

A multivariate statistical analysis of EEG signals for differentiation of musicians and non-musicians

Estela Ribeiro¹, Carlos Eduardo Thomaz¹

¹Departamento de Engenharia Elétrica – Centro Universitário FEI (FEI)
Avenida Humberto Alencar Castelo Branco – São Bernardo do Campo – SP – Brazil

estela.eng@hotmail.com, cet@fei.edu.br

Abstract. *It is possible to reveal whether a subject received musical training through the neural activation patterns induced in response to music listening. We are particularly interested in analyzing the brain data on a global level, considering its activity registered in electroencephalogram electrodes signals. Our experiments results, with 13 musicians and 12 non-musicians who listened the song Hungarian Dance No 5 from Johannes Brahms, have shown that it is possible to differentiate musicians and non-musicians with high classification accuracy (88%). Given this multivariate statistical framework, it has also been possible to highlight the most expressive and discriminant changes in the participants brain according to the acoustic features extracted from the audio.*

1. Introduction

Music plays an important role in increasing the interconnection between our two brain hemispheres [Bennet and Bennet 2008], promoting neuroanatomical differences between musicians and non-musicians [Peretz and Zatorre 2004] and protection against cognitive decline related to aging, preserving perceptual auditory functions [Baird and Samson 2015].

In the last decade we have witnessed a significant increase in the number of scientific studies related to music perception, focused on understanding the neural processing of mainly artificial sounds. Such music processing relies on experiments that artificially manipulate the musical features according to the specifications of each particular experiment [Poikonen et al. 2016a, Poikonen et al. 2016b, Alluri et al. 2012, Sheng-Fu et al. 2011]. However, music has a complex and multidimensional nature, requiring a large neural demand [Vusst et al. 2014], presented dynamically and with a continuous flow of information [Alluri et al. 2012]. Thus, the most recent works have proposed using music in its naturalistic way, where the volunteer has listened to entire pieces of musics or some excerpts of musics [Alluri et al. 2012, Alluri et al. 2013, Abrams et al. 2013, Poikonen et al. 2016a, Poikonen et al. 2016b, Markovic et al. 2017]. Although these works have interestingly indicated that other brain areas are recruited to process music beyond those traditional findings, to the best of our knowledge all the statistical processing of these signals so far has been made on a local level [Sheng-Fu et al. 2011, Virtala et al. 2014, Vusst et al. 2014, Poikonen et al. 2016a, Poikonen et al. 2016b, Rigoulot et al. 2015, Markovic et al. 2017], looking for specific electrodes or regions of interest on the brain, based on the Event-Related Potentials (ERPs).

To the best of our knowledge, only a few works have used machine learning methods on data collected during music listening to predict whether the volunteers were musicians or not. Using EEG (electroencephalogram), it was possible to achieve a 97% accuracy [Sheng-Fu et al. 2011] in a traditional approach on a local level, using artificial stimuli (relation between consonant and dissonant intervals in music theory) and the data recorded at only one electrode to discriminate musicians and non-musicians. More recently, using fMRI (functional Magnetic Resonance Imaging), it was possible to achieve a 77% accuracy [Saari et al. 2018] in a naturalistic approach, acquiring the data from musicians and non-musicians during listening of three musical pieces from different genres, indicating the existence of musicianship-specific representation of music processing during listening.

In this work we have addressed this problem in a different way through a multivariate statistical perspective. In other words, we are interested in analyzing the brain data on a global level, considering its activity registered in all electrodes on a given time instant, and using this global information to differentiate musicians and non-musicians. Our main contribution is to understand how this information is changing in the original space of the presented brain data, looking not only for the most expressive changes, but also for the most discriminant ones, according to the acoustic features extracted from the audio.

The paper is organized as follows. In section 2 we describe the four-step multivariate statistical framework proposed here for differentiate the two-sampled groups of volunteers. In section 3, a brief description of the data acquisition and the experimental results carried out are discussed. Finally, in section 4, we summarize the main contributions of the paper and discuss future improvements.

