2025 47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) | 979-8-3315-8618-8/25/$31.00 ©2025 IEEE | DOI: 10.1109/EMBC58623.2025.11251651

Brain-Body Coupling in Listening to Metronomic Sounds and Music

Abhinav Uppal[1,*], Min Suk Lee[1], Teng Fei[2], Adyant Balaji[1], Zhaoyi Liu[3], Tzyy-Ping Jung[4,5], Lara M. Rangel[2], Virginia R. de Sa[2,5,6], Rahul Parhi[3], Akshay Paul[5], Yuchen Xu[5], and Gert Cauwenberghs[1,5]

***Abstract*— Wearables continue to expand their ability to sample more signals from the brain and body while reducing form factors and limiting on-device computing to meet real-world comfort and power constraints. Where evolving multimodal sensing platforms can combine electrophysiological (ExG), optical, chemical, and mechanical sensors for holistic brain-body state estimation, understanding of cross-modal coupling mechanisms remains limited. Here, we collect and analyze bio-signals in a brain-body coupling experiment using simultaneous electroencephalogram (EEG), electrocardiogram (ECG), and respiratory measurements. Our experimental paradigm contrasted listening to self-selected music, preceded by tempo-matched isochronous cues. The results show a clear entrainment of cardiac rhythms with the underlying beat of auditory stimuli for both metronome cues and subject-selected music. Magnitude-squared coherence analysis showed frequency coupling across recorded modalities, with ECG, EEG, and breath, exhibiting peak coherence near heard or underlying musical beat or its harmonics. We present minimally pre-processed data to motivate future methodological explorations of multiscale brain-body entrainment to rhythmic stimuli.**

***Clinical relevance*—These findings have implications for designing music therapy and biofeedback interventions that could adapt to ongoing brain-body rhythms. Understanding brain-body entrainment mechanisms and validating simplified measurement approaches could enable more accessible and effective rhythmic interventions in clinical settings, particularly for conditions benefiting from audio-based therapies or cardiac rehabilitation.**

I. Introduction

Coupling between oscillating systems can be unidirectional or bidirectional, sometimes contrasted as entrainment or synchronization [1]. More generally, an oscillating system can dynamically entrain to an external rhythm, or two oscillating systems can bidirectionally couple to synchronize in phase [2]. Entrainment phenomena in the brain and body have been reported for environmental stimuli and internal signaling. For example, audio can entrain breath [3], breath can affect EEG [4], [5], and EEG can track the audio envelope [6]. Therefore, researchers hypothesize that brain and body rhythms follow a dyadic structure, where slower oscillations couple with faster oscillations via phase-amplitude or frequency-frequency (phase-phase) coupling [7].

[1]Shu Chien-Gene Lay Dept. of Bioengineering, [2]Dept. of Cognitive Science, [3]Dept. of Electrical and Computer Engineering, [4]Swartz Center for Computational Neuroscience, [5]Institute for Neural Computation, [6]Halıcıoğlu Data Science Institute, UC San Diego, La Jolla, USA.
*Contact: auppal@ucsd.edu

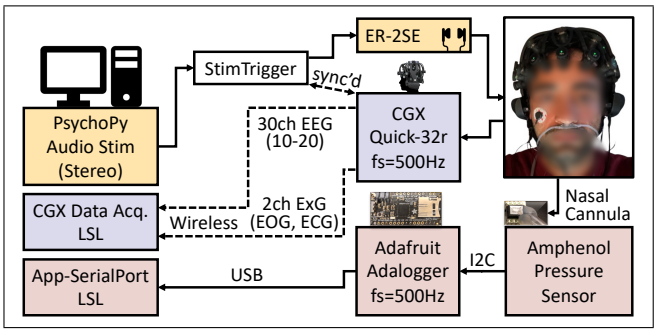

Fig. 1. Data acquisition (acq.) setup for synchronized recording of electrophysiology (ExG) signals: dry electroencephalogram (dry EEG), electrocardiogram (ECG), electrooculogram (EOG), combined with nasal respiration during eyes-closed music listening. Stereo audio stimuli (stim) were played in a PsychoPy experiment over wired, perceptually flat earbuds (Etymotic ER-2SE), routed through a trigger box (CGX StimTrigger) for synchronized (sync'd) event marking of collected data. ExG signals were recorded at a 500 Hz sampling frequency (fs) using a wireless dry EEG headset (CGX Quick-32r) with 30 EEG channels (ch) in standard 10-20 montage (reference: A1), and two auxiliary channels used as chest ECG, and EOG below the right eye. A nasal cannula connected to a digital pressure sensor (Amphenol ELVH) recorded breath synchronized with ExG using Lab Streaming Layer (LSL) based PC applications.

