CENTRO UNIVERSITÁRIO FEI

ESTELA RIBEIRO

UM ESTUDO SOBRE PREDIÇÃO DE MUSICALIDADE POR MEIO DA ANÁLISE DE SINAIS DE EEG

São Bernardo do Campo
2017

ESTELA RIBEIRO

UM ESTUDO SOBRE PREDIÇÃO DE MUSICALIDADE POR MEIO DA ANÁLISE DE SINAIS DE EEG

Dissertação de Mestrado, apresentada ao Centro Universitário FEI para obtenção do título de Mestre em Engenharia Elétrica. Orientada pelo Prof. Dr. Carlos Eduardo Thomaz.

São Bernardo do Campo
2017

Ribeiro, Estela.
Um estudo sobre predição de musicalidade por meio da análise de sinais de EEG / Estela Ribeiro. São Bernardo do Campo, 2017.
114 p. : il.

Dissertação - Centro Universitário FEI.
Orientador: Prof. Dr. Carlos Eduardo Thomaz.

1. EEG. 2. k-NN. 3. Redes Neurais. 4. Música. I. Eduardo Thomaz, Carlos, orient. II. Título.

Elaborada pelo sistema de geração automática de ficha catalográfica da FEI com os dados fornecidos pelo(a) autor(a).

centro universitário FEI

APRESENTAÇÃO DE DISSERTAÇÃO
ATA DA BANCA EXAMINADORA

Programa de Pós-Graduação Stricto Sensu em Engenharia Elétrica

Mestrado

PGE-10

Aluno: Estela Ribeiro **Matrícula:** 115310-5

Título do Trabalho: Um estudo sobre predição de musicalidade por meio da análise de sinais de EEG.

Área de Concentração: Processamento de Sinais

Orientador: Prof. Dr. Carlos Eduardo Thomaz

Data da realização da defesa: 12/09/2017

ORIGINAL ASSINADA

Avaliação da Banca Examinadora:

São Bernardo do Campo, 12 / 09 / 2017.

MEMBROS DA BANCA EXAMINADORA	
Prof. Dr. Carlos Eduardo Thomaz	Ass.: _____
Prof.ª Dr.ª Ilza Zenker Leme Joly	Ass.: _____
Prof. Dr. Ivandro Sanches	Ass.: _____

A Banca Julgadora acima-assinada atribuiu ao aluno o seguinte resultado:

APROVADO ☒ REPROVADO ☐

VERSÃO FINAL DA DISSERTAÇÃO

APROVO A VERSÃO FINAL DA DISSERTAÇÃO EM QUE FORAM INCLUÍDAS AS RECOMENDAÇÕES DA BANCA EXAMINADORA

_____

Aprovação do Coordenador do Programa de Pós-graduação

_____

Prof. Dr. Carlos Eduardo Thomaz

À minha família.

AGRADECIMENTOS

A DEUS, que tem sido a minha Rocha em todos os tempos, pela oportunidade de cursar este mestrado e por me dar forças para nunca desistir, por me dar direção nas horas incertas, me ensinando que tudo é possível bastando apenas acreditar e por estar ao meu lado o tempo todo.

Aos meus pais, Edivaldo e Dilce, que sempre estiveram presentes, me incentivaram, me apoiaram e acreditaram em mim, me ensinando o quão importante é estudar e se esforçar para alcançar os meus sonhos. Eu não teria chegado até aqui se não fossem vocês.

A minha irmã Angela, que sempre acreditou e se orgulhou de mim, me incentivando a seguir em frente, por mais difícil que as coisas fossem. Obrigada por estar sempre ao meu lado.

Ao meu orientador, Prof. Dr. Carlos Eduardo Thomaz, pela disponibilidade, apoio, confiança, paciência, motivação e sugestões, me dando uma "luz" quando os problemas pareciam difíceis demais para encontrar uma solução. Foi o senhor que me inspirou a seguir essa linha de pesquisa e a sua participação foi fundamental para a realização deste trabalho.

Aos professores do Centro Universitário da FEI pelos ensinamentos e pelas aulas ministradas e por estarem sempre disponíveis. Meu sincero agradecimento.

Ao meu querido maestro Roberto Farias que tem sido fonte de inspiração em todo o tempo, me motivando cada vez mais a seguir esta linha de pesquisa e me fazendo enxergar a beleza da música e os desafios que os músicos enfrentam na atualidade.

Aos meus amigos, colegas e familiares que me apoiaram e acreditaram em mim, em especial a Erika e a Thais que me ajudaram durante o processo de revisão desta dissertação. Eu não teria terminado sem a sua ajuda, muito obrigada.

Aos meus amigos do mestrado, pelos momentos divididos juntos, encontrando dificuldades a nossa frente, mas superando todas elas. Foi bom poder contar com vocês.

A todos os voluntários que contribuíram para a minha pesquisa, disponibilizando o seu tempo para participarem dos meus experimentos. Meus sinceros agradecimentos.

A CAPES e o Centro Universitário FEI pela bolsa de estudos de mestrado.

“Never stop fighting until you arrive at your destined place - that is, the unique you. Have an aim in life, continuously acquire knowledge, work hard, and have perseverance to realise the great life.”

A. P. J. Abdul Kalam

RESUMO

Devido a sua natureza multidimensional e complexa, a música requer uma grande demanda neural e o processo de plasticidade faz com que haja diferenças estruturais entre o cérebro de pessoas que praticam (músicos) e não praticam (não-músicos) instrumentos musicais. Estudos recentes têm sido realizados a fim de compreender estas diferenças e a forma como o cérebro processa a música. Músicas são capazes de extrair os componentes encontrados em estudos convencionais de potencial evocado em Eletroencefalograma (EEG) e, a partir da seleção de características acústicas extraídas dos áudios musicais, é possível determinar os instantes no tempo da música onde há alto contraste (*triggers*), considerados como estímulos na análise do sinal de EEG. O principal objetivo deste trabalho é desenvolver um arcabouço computacional capaz de classificar, por meio de aprendizado supervisionado, os sinais de EEG registrados de um grupo de voluntários em músicos e não-músicos. Este arcabouço é composto pelas seguintes 6 etapas: (1-2) Extração e seleção de características acústicas; (3) Seleção de *triggers*; (4-5) Processamento e análise dos sinais de EEG; e (6) Classificação. Experimentos foram realizados com 26 voluntários, sendo 13 não-músicos e 13 músicos amadores, que ouviram duas músicas instrumentais clássicas (Dança Húngara No. 5, de Johannes Brahms, e O Barbeiro de Sevilha - Abertura, de Gioachino Rossini), enquanto era realizada a aquisição dos sinais de EEG, utilizando eletrodos a seco. Para a classificação foram propostos dois cenários que compuseram a entrada dos classificadores, variando a latência em que se observava o sinal de EEG. Como classificadores, foram utilizados o k-NN e a Rede Neural. Foi possível obter como resultado uma taxa de acerto na classificação acima de 80% para ambas as músicas utilizando o k-NN.
Palavras-chave: EEG. k-NN. Redes Neurais. Música.

ABSTRACT

Owing to its multidimensional nature and complexity, music requires a higher neural demand and the process of plasticity causes structural differences between the brains of people who practice (musicians) and do not practice (non-musicians) musical instruments. Recent studies have been conducted in order to understand these differences and the way the brain processes music. Musics are able to extract the components found in conventional studies of evoked potential in Electroencephalogram (EEG) and, from the selection of acoustic features extracted from the audio signals, it is possible to determine the time instants of the music where there is a high contrast (triggers), considered as stimuli in EEG signal analysis. The main objective of this work is to develop a computational model capable of classifying, through supervised learning, the EEG signals recorded by a group of volunteers in musicians and non-musicians. This model is composed of the following 6 steps: (1-2) Extraction and selection of acoustic features; (3) Selection of triggers; (4-5) Processing and analysis of EEG signals; and (6) Classification. Experiments were carried out with 26 volunteers, 13 non-musicians and 13 amateur musicians, who listened to two classical music (Hungarian Dance No.5, from Johannes Brahms, and The Barber of Seville – Overture, from Gioachino Rossini), while it was performed the acquisition of EEG signals using dry electrodes. For the classification were proposed two scenarios that composed the input data of the classifiers, varying the latency in which the EEG signal was observed. As classifiers, the k-NN and the Neural Network were used. It was possible to obtain as a result an assertive rate above 80% for both songs using the k-NN.
Keywords: EEG. k-NN. Neural Network. Music.

LISTA DE ILUSTRAÇÕES

Figura 1 – Composição do Encéfalo 34
Figura 2 – Estrutura de um neurônio biológico 37
Figura 3 – Tipos de Neurônios do SNC classificados de acordo com sua Morfologia . 38
Figura 4 – O Potencial de Ação 40
Figura 5 – Transmissão Sináptica 42
Figura 6 – Geração do Sinal Elétrico do EEG. 44
Figura 7 – Sistema 10-20 de posicionamento de Eletrodos 45
Figura 8 – Registro Referencial 46
Figura 9 – Sequência típica dos componentes ERP a partir de estímulos auditivos . . 50
Figura 10 – Exemplo de aplicação da média do sinal de EEG entre testes 52
Figura 11 – Exemplo de aplicação da média do sinal de EEG entre sujeitos 53
Figura 12 – Produção do som pela variação da pressão do ar 55
Figura 13 – Orelha externa, média e interna 56
Figura 14 – A estrutura da Cóclea 57
Figura 15 – Vias auditivas centrais 59
Figura 16 – Fluxograma da metodologia de classificação de sinais de EEG 63
Figura 17 – Decomposição em janelas da forma de onda de um sinal de áudio 65
Figura 18 – Frequência de corte para a determinação do *Brigtness* 67
Figura 19 – Resposta do filtro gerador das 10 Sub-bandas do Spectral Flux 70
Figura 20 – Ilustração da ocorrência de um *trigger*, assumindo as condições propostas 74
Figura 21 – Ilustração de um sinal de EEG dividido em 6 épocas, delimitadas em 200ms pré-estímulo e 800ms pós-estímulo, para 3 características acústicas distintas 75
Figura 22 – Ilustração do funcionamento do algoritmo: "Moving Window peak-to-peak Threshold" 76
Figura 23 – Artefatos marcados para remoção no sinal de EEG 76
Figura 24 – Arquitetura da rede *MultiLayer Perceptron* com 2 camadas escondidas . . 78
Figura 25 – Placa Cyton e módulo Daisy conectados e módulo USB 81
Figura 26 – Conexão dos eletrodos na Placa Cyton e módulo Daisy 82
Figura 27 – Ultra Cortex IV com Placa Cyton e módulo Daisy conectados 82
Figura 28 – Posicionamento dos eletrodos de EEG de acordo com a plataforma OpenBCI para 16 eletrodos 83
Figura 29 – Gráfico de dispersão das características acústicas nos fatores da Análise Fatorial 87
Figura 30 – Forma de onda da música “Dança Húngara No.5” de Johannes Brahms com os *triggers* encontrados 88

Figura 31 – Forma de onda da música “O Barbeiro de Sevilha - Abertura” de Gioachino Rossini com os *triggers* encontrados 89
Figura 32 – Sinal médio da região central de músicos e não-músicos para a música: Dança Húngara No.5, de Johannes Brahms 90
Figura 33 – Sinal médio da região central de músicos e não-músicos para a música: O Barbeiro de Sevilha, de Gioachino Rossini 90
Figura 34 – Mapa topográfico cortical da música Dança Húngara No.5, de Johannes Brahms, para a componente N100 92
Figura 35 – Mapa topográfico cortical da Dança Húngara No. 5, de Johannes Brahms, para a componente P200 92
Figura 36 – Mapa topográfico cortical da música O Barbeiro de Sevilha - Abertura, de Gioachino Rossini, para a componente N100 93
Figura 37 – Mapa topográfico cortical da música O Barbeiro de Sevilha - Abertura, de Gioachino Rossini, para a componente P200 93
Figura 38 – Taxa de acerto na classificação do Cenário 1 para o classificador k-NN . . 97
Figura 39 – Taxa de acerto na classificação do Cenário 1 para a rede neural MLP . . . 98
Figura 40 – Gráfico de Barras, apresentando a classificação dos voluntários para o Cenário 1 99
Figura 41 – Taxa de acerto na classificação do Cenário 2 para o classificador k-NN . . 101
Figura 42 – Taxa de acerto na classificação do Cenário 2 para a rede neural MLP . . . 102
Figura 43 – Gráfico de Barras, apresentando a classificação dos voluntários para o Cenário 2 103

LISTA DE TABELAS

Tabela 1 – Intervalo de Frequências das Sub-Bandas do *Spectral Flux* 70
Tabela 2 – Parâmetros do filtro Butterworth 75
Tabela 3 – Relação de Voluntários 84
Tabela 4 – Componentes principais, autovalores, porcentagem da variância explicada e proporção acumulada pelos componentes 86
Tabela 5 – Carregamento das variáveis nos fatores 86
Tabela 6 – Apresentação dos parâmetros utilizados para determinação dos *triggers* de cada característica e o número de *triggers* encontrados 88
Tabela 7 – Resultado da classificação dos voluntários para o Cenário 1 para os classificadores k-NN e a Rede Neural MLP 99
Tabela 8 – Latência média dos grupos de voluntários para cada característica acústica apresentada. 100
Tabela 9 – Resultado da classificação dos voluntários para o Cenário 2 para os classificadores k-NN e a Rede Neural MLP. 102

LISTA DE ABREVIATURAS

kNN	k Nearest Neighbors.
CP	Componente Principal.
EEG	Eletroencefalograma.
ERP	Event Related Potential.
FA	Factor Analysis.
fMRI	Functional Magnetic Resonance Imaging.
ISI	Inter-Stimulus Interval.
MEG	Magnetoencefalograma.
MIR	Music Information Retrieval.
MLP	MultiLayer Perceptron.
MoRI	Magnitude of the Rapid Increase.
MRI	Magnetic Resonance Imaging.
NGM	Núcleo Geniculado Medial.
PA	Potencial de Ação.
PCA	Principal Component Analysis.
PEPS	Potencial Excitatório Pós-Sináptico.
PET	Positron Emission Tomography.
PIPS	Potencial Inibitório Pós-Sináptico.
PLFP	Preceding Low-Feature Phase.
REM	Rapid Eye Movement.
RMS	Root Mean Square.
SNC	Sistema Nervoso Central.
SNP	Sistema Nervoso Periférico.
ZCR	Zero Crossing Rate.

LISTA DE SÍMBOLOS

$[ion]_i$	Concentração iônica intracelular.
$[ion]_o$	Concentração iônica extracelular.
Ca^{2+}	Íon de Cálcio.
d_j	Distância entre e e e_{j}.
e	Coordenada de um ponto arbitrário no escalpo.
e_j	Coordenada de um eletrodo no escalpo.
E_{ion}	Potencial de equilíbrio.
F	Constante de Faraday.
H	Número de neurônios na camada escondida.
h	Sobreposição da janela.
k	Número de vizinhos mais próximos.
K^+	Íon de Potássio.
m	Ordem da interpolação.
μ	Média.
N	Comprimento da janela.
Na^+	Íon de Sódio.
R	Constante dos gases.
σ	Desvio padrão.
T	Temperatura absoluta.
V_{p+}	Limiar superior.
V_{p-}	Limiar inferior.
Vm	Resultado da interpolação.
X	X-fold cross-validation.
$X(k)$	FFT do sinal de áudio.
$x(n)$	Sinal de áudio.
Z	Matriz de dados normalizados.
z	Carga do íon.
z_j	Valor do potencial elétrico no eletrodo e_{j}.

SUMÁRIO

1 INTRODUÇÃO . . . 27
1.1 TRABALHOS RELACIONADOS . . . 28
1.2 MOTIVAÇÃO . . . 30
1.3 OBJETIVO . . . 30
1.4 ORGANIZAÇÃO DO TRABALHO . . . 31
2 CONCEITOS FUNDAMENTAIS . . . 33
2.1 A NATUREZA DO SINAL ELÉTRICO NO CÉREBRO . . . 33
2.1.1 O Sistema Nervoso . . . 33
2.1.1.1 O Sistema Nervoso Central (SNC) . . . 34
2.1.1.2 O Sistema Nervoso Periférico (SNP) . . . 36
2.1.2 O Neurônio . . . 36
2.1.2.1 O Potencial de Ação . . . 38
2.1.2.2 A Transmissão Sináptica . . . 41
2.1.3 Considerações Complementares . . . 42
2.2 ELETROENCEFALOGRAMA . . . 42
2.2.1 Geração Do Sinal De EEG . . . 43
2.2.2 Fisiologia Do EEG . . . 45
2.2.2.1 Mapas Topográficos Corticais . . . 46
2.2.2.2 Características dos ritmos em EEG . . . 48
2.2.2.3 Artefatos em EEG . . . 49
2.2.3 Potencial Relacionado ao Evento . . . 49
2.2.4 Considerações Complementares . . . 54
2.3 O SISTEMA AUDITIVO . . . 54
2.3.1 O Som . . . 54
2.3.2 Estrutura Do Sistema Auditivo . . . 55
2.3.2.1 A Cóclea e o Processamento Auditivo . . . 56
2.3.3 Processamento Musical . . . 60
2.3.4 Considerações Complementares . . . 62
3 METODOLOGIA . . . 63
3.1 EXTRAÇÃO DE CARACTERÍSTICAS ACÚSTICAS . . . 64
3.1.1 Root Mean Square Energy . . . 66
3.1.2 Zero Crossing Rate . . . 66
3.1.3 Brightness . . . 67
3.1.4 Spectral Centroid . . . 67
3.1.5 Spectral Spread . . . 67
3.1.6 Spectral Flatness . . . 68
3.1.7 Spectral Skewness . . . 68

3.1.8 Spectral Kurtosis 69
3.1.9 Spectral Rolloff 69
3.1.10 Spectral Flux 69
3.1.11 Spectral Roughness 71
3.1.12 Spectral Entropy 71
3.2 SELEÇÃO DE CARACTERÍSTICAS ACÚSTICAS 72
3.3 SELEÇÃO DE TRIGGERS 73
3.4 PROCESSAMENTO DO SINAL DE EEG 74
3.5 ANÁLISE DOS SINAIS DE EEG 77
3.6 CLASSIFICAÇÃO 77
4 EXPERIMENTOS E RESULTADOS 81
4.1 AQUISIÇÃO DOS SINAIS DE EEG 81
4.2 PARTICIPANTES 83
4.3 ESTÍMULOS 85
4.4 PROCEDIMENTOS 85
4.5 EXTRAÇÃO E SELEÇÃO DAS CARACTERÍSTICAS ACÚSTICAS 85
4.6 PROCESSAMENTO E ANÁLISE DO SINAL DE EEG 89
4.7 CLASSIFICAÇÃO 95
4.7.1 Classificação de acordo com Cenário 1 96
4.7.2 Classificação de acordo com Cenário 2 99
5 CONCLUSÃO 105
REFERÊNCIAS 107
ÍNDICE 112
APÊNDICE A – Termo de Consentimento Livre Esclarecido 113

1 INTRODUÇÃO

O ser humano está em constante desenvolvimento e os processos de aprendizado passados por ele produzem mudanças físicas no cérebro, chamada de plasticidade, onde áreas utilizadas mais frequentemente são expandidas, enquanto que áreas raramente utilizadas são reduzidas (A. BENNET; D. BENNET, 2008; ZULL, 2004).

Durante o processo evolucionário, o sistema sensorial do ser humano foi modificado pelo meio ambiente de tal forma que o córtex auditivo passou a processar, além dos sinais do meio ambiente, sinais acústicos complexos como a fala e a música (ANGULO-PERKINS et al., 2014). A música tem um papel importante, tanto no processo evolucionário quanto no crescimento do ser humano, que é aumentar a interconexão entre os dois hemisférios cerebrais (A. BENNET; D. BENNET, 2008).

Tocar um instrumento musical é uma das habilidades mais especializadas que o ser humano é capaz de adquirir, devido a grande demanda neural que ela requer (VUSST et al., 2014). Para Muszkat, Correia e Campos (2000) as funções musicais são “complexas, múltiplas e localizadas de forma assimétrica no cérebro", onde o hemisfério direito está relacionado a altura, ao timbre e a discriminação melódica, e o hemisfério esquerdo está relacionado ao ritmo, a identificação de melodias e ao processamento temporal e sequencial de sons.

Desta forma, o estudo do processamento musical pelo cérebro traz informações importantes sobre os aspectos da música, revelando como o cérebro processa suas características como harmonia, ritmo, timbre, entre outros, permitindo também compreender como o som é transformado em uma percepção musical.

Segundo Poikonen et al. (2016), as pesquisas de música no campo da neurociência, utilizando o Eletroencefalograma (EEG) e o Magnetoencefalograma (MEG), têm se focado em entender o processamento neural de sons artificiais, ou seja, estas pesquisas têm se apoiado em experimentos controlados onde as características musicais são apresentadas de forma isolada e artificialmente manipuladas comparando, por exemplo, tons puros *versus* tons complexos, acordes consonantes *versus* acordes dissonantes, entre outros (BESSON; FAÏTA, 1995; KOELSCH et al., 2000; GEORGE; DONNA, 2011; VIRTALA et al., 2014; VUSST et al., 2014; SHENG-FU et al., 2011).

Por meio destes experimentos, têm sido possível obter informações importantes sobre o processamento de elementos individuais da música pelo cérebro, porém, é necessário analisar também os sons naturais que são caracterizados pela sua espontaneidade, impureza, interação e fluxo contínuo de sobreposição de notas, permitindo a extração de mais informações sobre como o cérebro processa a música como um todo (POIKONEN et al., 2016; ALLURI et al., 2012).

