

Spectral Amplitude Modulation in EEG: Potential Correlations with Musical Stimuli

Brad Caldwell, BSCE

September 26, 2024

Abstract

The study of neural responses to musical stimuli offers valuable insights into how the brain encodes complex auditory information. This paper compares spectral amplitude modulations from various EEG electrodes, time-locked to attended music, to explore potential correlations. Instead of using traditional 2D TF plots, we utilized a 3D time-frequency-amplitude (TFA) fabric over a 500 ms window, synchronized with the audio of the song. This approach was applied to five songs, examining the TFA fabric across different electrodes. The TFA fabric refers to the continuous manifold created by the frequency-amplitude (FA) curve over time. This manifold, or fabric, is characterized by its inherent smoothness and contiguity. Even spatial movements across the cortex appear to alter the morphology of the FA curve (and therefore the TFA fabric) in a contiguous manner.

Both excitation and inhibition contribute to pulling the FA curve away from an approximately $1/f$ baseline, effectively causing perturbations that determine its complex modulation. This interaction causes the spectral density curve to move in synchrony with the beat of the music, creating crests and troughs in the TFA fabric that show possible correlations with the rhythm and flow of the attended auditory stimuli. The effect appears to be more pronounced in the 30–40+ Hz region, where the vibrato of the voice, the texture of violins, and the hit of drums or cymbals appear to show up in prominent spectral amplitude crests (and sometimes troughs). This study suggests that the brain’s oscillatory activity could be organized around **(spectral) amplitude modulation**. If so, the mechanism may extend beyond music, potentially playing a broader role in the encoding of attention.

1 Introduction

Understanding how the brain processes and encodes complex auditory stimuli, such as music, is crucial for unraveling the neural mechanisms underlying perception, attention, and memory. Music, with its rich temporal and spectral structure, provides a unique opportunity to explore these processes. Previous research has demonstrated that the brain’s oscillatory activity, as measured by electroencephalography (EEG), is sensitive to various elements of musical stimuli, including rhythm, melody, and harmony. However, whether these elements correlate with elements in the time-frequency-amplitude (TFA) fabric of the local field potential (LFP, as recorded by EEG electrodes) remains largely unexplored.

In this study, we focus on potential correlations between features of audio and features in the TFA fabric as recorded from various cortical locations via EEG (see Figure 1 for an overview of TFA Fabric Terminology). We also noted ‘pit points’-recurring places in the fabric where a particular frequency dipped to near-zero amplitude for roughly 1 ms (0.1 Hz x 1 ms ‘point’). These pit points were also seen to occur in analysis of the downsampled audio signal itself, although in a more frequent, aligned, and apparently rhythmic or consistent manner (see Figure 2 to see the pit points in both audio and brain signals). In terms of the neural signal, both the pit points and modulation of the FA curve continue in the absence of a musical stimulus.

We analyzed EEG data collected from a participant while they listened to five distinct musical pieces spanning different genres and tempos. The songs chosen for this study were:

- “Use Me” by PVRIS, electrode O1: [S1](#) (all 16 electrodes simultaneously: [S6](#)).
- “Smoke” by PVRIS, electrode Fz: [S2](#).

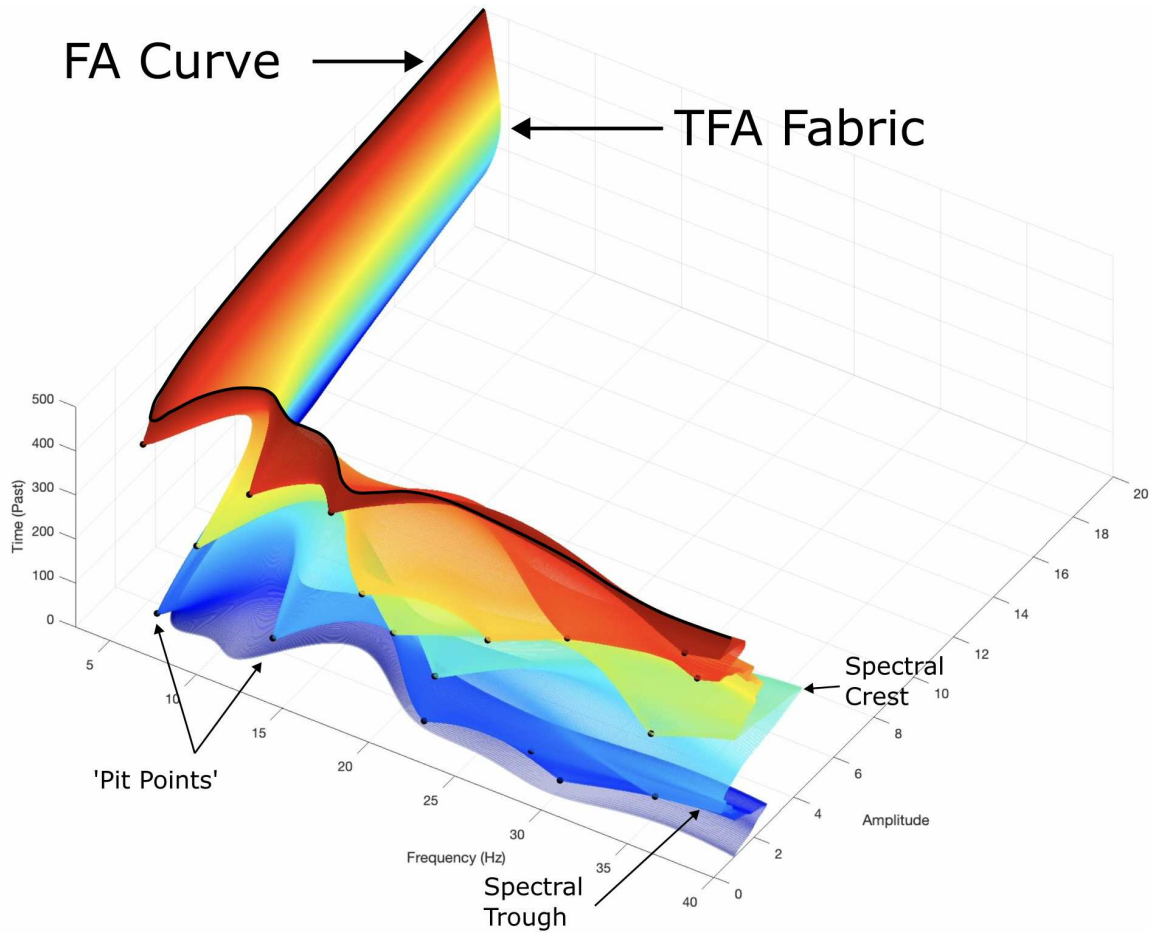


Figure 1: TFA fabric terminology.

