UNIVERSITY OF CENTRAL OKLAHOMA

Edmond, Oklahoma

Jackson College of Graduate Studies

Emotional and Neurological Responses to Timbre

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for

the degree of

MASTER OF ARTS IN EXPERIMENTAL PSYCHOLOGY

By

SEPHRA H. SCHEUBER

Moore, Oklahoma

2019

EMOTIONAL AND NEUROLOGICAL RESPONSES TO TIMBRE

A THESIS

APPROVED FOR THE DEPARTMENT OF PSYCHOLOGY

December 9, 2019

M Vanhoy

Dr. Mickie Vanhoy, Committee Chairperson

Mark Hamlin

Dr. Mark Hamlin, Committee Member

Dennis Borycki

Dennis Borycki, Committee Member

Acknowledgments

To all of the people who helped me in my research by asking difficult questions or providing answers to my own difficult questions, thank you.

To all of the friends and family who have encouraged me and loved me through the process, thank you. I would especially like to thank my husband, Matt, who has read many drafts, listened to hours of ramblings, been patient with my distractions, and encouraged me when I needed it most.

Thank you to Ted Mofle, Hannah Whitten, Kurtis Roberts, and all of the others around the lab who were so kind and helpful throughout this process. Thank you so much to Anita Kuchera and Dalton Tindall, whose incredible voice acting made for fantastic stimuli.

Thank you to the UCO Office of High-Impact Practices for funding and support, the College of Education and Professional Studies, the Jackson School of Graduate Studies, and the Psychology Department.

To my thesis advisory committee made up of Dr. Mickie Vanhoy, Dr. Mark Hamlin, and Dennis Borycki: you have each been invaluable sources of encouragement, information, and wisdom to me throughout this process. I cannot thank you all enough for believing in my potential and pushing me to close in on it.

Emotional and Neurological Responses to Timbre

Sephra H. Scheuber

University of Central Oklahoma

Author Note.

Sephra Scheuber, Department of Psychology, University of Central Oklahoma.

This research was supported in part by a grant from the UCO Office of High-Impact Practices.

Correspondence concerning this study should be addressed to Sephra Scheuber, Department of Psychology, University of Central Oklahoma, 100 North University Drive, Edmond, OK 73034.

E-mail: sscheuber@uco.edu

Abstract

Humans detect and react to characteristics of timbre in speech and instrumental sounds, but the relationship between emotions conveyed by timbre in non-verbal vocalizations and those conveyed by electric guitar sounds is unknown. For this study, I created a series of sounds with varying timbre characteristics: non-verbal voice sounds and electric guitar sounds. Sounds were validated through categorization of emotion and ratings of intensity and believability. I found that sounds with low (slow onset) attack slope were most likely to be categorized as angry, while sounds with high (fast onset) attack slope were most likely to be categorized as happy. I collected EEG data from participants while they made judgements on the emotional similarity of guitar sounds (primes) when compared with vocal sounds (targets). I conducted Multifractal Detrended Fluctuation Analysis and MANOVAs, and I found systematic differences of the multifractal spectrum of EEG responses between conditions (emotion and sound type) that would be obscured by other forms of analysis. This information could be applied to development of more effective and psychologically healthy entertainment (music, film, etc.). It is also applicable in therapy situations in which there is a need to induce a certain neurological or emotional state.

Keywords: music therapy, timbre, emotion, sound, audio, EEG, wavelet, guitar, empirical mode decomposition, complexity, multifractal detrended fluctuation analysis, fractal

Table of Contents

Emotional and Neurological Responses to Timbre

People have considered music a communicator of emotion for a long time (Spencer, 1911, p.603-642), although it is usually considered more subjective than linguistic communication. However, listeners often agree regarding emotional content of musical stimuli (Eerola, Ferrer, & Alluri, 2012; Filipic, Tillmann, & Bigand, 2010). Instrumental sounds are used to express emotion, similarly to vocal sounds (Eerola, Ferrer, and Alluri, 2012; Ilie &Thompson, 2006). Many aspects of this expression have been studied: fundamental frequency (or "pitch"), amplitude (or "volume"), timbre (or "tone quality"), etc. However, timbre has been studied considerably less than other aspects of sonic expression.

Timbre has historically been most commonly defined as what it is not rather than what it is. In 1960, the American Standards Association defined timbre as "that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar" (American Standards Association, 1960, p.47). More positive definitions mostly involve different levels of dissection of the measurable elements that make up timbre. Broadly, these include spectral elements, temporal elements, and spectro-temporal elements. Within these broad categories are more specific measurable qualities such as brightness, attack time, spectral centroid, spectral flux, and many others (Hailstone et al., 2009; McAdams & Cunible, 1992; Caclin, McAdams, Smith, & Winsberg, 2005; Griffiths & Warren, 2004). For the purposes of the study presented in this document, timbre will be defined as an emergent characteristic of sound based on the interplay of multiple spectral, temporal, and spectro-temporal traits by which a listener can distinguish differences between sounds that are perceptually the same in pitch, loudness, and duration.

Vocal timbre offers important information about the state of an environment, fostering emotions in the listener. For example, vocal sounds of aggression or repeated loud sounds are generally met with feelings of fear in the listener and a response of the sympathetic nervous system. This response of the sympathetic nervous system includes several physiological changes (Kato et al., 2014). The parasympathetic nervous system (which activates during times of rest) tends to produce inverse physiological changes. Threatening sounds, such as the sound of a predator or an angry human voice, and relaxing sounds, such as the sound of moving water or infant-directed speech, within the natural environment can cause these inverse types of physiological responses (Santesso, Schmidt, & Trainor, 2007); instrumental sounds might cause similar responses, depending upon their timbral characteristics.

Previous studies of timbre have focused on classical instruments. This makes sense, as composers of classical styles use timbre as a primary expressive tool in the choice of instruments (sometimes called “arrangement”) for a given piece. However, this has presented problems for the study of timbre. Manipulations of timbre in recordings of classical instruments do not necessarily have the expected effect on the listener. This might be because listeners have a fixed template for the given instrument and ignore the deviations from that template (Eerola, Ferrer, & Alluri, 2012). In contrast, timbre is so frequently and drastically manipulated in electric guitar that it is unlikely that listeners would have a strict template for the timbre they expect from this instrument. The use of the electric guitar also allows for a more controlled study of timbre, since only one instrument is being used. The electric guitar is a common instrument used in various settings, so this knowledge could benefit multiple communities (audio engineers and producers, sound designers for film and video games, music therapists, musicians, etc.). This instrument’s purposeful timbre manipulation in many commercial recordings and performances is used to express genre, style, and emotion; this is comparable to the way that vocal sounds are manipulated by speakers to convey emotional content to the listener. In fact, several devices have been invented to create the sound of a voice in conjunction with an electric guitar (the “talk box,” “singing guitar,” etc.). The common use of the electric guitar in popular recordings and the frequent tone manipulation makes this instrument an ideal choice for the study and application of effects of timbre.

Body size projection theory suggests that vocal timbre is a tool for communicating body size and, consequently, a tool for communicating threat level (Ohala, 1984; Xu, Lee, Wing-Li, Liu, & Birkholz, 2013; Morton, 1977). This theory suggests that sounds with a more “pure tone” (less noise-like, more singular in frequency content) are used to indicate a submissiveness/lack of threat; it also suggests that sounds that are “rougher” (more noise-like, less singular in frequency content) are used to indicate presence of threat. Evidence for this theory can be observed in everyday life; for example, birds and human infants tend to have more “pure” timbre characteristics in their communication sounds, while lions and bears tend to have more “rough” timbre characteristics (growling, for example). Morton specifically suggests that animals use lower frequencies and more harmonics (“harshness”) to suggest larger body size and communicate aggression, similar to the way that animals might use fur or feathers to suggest larger body size (1977). Morton created a list of what he called “motivation-structural rules” (MS) around these observations. These physiological mechanisms may be related to the purpose music plays in our society. Emotional responses to auditory stimuli are tied to these physiological responses and are likely to incur related changes in response to musical stimuli.

Instrumental sounds are used to express emotion, similarly to vocal sounds. The ratio of high-frequency to low-frequency energy in a sound is negatively correlated with the valence rating of the listener, and energetic sounds tend to have fast attacks and more high-frequency energy than low-frequency energy (Eerola, Ferrer, & Alluri, 2012). Listeners have different valence ratings dependent on pitch in both music and speech, different tense arousal ratings dependent on intensity of sounds in both music and speech, as well as other main effects (Ilie & Thompson, 2006). Arousal and valence of non-linguistic sounds have even been shown to have a measurable effect on electrodermal activity (EDA) of listeners (Greco, Valenza, Citi, & Scilingo, 2017; Tajadura-Jiménez et al., 2010). Taken together, these results indicate the possibility that humans might have emotional and physiological reactions to certain sonic timbres across cultures.

Timbre of isolated instrument sounds can convey emotional information in a similar way to emotional speech prosody. In one study, participants judged emotional categories (sad, happy, angry, or neutral) and intensity levels of instrument sounds based on differences in timbre. Different classical instruments were used for production of different timbres. Participants were then asked to judge similarity between primes (instrument sounds) and targets (emotional speech sounds) while EEG data was recorded. People produced a significantly larger N400 ERP response at the Cz electrode location when the stimulus pair was incongruent. For example, with angry targets, there was a stronger N400 when the prime was happy or sad than when the prime was angry (Liu, Xu, Alter, & Tuomainen, 2018). These results also supported the size code theory (Ohala, 1984), in that sounds and speech that were perceived as angry were found to be "rough" in timbre (higher attack slope, higher spectral centroid, higher high-frequency to low-frequency ratio, and lower spectral flux), while happy sounds and speech were "pure" in timbre (lower attack slope, lower spectral centroid, lower high-frequency to low-frequency ratio, and higher spectral flux) (Liu et al., 2018). While this research is extremely valuable, it again uses different classical instruments as the way to control timbre variation. This makes sense for use in traditional music composition, which primarily uses arrangement of instrument entrance and exit times as an overarching timbre control to affect emotion, but there is an additional level of knowledge that would be very useful in the contemporary music industry. In the contemporary industry, direct tone manipulation of certain instruments (electric guitar being arguably the most common other than the voice) is one of the primary ways of controlling the emotional impact of a song on the listener.

While responses to timbre could be measured in various ways, I decided to measure responses via electroencephalography (EEG) due to the speed at which an EEG signal can fluctuate and be captured. Recent research has indicated that neurological data is multifractal in its dimension. While a line is one-dimensional, a square is two-dimensional, and an object such as a pen is three-dimensional, some things have fractional dimensions. For example, the dimension of a Sierpinski triangle is between one and two, about 1.585 (Figure 1). Other things, however, have multiple fractional dimensions: these are called multifractals. Multifractals require multiple fractal exponents to represent their dimension. Much of the research considering psychophysiological responses to sound lacks analysis styles that take the inherent complexity of EEG signals into account. While there has been some research done in this area (Maity et al., 2015), EEG responses to non-linguistic affective vocal sounds and timbral manipulations of electric guitar have yet to be studied in this way to the researcher's knowledge.

The most common way to analyze EEG in response to given stimuli is through ERPs (event-related potentials). While this type of analysis offers valuable insight, time-series analysis techniques offer a different perspective on EEG data. Time series analysis techniques take into account the lack of independence in EEG data points. Unlike some other data types, in EEG data (and many other types of physiological data) each data point is related to the point before and after it; the data is nonstationary. This type of data also often has self-similarity at different scales; it is multifractal. Time series approaches to EEG analysis are a good alternative to a more traditional wave analysis like Fourier Transforms, which are only accurate if a waveform is linear and stationary.

A sound waveform or EEG waveform looks random if you try to model it with a line. A sine wave is a better idea, but not as good an idea as a wavelet, or better still, a mathematical tool called detrended fluctuation analysis. What it does is detect patterns that are distributed across multiple scales. These patterns are not whole numbered dimensional patterns like 1D or 2D objects. These objects have fractional dimensions. For example, a Sierpinski triangle has a fractional dimension of approximately 1.585. If you could shrink yourself down like an ant to stand on it, it would be self-similar in every direction at every scale. It is a perfect mathematical object.

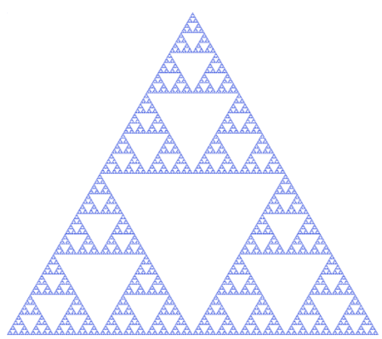

Figure 1. A Sierpinski triangle, which has a dimension of 1.585. This is an example of a monofractal. Image by Beojan Stanislaus.

Objects in nature are not perfect like this, however, because physical objects cannot be infinite in three-dimensional space. The same fractional dimension is not seen throughout the entire Romanesco broccoli flower, but the fractional dimension varies a bit according to the complexity of the physical system. A system that has only one fractional dimension is monofractal. A system that is not perfectly the same in every direction at all scales but still has fractional dimension is multifractal. That system can be described with detrended fluctuation analysis better than it can be described by other means.

Figure 2. One example of a multifractal object, Romanesco broccoli. Photo by Artiom Vallat on Unsplash.

Figure 3. Another example of a multifractal object, lightning. Photo by Mélody P on Unsplash

Prior to analysis of EEG data, it is generally necessary to smooth the data for removal of noise in the signal. Empirical Mode Decomposition (EMD) is one such method, which smooths the data by sifting it into Intrinsic Mode Functions (IMFs). EMD is a smoothing method for nonlinear data that uses local minima and maxima of the waveforms to sift out different pieces that make up the waveform. These pieces are separated based on the energy in the signal at different time scales and different frequencies. This produces a number of largely orthogonal IMFs that make up the original signal. This then allows for certain noise components to be extracted and removed from the waveform.

It is also often necessary to extract certain features/rhythms of waveforms in order to analyze them separately (alpha, beta, delta, theta, etc.). The Wavelet Transform (WT) is a method for this process that takes into account the nonlinearity of EEG data. This involves choosing a wavelet and comparing it to one portion of the signal at a time. This allows for multiresolution analysis in frequency, amplitude, and timing of the waveform, leading to a more accurate deconstruction of the frequency components of the waveform than a more traditional analysis style such as a Fourier Transform.

After these components are extracted, the EEG signals need to be analyzed, in this case using Multifractal Detrended Fluctuation Analysis (MFDFA) due to the multifractal nature of the data. When graphing the multifractal spectrum, Dq is the q-order dimension; hq is the q-order singularity exponent (also known as the Hölder exponent, or α) (Figure 4). A higher Dq value indicates a higher dimension and a higher level of chaos within the system. A monofractal such as the Sierpinski triangle (Figure 13) would only need a single point along the y-axis (Dq) at 1.585 to represent its dimension, while a multifractal needs multiple points along the y-axis because it has multiple fractional dimensions. A higher hq value indicates higher persistence of the multifractal spectrum. A persistent signal is one that continues in the same direction from one data point to the next, while an anti-persistent signal is one that changes direction from one data point to the next.

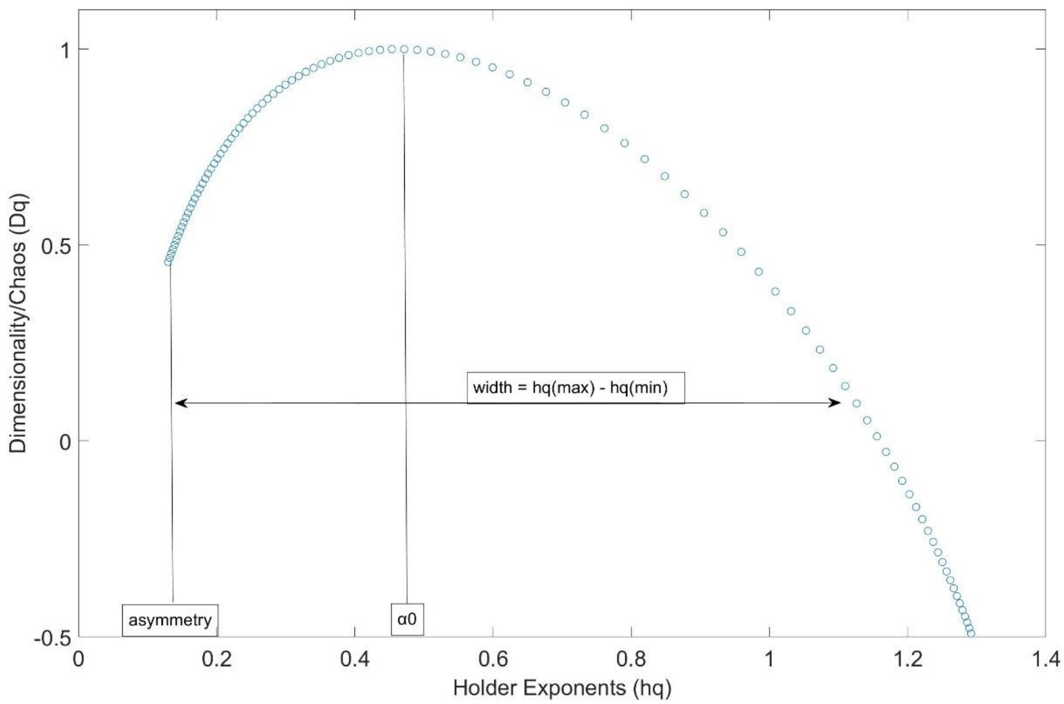

Figure 4. An example graph of the multifractal spectrum of one participant while listening to a guitar sound categorized as angry, with parameters of width, α_0 (hq at the curve's apex, where q=0), and asymmetry marked.