Além disso, evidências sugerem que o ser humano possui habilidades musicais implícitas (KOELSCH et al., 2000), ou seja, apesar dos não-músicos não possuírem um aprendizado formal em música, eles são musicais e capazes de extrapolarem sons que poderão ocor-

rer ao longo de uma música, de acordo com o contexto da teoria musical, mesmo que não-intencionalmente. Porém, apesar dos não-músicos serem musicais e apresentarem um certo nível de processamento musical, a musicalidade, ou a prática instrumental, está associada a diferenças estruturais e funcionais entre o cérebro de músicos e não-músicos (VIRTALA et al., 2014).

Sendo assim, a hipótese levantada neste trabalho sugere que é possível diferenciar um músico de um não-músico por meio das rápidas mudanças que ocorrem nas características acústicas extraídas de músicas apresentadas como estímulo no experimento proposto, utilizando-se de métodos multivariados estatísticos e conexionistas para o processamento dos sinais de áudio e de EEG.

1.1 TRABALHOS RELACIONADOS

Neste contexto, diversas pesquisas têm sido feitas sobre este tema, analisando as diferenças estruturais existentes no cérebro de músicos e não-músicos, buscando compreender como o cérebro processa a música.

Em um estudo com dois grupos, um possuindo treinamento musical e um grupo de controle, Sarnthein et al. (1997) observou que no primeiro grupo, conforme havia o aumento do grau de treinamento musical, era exibido maior coerência no sinal de EEG entre os dois hemisférios cerebrais.

Segundo George e Donna (2011), a prática musical é capaz de melhorar a memória de trabalho[1]. Seus resultados foram obtidos através de experimentos comportamentais e neurais, por meio do potencial evocado P300 em músicos e não-músicos, onde o componente P300 ocorreu de forma antecipada em músicos, indicando que, a longo prazo, a prática musical gera melhorias na memória de trabalho do grupo de músicos em relação aos não-músicos.

Com o intuito de classificar músicos e não-músicos por meio de sinais de EEG, Sheng-Fu et al. (2011) desenvolveram um experimento utilizando um grupo de não-músicos e de músicos profissionais, para os quais foram apresentados alguns estímulos artificias, havendo a necessidade do voluntário julgar se o estímulo apresentado era consonante ou dissonante. Sheng-Fu et al. (2011) utilizaram uma rede neural para classificar os dados obtidos, tendo como sinal de entrada o valor médio do sinal de EEG amostrado de uma época de 0ms a 350ms pós-estímulo do eletrodo Cz, a uma frequência de amostragem de 250Hz e utilizando a Análise de Componentes Principais (PCA) para redução de dimensionalidade, obtendo como resultado uma taxa de acerto na classificação superior a 90%.

Alluri et al. (2012) desenvolveram um novo paradigma para obtenção de imagens de *Functional Magnetic Resonance Imaging* (fMRI) enquanto os participantes ouviam o som de músicas (o tango "Adios Noninõ", de Astor Piazzola), encontrando resultados que corroboram

[1] Também chamada de Memória de Curto Prazo, que permite temporariamente o armazenamento e a manipulação de informações enquanto uma determinada tarefa está sendo realizada (MOURÃO; MELO, 2011).

com experimentos de Potencial Relacionado ao Evento (ERP), ou Potencial Evocado, tradicionais realizados anteriormente que se utilizavam de estímulos artificiais. De acordo com seu estudo, durante a audição de músicas ocorre o recrutamento de áreas do sistema límbico e de recompensa no processamento da pulsação, de regiões relacionadas à emoção no processamento da tonalidade e do cerebelo envolvendo o processamento do timbre.

De forma similar, Alluri et al. (2013) utilizaram o mesmo paradigma proposto anteriormente para predizer a evolução temporal das repostas cerebrais para músicas instrumentais e músicas cantadas. Os dados foram obtidos por fMRI através da apresentação de estímulos naturais para predizer qual atividade cerebral é gerada em resposta ao processamento de uma determinada característica acústica apresentada ao voluntário.

Utilizando fragmentos de músicas e estímulos pseudo-artificiais, Abrams et al. (2013) utilizaram o *functional Magnetic Ressonance Imaging* (fMRI) para examinar a sincronização das respostas cerebrais entre ouvintes não-músicos. Para realizar a sua análise, eles utilizaram o *Inter-Subject Synchronization* (ISS), cujas estruturas cerebrais que são sincronizadas entre sujeitos, durante um estímulo extenso, constituem em um conjunto de regiões cerebrais responsáveis pelo processamento dos elementos estruturais contidos naquele estímulo durante o tempo. Como resultado, Abrams et al. (2013) mostraram que houve uma maior sincronização ao ouvir músicas em relação aos estímulos pseudo-artificiais. As respostas cerebrais foram encontradas de maneira bilateral no mesencéfalo, no tálamo, no A1 e em áreas associativas, e de maneira lateralizada no hemisfério direito em estruturas no córtex frontal e parietal e em regiões de planejamento motor, associadas a um alto nível de função cognitiva. De acordo com seu estudo, o conteúdo estrutural de sequências sonoras é suficiente para alterar dramaticamente a sincronização dentro da extensa rede neural.

A partir de um experimento comportamental e neural expondo músicos e não-músicos a um conjunto de estímulos contendo acordes maiores, menores e invertidos, Virtala et al. (2014) encontraram que o grupo de músicos apresentaram maior amplitude no componente N1 de ERP do que os não-músicos, e que apenas os músicos apresentaram o *Mismatch Negativity* (MMN) para acordes menores, apesar dos dois grupos serem capazes de discriminar os acordes maiores e menores. Isso indica que possivelmente exista a necessidade de um certo nível de conhecimento musical para que estes potenciais evocados ocorram.

Vusst et al. (2014), propuseram diferenciar não-músicos, que não possuíam nenhum conhecimento musical, e músicos profissionais, que possuíam estilos musicais diferentes (clássico, jazz e rock), a partir de um experimento tradicional de ERP, o MMN, apresentando estímulos com algum desvio (altura, timbre, intensidade, ritmo, etc.). Como resultado, foi encontrado que existe uma diferença no processamento musical dependente do estilo musical de cada voluntário. Ou seja, existem diferenças entre não-músicos e músicos, mas também existem diferenças entre músicos dependentes do instrumento e do estilo musical tocado.

Utilizando o método proposto por Alluri et al. (2012), porém com o EEG, Poikonen et al. (2016) propôs revelar as respostas cerebrais de não-músicos induzidas por músicas, a

partir das rápidas alterações das características musicais que são caracterizadas por regiões de alto-contraste musical, precedidas por regiões de baixo-contraste, denominadas *Preceding Low-Feature Phase* (PLFP). Foram investigadas as respostas do ERP para três músicas de gêneros distintos e, como resultado, Poikonen et al. (2016) encontraram que os componentes N1 e P2 de ERP são extraídos não apenas por sons simples com tempo precisamente definido de silêncio denominado *Interstimulus Interval* (ISI), mas também por uma dinâmica de estímulos naturais contínuos como peças musicais. Porém, estes estímulos precisam possuir fortes contrastes acústicos para serem capazes de extrair os componentes de ERP convencionais.

De fato, a música modifica estruturalmente o cérebro através da plasticidade cerebral. Essas diferenças podem ser atribuídas ao extenso treinamento e não necessariamente por diferenças inatas entre músicos e não-músicos (VIRTALA et al., 2014).

1.2 MOTIVAÇÃO

A música está em todos os lugares e saber como o cérebro processa essa informação é uma questão que vem sendo estudada cada vez mais nos últimos tempos. Sabe-se que, devido ao processo de plasticidade cerebral, a forma de processar a música por um músico é diferente de um não-músico. Tendo em vista esta diferença entre os dois grupos, a motivação deste trabalho é a possibilidade de classificar músicos e não-músicos utilizando-se de músicas como estímulo em vez de estímulos artificiais.

1.3 OBJETIVO

O principal objetivo deste trabalho é desenvolver um arcabouço computacional capaz de classificar, por meio de aprendizado supervisionado, os sinais de EEG registrados de um grupo de voluntários em músicos e não-músicos. Mais especificamente, pretende-se:

a) Estender a metologia proposta por Poikonen et al. (2016) para verificar a possibilidade de predição de musicalidade por meio de sinais de EEG;

b) Realizar experimentos de mapeamento cerebral, aproximando-os de um ambiente natural, como se o voluntário estivesse ouvindo músicas normalmente em sua rotina diária;

c) Comparar dois cenários de classificação, um utilizando o tempo como fator determinístico para formação do vetor de entrada e outro utilizando o resultado do potencial evocado para a formação do vetor de entrada;

d) Comparar os resultados apresentados por dois classificadores, um não-paramétrico que se utiliza da distância entre os seus vizinhos para a classificação, e outro que busca uma relação não-linear do conjunto de dados apresentado.

1.4 ORGANIZAÇÃO DO TRABALHO

Este trabalho está organizado da seguinte forma: No Capítulo 2 apresenta-se uma revisão conceitual e bibliográfica sobre o tema apresentado, descrevendo como o sinal elétrico é gerado no cérebro, como ele é medido pelo EEG e descrevendo o funcionamento do sistema auditivo. No Capítulo 3 descreve-se a metodologia proposta para o desenvolvimento do trabalho, apresentando detalhadamente todas as etapas a serem executadas. No Capítulo 4 são apresentados os experimentos e resultados encontrados neste trabalho e, por fim, no Capítulo 5 apresenta-se a conclusão da presente dissertação e são apresentadas as possibilidades de trabalhos futuros.

2 CONCEITOS FUNDAMENTAIS

Este capítulo apresenta os conceitos fundamentais para a compreensão dos métodos aplicados neste trabalho, iniciando a primeira seção com a descrição da natureza do sinal elétrico no cérebro, destacando as suas principais fontes geradoras de potenciais elétricos, os Potenciais de Ação (PA) e as Sinapses, que são registradas pelo Eletroencefalograma (EEG), que é descrito na segunda seção deste capítulo, onde são apresentadas as suas principais características e as ferramentas para a análise do sinal registrado, através dos Mapas Topográficos e dos Potenciais Relacionados ao Evento (ERP).

A seguir, na terceira seção, o Sistema Auditivo é apresentado, descrevendo como o som percorre pelas vias auditivas até atingir o Córtex Auditivo Primário (A1), e então esta informação é distribuída para as áreas associativas do córtex, gerando a percepção auditiva e, por fim, apresentando como o cérebro processa a música.

2.1 A NATUREZA DO SINAL ELÉTRICO NO CÉREBRO

Nesta seção é apresentada a natureza do sinal elétrico no cérebro, mostrando de forma simplificada como o sistema nervoso é constituído e a estrutura dos neurônios, que são células altamente especializadas e essenciais para o sistema nervoso, capazes de propagar sinais transmitindo informações por meio de sequências de pulsos elétricos chamados Potenciais de Ação (PA), e capazes de se comunicarem entre si por meio da transmissão sináptica.

2.1.1 O Sistema Nervoso

O sistema nervoso é fundamental para a vida e é ele quem permite ao ser humano sentir, pensar e se mover. Algumas evidências sugerem que nossos ancestrais pré-históricos já entendiam que o cérebro era essencial para a vida (BEAR; CONNORS; PARADISO, 2006) e, com o passar do tempo, na tentativa de compreender o funcionamento do cérebro, surgiram várias ideias sobre a sua função e estrutura.

Aristóteles (384-332 a.C.) acreditava que o coração era o centro do intelecto e que o cérebro funcionava como um radiador, servindo para resfriar o sangue superaquecido do coração. Segundo ele, o temperamento dos humanos estava relacionado com a sua capacidade de resfriamento do cérebro. Hipócrates (469-379 a.C.) acreditava que o cérebro não estava apenas envolvido com as sensações, mas era também a sede da inteligência (BEAR; CONNORS; PARADISO, 2006).

Quando, em 1859, o biólogo inglês Charles Darwin publicou “A origem das espécies” falando sobre a seleção natural, ele incluiu o comportamento dentre os traços que eram herdados durante a evolução das espécies, mesmo que ele seja um reflexo da atividade do sistema

nervoso (BEAR; CONNORS; PARADISO, 2006). Sendo assim, os mecanismos que formam um determinado comportamento deveriam ser similares nas espécies.

Com a melhoria dos microscópios no início do século XIX, os cientistas puderam examinar tecidos animais em escalas maiores. Foi em 1839 que o zoologista alemão Theodor Schwann propôs a "teoria celular", afirmando que todos os tecidos são compostos por unidades microscópicas chamadas células (BEAR; CONNORS; PARADISO, 2006).

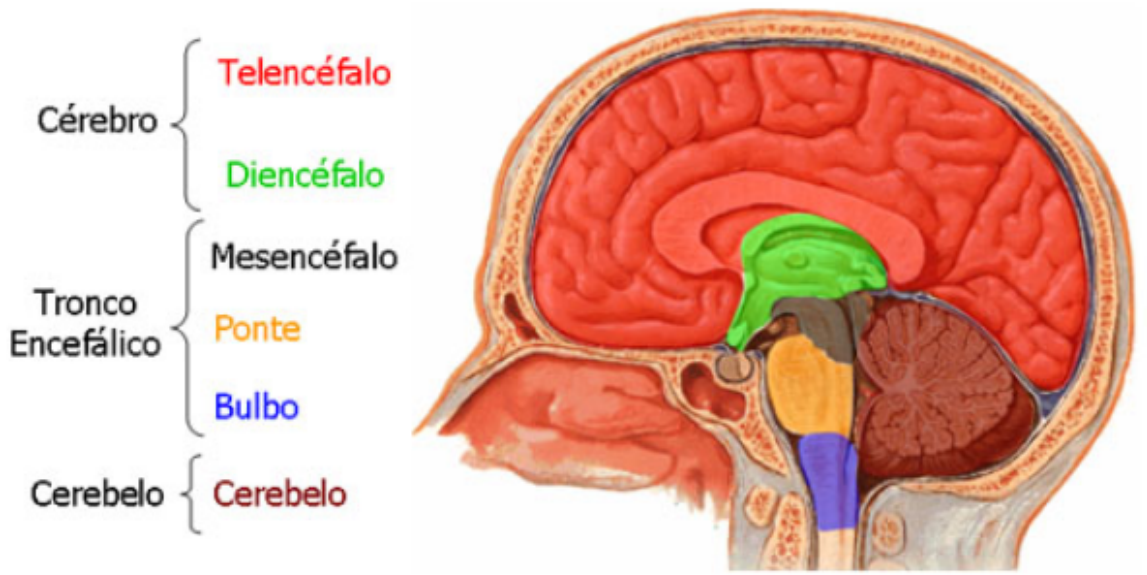
O que permite ao ser humano interagir com o ambiente é o sistema nervoso. Ele pode ser dividido em Sistema Nervoso Central (SNC), composto pelo encéfalo e medula espinhal, e o Sistema Nervoso Periférico (SNP), composto por todos os nervos e seus componentes fora do SNC (KANDEL et al., 2013). É no sistema nervoso central que está a grande maioria dos neurônios nos animais.

2.1.1.1 O Sistema Nervoso Central (SNC)

O tecido do SNC é muito delicado e se encontra protegido por três membranas, chamadas de meninges, que são estruturas importantes para a defesa do SNC, sendo elas: a dura-máter, a membrana aracnoide e a pia-máter. Dentre elas, a dura-máter é a mais externa e a pia-máter a mais interna.

O sistema nervoso central pode ser dividido em medula espinhal e encéfalo, onde a Figura 1 apresenta a composição do encéfalo (WECKER; SOARES; NEMOS, 2001).

Figura 1 – Composição do Encéfalo



Fonte: Wecker, Soares e Nemos, 2016.

A Medula Espinhal está envolvida pelos ossos da coluna vertebral e está em continuidade com o tronco encefálico. Ela é o maior condutor de informações, recebendo e processando

informações sensoriais da pele, juntas e músculos e controlando o movimento, sendo subdividida em cervical, torácica, lombar e sacral (KANDEL et al., 2013).

A medula espinhal se comunica com o corpo por intermédio dos nervos espinhais que emergem dela por meio de espaços existentes entre cada vértebra da coluna vertebral (BEAR; CONNORS; PARADISO, 2006).

O Encéfalo pode ser dividido em: hemisférios cerebrais, diencéfalo, cerebelo e tronco encefálico.

O Tronco Encefálico é formado pelo bulbo, a ponte e o mesencéfalo. Ele recebe informações sensoriais da pele e dos músculos da cabeça promovendo o controle motor de sua musculatura. Ele transmite informações da medula espinhal ao cérebro e no sentido contrário, regulando também os níveis de excitação e sensibilização (KANDEL et al., 2013).

O tronco encefálico é considerado a porção mais primitiva do encéfalo dos mamíferos, e também regula funções vitais como a respiração, o estado de alerta consciente e a temperatura corporal (BEAR; CONNORS; PARADISO, 2006).

O Bulbo inclui alguns centros responsáveis por funções autônomas vitais como a digestão, respiração e o controle da frequência cardíaca. Nele também estão o centro vasomotor e o centro do vômito. A Ponte transmite informações sobre movimento dos hemisférios do cérebro para o cerebelo e tem papel fundamental na regulação do padrão e ritmo respiratório. O Mesencéfalo controla várias funções motoras e sensoriais, incluindo o movimento dos olhos e a coordenação dos reflexos auditivos e visuais (KANDEL et al., 2013).

O Cerebelo é conectado ao tronco encefálico por uma série de fibras. Ele é um centro para o controle do movimento, modulando sua força e o alcance, equilíbrio e coordenação, e está envolvido com o aprendizado de habilidades motoras (KANDEL et al., 2013; BEAR; CONNORS; PARADISO, 2006).

O Diencéfalo contém duas estruturas: O Tálamo, que processa maior parte da informação que chega ao córtex cerebral do restante do SNC, além da motricidade e do comportamento emocional, e o Hipotálamo, que regula as funções endócrinas e viscerais (KANDEL et al., 2013).

Os Hemisférios Cerebrais possuem: o Telencéfalo, os Núcleos da Base, a Amígdala, as Formações Hipocampais e o Corpo Caloso.

O Telencéfalo é dividido em dois hemisférios (direito e esquerdo) separados pela fissura longitudinal, interligados pelo Corpo Caloso, unindo áreas simétricas do córtex de cada hemisfério. O córtex cerebral é dividido em quatro áreas: frontal, parietal, occipital e temporal (KANDEL et al., 2013).

O tecido cerebral é dividido em duas partes: a substância cinzenta e a substância branca. Na substância cinzenta existem altas concentrações de corpos celulares de neurônios. Ela possui uma coloração rosada, porém após ser retirado de um animal, perde a sua cor, tornando-se acinzentado. A substância branca contém quase exclusivamente axônios de neurônios. Ór-

gãos frescos possuem a cor esbranquiçada devido à grande quantidade de mielina presente nela (BEAR; CONNORS; PARADISO, 2006).

De acordo com Kandel et al. (2013), "o neocórtex – região do córtex mais perto da superfície cerebral – é organizado em camadas e colunas". Cada camada possui diferente aferências e eferências, que se referem ao fluxo de informação, onde informações aferentes têm origem na periferia e se dirigem ao SNC e informações eferentes têm origem no SNC e se dirigem à periferia. A maior parte do neocórtex possui seis camadas que são numeradas da parte mais externa do córtex até a substância branca. A espessura de cada camada e a sua organização funcional varia ao longo do córtex.

Existem 12 pares de nervos cranianos, além dos que se originam na medula espinhal e inervam no corpo, que são originários no tronco encefálico e que inervam, principalmente, a cabeça. Cada um deles possui um nome e um número associado. Muitos dos nervos cranianos possuem um complexo conjunto de axônios que realizam diversas funções (BEAR; CONNORS; PARADISO, 2006).

2.1.1.2 O Sistema Nervoso Periférico (SNP)

Toda a parte do sistema nervoso que não seja o encéfalo e a medula espinhal é chamada de sistema nervoso periférico, podendo ser dividido em duas partes, o SNP Somático e o SNP Visceral.

No SNP Somático, todos os nervos que inervam a pele, articulações e músculos estão sobre controle voluntário. O soma, ou corpo do neurônio, é localizado dentro do SNC, mas seus axônios estão predominantemente no SNP. Os axônios sensoriais coletam informações deles e entram na medula espinhal pelas raízes dorsais (BEAR; CONNORS; PARADISO, 2006).

No SNP Visceral, os neurônios inervam os órgãos internos, vasos sanguíneos e glândulas. Ele é controlado por nervos da medula, tronco cerebral e hipotálamo e opera de forma autônoma. Os axônios sensoriais viscerais carregam ao SNC informações sobre as funções viscerais, como pressão e conteúdo de oxigênio no sangue (BEAR; CONNORS; PARADISO, 2006).

