

# AUTOMATIC ANALYSIS OF GLOBAL MUSIC RECORDINGS SUGGESTS SCALE TUNING UNIVERSALS

Meng-Jou Ho<sup>1</sup> Shoichiro Sato<sup>1</sup> Jiei Kuroyanagi<sup>1</sup> Joren Six<sup>2</sup> Steven Brown<sup>3</sup> Shinya Fujii<sup>1</sup> Patrick Savage<sup>1</sup>

<sup>1</sup>Keio University, <sup>2</sup>Ghent University, <sup>3</sup>McMaster University  
{melodyho, shoichirosato-0107, jieikeio}@keio.jp  
joren.six@ugent.be, stebro@mcmaster.ca  
{sfujii, psavage}@sfc.keio.ac.jp

## ABSTRACT

The structure of musical scales has been proposed to reflect universal acoustic principles based on simple integer ratios. However, some studying tuning in small samples of non-Western cultures have argued that such ratios are not universal but specific to Western music. To address this debate, we applied an algorithm that could automatically analyze and cross-culturally compare scale tunings to a global sample of 50 music recordings, including both instrumental and vocal pieces. Although we found great cross-cultural diversity in most scale degrees, these preliminary results also suggest a strong tendency to include the simplest possible integer ratio within the octave (perfect fifth, 3:2 ratio, ~700 cents) in both Western and non-Western cultures. This suggests that cultural diversity in musical scales is not without limit, but is constrained by universal psycho-acoustic principles that may shed light on the evolution of human music.

## 1. BACKGROUND

Music, like language, is a human universal found in every known culture throughout world history. Although music takes many different forms cross-culturally, scientists have identified numerous “statistical universals” that predominate in most – but not all – of the world’s music [5]. One hypothesis to explain musical universals proposes that the striking cross-cultural convergence of the structure of musical scales (e.g., pentatonic scales; see Fig. 1) may reflect universal acoustic principles based on simple integer ratios, because such ratios maximize the overlap among harmonic frequency spectra and thus sound more consonant [1]. However, cross-cultural analyses of both instrument tunings [2] and perceptions of consonance [3] in small samples of non-Western cultures concluded that no such cross-cultural preferences for simple-integer ratios exist. This debate remains unresolved because of a lack of objective global data on scale tunings. Our goal was to address this lack of data by taking advantage of new algorithms for automatic scale tuning analysis [6]. Based on the integer ratio hypothesis, we predicted that the simplest ratios such as the perfect 5<sup>th</sup> (3:2) and perfect

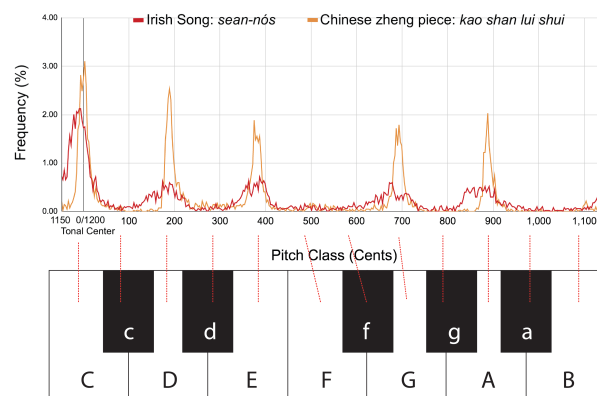
©Meng-Jou Ho, Shoichiro Sato, Jiei Kuroyanagi, Joren Six, Steven Brown, Shinya Fujii, Patrick Savage. Licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). **Attribution:** Meng-Jou Ho, Shoichiro Sato, Jiei Kuroyanagi, Joren Six, Steven Brown, Shinya Fujii, Patrick Savage. “Automatic Analysis of Global Music Recordings Suggests Scale Tuning Universals” Extended Abstracts for the Late-Breaking Demo Session of the 19th International Society for Music Information Retrieval Conference, 2018.

4<sup>th</sup> (4:3) would predominate cross-culturally.

## 2. METHOD

### 2.1. Pitch Class Histogram

We used Tarsos [6] to analyze and compare tunings because, unlike most MIR algorithms that are based on scale models incorporating Western 12-tone equal temperament, Tarsos was designed explicitly for automatic analysis of any music from around the world without imposing such culture-specific theories. Tarsos first extracts the pitch histogram, then collapses this pitch histogram across octaves to create a pitch class histogram, expressed in cents [2] ranging from 0-1200. We used Tarsos’s default YIN pitch estimation algorithm.



**Figure 1.** A traditional Irish song and a Chinese instrumental piece demonstrate similar pentatonic scales.

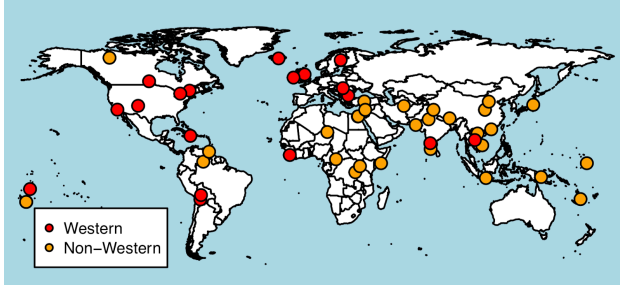
#### 2.1.1 Normalization to the Tonal Center

In order to meaningfully compare scales between different songs that use different absolute pitches, we attempted to normalize each song to a shared tonal center by setting the pitch class of its final note to 0 cents. When the final note is not relevant (e.g., fade outs, excerpts), we instead normalized to the most frequent note (in the future we will explore the effect of these different methods of normalization). Therefore, all scales begin with the first scale degree centered at 0 cents.