Parameters of the multifractal spectrum that were analyzed are somewhat new to neurological study, and have not yet been studied in regard to responses to sound to the researcher's knowledge. There is little consensus on what they mean for neural dynamics and overall mental and emotional health. These parameters can be discussed in general terms, but further study is needed in understanding the practical meanings behind them. As asymmetry of the multifractal spectrum decreases, extreme neurological events are more frequent and play a more important role in the signal. α_0 of the multifractal spectrum (the value of hq at the apex of the spectrum) is representative of the persistence of the multifractal spectra over time, or its regularity. As α_0 goes up, the signal is more persistent (more likely to go in the same direction from one data point to the next rather than change direction). Width is the range of the Holder exponents (maximum of hq-minimum of hq), and it is the characteristic most commonly addressed in current EEG research regarding multifractal spectra. Multifractal spectrum width indicates the level of self-similarity in the signal (sometimes called complexity or "richness").

Neural complexity has been shown to be high for individuals with schizophrenia when compared to controls, and severity of schizophrenia symptoms appears to be associated with complexity. After treatment, the complexity of neurological signals for those with schizophrenia decreases in the fronto-temporal regions of the brain (Sokunbi et al., 2014). Those suffering with depression also have higher neural complexity, with a decrease post-treatment (Li et al., 2008; Méndez et al., 2012). However, not all pathological states share this pattern. People with post-stroke depression have lower overall neural complexity than controls (Zhang et al., 2015), and there is conflicting evidence for whether those with Alzheimer's have higher or lower complexity than controls (van Cappellen van Walsum et al., 2003; Grieder, Wang, Dierks, Wahlund, & Jann, 2018).

During timbre judgements, people show maximal N400 ERP (Event-Related Potential) in the mid-central posterior region of the brain, which includes the Cz electrode site. A graphic of the neurological signals at the Cz electrode was used to show the typical N400 responses to differing emotions for this recent study (Liu et al., 2018). People listening to a drone that subtly changes in timbre over time show maximal increase in multifractality of alpha brain waves at the frontal midline (Fz) electrode (Maity et al., 2015). Listening to music is associated with changes in alpha and theta brain wave rhythms (Kabuto, Kageyama, & Nitta, 1993; Maity et al., 2015). It is possible that central electrodes show especially prominent alterations during sound processing due to corpus callosum involvement, as musicians have a larger anterior corpus callosum than non-musicians (Schlaug, Jäncke, Huang, Staiger, & Steinmetz, 1995).

In this study, I test the research hypothesis that there are systematic differences between conditions (emotion and sound type of sound being heard) that would be obscured by other forms of analysis. More specifically: (1) Emotional categorizations of electric guitar sounds and non-linguistic vocal sounds differ as a function of alterations in timbre. (2) The multifractal spectrum of EEG responses (width, apex, and asymmetry) differs as a function of the emotional category of a sound.

Method

Participants

Sixteen students with normal hearing abilities (as defined by UCO's Speech and Hearing Clinic) from the University of Central Oklahoma participated in the experiment. People who participated in the stimuli validation were not allowed to participate in the EEG experiment. Due to poor electrode connection, data from six participants was not used, leaving ten data sets for analysis.

Apparatus

Appropriate audio files had to be created for use as stimuli prior to EEG data collection. Two types of stimuli were created: non-linguistic affective vocalizations and electric guitar sounds with timbre alterations.

The emotions portrayed by the vocalizations were anger, happiness, and sadness. Twelve ten-second examples of each emotion were recorded, six performed by a female voice actor and six by a male voice actor. They were recorded using Pro Tools audio recording software (Avid, 2018). These stimuli were normalized for perceptual loudness and precise length.

Twenty-eight ten-second electric guitar sounds were recorded with variation in timbral characteristics using a guitar, guitar amplifier, and guitar pedals. These were also recorded within Pro Tools audio recording software (Avid, 2018). All guitar sounds were recorded playing the D#4 note for the sake of comparability with recent research and because this is the mean note found to be present in a large collection of written musical scores (Liu et al., 2018; Eerola, Ferrer, & Alluri, 2012; Huron, 2001). These sounds were then normalized for perceptual loudness and length to match that of the vocal sounds.

Ten students with normal hearing abilities (as defined by the UCO Speech and Hearing Clinic) from the University of Central Oklahoma participated in the validation process. After reading and signing a consent form, students were seated in front of a computer and given Audio Technica ATH-M40x headphones due to their excellent frequency reproduction ability and sound isolation (Audio Technica, 2018). People heard each sound in a randomized order, and they were asked to answer questions after each sound. The first question was regarding emotion categorization (anger, happiness, sadness, or other/none), the second question was regarding the emotion intensity on a 5-point scale (1= *very weak* to 5= *very strong*), and the third was regarding how believable the emotion was on a 5-point scale (1= *not believable at all* to 5 = *very believable*). The believability rating was only collected for the vocal sounds.

Responses regarding emotion categorization agreed on all vocal sounds at 80% or above. The two most believable and intense non-linguistic vocal sounds as judged by the listeners for each emotion and each sex were chosen for use in the EEG experiment (four total vocal sounds for each emotion, two per sex). While there were trends among the participant responses to the guitar sounds, they were much more subtle. Entropy was calculated for emotion categorization responses for each of the guitar sounds, and the four guitar sounds with the lowest entropy for each emotional category were chosen for use in the experiment (Table B1). Twenty-four sounds were chosen, twelve per sound type (guitar and voice).

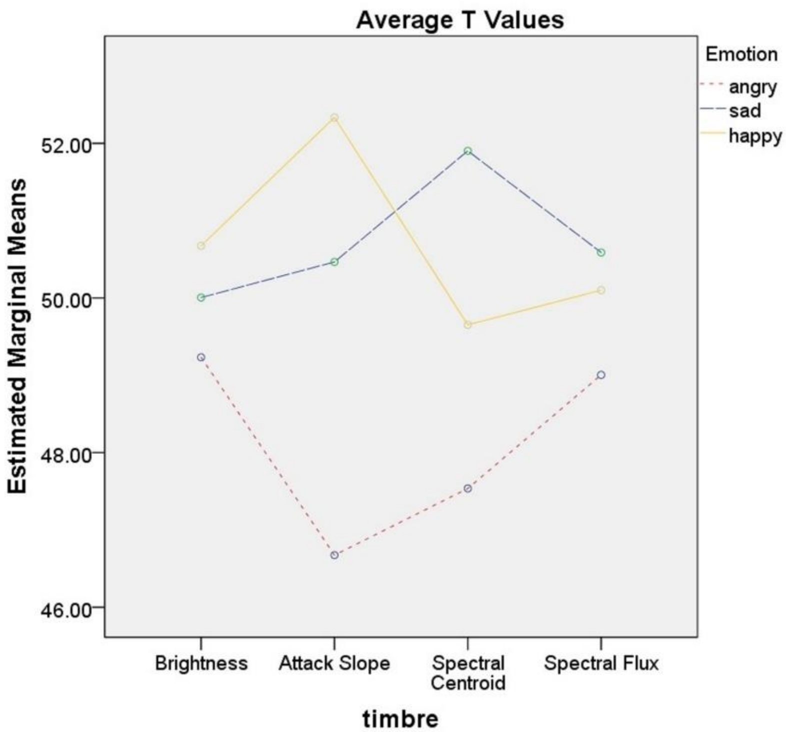

Figure 5. Comparing means of standardized scores of each of the timbre characteristics for different emotions.

After initial sonic analysis in using MIRtoolbox (Lartillot, 2014; Lartillot & Toiviainen, 2007) in MATLAB (MATLAB, 2014), a multivariate analysis of variance (MANOVA) was run comparing emotion categories for vocalizations and guitar sounds on standardized scores (T scores) for each of the four timbral characteristics (brightness, attack slope, spectral centroid, and spectral flux). The results of this MANOVA show a significant effect of emotion on the dependent variables (DVs, timbre characteristics) (F (8, 98) = 3.812, p = .001; Wilks' Λ = 0.582), while the effect of sound (guitar vs. voice) (F (4, 49) = 0.453, p = .770; Wilks' Λ = 0.964), and the interaction of sound and emotion (F (8, 98) = 1.602, p = .134; Wilks' Λ = 0.782) were non-significant (Table B2). Emotional categorizations of electric guitar sounds and non-linguistic vocal sounds differ reliably, depending on their timbre characteristics. Further inspection of analysis results show that emotion categorization only has a significant relationship with the timbre characteristic of attack slope (F (2, 52) = 10.224, p < .001). Specifically, angry sounds have a low (slow) attack slope. Emotion categorization did not have a significant relationship with brightness (F (2, 52) = 0.406, p = .668), centroid (F (2, 52) = 0.674, p = .514), or flux (F (2, 52) = 1.093, p = .343) (Table B3).

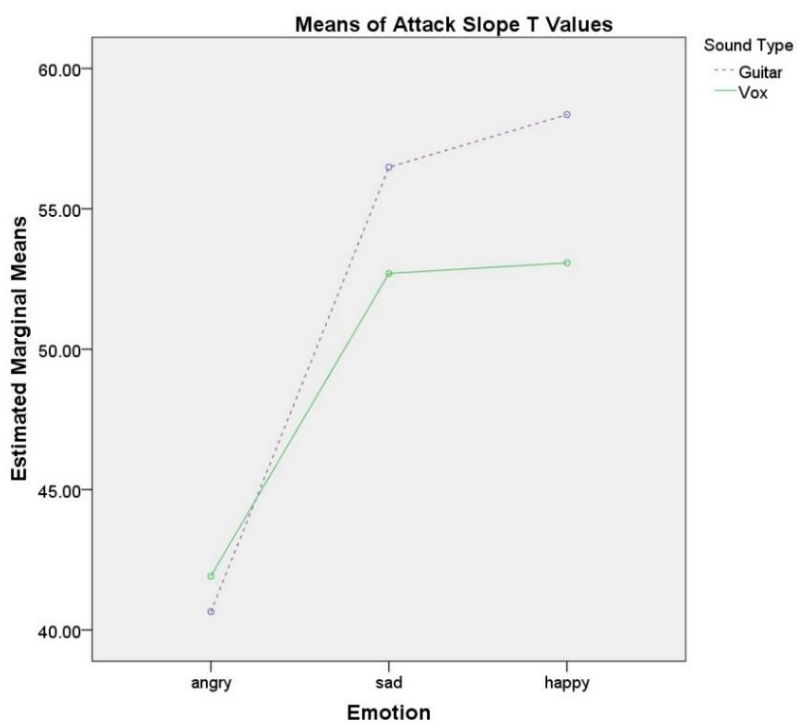

Figure 6. Standardized Attack Slope means for guitar and voice on different emotional categories.

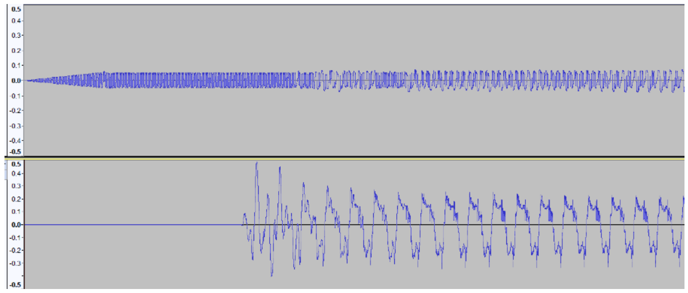

Figure 7. Attack portions of waveforms of a guitar sound that was judged to be “angry” (top) and one that was judged to be “happy” (bottom).

BioPac’s MP150, Biopac Student Lab Professional Version 3.7.7, and an electrode cap were used to collect EEG data while participants listened to stimuli on professional-grade Audio Technica ATH-M40x headphones (Biopac Systems, Inc., 2010; Electro-Cap International, Inc.; Audio Technica, 2018). Responses were recorded using PEBL (Mueller & Piper, 2014; Mueller, 2012).

Procedure

Participants were taken to one of the rooms in the UCO Psychology Lab and instructed to sit comfortably in a chair with feet flat on the floor and their hands resting on their legs after reading and signing a consent form (Appendix A). Participants were fitted with an electrode cap, and one reference electrode was placed on the mastoid bone behind the right ear. Gentle abrasion was done at each electrode site, and electrodes were filled with conductive paste. I recorded signal at Cz and Fz according to the 10:20 method, with one ground. After two minutes of resting time, the participants were asked to sit as still as possible and given two minutes to practice and familiarize themselves with the experiment task before the experiment began. Participants were tested one at a time in a quiet laboratory setting with an ambient noise background of less than 50 decibels SPL (sound pressure level). Each trial consisted of a pair of sounds (pairs were randomized): a guitar sound (prime), one thousand milliseconds (1000 ms) of silence, then a vocal sound (target). Participants then had one thousand milliseconds (1000 ms) to indicate whether the sounds were congruent (similar) via a left mouse click or incongruent (not similar) via a right mouse click. There were nine possible types of stimulus pairs (for example: angry guitar and angry voice, or angry guitar and sad voice), each with four possible guitar sounds and four possible vocal sounds. The experiment consisted of four ten-minute sessions (27 trials each), with short breaks (two to five minutes) between each session for rest and movement. EEG data was recorded at 250 samples per second, resulting in collection of over one million data points collected for each participant.

Data Denoising and Extraction

I divided EEG data into chunks for each participant to isolate data that was recorded while they listened to sounds of a given emotion. After this, I averaged data for all trials of a given emotion (angry, happy, or sad) by electrode (Cz or Fz) and sound type (guitar or voice) for each participant. Then the data was formatted. The time column had to be formatted to “Time”, specifically the format shown as “37:30:55” in Excel. After this, I selected “Custom,” where I typed “.000” after “ss” to retain the highest resolution. After formatting, the data was imported into MATLAB, where the Signal Multiresolution Analyzer app was used to conduct Empirical Mode Decomposition (EMD) for denoising. EMD involves breaking a signal down into components called Intrinsic Mode Functions (IMFs). IMFs containing lower frequencies in which most blink artifacts reside were discarded (<3.5 Hz), and the remaining IMFs were used to reconstruct the signal. The same app was then used to conduct the Wavelet Transformation (WT) for extraction of alpha and theta rhythms. The db4 (Daubechies) wavelet was used to break the signal down, after which the alpha range (~8-12 Hz) and theta range (~4-8 Hz) were extracted and exported as separate MATLAB objects (see Appendix D for code samples).

Results

Multifractal Detrended Fluctuation Analysis was conducted on each object using the toolbox provided by Espen Ihlen (2012). I used q (magnification) values from negative five to positive five, and I used a scale maximum of 250, a little less than one-tenth of the number of samples in each chunk of data (2750 samples in each chunk). I calculated the width of the multifractal spectrum, the α_0 (apex) of the spectrum (the value of hq where q = 0 and Dq=1), and the asymmetry (sometimes referred to as α_s) of the spectrum for each (see Appendix D for more code details).

I ran four repeated measures MANOVAs with width, α_0, and asymmetry as dependent variables (DVs) and sound type and emotion as independent variables (IVs) (Table B4; Table B5; Table B6; Table B7). Four MANOVAs were run to allow for these comparisons on Cz and Fz electrodes and on alpha and theta brain wave rhythms (one for alpha rhythms from the Cz electrode, one for theta rhythms from the Cz electrode, one for alpha rhythms from the Fz electrode, and one for theta rhythms from the Fz electrode). This was done so that the effects of sound type and emotion might be seen on each of these separately, as they are separate channels of output data.

