Are pitches discrete? An information-theoretic framework and a corpus study

John M. McBride[1], Adam T Tierney[5], Peter Pfordresher[3], Joren Six[4], Shinya Fuji[2], Patrick E Savage[2]

[1]*Insitute for Basic Science, Korea, Republic of (South Korea)*
[2]*Keio University, Japan*
[3]*University at Buffalo, NY, USA*
[4]*Ghent University, Belgium*
[5]*Birkbeck, University of London, UK*
jmmcbride@protonmail.com

Background:

Human melodies are generally thought of as sequences of discrete pitches: discrete *categories*, and also discrete / *level* pitches. Western musical theory does have terms for deviations from discreteness (vibrato, glissandi, scoops), but melodies, when recorded in symbolic notation, are typically defined as a sequence of discrete pitches. Here we ask, is this an appropriate description of music? Are pitches really discrete? Does the degree of discreteness vary across cultures? Is there something special about the perception of human music, or is discreteness common to other domains (bird song, human speech)? These questions are hindered by the fact that pitch-discreteness lacks a rigorous quantitative definition.

Aims:

To establish a rigorous, information-theoretic framework in which we can quantify the degree of discreteness. To validate this framework using human perceptual data, and to use the framework to study the above questions.

Methods:

Our framework models pitch perception as a combination of Bayesian inference and rate-distortion optimization, such that an information-theoretic observer (*e.g.*, a human listening to music) can be defined by two parameters: a time-frequency uncertainty limit, γ, and λ, the ratio of cost of information to the cost of inaccuracy in the pitch percept. Using this framework we study pitch-discreteness in cross-domain and cross-cultural corpora of human song, human speech, and bird song. To find an information-theoretic observer that best approximates humans, we: (i) fit this model to a set of psychophysics experiments (Micheyl et al., 2012); (ii) create an (effectively a transcription) algorithm using this framework that decomposes a fundamental frequency vs time plot into segments which are labelled discrete or non-discrete by the "observer", and then compare the discrete segments to manual transcriptions of human song (Mehr et al., 2019), and to automatic transcriptions by a state-of-the-art algorithm, TONY (Mauch et al., 2015). This allows us to approximate humans to some extent, but more importantly, by fixing γ and λ we can observe different domains / societies in an unbiased way.

For cross-domain comparisons we use ~1,500 vocal songs from the Global Jukebox collection (Wood et al., 2021), 62 bird songs (Tierney et al., 2011) and several large speech corpora (Valk & Alumäe, 2021).

Results:

We show that human pitch perception appears to be consistent with an information-theoretic observer (through perceptual experiments, and transcriptions of melodies). When we compare human music and speech, on average, we find that speech and music contain similar levels of "discreteness"; however, when we control for segment duration, we find that longer segments are considerably more discrete in music (Figure 1). Compared to the differences between music and speech, cross-cultural differences in music are negligible. We need more bird song before drawing strong conclusions.