2. Methods

Our framework consists of the following four parts: Acoustic features extraction; Trigger selection; EEG signal processing; and Multivariate statistical analysis. This framework, illustrated in Figure 1, was built from previous works of EEG music processing [Poikonen et al. 2016a, Lartillot 2014, Lerch 2012] and high-dimensional data analysis [Sato et al. 2008, Xavier et al. 2015, Gregori et al. 2017].

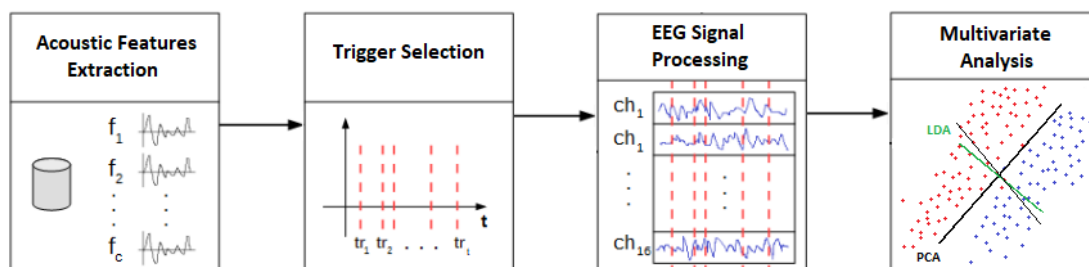


Figure 1. A schematic illustration of the four main parts of the framework proposed to statistically analyze the EEG high-dimensional brain data, been them: (I) Acoustic features extraction; (II) Trigger selection; (III) EEG signal processing; and (IV) Multivariate statistical analysis.

2.1. Acoustic Feature Extraction

In this step the low-level acoustic features that describe the audio signal used in this work are extracted using MIRtoolbox (version 1.6.1) [Lartillot 2014]. To perform such extraction, the signal is decomposed into a 50 ms windows with 50% of overlap. In this way, instead of analyzing the audio signal as a whole, we will analyze some features that compose and describe it quantitatively, which do not necessarily present properties that can be directly interpreted [Lerch 2012], but the changes that occur in them are capable of generating significant neural responses (ERPs) [Poikonen et al. 2016a].

Following [Poikonen et al. 2016a], we have used the acoustic feature called brightness. This feature is defined as the amount of spectral energy above a threshold value (default: 1500 Hz) [Lartillot 2014] where high values of brightness mean that a high percentage of spectral energy is concentrated in the higher end of the frequency spectrum [Poikonen et al. 2016a], whereas low values indicates moments where most of the sound energy is on the low-frequency spectrum.

2.2. Trigger Selection

According to [Poikonen et al. 2016a], rapid changes on the acoustic features extracted from the audio are capable to elicit sensory components similarly to those found in traditional methods. These specific changes, called triggers, are instants in the time series generated by the extracted features where occurs a high-contrast.

To identify the triggers, the upper threshold V_{p+} and lower threshold V_{p-} are determined from the mean values of the acoustic feature, where a trigger occurs when the signal remains below V_{p-} for a minimum interval of time called Preceding Low-Feature Phase (PLFP), followed by an ascendant phase where the signal reaches V_{p+} [Poikonen et al. 2016a]. The parameters used for trigger selection can be changed and the more rigid they are, the greater the acoustic contrast identified and, as a consequence, the lower their probability of occurrence.

2.3. EEG Signal Processing

We have used the OpenBCI EEG device to acquire the brain electrical signals. This device is composed of 16 dry electrodes, positioned according to the international 10-20 system [Marcuse et al. 2016], and 2 electrodes placed on each earlobe as reference, with a sampling rate of 127 Hz. The EEG signal is preprocessed using a bandpass Butterworth filter of 1-30 Hz. The continuous EEG data were separated into epochs according to the triggers, and averaged separately for each individual and electrode. The epochs started 100 ms before the trigger and ended 300 ms after the trigger. The baseline has been defined according to the 100 ms time period before the trigger. Visually detected EEG channels with noise were removed from analysis and to remove artifacts, the epochs with amplitudes above than $100\mu\text{V}$ (in absolute values) have been rejected.

