with well-developed auditory scene analysis capacities, this would support the auditory scene analysis hypothesis. If these phenomena were shown to be mostly specific to species who make pitched vocalizations, this would

support the vocal similarity hypothesis. If the phenomena were shown to be specific to species who make synchronized vocalizations, this would support the social bonding hypothesis. Unfortunately, the existing research does not point to any of these outcomes: as of yet, there is no compelling evidence that any non-human animals perceive consonance in the valenced way that Western listeners do (see Toro & Crespo-Bojorque, 2017, for a review and Harrison & Pearce, 2020, for a discussion).

Large-scale cross-cultural and cross-species studies present significant logistical challenges. In the meanwhile, some progress could be made in disentangling the hypotheses by examining additional chord features that might predict consonance judgments. For example, if consonance judgments were shown to be meaningfully influenced by additional acoustic features specific to auditory scene analysis, this would provide some support for the auditory scene analysis hypothesis. Wright and Bregman (1987) present interesting arguments in this direction that still lack comprehensive empirical investigation (though see Bonin, Trainor, Belyk, & Andrews, 2016, and Kowalewski, Friedman, Zavoyskiy, & Neill, 2019). Alternatively, if consonance judgments were shown to be meaningfully influenced by additional features specific to vocalizations, this would support the vocal similarity hypothesis. One must acknowledge, however, that this work cannot really determine whether preferences for these features truly evolved, rather than being wholly culturally determined: for this we still need cross-cultural studies.

On the one hand, this is a negative conclusion: properly understanding the evolutionary bases of consonance will require logistically complex cross-cultural or cross-species studies rather than comparatively straightforward studies with Western-enculturated listeners (e.g., Bowling et al., 2018; Friedman et al., 2021). On the other hand, the current growth of interest in large-scale cross-cultural studies provides good reasons to be optimistic for the future (e.g., Mehr et al., 2019; Savage, Brown, Sakai, & Currie, 2015). In the meanwhile, we must treat these kinds of evolutionary questions with caution.

#### Author Note

*Correspondence concerning this article should be addressed to Peter M. C. Harrison, Max-Planck-Institut für empirische Ästhetik, Grüneburgweg 14, 60322 Frankfurt am Main, Germany. E-mail: peter.harrison@ae.mpg.de*

<sup>1</sup>In fact, McDermott et al. (2016) state a slightly different version of this hypothesis. They claim that only harmonicity, not interference, contributes to consonance; furthermore, they do not claim that aversion to interference is not universal, only that liking of harmonicity is not universal.