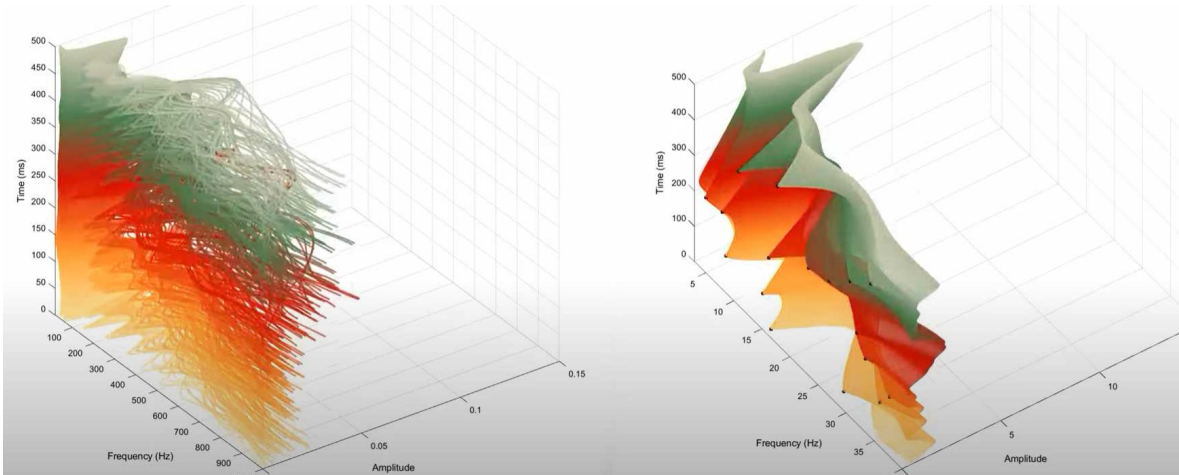


Figure 2: Comparing audio (left) and brain signals at Fp1 (right), at the same time window in the song “Reeling in the Years” by Steely Dan. Audio was first downsampled from 48 kHz to 2 kHz, whereas brain signal from electrode Fp1 was upsampled from 125 Hz to 1 kHz.

- “Reeling in the Years” by Steely Dan, electrode Fp1: [S3](#). ([S7](#) for audio to brain signal comparison of TFA fabrics of entire song).
- “Crystal Ball” by Styx, electrode T5: [S4](#).

- “Aja” by Steely Dan, electrode Fz: S5.

These songs were selected for their complex and varied auditory features, providing a rich dataset for examining how features in the TFA fabric for some cortical region’s LFP (as captured by EEG electrode) might correlate with features in music. The analysis focused on time-aligning audio and brain signals (as visualized as TFA fabric) in video format whereby potential correlations could be seen. Occasionally, a brief lead or lag was allowed (up to +/- 33 ms) if it appeared to align the features in the TFA fabric with the features in the audio of the song. In general, frontal cortex (Fp1/2, F7/3/4/8) TFA fabric features appeared most aligned with audio features if given a 0-33 ms lead relative to the audio (as though predictive), whereas occipital cortex (O1/2) TFA fabric features appeared most aligned with audio features if given a 33 ms lag relative to the audio (still slightly predictive, but closer to the time that audio signals would be expected to arrive to the cortex).

Itoh et. al. [1] said, “A major generator of the human P1, which peaks around 50 ms after sound onset, is located in the primary auditory area[2]. Considering that single neurons in the human auditory core fire as early as 10–20 ms in latency following auditory stimulus[3][4], the delayed latency of P1 at approximately 50 ms provides evidence of the first stages of neuronal processing in the neocortical circuit.” Hillyard et. al. [5] also examine timing of auditory signals.

2 Methods

- **Participant:** Male, 41 years old, familiar with all the chosen songs.
- **Materials:** OpenBCI 16-channel EEG cap with sintered Ag/AgCl electrodes, Cyton + Daisy Biosensing Board (125 Hz sampling rate) transmitting wirelessly to a MacBook Pro, Fender Acoustasonic amp playing songs from iPhone YouTube Music app using aux input, Canon 5D Mark IV DSLR camera to record and align timestamps on OpenBCI software running on MacBook Pro during the session with the timing of the music.
- **Procedure:** The EEG cap was fitted, and the electrodes were filled with conductive gel. Impedances were reduced to $\leq 5\text{-}10\text{ k}\Omega$. During the recording, the participant remained as still as possible, with eyes mostly closed, and focused on the music. After the session, timestamps marking the onset of silence and the start of the audio for each song were recorded as precisely as possible for subsequent use in Matlab time-frequency fabric videos. The clean audio was then synchronized with the Matlab-generated time-frequency fabric videos using Adobe Premiere Pro, with an expected alignment precision of ± 10 ms relative to the voltage level time points. For the video of electrode O1, the audio was deliberately offset by 33 ms (equivalent to 2 frames in a 60 Hz video) to better highlight the correlation between the SDC modulations and the musical rhythms and vocals. The test was performed on 9/17/24.
- **Data Analysis:** OpenBCI voltage values from each electrode were timestamped and were expected to be spaced evenly at 8 ms intervals (125 Hz). However, it was observed that while the samples averaged 8 ms apart, they varied from 5 ms to 11 ms. To address this, timestamps were used to plot points accurately in time, a spline was fitted, and voltage levels were interpolated at precise intervals. The data points were then upsampled to 1000 Hz, meaning interpolation values were read 1 ms apart. Once the interpolated voltages were obtained in Matlab, 371 complex Morlet wavelets, spaced from 3 Hz to 40 Hz in 0.1 Hz steps, were used to convolve the EEG data, yielding highly precise frequency information as an array of complex numbers at each 1 ms time point. According to Cohen [6], the convolution of EEG data with Morlet wavelets is a powerful method for analyzing time-frequency characteristics.

Mathematical Formulation of Complex Morlet Wavelets

1. Sine Wave Component

The sine wave component of the Morlet wavelet is expressed as:

$$\psi_{\text{sine}}(t, f) = e^{i \cdot 2\pi f t} \tag{1}$$

where $\psi_{\text{sine}}(t, f)$ is the sine wave at time t and frequency f .

2. Gaussian Window

The Gaussian window that modulates the sine wave is given by:

$$G(t, f) = e^{-\frac{t^2}{2s^2}} \quad (2)$$

where s is the standard deviation of the Gaussian, related to the frequency by $s = \frac{5}{2\pi f}$.

3. Complex Morlet Wavelet

The final Complex Morlet Wavelet is the product of the sine wave and the Gaussian window:

$$\psi_{\text{CMW}}(t, f) = \psi_{\text{sine}}(t, f) \cdot G(t, f) = e^{i \cdot 2\pi f t} \cdot e^{-\frac{t^2}{2s^2}} \quad (3)$$

where $\psi_{\text{CMW}}(t, f)$ is the Complex Morlet Wavelet at time t and frequency f .