Emotion was the only significant effect for both alpha wave rhythms at the Cz electrode (F (6, 32) = 4.275, p = .003; Wilks' Λ = 0.308, partial η^2 = .445) (Table 1) and theta wave rhythms at the Cz electrode (F (6, 32) = 3.687, p = .007; Wilks' Λ = 0.350, partial η^2 = .409) (Table 2). Happy sounds induce higher α_0 than sad and angry sounds at the Cz electrode for both alpha and theta rhythms (Figure 8; Figure 9). The widths of the multifractal spectra differ for alpha at the Cz electrode, with responses to angry sounds having the highest width, followed by sad sounds, and happy sounds having the lowest width. Asymmetry of the multifractal spectra also differ for alpha at the Cz electrode, with the same pattern (A>S>H).

Table 1

Results of MANOVA at electrode site Cz on brain wave rhythm alpha with multifractal spectrum characteristics (width, α_0, asymmetry) as DVs.

Within Subjects Effect	*Wilks' λ*	*F*	*Hypothesis df*	*Error df*	*Sig.*	*Partial η^2*	*Observed Power*
emotion	.308	4.275	6.000	32.000	.003*	.445	.953
sound	.909	.234	3.000	7.000	.870	.091	.077
emotion * sound	.731	.903	6.000	32.000	.505	.145	.303

Note. * denotes significant p values at $\alpha = .05$.

Table 2

Results of MANOVA at electrode site Cz on brain wave rhythm theta with multifractal spectrum characteristics (width, α_0, asymmetry) as DVs.

Within Subjects Effect	*Wilks' λ*	*F*	*Hypothesis df*	*Error df*	*Sig.*	*Partial η^2*	*Observed Power*
emotion	.350	3.687	6.000	32.000	.007*	.409	.915
sound	.886	.300	3.000	7.000	.825	.114	.084
emotion * sound	.616	1.462	6.000	32.000	.223	.215	.485

Note. * denotes significant p values at $\alpha = .05$.

The results at the Fz electrode are more varied, with a significant effect of only sound type on alpha dependent variables (F (3, 7) = 5.348, p = .031; Wilks' Λ = 0.304, partial η^2 = .696) (Table 3) and a significant interaction of emotion and sound on theta dependent variables (F (6, 32) = 3.840, p = .005; Wilks' Λ = 0.338, partial η^2 = .419) (Table 4). Guitar sounds have a higher α_0 than vocal sounds for alpha waves at Fz (Figure 10). Angry guitar sounds elicit the highest α_0 for theta waves at Fz, followed by happy sounds, then followed by sad sounds (Figure 11). The multifractal spectra for responses to vocal sounds for theta at Fz are very similar across emotions, with responses to sadness having the lowest width and responses to happiness having the highest width (Figure 12).

Table 3

Results of MANOVA at electrode site Fz on brain wave rhythm alpha with multifractal spectrum characteristics (width, α_0, asymmetry) as DVs.

Within Subjects Effect	*Wilks' λ*	*F*	*Hypothesis df*	*Error df*	*Sig.*	*Partial η^2*	*Observed Power*
emotion	.563	1.775	6.000	32.000	.136	.250	.579
sound	.304	5.348	3.000	7.000	.031*	.696	.718
emotion * sound	.640	1.333[c]	6.000	32.000	.271	.200	.444

Note. * denotes significant *p* values at $\alpha = .05$.

Table 4

Results of MANOVA at electrode site Fz on brain wave rhythm theta with multifractal spectrum characteristics (width, α_0, asymmetry) as DVs.

Within Subjects Effect	*Wilks' λ*	*F*	*Hypothesis df*	*Error df*	*Sig.*	*Partial η^2*	*Observed Power*
emotion	.429	2.813	6.000	32.000	.026*	.345	.811
sound	.140	14.357	3.000	7.000	.002*	.860	.989
emotion * sound	.338	3.840	6.000	32.000	.005*	.419	.927

Note. * denotes significant *p* values at $\alpha = .05$.

A conservative correction for familywise error would mean setting significance at α = 0.0125 (.05/4). This would leave all results previously mentioned still significant except the effect of sound type on Fz alpha and the main effect of emotion on Fz theta (though the effect of the interaction remains).

After each participant's multifractal spectra were calculated for each type of sound, graphs were made for individual participants (Appendix C). I then averaged hq and Dq values for all participants for each sound type, electrode, and brain wave rhythm. Then, graphs were created by plotting these hq averages by the Dq averages for each significant comparison (Figures 8-12).

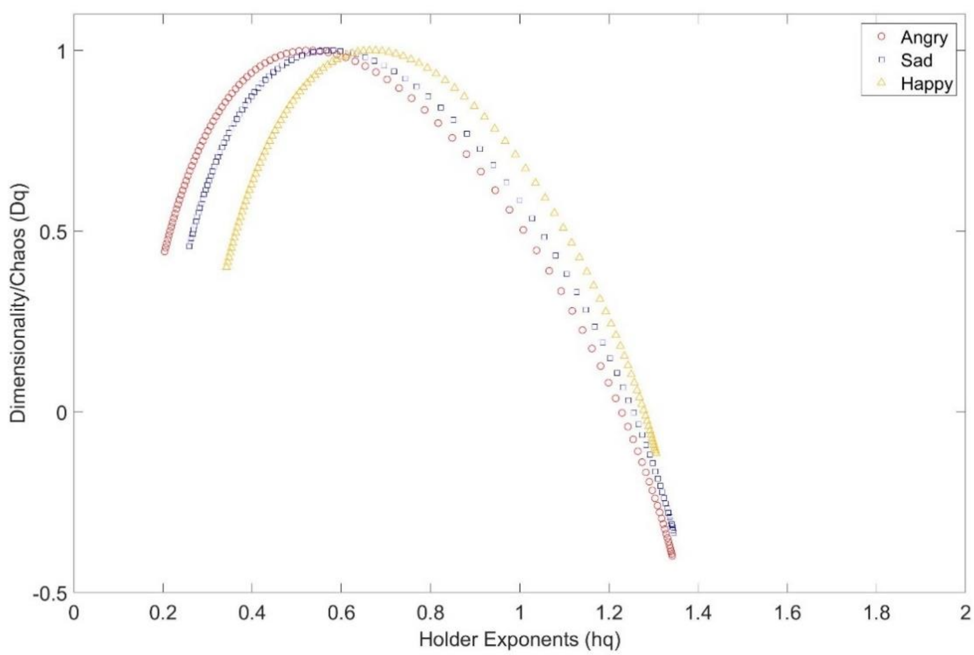

Figure 8. Multifractal spectra averages for alpha at Cz, averaged across guitar and voice.

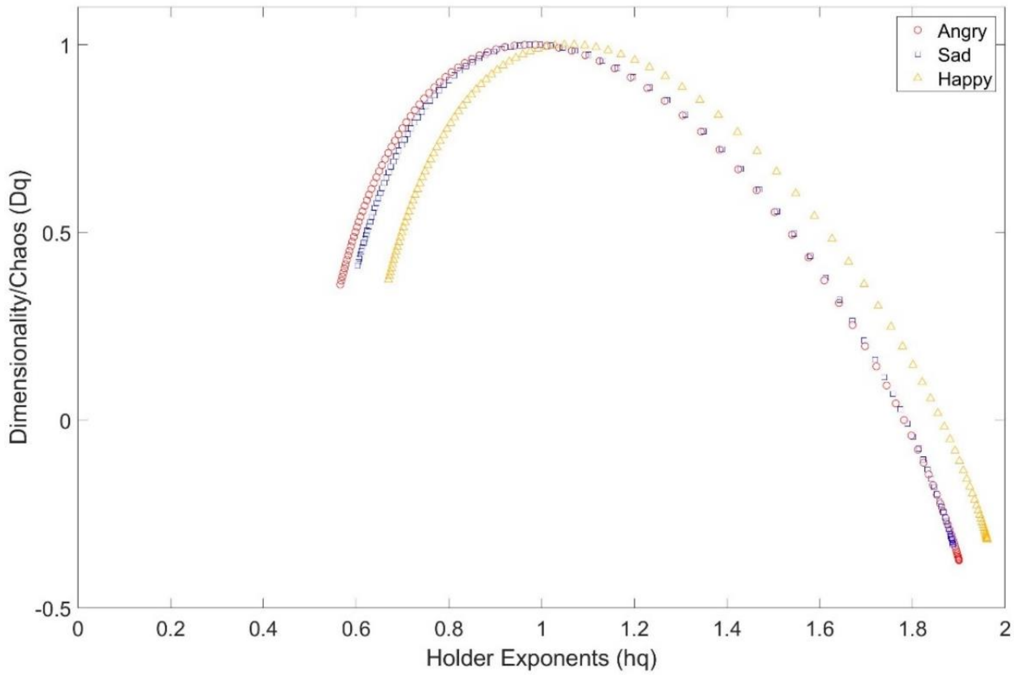

Figure 9. Multifractal spectra averages for theta at Cz, averaged across guitar and voice.

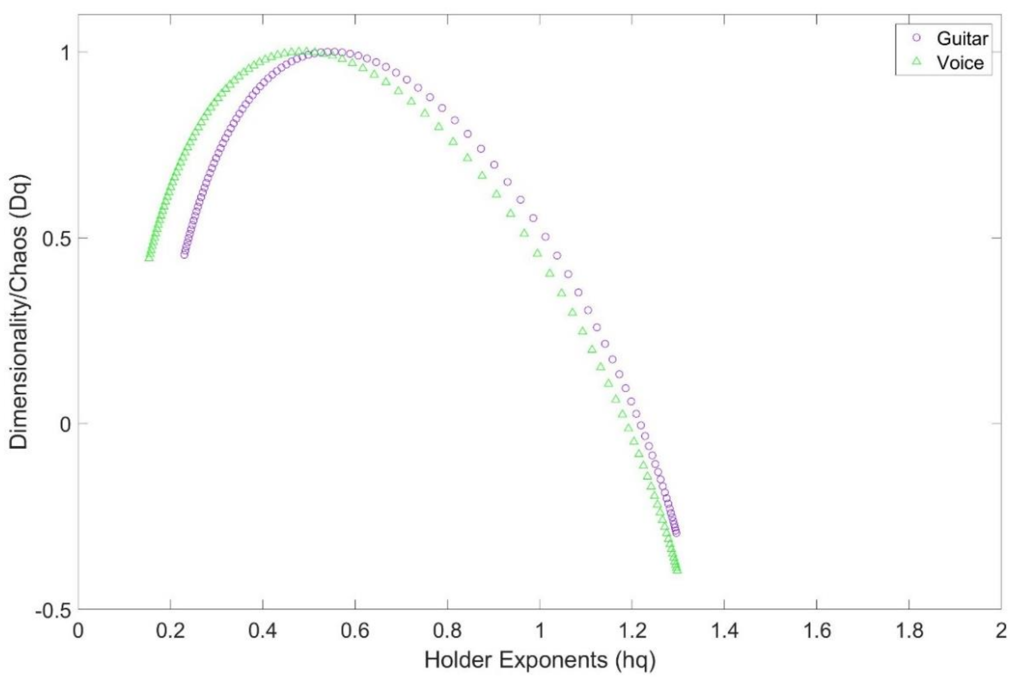

Figure 10. Multifractal spectra averages for alpha at Fz, averaged across emotions.

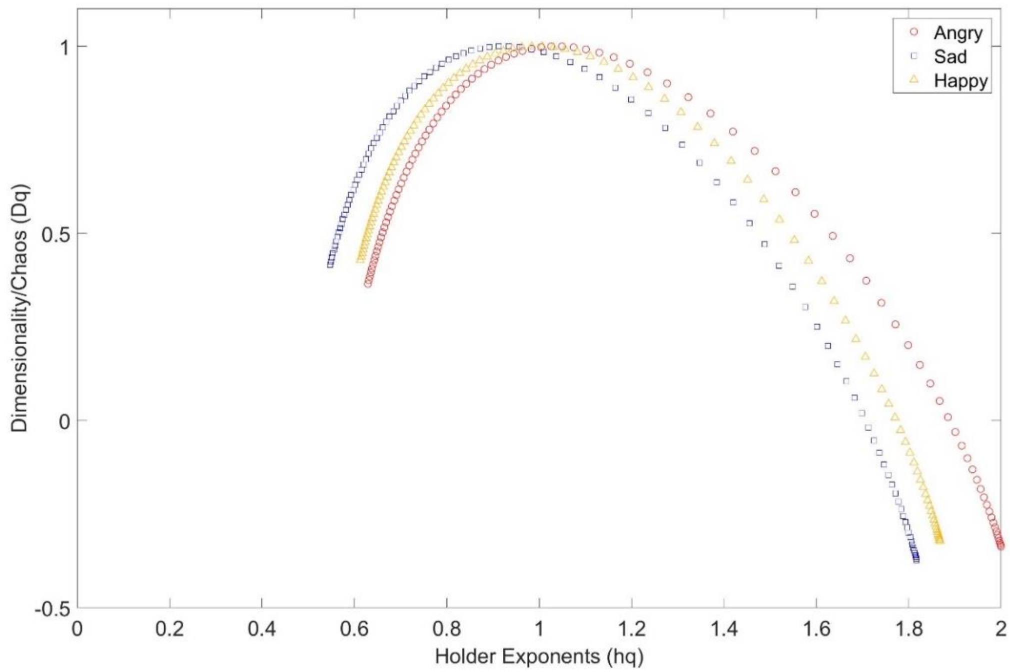

Figure 11. Multifractal spectra averages for theta at Fz for guitar.

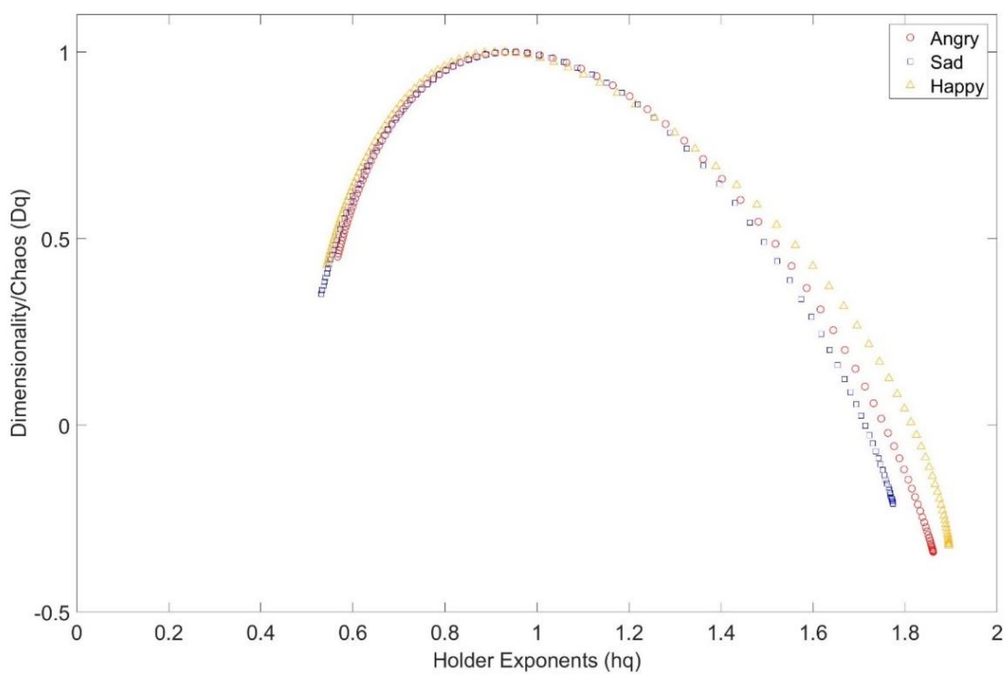

Figure 12. Multifractal spectra averages for theta at Fz for voice.

To expound upon these results, univariate tests were also completed with a Bonferroni correction (Table B8; Table B9; Table B10; Table B11). These results show that α_0 is the primary dependent variable with significant relationships to the emotion category; Cz alpha (F (2, 18) = 9.820, p = .001; partial η^2 = .522); Cz theta (F (2, 18) = 5.174, p = .017; partial η^2 = .365); Fz alpha (F (2, 18) = 5.638, p = .013; partial η^2 = .385); Fz theta (F (2, 18) = 6.357, p = .008; partial η^2 = .414). However, the alpha rhythms from the Cz electrode show significant relationships to all three dependent variables; width (F (2, 18) = 10.683, p = .001; partial η^2 = .543); asymmetry (F (2, 18) = 5.624, p = .013; partial η^2 = .385). To further investigate these relationships, means of each property were calculated for each emotion and each instrument.

Means of asymmetry were higher for angry sounds than for happy sounds when listening to both guitar and vocal sounds. The patterns for asymmetry in EEG responses differed between voice and guitar on sad sounds (Figure 13).