2.4. Multivariate Analysis

In this step we firstly describe the data matrix \mathbf{X} of the epoch EEG signals at 100 ms post stimulus in a way that the corresponding sampled input data can be treated as a high-dimensional point in a multivariate space, as follows:

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_N \end{bmatrix} = \begin{bmatrix} x_{11} & \dots & x_{1n} \\ x_{21} & \dots & x_{2n} \\ \vdots & \vdots & \vdots \\ x_{N1} & \dots & x_{Nn} \end{bmatrix}, \quad (1)$$

where N is the number of volunteers and n is the number of electrodes ($n = 16$ here).

With such data representation we are assuming that our EEG multivariate statistical analyses might include the electrical potentials at the predefined timestamps to understand the complexity of the brain activity not only in terms of the N sampling, but also in terms of the n electrodes simultaneously.

2.4.1. Principal Component Analysis

Analogously to others [Gregori et al. 2017, Xavier et al. 2015, Sato et al. 2008] we have used PCA to highlight the most expressive changes in terms of variance information of the electrical potentials.

Let the $N \times n$ matrix \mathbf{X} described in equation (1). The mathematical equation relating the sample covariance matrix \mathbf{S} of \mathbf{X} to its eigenvectors \mathbf{P} and eigenvalues Λ is

$$\mathbf{P}^T \mathbf{S} \mathbf{P} = \Lambda. \quad (2)$$

Such a set of eigenvectors ranked in decreasing order of their corresponding eigenvalues ($\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$) is known as the principal components $\mathbf{P} = (\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m)$, where $m \leq n$.

After obtaining the m principal components, it is possible to navigate along the axes of such dimensions to capture the most expressive changes of the data matrix \mathbf{X} [Xavier et al. 2015]. This navigation process can be mathematically described as

$$\mathbf{y} = \bar{\mathbf{x}} + (j\sqrt{\lambda_i}) \times \mathbf{p}_i, \quad (3)$$

where $j \in \{-3, -2, -1, 0, 1, 2, 3\}$ and $\bar{\mathbf{x}}$ is the sample mean vector given by

$$\bar{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^N \bar{\mathbf{x}}_i. \quad (4)$$

2.4.2. Linear Discriminant Analysis

A well-known multivariate statistical method that identifies the most discriminant dimension for separating two-sample groups like musicians and non-musicians is the Linear Discriminant Analysis (LDA) [Fukunaga 1994].

Let the scatter matrices between-class \mathbf{S}_b and within-class \mathbf{S}_w be defined, respectively, as

$$\mathbf{S}_b = \sum_{i=1}^g N_i (\bar{\mathbf{x}}_i - \bar{\mathbf{x}})^T (\bar{\mathbf{x}}_i - \bar{\mathbf{x}}), \quad (5)$$

$$\mathbf{S}_w = \sum_{i=1}^g (N_i - 1) \mathbf{S}_i = \sum_{i=1}^g \sum_{j=1}^{N_i} (\mathbf{x}_{i,j} - \bar{\mathbf{x}}_i)^T (\mathbf{x}_{i,j} - \bar{\mathbf{x}}_i), \quad (6)$$

where $\mathbf{x}_{i,j}$ is the n -dimensional sample j from class i , N_i is the number of training samples from class i , g represents the total number of classes ($g = 2$, $N_1 = 13$ and $N_2 = 12$, here) and $\bar{\mathbf{x}}_i$ is the sample group mean vector.

The Fisher's criterion is maximized when the projection vector \mathbf{w} is composed by the leading eigenvector of $\mathbf{S}_w^{-1} \mathbf{S}_b$ with nonzero corresponding eigenvalue [Fukunaga 1994, Xavier et al. 2015].