Convolution Process

1. Fourier Transform of EEG Data

The EEG data $x(t)$ is transformed into the frequency domain using the Fast Fourier Transform (FFT):

$$X(f) = \text{FFT}\{x(t)\} \quad (4)$$

where $X(f)$ is the Fourier-transformed EEG data.

2. Fourier Transform of Morlet Wavelets

For each frequency f_i , the Morlet wavelet $\psi_{\text{CMW}}(t, f_i)$ is transformed into the frequency domain:

$$\Psi(f_i) = \frac{\text{FFT}\{\psi_{\text{CMW}}(t, f_i)\}}{\max(|\text{FFT}\{\psi_{\text{CMW}}(t, f_i)\}|)} \quad (5)$$

where $\Psi(f_i)$ is the normalized Fourier-transformed Morlet wavelet for frequency f_i .

3. Convolution in the Frequency Domain

The convolution of the EEG data with the Morlet wavelet is performed in the frequency domain by multiplying the transformed EEG data with the transformed wavelet:

$$Y(f_i) = X(f) \cdot \Psi(f_i) \quad (6)$$

where $Y(f_i)$ is the result of the convolution in the frequency domain for frequency f_i .

4. Inverse Fourier Transform to Time Domain

The result of the convolution is transformed back to the time domain using the Inverse Fast Fourier Transform (IFFT):

$$y(t, f_i) = \text{IFFT}\{Y(f_i)\} \quad (7)$$

where $y(t, f_i)$ is the convolved EEG signal at time t for frequency f_i .

5. Truncation of Edge Artifacts

Finally, the resulting time-domain signal is truncated to remove edge artifacts introduced by the convolution process:

$$y_{\text{truncated}}(t, f_i) = y(t, f_i)[k + 1 : (n - k)] \quad (8)$$

where k represents the half-length of the kernel (Morlet wavelet), and n is the total length of the signal.

Frequency-Amplitude (FA) Curve

Finally, the frequency-amplitude (FA) curve is calculated as:

$$A(f) = \sqrt{\text{real}(f)^2 + \text{imag}(f)^2} \quad (9)$$

where $A(f)$ is the amplitude of the FA curve at frequency f .

3 Results

3.1 Are These FA Curve Modulations Merely Artifacts?

While interpreting the significance of FA curve modulations, it is crucial to consider the possibility of artifacts. However, several factors suggest that the observed modulations are not artifacts. First, the features of the TFA fabric for O1/2 appear to align best with features in the audio when lagged by 33 ms (invariant to frequency), which is not typical of common EEG artifacts, such as muscle movements or electrical noise. Moreover, the participant remained still during the recordings, minimizing the likelihood of muscle artifacts. Additionally, certain noises and signals were notably absent, possibly when the participant was not attending to them. Features in a song (with which features in the TFA fabric appear to align or possibly correlate) seem to include pitch shifts, such as the bending of a guitar note upwards. Finally, crests that align with and may correlate with drum hits occasionally show a lag relative to other crests; such that not all crests could be amplitude artifacts from the audio itself.

3.2 TFA Fabric Uncovers Hidden Details in Traditional TF Plots

Traditional time-frequency (TF) plots, while useful, often display ‘black holes,’ especially in the low-frequency range, that correspond to the ‘pit points,’ albeit with much lower resolution and visibility. Figure 3 demonstrates this phenomenon. TFA fabric visualization offers a valuable method for examining spectral data over time and may represent a somewhat novel approach in the field, although Pielemeier et. al. [7] also considered a 3D plot of the volume under a time-frequency-amplitude surface to be useful in musical signal analysis. The idea is, instead of using color/luminosity to represent the strength or weakness of a frequency at any given time, the TFA fabric approach uses **displacement in 3D depth** to make amplitude fluctuations more apparent.

3.3 Pit Point Spectral Density

Analysis across all electrodes revealed an average pit point density of 40 points per second (in the 3–40 Hz range), except for electrode Fz, which showed a significantly higher density of 80 points per second (Fz was used to replace T6, which did not reach an acceptable impedance during initialization). The spectral distribution of these points is shown in Figures 4 and 5. These densities are not perfectly reflective of conic shaped pit point densities, as lines of several local minima were sometimes captured.

3.4 Spectral Downshifting

Humans perceive audio in the range of 20–20,000 Hz, yet neurons typically fire within the 2–200 Hz range. This discrepancy suggests a downshifting of frequency during neural encoding, which could explain the mapping of vocal elements to 40 Hz modulations within the SDC spectrum. This downshifting may be indicative of how the brain processes and perceives complex auditory information.