Given the potential to entrain brain-body signals with rhythmic stimuli, rhythmic entrainment has been adopted for brain rehabilitation in, for example, speech and motor systems [8], Alzheimer's [9], and schizophrenia [10]. Entrainment effects, however, vary across individuals [11] and are not always clear or significant [12], [13]. This variability emphasizes the need to develop acquisition and analysis methods for within-subject, longitudinal studies of brain-body coupling in beyond-the-lab settings. This report introduces a multimodal acquisition setup whereby a subject could be trained to adorn and perform longitudinal data collection. This setup simultaneously records electroencephalogram (EEG), electrooculogram (EOG), electrocardiogram (ECG), nasal respiration, and synchronization triggers (event markers) for aligning audio stimulus delivery with electrophysiological data, as shown in Fig. 1. We present initial results from a brain-body coupling experiment comparing two audio stimuli: self-selected music and tempo-matched metronomic pulses.

The remainder of this report is organized as follows: Sec. II describes the experimental paradigm, including listening tasks, data collection hardware, and phase-locking analysis methods. Sec. III shows brain-body signals collected during audio stimulation. Sec. IV discusses the results and proposes follow-up data collection and analysis studies.

II. Methods

A. Experimental Paradigm

The overall paradigm (Fig. 2) contrasted three different conditions: listening to self-selected music, listening to metronomic (isochronous) pulses of similar tempo, and tapping along with metronomic pulses. We analyzed and reported only listening data. The Institutional Review Board approved the experimental procedures.

1) Subject: One right-handed subject (male, Age: 30) with prior musical training was recorded longitudinally for multiple recording sessions, each lasting roughly an hour, including setup time. The session analyzed here was recorded early in the morning, with the subject instructed to avoid caffeine before recording to avoid the confounding additive effects of music and caffeine on EEG [14].

2) Audio Stimuli: Before the experiment, the subject provided a YouTube playlist of music they were familiar with and enjoyed. Tempo-matched metronomic stimuli were generated for each song on the playlist, consisting of 10 ms duration raised-cosine shaped white noise bursts, repeated 60 times at the song's BPM. For the presented analysis, we selected a listening trial in which the subject selected an uncued, live solo recording consisting of vocals and synth, averaging 85 BPM with a 4/4 metrical structure (LAUV - I Like Me Better, Sofar London, YouTube).

3) Experiment: PsychoPy [15] was used for displaying instructional text and playing audio stimuli for the tasks shown in Fig. 2. A simple song menu prompted the subject to press an alphanumeric key to select a song from their pre-provided playlist. The subject then listened to the metronome cues, followed by their chosen music. Brief noise bursts marked the start and end of each trial for offline analysis. Recordings took place in a dark room to minimize distractions and line noise (60 Hz) interference. Displayed instructions directed the subject to keep their eyes closed whenever audio played.

B. Synchronized Data Acquisition

1) ExG Sensing: We used the CGX Quick-32r wireless dry EEG headset to record 30-channel EEG data in the standard 10-20 montage, with reference and ground on either side of the left earlobe. Two auxiliary channels provided by the headset recorded ECG (auxiliary gel electrode placed on the left chest, shared earlobe reference with EEG), and EOG (auxiliary gel electrode placed diagonally below the right eye). We acquired 500 Hz samples wirelessly using CGX Data Acquisition Software, streamed to Lab Streaming Layer [16], and recorded using App-LabRecorder. CGX Wireless StimTrigger synchronized audio output with ExG data acquisition.

