LETTER

Little evidence for the vocal similarity hypothesis

Jack Goffinet[a,1]

Bowling et al. (1) report musical consonance judgements of various musical chords and a model designed to assess their "vocal similarity," claiming the fit between model and data provides evidence that the perceived consonance of tone combinations is explained in part by their "relative similarity to voiced speech sounds." I believe, however, this claim is unjustified.

The Bowling et al. (1) report states that consonance judgements are tested for "every possible chromatic dyad, triad, and tetrad within a single octave," yet because a just intonation tuning system is employed rather than the standard equal-temperament system, the actual number of possible chords is more than triple the number tested. Confusion between tuning systems also has significant consequences for the model. For example, one dyad is discussed as the tritone, but in the tuning system utilized there are seven distinct tritones within the octave, with modeled consonance ranks ranging widely from 39th to 71st percentiles among dyads.

Analysis of the fit between model and experiment is restricted to pairwise comparisons of chords with significantly different experimental ratings, only 32% of total pairs of experimentally tested chords (3.1% of pairs of possible chords). This inflates the perceived accuracy of the model by excluding pairs with similar ratings. An isotonic regression, which gives an upper bound on the variance in the data explainable by the model, provides a much more informative assessment. Using this method one finds upper bounds on explained variance of mean consonance ratings of 87.8%, 69.3%, and 56.7% for dyads, triads, and tetrads, respectively. The Bowling et al. (1) paper states that, "compared with previous models …, [the vocal similarity model] accords more closely with the available empirical data," although no comparison with alternative models is presented apart from the relative consonance of a single triad compared with the results of a single model. For more perspective, a simple roughness-based model of consonance introduced by William Sethares (2, 3) explains upwards of 84.4%, 77.0%, and 69.5% of variance with a single tuned parameter, outperforming the vocal similarity metric in the two chord categories least likely affected by cultural factors.

The vocal similarity model (1) purports to estimate the relative similarity of musical chords to "conspecific vocalizations," yet only two properties of speech are taken into consideration: that voiced speech is harmonic and that the lower limit of the male vocal range is 50 Hz. An advantage of this approach is that it is easily understood and readily implemented. However, it is unclear how the model performs as a vocal similarity metric, especially considering the importance of formants in predicting biologically relevant speaker information (4, 5). A more straightforward approach to testing the vocal similarity hypothesis would make use of speech data. For example, a probabilistic classifier trained to discriminate speech from other harmonic environmental sounds would provide a natural model of vocal similarity.

The phenomenon of musical consonance has resisted simple explanation for centuries. The vocal similarity hypothesis, although worthy of further investigation, currently offers no exception.

1 Bowling DL, Purves D, Gill KZ (2018) Vocal similarity predicts the relative attraction of musical chords. *Proc Natl Acad Sci USA* 115:216–221.
2 Sethares WA (1993) Local consonance and the relationship between timbre and scale. *J Acoust Soc Am* 94:1218–1228.
3 Sethares WA (2005) *Tuning, Timbre, Spectrum, Scale* (Springer, London).
4 Hillenbrand JM, Clark MJ (2009) The role of f (0) and formant frequencies in distinguishing the voices of men and women. *Atten Percept Psychophys* 71:1150–1166.
5 Smith DR, Patterson RD (2005) The interaction of glottal-pulse rate and vocal-tract length in judgements of speaker size, sex, and age. *J Acoust Soc Am* 118:3177–3186.

[a]Private address, Athens, GA 30605
Author contributions: J.G. wrote the paper.
The author declares no conflict of interest.

[1]Email: jackgoffinet@gmail.com.