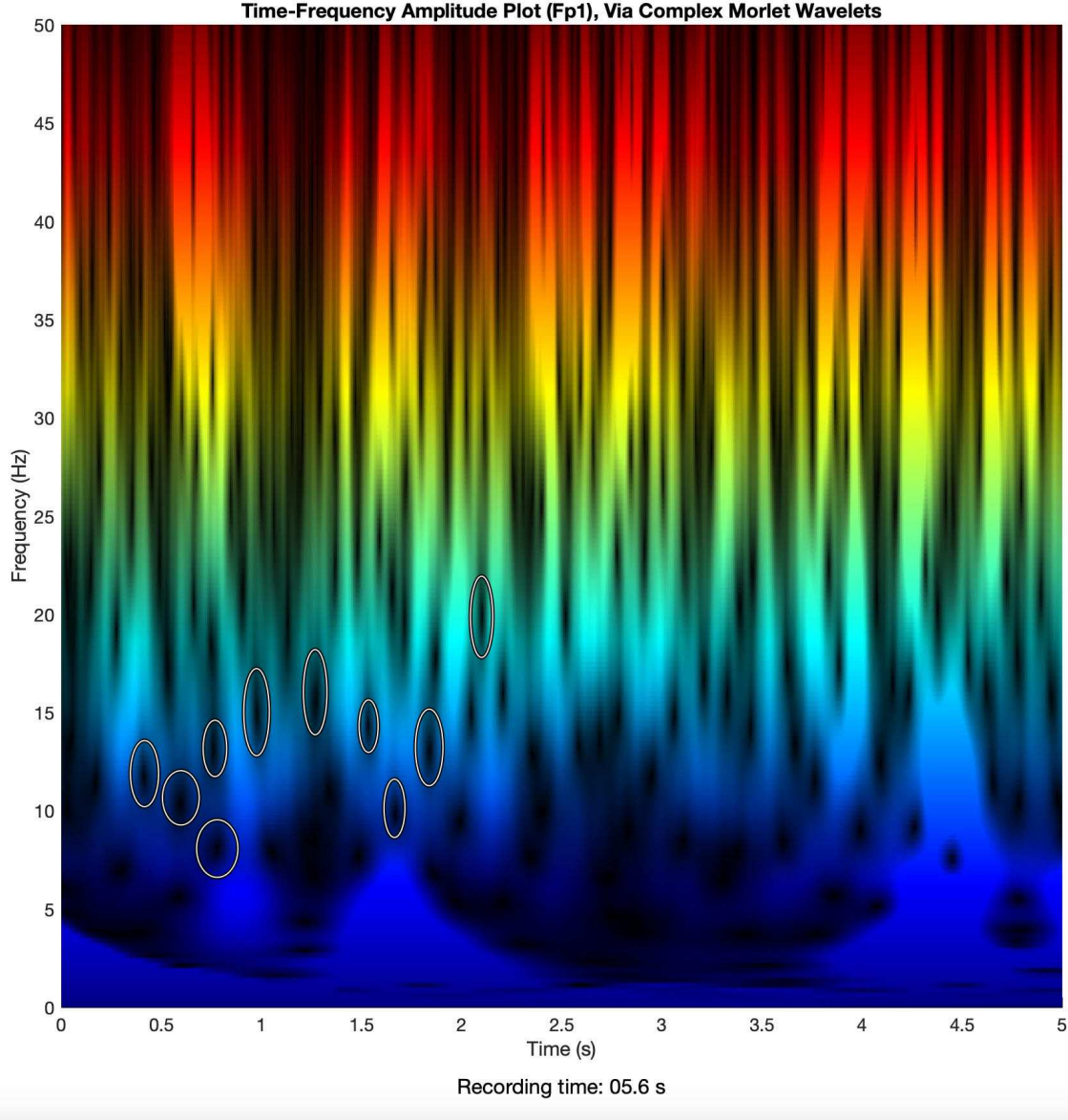


Figure 3: Traditional TF plot showing low-resolution ‘black holes’ (circled) corresponding to pit points.

3.5 Correlation Between Spectral Fabric Morphology and Musical Elements

This section presents the main finding of our study: the morphology of the spectral fabric consistently correlates with elements within the music, revealing a pattern that is both intriguing and significant.

Interestingly, for ‘Use Me,’ the effect was very prominent over the occipital cortex electrodes (O1/2), although strong correlations with the music were also observed at Fz, which seemed to play a more predictive role. For the song ‘Use Me,’ Fz aligned best with the music when placed 33 ms ahead of it, whereas O1 aligned best when placed 33 ms behind the music. In other songs, the timing of the chosen electrode (T5, Fz) seemed to align fine with zero offset.

Several electrodes, including Fp1/2, F3/7/8/4, C3/4, and P3/4, exhibited less sensitivity to the musical stimuli. Interestingly, the spectral fabric at Fz exhibited approximately twice as many folds compared to other electrodes, indicating a heightened level of engagement with the musical stimuli, as shown in Figure 6.

Electrodes Fp1 and Fp2 exhibited strong morphological coupling in their TFA fabrics, with minimal

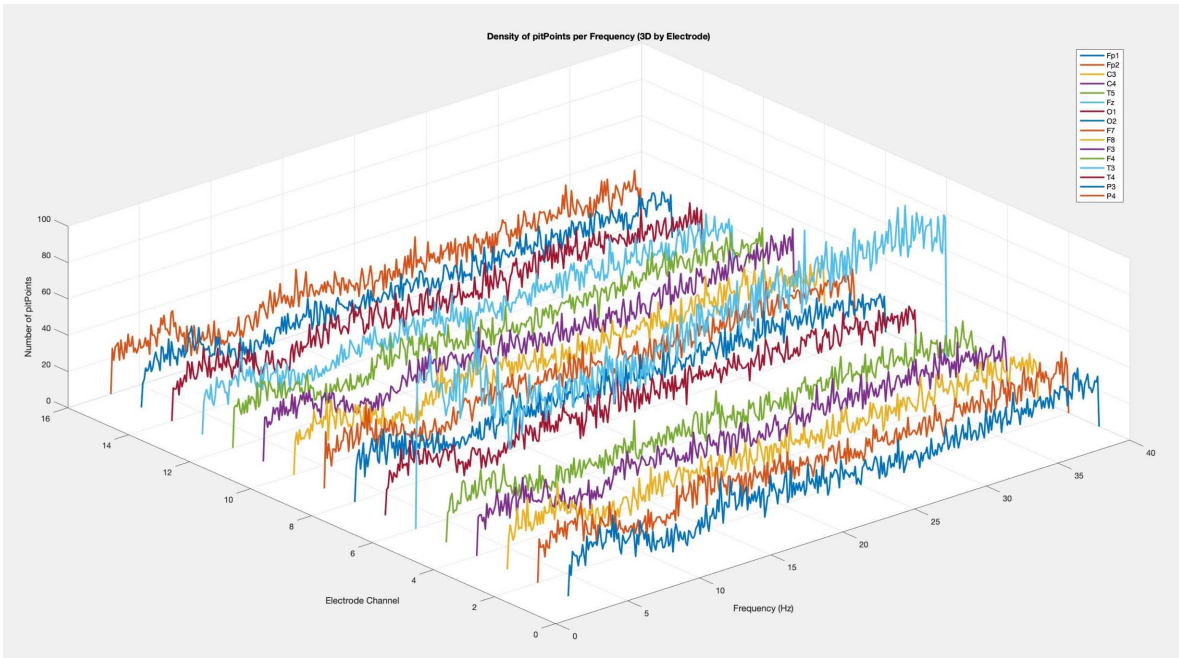


Figure 4: 3D plot of pit point spectral density across electrodes, highlighting the increased density at electrode Fz, over entire 3.5-minute song ‘Use Me’ by PVRIS.