2) Respiration Sensing: A nasal breath sensor adapted from [17] was simplified to use a digital output pressure sensor from the same manufacturer, combined with LSL synchronization of output samples. A nasal cannula recorded nasal airflow using the top port on an Amphenol AllSensors ELVH-L01D-HRRD-C-N2A4 differential pressure sensor (1 mm H_2O, 3.3 V supply, I^2C digital output), with the bottom port left open (ambient pressure). The pressure sensor output was sampled using a 500 Hz timer interrupt on an SAMD21 microcontroller (Adafruit Adalogger), and pushed over serial (USB) into an LSL stream using App-SerialPort, v1.1, saving all acquired streams to an XDF file.

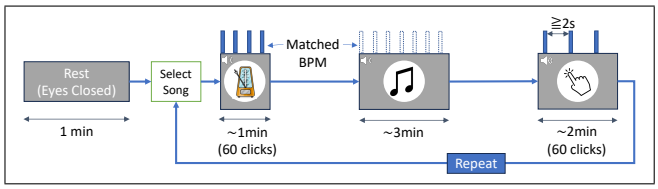

Fig. 2. The paradigm included a baseline task (1 minute of eyes-closed resting), song selection prompting, listening, and tapping trials with simultaneous bio-signal recording. Metronome icon from vectorportal.com.

3) ExG-Respiration Acquisition Latency: Latency between wireless ExG recording and wired respiration recording was determined by playing 10-ms duration, raised-cosine shaped noise bursts (identical to the metronome cues) at a repetition rate of 60 BPM. PsychoPy's analog audio output was sent to the Wireless StimTrigger, and subsequently split two ways using a DJ cable (Native Instruments). One split was input as an auxiliary channel to CGX Quick-32r headset, with the headset's REF and GND shorted to the audio's ground (with no subject in the loop for safety). The other split was used to simultaneously deliver acoustic input to the ELVH pressure sensor's top port using one of the ER-2SE (Etymotic Research) earbuds. Comparing peak audio onsets between the ExG and respiration streams showed a mean latency of approximately 28 ms, averaged over 1 minute of data collection, with the wired nasal sensor leading the wireless EEG headset.

C. Data Preprocessing

Acquired LSL streams for ExG and the respiration data were imported into Python using the `pyxdf` package, and analyzed using MNE-Python 1.9.0 [18], `neurokit2` [19], and custom scripts.

1) ExG: In this exploratory analysis, ExG data remained unfiltered to prevent accidental removal of slow cortical activity [20] or the introduction of high-pass filterilter artifacts [21]. Time domain analysis applied linear de-trending to generate epoch images. Time-scattering analysis (described below) was run on raw data, without de-trending.

2) Respiration: 14-bit pressure samples output by the pressure sensor underwent normalization and band-pass filtering using `neurokit2.rsp.rsp_clean`, which implemented a 0.05 Hz–3 Hz second-order Butterworth filter. Although the ExG and respiration LSL streams had the same nominal sampling rate of 500 Hz, LSL time-stamps and effective sampling differed because of the two separate stream sources: CGX Acquisition Software and App-SerialPort. We used cubic spline interpolation from `scipy` to interpolate the band-pass filtered respiration stream to match ExG timestamps. Nasal pressure data were integrated (cumulatively summed) and band-pass filtered again to obtain

a surrogate for respiratory volume, as typically measured by chest or abdominally worn respiration belts.

3) Visual Inspection and Pulse Artifacts: ExG and respiration samples, now aligned to the same time basis, were visually inspected to assess signal quality. Temporal channels T7 and T8 displayed pulse artifacts [22]. T7 showed larger pulse artifacts for the tested subject and served as a proxy to measure the pulse of the scalp.

D. Event-Locked Epoching and ECG-to-Beat Phase

1) Data Epoching: We captured audio triggers in the ExG stream to extract events and segment the data into metronomic and music-listening trials. The metronomic listening trial was epoched into approximately 706 ms segments, corresponding to 85 BPM repetition rate, with t=0 locked to audio trigger onsets. For the music listening trial, we used BeatNet 1.1.3 [23] to extract the underlying beat of the music stimulus, which identified quarter-note ¼-note beat onsets across the entire piece. We generated music-listening epochs by selecting the first onset in each group of four beats, resulting in longer segments of approx. 2.8 s (four ¼-notes). No epoch rejection criterion was applied to preserve all epochs in experimental time order for subsequent analysis.

