

Review of “Comparing the rhythm and melody of speech and music: The case of British English and French”

Abhijit Bhattacharjee

In human expression, both speech and music contain rhythmic and melodic patterns intended to communicate a certain intent. These patterns characterize *prosody*, a feature inherent to both domains. Researchers have long theorized that the linguistic prosody in an individual culture’s native language influences the melodies and rhythms of its music. Is there a way to demonstrate that musical patterns reflect the prosody of their creators’ languages?

Authors Aniruddh D. Patel, et al. believe they have the answer. By studying the melodies and rhythms of English and French speakers, and comparing them to classical instrumental music in both the English and French traditions, the authors found strong evidence to suggest that music evinces the prosody of a composer’s mother tongue. They also believe that the experimental methods they developed to perform this study can apply to all languages and musical traditions in the world.

To compare melodies and rhythms across speech and music for two different languages, it is first necessary to devise quantitative metrics for objective comparison. Patel, et al. did this by calculating the normalized pairwise variability index (nPVI) and coefficient of variation (CV) for rhythmic patterns, in addition to the *prosogram* and pitch coefficient of variation (also CV) for melodic patterns. The nPVI, which measures durational contrast between successive vowels in a spoken sentence, is also a strong metric for rhythm in music since musical notes can be roughly compared to syllables, which are built from vowels. For melody, the prosogram is an appropriate tool because it tries to capture the intonation of sounds as perceived by human ears, providing a series of level contrasts between successive pitches. The CV for both rhythm and melody is a derived parameter that exposes overall variability in each respective feature.

Using a test dataset containing English and French speech samples as well as British and French instrumental music from early 20th-century composers, the authors gathered the measurements mentioned above. Each sample was processed for rhythm and melody features such as nPVI and CV and presented analytically in a rhythm-melody (RM) feature space. The results of the analysis showed a clear similarities between speech and music in both intersyllabic (or inter-note) duration and pitch interval distances for both English and French. This confirmed the intuition that musical prosody is influenced by the native language of its composer.

In this study, there was one small part of the methodology that felt unsound. To analyze rhythm nPVI for music, the authors calculated durational contrast from written score, as opposed to a performance—the implication being that there would be “undue” variation from the expressiveness of the performer. However, they calculated durational contrast for speech directly from acoustic samples. An orator could be considered a “performer,” so wouldn’t an acoustical performance of the music be a more apt comparison? After all, the question here is about the correlation of native language speakers on music; one could have a musician of the same culture and language perform the piece to be compliant with the study. Another issue is that the authors claim the quantitative methods to be generally applicable to any cultural background. Many other cultures do not have a musical notation that reveals inter-note durations; thus it would be obligatory to calculate rhythm nPVI upon acoustic samples directly.