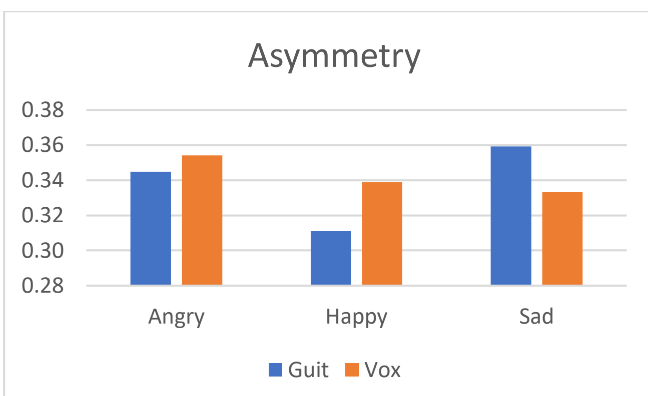

Figure 13. Means of the asymmetry of the multifractal spectra while participants heard each type of sound.

Means of α_0 were higher for happy sounds than for angry sounds when listening to both guitar and vocal sounds. The patterns for α_0 in EEG responses also differed between voice and guitar on sad sounds (Figure 14).

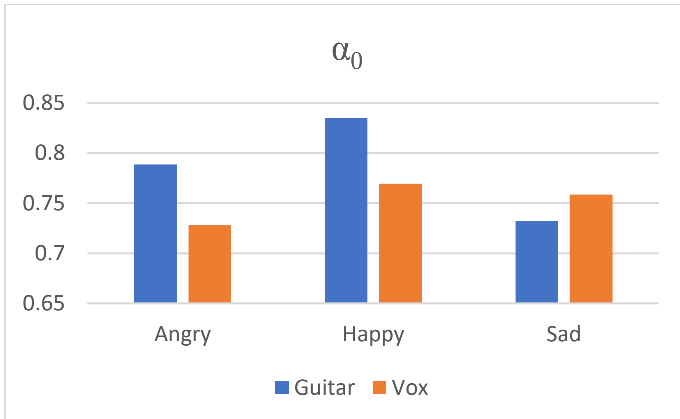

Figure 14. Means of α_0 of the multifractal spectra while participants heard each type of sound.

Means of width of the multifractal spectra were higher for angry sounds than happy sounds in responses to both guitar and voice. Again, the pattern differed on sad sounds (Figure 15).

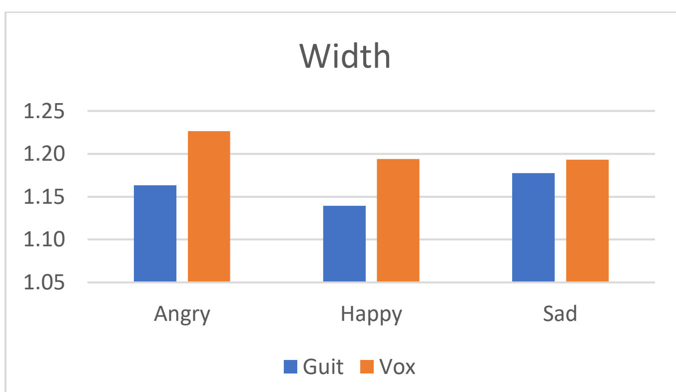

Figure 15. Means of the width of the multifractal spectra while participants heard each type of sound.

Discussion

In this study, I tested the hypothesis that there are systematic differences of the multifractal spectrum of EEG responses (width, apex, asymmetry) between conditions (emotion and sound type of sound being heard) that would be obscured by other forms of analysis.

To address this question, a collection of high-quality sounds was recorded. Non-verbal vocal sounds were recorded for the emotions of sadness, happiness, and anger. Electric guitar sounds were recorded with differing timbre characteristics. Listeners validated these sounds for their emotional content by making judgements about what emotion they conveyed, how intense the emotion was, and (for vocal sounds) how believable the emotion was. The results of this process indicated that certain parameters of electric guitar timbre can predictably convey emotional information to listeners. They also indicated that there is a relationship between the emotional information conveyed through timbre in electric guitar sounds and that conveyed through non-linguistic vocal sounds. Specifically, a high mean attack slope (which might also be called a fast onset time) for a given electric guitar or non-linguistic vocal sound indicates a "happy" emotion, a low mean attack slope indicates "angry," and an attack slope between those two indicates "sad." This is an interesting finding, as some studies have found anger to have a particularly high attack slope. This could be related to more subtle differences in types of anger (for example, threatening anger or more passive anger). To my knowledge, the finding that attack slope alone can predict the emotional category of a sound (at least in regard to these three categories) and, to some degree, the neurological response of the listener, is unique to this study.

While self-report measures of perceived emotion were helpful in obtaining listeners' subjective perceptions about the sounds, the EEG data collected during the experiment give further insight into how hearing these sounds with perceived emotional content affect the underlying neurological processes of the brain. Since the processes of the brain have been found to be complex and multifractal in several studies including the current study, it is appropriate to use multifractal analysis techniques and inappropriate to use other analysis techniques that do not account for this complexity. It was found that the singularity spectrum/multifractal spectrum of EEG responses (width, apex, and asymmetry) differs as a function of the emotional category of a sound at the Cz electrode. At the Fz electrode, the multifractal spectra of the alpha rhythm waves differ as a function of the sound type. The multifractal spectra of the theta rhythm at Fz differ as a function of the interaction of sound type and emotion.

There are some trends across electrode sites and brain wave rhythms. Where emotion had a significant contribution to the differences, multifractal spectra of EEG responses to angry sounds have a higher width than spectra for sad sounds, indicating that neurological responses at these electrodes are higher in complexity when listening to angry sounds than when listening to sad sounds. Multifractal spectra of EEG responses where emotion was significant have a higher α_0 for happy sounds than for sad sounds, indicating that neurological responses at these electrodes are more persistent when listening to happy sounds than when listening to sad sounds. Asymmetry for theta rhythms at the Fz electrode is higher in response to happy sounds than to angry sounds, for both guitar and voice. This indicates that extreme events play a less prominent role in neurological responses to angry sounds than in neurological responses to happy sounds when considering the theta rhythm at Fz. Neurological responses to angry sounds rely more heavily on smaller fluctuations, while responses to happy sounds are more balanced in their reliance on both small and large fluctuations.

These results have direct implications for the entertainment industry. The results support the current practice of often using low attack slopes to build suspense by representation of anger or threat in sound for film. Sound designers for a film or video games often utilize sounds with low (slow) attack slopes if they want to convey anger to the consumer. Conversely, a sound designer would want to use sounds with high (quick) attack slopes to convey happiness. This could also apply to singers and voice actors, who could slowly increase the amplitude of a note or phrase to more accurately portray anger. Further research involving attack slopes would be necessary to understand how attack slopes are used in different styles of music, but according to the present study it would be likely that slow (low) attack slopes would be most effective in conveying anger in music, especially when considering the electric guitar. It is tempting when discussing music to equate attack slopes involved in timbre of a sound with the tempo of a song; however, these are two separate features. A fast song in the metal genre, for example, might use guitar sounds with relatively low attack slopes. Through the song itself is moving quickly, each individual note might be relatively slow in its rise in amplitude when compared to other guitar timbres. While evidence toward the overarching concept that listeners perceive different emotional communication based only on changes in timbre is in itself important, more research is needed to add further ecological validity to these findings.

These results also apply directly to music therapy and other forms of sonic therapy. For example, if there is a need to raise multifractal spectral width, listening to a sound with a low (slow) attack slope would encourage this. This might be helpful in altering neurological systems for patients who struggle with certain pathologies such as poststroke depression and have a particularly low multifractal spectral width. Conversely, it is possible that sounds with high (quick) attack slopes might decrease multifractal spectral width for people experiencing schizophrenia, epileptic seizures, or major depression by encouraging the neurological system to become less chaotic and more persistent. Further research needs to be done to confirm whether these results might generalize to those who suffer from these conditions, and whether they might help with symptoms in a practical way.

The results of this study could be expanded upon in many ways. One limitation of the current study is that direct comparison between the vocal sounds and the guitar sounds may have been difficult for participants, since the guitar sounds contained only one note (fundamental frequency) while the vocal sounds contained fluctuating fundamental frequencies. While this limitation was necessary to isolate the effects of timbre on the perception of the guitar sounds, a study in which the sounds are more directly comparable could reveal new insights. Using vocal sounds that are consistently categorized into different emotions while not containing pitch fluctuation would allow for a more direct comparison and might reveal more about the relationship between communication through timbre in voice and communication through timbre in electric guitar. However, this would be fairly unnatural and would therefore lack much in terms of ecological validity. Another way to improve the comparability of sounds might be to imitate the vocal fluctuations of fundamental frequency with a guitar. This would be somewhat difficult to do with a high level of accuracy, but would have higher ecological validity since it is rare that vocalists or guitarists only use a single fundamental frequency in communication. Also, while the results of this study showed some trends, there are individual differences that can be seen between participants (Appendix C). This indicates that there are other factors at play in these neurological responses that were not considered in this study and could be used as variables in future research. For example, how might neurological responses of women differ from those of men? Beyond differences in the sex of the listener, differences in responses depending on the sex of the speaker might be explored. I controlled for this by utilizing a male and a female speaker, but this could instead be treated as its own independent variable (IV) in future studies to determine whether there are systematic differences between multifractal neurological responses to male vocal sounds and responses to female vocal sounds. I think that there could be a systematic difference between responses to vocal sounds of anger dependent on whether the sex of the listener matches that of the speaker. For example, when the sex of the speaker and listener match, the observable reaction might be more aggressive than it would be if the sex is mismatched. This might relate to a difference in the underlying neurological state as well. Based on my research, I speculate that there would be an interaction between the sex of the speaker and the sex of the listener on neurological responses. Responses to same-sex vocal sounds of anger might have higher multifractal width, lower α_0, and higher asymmetry than mismatched-sex vocal sounds. Multifractal characteristics would also likely differ between professional audio engineers (for music, film, video games, etc.) and those without any audio work experience, since a part of an audio engineer's job is to listen for and adjust timbre differences. I suggest that audio engineers might show more pronounced differences in neurological responses to sounds with differing perceived emotional content. Another potential area for expansion is in observing behavioral and cognitive results of the alteration of the multifractal spectrum to offer further insight into how this information might be applied effectively in music therapy. For example, the following questions might be asked: Can the induced neurological state be maintained over time? If so, how might changes in this state be reflected in cognitive tasks? How might someone convey fear or calmness using attack slopes? I predict that sounds that convey fear might have a relatively high attack slope, since fear is an emotion that is more associated with submissiveness rather than aggression (such as anger). The multifractal neurological responses to these sounds of fear might be similar to those of sadness, since these sounds both imply submissiveness and need for social support. However, some have speculated that fear (along with some other emotions) may be difficult to express utilizing only timbre and could rely on other elements of sound. These results could also be expanded upon from a musicological perspective by studying different pieces or genres of music to determine average attack slopes. In combination with collection of EEG data and/or self-report data, the extent to which the current findings might relate to everyday music listening experiences might be discovered.

References

American Standards Association (1960). *American standard acoustical terminology: (including mechanical shock and vibration).* [New York].

Audio Technica (2018). Retrieved from https://www.audio-technica.com/cms/headphones/75b2f282c93a7651/

Avid (2018). Pro Tools Version 2018.4 [Computer Software].

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology, 70*(3), 614-36.

Biopac Systems, Inc. (2010). Biopac Student Lab PRO Software and MP35 Hardware. Further information can be found at https://www.biopac.com/

Bowman, C., & Yamauchi, T. (2016). Perceiving categorical emotion in sound: The role of timbre. *Psychomusicology: Music, Mind, and Brain, 26*(1), 15-25.

Caclin, A., McAdams, S., Smith, B., & Winsberg, S. (2005). Acoustic correlates of timbre space dimensions: A confirmatory study using synthetic tones. *The Journal of the Acoustical Society of America, 118*(1), 471-482.

Department of Psychology (2008). G*Power [Computer Software]. Retrieved from http://download.cnet.com/G-Power/3000-2054_4-10647044.html

Eerola, T., Ferrer, R., & Alluri, V. (2012). Timbre and affect dimensions: evidence from affect and similarity ratings and acoustic correlates of isolated instrument sounds. (2012) *Music Perception: An Interdisciplinary Journal*, *30*(1), 49-70.

Electro-Cap International, Inc. [Electrode Cap and Supplies].

Filipic, S., Tillmann, B., & Bigand, E. (2010). Judging familiarity and emotion from very brief musical excerpts. *Psychonomic Bulletin & Review, 17*(3), 335-341.

Greco, A., Valenza, G., Citi, L., & Scilingo, E. (2017). Arousal and valence recognition of affective sounds based on electrodermal activity. *Sensors Journal, IEEE, 17*(3), 716-725.

Grieder, M., Wang, D., Dierks, T., Wahlund, L., & Jann, K. (2018). Default Mode Network complexity and cognitive decline in mild Alzheimer's Disease. *Frontiers in Neuroscience, 12*, 770.

Griffiths, T., & Warren, J. (2004). What is an auditory object? *Nature Reviews. Neuroscience, 5*(11), 887-92.

Hailstone, J., Omar, R., Henley, S., Frost, C., Kenward, M., & Warren, J. (2009). It's not what you play, it's how you play it: Timbre affects perception of emotion in music. *The Quarterly Journal of Experimental Psychology, 62*(11), 2141-2155.

Handtronix Online Hearing Test. Retrieved from www.hearingscreeners.com/onlineTest.html

Huron, D. (2001). Tone and voice: A derivation of the rules of voice-leading from perceptual principles. *Music Perception, 19*, 1–64.

Ihlen, E. (2012). Introduction to multifractal detrended fluctuation analysis in matlab. *Frontiers in Physiology, 3*, 141.

Ilie, G., & Thompson, W. (2006). A comparison of acoustic cues in music and speech for three dimensions of affect. *Music Perception: An Interdisciplinary Journal, 23*(4), 319-330.

Jarvis, B. G. (2016). MediaLab (Version 2016) [Computer Software]. New York, NY: Empirisoft Corporation.

Kabuto, M., Kageyama, T., & Nitta, H. (1993). EEG power spectrum changes due to listening to pleasant music and their relation to relaxation effects. Nihon Eiseigaku Zasshi. *Japanese Journal of Hygiene, 48*(4), 807-818.

Kato, F., Iwanaga, R., Chono, M., Fujihara, S., Tokunaga, A., Murata, J., Tanaka, K., Nakane, H., Tanaka, G. (2014). Relationship between sympathetic skin responses and auditory hypersensitivity to different auditory stimuli. *Journal of Physical Therapy Science, 26*(7), 1087-91.

Lartillot, O. (2014). *MIRtoolbox User's Guide 1.6.1*. Available at: https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox/MIRtoolbox1.6.1guide

Lartillot, O., Toiviainen, P. (2007). A Matlab toolbox for musical feature extraction from audio. *International Conference on Digital Audio Effects*, Bordeaux.

Laukka, P., Elfenbein, H. A., Chui, W., Thingujam, N. S., Iraki, F. K., Rockstuhl, T., Althoff, J. (2010). "Presenting the VENEC corpus: development of a cross-cultural corpus of vocal emotion expressions and a novel method of annotating emotion appraisals," in *Proceedings of the LREC 2010 Workshop on Corpora for Research on Emotion and Affect*, eds L. Devillers, B. Schuller, R. Cowie, E. DouglasCowie, and A. Batliner (Valletta: European Language Resources Association), 53–57.

Laukka, P., Elfenbein, H., Söder, N., Nordström, H., Althoff, J., Chui, W., Kang'Ethe Eiraki, F., Erockstuhl, T., Thingujam, N. (2013). Cross-cultural decoding of positive and negative non-linguistic emotion vocalizations. *Frontiers in Psychology, 4*, 353.

Li, Y., Tong, S., Liu, D., Gai, Y., Wang, X., Wang, J., Qiu, Y., Zhu, Y. (2008). Abnormal EEG complexity in patients with schizophrenia and depression. *Clinical Neurophysiology, 119*(6), 1232-1241.

Lima, C., Castro, F., & Scott, S. (2013). When voices get emotional: A corpus of nonverbal vocalizations for research on emotion processing. *Behavior Research Methods, 45*(4), 1234-1245.

Liu, X., Xu, Y., Alter, K. & Tuomainen, J. (2018). Emotional connotations of musical instrument timbre in comparison with emotional speech prosody: Evidence from acoustics and event-related potentials. *Frontiers in Psychology, 9*, 737.