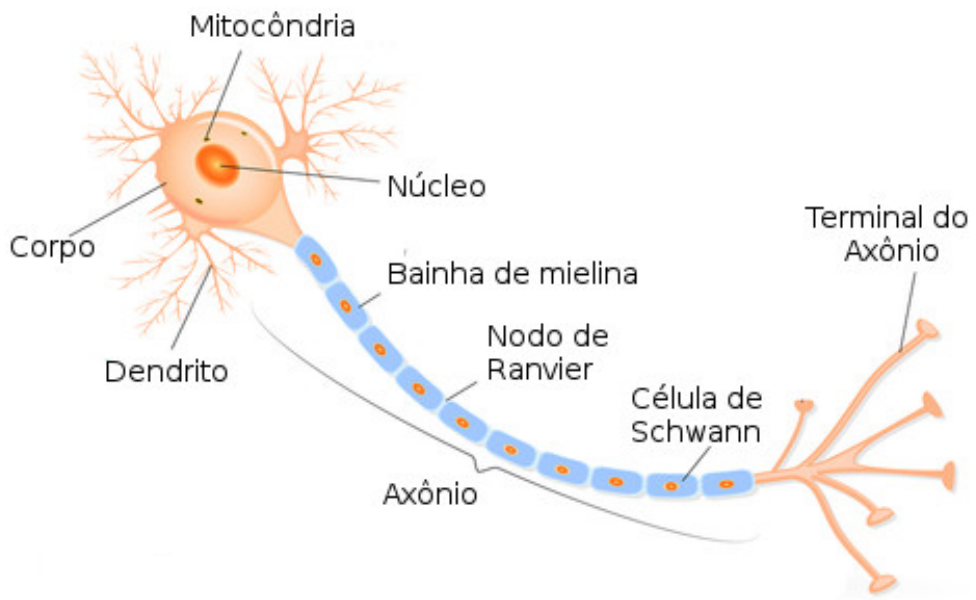
2.1.2 O Neurônio

O cérebro humano é uma rede com mais de 100 bilhões de neurônios interconectados em sistemas que constroem a nossa percepção do mundo externo (KANDEL et al., 2013). Os neurônios são elementos essenciais para o sistema nervoso e são um tipo de célula altamente especializada (SQUIRE et al., 2013). Todos os processos neurológicos são dependentes de uma complexa interação entre células e neurônios.

Um neurônio típico possui 4 regiões morfologicamente definidas, cada uma delas possuindo um papel distinto na geração de sinais e na comunicação com outras células nervosas.

Estas regiões são: (1) Corpo Celular, ou Soma, (2) Dendritos, (3) Axônios e (4) Terminais Pré-sinápticos, ou terminal do axônio (OLIVEIRA, 2006). A Figura 2 mostra uma estrutura de um neurônio biológico de acordo com a sua composição.

Figura 2 – Estrutura de um neurônio biológico



Fonte: Oliveira, 2017.

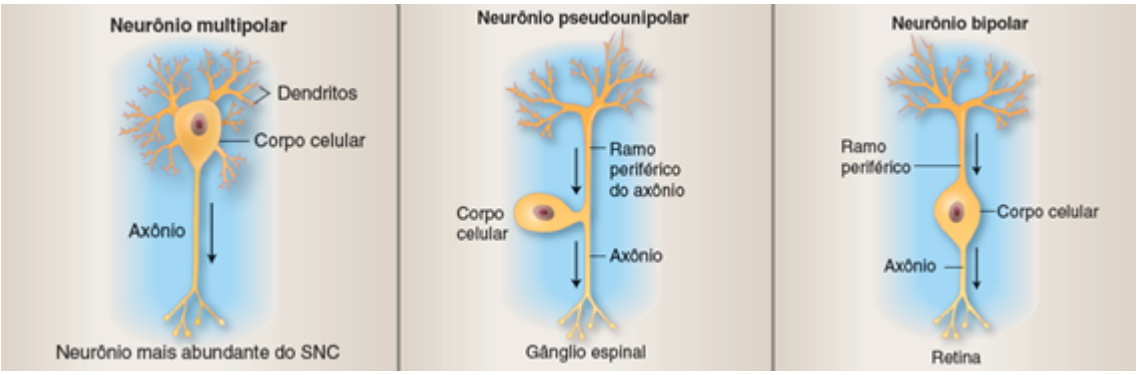
O Corpo Celular, ou Soma, é o centro metabólico do neurônio, onde estão concentradas as principais organelas intracelulares. Um neurônio geralmente possui um axônio e um número variado de dendritos. Os dendritos recebem sinais vindos de outras células nervosas, enquanto que o axônio transmite sinais a outros neurônios. Um axônio pode transmitir sinais elétricos a distâncias de 0,1 mm a 2 m (KANDEL et al., 2013).

Existem diversos tipos de neurônios no SNC. Eles podem ser classificados de acordo com o seu tamanho, morfologia ou conforme os neurotransmissores que utilizam. A Figura 3 mostra uma classificação básica dos neurônios conforme a sua morfologia.

O conteúdo interno do neurônio é separado do meio externo por uma membrana com bicamada fosfolipídica, que não o isola completamente do meio extracelular, mas a sua permeabilidade seletiva permite a troca de íons. Esta é a propriedade que torna possível a geração e propagação de sinais bioelétricos (LENT, 2001).

Algumas proteínas associadas à membrana bombeiam substâncias de dentro para fora da célula enquanto que outras formam poros que regulam quais substâncias podem acessar o interior do neurônio. De acordo com Bear, Connors e Paradiso (2006), dependendo da região da célula, seja ela o soma, o dendrito ou o axônio, a composição proteica da membrana varia e esta é uma importante característica dos neurônios. Portanto, as proteínas fornecem caminhos para os íons cruzarem a membrana neuronal.

Figura 3 – Tipos de Neurônios do SNC classificados de acordo com sua Morfologia



Fonte: Adaptado de Krebs, Weiberg e Akesson, 2012, p.5.

As diferentes proteínas possuem formas, tamanhos e características químicas diferentes para desempenhar as suas funções nos neurônios. Elas são moléculas formadas a partir de inúmeras combinações de aminoácidos diferentes (BEAR; CONNORS; PARADISO, 2006).

Os canais iônicos são formados por moléculas proteicas que se estendem através da membrana. Uma propriedade importante dos canais iônicos é a existência de "portões" que podem ser abertos/fechados por alterações locais na membrana. Desta forma, canais de potássio são seletivamente permeáveis ao íon de potássio (K^+), canais de sódio ao íon de sódio (Na^+) e assim por diante, permitindo que haja um caminho entre o meio intra e extracelular.

2.1.2.1 O Potencial de Ação

Os neurônios são notáveis dentre outras células do corpo pela sua habilidade de propagar sinais rapidamente e através de longas distâncias (DAYAN; ABBOTT, 2005). A informação pode fluir em duas direções: as aferentes, da periferia para o SNC, e as eferentes, do SNC para a periferia. Os neurônios transmitem informação gerando pulsos elétricos característicos, chamado Potencial de Ação (PA), onde uma sequência de disparos formam vários padrões temporais.

Todas as células estão envolvidas por uma bicamada fosfolipídica, mantendo uma concentração de íons diferente no espaço intra e extracelular. Porém, os neurônios e músculos são as únicas células capazes de enviar sinais ao longo de sua superfície ou de explorar essas diferenças iônicas para gerar sinais elétricos (KREBS; WEIBERG; AKESSON, 2012).

O Potencial de Ação é inicializado em uma região próxima da origem do axônio, no cone de inserção, de onde ele propaga o sinal elétrico a uma velocidade que pode variar de 1 a 100m/s, dependendo do diâmetro axonal (KANDEL et al., 2013), sem falha na transmissão e sem distorção na sua velocidade.

A amplitude do sinal elétrico produzido no PA é constante e possui o valor aproximado de 100mV, por ser uma resposta do tipo "tudo ou nada" (KANDEL et al., 2013). Durante a sua trajetória pelo axônio, este sinal pode se regenerar em intervalos regulares.

Segundo Kandel et al. (2013), a informação transmitida pelo PA é determinada pelo caminho que o sinal elétrico atravessa através do cérebro e não pela forma de seu sinal elétrico.

A região intracelular possui uma carga elétrica negativa comparada com a extracelular. Essa diferença de cargas na membrana é chamada de Potencial de Repouso que, em um neurônio típico, é cerca de -65 mV (LENT, 2001). O movimento de íons através da membrana gera um gradiente elétrico para cada íon. A soma de todos esses gradientes é o Potencial de Membrana (Vm) (KREBS; WEIBERG; AKESSON, 2012).

A membrana pode alterar ativamente a sua permeabilidade para diferentes íons através dos canais iônicos, que são proteínas da membrana que permitem a passagem de íons, gerando um fluxo de corrente (KREBS; WEIBERG; AKESSON, 2012). Isso altera o potencial de membrana, fazendo com que o interior dela se torne positivamente carregado em relação ao exterior, gerando assim um Potencial de Ação (LENT, 2001).

Os Potenciais de Ação (PA) são, portanto, "impulsos elétricos ou alterações no potencial da membrana, que são dependentes de proteínas especiais que atravessam a bicamada lipídica" (KREBS; WEIBERG; AKESSON, 2012), p.10.

Cada íon possui o seu próprio potencial de equilíbrio. Desta forma, sabendo a carga elétrica de um íon e a diferença de concentração intra e extracelular, é possível calcular os valores do potencial de equilíbrio de cada íon utilizando a equação de Nernst (Equação (1)), sendo que o equilíbrio é o balanço entre duas influências: a difusão, que força o íon a favor do seu gradiente de concentração, e a eletricidade, que faz com que o íon seja atraído por cargas opostas e repelido por cargas iguais (BEAR; CONNORS; PARADISO, 2006), tal que:

$$E_{ion} = 2.303 \left(\frac{R \times T}{z \times F} \right) \times \log \left(\frac{[ion]_o}{[ion]_i} \right), \tag{1}$$

onde E_{ion} é o potencial de equilíbrio, R é a constante dos gases, T é a temperatura absoluta, z é a carga do íon, F é a constante de Faraday, log é o logaritmo na base 10, $[ion]_o$ é a concentração iônica extracelular e $[ion]_i$ é a concentração iônica intracelular.

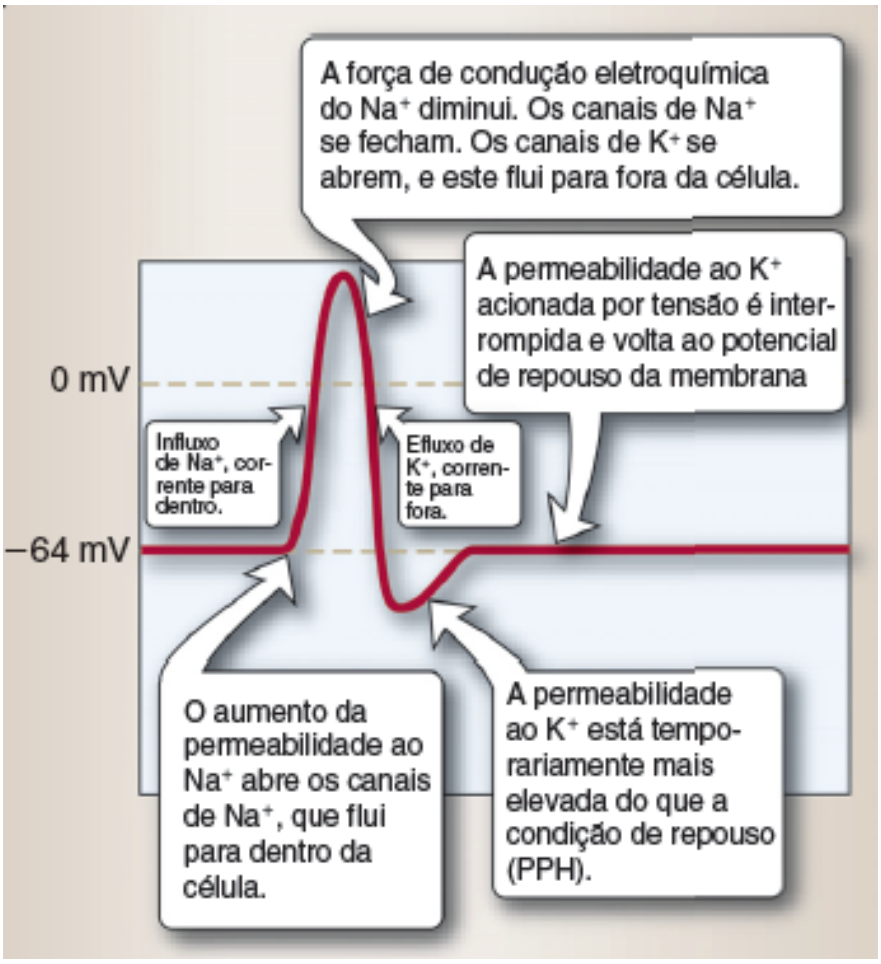
De acordo com Bear, Connors e Paradiso (2006), o K^+ está mais concentrado no meio intracelular que no meio extracelular, enquanto que o Na^+ e o Ca^{2+} estão mais concentrados no meio extracelular que no meio intracelular.

A entrada de sódio no meio intracelular despolariza a membrana, ou seja, a superfície interna da membrana se torna menos negativa, em resposta a mudanças no potencial da membrana (KANDEL et al., 2013). Quando esta despolarização atinge um determinado limiar, a membrana irá gerar um PA. Portanto, potenciais de ação são causados pela despolarização da membrana além do limiar (BEAR; CONNORS; PARADISO, 2006).

Durante o potencial de ação, ocorre uma redistribuição de cargas elétricas pela membrana. A repolarização é provocada pela saída de íons de potássio do meio intracelular, até que o interior fique negativamente carregado, levando o potencial de membrana de 0mV a -80mV, hiperpolarizando a membrana (BEAR; CONNORS; PARADISO, 2006).

Como o potencial de membrana se torna mais negativo em relação ao potencial de equilíbrio, os íons de Na^+ tornam a entrar no meio intracelular, atingindo então o potencial de equilíbrio de aproximadamente -65 mV, conforme ilustrado na Figura 4.

Figura 4 – O Potencial de Ação



Fonte: Adaptado de Krebs, Weiberg e Akesson, 2012, p.11.

Em princípio esta teoria é bastante simples, mas realizar esta medição na prática se tornou um desafio. Os experimentos utilizando o método de *voltage clamp* realizados pelos fisiologistas Alan Hodgkins e Andrew Huxley, em 1950, permitiu a eles calcularem as mudanças que ocorriam nas condutâncias em diferentes potenciais de membrana, medindo as correntes que fluíam através dela, fazendo com que eles ganhassem o Prêmio Nobel em 1952.

Eles desenvolveram o modelo, hoje conhecido como Modelo de Hodgkins-Huxley, que evidencia o papel dos canais iônicos dependentes da voltagem e a descrição da dinâmica de operação dos canais iônicos e da permeabilidade da membrana.

Sendo assim, Potenciais de Ação ocorrem apenas nos axônios e eles são sempre iguais no mesmo neurônio, mas, devido à variabilidade entre as células, eles são diferentes de um neurônio para outro, mudando a sua velocidade de transmissão, a quantidade de canais, etc.

2.1.2.2 A Transmissão Sináptica

Os neurônios se comunicam entre si por meio de sinapses, formando uma rede funcional para o processamento e armazenamento de informação. Essa transferência de informação é chamada de Transmissão Sináptica (KREBS; WEIBERG; AKESSON, 2012). Após ser gerado o potencial de ação no neurônio pré-sináptico, a informação é transmitida pelo axônio que irá enviá-la ao neurônio pós-sináptico.

A membrana pré e pós-sináptica são separadas por uma fenda chamada de Fenda Sináptica que possui uma largura de 20 a 40 nm (KANDEL et al., 2013).

Portanto, a sinapse apresenta dois lados que indicam a direção do fluxo de informação, o pré e o pós-sináptico. O lado pré-sináptico geralmente consiste de um terminal axônico, enquanto que o pós-sináptico pode ter um dendrito ou um soma de outro neurônio.

As sinapses podem ser do tipo elétrica, que possui uma resposta rápida e de curta duração, tendo uma comunicação citoplasmática com uma distância entre o neurônio pré e pós-sináptico de apenas 4nm, ou elas podem ser do tipo química.

A transmissão de sinapses químicas depende da difusão do neurotransmissor na fenda sináptica. Os locais de liberação de neurotransmissores são denominados Zonas Ativas. O terminal pré-sináptico contém de 100 a 200 vesículas sinápticas, cada uma delas contendo milhares de moléculas de neurotransmissores (KANDEL et al., 2013).

Durante o potencial de ação pré-sináptico, canais de Ca^{2+} dependentes de voltagem na zona ativa se abrem, permitindo que os íons de cálcio entrem no terminal pré-sináptico.

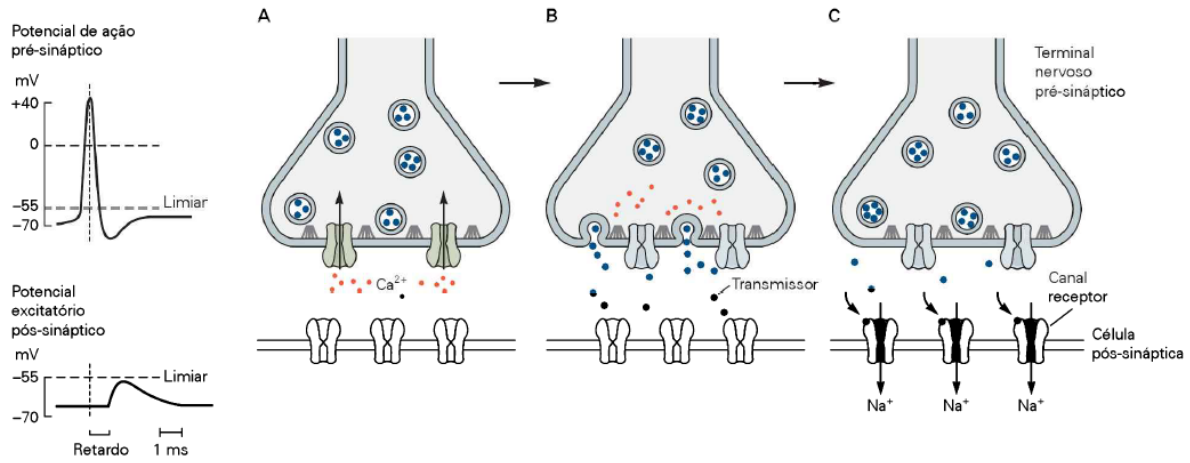
Com o aumento da concentração dos íons de cálcio no terminal pré-sináptico ocorre uma reação bioquímica que faz com que as vesículas sinápticas se unam à membrana pré-sináptica, liberando os neurotransmissores na fenda sináptica por um processo chamado Exocitose (KANDEL et al., 2013).

A membrana pós-sináptica possui receptores para os neurotransmissores. Quando os neurotransmissores chegam à fenda sináptica eles ativam os receptores no terminal pós-sináptico, causando a abertura/fechamento dos canais iônicos. O fluxo de íons altera a condutância da membrana e o potencial pós-sináptico da célula (KANDEL et al., 2013).

A Figura 5 descreve este processo, onde é possível verificar que (A) o PA no terminal pré-sináptico causa a abertura de canais de cálcio, permitindo a entrada dos íons de cálcio no terminal. O aumento da concentração de íons de cálcio (B) no terminal pré-sináptico faz com que as vesículas sinápticas se unam à membrana celular, liberando neurotransmissores na fenda sináptica. O neurotransmissor atravessa a fenda sináptica (C), causando a abertura dos canais iônicos no terminal pós-sináptico, alterando a condutância da membrana e do potencial pós-sináptico.

Apesar de a transmissão sináptica química não ser tão rápida quanto a elétrica, ela possui uma importante propriedade de amplificação, pois, uma vesícula sináptica libera milhares de

Figura 5 – Transmissão Sináptica



Fonte: Adaptado de Kandel et al., 2013, p.185.

neurotransmissores que podem abrir vários canais iônicos na célula-alvo (BEAR; CONNORS; PARADISO, 2006).

Quando a liberação de neurotransmissores causa uma despolarização na membrana pós-sináptica, ocorre então um Potencial Excitatório Pós-Sináptico (PEPS). A abertura de canais iônicos por acetilcolina e glutamato causam PEPS (BEAR; CONNORS; PARADISO, 2006).

Quando a liberação de neurotransmissores causa uma hiperpolarização na membrana pós-sináptica, ocorre então um Potencial Inibitório Pós-Sináptico (PIPS). A abertura de canais iônicos por glicina ou GABA causam PIPS (BEAR; CONNORS; PARADISO, 2006).

2.1.3 Considerações Complementares

Portanto, conforme foi descrito, a origem do sinal elétrico no cérebro ocorre devido às interações entre os neurônios que, individualmente, contribuem para a geração do potencial elétrico através dos potenciais de ação, que transmitem a informação pelo axônio até o terminal pré-sináptico, e da transmissão sináptica, onde ocorre a transferência de informação de um neurônio a outro.

O sinal gerado pelos neurônios no cérebro precisa passar por várias camadas de tecido não neural, sendo elas as meninges (pia-máter, aracnoide e dura-máter), os ossos do crânio e a pele, para alcançar os eletrodos responsáveis pelo registro do sinal elétrico que ficam posicionados no escalpo.

2.2 ELETROENCEFALOGRAMA

Nesta seção é apresentado o eletroencefalograma (EEG), mostrando como o sinal elétrico é gerado no cérebro e descrevendo o equipamento de EEG, o posicionamento dos eletro-

dos, os artefatos contidos no sinal e as características dos ritmos de EEG. Também é descrito o Potencial Relacionado ao Evento (ERP), que é uma técnica utilizada para o registo de potenciais elétricos evocados a partir de um determinado estímulo, e os Mapas Topográficos, utilizados para a visualização da atividade elétrica.

2.2.1 Geração Do Sinal De EEG

O Eletroencefalograma (EEG) é uma técnica não-invasiva que realiza o registro das oscilações do potencial elétrico do cérebro, a partir de eletrodos fixados no escalpo (NUNEZ; SRINIVASAN, 2006).

Seu surgimento ocorreu em 1875, através do trabalho realizado por Richard Caton, um fisiologista inglês que, utilizando um mecanismo sensível à voltagem na superfície exposta do encéfalo de cães e coelhos, descobriu a existência de correntes elétricas no cérebro, e realizou o seu registro (BEAR; CONNORS; PARADISO, 2006).

Hans Berger, um psiquiatra austríaco, foi o primeiro a fazer a descrição do EEG em humanos, em 1929. Em seu experimento, ele observou que havia diferenças no sinal de EEG entre o estado de sono e de vigília, descobrindo assim as ondas Alfa.