After obtaining the eigenvector \mathbf{w} , it is possible to navigate along the projection of the corresponding vector and extract the discriminant differences from the data set, assuming its dispersion follows a Gaussian distribution [Sato et al. 2008]. This navigation process in the most discriminating projection is given by

$$\mathbf{y} = \bar{\mathbf{x}} + (j\sigma_i + \bar{x}_i) \times \mathbf{w}, \quad (7)$$

where $j \in \{-1, 0, 1\}$, $\bar{\mathbf{x}}$ is the global mean, and σ_i and \bar{x}_i are the standard deviation and the mean of each class $i \in \{1, 2\}$ on the LDA space, respectively.

3. Results

In this work we used the 10-fold cross-validation scheme, where 10% of the data was used as test and 90% as training, performing a simple Euclidean mean classifier on the data projected on the most discriminant axis found at each iteration of the LDA.

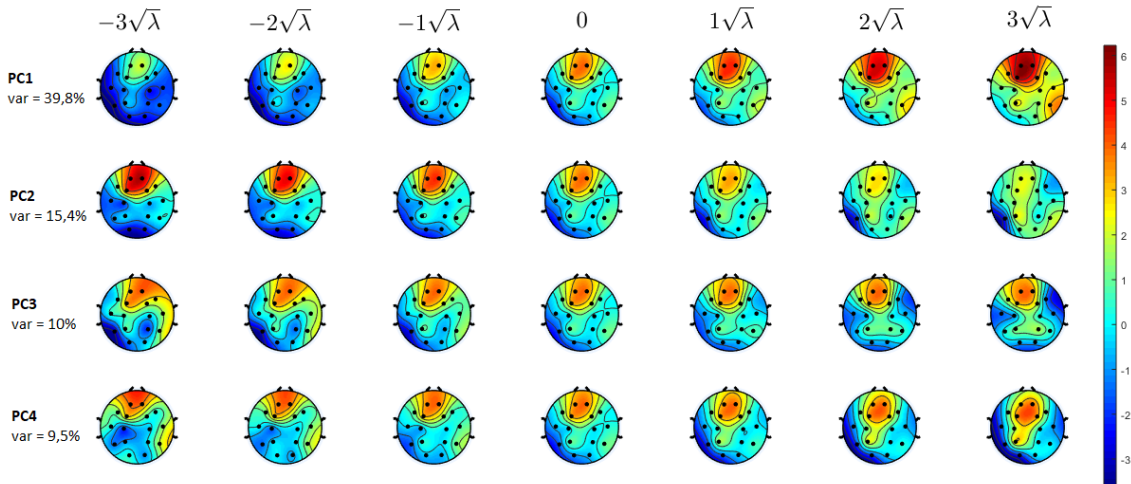


Figure 2. Brain mapping of each principal component (PC). Hyperplane navigation of the first four principal components captured by PCA at the latency of 100 ms post stimulus.