Maity, A.K., Pratihar, R., Mitra, A., Dey, S., Agrawal, V., Sanyal, S., Banerjee, A., Sengupta, R., Ghosh, D. (2015). Multifractal Detrended Fluctuation Analysis of alpha and theta EEG rhythms with musical stimuli. *Chaos, Solitons and Fractals: The Interdisciplinary Journal of Nonlinear Science, and Nonequilibrium and Complex Phenomena, 81*, 52-67.

MATLAB 2014, The MathWorks, Natick, 2014.

Mcadams, S., & Cunible, J. (1992). Perception of timbral analogies. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 336*(1278), 383-389.

Méndez, M., Zuluaga, P., Hornero, R., Gómez, C., Escudero, J., Rodríguez-Palancas, A., Ortiz, T., Fernández, A. (2012). Complexity analysis of spontaneous brain activity: Effects of depression and antidepressant treatment. *Journal of Psychopharmacology, 26*(5), 636-643.

Ming-Zher Poh, N. C., Swenson, R. W., & Picard, R. W. (2010). A wearable sensor for unobtrusive, long-term assessment of electrodermal activity. *Biomedical Engineering, IEEE Transactions on, 57*(5), 1243-1252.

Morton, E. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *The American Naturalist, 111*(981), 855-869.

Mueller, S. T., & Piper, B. J. (2014). The Psychology Experiment Building Language (PEBL) and PEBL test battery. *Journal of Neuroscience Methods* (222), 250–259.

Mueller, S. T. (2012). The Psychology Experiment Building Language, Version 0.13. Retrieved from http://pebl.sourceforge.net.

Ohala, J. J. (1984). An ethological perspective on common cross-language utilization of F_0 of voice. *Phonetica, 41,* 1-16.

Paquette, S., Peretz, I., & Belin, P. (2013). The musical emotional bursts: A validated set of musical affect bursts to investigate auditory affective processing. *Frontiers in Psychology, 4*, 509.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rendall, D., Vokey, J., Nemeth, C., & Humphreys, Glyn W. (2007). Lifting the curtain on the Wizard of Oz: Biased voice-based impressions of speaker size. *Journal of Experimental Psychology: Human Perception and Performance, 33*(5), 1208-1219.

RStudio Team (2015). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/

Santesso, D. L., Schmidt, L. A., & Trainor, L. J. (2007). Frontal brain electrical activity (EEG) and heart rate in response to affective Infant-Directed (ID) speech in 9-month-old infants. *Brain and Cognition, 65*(1), 14-21.

Scherer, K., Banse, R., & Wallbott, H. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology, 32*(1), 76-92.

Schimmack, U., & Grob, A. (2000). Dimensional models of core affect: A quantitative comparison by means of structural equation modeling. *European Journal of Personality*, *14*(4), 325-345.

Schlaug, G., Jäncke, L., Huang, Y., Staiger, J., & Steinmetz, H. (1995). Increased corpus callosum size in musicians. *Neuropsychologia, 33*(8), 1047-1055.

Shiota, M., Neufeld, S., Yeung, W., Moser, S., Perea, E., & Phelps, E. A. (2011). Feeling good: autonomic nervous system responding in five positive emotions. *Emotion, 11*(6), 1368-1378.

Sokunbi, M., Gradin, V., Waiter, G., Cameron, G., Ahearn, T., Murray, A., Steele, D., Staff, R. (2014). Nonlinear complexity analysis of brain fMRI signals in schizophrenia. *PLoS ONE, 9*(5), e95146.

Spencer, H. (1911). *Essays on Education and Kindred Subjects*. Auckland: The Floating Press.

Tajadura-Jiménez, A., Larsson, P., Väljamäe, A., Västfjäll, D., Kleiner, M., & Phelps, E. A. (2010). When room size matters: Acoustic influences on emotional responses to sounds. *Emotion, 10*(3), 416-422.

Van Cappellen van Walsum, A., Pijnenburg, Y., Berendse, H., Van Dijk, B., Knol, D., Scheltens, P., & Stam, C. (2003). A neural complexity measure applied to MEG data in Alzheimer's disease. *Clinical Neurophysiology, 114*(6), 1034-1040.

Vanhoy, M. (2008). An entropy estimator of population variability in nominal data. *Journal of Scientific Psychology, June*, 25-30.

Xu, Y., Lee, A., Wing-Li, W., Liu, X., & Birkholz, P. (2013). Human vocal attractiveness as signaled by body size projection. *PLoS One, 8*(4), E62397.

Zhang, Y., Wang, C., Sun, C., Zhang, X., Wang, Y., Qi, H., He, F., Zhao, X., Wan, B., Du, J., Ming, D. (2015). Neural complexity in patients with poststroke depression: A resting EEG study. *Journal of Affective Disorders, 188*, 310-318.

Appendices

Appendix A: Consent Form

UNIVERSITY OF CENTRAL OKLAHOMA

INFORMED CONSENT FORM

Research Project Title: Physiological and Emotional Reactions to Timbre in Voices and Electric Guitar

Researcher (s): Sephra Scheuber, Dr. Mickie Vanhoy

A. **Purpose of this research:** This research is designed to explore emotional and physiological reactions to sounds and to build scientific knowledge about how sounds can affect people physiologically and psychologically.
B. **Procedures/treatments involved:** First, participants will receive a basic hearing test to ensure that hearing is adequate for participation. If the hearing test is not passed, the participant will not be able to complete the study but will still receive course credit. If the hearing test is passed, participants will be seated comfortably in a laboratory room in front of a computer with feet flat on the floor. Four electrodes will be placed on the participant's scalp for measurement of EEG brain signals. This requires mild scalp abrasion. The participant will be provided with headphones to put on. They will hear several sounds while sitting still and will be asked to answer a question after each pair of sounds via a mouse click. The questions will be based on the subjective experience of the participant; there are no "right" or "wrong" answers. The experimenter or lab assistant will be nearby to deal with any technical malfunctions.
C. **Expected length of participation:** 1.5 hours total
D. **Potential benefits:** Course credit and the opportunity to contribute to scientific knowledge about how sounds affect human psychology and physiology. Participants who do not pass the hearing test will be unable to complete the experiment but will still receive course credit.
E. **Potential risks or discomforts:** Though improbable, there is the possibility of equipment malfunction leading to sounds being too loud and hurting the ears or a small electrical

APPROVED
July 07, 2019
UCO IRB

shock. There is also a possibility of sympathetic nervous system arousal, but this is not likely to exceed the levels that would happen while listening to music or an emotional conversation. As electrode placement requires very mild skin abrasion, there is also a possibility of mild scalp irritation.

F. **Medical/mental health contact information (if required):** N/A

G. **Contact information for researchers:** Sephra Scheuber: sscheuber@uco.edu

H. **Contact information for UCO IRB:** irb@uco.edu or (405) 974-5497

I. **Explanation of confidentiality and privacy:** No identifying data will be kept, only a list of who participated in order to give course credit.

J. **Assurance of voluntary participation:** At any time, you can choose to walk away and withdraw your participation in this study, but you will not receive course credit. You will not be penalized for withdrawing. This is completely voluntary.

AFFIRMATION BY RESEARCH SUBJECT

I hereby voluntarily agree to participate in the above listed research project and further understand the above listed explanations and descriptions of the research project. I also understand that there is no penalty for refusal to participate, and that I am free to withdraw my consent and participation in this project at any time without penalty. I acknowledge that I am at least 18 years old. I have read and fully understand this Informed Consent Form. I sign it freely and voluntarily. I acknowledge that a copy of this Informed Consent Form has been given to me to keep.

Research Subject's Name:

Signature: ______________________________ Date

APPROVED
July 07, 2019
UCO IRB

Appendix B: Tables

Table B1

Entropy calculations for each guitar sound.

File Name	Angry	Sad	Happy	Other	Entropy	Most Common Emotion Category
AMP4DISTRD*	3.03%	6.06%	18.18%	72.73%	0.817412	HAPPY
AMP5SW*	6.90%	6.90%	31.03%	55.17%	1.060088	HAPPY
AMP8SW*	0.00%	53.85%	0.00%	46.15%	0.690186	SAD
AMPCln	0.00%	44.44%	11.11%	44.44%	0.964963	SAD
Guit1CutHarsh	3.45%	20.69%	20.69%	55.17%	1.096174	SAD/HAPPY SPLIT
Guit1Dark*	3.45%	27.59%	0.00%	68.97%	0.727635	SAD
Guit1HiCutLowShlf	3.70%	37.04%	0.00%	59.26%	0.800012	SAD
Guit1LowCutHiShlf	10.71%	7.14%	10.71%	71.43%	0.907468	ANGRY/HAPPY SPLIT
Guit1121NL1*	0.00%	25.00%	0.00%	75.00%	0.562335	SAD
Guit1121NL5	4.00%	48.00%	0.00%	48.00%	0.833365	SAD
Guit1121NL7	3.70%	37.04%	0.00%	59.26%	0.800012	SAD
Guit1121NL10*	15.38%	7.69%	0.00%	76.92%	0.687092	ANGRY
Guit1121NL15	0.00%	18.18%	9.09%	72.73%	0.759547	SAD
Guit1121NL16	0.00%	35.71%	21.43%	42.86%	1.060944	SAD
Guit1121NL19	8.33%	41.67%	0.00%	50.00%	0.918428	SAD
Guit1121NL23	11.11%	7.41%	22.22%	59.26%	1.08124	ANGRY
Guit1121NL27	11.54%	15.38%	11.54%	61.54%	1.085086	SAD
Guit1121NL30*	6.67%	13.33%	0.00%	80.00%	0.627705	SAD
Guit1121NL31	12.50%	25.00%	12.50%	50.00%	1.213008	SAD
Guit1121NL35*	53.33%	20.00%	0.00%	26.67%	1.009614	ANGRY
Guit1121NL38*	50.00%	0.00%	50.00%	0.00%	0.693147	ANGRY
Guit1121NL40	13.64%	36.36%	13.64%	36.36%	1.2791	SAD
Guit1121NL43	53.33%	0.00%	20.00%	26.67%	1.009614	ANGRY
Guit1121NL46	4.35%	60.87%	0.00%	34.78%	0.805827	SAD
Guit1121NL51*	3.57%	21.43%	32.14%	42.86%	1.177045	HAPPY
Guit1121NL52*	23.81%	19.05%	0.00%	57.14%	0.97732	ANGRY
Guit1121NL55*	10.34%	0.00%	20.69%	68.97%	0.816915	HAPPY
Guit1121NL60	10.71%	7.14%	10.71%	71.43%	0.907468	ANGRY/HAPPY SPLIT

Note. * indicates sounds that were chosen.

Table B2

Results of MANOVA with timbre characteristics (brightness, attack slope, spectral centroid, and spectral flux) as dependent variables.

Effect	*Wilks' λ*	*F*	*Hypothesis df*	*Error df*	*Sig.*
emotion	.582	3.812	8.000	98.000	*.001
sound	.964	.453	4.000	49.000	.770
emotion * sound	.782	1.602	8.000	98.000	.134

Note. * denotes significant *p* values at $\alpha = .05$.

Table B3

Results of Between-Subjects Effects for MANOVA with timbre characteristics as dependent variables.

Effect	*SS*	*df*	*MS*	*F*	*Sig.*
brightness	88.294	2	44.147	.406	.668
attack slope	1579.422	2	789.711	10.224	.000*
centroid	167.327	2	83.663	.674	.514
flux	262.189	2	131.094	1.093	.343

Note. * denotes significant p values at $\alpha = .05$.

Table B4

Results of MANOVA at electrode site Cz on brain wave rhythm alpha with multifractal spectrum characteristics (width, α_0, asymmetry) as DVs.

Within Subjects Effect	*Wilks' λ*	*F*	*Hypothesis df*	*Error df*	*Sig.*	*Partial* η^2	*Observed Power*
emotion	.308	4.275	6.000	32.000	.003*	.445	.953
sound	.909	.234	3.000	7.000	.870	.091	.077
emotion * sound	.731	.903	6.000	32.000	.505	.145	.303

Note. * denotes significant *p* values at $\alpha = .05$.

Table B5

Results of MANOVA at electrode site Cz on brain wave rhythm theta with multifractal spectrum characteristics (width, α_0, asymmetry) as DVs.

Within Subjects Effect	*Wilks' λ*	*F*	*Hypothesis df*	*Error df*	*Sig.*	*Partial η^2*	*Observed Power*
emotion	.350	3.687	6.000	32.000	.007*	.409	.915
sound	.886	.300	3.000	7.000	.825	.114	.084
emotion * sound	.616	1.462	6.000	32.000	.223	.215	.485

Note. * denotes significant *p* values at $\alpha = .05$.

Table B6

Results of MANOVA at electrode site Fz on brain wave rhythm alpha with multifractal spectrum characteristics (width, α_0, asymmetry) as DVs.

Within Subjects Effect	*Wilks' λ*	*F*	*Hypothesis df*	*Error df*	*Sig.*	*Partial* η^2	*Observed Power*
emotion	.563	1.775	6.000	32.000	.136	.250	.579
sound	.304	5.348	3.000	7.000	.031*	.696	.718
emotion * sound	.640	1.333[c]	6.000	32.000	.271	.200	.444

Note. * denotes significant *p* values at α = .05.

Table B7

Results of MANOVA at electrode site Fz on brain wave rhythm theta with multifractal spectrum characteristics (width, α_0, asymmetry) as DVs.

Within Subjects Effect	*Wilks' λ*	*F*	*Hypothesis df*	*Error df*	*Sig.*	*Partial η^2*	*Observed Power*
emotion	.429	2.813	6.000	32.000	.026*	.345	.811
sound	.140	14.357	3.000	7.000	.002*	.860	.989
emotion * sound	.338	3.840	6.000	32.000	.005*	.419	.927

Note. * denotes significant *p* values at α = .05.

Table B8

Results of univariate tests with Bonferonni correction at electrode site Cz on brain wave rhythm alpha with multifractal spectrum characteristics (width, α_0, asymmetry) as DVs.

Source	Measure	*SS*	*df*	*MS*	*F*	*Sig.*	*Partial* η^2
emotion	Alpha0	.209	2	.105	9.820	.001*	.522
	Width	.323	2	.162	10.683	.001*	.543
	Asymmetry	.116	2	.058	5.624	.013*	.385
Error(emotion)	Alpha0	.192	18	.011			
	Width	.272	18	.015			
	Asymmetry	.185	18	.010			
sound	Alpha0	.001	1	.001	.293	.602	.032
	Width	.006	1	.006	.134	.723	.015
	Asymmetry	.003	1	.003	.564	.472	.059
Error(sound)	Alpha0	.043	9	.005			
	Width	.423	9	.047			
	Asymmetry	.056	9	.006			
emotion * sound	Alpha0	.024	2	.012	1.641	.221	.154
	Width	.051	2	.025	1.270	.305	.124
	Asymmetry	.017	2	.009	1.055	.369	.105
Error(emotion*sound)	Alpha0	.134	18	.007			
	Width	.360	18	.020			
	Asymmetry	.146	18	.008			

Note. * denotes significant *p* values at $\alpha = .05$.

Table B9

Results of univariate tests with Bonferonni correction at electrode site Cz on brain wave rhythm theta with multifractal spectrum characteristics (width, α_0, asymmetry) as DVs.

Source	Measure	*SS*	*df*	*MS*	*F*	*Sig.*	*Partial η^2*
emotion	Alpha0	.086	2	.043	5.174	.017*	.365
	Width	.029	2	.014	.345	.713	.037
	Asymmetry	.001	2	.000	.045	.956	.005
Error(emotion)	Alpha0	.149	18	.008			
	Width	.756	18	.042			
	Asymmetry	.176	18	.010			
sound	Alpha0	.002	1	.002	.819	.389	.083
	Width	1.323E-7	1	1.323E-7	.000	.998	.000
	Asymmetry	7.125E-5	1	7.125E-5	.017	.900	.002
Error(sound)	Alpha0	.024	9	.003			
	Width	.189	9	.021			
	Asymmetry	.038	9	.004			
emotion * sound	Alpha0	.029	2	.015	3.328	.059	.270
	Width	.023	2	.012	.646	.536	.067
	Asymmetry	.027	2	.014	1.586	.232	.150
Error(emotion*sound)	Alpha0	.079	18	.004			
	Width	.322	18	.018			
	Asymmetry	.154	18	.009			

Note. * denotes significant *p* values at α = .05.

Table B10

Results of univariate tests with Bonferonni correction at electrode site Fz on brain wave rhythm alpha with multifractal spectrum characteristics (width, α_0, asymmetry) as DVs.

