A Computational Method for Empirically Validating Synchronisation Between Musical Phrase Arcs and Autonomic Variables

Natalia Cotic[1], Vanessa Pope[1], Mateusz Soliński[1], Pier Lambiase[1,2], Elaine Chew[1]

[1] School of Biomedical Engineering & Imaging Sciences and Department of Engineering
King's College London, United Kingdom
[2] University College London, United Kingdom

Abstract

Prior research suggests listeners' autonomic variables may entrain with musical phrase arcs. Here, we apply a computational technique for automatic extraction of probabilistic phrase arc boundaries from recorded music audio and compare them to envelopes derived from physiological signals (respiration and RR intervals). The objective is to automate the evaluation of synchronisation of autonomic reactions to musical phrases, thereby assessing music's ability to regulate physiological states. 20 participants' (10 women, aged 39±11 years old) physiological signals (respiration and RR intervals) were recorded whilst listening to Prokofiev playing Prokofiev's Gavotte Op.12 No.2 rendered on a reproducing piano. Phrase arcs are computed from the loudness profile using a novel Bayesian approach incorporating dynamic programming. Results show increased curve similarity between the loudness phrase arcs and the physiological signal envelopes for the original data compared to the surrogate, with p-values of 0.057 and 0.021 based on the binomial test. We have presented a fully automated data-driven technique for testing hypotheses about entrainment between musical phrase arcs and autonomic variables. Preliminary findings support the potential of employing music with specific structural qualities to modulate physiological signals.

1. Introduction

Cardiovascular disease is the number one cause of death [1]. Studies have shown that music interventions improve cardiovascular outcomes [2]. However, specifics of how music modulates cardiovascular variables are not well understood, hindering the use of music in precision medicine. By analysing individuals' autonomic responses that are synchronised with music structures during music engagement, we can identify the music properties that elicit a reaction from them. This knowledge can then be used to create cardiovascular therapies tailored to the person's response. Building a computational method to analyse people's response will allow music to be integrated into the design of digital cardiovascular therapies and also enable these therapies to be applied at scale.

An important mechanism through which music works on the cardiovascular system is via entrainment [3]. In 2009, Bernardi et al. [4] discovered that listeners' autonomic variables "tracked the music emphasis and rhythmic phrases" that have 10-second periods, like Mayer waves. Here, our goal is to design a computational system to automate the process of phrase finding and evaluation of music-based autonomic entrainment to test the extent to which music entrains the autonomic variables.

In Music Information Research (MIR), musical phrases have been shown to be marked by variations in tempo and loudness, with impact on perceived segments [5, 6]. Musical phrases are indicated in music performance by arc-like increases and decreases in tempo [7, 8] and/or loudness [9–11]. This has been operationalised in phrase arc finding algorithms [5, 12] based on a Bayesian approach. We use [12], which gives a credence profile of likely phrase arc boundaries based on some priors of the distribution of arc parameters. The results of this analysis on both musical features and physiological signals are compared to provide an assessment of entrainment.

2. Methodology

2.1. Data Collection

The data analysed is obtained from the HeartFM study [13], in which 8 piano performances were computationally altered to make 30 loud/soft, fast/slow versions. ECG, respiration, continuous blood pressure (BP) from 92 participants were recorded whilst listening to music playlists comprising of 9 renditions of the 8 pieces. The music was rendered through a reproducing grand piano, which enables music to be a controlled variable by ensuring consistent playback through the electronic mechanism

 ISSN: 2325-887X DOI: 10.22489/CinC.2024.380

that powers the piano whilst letting participants listen to live music instead of recordings. For testing the method we looked at the subset of participants that listened to one version of Prokofiev playing Prokofiev's Op.12 N2 Gavotte, as it contains well defined phrases and it had a high number of people that listened to it throughout the study. Out of the 24 participants who listened to version 1 of the Gavotte, data from 4 was not considered due to low quality signals.

Music Data – The MIDI file of Prokofiev playing the Gavotte and the sound recording of the music rendered on the reproducing piano are aligned. Together with the digital score, the music files were processed using the music alignment and feature extraction tool cosmodoit [14]. Two of the automatically extracted features were tempo and loudness. Beat misalignments were corrected manually.