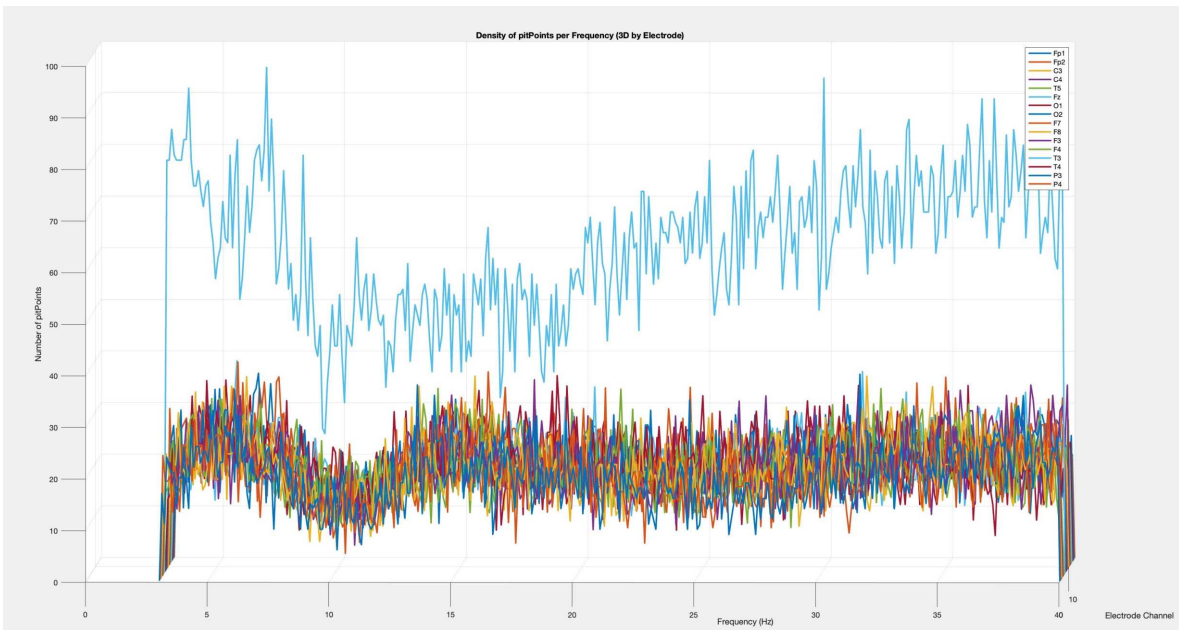


Figure 5: 2D plot of pit point spectral density across electrodes, highlighting the increased density at electrode Fz, over entire 3.5-minute song ‘Use Me’ by PVRIS.

average lag. In the frontal cortex (F7, F3, F4, F8), a similar morphology is often observed, though with a consistent phase lag relative to Fp1/2—approximately 4 ms across all frequencies. This results in a spectrum of phase lags (less for lower frequencies like 2 Hz, more for higher frequencies like 40 Hz). This suggests that the TFA fabric morphology at Fp1/2 is possibly being ‘telegraphed’ to F7/3/4/8 with a slight delay, implying that Fp1/2 may act as the generators of this waveform. Interestingly, Fz does not exhibit this coupling.

Figures 7, 8, 9, 10, and 11 illustrate examples of possible correlation between modulation features of the FA curve and features in the audio.

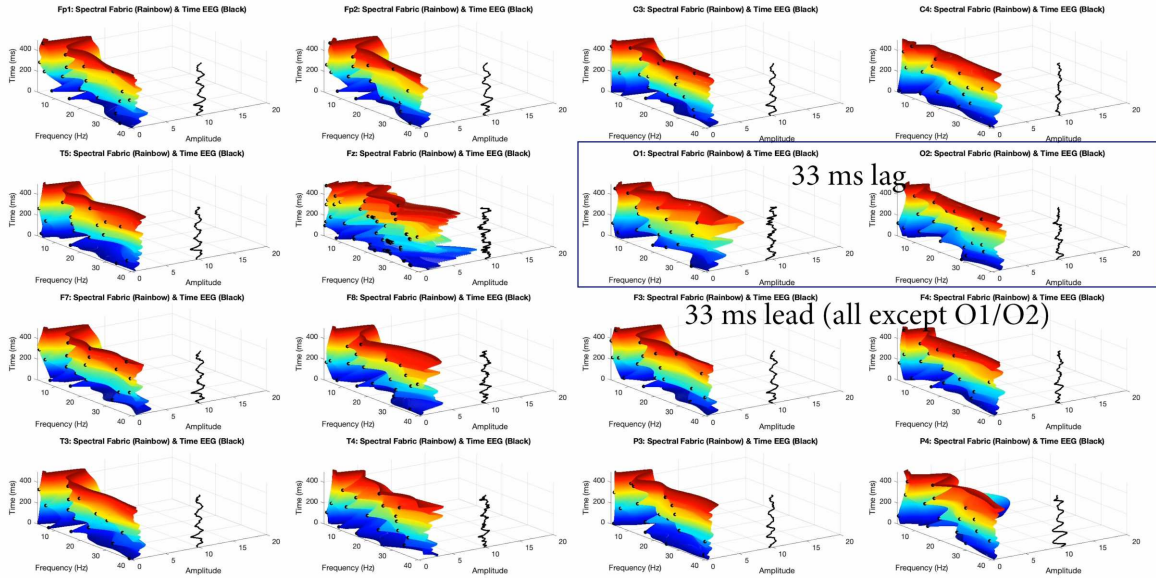


Figure 6: Ongoing FA Curve Modulation at 16 Electrodes While Participant Listened to ‘Use Me’ by PVRIS

4 Discussion

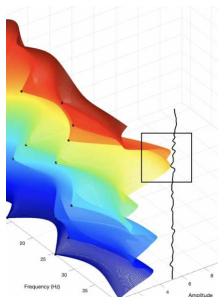
4.1 Excitation/Inhibition Balance

A study by Eisen et al. [8] explores how propofol anesthesia disrupts the brain’s dynamic stability, a balance between excitability and controllability that is critical for consciousness. Using a novel method called Delayed Linear Analysis for Stability Estimation (DeLASE), the researchers quantified changes in dynamic stability within macaque cortex local field potentials (LFPs). Their findings revealed that neural dynamics become more unstable during unconsciousness, with cortical trajectories reflecting patterns consistent with destabilized systems. By increasing inhibitory tone in simulated neural networks, they observed similar destabilization, suggesting that anesthesia-induced unconsciousness is linked to disrupted stability mechanisms within the brain. This may align with the concept of FA curve modulation by spectral-amplitude inhibition processes (spectral-amplitude troughs, and possibly including pit points) and spectral-amplitude excitation processes (spectral-amplitude crests).

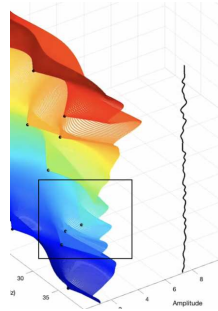
4.2 Morphology Coupling and Lag

As mentioned, Fp1 and Fp2 exhibit a high degree of coupling in the morphology of the fabric, with little to no average lag. There is less coupling or telegraphing of morphology to F7/3/4/8, which shows a 4 ms lag, suggesting possible Granger causality. Coupling with posterior cortical sites occurs even less frequently and with less precision, showing greater lags, possibly ranging from 33 to 66 ms to reach O1/O2. Fp1 and Fp2 seem to demonstrate a sense of solidarity and resolve, as if somewhat aloof to the music or perhaps on guard against other potential threats. Alternatively, Fp1/2 may exhibit a more refined and stable encoding of music, while Fz appears to capture a greater amount of detail but with less focus, potentially reflecting a different mode of processing. Fp1/2 may represent a more controlled, consistent response, while Fz demonstrates a more variable, dynamic engagement with the stimuli.