Source	Measure	*SS*	*df*	*MS*	*F*	*Sig.*	*Partial* η^2
emotion	Alpha0	.027	2	.013	5.638	.013*	.385
	Width	.004	2	.002	.244	.786	.026
	Asymmetry	.004	2	.002	.566	.578	.059
Error(emotion)	Alpha0	.043	18	.002			
	Width	.142	18	.008			
	Asymmetry	.067	18	.004			
sound	Alpha0	.057	1	.057	9.827	.012*	.522
	Width	.092	1	.092	2.845	.126	.240
	Asymmetry	2.998E-5	1	2.998E-5	.024	.881	.003
Error(sound)	Alpha0	.052	9	.006			
	Width	.291	9	.032			
	Asymmetry	.011	9	.001			
emotion * sound	Alpha0	.021	2	.011	3.993	.037*	.307
	Width	.001	2	.000	.016	.984	.002
	Asymmetry	.004	2	.002	.544	.589	.057
Error(emotion*sound)	Alpha0	.048	18	.003			
	Width	.378	18	.021			
	Asymmetry	.068	18	.004			

Note. * denotes significant *p* values at $\alpha = .05$.

Table B11

Results of univariate tests with Bonferonni correction at electrode site Fz on brain wave rhythm theta with multifractal spectrum characteristics (width, α_0, asymmetry) as DVs.

Source	Measure	*SS*	*df*	*MS*	*F*	*Sig.*	*Partial* η^2
emotion	Alpha0	.033	2	.016	6.357	.008*	.414
	Width	.062	2	.031	.901	.424	.091
	Asymmetry	.012	2	.006	1.023	.379	.102
Error(emotion)	Alpha0	.046	18	.003			
	Width	.619	18	.034			
	Asymmetry	.104	18	.006			
sound	Alpha0	.038	1	.038	15.276	.004*	.629
	Width	9.203E-5	1	9.203E-5	.007	.934	.001
	Asymmetry	.000	1	.000	.026	.876	.003
Error(sound)	Alpha0	.023	9	.003			
	Width	.114	9	.013			
	Asymmetry	.077	9	.009			
emotion * sound	Alpha0	.038	2	.019	8.685	.002*	.491
	Width	.085	2	.042	.992	.390	.099
	Asymmetry	.031	2	.016	8.029	.003*	.471
Error(emotion*sound)	Alpha0	.039	18	.002			
	Width	.770	18	.043			
	Asymmetry	.035	18	.002			

Note. * denotes significant *p* values at $\alpha = .05$.

Appendix C: MFDFA Spectra Averages by Participant

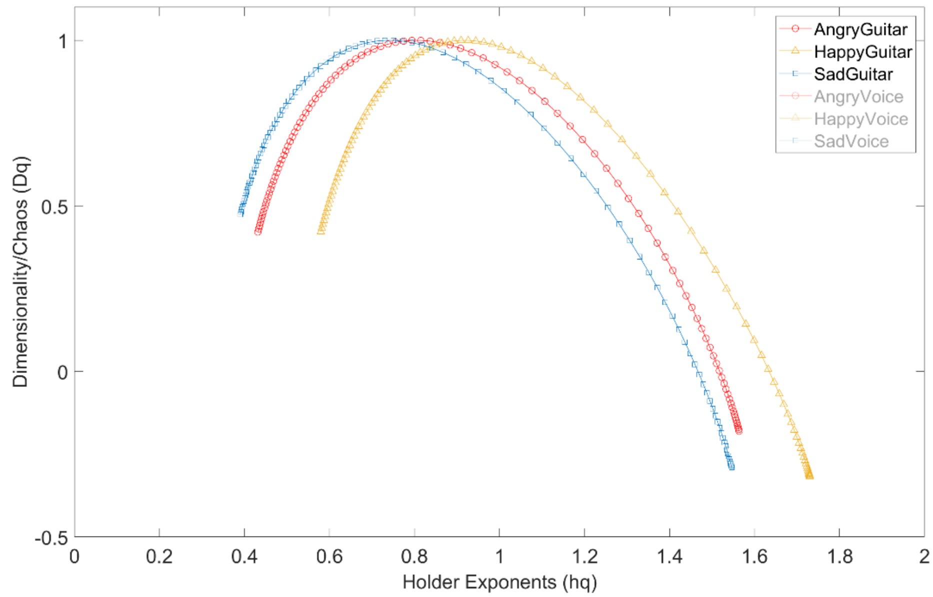

Figure 16. Multifractal spectra of EEG responses to guitar sounds of different emotions for participant 1.

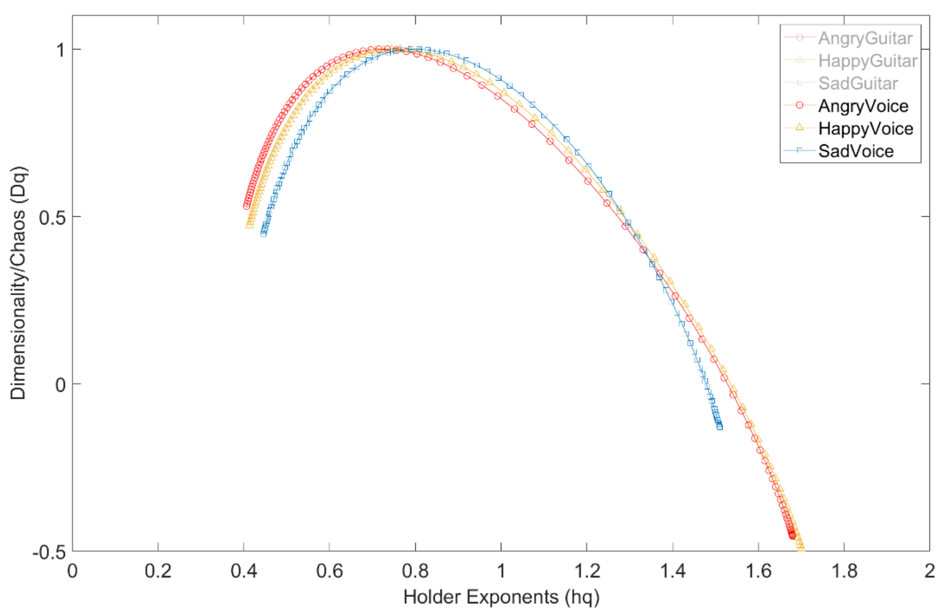

Figure 17. Multifractal spectra of EEG responses to voice sounds of different emotions for participant 1.

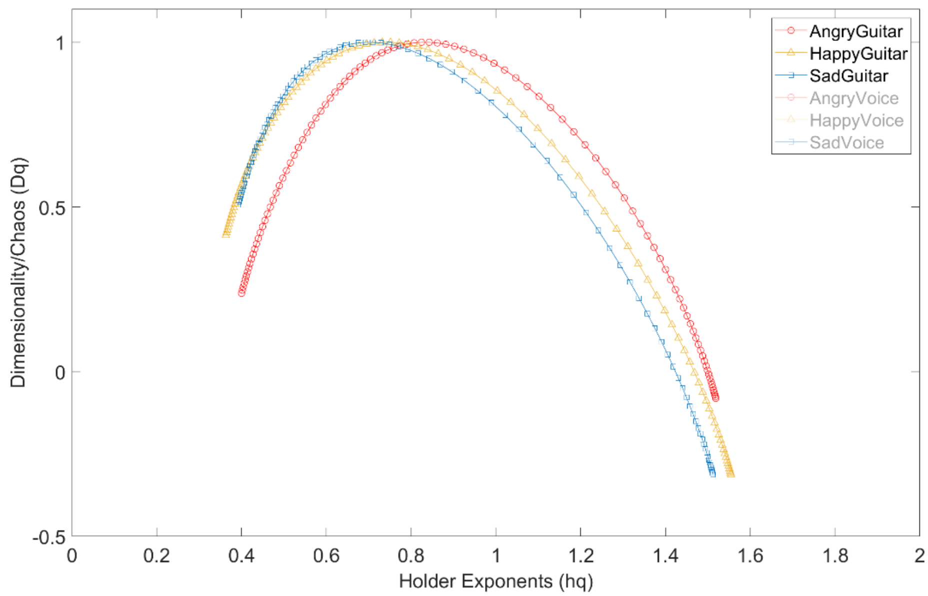

Figure 18. Multifractal spectra of EEG responses to guitar sounds of different emotions for participant 2.

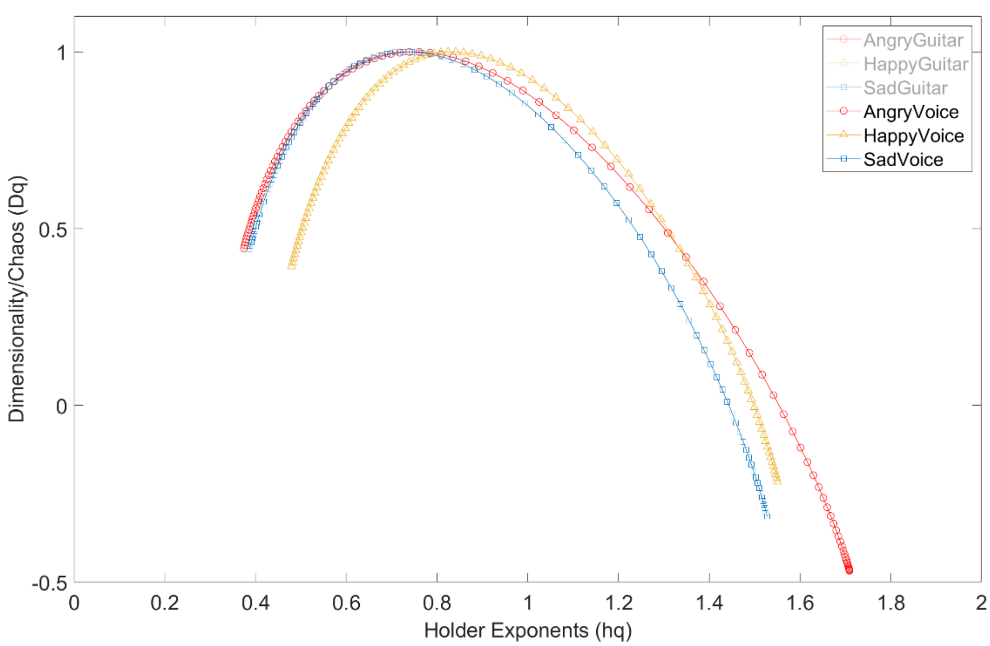

Figure 19. Multifractal spectra of EEG responses to voice sounds of different emotions for participant 2.

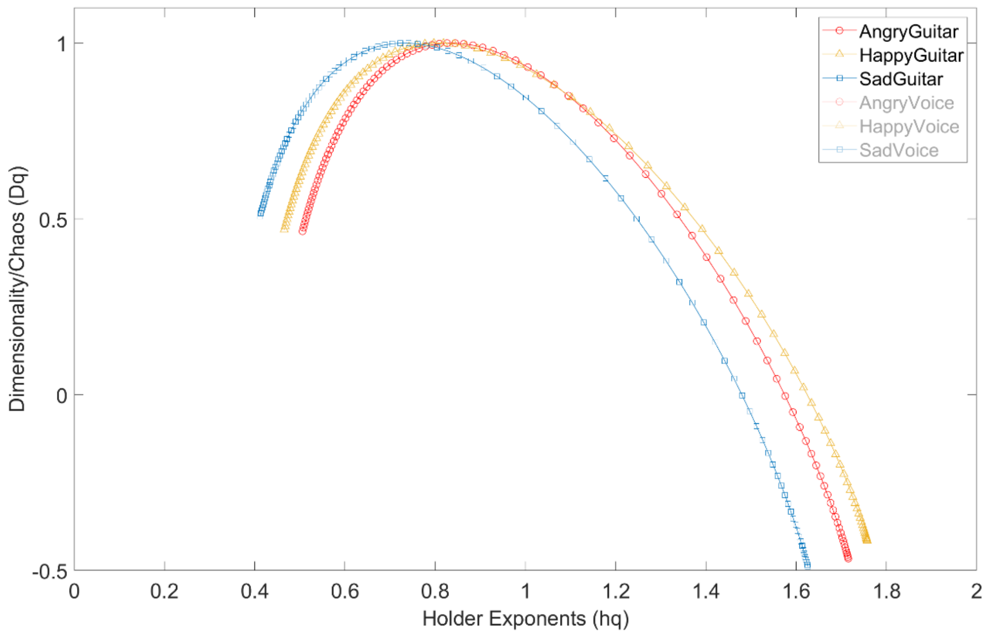

Figure 20. Multifractal spectra of EEG responses to guitar sounds of different emotions for participant 3.

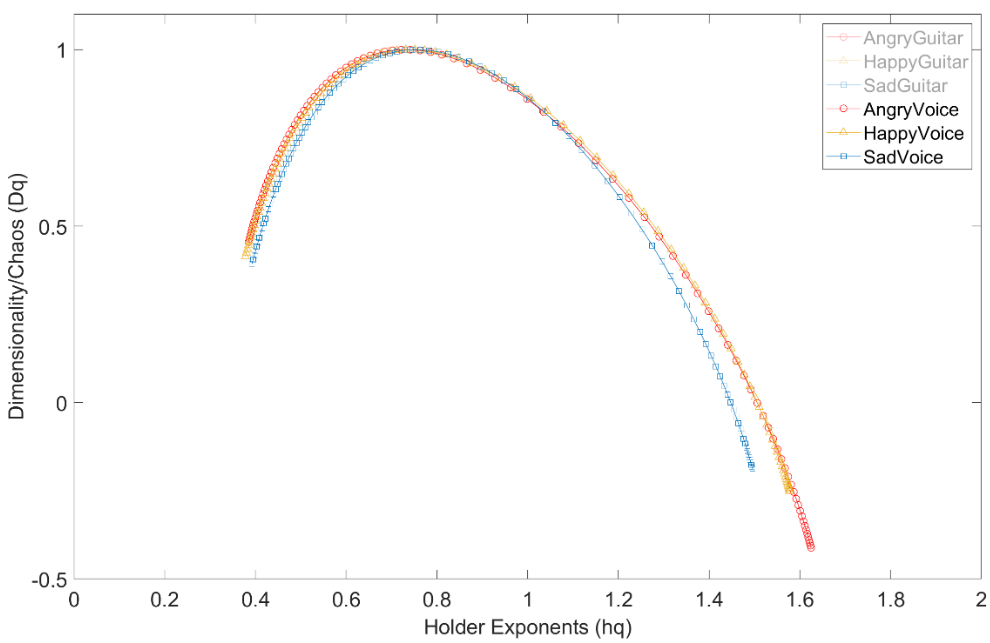

Figure 21. Multifractal spectra of EEG responses to voice sounds of different emotions for participant 3.

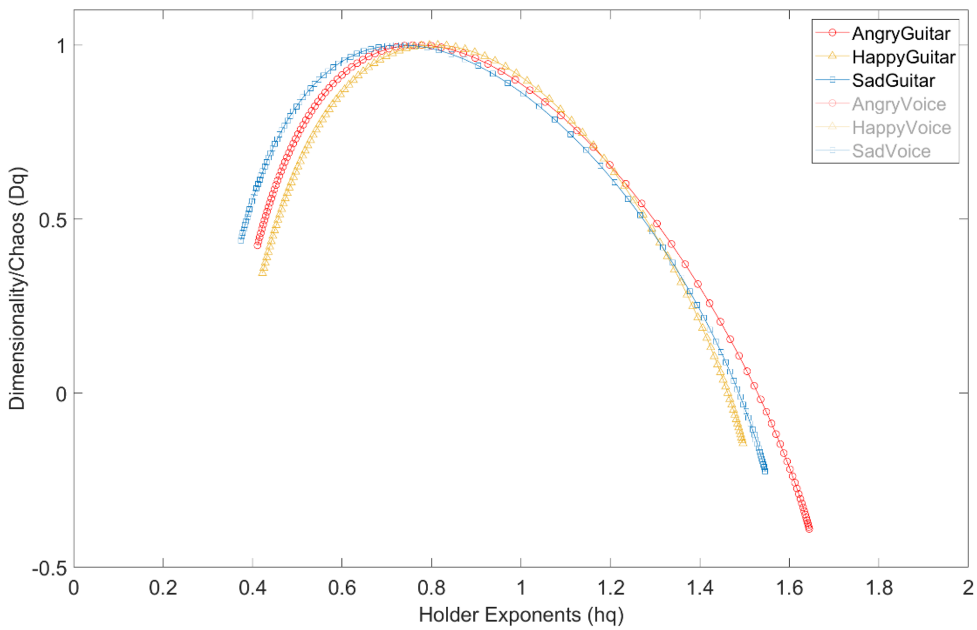

Figure 22. Multifractal spectra of EEG responses to guitar sounds of different emotions for participant 4.