### 2.2. Music Samples

For this preliminary analysis, a subset of 50 monophonic pieces from nine regions was selected from the full 304 recordings from *Garland Encyclopedia of World Music* that were previously analyzed manually [5]. 4-6 pieces

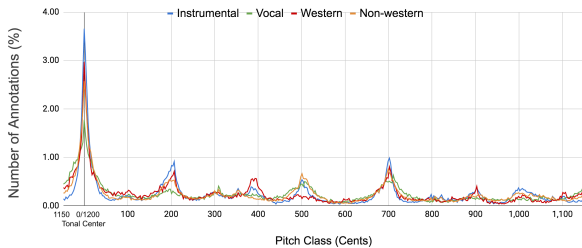
were chosen from each of the nine regions, half instrumental and half vocal. We excluded polyphonic recordings or recordings with loud background noise because automatic pitch estimation cannot yet be performed accurately for such pieces.



**Figure 2.** Map of the 50-recording sample.

### 3. PRELIMINARY RESULTS

Figure 3 shows average tunings for all 50 pieces, separated into “Western” (“Eurogenetic” music performed by or heavily influenced by speakers of European languages) and “non-Western” (all other) music. In the future we will use a larger sample to perform more detailed regional analyses.



**Figure 3.** Average scale tunings in Western and non-Western regions and Vocal and Instrumental instrumentations (see Fig. 2). The x-axis begins and ends at 1150 cents in order to show the distribution around the tonal center (set to 0 cents).

Figure 3 shows the common use of perfect 5ths (~700 cents) and major 2nds (~200 cents) in both Western (red) and non-Western (orange) music. It also shows a clear split in preferences towards major thirds (~400 cents) and perfect 4ths (~500 cents) in Western and non-Western regions, respectively. Additional analyses (not shown) confirm that the interval of approximately 700 cents is consistently preferred in all nine regions, for both vocal and instrumental music (although the range around intervals tends to be less precise for vocal than instrumental music).

### 4. DISCUSSION

Throughout different analyses, the emphasis on the perfect 5th (~700 cents) remains constant. The fact that the simplest possible ratio within the octave (perfect fifth; 3:2 ratios) was the most universal supports the integer ratio hypothesis [1].

In contrast, the fact that the next most simple possible ratio (perfect fourth; 4:3 ratio) was not universal, but the more complex major second (9:8 ratio) was relatively universal, is not clearly explained by the integer ratio hypothesis. We suspect that a theory combining the perception-based integer ratio hypothesis with the pro-

duction-based vocal mistuning hypothesis [3] and the need for step-wise motion using small melodic intervals [5] may better explain these tendencies.

### 5. FUTURE WORK

Currently, our analysis assumes octave equivalence by collapsing pitch histograms into pitch class histograms. However, in the future we could empirically test for the universality of octave equivalence, especially since the integer ratio hypothesis predicts that the octave (2:1 ratio) should be the most universal interval of all. We will also explore the effect of using the final pitch class vs. most frequent pitch class for normalization, and do more fine-grained regional analysis beyond a simple “Western” / “non-Western” dichotomy.

Currently, our analyses requires manual screening of recordings that are not analyzable using Tarsos and manual normalization to a tonal center. Automating this process would allow us to potentially expand this preliminary analysis to include thousands or even millions of recordings. Comparing such automated musical analysis against samples of speech and birdsong should allow us to determine whether aspects of musical scales are specific to human music. Meanwhile, comparison against cross-cultural data for perceptions of consonance should help to reveal the causal mechanisms underlying scale universals.

### 6. ACKNOWLEDGEMENTS

This work was funded by a Keio Research Institute at SFC Startup Grant and a Keio Gijuku Academic Development Fund Individual Grant to P.E.S.

### 7. REFERENCES

- [1] D. L. Bowling, and D. Purves. “A Biological Rationale for Musical Consonance,” *Proceedings of National Academy of Sciences*, vol.112, no.36, 2015.
- [2] A. J. Ellis, “On the Musical Scales of Various Nations,” *Journal of the Society of Arts*, vol.23, no. 1688, pp. 435-527, 1885.
- [3] J. H. McDermott, A. F. Schultz, E. A. Undurraga, and R. A. Godoy, “Indifference to Dissonance in Native Amazonians Reveals Cultural Variation in Music Perception,” *Nature*, vol. 535, pp. 547–550, 2016.
- [4] P. Q. Pfordresher and S. Brown, “Vocal Mistuning Reveals the Origin of Musical Scales,” *J. Cogn. Psychol.*, vol. 29, no. 1, pp. 35–52, 2017.
- [5] P. E. Savage, S. Brown, E. Sakai, & T. E. Currie, “Statistical Universals Reveal the Structures and Functions of Human Music,” *Proceedings of the National Academy of Sciences of the United States of America*, vol.112, no.29, pp. 8987–8992, 2015.
- [6] J. Six, O. Cornelis, and M. Leman, “Tarsos, a Modular Platform for Precise Pitch Analysis of Western and Non-Western Music,” *Journal of New Music Research*, vol.42, no.2, pp. 113-129, 2013.