Physiological Data – Synchronised electrocardiographic (ECG) and respiration signals were collected via the heartfm mobile app [15]. The ECG data was collected with a Polar H10 sensor sampled at 1000Hz and RR intervals were derived from the ECG. The RR intervals were also detected with Polar's in-built software which also downsamples the signal to 130Hz. Respiratory signals were obtained from the BIOPAC respiration belt that captures changes in participants' thoracic circumference. The signal was sampled with a frequency of 13Hz. The respiration data was de-trended through a rolling average calculation with window size of 200 data points, followed by a 4th order Butterworth filter and 1Hz cutoff frequency. These steps ensured the data was cleaned without affecting the morphology and natural structure of the breathing pattern.

2.2. Data Preparation

Both the music and physiological signals were interpolated to the music beats to facilitate plotting and comparisons along the same beat-time axis. Following interpolation, all data was scaled with the median and quantiles for easier comparison.

2.3. Data Analysis

2.3.1. Phrase Arcs

To create a computational pipeline for this analysis, the arc boundary algorithm built in [12] located likely phrase arc boundaries through a Bayesian algorithm, which were then compared with the autonomic variables analysed. The priors required were estimated from the first 10 arc boundaries which were manually annotated, by listening to the piece. Quadratic functions were then fitted to the loudness signal values between each pair of boundaries. The estimated coefficient and noise values were then used as priors for the algorithm. This Gavotte was played with minimal tempo variation; phrase information was mainly conveyed through loudness variation. Thus, arcs were obtained by processing only the loudness signal with the algorithm. Based on the priors, the Bayesian program calculated a credence value for there being an arc boundary at each individual beat index. These credence values were then used for the analysis.

2.3.2. Envelopes

The envelopes of the physiological signals were also computed to test if shifts in the signals occurred at the beats where the algorithm assessed a high loudness boundary credence. The Hilbert transform was used for processing the physiological signals, respiration and RR intervals, as it provides smooth, continuous, and frequency-independent envelopes while preserving signal characteristics. Spline interpolation was then used to create the envelope from the detected peaks of the Hilbert transform.

2.3.3. Curve Differences

Curve difference calculation was implemented to calculate the similarity between the music and physiological signal. Earth Mover's Distance (EMD) measures the cost of transforming one distribution into another. EMD considers the shape and spread of the data, an advantage for measuring if the physiological signals synchronise with the music, and the similarity of the resulting profiles. This implementation uses the Wasserstein Distance function in Python.

2.4. Surrogate data

The same analysis steps were applied to surrogate data to test the significance of the results. Surrogate data was generated for each individual separately, with the aim of having test data that matched the unique physiological behaviour of the individual, while allowing for comparison against their reaction during music. For each participant, the physiological signals for all music listening segments from their entire session except for the track analysed were concatenated. Segments of length matching the number of beats of the Gavotte were selected at random from the concatenated signals. The track order was shuffled at each iteration to ensure different segment combinations were being analysed each time. The envelopes of these segments had their EMD computed against the loudness credence boundaries of the Gavotte to determine if participants' reactions during the Gavotte was more similar to the musical stimuli than their physiological recordings during other music tracks were to the Gavotte.

3. Results

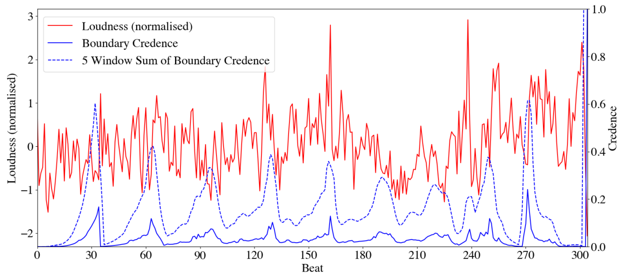


Figure 1: Loudness in sones (red), loudness boundary credence profile (blue) and its five-point sum (dotted blue), and phrase arcs (orange) detected by the Bayesian algorithm.