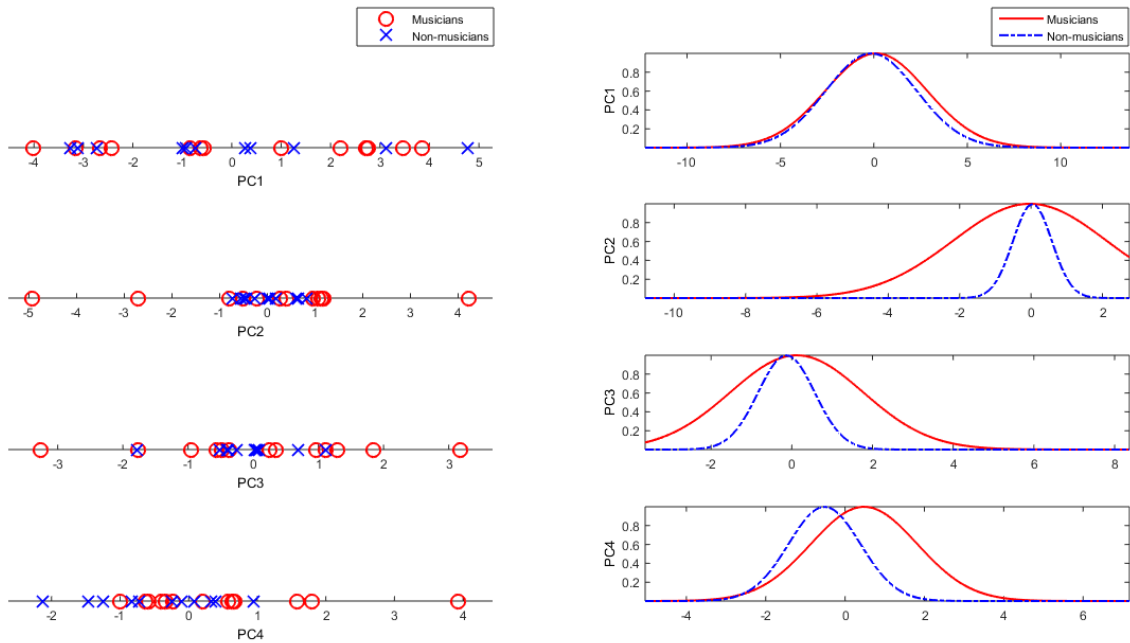


Figure 3. Projection of the data on each principal component (PC) axes. Hyper-plane navigation of the first four principal components captured by PCA at the latency of 100 ms post stimulus.

Twenty five native Brazilians took part in the experiment (16 male and 9 female), 13 amateur musicians (26, 8 ± 6 , 3 years) and 12 non-musicians (31, 3 ± 7 , 8 years). All participants gave a written informed consent to participate in the study. The group of musicians have played their main instruments on average for $14,6 \pm 5,79$ years, their musical styles varied among classic, pop and rock, and their main instruments varied among drums, guitar, flute, violin and saxophone.

As stimulus, the 3.2 min classical music Hungarian Dance n.5 of Johannes Brahms were used, recorded in a concert by the Budapest Symphonic Orchestra, presented via intra earphone. The subjects were instructed to listen to music as still as possible, while their EEG signals were recorded.

We have applied a $\pm 20\%$ of the mean value of the acoustic feature brightness to define the upper threshold V_{p+} and lower threshold V_{p-} , respectively, to determine the triggers, finding 8 triggers along the music used as stimulus.

The topographic maps and the projection of the data generated by PCA are shown in Figures 2 and 3. At Fig. 2 we can see the changes of the brain responses along each principal component (PC), disclosing major electrical potential differences in several parts of the brain. However, when the data are projected on the axis of each principal component for classification, as can be seen in Fig. 3, using a simple Euclidean mean classifier, it's not possible to see a distinction between the musicians and non-musicians sample groups (classification rates around chance level).

Therefore, although PCA finds the directions that the data most variate, they are not necessarily the ones that best discriminate the sample group of volunteers. Never-