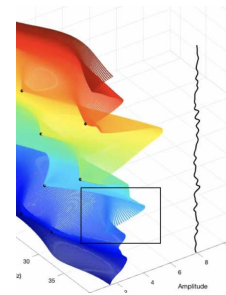
Also, Fp1 and Fp2 showed strong visual coupling of phase-frequency-amplitude helical ramp fabric, both in terms of morphology and phase of fabric; Fp1 and Fz showed little visual coupling (see Figure 12). [Supplemental video S8](#) shows phase-frequency diagrams for a short portion of music for Fp1 (red) vs Fp2 (blue); and [Supplemental video S9](#) shows downsampled phase-frequency diagrams for an entire song for Fp1 (red) vs Fz (blue in this video). [Supplemental video S10](#) shows an entire song with



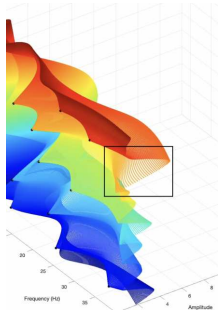
(a) Flam, 11:42 in S1.



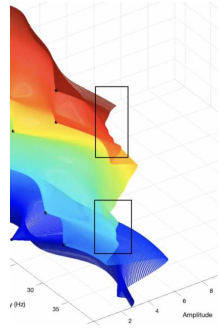
(b) 'Think' cut short, 12:19 in S1.



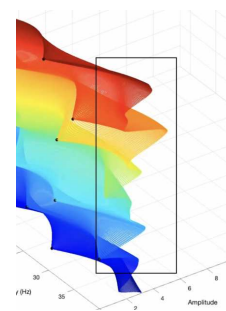
(c) Percussive plosive, 12:31 in S1.



(d) Sigh sound, 15:33 in S1.



(e) Texture of violins/strings, 23:13 in S1.



(f) Melodic plucked tones, 23:16 in S1.

Figure 7: 'Use Me' by PVRIS (electrode O1, lagged an estimated 33 ms [\pm 10 ms] relative to audio)-Instances where features of music might correlate with FA curve modulation: (a) Flam; (b) 'Think' cut short; (c) Percussive plosive (letter 'b'); (d) Sigh exhalation; (e) Texture of violins/strings; (f) Melodic plucked tones.

second-order TFA fabric (amplitudes were summed and normalized across all frequencies, at each time point, and a second operation of complex Morlet wavelets were run to see the TFA fabric).

4.3 High Vs Low Frequency Modulation of the FA Curve

Fz was found to modulate the FA curve at roughly twice (or higher) the rate of other electrodes. Further investigation is needed to confirm that this finding is not due to a faulty electrode on the EEG cap or some other error. If this region of the cortex does indeed behave differently from the rest, it warrants further study. As observed, other brain regions appear to be limited in their ability to modulate the SDC, operating at a relatively low rate (sometimes as low as 10 Hz) in the seemingly critical 30-40+ Hz range of the curve. In other words, these typical brain regions sometimes only/mostly modulate the 30-40 Hz segment of the curve up and down in amplitude ten or twenty times a second (10-20 Hz). This is 10-20 Hz modulation of the amplitude of 30-40 Hz. Fz shows an ability to consistently modulate the amplitude of 30-40 Hz at a rate of around 40 Hz, if the findings are accurate.

4.4 Observations on Possible Correlation between Features in Music and Features in the Modulation of the FA Curve

Caution is warranted in reading too much into the visual patterns seen from the FA curve's modulation; nevertheless, several instances where features appear to possibly correlate have been noted. Further study, including with the use of statistical analysis, is needed to make any solid confirmation.

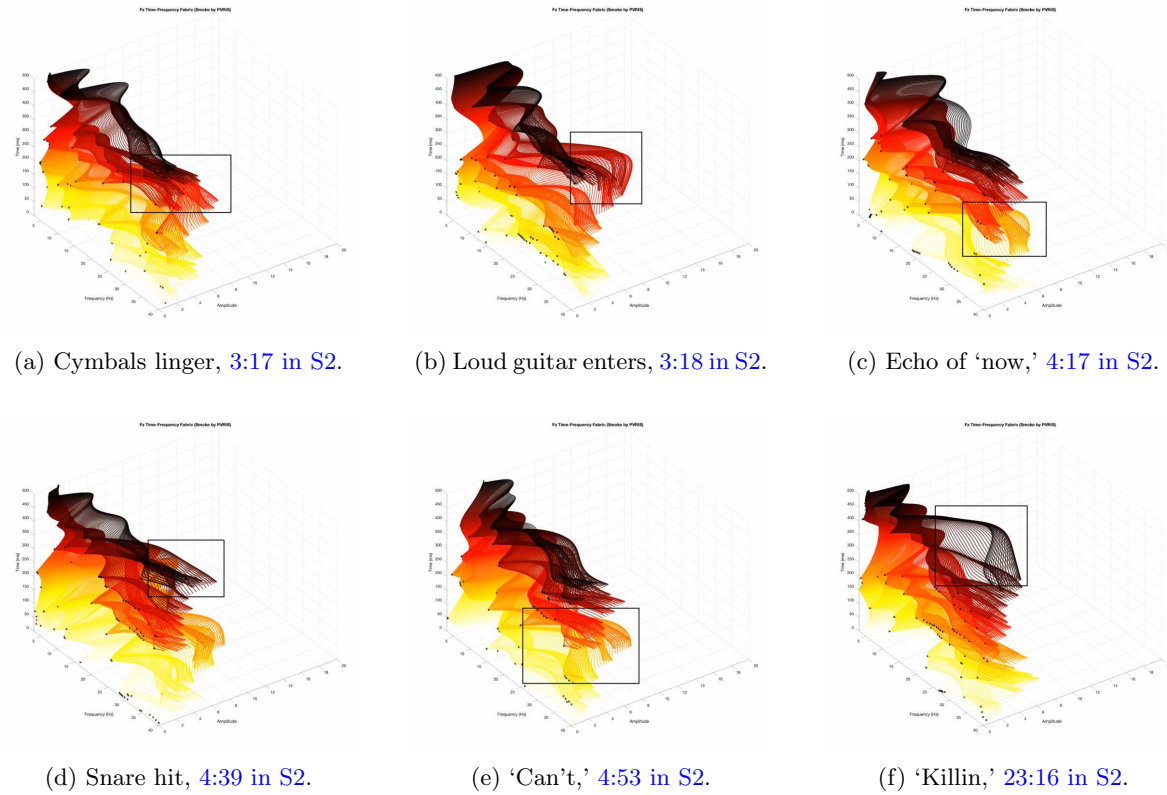


Figure 8: ‘Smoke’ by PVRIS (electrode Fz, no lag)-Instances where features of music might correlate with FA curve modulation: (a) Cymbals linger; (b) Loud guitar enters; (c) Echo of the word ‘now,’ while not loud, is useful to humans, and possibly correlates with a crest in TFA fabric; (d) Snare hit possibly correlates with the inversion of the FA curve waveform; (e) The word ‘can’t,’ while not loud, is useful to humans, and possibly correlates with a crest in TFA fabric; (f) The word ‘killin,’ while not loud, is useful to humans, and possibly correlates with a crest in TFA fabric. Note also that electrode Fz here modulates the amplitude of 20-40 Hz at a much faster rate than other electrodes.

