Evidence supporting a link between harmonicity and the attractiveness of simultaneous tone combinations has emerged from an experiment designed to mitigate effects of musical enculturation. I examine the analysis undertaken to produce this evidence and clarify its relation to an account of tonal aesthetics based on the biology of auditory-vocal communication.

Received: December 15, 2020, accepted December 16, 2020.

Key words: Consonance, harmonicity, roughness, vocalization, evolution, aesthetics

In this issue of Music Perception, Friedman, Kowalewski, Vuvan, and Neill (2021) describe novel evidence that tonal stimuli with high spectral similarity to harmonic series’ (harmonicity) are perceived as attractive (consonant). They interpret this evidence in partial support of the biological hypothesis that consonance derives from a generalized affective attraction to voice-like harmonic spectra, canalized in humans by the benefits of auditory-vocal communication (Bowling & Purves, 2015; Terhardt, 1984). Their experimental approach is designed to circumvent the cultural hypothesis that consonance derives instead from lifetime learning, supported by the psychological maxim that objects made familiar become preferred (McDermott, Schultz, Undurraga, & Godoy, 2016; Zajonc, 1968). Specifically, Friedman et al. argue that because cultural familiarity is correlated with vocal similarity in stimuli derived from popular tuning systems like the chromatic scale, a pure test of the relationship between vocal similarity and consonance requires unfamiliar stimuli. Accordingly, Friedman et al. base their work on an obscure musical tuning system, the Bohlen-Pierce (BP) scale.

Friedman et al.’s (2021) results begin by reporting that harmonicity and consonance are not significantly related across all possible two- and three-tone chords derived from the BP scale (“dyads” and “triads”), thus marking a difference with chromatic dyads and triads (cf. Figure 1A & 1B). Rather than ending their investigation here, however, they address the cause of this discrepancy through two additional analyses, both of which show significant relationships between harmonicity and consonance in subsets of BP chords (Figure 1B). Although skeptics may question these analyses as examples of “post hoc” data dredging, I argue that both are better considered as remedies for experimental design flaws, which otherwise confound the aims of Friedman et al.’s study. The first additional analysis excludes BP chords comprising chromatic intervals. BP intervals #2, #11, and #13 closely correspond to an equally tempered minor third, major tenth, and perfect twelfth, differing by less than two cents on average (mean = 1.31¢). When these “chromatic” BP intervals are combined into chords, the results include octave-expanded versions of the chromatic major and minor triads (arguably the two most popular chords of all time; Parncutt, Reisinger, Fuchs, & Kaiser, 2019) among others. Accordingly, the exclusion of chords comprising chromatic BP intervals is justified by Friedman et al.’s aim to circumvent cultural familiarity. The second analysis in question excludes BP chords with very low harmonicity. A full 85% of BP dyads and triads fall below the first harmonic similarity quartile of chromatic dyads and triads (“q1”)², revealing a stark difference in the distribution of harmonicity between BP and chromatic tuning systems (cf. Figure 1C & D). This difference undermines Friedman et al.’s aim to compare the relation between harmonicity and consonance between BP and chromatic tuning systems, which appear in their totality here as apples and oranges. Excluding BP chords with exceptionally low harmonicity renders a more direct comparison, also in accord with predetermined aims.

A further aspect of Friedman et al.’s (2021) analysis leverages incon—an R software package instantiating models of consonance (Harrison & Pearce, 2020)—to examine BP consonance in relation to auditory

1 Differential pitch sensitivity at the center frequency of Friedman et al.’s chords (middle C) has been estimated at ~10-20¢ (~2-3 Hz; Shower & Biddulph, 1931; Zwicker, Flottorp, & Stevens, 1957).

2 “Harmonic similarity” is a simple measure of harmonicity corresponding to the degree to which a chord’s spectrum represents a single harmonic series defined by its greatest common divisor (Bowling, Purves, & Gill, 2018; Gill & Purves, 2009).
roughness. Although I agree with Friedman et al. (and Harrison & Pearce, 2020) that consonance may be best understood as a composite of these factors (Terhardt [1984] traces a similar idea to Helmholtz, 1877/1954), any claim that roughness outweighs harmonicity made on the basis of comparing regression coefficients should be treated with skepticism. The reason is that roughness and harmonicity are strongly correlated in tonal stimuli (Figure 2; see Graham, 2003). An alternative approach in a composite model of consonance would be to differentiate the ranges of harmonicity over which each factor maximally operates, with harmonicity driving our attraction to chords at moderate and high levels, and roughness driving our aversion to chords at low levels. This composite view has the benefit of accounting for variance in the consonance of chords with “virtual” pitches below the range of typical phonation⁵.

A final point concerns VST and ongoing tension between biological and cultural accounts of consonance. The value of VST lies in providing a framework for understanding human tone perception in biological

---

⁵The average greatest common divisor frequency (gcd_f0) of chords below q1 is 31 Hz in Figure 1A, and only 4 Hz in Figure 1B. Some of these chords may have additional relevant virtual pitches above gcd_f0 a result that is not explained by consonance based purely on similarity to harmonic vocal spectra. Furthermore, because auditory roughness in an important feature of mammalian vocalizations conveying intensely negative affect—e.g., in alarm, distress, and aggression; Arnal, Flinker, Kleinschmidt, Giraud, & Poeppel, 2015; Arnal, Kleinschmidt, Spinelli, Giraud, & Mégévand, 2019; Fitch, Neubauer, & Herzel, 2002; Li et al., 2018)—its integration with harmonicity in a unified biological account of consonance is fully consistent with vocal similarity theory (VST).

A final point concerns VST and ongoing tension between biological and cultural accounts of consonance. The value of VST lies in providing a framework for understanding human tone perception in biological

---

(Friedman et al., 2021).
terms, as it has already done in the context of musical affect (Bowling, Sundararajan, Han, & Purves, 2012; Briefer, 2012; Filippi et al., 2017; Juslin & Laukka, 2003). What theories of consonance rooted in cultural familiarity often obscure is that the affective consequences of repeated exposure provide no explanation of which musical intervals become attractive. This is the object of explanation in a VST account of consonance. Accordingly, cultural learning is not logically opposed to VST. A cultural account of consonance can be made to oppose VST, however, by additionally positing that initial affective susceptibility to tonal stimuli is entirely undifferentiated pre-exposure, making any developed attraction wholly contingent on initial conditions. While this would render cultural familiarity a sound alternative to VST, it is at odds with much musical phenomenology and, arguably, with the neurobiology of affect (Berridge & Kringelbach, 2015; Bowling, Hoeschele, Gill, & Fitch, 2017).

In sum, Friedman et al. (2012) provide compelling evidence that harmonicity and consonance are linked in a way that transcends mere exposure to stimuli. Their paper is an important contribution to understanding how and why we hear tones in the way that we do, with implications for a better understanding of ourselves and the specific ways that music may improve our condition.

**Author Note**

Correspondence concerning this article should be addressed to Daniel L. Bowling, Stanford University, 1201 Welch Rd. MSLS P-106, Mail core 5485, Stanford, CA 94306. E-mail: dbowling@stanford.edu

**References**


---

*The meaning of “pre-exposure” is ill-defined in consonance research. Although “before exposure to music” is often implied, there is no clear line that differentiates what we call music from other forms of tone-based auditory-vocal communication, and the auditory system receives harmonic stimulation from mother’s voice as soon as it comes online, which it has presumably done for hundreds of millions of years (Bowling et al., 2020; Chen & Wiens, 2020; Chiandetti & Vallortigara, 2011; Negus, 1949).*