O sinal gerado no cérebro deve passar por várias camadas de tecido não neural, incluindo as meninges (pia-máter, aracnoide e dura-máter), os ossos do crânio e a pele para alcançar os eletrodos localizados no escalpo e ser registrado. Por este motivo, a contribuição de um único neurônio cortical isolado é muito pequena. São necessários milhares de neurônios subjacentes ativados em conjunto para gerar um sinal grande o suficiente para ser visualizado (BEAR; CONNORS; PARADISO, 2006).

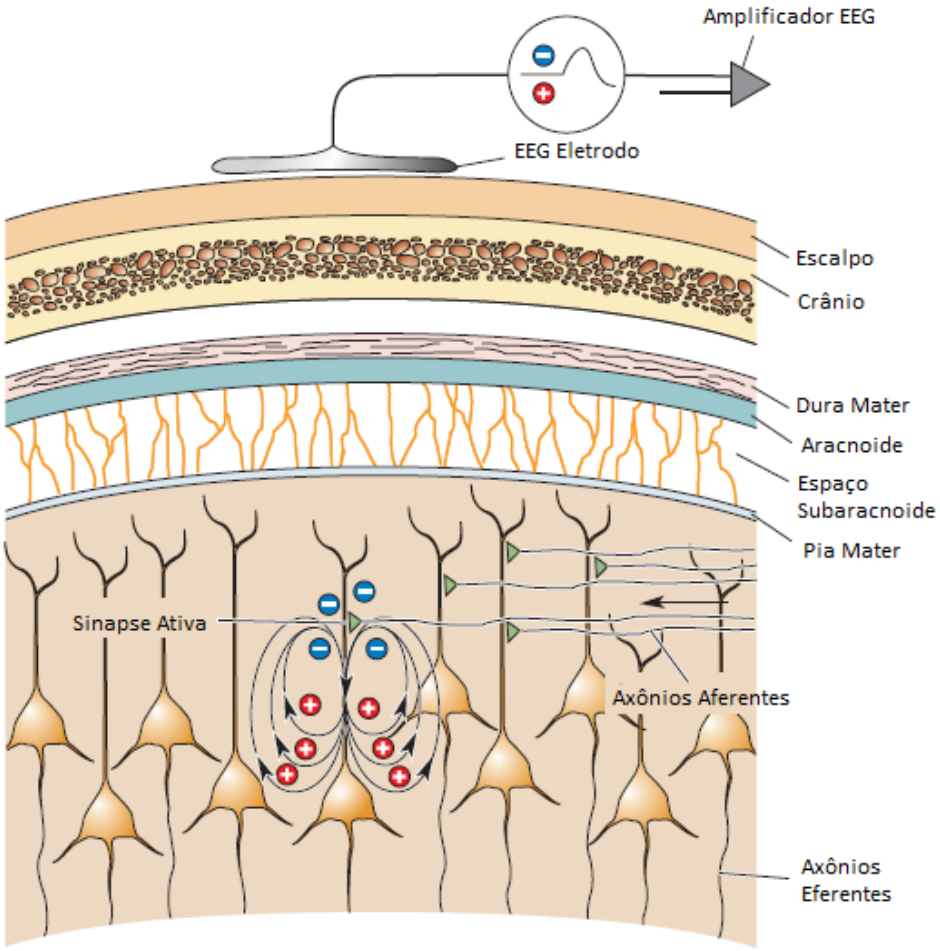
De acordo com Buzsáki (2006), os sinais medidos em EEG refletem uma ação cooperativa entre neurônios, ou seja, o que é medido no meio extracelular reflete um comportamento médio de vários neurônios interagindo uns com os outros. Segundo Bear, Connors e Paradiso (2006), a amplitude do sinal de EEG depende do quão síncrona é a atividade dos neurônios subjacentes.

Porém, qualquer membrana excitável pode contribuir para o campo extracelular. Segundo Buzsáki (2006), não somente neurônios, mas também a glia e até vasos sanguíneos podem contribuir com a medida de campo de potencial elétrico, onde este campo é uma superposição de todos os processos iônicos, desde rápidos potenciais de ação até lentas flutuações na glia (BUZSÁKI; ANASTASSIOU; KOCH, 2012),.

Como as correntes extracelulares de vários elementos individuais devem se sobrepor, ou seja, estarem sincronizadas, para produzirem um sinal grande o suficiente para ser medido no EEG, os eventos lentos, como as correntes sinápticas, são a maneira mais simples de serem medidas (BEAR; CONNORS; PARADISO, 2006; KANDEL et al., 2013).

Os Potenciais de Ação são eventos rápidos, que ocorrem em um intervalo de tempo menor que 10ms e raramente disparam de forma síncrona, enquanto que as correntes sinápticas,

Figura 6 – Geração do Sinal Elétrico do EEG.



Fonte: Adaptado de Bear, Connors e Paradiso, 2006, p.587.

produzidas pelos potenciais excitatórios pós-sinápticos (PEPS) e os inibitórios (PIPS) são eventos mais lentos e ocorrem em um número muito maior que os PAs em um intervalo de tempo consideravelmente maior, entre 50 a 200 ms (MARCUSE; FIELDS; YOO, 2016).

Sendo assim, um campo de potencial é gerado durante a somação do PIPS e PEPS na rede neural, formando correntes elétricas que fluem dentro e ao redor das células e esse fluxo de corrente cria um campo que se espalha ao redor da origem do sinal elétrico (MARCUSE; FIELDS; YOO, 2016). O efeito deste campo de potencial diminui conforme aumenta a distância entre a fonte do sinal.

Portanto, as correntes elétricas que são registradas pelo EEG de escalpo amostram, principalmente, as atividades sinápticas que ocorrem nas camadas superficiais do córtex por neurônios piramidais localizados nas camadas III, V e VI (OLEJNICZAK, 2006), quando axônios aferentes ipsilaterais fazem sinapse com os dendritos destes neurônios, como mostra a Figura 6.

2.2.2 Fisiologia Do EEG

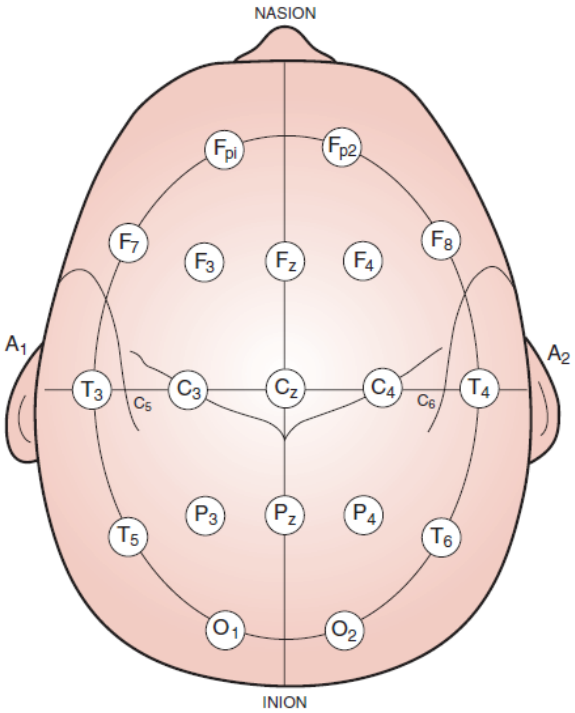
O EEG apresenta uma boa resolução temporal, na ordem de milissegundos, porém com baixa resolução espacial, diferente do *Magnetic Resonance Imaging* (MRI) e o *Positron Emission Tomography* (PET) que possuem boa resolução espacial, mas baixa resolução temporal (BUZSÁKI, 2006).

O eletrodo é o elemento inicial para o registro de biopotenciais, pois é através dele que é possível converter um sinal gerado por um biopotencial em um sinal elétrico capaz de ser analisado (USAKLI, 2010).

Os eletrodos de EEG são feitos de metal, podendo ser produzidos em forma de disco, agulha ou microeletrodos (usados para medir potenciais intracranianos), sendo o Cloreto de Prata (AgCl) o material mais utilizado em aplicações de neurociência (USAKLI, 2010).

O posicionamento dos eletrodos deve seguir a convenção 10-20 recomendada pela *International Federation of Societies for Electroencephalografy and Clinical Neurophysiology*, sendo posicionados de acordo com marcos anatômicos determinados, sendo 10% ou 20% o intervalo dos eletrodos.

Figura 7 – Sistema 10-20 de posicionamento de Eletrodos



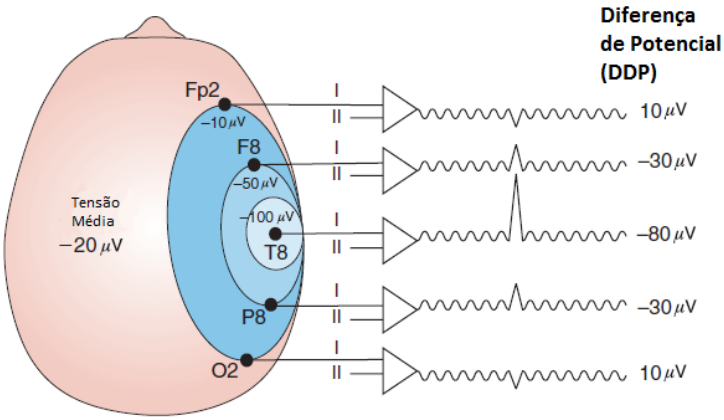
Fonte: Marcuse, Fields e Yoo, 2016, p.3.

A Figura 7 ilustra um exemplo do posicionamento de eletrodos seguindo o padrão 10-20, onde é possível verificar que os eletrodos designados com números ímpares são colocados

no lado esquerdo, enquanto que os eletrodos designados com números pares são colocados do lado direito.

Segundo Malmivuo e Plonsey (1995), a atividade espontânea produzida pelo cérebro capturada pelo EEG tem uma amplitude de cerca de 100 μV e 100 Hz quando medido pelo escalpo e entre 1-2 mV quando medido na superfície do cérebro (intracraniana).

Figura 8 – Registro Referencial



Fonte: Adaptado de Marcuse, Fields e Yoo, 2016, p.10.

Existem duas formas para registrar os sinais de EEG, sendo eles o Registro Bipolar, onde é gravada a diferença de potencial entre dois eletrodos posicionados no escalpo, e o Registro Referencial (Figura 8), onde é gravada a diferença de potencial entre um eletrodo e um ponto de referência que não é afetado pelos potenciais gerados pela atividade cerebral como, por exemplo, as orelhas.

2.2.2.1 Mapas Topográficos Corticais

Após ser realizado o registro da atividade elétrica são gerados os mapas topográficos corticais. Este mapeamento dos sinais de EEG começou a ser demonstrado nos anos 1950s e 1960s (DUFFY, 1989).

Para gerar os mapas topográficos corticais, é necessário fazer uma interpolação para haver uma melhor visualização da atividade elétrica nas regiões do escalpo que estão entre os eletrodos. A interpolação tem por objetivo utilizar os valores conhecidos do potencial elétrico registrado nos eletrodos, para estimar os valores que são desconhecidos em outros pontos no escalpo.

Segundo Hinrichs, Simon e Künkel (1989), existem, em geral, duas classes de algoritmos de interpolação para a geração dos mapas topográficos corticais, os globais, onde a interpolação depende dos valores obtidos em todos os eletrodos no escalpo, e os locais, onde a interpolação depende apenas dos valores obtidos entre os eletrodos vizinhos ao observado. Porém é importante ressaltar que todas as técnicas de interpolação são aproximações e os métodos

existentes variam no número de eletrodos utilizados e no fator de peso aplicado a cada eletrodo (DUFFY, 1989).

Um dos métodos mais comuns para interpolação é o método de interpolação linear dos 3 vizinhos mais próximos, desenvolvido em 1979 (DUFFY, 1989), onde "áreas triangulares formadas por eletrodos são preenchidas por combinações lineares dos valores registrados pelos eletrodos naquela área" (KOLES; PARANJAPE, 1988).

O método de interpolação dos k-vizinhos mais próximos é calculado de acordo com a Equação (2) (PERRIN et al., 1987; SIVANANDAN, 2013), considerando $e = (x,y)$ a coordenada de um ponto arbitrários no escalpo, $e_j = (x_j,y_j)$ a coordenada do eletrodo j, onde j varia de 1 ao número de eletrodos, k é o número de vizinhos mais próximos, z_j é o valor do potencial elétrico no eletrodo e_j, m é a ordem da interpolação e d_j é a distância entre e e e_j, ou seja,

$$V_m(x,y) = \frac{\sum_{j=1}^{k} z_j d_j^m}{\sum_{j=1}^{k} d_j^m}, \tag{2}$$

onde,

$$d_i^2 = (x - x_j)^2 + (y - y_j)^2. \tag{3}$$

O uso de cores no mapeamento é uma ferramenta útil para enfatizar distinções, destacando os limiares no mapeamento da superfície do córtex e provendo uma relação de gradiente da superfície (DUFFY, 1989).

Quando ocorre um potencial pós-sináptico em um neurônio, um pequeno dipolo é criado e o sinal de EEG é obtido a partir da soma dos dipolos de milhares de neurônios (LUCK, 2014), podendo ser registrado nos eletrodos valores positivos ou negativos do potencial elétrico. A toolbox ERPlab apresenta como padrão para a geração dos mapas a cor Azul para valores mais negativos e a cor Vermelha para valores mais positivos do potencial elétrico registrado no eletrodo.

Os mapas topográficos podem trazer informações importantes sobre a localização cortical de uma fonte geradora. Para realizar uma comparação entre mapas topográficos, os mapas devem ser escalados em uma unidade de medida global (D. LEHMANN, 1989).

Segundo Murray, Brunet e Michel (2008), diferenças topográficas possuem uma interpretabilidade neurofisiológica direta e mudanças do potencial elétrico no escalpo apenas podem ser causadas por mudanças nas configurações cerebrais que ocorrem de fontes intracranianas.

Utilizando a Toolbox ERPlab, é possível gerar mapas topográficos corticais, que utiliza o método de interpolação do vizinho mais próximo, sendo gerados a partir do valor instantâneo do potencial elétrico registrado nos eletrodos, ou do potencial médio entre dois instantes fixados no tempo em um sinal de EEG, gerando um mapa topográfico com uma escala de cor que apresenta os valores mínimos e máximos dos valores registrados nos eletrodos.

2.2.2.2 Características dos ritmos em EEG

No sistema nervoso central é possível observar oscilações neurais que são geradas por grupos de neurônios e estas oscilações são caracterizadas pela sua frequência, amplitude e fase. De acordo com Kamel e Malik (2014), funções cognitivas como transferência de informação, percepção, controle motor e memória estão relacionadas com as oscilações neurais e a sincronização.

Os neurônios podem gerar PAs em padrões rítmicos que são fundamentais para a codificação da informação (KAMEL; MALIK, 2014) e o sinal de EEG é descrito em termos de bandas de frequência.

As bandas de frequências do EEG são normalmente subdivididas em faixas de frequência e podem estar relacionadas a diversos estados físicos e comportamentais (COLLURA, 1997; BUZSÁKI, 2006; MULERT; LEMIEUX, 2010):

a) Ondas Delta (0,5 a 4 Hz): São mais encontradas em crianças e, conforme envelhece, o ser humano passa a produzir menos ondas Delta. As ondas Delta estão relacionadas ao sono não-REM (*Rapid Eye Moviment*), e ocorrem de forma síncrona em todo o córtex, estando associadas ao estado de inconsciência e de sono profundo;

b) Ondas Teta (4 a 8 Hz): São normalmente encontradas em fases do sono em qualquer idade. Ocorrem de forma contínua enquanto um comportamento é executado, estando relacionadas a um estado de atenção, criatividade, intuição;

c) Ondas Alfa (8 a 12 Hz): São encontradas em todos os grupos de pessoas. Ocorrem quando são fechados os olhos e aparecem de forma mais prominente nas regiões occipitais do córtex, estando relacionada a um estado de relaxamento e consciência, mas pode ser bloqueada por vários meios, como movimento dos olhos, imaginação visual e atividade mental, como calculo aritmético;

d) Ondas Beta (12 a 30 Hz): São encontradas em todos os grupos e idades de pessoas. São caracterizadas por ondas de baixa amplitude e alta frequência, observadas enquanto se está acordado. As ondas Beta estão relacionadas a um estado de consciência, atividade mental, pensamento lógico, planejamento, ou seja, são apresentadas durante o dia a dia para completar tarefas conscientes;

e) Ondas Gama (>30 Hz): As ondas Gama estão relacionadas com o alto nível de processamento de informação, sendo importante para processos de aprendizado e memória.

2.2.2.3 Artefatos em EEG

Os sinais de EEG possuem artefatos que, de acordo com Anghinah et al. (2006), são definidos como "todo potencial elétrico proveniente de outra fonte que não seja o cérebro". Algumas fontes de artefatos são (ARGOUD, 2001):

a) O equipamento de EEG: Gerados pelo mau funcionamento do equipamento;
b) Interferências elétricas externas: A mais comum é a frequência de oscilação da rede (60 Hz no Brasil);
c) O eletrodo: A qualidade do eletrodo, do gel condutor, o estado da pele, etc. alteram a impedância do eletrodo;
d) O Paciente: O próprio paciente pode ser gerador de artefatos, como o movimento dos olhos, piscar dos olhos, movimento do corpo, etc.

2.2.3 Potencial Relacionado ao Evento

A cada instante no tempo o cérebro apresenta um estado funcional global resultante da interação entre as novas informações recebidas pelo cérebro e a atividade espontânea cerebral, tendo como consequência a iniciação de uma série de diferentes e breves estados cerebrais que manifestam diferentes passos e aspectos no processamento da informação (D. LEHMANN, 1989).

Os Potenciais Relacionados ao Evento, ou *Event-Related Potential* (ERP), ou Potencial Evocado, são "pequenos potenciais elétricos gerados nas estruturas cerebrais em resposta a eventos específicos ou estímulos" (SUR; SINHA, 2009). A análise de ERP é uma técnica que registra os potenciais elétricos que são evocados a partir de um determinado estímulo. Segundo Luck (2014), o ERP quase sempre reflete potenciais pós-sinápticos em vez de potenciais de ação. A exceção a esta regra são as respostas auditivas evocadas no tronco encefálico, que refletem o disparo sincronizado de PAs na ordem de poucos milissegundos.

O ERP possuí uma excelente resolução temporal, ao custo da sua resolução espacial e, segundo Woodman (2010), "o ERP nos permite observar uma série de operações cognitivas que ocorrem quando uma informação sensorial é enviada pelo sistema nervoso central, antes mesmo que uma resposta comportamental seja realizada".

Aos diferentes aspectos do processamento das novas informações que são apresentadas ao cérebro é dado o nome de "componentes" (D. LEHMANN, 1989). Os componentes de ERP são respostas sensoriais obtidas pelo ERP caracterizadas pela sua latência, morfologia e topografia (KAMEL; MALIK, 2014). De acordo com Luck (2014), p.68, os componentes de ERP podem ser definidos operacionalmente como "um conjunto de mudanças de voltagem que variam sistematicamente com relação ao tempo, aos indivíduos, a amplitude, etc.".

A maioria dos componentes de ERP são representados pela letra N, quando a polaridade é negativa, e P, quando a polaridade é positiva, seguidos por um número que representa ou a

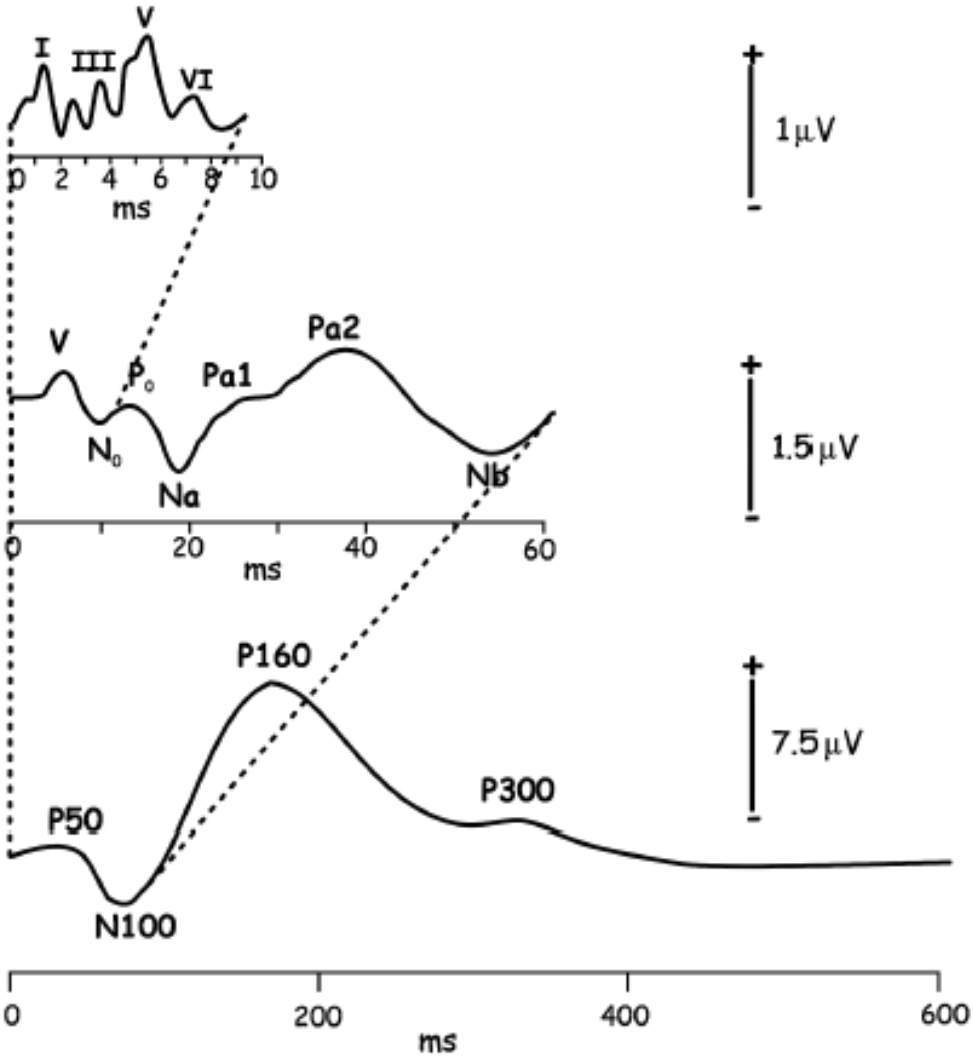
posição temporal do pico na forma de onda (e.g. N300 para um pico negativo em 300ms) ou a posição ordinal do pico na onda (e.g. N1 para o primeiro pico negativo).