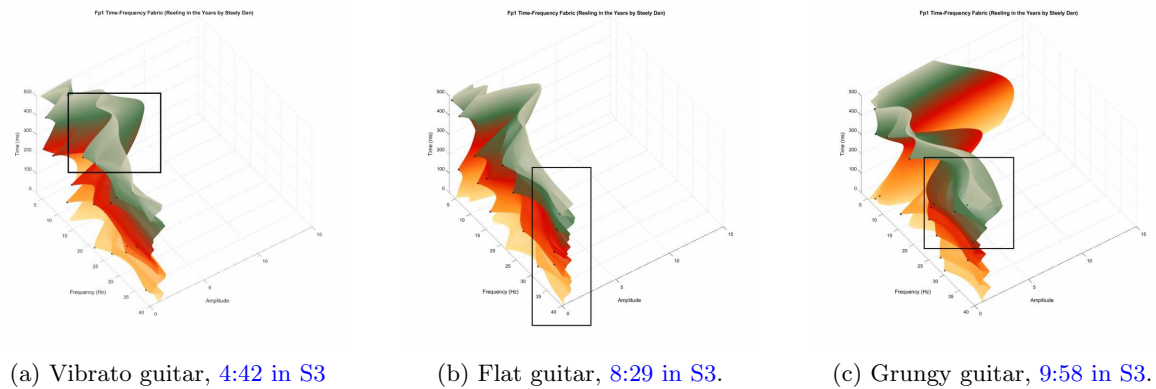


Figure 9: ‘Reeling in the Years’ by Steely Dan (electrode Fp1, no lag)-Instances where features of music might correlate with FA curve modulation: (a) Vibrato of guitar; (b) Flat distorted guitar timbre seems encoded by flatness in SDC; (c) Guitar gets briefly grungy.

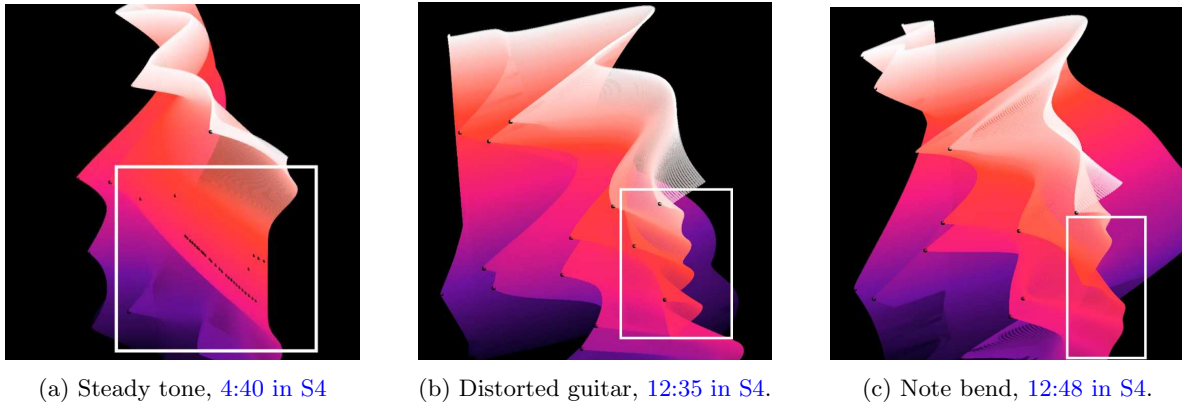


Figure 10: ‘Crystal Ball’ by Styx (electrode T5, no lag)-Instances where features of music might correlate with FA curve modulation: (a) Steady/repeated tone; (b) Distorted guitar; (c) Note bend in guitar solo.