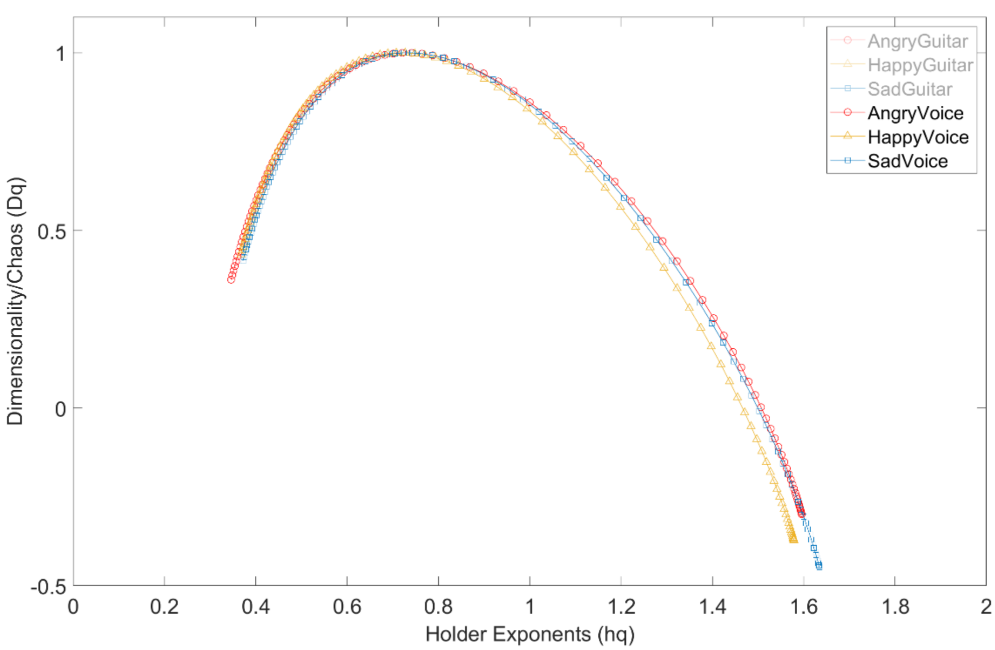

Figure 23. Multifractal spectra of EEG responses to voice sounds of different emotions for participant 4.

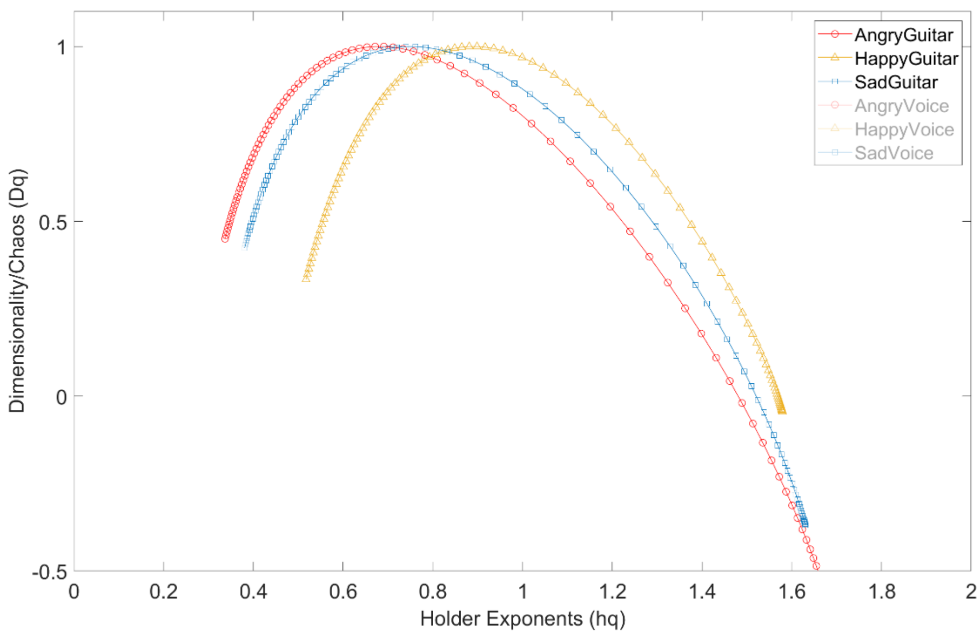

Figure 24. Multifractal spectra of EEG responses to guitar sounds of different emotions for participant 5.

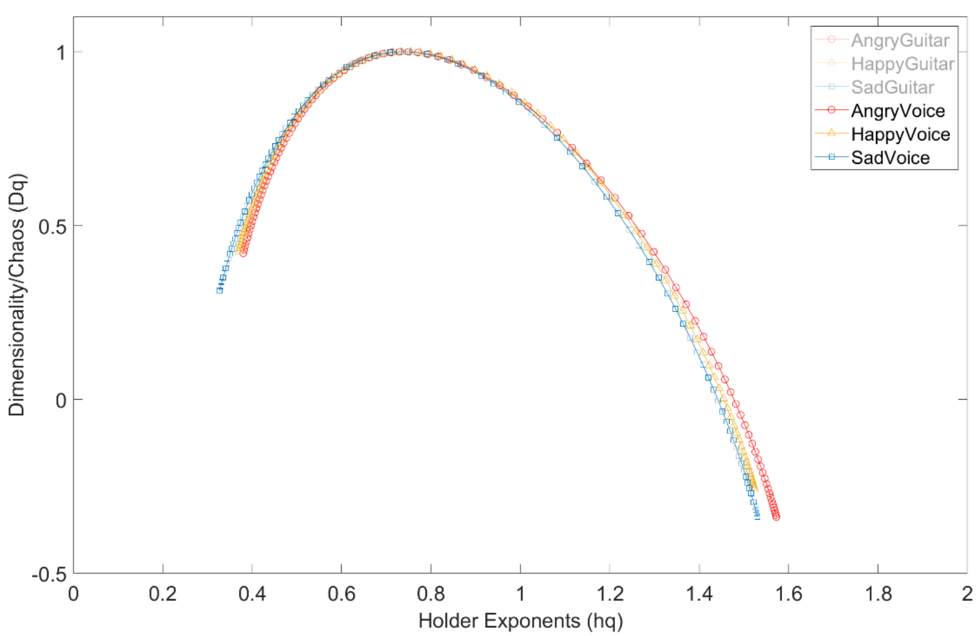

Figure 25. Multifractal spectra of EEG responses to voice sounds of different emotions for participant 5.

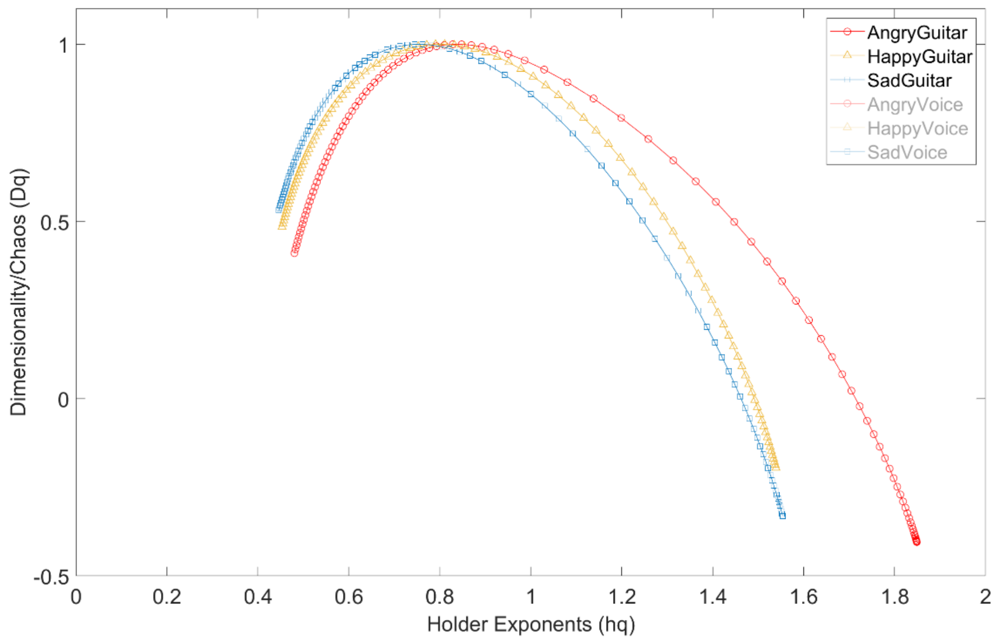

Figure 26. Multifractal spectra of EEG responses to guitar sounds of different emotions for participant 6.

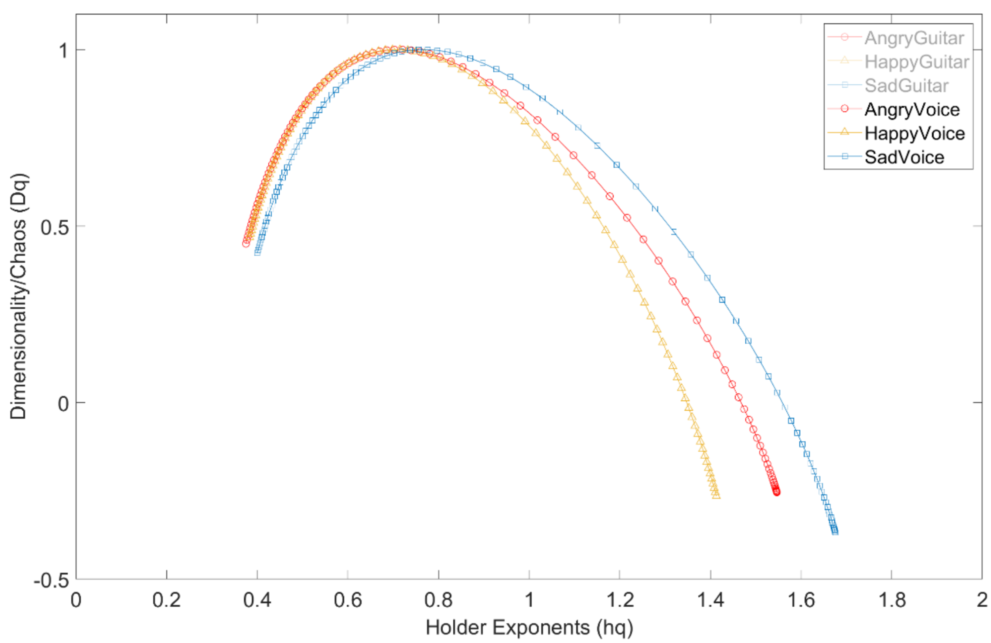

Figure 27. Multifractal spectra of EEG responses to voice sounds of different emotions for participant 6.

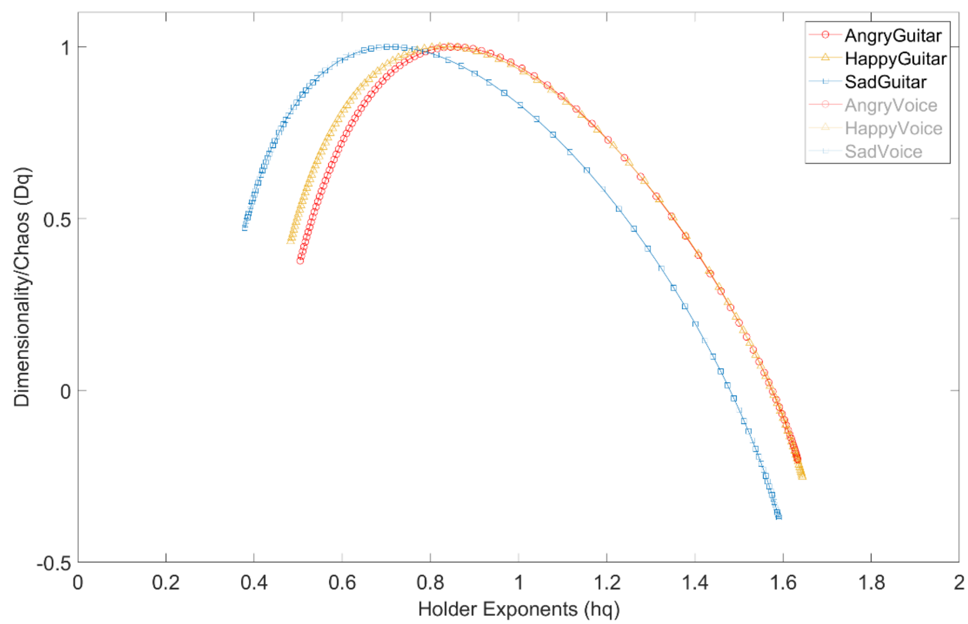

Figure 28. Multifractal spectra of EEG responses to guitar sounds of different emotions for participant 7.

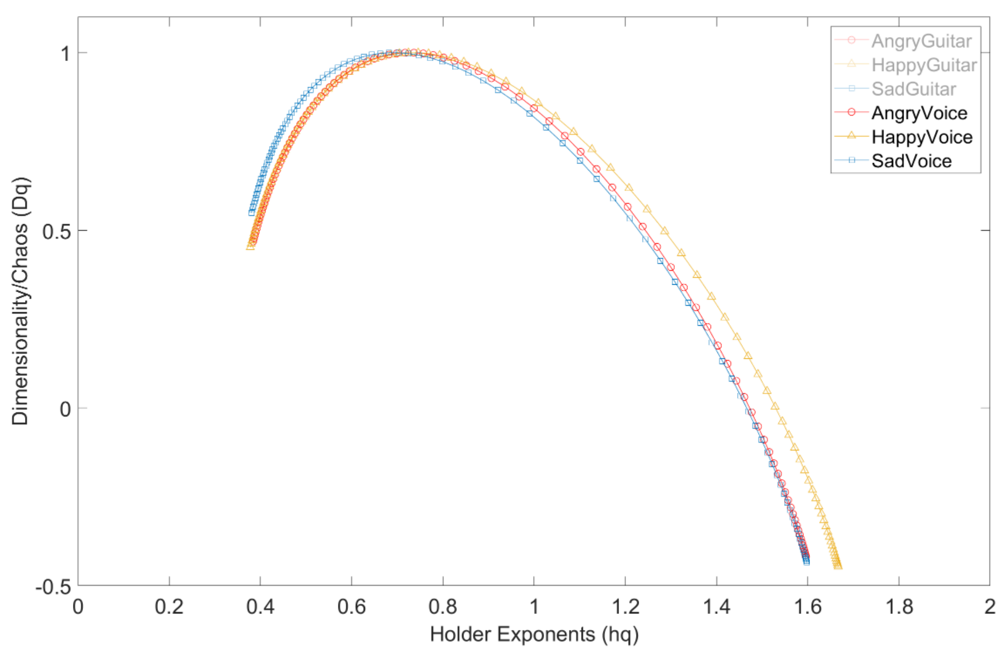

Figure 29. Multifractal spectra of EEG responses to voice sounds of different emotions for participant 7.

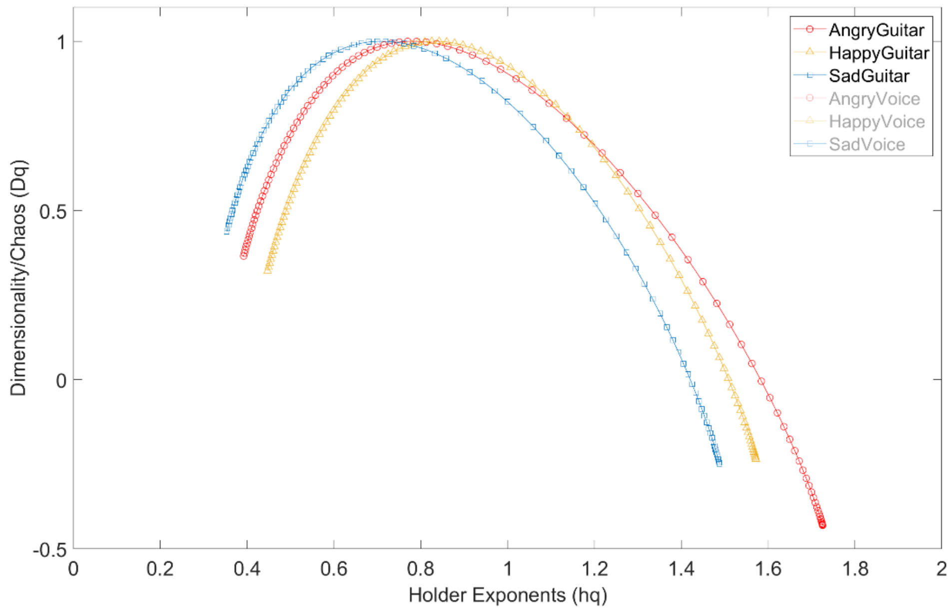

Figure 30. Multifractal spectra of EEG responses to guitar sounds of different emotions for participant 8.