Segundo Luck (2014), os componentes de ERP são divididos em 3 categorias principais:

a) Componentes Sensoriais Exógenos: Ativados pela presença de um estímulo;
b) Componentes Endógenos: Refletem o processamento neural dependente de uma tarefa;
c) Componentes Motores: Relacionados com a preparação e execução de uma resposta motora.

A Figura 9 ilustra os componentes ERP evocados a partir da apresentação de um estímulo auditivo.

Figura 9 – Sequência típica dos componentes ERP a partir de estímulos auditivos



Fonte: Adaptado de Luck, 2014, p.80.

O Potencial Evocado Auditivo de Tronco Encefálico (PEATE) é nomeado com algarismos romanos e possui picos nos primeiros 10 ms após o estímulo. Eles refletem o fluxo de informação vindo da cóclea ao tronco encefálico e ao tálamo, normalmente utilizados para analisar a integridade da via auditiva. São normalmente chamados de Potenciais de Curta Latência. O Potencial de Média Latência possui picos entre 10 e 50 ms e surgem em parte pelo Núcleo Geniculado Medial (NGM) e pelo córtex auditivo primário.

A transmissão da informação ao longo da via auditiva é muito rápida. Segundo Luck (2014), a partir da perspectiva do processamento sensorial auditivo, 100 ms é relativamente tarde. As respostas dos Potenciais de Longa Latência (e.g. P1, N1, P2) podem ser influenciadas por fatores de alto nível, como a atenção e a excitação.

A resposta N1-P2 é um componente sensorial exógeno, ou seja, não envolve o processamento cognitivo, e é afetado pelo nível de excitação, atenção e estado de alerta. A componente N1 reflete o estado de atenção ao estímulo apresentado, que ocorre quando uma série de estímulos auditivos idênticos são apresentados, provocando uma onda negativa próxima de 100 ms pós-estímulo (BROSSI et al., 2007), que precede o componente P2, onde ocorre uma onda positiva próxima de 200 ms pós-estímulo.

O *Mismatch Negativity* (MMN) ocorre apenas no sistema auditivo e é gerado quando um estímulo raro é apresentado, sem requerer a atenção do voluntário, com latência entre 150 a 275 ms pós-estímulo. Segundo Brossi et al. (2007), quando uma série de estímulos, que provocam o N1, é apresentado ao voluntário, e um estímulo diferente for inserido nesta série, aparecerá um pico negativo adicional, que permanece por outros 100 ms, o MMN.

O componente P3 é extraído durante a realização de tarefas específicas. A amplitude do P3 é maior quando o sujeito demanda esforço para realizar a tarefa, mas é menor quando há uma incerteza sobre o estímulo (LUCK, 2014).

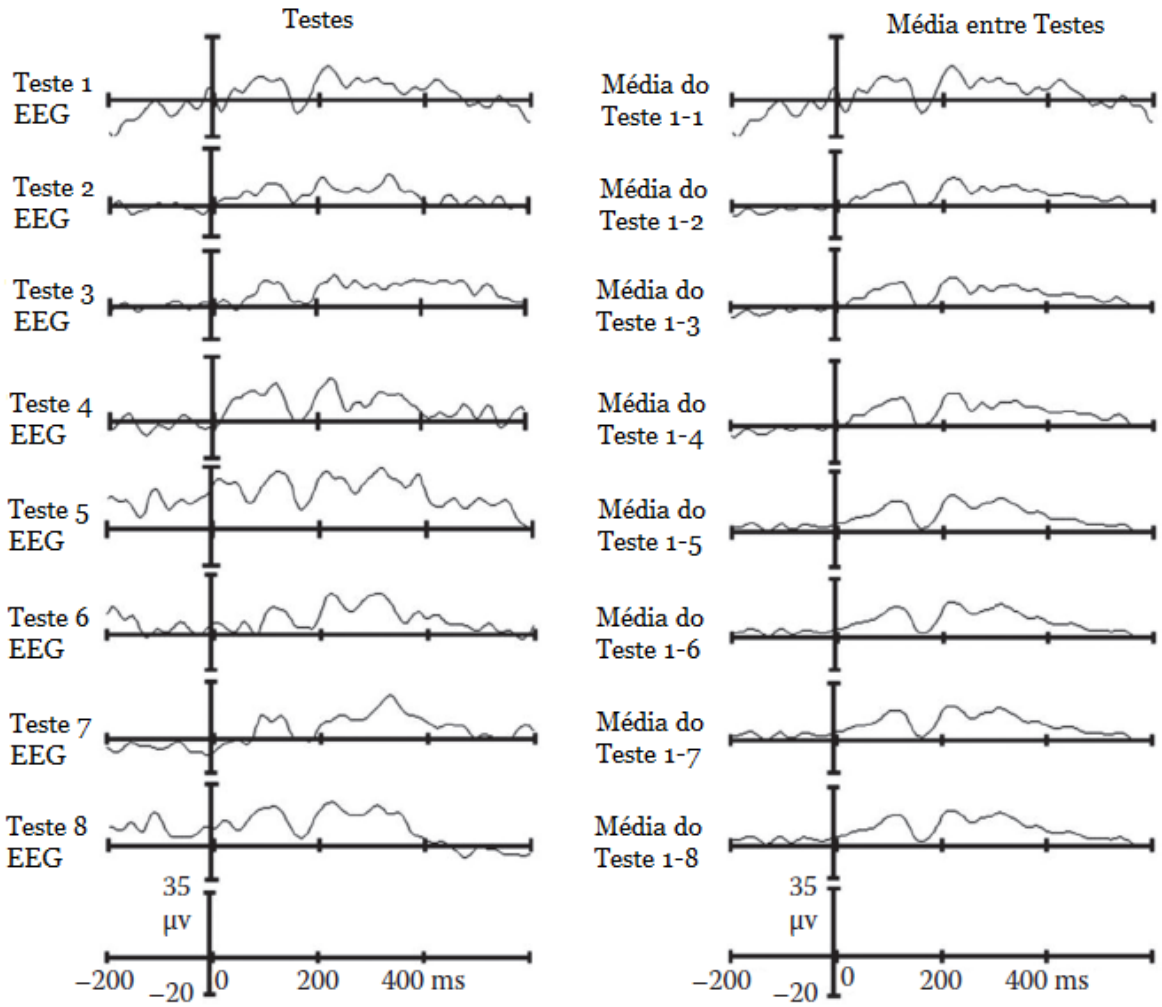
De acordo com Kamel e Malik, 2014, p.87:

> O ERP permite compreender as características temporais e espaciais da atividade neural relacionada aos processos que compõem o comportamento como a atenção seletiva, a codificação de informação, a resposta seletiva, o controle inibitório e o monitoramento de performance. O ERP proporciona um nível de análise mais profundo e objetivo sobre o comportamento.

Quando os estímulos são apresentados repetidamente, é comum realizar a média dos potenciais elétricos (WOODMAN, 2010), sendo eles sincronizados de acordo com o instante em que ocorre o estímulo, pois espera-se que para este estímulo apresentado as respostas neurais sejam semelhantes entre si, tendo como objetivo eliminar os ruídos e artefatos presentes no sinal de EEG e destacar as respostas que são comuns entre cada estímulo e atenuar as respostas que sejam diferentes.

As médias são realizadas entre os estímulos apresentados a um voluntário, mas em alguns casos também é comum realizar uma média geral entre a resposta de todos os voluntários ou de um grupo de voluntários, para comparar a resposta entre os grupos.

Figura 10 – Exemplo de aplicação da média do sinal de EEG entre testes



Fonte: Adaptado de Kamel e Malik, 2014, p. 74.
Legenda: Coluna da esquerda representa os testes e da direita a média entre os testes.

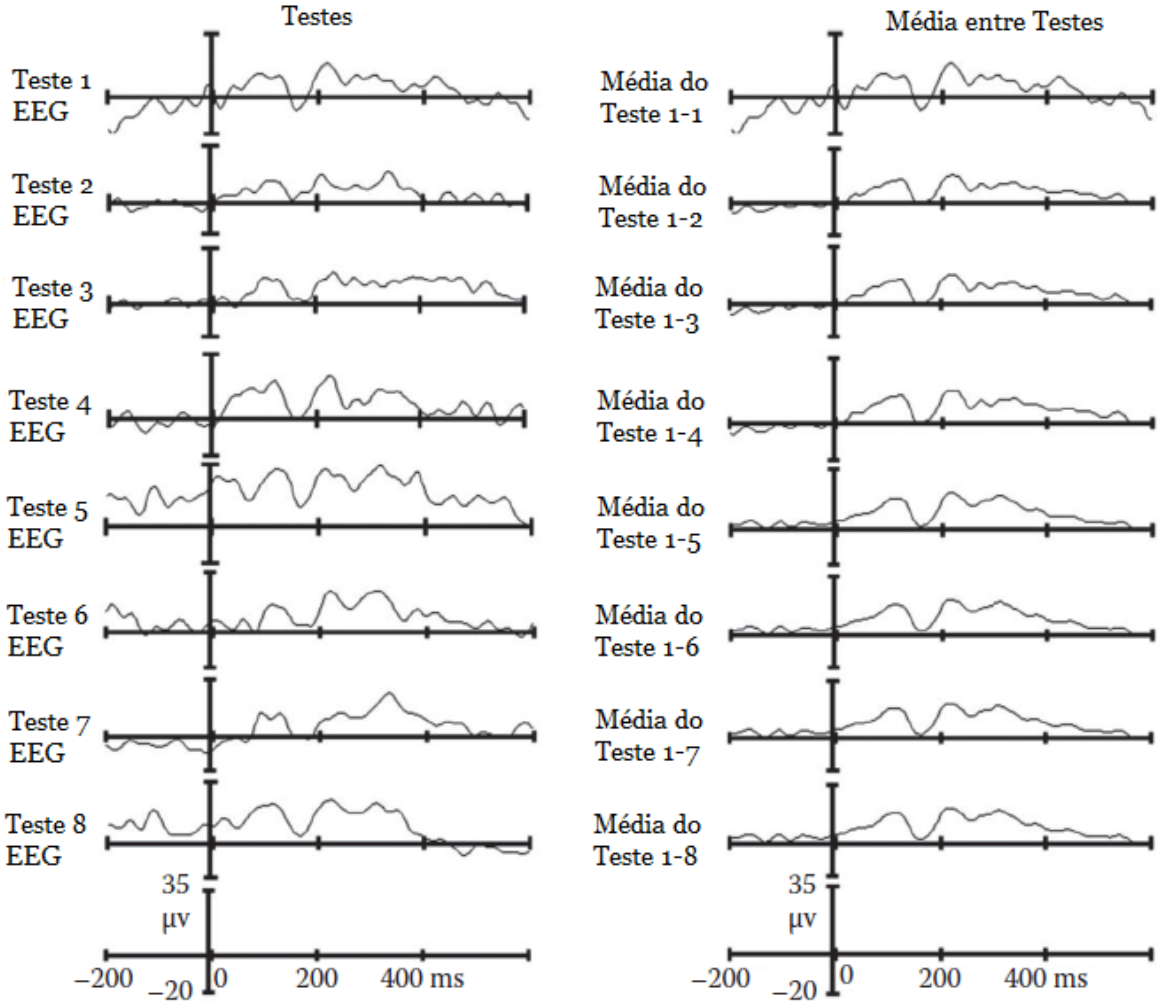
Na Figura 10 é possível verificar a aplicação da média de uma janela (chamada época) do sinal de EEG, sendo o instante zero considerado como o instante em que ocorreu o estímulo e determinando uma janela de 200 ms antes de ocorrer o estímulo e 600 ms após a ocorrência do estímulo.

A coluna a esquerda apresenta os testes, ou estímulos, apresentados a um único voluntário. É possível verificar que para cada estímulo apresentado a resposta neural do voluntário é distinta, possuindo alguns pontos em comum, como um vale próximo de 100 ms após o estímulo, seguido de um pico próximo de 200 ms.

Na coluna a direita é realizada a média dos testes, ou estímulos, apresentados ao voluntário. Note que conforme é acrescentado testes ao resultado da média, o sinal resultante se torna mais característico, ou mais próximo ao esperado, removendo ruídos dos sinais anteriores e destacando os comportamentos que eram comuns a todos os testes.

Sendo assim, ao realizar a média entre os estímulos apresentados a um único voluntário, é possível observar e evidenciar o que varia em comum nas respostas cerebrais do voluntário para os estímulos apresentados.

Figura 11 – Exemplo de aplicação da média do sinal de EEG entre sujeitos



Fonte: Adaptado de Kamel e Malik, 2014, p. 77.
Legenda: Coluna da esquerda representa a média dos sujeitos e da direita a média destes sinais.

A Figura 11 apresenta o mesmo padrão de janelamento da figura anterior, porém, a média está sendo realizada entre voluntários, em um passo posterior ao realizado na Figura 11, onde a média dos estímulos de cada voluntário já foram calculadas individualmente e agora é realizada a média do sinal resultante de cada voluntário.

A coluna a esquerda apresenta a média do sinal de EEG de cada voluntário e a coluna a direita apresenta a média entre todos os voluntários. Da mesma forma que o anterior, conforme são acrescentados os sinais de cada voluntário ao resultado da média, o sinal resultante destaca os comportamentos que são comuns entre todos os voluntários.

Desta forma, ao realizar a média entre os voluntários é possível observar e evidenciar quais são as respostas cerebrais que variam em comum entre todos os voluntários, podendo

assim demonstrar um certo padrão de ativação cerebral para o estímulo apresentado para aquele conjunto de voluntários.

É importante observar que ao realizar as médias, os sinais de EEG devem estar sincronizados de acordo com o instante em que os estímulos são apresentados aos voluntários, havendo assim um alinhamento temporal para a realização das médias.

2.2.4 Considerações Complementares

Conforme foi descrito, o sinal registrado pelo EEG amostra principalmente as transmissões sináptica nas camadas superficiais do córtex, através de eletrodos posicionados no escalpo de acordo com o Sistema 10-20. Sua vantagem em frente a outros equipamentos é a sua resolução temporal, na ordem de milissegundos.

Utilizando o EEG é possível utilizar a técnica de ERP, que permite analisar os potenciais elétricos evocados que são registrados a partir de um determinado estímulo.

2.3 O SISTEMA AUDITIVO

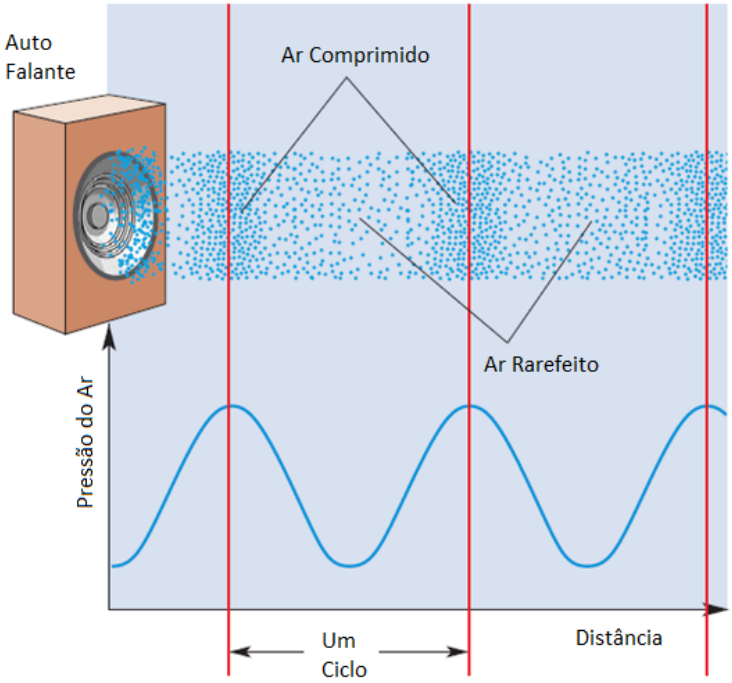
Nesta seção é apresentado o sistema auditivo, mostrando as estruturas que o compõem e como o som é processado por ele, descrevendo de forma simplificada a via auditiva desde a orelha externa até o córtex auditivo e é apresentado também alguns estudos sobre o processamento musical.

2.3.1 O Som

Em um alto-falante, o deslocamento do cone, que é colocado para vibrar por uma bobina eletromagnética, desloca-se para frente e para trás repetidamente, deslocando as partículas que constituem o ar e ao se mover para frente o cone comprime essas partículas e ao se mover para trás ele as descomprime, conforme mostra a Figura 12. Semelhantemente, o som é gerado por vibrações do ar e o nosso sistema auditivo é capaz de perceber estas vibrações (BEAR; CONNORS; PARADISO, 2006).

Esta vibração se propaga em todo o espaço a uma velocidade de aproximadamente 340 m/s (KANDEL et al., 2013). Essas vibrações do ar que produzem os sons são chamadas de Ondas Sonoras. Os sons são, então, vibrações capazes de estimular o sistema auditivo, provocando uma percepção (LENT, 2001).

Figura 12 – Produção do som pela variação da pressão do ar



Fonte: Adaptado de Bear, Connors e Paradiso, 2006, p. 345.

2.3.2 Estrutura Do Sistema Auditivo

O aparelho auditivo é formado por três estruturas conhecidas como: orelha externa, média e interna. Na Figura 13 é possível ver de forma geral os componentes que compõe o aparelho auditivo.

A orelha externa é composta por três estruturas chamadas: Pavilhão Auricular, Concha e Meato Auditivo Externo. De acordo com Kandel et al. (2013), o pavilhão auricular atua como um refletor, permitindo que o som seja capturado de forma efetiva e então focando-o para o meato auditivo externo, ou seja, para dentro do canal auditivo. O meato auditivo externo termina no tímpano, que começa a vibrar quando um estímulo sonoro incide sobre ele.

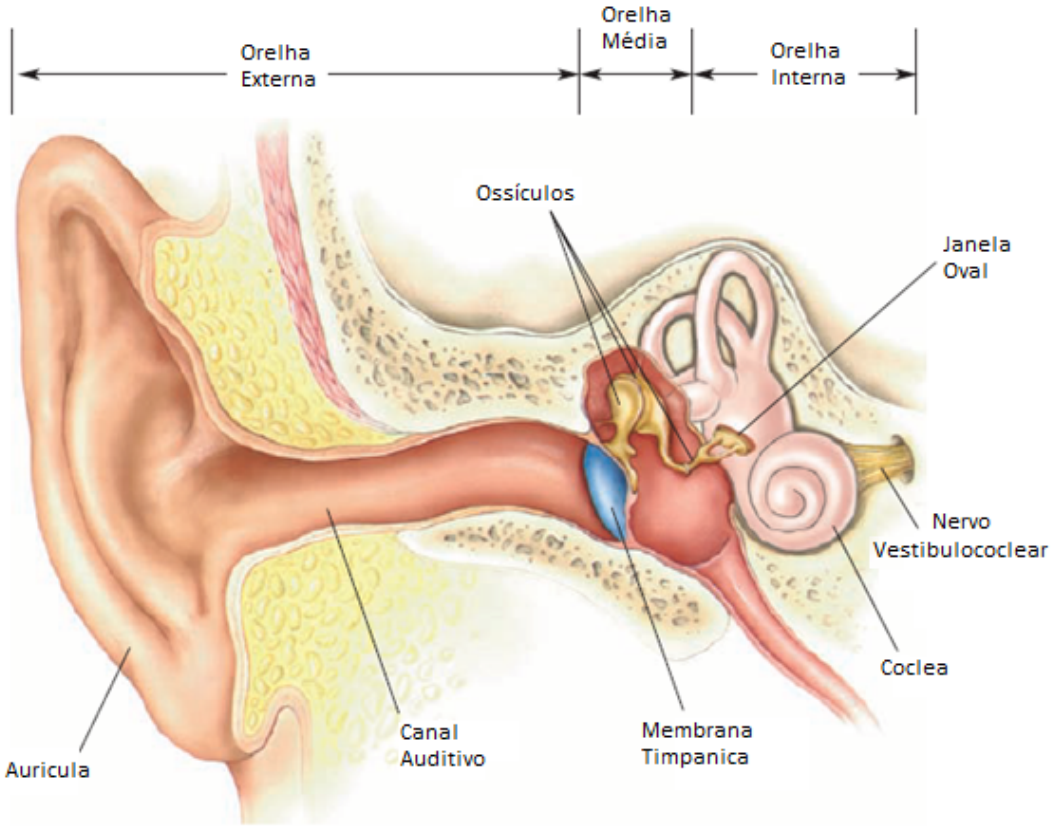
O tímpano separa a orelha externa da orelha média que é uma cavidade cheia de ar que contêm três ossículos articulados entre si: o Martelo, a Bigorna e o Estribo. Eles são capazes de transmitir as vibrações do tímpano para uma segunda membrana, que cobre um orifício no osso do crânio, chamada de Janela Oval.

