



## Strong evidence for cross-cultural regularities in music and speech

A recommendation by [Chris Chambers](#)  based on peer reviews by [Bob Slevc](#) and [Nai Ding](#) of the STAGE 2 REPORT:

Yuto Ozaki, Adam Tierney, Peter Pfordresher, John McBride, Emmanouil Benetos, Polina Proutskova, Gakuto Chiba, Fang Liu, Nori Jacoby, Suzanne Purdy, Patricia Opondo, Tecumseh Fitch, Shantala Hegde, Martín Rocamora, Rob Thorne, Florence Ewomazino Nweke, Dhwani Sadaphal, Parimal Sadaphal, Shafagh Hadavi, Shinya Fujii, Sangbuem Choo, Marin Naruse, Utae Ehara, Latyr Sy, Mark Lenini Parselelo, Manuel Anglada-Tort, Niels Chr. Hansen, Felix Haiduk, Ulvhild Færøvik, Violeta Magalhães, Wojciech Krzyżanowski, Olena Shcherbakova, Diana Hereld, Brenda Suyanne Barbosa, Marco Antonio Correa Varela, Mark van Tongeren, Polina Dessiatnitchenko, Su Zar Zar, Iyadh El Kahla, Olcay Muslu, Jakelin Troy, Teona Lomsadze, Dilyana Kurdova, Cristiano Tsope, Daniel Fredriksson, Aleksandar Arabadjiev, Jehoshaphat Philip Sarbah, Adwoa Arhine, Tadhg Ó Meachair, Javier Silva-Zurita, Ignacio Soto-Silva, Neddriel Elcie Muñoz Millalonco, Rytis Ambrazevičius, Psyche Loui, Andrea Ravignani, Yannick Jadoul, Pauline Larrouy-Maestri, Camila Bruder, Tutushamum Puri Teyxokawa, Urise Kuikuro, Rogerdison Natsitsabui, Nerea Bello Sagarzazu, Limor Raviv, Minyu Zeng, Shahaboddin Dabaghi Varnosfaderani, Juan Sebastián Gómez-Cañón, Kayla Kolff, Christina der Nederlanden, Meyha Chhatwal, Ryan Mark David, I Putu Gede Setiawan, Great Lekakul, Vanessa Nina Borsan, Nozuko Nguqu, Patrick E. Savage (2023) Globally, songs and instrumental melodies are slower, higher, and use more stable pitches than speech [Stage 2 Registered Report]. PsyArXiv, ver. 11, peer-reviewed and recommended by Peer Community in Registered Reports.

<https://doi.org/10.31234/osf.io/jr9x7>

Submitted: 16 May 2023, Recommended: 03 July 2023

**Cite this recommendation as:**

Chambers, C. (2023) Strong evidence for cross-cultural regularities in music and speech. *Peer Community in Registered Reports*, 100469. [10.24072/pci.rr.100469](https://doi.org/10.24072/pci.rr.100469)

Published: 03 July 2023

Copyright: This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

---

For centuries, the ubiquity of language and music across human societies has prompted scholars to speculate about their cross-cultural origins as well as their shared and unique characteristics. Depending on the extent

to which contemporary theories emphasise the role of biology vs. culture, a range of hypotheses have been proposed concerning expected similarities and differences in song and speech. One class of hypotheses stemming from cultural relativism assumes a lack of universal regularities in song and speech, and therefore predicts no systematic cross-cultural relationships. On the other hand, more recent evolutionary hypotheses such as the social bonding hypothesis, motor constraint hypothesis, and sexual selection hypothesis all predict differences or similarities in specific characteristic of vocalisations, such as pitch regularity, pitch interval size, and melodic contour. Existing results are mixed in their support of these predictions. In the current study, Ozaki et al. (2023) elucidated cross-cultural similarities and differences between speech and song in 75 different linguistic varieties spanning 21 language families. Understanding precisely how song and speech are related is methodologically challenging due to the multitude of confounds that can arise in comparing natural recordings. Here the authors overcame these difficulties with four types of carefully controlled recordings: singing, recitation of sung lyrics, spoken description of the song, and instrumental version of the sung melody. The authors then examined six features that are amenable to reliable comparison, including pitch height, temporal rate, pitch stability, timbral brightness, pitch interval size, and pitch declination. With this data in hand, the authors asked which acoustic features differ reliably between song and speech across cultures, with the expectation that song would exhibit higher pitch, slower rate and more stable pitch than speech. At the same time, the authors expected song and speech to be reliably similar in the characteristics of timbral brightness, pitch intervals and pitch contours. The findings provided strong support for the preregistered hypotheses. Relative to speech, songs exhibited higher pitch, slower temporal rate, and more stable pitches, while both songs and speech had similar pitch interval size and timbral brightness. Only one hypothesis was unsupported, with the comparison of pitch declination between song and speech turning out inconclusive. To overcome potential sources of analytic bias, the authors undertook additional robustness checks, including reanalysis of a previously published dataset of over 400 song/speech recordings; this exploratory analysis corroborated the conclusions from the confirmatory analysis. Overall this study offers a unique insight into the shared global characteristics of language and music, with implications for understanding their cultural and biological (co)evolution. The Stage 2 manuscript was evaluated over one round of in-depth review. Based on responses to the reviewers' comments, the recommender judged that the manuscript met the Stage 2 criteria and therefore awarded a positive recommendation. **URL to the preregistered Stage 1 protocol:** <https://osf.io/jdhtz> **Level of bias control achieved: Level 2.** *At least some data/evidence that was used to answer the research question had been accessed and partially observed by the authors, but the authors certify that they had not yet observed the key variables within the data that were be used to answer the research question AND they took additional steps to maximise bias control and rigour.*