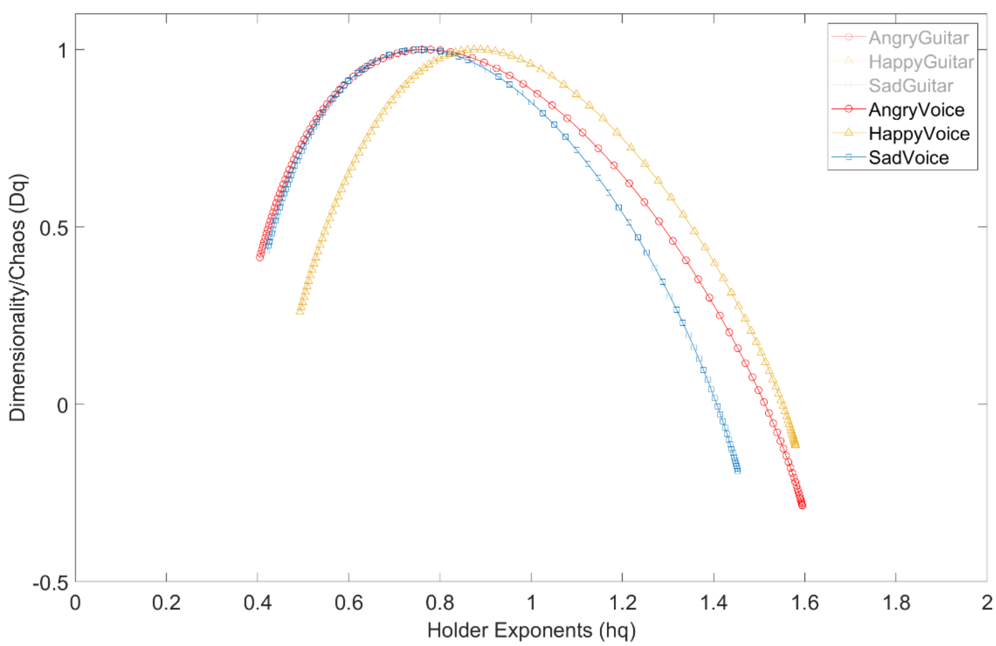

Figure 31. Multifractal spectra of EEG responses to voice sounds of different emotions for participant 8.

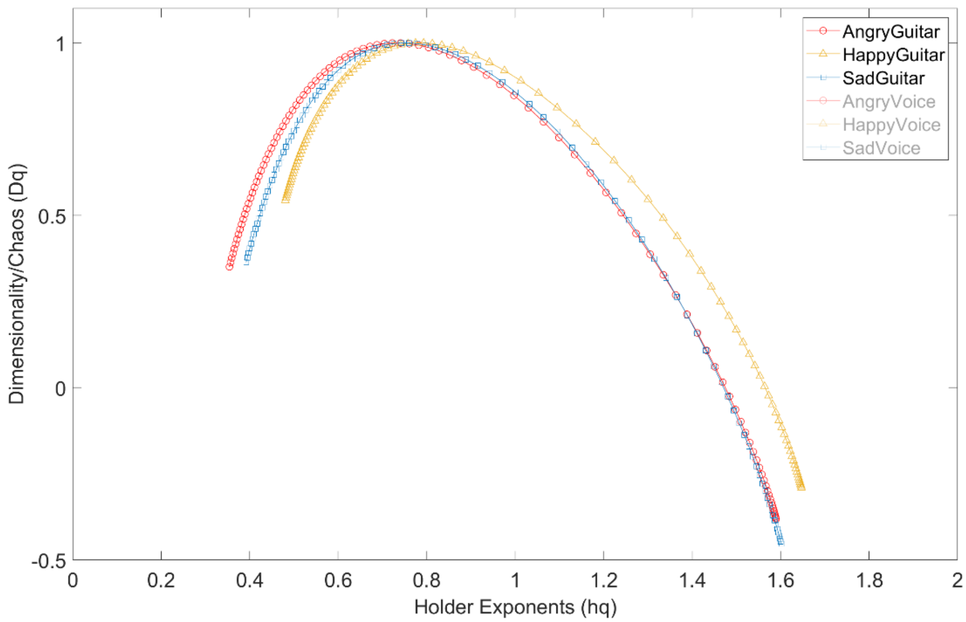

Figure 32. Multifractal spectra of EEG responses to guitar sounds of different emotions for participant 9.

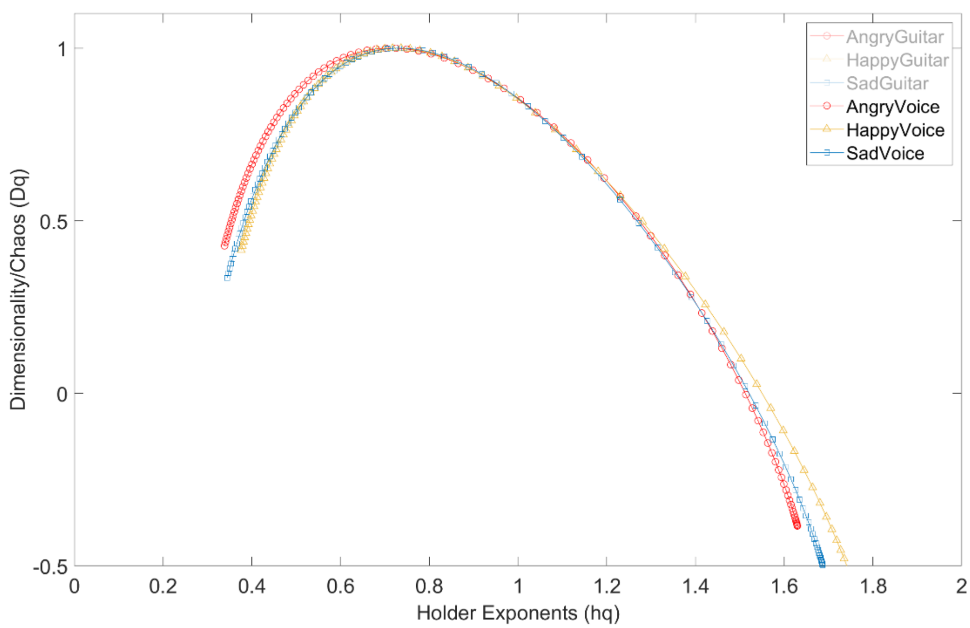

Figure 33. Multifractal spectra of EEG responses to voice sounds of different emotions for participant 9.

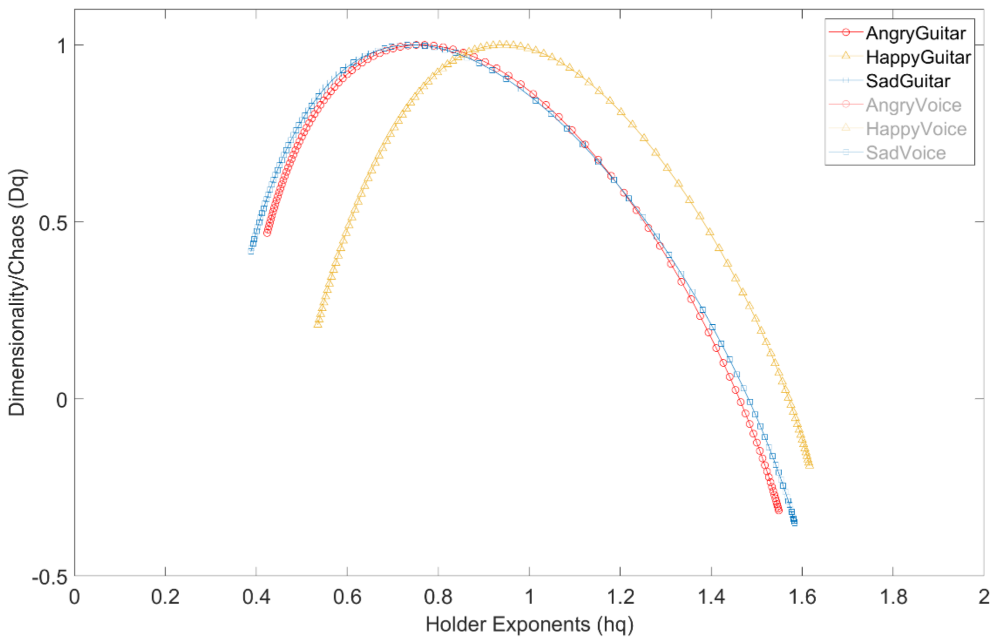

Figure 34. Multifractal spectra of EEG responses to guitar sounds of different emotions for participant 10.

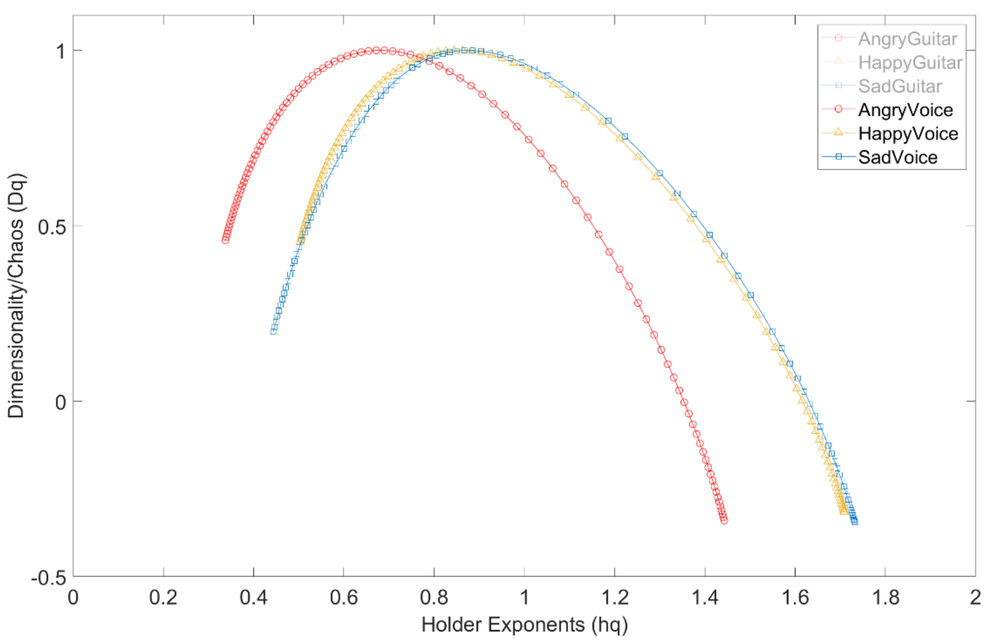

Figure 35. Multifractal spectra of EEG responses to voice sounds of different emotions for participant 10.

Appendix D: Algorithms and Code

MATLAB Code for Empirical Mode Decomposition (sample)

```
> levelForReconstruction = [true, true, true, false, false, false, false, false, true];
> [imf, residual] = emd(GuitAP15T10CzARRAY, ...
    'SiftRelativeTolerance', 0.25, ...
    'SiftMaxIterations', 100, ...
    'MaxNumIMF', 10, ...
    'MaxNumExtrema', 1, ...
    'MaxEnergyRatio', 20, ...
    'Interpolation', 'spline');
> mra = [imf residual].';
>GuitAP15T10CzARRAY2 = sum(mra(levelForReconstruction,:),1);
```

MATLAB Code for Wavelet Transformation (sample)

```
> levelForReconstruction = [false, false, false, true, false, false, false, false, false, false, true];
> wt = modwt(GuitAP15T10CzARRAY2, 'db4', 10);
> mra = modwtmra(wt, 'db4');
> GuitAP15T10CzAlpha = sum(mra(levelForReconstruction,:),1);
> levelForReconstruction = [false, false, false, false, true, false, false, false, false, false, true];
> wt = modwt(GuitAP15T10CzARRAY2, 'db4', 10);
> mra = modwtmra(wt, 'db4');
> GuitAP15T10CzTheta = sum(mra(levelForReconstruction,:),1);
```

MATLAB Code for MFDFA (sample)

```
> scmin=10;
> scmax=250;
> scres=10;
> exponents=linspace(log2(scmin), log2(scmax),scres);
> scale=round(2.^exponents);
> q=linspace(-5,5,101);
> [GuitAP10T10CzAlphaHq,GuitAP10T10CzAlphatq,GuitAP10T10CzAlphahq,
GuitAP10T10CzAlphaDq,GuitAP10T10CzAlphaFq]=MFDFA1(GuitAP10T10CzAlpha,
scale,q,1,1)
```

PEBL Code (sample)

```
define WaitForDownClickDuring()
{
part5data<-FileOpenAppend("EEGParticipant5.txt")
a<-GetTime()
click<-WaitForMouseButtonWithTimeout(1000)
b<-GetTime()
if((b-a)<1000)
{
Wait(1000-(b-a))
}
FilePrint(part5data,click)
}
```

```
define Start(p)
{
##PRACTICE
part5data<-FileOpenAppend("EEGParticipant5.txt")
win <- MakeWindow ()
inst <- EasyTextBox("",350,300,win,22,700,300)
inst.text <- "First, you will do a few rounds of practice. You will hear a series of sounds, in pairs. The first sound in each pair will be a guitar, and the second sound will be a voice. Each sound represents happiness, sadness, or anger. After hearing two (2) sounds, indicate whether the pair of sounds you just heard is emotionally similar (right click) or dissimilar (left click). Please be as still as possible. After responding to each pair of sounds, you may blink. Let the researcher know if you have questions before you begin the practice round. Click the mouse to continue."
WaitForDownClick ()
RemoveObject (inst,win)

practiceonset<-GetTime()
practicestamp<-Timestamp()
FilePrint(part5data,practiceonset)
FilePrint(part5data,practicestamp)
prime <- EasyTextBox("",400,300,win,22,600,200)
prime.text <- "                    WAIT


dissimilar (left click)              similar (right click)"
```

```
practice <-LoadSound("EEGAudioSessionPRACTICE.wav")

PlayBackground(practice)

Draw()

Wait(21000)


resp <- EasyTextBox("",400,300,win,22,600,200,"black","green")

resp.text <- "                              RESPOND


dissimilar (left click)                   similar (right click)"

Draw()

WaitForDownClickDuring()


prime <- EasyTextBox("",400,300,win,22,600,200)

prime.text <- "                              WAIT


dissimilar (left click)                   similar (right click)"

Draw()

Wait(21000)


resp <- EasyTextBox("",400,300,win,22,600,200,"black","green")

resp.text <- "                              RESPOND


dissimilar (left click)                   similar (right click)"
```

Draw()

WaitForDownClickDuring()

Appendix E: Supplementary Disc

See attached disc for the following data:

1. Hurst exponents for each participant (Hq)
2. Hölder exponents for each participant (hq)
3. Dimension values for each participant (Dq)
4. Raw Cz and Fz data for each participant
5. Sound files
6. Audio analysis data
7. Trial Information

Appendix F: "Letter to a Thesis Writer"

Creating and writing this thesis is one of the most challenging things I've done so far, so I thought it would be pertinent to share a few bits of wisdom with those who come after me (yes, it was my advisor's idea).

First, you are likely to feel some form of self-doubt, sometimes referred to as "imposter syndrome." In some ways, this can help you to do your work with more rigor, but only if you believe you are capable of doing it well. One professor during my time here told us all that we would not have made it into the program if someone did not believe we could complete it (thank you, Dr. Jeyaraj-Powell). I needed to hear that long before I did, so I'm telling you now.

Second, you must do your work to your own standards. The experience of many students is of teachers forcing them to do their best work at each step along the way. When you are a graduate student, that job is yours. You are the one who is most invested in your project, and even if others say it's great (or it stinks), you are the one who knows why it is interesting and valuable research. You will not be happy with your research if you know it is not as good as it could be. If you question whether you should make it just a little better, you should! The balance of that is that you have to know when you need to rest. Good advisors and teachers will push you, because they want you to do as well as you possibly can. Only you know where the line is between doing your best and having a panic attack. It might be a good idea to get a counselor if you don't have one yet; hopefully they can help you find that line. Be as consistent as you can with the things you know are important. Make time for friends, family, and faith (if you are of that persuasion) when you can. It won't be as much as you would like, but those supports will get you through. You can do this!