Ligado ao tímpano está o martelo, conectado com a bigorna que forma uma conexão flexível com o estribo. O estribo possui uma porção achatada que se move para dentro e para fora da janela oval transmitindo as vibrações sonoras aos fluídos da cóclea na orelha interna (BEAR; CONNORS; PARADISO, 2006).

A membrana da janela oval separa a orelha média da orelha interna e atrás dela está a Cóclea que é preenchida por um fluido, onde se encontram os receptores auditivos (LENT,

2001). Na orelha interna é gerado uma resposta neural ao som e este sinal é processado por uma série de núcleos no tronco encefálico. De acordo com Bear, Connors e Paradiso (2006), este sinal é enviado ao Núcleo Geniculado Medial (NGM), que é um núcleo de retransmissão no tálamo e, por fim, o NGM se projeta ao Córtex Auditivo Primário, ou A1, no lobo temporal.

Figura 13 – Orelha externa, média e interna



Fonte: Adaptado de Bear, Connors e Paradiso, 2006, p. 347.

2.3.2.1 A Cóclea e o Processamento Auditivo

A cóclea possui uma forma em espiral que lembra uma concha de caracol. A sua estrutura diminui de diâmetro progressivamente, tendo três câmaras no seu interior preenchidas por um fluído: a Escala Vestibular e Timpânica, preenchidas por um fluido chamado Perlinfa, e a Escala Média, preenchida por um fluido chamado Endolinfa.

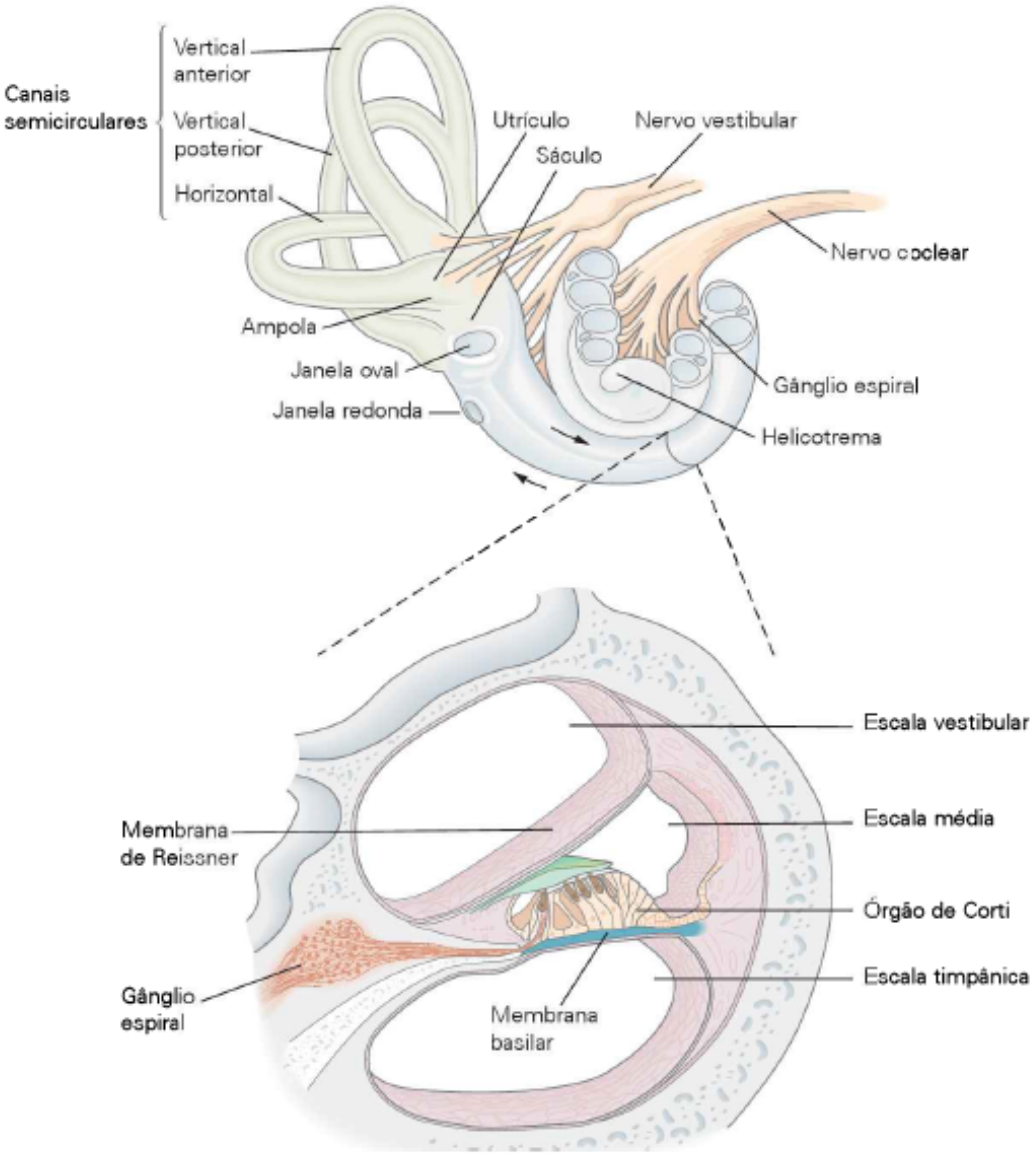
A Membrana de Reissner separa a Escala Vestibular da Escala Média e a Membrana Basilar separa a Escala Timpânica da Escala Média. No ápice da cóclea, a Escala Média está fechada e a Escala Timpânica se torna contínua com a Escala Vestibular através de um orifício nas membranas, chamado de Helicotrema.

Segundo Bear, Connors e Paradiso (2006), p.351, “a Membrana Basilar é cerca de cinco vezes mais larga no ápice do que na base e a rigidez da membrana diminui da base em direção ao ápice, sendo que a base é aproximadamente 100 vezes mais rígida”.

Apoiado na Membrana Basilar e coberto pela Membrana Tectorial está o Órgão de Corti, onde se encontram os neurônios receptores auditivos (SQUIRE et al., 2013). É possível ver as estruturas que compõem a cóclea na Figura 14.

Como os fluidos que preenchem a cóclea são incompressíveis, no momento que os ossículos movem a membrana que cobre a janela oval há um aumento ou diminuição de pressão no interior da cóclea, dependendo do movimento realizado. Como as Membranas Basilar e de Reissner são flexíveis, elas se movimentam em resposta ao som (BEAR; CONNORS; PARADISO, 2006).

Figura 14 – A estrutura da Cóclea



Fonte: Adaptado de Kandel et al., 2013, p. 657.

Qualquer pressão diferente entre a Escala Vestibular e a Escala Timpânica gera movimento na Membrana Basilar (KANDEL et al., 2013). Este movimento gerado faz com que a Membrana Basilar se mova próxima à sua base, iniciando a propagação de uma onda em direção ao seu ápice. Dependendo da frequência do som gerado, a onda percorrerá uma distância diferente na Membrana Basilar. A resposta da Membrana Basilar estabelece um código de localização chamado Tonotopia (BEAR; CONNORS; PARADISO, 2006).

No Órgão de Corti encontram-se os mecanorreceptores responsáveis por transformar a energia do estímulo em energia elétrica. Os mecanorreceptores da cóclea são as Células Ciliadas, que são células epiteliais polarizadas.

As células ciliadas possuem Estereocílios no seu lado apical. As células ciliadas externas mantêm suas extremidades na membrana tectorial, enquanto que as células ciliadas internas mantêm suas extremidades abaixo da membrana tectorial. A parte apical das células ciliadas estão imersas na endolinfa, enquanto que a parte basal está imersa na perilinfa.

Quando o estribo se move após um estímulo sonoro, a Membrana Basilar se move em resposta e toda a estrutura que sustenta as células ciliadas se move em resposta. Esta movimentação gera um deslocamento nos estereocílios para um lado ou para o outro (BEAR; CONNORS; PARADISO, 2006).

Quando os cílios se deslocam, ocorrem alterações no potencial de receptor da célula ciliada. Ao mover-se em uma direção, ocorre a abertura de canais de potássio na extremidade dos estereocílios, aumentando a corrente de entrada de potássio neles. Ao mover-se na direção contrária estes canais de potássio se fecham completamente, prevenindo o influxo de potássio (BEAR; CONNORS; PARADISO, 2006).

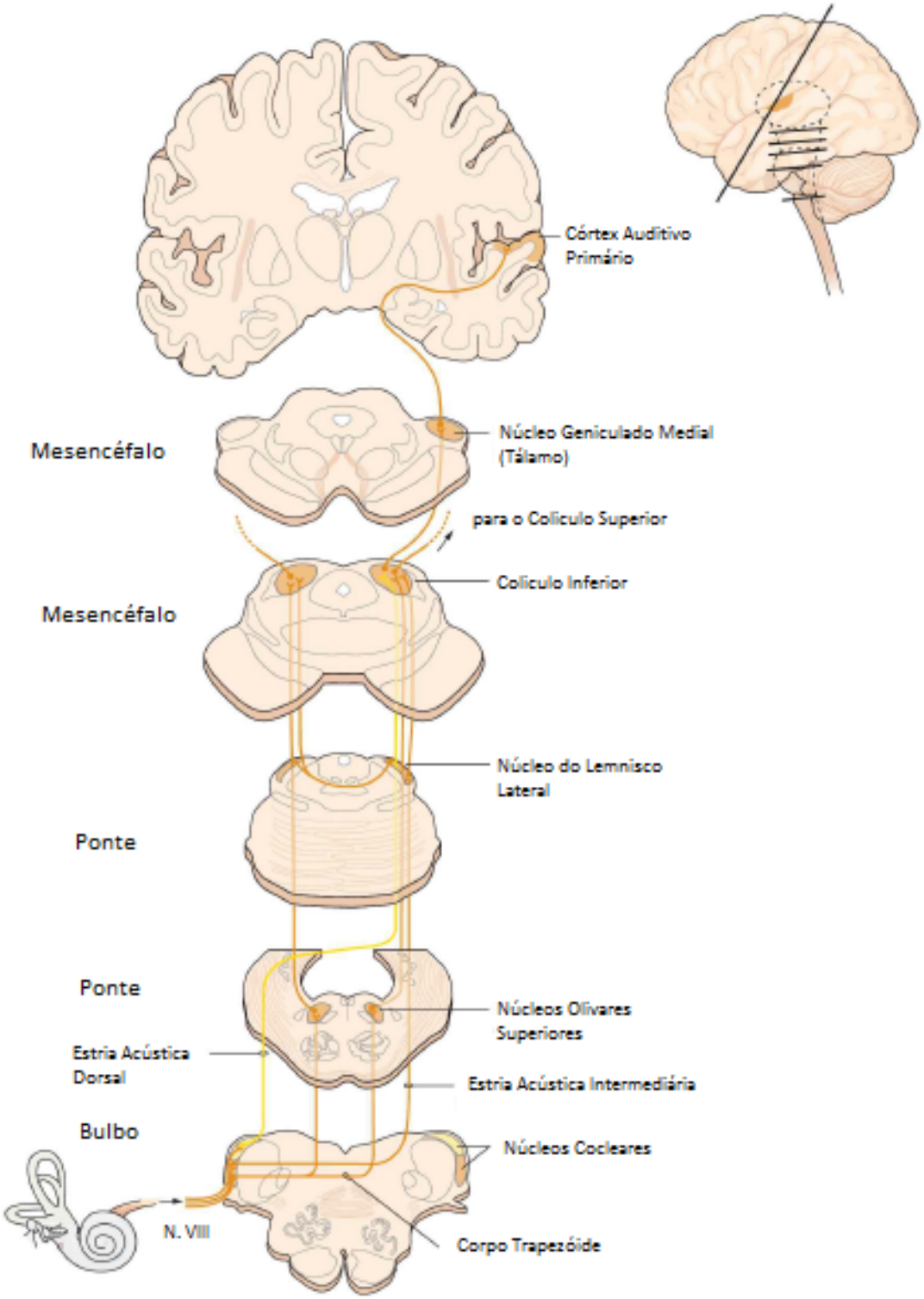
A entrada de potássio na célula ciliada causa a sua despolarização, ativando canais de cálcio dependentes de voltagem que, por sua vez, dispara a liberação do neurotransmissor glutamato, provocando, no neurônio bipolar, potenciais pós-sinápticos excitatórios e salvas de potenciais de ação propagados através das fibras auditivas (LENT, 2001).

Os neurônios, cujos corpos celulares estão no gânglio espiral, estabelecem sinapses com as células ciliadas e seus axônios entram no Nervo vestíbulo-coclear (VIII Nervo Craniano), que se projeta aos núcleos cocleares no Bulbo (BEAR; CONNORS; PARADISO, 2006).

A informação auditiva codificada é conduzida através do nervo auditivo, e penetra bilateralmente no nível do bulbo ao Núcleo Coclear Dorsal e Ventral. Cada axônio se ramifica, estabelecendo sinapses em neurônios de ambos os núcleos cocleares (LENT, 2001).

Segundo Bear, Connors e Paradiso (2006), p.364, “as células do Núcleo Coclear Ventral projetam seus axônios à Oliva Superior (ou núcleo olivar superior), de ambos os lados do tronco encefálico”. Estes neurônios sobem pelo Lemnisco Lateral (conjunto de axônios) e inervam no Colículo Inferior, no mesencéfalo (LENT, 2001). Os neurônios do Colículo Inferior enviam seus neurônios ao NGM do tálamo que, então, projeta-se ao Córtex Auditivo. O percurso feito pelas vias auditivas é ilustrado na Figura 15.

Figura 15 – Vias auditivas centrais



Fonte: Adaptado de Kandel et al., 2013, p. 685.

As vias auditivas terminam no córtex auditivo primário (A1 ou área 41 de Brodmann), onde as aferências fazem sinapses com as camadas corticais IIIb e IV. Nele há uma representação tonotópica de frequências distribuída ao longo da estrutura bidimensional do córtex, com

um leve gradiente de frequência em uma direção e contornos de isofrequências ao longo da outra direção (KANDEL et al., 2013; BEAR; CONNORS; PARADISO, 2006).

Na superfície de A1 são mapeados alguns parâmetros como: a latência de resposta neuronal, a percepção do volume e sua modulação e a taxa e duração da modulação da frequência (KANDEL et al., 2013).

Porém, de acordo com Kandel et al. (2013), p.710, "o reconhecimento do som é menos compreendido, e a multiplicidade de áreas auditivas corticais envolvidas sugere que sua análise será um grande desafio". Isso significa que ainda há muito para se compreender sobre o processamento auditivo.

Sendo assim, uma vez que o som foi transformado em uma resposta elétrica na cóclea, a percepção auditiva é gerada a partir de uma rica hierarquia de circuitos auditivos que analisam e processam estes sinais (KANDEL et al., 2013).

2.3.3 Processamento Musical

O processamento da música pelo cérebro tem chamado a atenção dos pesquisadores nos últimos tempos. Segundo Peretz e Zatorre (2004), um dos motivos para isso é o fato da música "oferecer uma oportunidade única de entender a organização do cérebro humano".

Ao ouvir uma música é possível perceber as suas características como:

a) Ritmo: Duração de uma série de notas e a combinação de seus valores no tempo;
b) Andamento: Velocidade de execução, o movimento rápido ou lento de execução da música;
c) Altura: Relacionado à frequência de uma nota em relação à sua posição relativa em uma escala musical;
d) Melodia: Combinação de sons sucessivos. É uma concepção horizontal da música;
e) Harmonia: Combinação de sons simultâneos. É uma concepção vertical da música;
f) Timbre: O que diferencia um instrumento de outro, produzido pelos harmônicos das vibrações de um instrumento.

A princípio, o estudo do processamento musical surgiu a partir de pacientes que possuíam algum tipo de anomalia ou lesão cerebral. Ao se comparar pacientes antes e após sofrerem uma determinada lesão, ou comparando um grupo de controle com pacientes com lesões, é possível descobrir qual o papel da área lesada, sendo esta uma grande fonte para compreender quais são as estruturas cerebrais que processam a música (WARREN, 2008).

De acordo com Ilari (2006), atualmente, acredita-se que o processamento musical é "um sistema neurocognitivo independente ou modular" e que "a música e os sons ambientes são processados por centros neuroanatômicos diferentes". Isso significa que não existe um "centro musical" que processa todas as informações musicais no cérebro (WARREN, 2008) e o estudo

da organização cerebral é uma grande ferramenta para compreender os mecanismos internos que fazem com que a música seja transformada em uma percepção.

Devido à sua natureza complexa, a música requer um alto nível de processamento neural para compreender todas as suas características. De acordo com Koelsch et al. (2000), o ser humano possui uma habilidade musical implícita, onde o cérebro humano (não-intencionalmente) antecipa os estímulos auditivos que estão por vir, "extrapolando expectativas que são consistentes com a teoria musical, mesmo em não-músicos".

Apenas uma pequena parcela de indivíduos se torna músico profissional, mas o que diferencia uma pessoa da outra é a intensidade com que cada um se dedica a diferentes tipos de atividades (A. C. LEHMANN; SLOBODA; WOODY, 2007). Segundo Peretz e Zatorre (2004), esta particularidade na distribuição de aquisição de habilidades confere ao músico um papel privilegiado na exploração da natureza e extensão da plasticidade cerebral, sendo que a plasticidade cerebral a capacidade do sistema nervoso de alterar a sua estrutura e funcionamento ao longo da vida de acordo com as necessidades e fatores do meio ambiente que está contido (LENT, 2001).

Para que o cérebro possa processar a música, o som causado pelo deslocamento das moléculas de ar deve fazer com que células ciliadas, localizadas na orelha interna, movam-se, transformando a informação sonora em um estímulo elétrico transmitido para os centros do tronco cerebral que se projetam para o córtex.

A frequência de vibração do som corresponde à localização das células ciliadas em uma relação tonotópica com a membrana basilar. A intensidade do som corresponde ao número de fibras ativadas (MUSZKAT; CORREIA; CAMPOS, 2000).

O nervo auditivo leva a informação sonora de maneira organizada ao córtex auditivo, localizado no Lobo Temporal. O lobo temporal direito tem um importante papel na decodificação da altura das notas (PERETZ; ZATORRE, 2004).

O córtex auditivo se conecta com o restante do encéfalo através de vários circuitos, como o hipocampo, que é uma área relacionada à memória e permite reconhecer a familiaridade das melodias e ritmos, o cerebelo e a amígdala, relacionados com a regulação motora e emocional, o núcleo acúmens, um pequeno núcleo de substância cinzenta relacionado ao sentido do prazer e recompensa (MUSZKAT; CORREIA; CAMPOS, 2000). Áreas do córtex motor e parietal estão relacionadas com a percepção do ritmo e com a ordem temporal (PERETZ; ZATORRE, 2004).

Vusst et al. (2014) cita vários trabalhos onde foi possível verificar mudanças estruturais no encéfalo de músicos e não-músicos, como diferenças morfológicas no Corpo Caloso, em áreas relacionadas à atividade motora, no aumento do volume cerebelar, aumento de massa cinzenta na região temporal em estruturas responsáveis pela linguagem e pela música e, em particular, em áreas de percepção auditiva.

Portanto, de acordo com Muszkat, Correia e Campos (2000), o hemisfério direito está relacionado com a discriminação da melodia, timbre e emoção nas áreas temporais e frontais,

enquanto que o hemisfério esquerdo está relacionado com o ritmo e a métrica. Havendo, portanto, uma especialização hemisférica para a música.

2.3.4 Considerações Complementares

Conforme foi descrito, o som passa por um longo circuito auditivo, vindo dos órgãos sensoriais, passando pelo tronco encefálico e o tálamo, até atingir o córtex auditivo primário que distribui a informação para as áreas associativas, para formar uma percepção.

A via auditiva é longa e ainda há muito a ser compreendido sobre o seu funcionamento. Seu circuito auditivo é capaz de transformar uma onda sonora, puramente mecânica, em uma percepção auditiva, respondendo a um código de localização sonoro, permitindo ao ser humano interagir com o meio em que se encontra.

Muito ainda precisa ser feito para se entender como o cérebro processa a música e existem muitos trabalhos sendo desenvolvidos com este intuito. Devido à plasticidade cerebral a música pode gerar alterações no cérebro que acredita-se que sejam computáveis para diferenciar músicos e não-músicos.