2) R-Peak Synchronization Analysis: Event-locked epoch images were generated to visually examine if and when ECG entrained to the metronome or beat onsets across metronomic and music listening trials. After visual examination, we used `neurokit2.ecg_process` to extract R-peaks from the ECG signal and performed a permutation-based analysis to ascertain ECG-to-beat synchronization for the music listening trial. We first defined a phase difference $\phi^{(i)}_{R,beat}$ between the ith R-peak at time-stamp $t^{(i)}_R$, and an ongoing stimulus beat with period T_{beat} as

$$\phi^{(i)}_{R,beat} = \frac{mod[t^{(i)}_R, T_{beat}]}{T_{beat}} \times 2\pi \tag{1}$$

The mod function effectively generated equally-spaced beat onsets spanning the recording session's duration, i.e., about 15 minutes. T_{beat} for this music stimulus is ≈706 ms (85 BPM). Next, to determine a baseline distribution of possible phase offsets $\phi^{(i)}_{R,beat}$, we permuted the timing of R-peaks $t^{(i)}_R$ by circularly shifting all $t^{(i)}_R$ time samples by the same randomized time shift, drawn at random from the range of all possible time steps between 0 and the duration of the experiment. Each circular shift produced a surrogate ECG series $s^{(k)}_R$ where the ECG's internal timing structure is preserved, but any phase relationship to the stimulus beat structure is randomized. We generated 10,000 such circularly shifted instances of $s^{(k)}_R$, resulting in a bootstrapped baseline population of physiologically plausible yet stimulus-decoupled ECG phase offsets. We compared this baseline distribution against actual (unperturbed) phase offsets $\phi^{(i)}_{R,beat}$ from the music listening trial.

E. Time-Scattering and Coherence Analysis

We also extracted time-frequency information for all bio-signals using the time scattering transform [24] as implemented in `kymatio` v0.3.0 [25]. We used a two-layer scattering network structure with 6 wavelets per octave (Q = 6) in the first layer and one wavelet per octave (Q = 1) in the second. The first octave used linearly spaced band-pass filters, while the remaining used logarithmically spaced center frequencies. Each wavelet (band-pass filter) is followed by a modulus nonlinearity and a low-pass filter (J = 6, corresponding to 64 samples @ 500 Hz, or 128 ms). We then computed magnitude-squared coherence between first-order scattering coefficients of each modality using `scipy.signal.coherence` with a segment length of 1024.

III. Results

A. Beat Onset-Locked Epochs and Evoked-Responses

1) Metronome Listening: Fig. 3 (a)-(f) plots beat onset-locked epochs during listening to metronomic cues at 85 BPM (706 ms period). The ECG and corresponding pulse artifacts in EEG electrode T7 reveal cardiac entrainment to the metronomic beat onsets. Cz is shown as a representative EEG electrode without pulse artifact, with alpha (10 Hz) activity visible as ripples. Motion artifacts are retained given the minimal pre-processing approach with only linear detrending (no filtering or epoch rejection) to preserve the temporal order of the listening trial.

2) Music Listening: Fig. 3 (g)-(l) shows epochs time-locked to the underlying musical beat extracted offline from the music stimulus. ECG shows phase-locking to the musical beat with some wander, visible as four bands with one heartbeat (R-peak) per musical beat. Pulse artifacts picked by T7 track ECG, and Cz shows representative EEG activity.

B. ECG-to-Musical Beat Phase Synchronization

Fig. 5 shows phase distributions from circular permutation analysis of ECG-to-musical beat phase $\phi^{(i)}_{R,beat}$. The blue bins show phase distributions from 10,000 surrogate binnings, which preserve the internal ECG timing structure but randomize phase offsets from 85 BPM beats. Comparing this baseline distribution to the actual $\phi^{(i)}_{R,beat}$ distribution from a music listening trial shows certain phase bins exceed baseline probabilities, indicating synchronization.