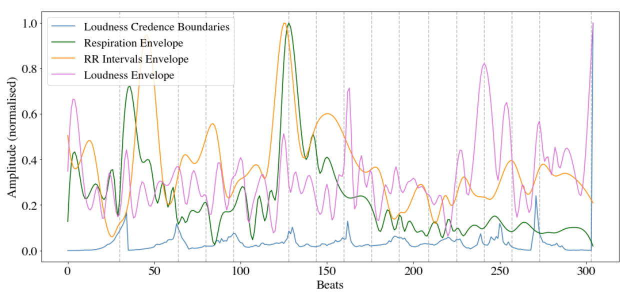
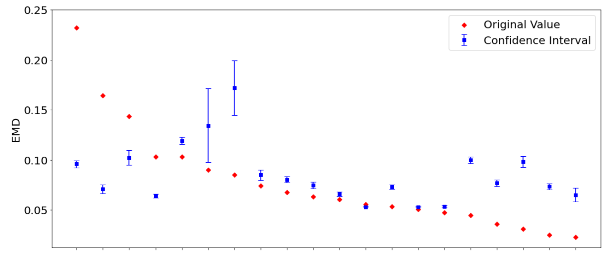


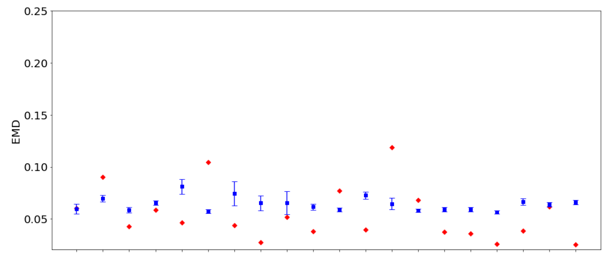
Figure 2: Loudness boundary credence profile and the physiological envelopes showing shifts around loudness credence peaks.

Figure 1 shows the loudness boundary credence profile (in blue) and the plausible phrase arcs detected by the Bayesian algorithm. Figure 2 shows the loudness boundary credence profile (blue) with the respiration envelope (green), and RR interval envelope (orange). Note that changes in the physiological envelopes occur near the boundary credence profile peaks. The Bayesian algorithm generally aligns with the manual annotations (grey vertical lines), indicating the algorithm agrees well with human perception. The algorithm identified fewer boundaries overall compared to the manual annotation, though for the ones identified by both methods there was an average 2 ± 2 beat difference between the boundary placements.

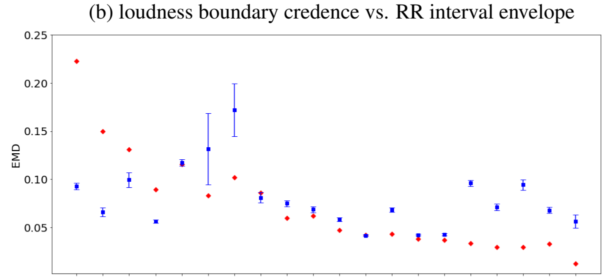
For most listeners, the EMD between the loudness credence profile and the respiration or RR envelope increases for the surrogate data, indicating greater similarity between the curve profiles during music listening. Figure 3 presents the curve difference results for both the original physiological data and the surrogate dataset shown as its 95% confidence interval. These differences were calculated between the loudness arc outputs and the envelopes of the physiological data. On average, the surrogate results



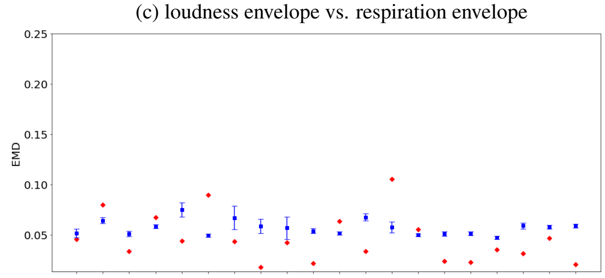
(a) Loudness boundary credence profile (blue) with the respiration envelope (green), RR interval envelope (orange), loudness envelope (purple).



(b) loudness boundary credence vs. RR interval envelope



(c) loudness envelope vs. respiration envelope



(d) loudness envelope vs. RR interval envelope

Figure 3: EMD values and corresponding surrogate data confidence intervals sorted by participant, showing increased similarity between curves during music listening.

had higher curve differences indicating reduced similarity, with values 0.02 and 0.06 higher for the loudness arcs with RR intervals and respiration envelopes, respectively. The values are sorted in descending order by participants in Figure 3a, and this order is maintained across the subsequent sub-figures.

Using the loudness boundary credence output instead of the envelope yields more statistically significant results (Table 1), partly due to the increased variation in the loudness envelope profile compared to the credence graph.