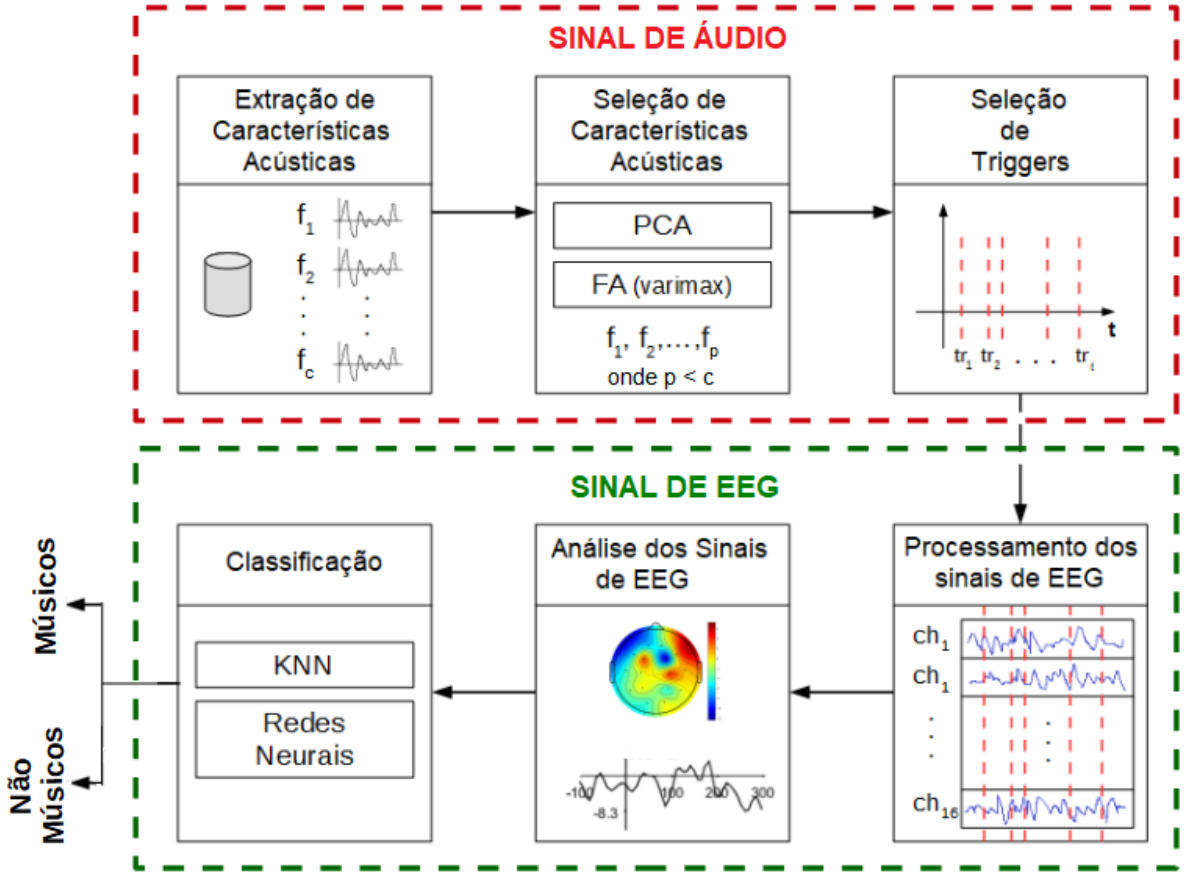
3 METODOLOGIA

Este capítulo descreve em detalhes a metodologia proposta e implementada para a realização deste trabalho.

A metodologia está dividida em 6 etapas, conforme a Figura 16, que apresenta o fluxograma das etapas propostas e implementadas.

Na Etapa 1, são descritas e extraídas c características acústicas que descrevem os sinais de áudio selecionados neste trabalho, utilizando a MIRtoolbox (versão 1.6.1) (LARTILLOT, 2014). Para reduzir a redundância destes dados, são selecionadas, na Etapa 2, por meio da Análise Fatorial (FA) (JOHNSON; WICHERN, 2007), as p características acústicas mais significativas (maior carga fatorial), onde $p \leq c$.

Figura 16 – Fluxograma da metodologia de classificação de sinais de EEG



Fonte: Autor.

Na Etapa 3 são determinados os *triggers* (POIKONEN et al., 2016), que são instantes no tempo da música onde há um alto-contraste nas p características acústicas selecionadas. Estes instantes são considerados como estímulos para o processamento do sinal de EEG na Etapa 4, onde os sinais são filtrados (1-30 Hz), demarcados por uma janela de tempo de 200 ms pré-estímulo, ou seja, 200 ms antes do instante em que os *triggers* ocorreram, até 800 ms

pós-estímulo e, por fim, é realizada uma média dentro destas janelas, resultando em um único sinal de 1000 ms de duração, para cada característica selecionada.

Deste sinal médio são gerados os mapas topográficos corticais dos grupos de voluntários, na Etapa 5, utilizando a toolbox ERPlab (LUCK, 2014), onde é feito uma análise dos sinais médios apresentados pelo grupo de músicos e não-músicos e, por fim, na Etapa 6 é realizada a classificação dos sinais de EEG de cada voluntário, utilizando os classificadores k-NN e Rede Neural.

3.1 EXTRAÇÃO DE CARACTERÍSTICAS ACÚSTICAS

A música possui uma natureza multidimensional, expressando uma série de características ao longo do tempo que variam em níveis de abstração. Perceber e processar estas características envolve um processo de segregação automática da informação musical no cérebro que deve processá-las em paralelo (ALLURI et al., 2012).

Os sinais de áudio possuem características que podem descrever o timbre, o ritmo e a tonalidade de uma música. Portanto, em vez de analisar o sinal de áudio como um todo, serão analisadas neste trabalho as características acústicas que o compõem e que o descrevem quantitativamente, para poder utilizá-las como métricas nos experimentos de EEG.

A MIRtoolbox (versão 1.6.1) (LARTILLOT, 2014) foi utilizada para processar as músicas selecionadas como estímulo neste trabalho. Esta biblioteca computacional fornece ferramentas para a extração de características acústicas de baixo nível, assim como ferramentas para análise tonal, estrutural e propriedades temporais de músicas.

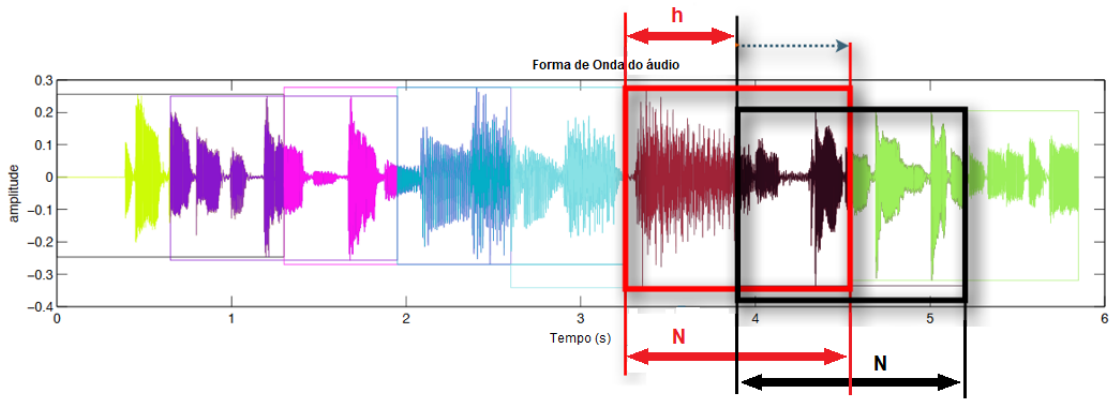
Segundo Kness e Schedl (2016), a resolução temporal do ouvido humano é de aproximadamente 10 milissegundos (ms), ou seja, o ouvido humano é capaz de perceber alterações de estímulos auditivos que ocorram em intervalos de tempo tão pequenos quanto 10 ms, portanto, não é necessário analisar o áudio em intervalos de tempo menores que este. Sendo assim, nesta etapa é realizado um janelamento do sinal de áudio, para que sejam calculadas características acústicas que sejam mais significativas para a percepção humana.

O comprimento do janelamento do sinal de áudio definido segue o padrão utilizado em pesquisas sobre Recuperação de Informação da Música, ou *Music Information Retrieval* (MIR), onde características de curta duração possuem janelas na ordem de milissegundos e características de longa duração possuem janelas na ordem de segundos, dependendo da aplicação (LERCH, 2012). Para evitar perda de informação devido ao janelamento é comum utilizar uma sobreposição no janelamento.

A Figura 17 mostra, de forma geral, a decomposição do sinal de áudio em janelas de comprimento l (em segundos), possuindo N amostras dentro de cada janela, com fator de sobreposição h.

As características acústicas podem ser denominadas como sendo de baixo nível (curta duração), ou de alto nível (longa duração).

Figura 17 – Decomposição em janelas da forma de onda de um sinal de áudio



Fonte: Adaptado de Lartillot, 2014, p.28.

As características baixo nível, são tipicamente calculadas diretamente do sinal bruto da forma de onda do áudio, apresentando normalmente descrições estatísticas da forma de onda (KNESS; SCHEDL, 2016), sem possuir necessariamente uma interpretação direta a percepção humana, podendo ser utilizadas para a construção descritores de características de alto nível (LERCH, 2012). Elas estão normalmente relacionadas com propriedades do timbre de um áudio (ALLURI et al., 2012), podendo apresentar outros aspectos como a tonalidade, e elas podem descrever propriedades estatísticas do sinal, a forma do espectro, a intensidade ou amplitude do sinal (LERCH, 2012).

As características de alto nível, descrevem a música em termos mais semelhantes aos da percepção humana (KNESS; SCHEDL, 2016), sendo extraídas normalmente em um segundo estágio no processamento do áudio (LERCH, 2012). Elas estão relacionadas com aspectos dependentes do contexto do áudio (ALLURI et al., 2012).

Neste trabalho serão utilizadas características acústicas de baixo nível (POIKONEN et al., 2016; ALLURI et al., 2012), que serão descritas a seguir, como sendo métricas extraídas diretamente do sinal de áudio, não necessariamente apresentando propriedades capazes de serem diretamente interpretadas à nível de percepção auditiva (LERCH, 2012), porém estas alterações que ocorrem nelas são capazes de gerar respostas neurais significativas (POIKONEN et al., 2016).

Tais características podem ser extraídas diretamente no domínio do tempo ou no domínio da frequência. Neste trabalho, o sinal de áudio é decomposto utilizando janelas de 50 ms, ou seja, $l = 0{,}05s$, com fator de sobreposição $h = 50\%$, e dentro de cada janela são aplicadas as métricas para a sua extração.

As características extraídas no domínio da frequência são descritas por:

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-\frac{2j\pi}{N}kn}, \tag{4}$$

sendo $k = 0, \ldots, N-1$, $X(k)$ é a Transformada Discreta de Fourier (DFT) do sinal $x(n)$ (LARTILLOT, 2014), de duração igual a N, representando um conjunto de frequências discretas no intervalo de 0 a 2π.

A DFT é implementada na MIRtoolbox através da Transformada Rápida de Fourier (FFT), que é um método computacionalmente mais eficiente de implementação da DFT (LARTILLOT, 2014), onde são utilizados os valores absolutos encontrados para a FFT.

Ao todo foram extraídas 12 características acústicas de baixo nível, sendo duas delas extraídas no domínio do tempo e as outras no domínio da frequência. Uma descrição detalhada de cada uma delas é apresentada nas subseções sequentes.

3.1.1 Root Mean Square Energy

O *Root Mean Square Energy* (RMS) é a primeira característica acústica apresentada extraída no domínio do tempo, sendo uma medida da energia do sinal de áudio, relacionada com a intensidade do som percebido (LERCH, 2012). De acordo com Kness e Schedl (2016), o RMS pode ser utilizado como um estimador de sonoridade e como um indicador de novos eventos em segmentação de áudio.

O valor do RMS pode ser extraído a partir do sinal de áudio $x(n)$, de acordo com a Equação (5), com uma janela de comprimento N, tal que:

$$RMS = \sqrt{\frac{1}{N} \sum_{n=0}^{\frac{N}{2}} x(n)^2}. \tag{5}$$

3.1.2 Zero Crossing Rate

O *Zero Crossing Rate* (ZCR) é a segunda característica apresentada extraída no domínio do tempo, sendo uma medida do número de mudanças do sinal no valor da amplitude do áudio (KNESS; SCHEDL, 2016), ou seja, registra o número de vezes que o sinal de áudio cruza o eixo de amplitude de positivo para negativo e vice-versa, sendo utilizado para a detecção de sons de percussão e de ruídos.

O valor de ZCR pode ser extraído a partir do sinal de áudio $x(n)$, de acordo com a Equação:

$$ZCR = \frac{1}{2N} \sum_{n=0}^{\frac{N}{2}} \left| sign(x(n)) - sign(x(n-1)) \right|, \tag{6}$$

onde a função $sign$ é definida como:

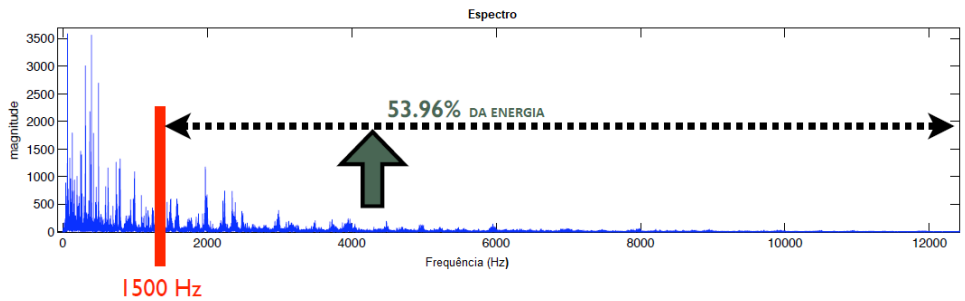
$$sign(x(n)) = \begin{cases} 1, & se \quad x(n) >= 0 \\ -1, & se \quad x(n) < 0 \end{cases}, \tag{7}$$

onde $x(n-1) = 0$ na inicialização ou quando tal valor não existe.

3.1.3 Brightness

O *Brightness* é a primeira característica apresentada extraída no domínio da frequência, sendo definido como a porcentagem da quantidade de energia acima de uma determinada frequência de corte (ex.: 1500Hz) assim como mostra a Figura 18, apresentando valores entre 0 e 1. Quanto maior o seu valor, maior a energia do sinal e maior a intensidade do brilho na janela em que é calculado (LARTILLOT, 2014).

Figura 18 – Frequência de corte para a determinação do *Brigtness*



Fonte: Adaptado de Lartillot, 2014, p.127.

3.1.4 Spectral Centroid

O *Spectral Centroid* (SC) representa o centro de gravidade do espectro do sinal, i.e., a banda de frequência onde a maior parte da energia está concentrada, sendo utilizada como uma medida de brilho e nitidez do áudio (KNESS; SCHEDL, 2016; LERCH, 2012).

O valor de SC é definido a partir da Equação (8), onde valores baixos indicam componentes de baixa frequência e, por consequência, baixo brilho.

$$SC = \frac{\sum_{k=0}^{\frac{N}{2}} kX(k)}{\sum_{k=0}^{N-1} X(k)}, \tag{8}$$

onde $X(k)$ é a FFT do sinal de áudio $x(n)$, conforme descrito na Equação 4.

3.1.5 Spectral Spread

O *Spectral Spread* (SS), também conhecido como *Bandwidth* (largura de banda), deriva do *Spectral Centroide*, e descreve a concentração da magnitude do espectro ao redor do SC, podendo ser interpretado como um desvio padrão da magnitude do espectro ao redor do *Spectral Centroide* (LERCH, 2012; KNESS; SCHEDL, 2016).

Em regiões de silêncio, onde não há energia espectral, o valor do SS é zero, assim como ocorre no SC, e em regiões onde há ruído branco [2], o valor do SS é alto. Matematicamente esta métrica pode ser calculada por:

$$SS = \sqrt{\frac{\sum_{k=0}^{\frac{N}{2}} |k - SC| \, X(k)}{\sum_{k-0}^{\frac{N}{2}} X(k)}}, \tag{9}$$

onde $X(k)$ é a FFT do sinal de áudio $x(n)$ e SC é o *Spectral Centroid* calculado pela equação anterior (8).

3.1.6 Spectral Flatness

O *Spectral Flatness* (SF) é uma medida do ruído em oposição à tonalidade. Valores baixos mostram um não achatamento do espectro do sinal (próximo de zero para tons puros), indicando que o espectro é formado por uma mistura de ondas senóides, enquanto que valores altos, indicam um achatamento do espectro, (próximo de 1 para ruído branco), indicando ruído (LERCH, 2012).

O SF é dado pela razão entre a média geométrica e a média aritmética da magnitude do espectro (LARTILLOT, 2014), dada por:

$$SF = \frac{\sqrt[N]{\prod_{k=0}^{\frac{N}{2}} |X(k)|}}{\frac{1}{N} \sum_{k=0}^{\frac{N}{2}} |X(k)|}, \tag{10}$$

onde $X(k)$ é a FFT do sinal de áudio $x(n)$.

3.1.7 Spectral Skewness

O *Spectral Skewness* (Skw) é uma métrica que mede a assimetria da distribuição da magnitude do espectro ao redor da sua média aritmética, calculado a partir do terceiro momento central (LERCH, 2012; LARTILLOT, 2014), tal que:

$$Skw = \frac{\frac{1}{N} \sum_{k=0}^{\frac{N}{2}} (X(k) - \bar{X})^3}{(\sqrt{\frac{1}{N} \sum_{k=0}^{\frac{N}{2}} (X(k) - \bar{X})^2})^3}, \tag{11}$$

onde $\bar{X}$ é definido como:

$$\bar{X} = \frac{1}{N} \sum_{k=0}^{N} X(k), \tag{12}$$

onde $X(k)$ é a FFT do sinal de áudio $x(n)$.

[2] O ruído branco é gerado por um conjunto de frequências com intensidade igualmente distribuídas ao longo de todo o range de audição (PIGEON, 2013).

Quando o valor de $|X(k)| = 0$, há silêncio no sinal. Quando o valor de $|X(k)| = const$, há ruído branco no sinal.

3.1.8 Spectral Kurtosis

O *Spectral Kurtosis* (Kur) mede se a forma da distribuição do valor da magnitude espectral é semelhante a uma distribuição Gaussiana ou não, podendo indicar transientes no sinal de áudio, calculando o quarto momento central (LERCH, 2012; LARTILLOT, 2014), tal que:

$$Kur = \frac{\frac{1}{N}\sum_{k=0}^{\frac{N}{2}}(X(k) - \bar{X})^4}{(\frac{1}{N}\sum_{k=0}^{\frac{N}{2}}(X(k) - \bar{X})^2)^2} - 3, \tag{13}$$

onde $\bar{X}$ é definido pela Equação (12) e $X(k)$ é a FFT do sinal de áudio $x(n)$.

3.1.9 Spectral Rolloff

O *Spectral Rolloff* (SRO) é definido como a frequência abaixo da qual a magnitude acumulada de $X(k)$ atinge uma determinada porcentagem β sobre a soma das magnitudes. Em outras palavras, o SRO é a frequência abaixo da qual a porcentagem β (ex.: $\beta = 85\%$) da magnitude da distribuição do espectro está concentrada (LARTILLOT, 2014), sendo

$$SRO = M, \tag{14}$$

tal que:

$$\sum_{k=0}^{\frac{M}{2}} |X(k)| = \beta \sum_{k=0}^{\frac{N}{2}} |X(k)|, \tag{15}$$

onde $X(k)$ é a FFT do sinal de áudio $x(n)$. Se M for o maior valor na qual esta equação é satisfeita, então essa frequência M é dita como *rolloff* (SUBRAMANIAN; RAO; ROY, 2004).

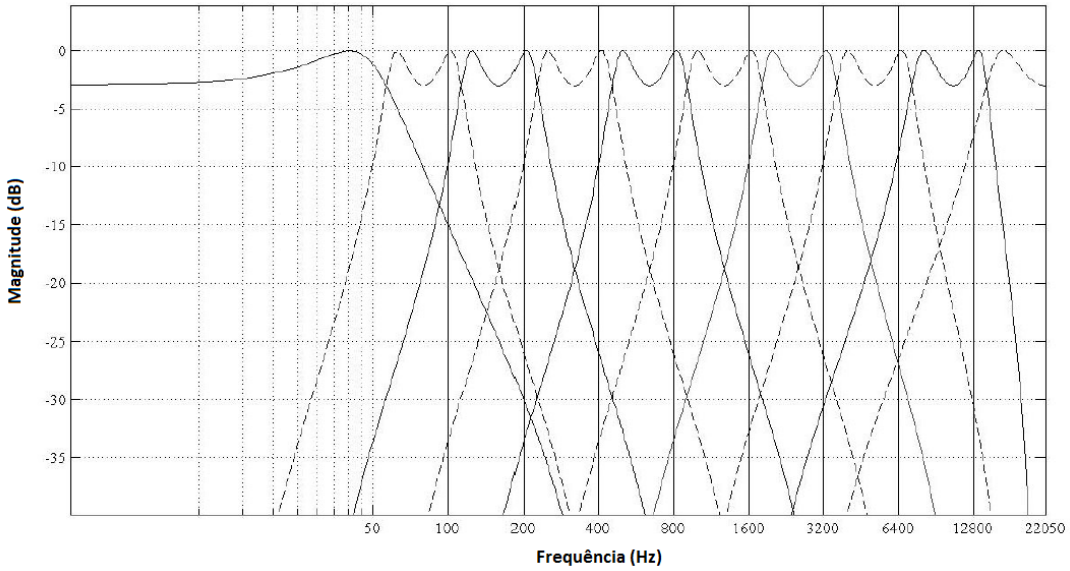
3.1.10 Spectral Flux

O *Spectral Flux* (SFlux) descreve a mudança na magnitude do espectro de frequência do áudio entre janelas consecutivas (KNESS; SCHEDL, 2016), através da Distância Euclidiana, de acordo com a Equação (16):

$$SFlux = \sum_{k=0}^{\frac{N}{2}} (|X(w,k) - X(w-1,k)|)^2, \tag{16}$$

onde $X(w,k)$ é a FFT da janela w e $X(w-1,k)$ é a FFT da janela $w-1$ do sinal de áudio $x(n)$.

Figura 19 – Resposta do filtro gerador das 10 Sub-bandas do Spectral Flux



Fonte: Alluri e Toiviainen (2010).

De acordo com Monteiro e Manzolli (2011), espera-se que valores de SFlux sejam baixos para sinais com envelope espectral estável e apresenta em picos para sinais onde o envelope espectral possui alta taxa de variação.

O SFlux pode ser dividido em Sub-bandas, onde mede-se a flutuação da frequência contida em 10 oitavas[3] musicais, escaladas em sub-bandas do espectro (LARTILLOT, 2014). Este novo conjunto de características foi introduzido por Alluri e Toiviainen (2010) que utilizou um banco de filtros elípticos de segunda ordem de 10 canais, cuja resposta é apresentada na Figura 19. As Sub-bandas estão delimitadas de acordo com a Tabela 1.