C. Magnitude-Squared Coherence

Fig. 4 shows magnitude-squared coherence for recorded bio-signals during the metronomic and music listening trials. Yellow vertical lines mark the audio's beat frequency (1.4 Hz for 85 BPM) and harmonics. Music listening (right column) reveals stronger synchronization between brain-body signals and the musical beat. EEG electrodes Cz and T7 show moderate coherence at the beat frequency (1.4 Hz) and alpha activity (10 Hz). Cz and ECG show weak coherence at the beat frequency and strong coherence for line noise (60 Hz). T7 and ECG show moderate coherence at the beat frequency and line noise, given that T7 picked pulse artifacts. Breath shows moderate coherence with ECG for the second harmonic of the beat frequency (2.8 Hz), compared to negligible coherence with Cz (EEG) or T7 (EEG with pulse artifact).

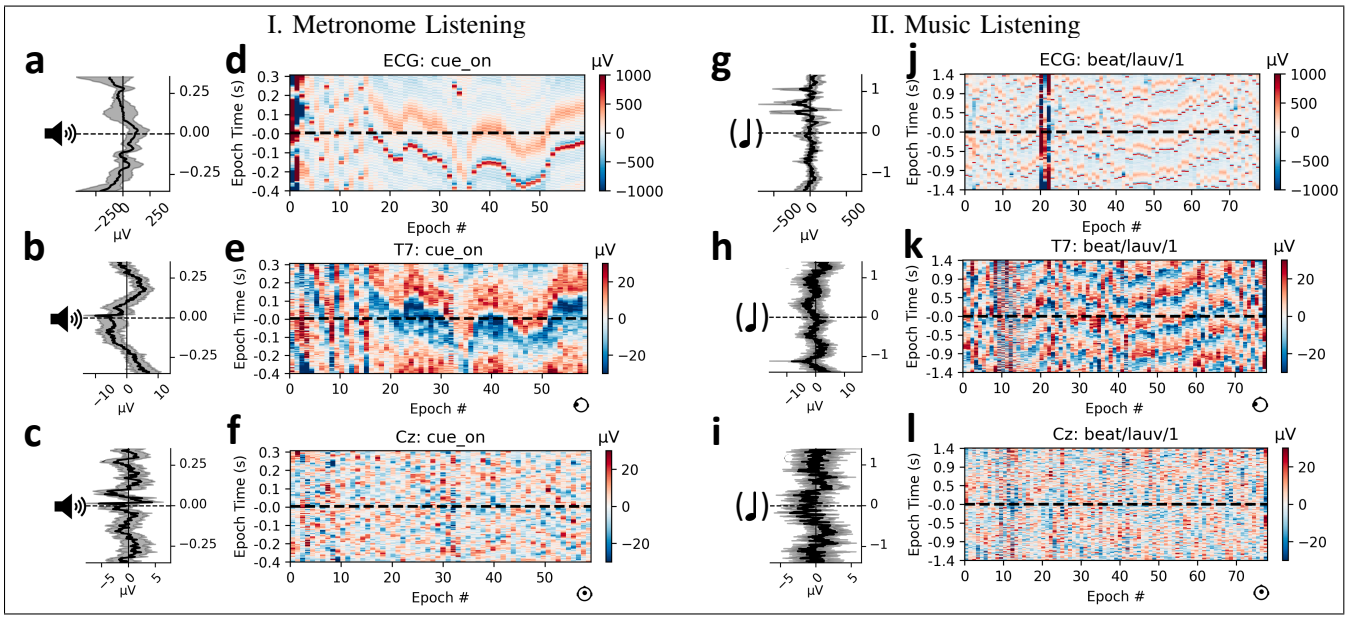

Fig. 3. Beat-onset evoked response potentials (ERPs) and contributing epochs as horizontally-stacked images, from eyes-closed listening to I. (a)-(f) 60 repetitions of periodic metronomic sounds (`cue_on` events), delivered at 85 beats per minute (BPM) over 42 s; II. (g)-(l) subject-selected music also averaging 85 BPM for 233 s. ECG: gel-pad auxiliary channel on the left chest, showing R-peaks in red, T-waves in orange, and 60 Hz line noise as background texture. T7: A dry EEG channel located over the left temporal lobe. This channel captures EEG and pulse artifacts for this subject, phase-locked to ECG with ≈200 ms latency. Cz: A representative dry EEG channel (vertex). All channels referenced to A1 (left earlobe) using dry contact. (a)-(f) t=0 corresponds to heard audio onsets. (g)-(l) t=0 corresponds to beat onsets extracted from music using BeatNet. (g)-(l) epochs are four times longer (≈3 seconds) than (a)-(f) epochs, corresponding to four ¼ beats (one measure of music).