**List of eligible PCI RR-friendly journals:**

- [Peer Community Journal](#)
- [PeerJ](#)
- [Royal Society Open Science](#)
- [Swiss Psychology Open](#)

**References:**

1. Ozaki, Y. et al. (2023). Globally, songs and instrumental melodies are slower, higher, and use more stable pitches than speech [Stage 2 Registered Report]. Acceptance of Version 11 by Peer Community in Registered Reports. <https://doi.org/10.31234/osf.io/jr9x7>

# Reviews

## Evaluation round #1

DOI or URL of the preprint: [10.31234/osf.io/jr9x7](https://doi.org/10.31234/osf.io/jr9x7)

Version of the preprint: 9

### Authors' reply, 30 June 2023

[Download author's reply](#)

[Download tracked changes file](#)

### Decision by [Chris Chambers](#) , posted 21 June 2023, validated 21 June 2023

#### Minor Revision

I have now received two evaluations of your Stage 2 submission by the reviewers were assessed the proposal at Stage 1. As you can see, both are very positive about the completed work, with one reviewer happy with the submission in its current state and the other raising some minor issues concerning terminology and clarification of the analyses. Please address these points in a revised manuscript and response, and we should be able to award a positive Stage 2 recommendation without further review.

### Reviewed by [Bob Slevc](#), 20 June 2023

I was excited about the first stage of this registered report, and have enjoyed seeing how the data turned out. For the stage 2 review process, I guess the most important question is whether the work was carried out as proposed in the stage 1 manuscript. This certainly seems to be the case! Changes from stage 1 are clearly described (and seem reasonable), the manuscript includes plenty of detail, and the recordings and analyses scripts are not only made openly available but are appropriately documented. I don't think it's really necessary for me to comment on the results and interpretation, but I will just note as an aside that my stage 1 concerns about the specific SESOI chosen appear to have been unfounded – effects are larger than I would have expected! That said, I do appreciate the inclusion of manipulated examples and discussion of how the SESOI was chosen (e.g., in section S7).

Overall, I enjoyed this paper and appreciate all the work that went into this project. I expect this will prove to be a really useful resource for many of us in the field(s)!

### Reviewed by [Nai Ding](#), 05 June 2023

In general, I think the paper is almost ready to publish. A number of issues, however, still need to be addressed and most of these issues are terminology issues. For example, it should be explicitly stated that the rhythm measures general refer to the rhythm of breath instead of the rhythm of sound (if I understood it correctly).

1. For the inter-onset interval, please specify the unit (i.e., the onset of what?). If it's the onset of a breath, I wonder why it can reflect the speed of speech or music. Suppose I take a breath after every syllable in one condition and take a break after a sentence in a second condition. I may breathe more frequently in condition 1, but the speech rate, e.g., measured by the number of syllables per second, may still be higher in condition two.

Similarly, in Fig. 8, what is called an onset and break annotation? Does it mean the duration within a breath? I don't know how the Fourier transform can be used to length normalization and interpolation.

phrase length -> breath duration

interval regularity -> this one is particularly confusing since interval also refers to the breath signal in IOI

loudness -> intensity or just short-term energy

2. Abstract: "relative to speech, songs consistently use" Consider to replace consistently with generally. If I understood the results correctly, there is high variability and the result is not consistent within every participant (e.g., Fig. 7).