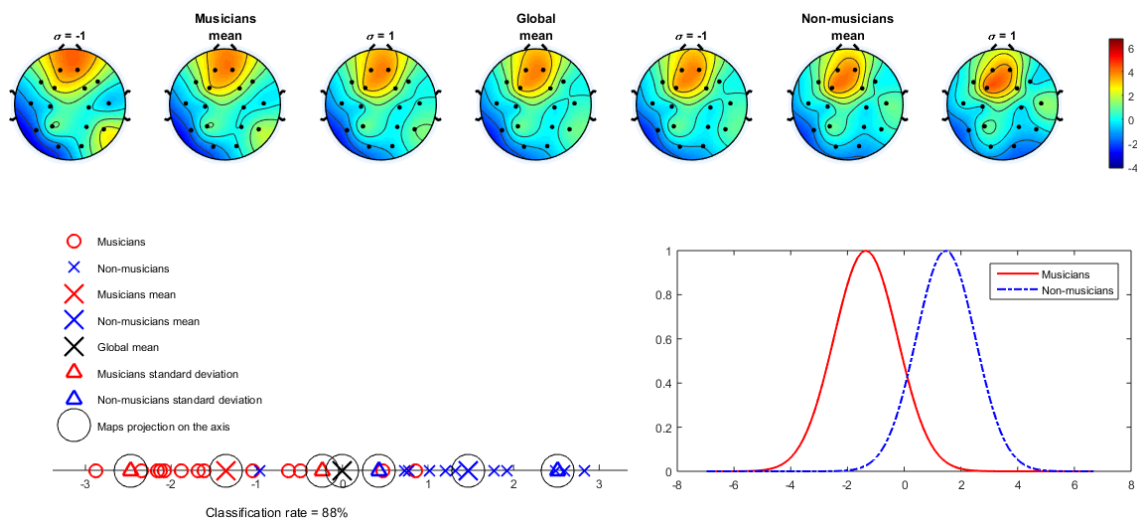


Figure 4. Hyperplane navigation of the most discriminant axis captured by LDA at the latency of 100 ms post stimulus: (Top) Reconstruction of the mean brain mapping when navigating along the most discriminant axis between the sample group of musicians (left) and non-musicians (right); (Bottom) Projection of the data on the discriminant axis.

theless, it's interesting to observe that in the first 4 principal components the data most variate at the pre-frontal and frontal regions of the cortex and at the 4th principal component it begins to show a separation on the sample groups. This results indicate that these regions are important for the processing of the acoustic feature extracted from the audio used as stimulus.

Differently from PCA, the topographic maps generated by the navigation on the most discriminant hyperplane of LDA, shown in Figure 4, present visually subtle electrical potential differences between the groups, but such differences are definitely discriminant (where musicians are placed at the left side of the axis and the non-musicians are placed at the right side) because using the same Euclidean mean classifier with a 10-fold cross-validation scheme it has been possible to achieve a mean classification rate equals to 88% ($\pm 4, 6$).

Although it would be expected a similar discriminability among the two methods (PCA and LDA) given that there is a similarity between their brain maps, these results show that what discriminates the sample groups is actually the pattern of cortical activation within the pre-frontal region. Thus, PCA show us where to look at the topographic maps and LDA show us how to look at it.

These results also revealed that the brain areas that best predict musicianship exist predominantly in the frontal areas of the brain, which corroborate to previous works [Saari et al. 2018] on how the brain processes certain acoustic characteristics of music.

Additionally, it's important to notice that the volunteer classified as the most distinct musician is a drummer that plays pop/rock music and the musician that was placed among the non-musicians is a pianist that plays jazz. Also, the non-musician that were placed among the musicians had piano classes for almost a year during her childhood. All the musicians in this work are amateurs and they don't play their instruments profession-

ally, but this results can suggest that even a short contact with music can change the way that the brain processes the music during listening.

4. Conclusion

This work proposed and implemented a multivariate statistical computational framework to differentiate musicians and non-musicians based on the most variant and discriminant features obtained through the analysis of the brain data on a global level. It describes in mathematical and computational terms the differences of the musical processing between musicians and non-musicians, producing a plausible linear separation among the groups based on high dimensional encoding of variance and discriminant information from sample groups.

Our results show that the Poikonen method [Poikonen et al. 2016a] to identify triggers in musics can be very helpful to differentiate the musicality of a volunteer based on the electrical potential register in the electrodes. Here we only used one acoustic feature (brightness), but there are other acoustic features that can describe an audio. Further research is needed to specify the influence of the acoustic features in this approach and the music-related background of the volunteers.

We believe that this multivariate statistical analysis generates a more comprehensive description of the EEG differences on the music processing information. Further experiments would be necessary though to extend the approach proposed exploring other acoustic features and understanding how they affect these results.

Acknowledgment

This work was supported by FEI and the Brazilian agency CAPES and CNPq (INCT MACC, process 465586/2014-7). We would like to thanks all the volunteers that took part of the experiments.