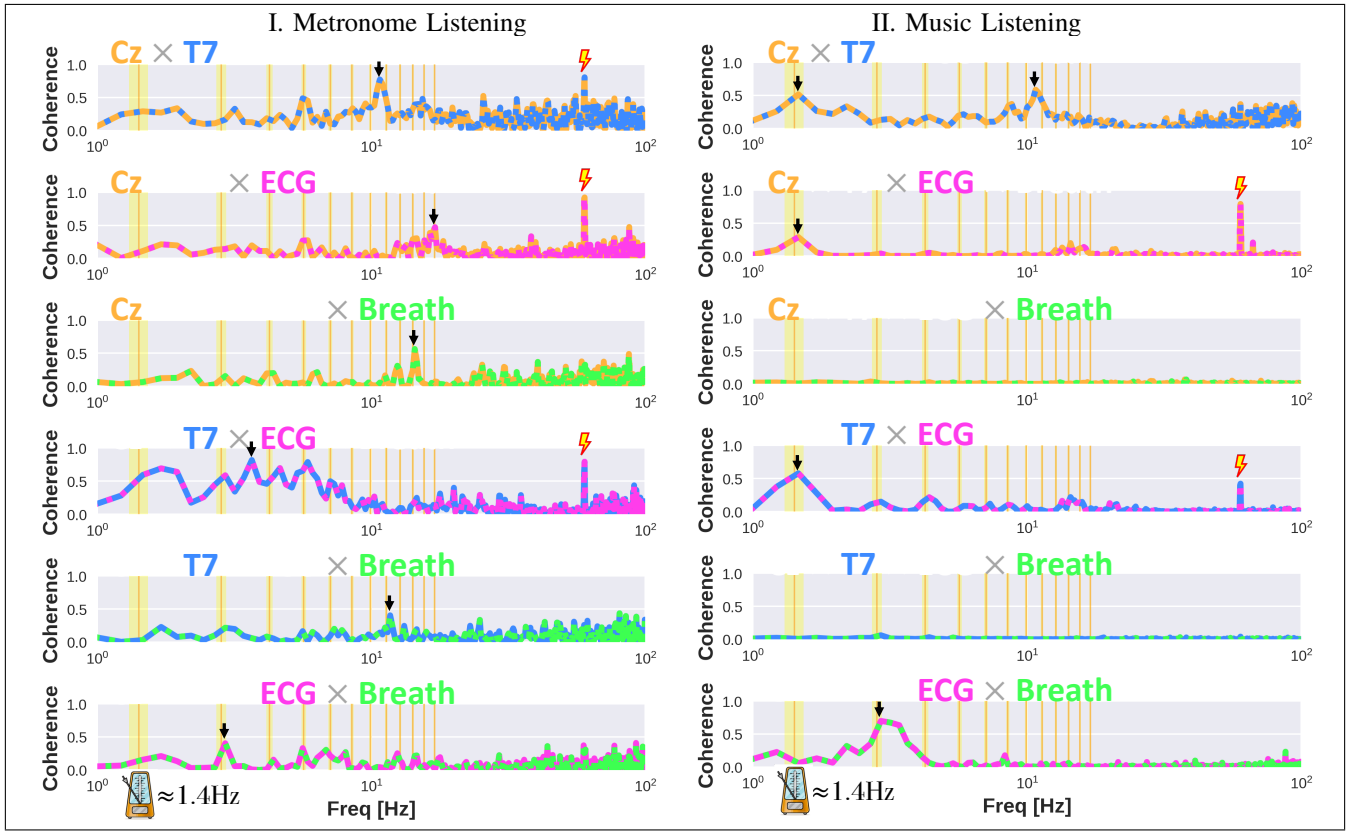

Fig. 4. Magnitude-squared coherence between first-order scattering spectra of pairs of brain-body signals during I. metronomic and II. music listening for approx. 85 BPM (1.4 Hz) underlying musical beat. Arrows show peak coherence frequencies, whereas thunderbolts mark artifactual coupling at 60 Hz (line noise). Cz: EEG electrode (vertex), T7: EEG electrode with pulse artifact, ECG: left chest, Breath: filtered and integrated nasal airflow.

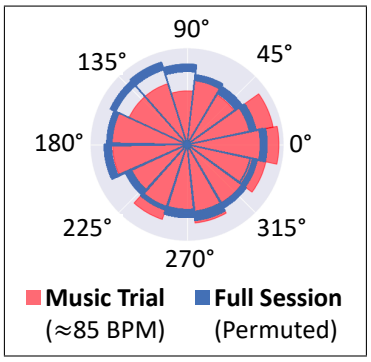

Fig. 5. ECG-to-musical beat phase corresponding to the music listening trial (blue) overshoots the baseline distribution (red) for certain phases, indicating above-baseline synchronization.

IV. Discussion and Conclusions

This study demonstrates the feasibility of capturing and analyzing brain-body coupling during rhythmic auditory tasks using a minimally invasive recording setup. We synchronized EEG, ECG, and breath recordings and demonstrated cardiac entrainment to rhythmic auditory stimuli in a musically trained subject, with visible phase-locking during metronome and music listening. We used permutation testing to infer ECG-to-musical beat synchronization during music listening. We used magnitude-squared coherence analysis to examine frequency coupling across brain-body signals, quantifying EEG interactions at the beat frequency and characteristic alpha activity. This analysis shows that this hardware setup can synchronously capture multiple physiological rhythms, making it suitable for longitudinal data collection under naturalistic tasks, such as music listening. Our findings suggest that music can be a powerful synchronizer of brain and body rhythms, potentially through both bottom-up sensory processing and top-down attentional mechanisms. The ability to capture multiple physiological signals with precise synchronization while maintaining minimal setup complexity supports applications in music therapy, cardiac rehabilitation, and cognitive behavioral interventions. Future work should explore translating these findings into real-time monitoring and adaptive therapeutic systems.

Acknowledgment

A.U. thanks Nicole Wells for helping with data acquisition; Amphenol AllSensors for providing ELVH pressure sensor samples; CGX support for providing a method for testing trigger timing; and Anthony Kim, Gina D'Andrea-Penna, Victor Minces, Andrea A. Chiba, Sarah Creel, Seana Coulson, María Florencia Assaneo, Benjamin L. Smarr, Alessandro D'Amico, Cassia Rizq, Srinivas Ravishankar, Ian Jackson, Eena L. Kosik, Keren Shao, Shlomo Dubnov, Shihab A. Shamma, and John R. Iversen for helpful discussions.