## References

- AMBRAZEVIČIUS, R. (2017). Dissonance/roughness and tonality perception in Lithuanian traditional Schwebungsdiaphonie. *Journal of Interdisciplinary Music Studies*, 8, 39–53. <https://doi.org/10.4407/jims.2016.12.002>
- BONIN, T. L., TRAINOR, L. J., BELYK, M., & ANDREWS, P. W. (2016). The source dilemma hypothesis: Perceptual uncertainty contributes to musical emotion. *Cognition*, 154, 174–181. <https://doi.org/10.1016/j.cognition.2016.05.021>
- BOWLING, D. L., PURVES, D., & GILL, K. Z. (2018). Vocal similarity predicts the relative attraction of musical chords. *Proceedings of the National Academy of Sciences*, 115(1), 216–221. <https://doi.org/10.1073/pnas.1713206115>
- EEROLA, T., & LAHDELMA, I. (2020). The anatomy of consonance/dissonance: Evaluating acoustic and cultural predictors across multiple datasets with chords. *OSF Preprints*. <https://doi.org/10.31219/osf.io/6aqhx>
- FLORIAN, G. (1981). The two-part vocal style on Baluan Island Manus Province, Papua New Guinea. *Ethnomusicology*, 25, 433–446.
- FRIEDMAN, R. S., KOWALEWSKI, D. A., VUVAN, D. T., & NEILL, W. T. (2021). Consonance preferences within an unconventional tuning system. *Music Perception*, 38(3), 313–330.
- HARRISON, P. M. C., & PEARCE, M. T. (2020). Simultaneous consonance in music perception and composition. *Psychological Review*, 127, 216–244. <https://doi.org/10.1037/rev0000169>
- HURON, D. (2001). Tone and voice: A derivation of the rules of voice-leading from perceptual principles. *Music Perception*, 19, 1–64. <https://doi.org/10.1525/mp.2001.19.1.1>
- KOWALEWSKI, D. A., FRIEDMAN, R. S., ZAVOYSKIY, S., & NEILL, W. T. (2019). A reinvestigation of the Source Dilemma Hypothesis. *Music Perception*, 36, 448–456. <https://doi.org/10.1525/mp.2019.36.5.448>
- MCDERMOTT, J. H., LEHR, A. J., & OXENHAM, A. J. (2010). Individual differences reveal the basis of consonance. *Current Biology*, 20, 1035–1041. <https://doi.org/10.1016/j.cub.2010.04.019>
- MCDERMOTT, J. H., SCHULTZ, A. F., UNDURRAGA, E. A., & GODOY, R. A. (2016). Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature*, 535, 547–550. <https://doi.org/10.1038/nature18635>
- MCPHERSON, M. J., DOLAN, S. E., DURANGO, A., OSSANDON, T., VALDÉS, J., UNDURRAGA, E. A., ET AL. (2020). Perceptual fusion of musical notes by native Amazonians suggests universal representations of musical intervals. *Nature Communications*, 11(1), 1–14. <https://doi.org/10.1038/s41467-020-16448-6>
- MEHR, S. A., SINGH, M., KNOX, D., KETTER, D. M., PICKENS-JONES, D., ATWOOD, S., ET AL. (2019). Universality and diversity in human song. *Science*, 366(6468), 1–17. <https://doi.org/10.1126/science.aax0868>
- SAVAGE, P. E., BROWN, S., SAKAI, E., & CURRIE, T. E. (2015). Statistical universals reveal the structures and functions of human music. *Proceedings of the National Academy of Sciences*, 112(29), 8987–8992. <https://doi.org/10.1073/pnas.1414495112>
- SAVAGE, P. E., LOUI, P., TARR, B., SCHACHNER, A., GLOWACKI, L., MITHEN, S., & FITCH, W. T. (2020). Music as a coevolved system for social bonding. *Behavioral and Brain Sciences*, 1–36. <https://doi.org/10.1017/S0140525X20000333>
- SMIT, E. A., MILNE, A. J., DEAN, R. T., & WEIDEMANN, G. (2019). Perception of affect in unfamiliar musical chords. *PLOS ONE*, 14(6). <https://doi.org/10.1371/journal.pone.0218570>
- TORO, J. M., & CRESPO-BOJORQUE, P. (2017). Consonance processing in the absence of relevant experience: Evidence from nonhuman animals. *Comparative Cognition and Behavior Reviews*, 12, 33–44. <https://doi.org/10.3819/CCBR.2017.120004>
- VASSILAKIS, P. N. (2005). Auditory roughness as a means of musical expression. In R. A. Kendall & R. H. Savage (Eds.), *Selected reports in ethnomusicology: Perspectives in systematic musicology* (Vol. 12, pp. 119–144). Los Angeles, CA: Department of Ethnomusicology, University of California.
- VENCOVSKÝ, V. (2016). Roughness prediction based on a model of cochlear hydrodynamics. *Archives of Acoustics*, 41, 189–201. <https://doi.org/10.1515/aoa-2016-0019>
- VYČINIENĖ, D. (2002). Lithuanian Schwebungsdiaphonie and its south and east European parallels. *The World of Music*, 44, 55–57.
- WRIGHT, J. K., & BREGMAN, A. S. (1987). Auditory stream segregation and the control of dissonance in polyphonic music. *Contemporary Music Review*, 2(1), 63–92.