	Loudness credence	Loudness envelope
RR Interval	0.021*	0.131
Respiration	0.057	0.131

Table 1: p-values for the binomial test for the results visualised in Figure 3 showing higher statistical significance for the loudness boundary credence profiles.

4. Discussions and Conclusions

For 15 out of the 20 listeners, the musical and physiological curves were more similar when congruent with the music selection than when compared to the surrogate data. This supports a degree of entrainment between the musical stimuli and the autonomic nervous system variables caused by a musical feature, loudness in this case.

Looking at data from 20 participants whilst they listened to Prokofiev's Op.12 No.2 Gavotte, a Bayesian algorithm identified the loudness arcs and the envelopes of the physiological signals were also computed. Statistical significance was achieved when using the loudness credence output from the algorithm compared to the envelope. The results of this computational method for entrainment quantification show that there is increased synchronicity between physiological signals and phrase boundaries during music. This provides further evidence for music's potential for modulating autonomic variables. Future work will look into the relation between entrainment of phrase arc boundaries and participants' heart rate variability metrics.

Acknowledgments

This result is part of the COSMOS and HEART.FM projects, which have received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement Nos. 788960 and 957532). NC is supported by a UK Research & Innovation (UKRI) Engineering & Physical Sciences Research Council (EPSRC) doctoral studentship through King's College London.

References

[1] WHO. Cardiovascular diseases, Jun 2021. URL https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds).

[2] Kulinski J, Ofori EK, Visotcky A, Smith A, Sparapani R, Fleg JL. Effects of music on the cardiovascular system. Trends in Cardiovascular Medicine August 2022; 32(6):390–398. URL https://doi.org/10.1016/j.tcm.2021.06.004.

[3] Cross I. Music and communication in music psychology. Psychology of Music 2014;42(6):809–819. URL https://doi.org/10.1177/0305735614543968.

[4] Bernardi L, Porta C, Casucci G, Balsamo R, Bernardi NF, Fogari R, Sleight P. Dynamic interactions between musical, cardiovascular, and cerebral rhythms in humans. Circulation Jun 2009;119(25):3171–3180. URL https://doi.org/10.1161/circulationaha.108.806174.

[5] Stowell D, Chew E. Maximum a posteriori estimation of piecewise arcs in tempo time-series. arXiv Cornell University January 2013;URL https://arxiv.org/abs/1302.0136.

[6] Povel DJ. Temporal structure of performed music. Acta Psychologica June 1977;41(4):309–320. URL https://doi.org/10.1016/0001-6918(77)90024-5.

[7] Todd NPM. The dynamics of dynamics: A model of musical expression. The Journal of the Acoustical Society of America June 1992;91(6):3540–3550. URL https://doi.org/10.1121/1.402843.

[8] Chuan CH, Chew E. A hybrid system for automatic generation of style-specific accompaniment. Computational Creativity 01 2007;.

[9] Cheng E, Chew E. A local maximum phrase detection method for analyzing phrasing strategies in expressive performances. In Communications in Computer and Information Science, volume 37. ISBN 978-3-642-04578-3, 01 2009; 347–353.

[10] Repp BH. Probing the cognitive representation of musical time: Structural constraints on the perception of timing perturbations. Cognition January 1992; 44(3):241–281. URL https://doi.org/10.1016/0010-0277(92)90003-z.

[11] Todd NPM. Motion in music: A neurobiological perspective. Music Perception An Interdisciplinary Journal 1999;17(1):115–126. ISSN 07307829, 15338312. URL http://www.jstor.org/stable/40285814.

[12] Guichaoua C, Lascabettes P, Chew E. End-to-end bayesian segmentation and similarity assessment of performed music tempo and dynamics without score information. Music Science 2024;7:20592043241233411. URL https://doi.org/10.1177/20592043241233411.

[13] ERC. Maximizing the therapeutic potential of music through tailored therapy with physiological feedback in cardiovascular disease, November 2020. URL https://doi.org/10.3030/957532.

[14] Guichaoua C, Bedoya D, Chew E. Cosmodoit: A python package for adaptive, efficient pipelining of features from performed music, In preparation.

[15] Chew E, Fyfe L, Picasso C, Lambiase P. Seeing music's effect on the heart. European Heart Journal 2024;.

Address for correspondence:

Natalia Cotic

Becket House, 1 Lambeth Palace Road, London, SE1 7EU, United Kingdom

natalia.cotic@kcl.ac.uk