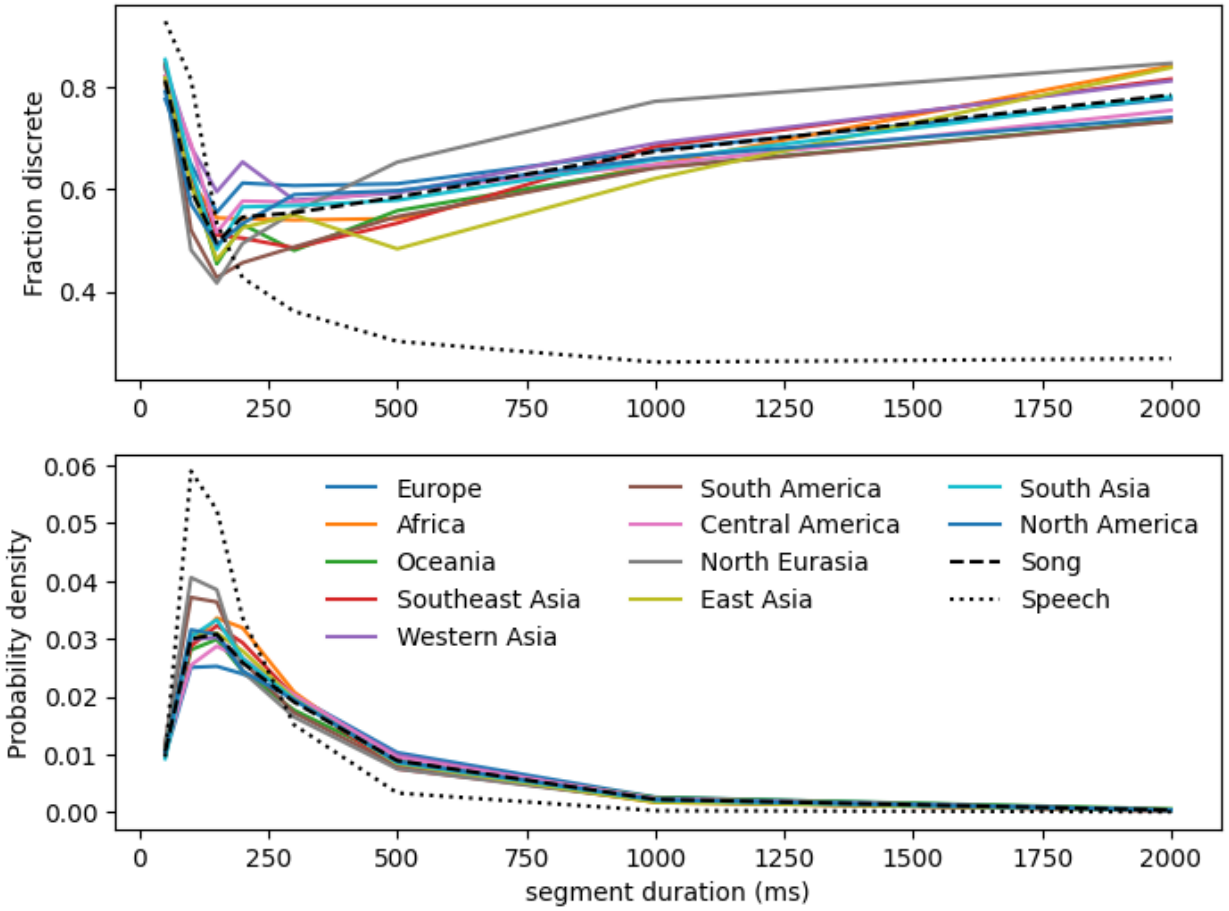

Figure 1. (Top) Fraction of discrete segments as a function of segment duration; (Bottom) segment duration distribution. Music data are shown for each continent group (colored); music (dashed line) and speech data (dotted line) are shown aggregated.

Conclusions:

We have developed a theoretical framework that enables quantitative examination of pitch-discreteness, allowing an unbiased investigation across domains/societies. We find that domains are surprisingly similar in terms of pitch-discreteness, and societies even more so. These results illuminate a hitherto-unexplored, fundamental assumption in human music. Beyond enhancing our basic understanding of music, these results have implications for the development of automated transcription of human music, and they guide the use of interval-analysis in the study of speech and birdsong. Finally, we can rigorously test the proposed framework, since it makes several quantitative predictions (effect of duration, stimuli type, context) that can be tested through perceptual experiments, providing a robust link between perceptual and corpus studies.

References

Mauch, M., Cannam, C., Bittner, R., Fazekas, G., Salamon, J., Dai, J., Bello, J., & Dixon, S. (2015, May). Computer-aided Melody Note Transcription Using the Tony Software: Accuracy and Efficiency. *Proceedings of the First International Conference on Technologies for Music Notation and Representation*.

Mehr, S. A., Singh, M., Knox, D., Ketter, D. M., Pickens-Jones, D., Atwood, S., Lucas, C., Jacoby, N., Egner, A. A., Hopkins, E. J., Howard, R. M., Hartshorne, J. K., Jennings, M. V, Simson, J., Bainbridge, C. M., Pinker, S., O\textquoterightDonnell, T. J., Krasnow, M. M., & Glowacki, L. (2019). Universality and diversity in human song. *Science*, *366*(6468). https://doi.org/10.1126/science.aax0868

Micheyl, C., Xiao, L., & Oxenham, A. J. (2012). Characterizing the dependence of pure-tone frequency difference limens on frequency, duration, and level. *Hearing Research*, *292*(1), 1–13. https://doi.org/https://doi.org/10.1016/j.heares.2012.07.004

Tierney, A. T., Russo, F. A., & Patel, A. D. (2011). The motor origins of human and avian song structure. *Proceedings of the National Academy of Sciences*, *108*(37), 15510–15515.

Valk, J., & Alumäe, T. (2021). Voxlingua107: a dataset for spoken language recognition. *2021 IEEE Spoken Language Technology Workshop (SLT)*, 652–658.

Wood, A., Kirby, K. R., Ember, C., Silbert, S., Daikoku, H., Mcbride, J., Passmore, S., Paulay, F., Flory, M., Szinger, J., & others. (2021). *The Global Jukebox: A public database of performing arts and culture*.

Keywords: pitch-discreteness, melody, cross-cultural analysis, cross-domain analysis, information theory