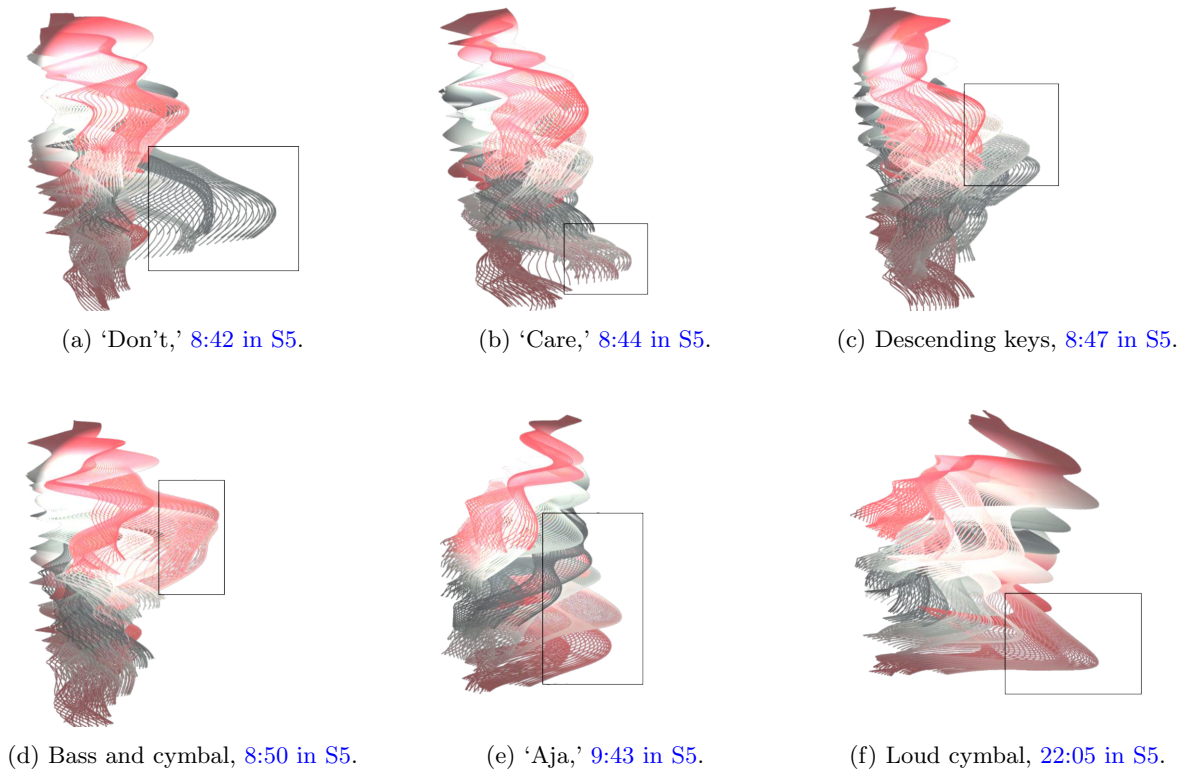


Figure 11: ‘Aja’ by Steely Dan (electrode Fz, no lag)-Instances where features of music might correlate with FA curve modulation: (a) Vocal ‘don’t;’ (b) Vocal ‘care;’ (c) Descending notes on keyboard; (d) Bass note goes up and down in addition to a cymbal hit; (e) Vocal ‘Aja;’ (f) Loud cymbal. Note also that electrode Fz here modulates the amplitude of 20-40 Hz at a much faster rate than other electrodes, making the fabric appear thinner; as at each millisecond, the new line covers a greater distance on average.

5 Further Study

5.1 Investigating Pit Points and Their Purpose

Pit points—spectro-temporally precise inhibitory markers—appear in the analysis of both neural and auditory signals, when using TFA fabric visualization. In neural data, they may take on more of

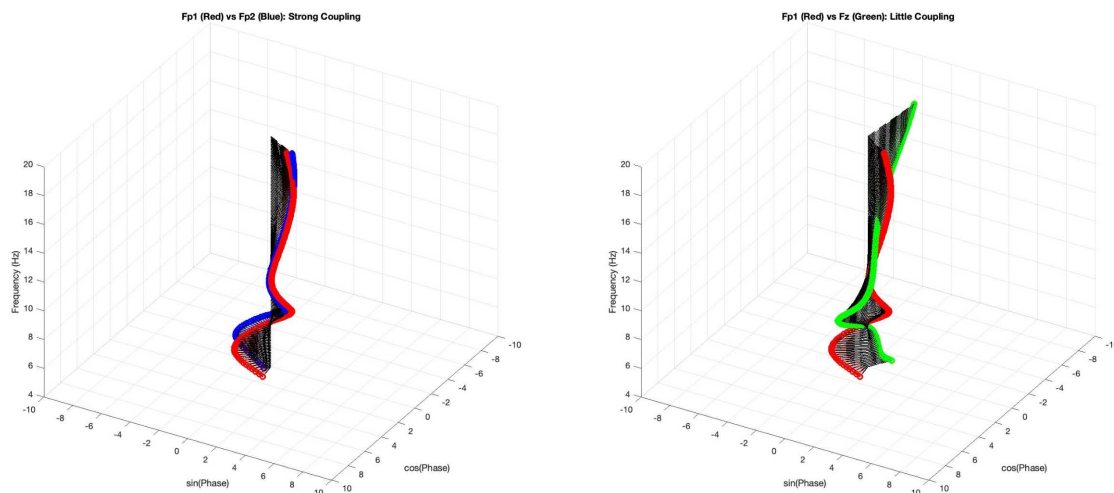


Figure 12: Phase-Frequency-Amplitude helical ramp fabrics for Fp1(red)/Fp2(blue) and Fp1(red)/Fz(green), at the same time point (2.5 seconds in) of the song “Reeling in the Years” by Steely Dan. Note the strong coupling (left) vs. weak coupling (right).

a trough line nature in the time direction at very low frequencies, and a trough line nature in the frequency direction at very high frequencies (30-40+ Hz). In neural data, the frequencies of the pit points appear to show no/little time-locking across frequency, whereas in audio data, multiple frequency pit points may occur at a single time, and recur at a specific and fairly uniform interval.

5.1.1 Generation Mechanisms

Future studies should focus on understanding the underlying neural mechanisms responsible for generating spectral-amplitude excitations and inhibitions. A study using deaf participants could explore whether spectral-amplitude modulation features appear less present or even absent when music is played, to rule out any possibility of these being audio/electrical artifacts.

5.2 Conclusion

The possible correlation of features in the modulation of the FA curve with features in audio has been explored and documented. Further study with statistical methods will be needed for any conclusive determination.

Author declares no competing interest.

References

- [1] Itoh, K., Konoike, N., Nejime, M., et al. (2022). Cerebral cortical processing time is elongated in human brain evolution. *Scientific Reports*, 12, 1103. <https://doi.org/10.1038/s41598-022-05053-w>.
- [2] Liégeois-Chauvel, C., Musolino, A., Badier, J. M., Marquis, P., and Chauvel, P. (1994). Evoked potentials recorded from the auditory cortex in man: Evaluation and topography of the middle latency components. *Electroencephalography and Clinical Neurophysiology Evoked Potentials Section*, 92, 204–214. [https://doi.org/10.1016/0168-5597\(94\)90064-7](https://doi.org/10.1016/0168-5597(94)90064-7).
- [3] Brugge, J. F., et al. (2008). Functional localization of auditory cortical fields of human: Click-train stimulation. *Hearing Research*, 238, 12–24. <https://doi.org/10.1016/j.heares.2007.11.012>.

- [4] Nourski, K. V., et al. (2014). Functional organization of human auditory cortex: Investigation of response latencies through direct recordings. *NeuroImage*, 101, 598–609. <https://doi.org/10.1016/j.neuroimage.2014.07.004>.
- [5] Hillyard, S., Hink, R., Schwent, V., and Picton, T. (1973). Electrical Signs of Selective Attention in the Human Brain. *Science (New York, N.Y.)*, 182, 177-180. <https://doi.org/10.1126/science.182.4108.177>.
- [6] Mike X. Cohen, *Analyzing Neural Time Series Data: Theory and Practice*, MIT Press, 2014.
- [7] W. J. Pielemeier, G. H. Wakefield, and M. H. Simoni, "Time-frequency analysis of musical signals," *Proceedings of the IEEE*, vol. 84, no. 9, pp. 1216-1230, Sept. 1996, doi: 10.1109/5.535242.
- [8] Eisen, A. J., Kozachkov, L., Bastos, A. M., Donoghue, J. A., Mahnke, M. K., Brincat, S. L., Chandra, S., Tauber, J., Brown, E. N., Fiete, I. R., and Miller, E. K. *Propofol anesthesia destabilizes neural dynamics across cortex*, *Neuron*, 2024, vol. 112, no. 16, pp. 2799–2813.e9, Elsevier. DOI: 10.1016/j.neuron.2024.06.011.