References

[1] P. Lakatos *et al.*, "A new unifying account of the roles of neuronal entrainment," *Current Biology*, vol. 29, no. 18, pp. R890–R905, 2019.
[2] L. Moumdjian *et al.*, "Entrainment and synchronization to auditory stimuli during walking in healthy and neurological populations: A methodological systematic review," *Frontiers in Human Neuroscience*, vol. 12, 2018.
[3] G. Leslie *et al.*, "Engineering Music to Slow Breathing and Invite Relaxed Physiology," in *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, Sep. 2019, pp. 1–7.
[4] A. B. L. Tort *et al.*, "Respiration-Entrained Brain Rhythms Are Global but Often Overlooked," *Trends in Neurosciences*, vol. 41, no. 4, pp. 186–197, Apr. 2018.
[5] M. C. Melnychuk *et al.*, "A Bridge between the Breath and the Brain: Synchronization of Respiration, a Pupillometric Marker of the Locus Coeruleus, and an EEG Marker of Attentional Control State," *Brain Sciences*, vol. 11, no. 10, p. 1324, Oct. 2021.
[6] S. Geirnaert *et al.*, "Fast, Accurate, Unsupervised, and Time-Adaptive EEG-Based Auditory Attention Decoding for Neuro-steered Hearing Devices," in *Brain-Computer Interface Research*, C. Guger *et al.*, Eds. Cham: Springer Nature Switzerland, 2024, pp. 29–40.
[7] W. Klimesch, "The frequency architecture of brain and brain body oscillations: an analysis," *European Journal of Neuroscience*, vol. 48, no. 7, pp. 2431–2453, 2018.
[8] M. H. Thaut *et al.*, "Neurobiological foundations of neurologic music therapy: rhythmic entrainment and the motor system," *Front. Psychol.*, vol. 5, 2015.
[9] D. Chan *et al.*, "Gamma frequency sensory stimulation in mild probable alzheimer's dementia patients: Results of feasibility and pilot studies," *PLOS ONE*, vol. 17, no. 12, p. e0278412, 2022.
[10] Y. Lin *et al.*, "Novel EEG-based neurofeedback system targeting frontal gamma activity of schizophrenia patients to improve working memory," in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 2022, pp. 4031–4035.
[11] M. F. Assaneo *et al.*, "Speaking rhythmically can shape hearing," *Nature Human Behaviour*, vol. 5, no. 1, pp. 71–82, 2021.
[12] H. Mütze *et al.*, "The effect of a rhythmic pulse on the heart rate: Little evidence for rhythmical 'entrainment' and 'synchronization'," *Musicae Scientiae*, vol. 24, no. 3, pp. 377–400, 2020.
[13] R. Nomura, "Reliability for music-induced heart rate synchronization," *Scientific Reports*, vol. 14, no. 1, p. 12200, 2024.
[14] B. Qiu *et al.*, "Effects of Caffeine Intake Combined with Self-Selected Music During Warm-Up on Anaerobic Performance: A Randomized, Double-Blind, Crossover Study," *Nutrients*, vol. 17, no. 2, p. 351, Jan. 2025.
[15] J. Peirce *et al.*, "PsychoPy2: Experiments in behavior made easy," *Behavior Research Methods*, vol. 51, no. 1, pp. 195–203, 2019.
[16] C. Kothe *et al.*, "The lab streaming layer for synchronized multimodal recording," *bioRxiv*, Feb. 2024.
[17] R. Kahana-Zweig *et al.*, "Measuring and Characterizing the Human Nasal Cycle," *PLOS ONE*, vol. 11, no. 10, p. e0162918, Oct. 2016.
[18] A. Gramfort *et al.*, "MEG and EEG data analysis with MNE-Python," *Frontiers in Neuroscience*, vol. 7, Dec. 2013.
[19] D. Makowski *et al.*, "NeuroKit2: A python toolbox for neurophysiological signal processing," *Behavior Research Methods*, vol. 53, no. 4, pp. 1689–1696, 2021.
[20] M. Brændholt *et al.*, "Breathing in waves: Understanding respiratory-brain coupling as a gradient of predictive oscillations," *Neuroscience & Biobehavioral Reviews*, vol. 152, p. 105262, 2023.
[21] A. de Cheveigné and D. Arzounian, "Robust detrending, rereferencing, outlier detection, and inpainting for multichannel data," *NeuroImage*, vol. 172, pp. 903–912, 2018.
[22] P. J. Allen *et al.*, "Identification of EEG events in the MR scanner: The problem of pulse artifact and a method for its subtraction," *NeuroImage*, vol. 8, no. 3, pp. 229–239, 1998.
[23] M. Heydari *et al.*, "BeatNet: CRNN and particle filtering for online joint beat downbeat and meter tracking," in *22th International Society for Music Information Retrieval Conference, ISMIR*, 2021.
[24] J. Andén and S. Mallat, "Deep Scattering Spectrum," *IEEE Transactions on Signal Processing*, vol. 62, no. 16, pp. 4114–4128, Aug. 2014.
[25] M. Andreux *et al.*, "Kymatio: Scattering transforms in python," *Journal of Machine Learning Research*, vol. 21, no. 60, pp. 1–6, 2020.