References

- Abrams, D. A., Ryali, S., Chen, T., Chordia, P., Khouzam, A., Levitin, D. J., and Menon, V. (2013). Inter-subject synchronization of brain response during natural music listening. *European Journal of Neuroscience*, 37:1458–1469.
- Alluri, V., Toiviainen, P., Jaaskelainen, I. P., Glerean, E., Sams, M., and Brattico, E. (2012). Large-scale brain networks emerge from dynamics processing of musical timbre, key and rhythm. *NeuroImage*, 59:3677–3689.
- Alluri, V., Toiviainen, P., Lund, T. E., Wallentin, M., Vuust, P., Nandi, A. K., Ristaniemi, T., and Brattico, E. (2013). From vivaldi to beatles and back: Predicting lateralized brain responses to music. *NeuroImage*, 83:627–636.
- Baird, A. and Samson, S. (2015). Music and dementia. *Progress in brain research*, 217:207–235.
- Bennet, A. and Bennet, D. (2008). The human knowledge system: music and brain coherence. *Information and knowledge management systems*, 38(3).
- Fukunaga, K. (1994). *Introduction to statistical pattern recognition*. Morgan Kaufmann, 2nd edition.

- Gregori, I. R. S., Sanches, I., and Thomaz, C. E. (2017). Clutch judder classification and prediction: A multivariate statistical analysis based on torque signals. *IEEE transactions on industrial electronics*, 64:4287–4295.
- Lartillot, O. (2014). *MIRtoolbox 1.6.1 Users Manual*. Department of Architecture, Design and Media Technology, 1st edition.
- Lerch, A. (2012). *An Introduction to audio content analysis. Applications in signal processing and music informatics*. IEEE Press, 1st edition.
- Marcuse, L. V., Fields, M. C., and Yoo, J. (2016). *Rowan's Primer of EEG*. Elsevier, 2nd edition.
- Markovic, A., Kuhnis, J., and Janche, L. (2017). Task context influences brain activation during music listening. *Frontiers in Human Neuroscience*, 1.
- Peretz, I. and Zatorre, R. J. (2004). Brain organization for music processing. *Annual Reviews Psychology*, 56:89–114.
- Poikonen, H., Alluri, V., Brattico, E., Lartillot, O., Tervaniemi, M., and Huotilainen, M. (2016a). Event-related brain responses while listening to entire pieces of music. *Neuroscience*, pages 58–73.
- Poikonen, H., Toiviainen, P., and Tervaniemi, M. (2016b). Early auditory processing in musicians and dancers during a contemporary dance piece. *Scientific Reports - Nature*.
- Rigoulot, S., Pell, M. D., and Armony, J. L. (2015). Time course of the influence of musical expertise on the processing of vocal and musical sounds. *Neuroscience*, 290.
- Saari, P., Burunat, I., Brattico, E., and Toiviainen, P. (2018). Decoding musical training from dynamic processing of musical features in the brain. *Scientific Report*, 708:1–12.
- Sato, J. R., Thomaz, C. E., Cardoso, E. F., Fujita, A., Martin, M. d. G. M., and Amaro, E. J. (2008). Hyperplane navigation: A method to set individual scores in fmri group datasets. *NeuroImage*, 42:1472–1480.
- Sheng-Fu, L., Tsung-Hao, H., Wei-Hong, C., and Kuei-Ju, L. (2011). Classification of eeg signals from musicians and non-musicians by neural networks. *World Congress on Intelligent Control and Automation, IEEE*, 8:865–869.
- Virtala, P., Huotilainen, M., Partanen, E., and Tervaniemi, M. (2014). Musicianship facilitates the processing of western music chords - an erp and behavioral study. *Neuropsychologia*, 61:247–258.
- Vusst, P., Brattico, E., Sspanen, M., Naatanen, R., and Tervaniemi, M. (2014). The sound of music: Differentiating musicians using a fast, musical multi-feature mismatch negativity paradigm. *Neuropsychologia*, 61:1432–1443.
- Xavier, I., Pereira, M., Giraldi, G., Gibson, S., Solomon, C., Rueckert, D., Gillies, D., and Thomaz, C. (2015). A photo-realistic generator of most expressive and discriminant changes in 2d face images. *International conference on emerging security technologies*, pages 80–85.