Tabela 1 – Intervalo de Frequências das Sub-Bandas do *Spectral Flux*

Sub-Banda	Intervalo de Frequência
Sub-Banda No. 1	0 - 50 Hz
Sub-Banda No. 2	50 - 100 Hz
Sub-Banda No. 3	100 - 200 Hz
Sub-Banda No. 4	200 - 400 Hz
Sub-Banda No. 5	400 - 800 Hz
Sub-Banda No. 6	800 - 1600 Hz
Sub-Banda No. 7	1600 - 3200 Hz
Sub-Banda No. 8	3200 - 6400 Hz
Sub-Banda No. 9	6400 - 12800 Hz
Sub-Banda No. 10	12800 - 22050 Hz

Fonte: Alluri e Toiviainen (2010).

[3] Uma oitava musical refere-se ao intervalo entre uma nota músical e outra apresentando uma razão de frequência de 2:1.

Alluri e Toiviainen (2010) correlacionaram as características extraídas do áudio com a percepção de timbre dos voluntários, associando as características com adjetivos que expressam alguma qualidade do timbre como, por exemplo, o brilho do áudio, a riqueza do áudio, etc.

Desta forma, as Sub-bandas No.2 e No.3 estão relacionadas com a percepção de riqueza (*Fullness*) do áudio, associadas às flutuações nas baixas frequências do espectro, enquanto que as Sub-bandas No.7 e No.8 estão relacionadas com a percepção de movimento (*Activity*) do áudio, associadas às flutuações nas altas frequências do espectro (ALLURI; TOIVIAINEN, 2010) (LARTILLOT, 2014).

3.1.11 Spectral Roughness

O *Spectral Roughness* (SR) é uma característica extraída no domínio da frequência que introduz o conceito de consonância e dissonância foi introduzido por Helmholtz, onde se dois tons puros soam em frequências muito próximas, então uma onda distinta aparece devido à interferência entre estes dois tons (SETHARES, 2005). Quando a diferença entre as duas frequências aumenta, a sensação de aspereza do som diminui, tornando-se mais agradável, mais consonante, enquanto que para pequenas diferenças entre as duas frequências ocorre o contrário, onde o som se torna mais áspero, mais dissonante (PLOMP; LEVELT, 1965).

Segundo Plomp e Levelt (1965), Helmholtz propôs que a dissonância atinja um nível máximo quando a diferença de frequência entre dois tons é de aproximadamente 30-40Hz. Segundo Sethares (2005), quando a diferença de frequência entre dois tons é de aproximadamente 32Hz, ocorre a dissonância.

Baseando-se nisso, Lartillot (2014) propuseram uma medida da dissonância, ao calcular os picos do espectro e então fazer a média dos pares de picos em que ocorre dissonância.

3.1.12 Spectral Entropy

Por fim, o *Spectral Entropy* (SE) indica se o espectro contém picos predominantes ou não, por exemplo, um tom puro possui mínima entropia e máximo ruído branco (ALLURI; TOIVIAINEN, 2010). De acordo com Lartillot (2014), se uma curva é extremamente achatada, correspondendo a uma situação de máxima incerteza, então a entropia é máxima. Esta métrica pode ser calculada pela equação:

$$SE = \frac{\sum_{k=0}^{\frac{N}{2}} X(k) \cdot logX(k)}{log\frac{N}{2}}, \tag{17}$$

onde $X(k)$ é a FFT do sinal de áudio $x(n)$.

3.2 SELEÇÃO DE CARACTERÍSTICAS ACÚSTICAS

As características acústicas extraídas dos áudios resultam em uma série temporal amostrada de acordo com o número de janelas existentes para cada música. Sendo assim, cada característica acústica de uma determinada música resultará em um vetor contendo w janelas, definidas de acordo com a Equação 18.

$$w = \frac{2t - hl}{l}, \tag{18}$$

onde t é o tempo de duração do áudio e l é o tamanho da janela, ambos em segundos e h é o fator de sobreposição.

A partir destas características extraídas do áudio, a redução de dimensionalidade é realizada por meio da Análise Fatorial (FA), utilizando a solução apresentada pela Análise de Componentes Principais (PCA), por meio das quais são selecionadas as características mais expressivas (de acordo com a ordem decrescente de variância) para representar o conjunto de áudios escolhidos para análise.

As características acústicas extraídas podem ser codificadas como uma matriz de dados $\mathbf{M}$ de dimensão $w \times c$, onde cada coluna representa uma característica acústica extraída do áudio e cada linha representa o valor encontrado para cada janela, ou seja:

$$\mathbf{M} = \begin{bmatrix} \mathbf{M_1} & \mathbf{M_2} & \dots & \mathbf{M_c} \end{bmatrix} = \begin{bmatrix} M_{1,1} & M_{1,2} & \dots & M_{1,c} \\ M_{2,1} & M_{2,2} & \dots & M_{2,c} \\ \vdots & \vdots & \vdots & \vdots \\ M_{w,1} & M_{w,2} & \dots & M_{w,c} \end{bmatrix} \tag{19}$$

A Análise Fatorial foi utilizada para descrever a associação entre as características acústicas extraídas do áudio. O FA tem por objetivo analisar a correlação entre as variáveis originais, estimando variáveis que não são observáveis diretamente, os fatores, que produzem o comportamento observado nos dados (JOHNSON; WICHERN, 2007; HAIR et al., 2014).

Para isso, primeiramente é realizado o PCA, a partir da matriz de dados normalizada $\mathbf{Z}$ possuindo média 0 e variância 1, selecionando as p componentes principais ($p \leq c$) que apresentam autovalores maiores que 1 (JOHNSON; WICHERN, 2007).

O FA é realizado utilizando o número p de fatores determinados pelo PCA, utilizando rotação *varimax*, onde idealmente espera-se encontrar um padrão de carregamento em que cada variável possua alta carga fatorial em apenas um fator e baixa carga fatorial nos fatores restantes (JOHNSON; WICHERN, 2007).

Sendo assim, são determinados os $\mathbf{F} = [\mathbf{f_1}, \mathbf{f_2}, \dots, \mathbf{f_p}]$ fatores, selecionando a característica acústica que apresente a maior carga fatorial dentro de cada fator, representando a variável mais significativa nele e, desta forma, são selecionadas as p características acústicas mais representativas dentro do conjunto de características extraídas dos áudios.

3.3 SELEÇÃO DE TRIGGERS

Após selecionar as p características mais relevantes para o estudo através da Análise Fatorial, é necessário selecionar os *triggers*, que são instantes no tempo onde há um alto-contraste na série temporal a ser analisada, determinados de acordo com o modelo proposto por Poikonen et al. (2016), que afirma que:

> As rápidas mudanças nas características musicais das músicas deverão extrair componentes sensoriais similares aos revelados nos métodos convencionais de estudo de ERP utilizando estímulos tonais, e a amplitude dos componentes de ERP serão dependentes da magnitude do rápido crescimento (MoRI) nos valores das características individuais, bem como da duração do período de tempo que precede valores de baixas características de maneira similar à amplitude de componentes de ERP tradicionais dependentes das características do estímulo apresentado.

De acordo com este modelo, a Magnitude do Rápido Crescimento, ou *Magnitude of Rapid Increase* (MoRI), define os limiares inferiores (V_{p-}) e superiores (V_{p+}) de cada p característica acústica selecionada anteriormente, correspondentes às regiões de baixo e alto-contraste, respectivamente.

Os limiares são definidos a partir dos valores médios das séries temporais para cada característica acústica, onde V_{p-} é definido como sendo igual a -20% da média e V_{p+} é definido como sendo igual a $+20\%$ da média da característica acústica analisada.

Em músicas não existem períodos de silêncios bem definidos, denominados Intervalo entre Estímulos, ou *Inter-Stimulus Interval* (ISI), como nos estudos convencionais de ERP. Portanto, Poikonen et al. (2016) definiram uma nova variável, denominada Fase Precedente de Baixa Característica ou *Preceding Low-Feature Phase* (PLFP), que corresponde a um determinado período de tempo em que os valores encontrados na série temporal de uma característica acústica não ultrapassam o valor de V_{p-}.

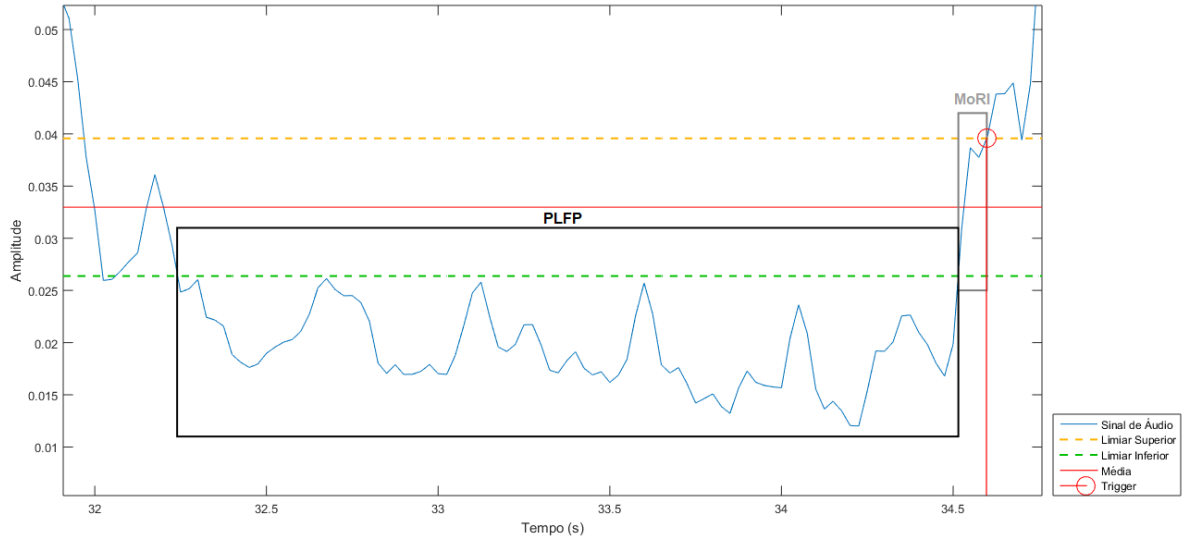
Neste trabalho, o comprimento do PLFP tem duração mínima de 400ms e os *triggers* válidos tiveram uma fase de crescimento (MoRI) em que o valor da característica crescia de V_{p-} a V_{p+} de duração máxima de 100ms.

A Figura 20 ilustra uma situação onde ocorre um *trigger*, a partir das condições estabelecidas, onde o sinal de áudio se mantém abaixo de V_{p-} por um período de tempo (PLFP) superior à 400ms, seguido por uma fase de crescimento (MoRI) inferior à 100ms.

Poikonen et al. (2016) encontraram que os componentes de ERP podem ser extraídos também por uma dinâmica de estímulos naturais, como as peças musicais, e não apenas por sons simples com tempo de silêncio definidos como os estímulos artificiais utilizados nos experimentos convencionais de ERP, sendo necessário, porém, que estes estímulos possuam fortes contrastes para que eles sejam capazes de extrair os componentes de ERP convencionais.

Portanto, um critério para determinar quais características acústicas serão de fato selecionadas é existir *triggers* para a característica analisada em todas as músicas utilizadas no experimento. Caso não exista nenhum trigger, a característica deverá ser substituída por aquela

Figura 20 – Ilustração da ocorrência de um *trigger*, assumindo as condições propostas



Fonte: Autor.

que possui um valor de carga fatorial imediatamente abaixo desta dentro do fator que ela representa.

Outro critério para a seleção das características acústicas é haver um número mínimo de *triggers* (tr), onde $tr \geq 4$. Caso contrário, deve-se escolher a característica que possui carga fatorial imediatamente abaixo dela e que cumpra com todos os critérios propostos.

Os parâmetros utilizados para a seleção de *triggers* podem ser alterados, sendo que quanto mais rígidos eles forem, menor a possibilidade de ocorrência de *triggers* para as músicas. Por outro lado, quanto menos rígidos eles forem, maior a possibilidade de ocorrência de *triggers*. Porém, diminuir a rigidez dos parâmetros significa diminuir o contraste acústico representado por cada *trigger*, podendo afetar as respostas neurais geradas por eles.

3.4 PROCESSAMENTO DO SINAL DE EEG

Todo o sinal adquirido de EEG é inicialmente pré-processado, aplicando um filtro digital Butterworth de ordem 10 Passa Banda de 1 Hz a 30 Hz, para remoção do nível DC e do sinal da rede elétrica, deixando apenas as frequências de interesse para a análise do sinal de EEG (POIKONEN et al., 2016). O filtro foi projetado de acordo com os parâmetros encontrados na Tabela 2.

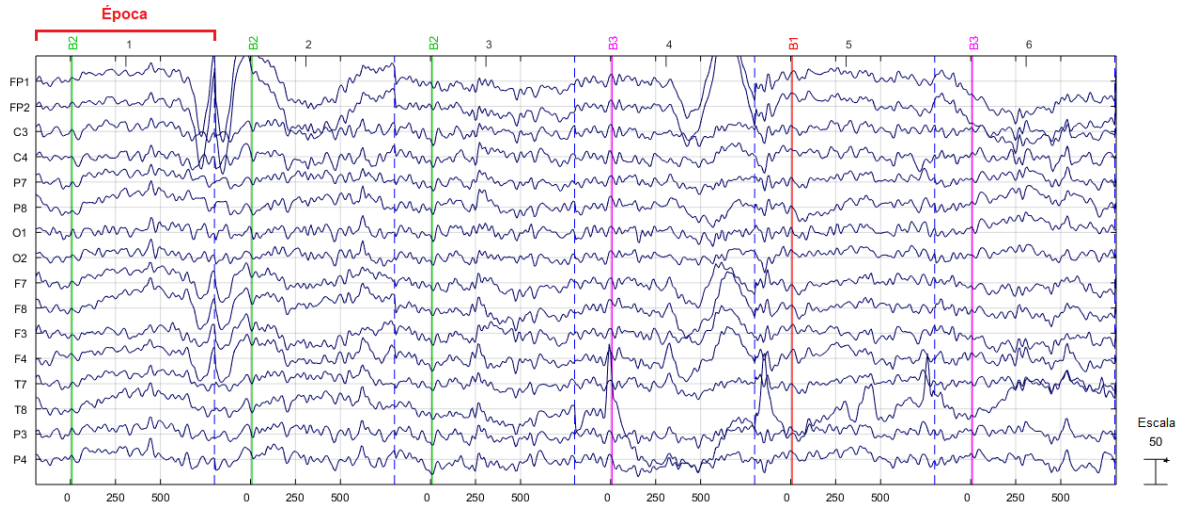
Após a filtragem, o sinal de EEG é pós-processado utilizando a Toolbox EEGLAB (v13.5.4b). Os *triggers* selecionados apresentam os tempos de interesse nas músicas para cada característica acústica selecionada, correspondendo a um evento/estímulo na análise de ERP.

Tabela 2 – Parâmetros do filtro Butterworth

fs	125	Frequência de Amostragem (Hz)
fn	fs/2	Frequência de Nyquist
Wp	[1 30]/fn	Banda de Passagem
Ws	[0.1 50]/fn	Banda de Rejeição
Ap	5	Atenuação na banda de passagem (dB)
As	50	Atenuação na banda de rejeição (dB)

Fonte: Autor.

Figura 21 – Ilustração de um sinal de EEG dividido em 6 épocas, delimitadas em 200ms pré-estímulo e 800ms pós-estímulo, para 3 características acústicas distintas



Fonte: Autor.

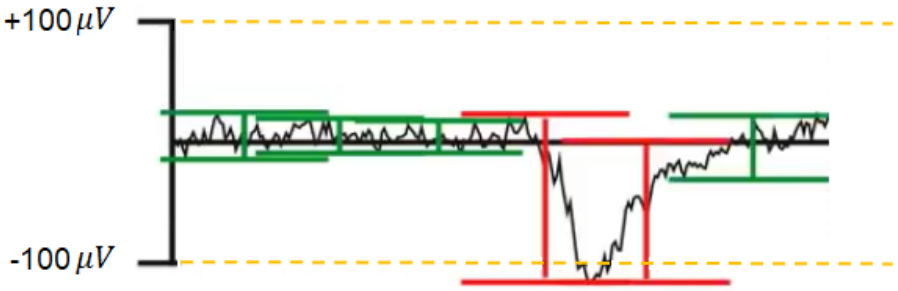
Após serem marcados os eventos no sinal de EEG, é feito um janelamento no sinal, denominado "época", delimitado por 200 ms pré-estímulo, ou seja, 200 ms antes da ocorrência do *trigger*, e 800 ms pós-estímulo, ou seja, 800 ms após a ocorrência do *trigger*.

A Figura 21 apresenta um sinal de EEG contendo 6 épocas, onde o *trigger* acontece no instante 0 ms de cada época, contendo 3 características acústicas distintas denominadas arbitrariamente como *B1, B2* e *B3*, que ocorrem em instantes diferentes no decorrer da música.

Tendo sido determinadas as épocas, é necessário remover os artefatos contidos no sinal de EEG dentro delas, principalmente artefatos referentes ao piscar de olhos, caracterizados por grandes picos no sinal que ocorrem principalmente nas regiões frontais do escalpo.

Para a remoção de artefatos foi utilizado o algoritmo *Moving Window peak-to-preak Threshold*, onde determina-se uma janela para a análise do sinal de EEG, verificando os máximos e mínimos dentro da janela, conforme a Figura 22. Caso o pico ou vale ultrapasse um limiar definido, aquela janela é marcada para remoção e todo o sinal contido na época rejeitada é excluído da análise. O limiar utilizado para a remoção de artefatos é de 100 μV, com uma janela de 200 ms com 50% de sobreposição.

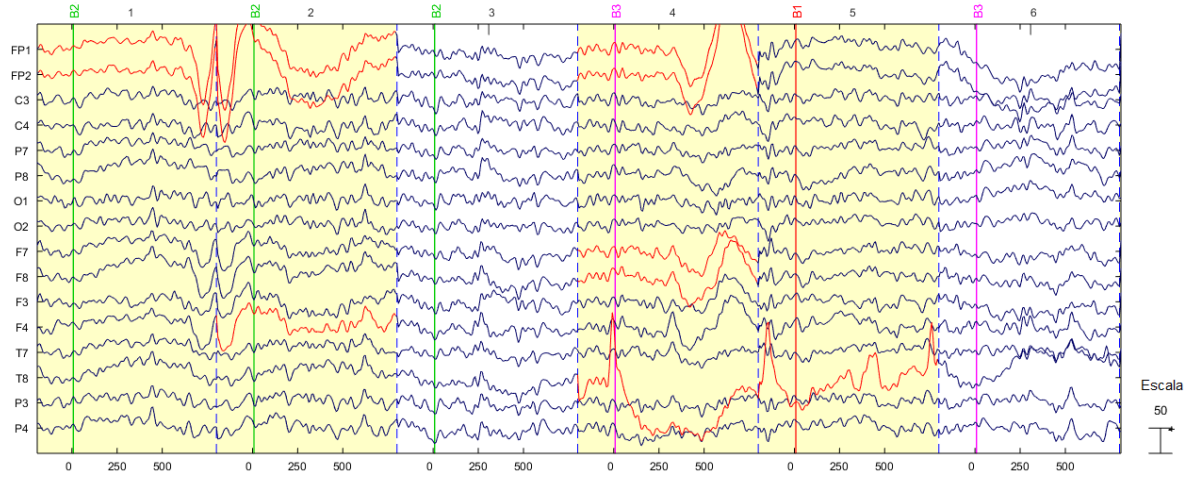
Figura 22 – Ilustração do funcionamento do algoritmo: "Moving Window peak-to-peak Threshold"



Fonte: Adaptado de Luck (2014).

A Figura 23 ilustra um resultado da remoção de artefatos no sinal de EEG, onde o sinal em vermelho está marcado para remoção. É possível perceber, portanto, o quão ruidoso pode ser o sinal de EEG devido a artefatos como o piscar dos olhos, movimentos musculares, entre outros, que devem ser removidos da análise.

Figura 23 – Artefatos marcados para remoção no sinal de EEG



Fonte: Autor.

Portanto, cada característica acústica selecionada gera um conjunto de *triggers* ao longo da música, onde espera-se encontrar uma ativação cortical similar entre cada um deles. Após a remoção de ruídos das épocas é calculada a média destes eventos, evidenciando o que varia em comum nas respostas cerebrais dos voluntários para os estímulos apresentados (LUCK, 2014).

3.5 ANÁLISE DOS SINAIS DE EEG

Para cada p característica acústica de cada música analisada é gerado uma única época de 1000 ms de duração (200 ms pré-estímulo e 800 ms pós-estímulo) correspondente ao sinal médio entre todos os *triggers* de cada característica para cada voluntário.

Dentro desta época média é realizada a análise deste sinal resultante, onde espera-se observar os componentes N100, caracterizado por um vale próximo de 100 ms (entre 80 – 150 ms) pós-estímulo, e o componente P200, caracterizado por um pico próximo de 200 ms (entre 150 – 275 ms) pós-estímulo (POIKONEN et al., 2016).

O componente N100 está relacionado com o estado de atenção do sujeito e é extraído por qualquer estímulo imprevisto, na ausência de uma tarefa específica a ser realizada. Sua ocorrência é em regiões centrais e frontais do escalpo. O componente P200 está relacionado com a comparação dos estímulos auditivos e apresenta alguns aspectos relacionados com o processamento da informação auditiva e a atenção. Sua ocorrência é em regiões centrais e frontais do escalpo, assim como regiões parietais e occipitais do escalpo.

Utilizando o sinal contido em cada época é possível gerar os mapas topográficos, que mostram a atividade neural do córtex como um todo, de acordo com o potencial elétrico registrado em cada eletrodo nos instantes desejados.

Desta forma, é analisado a ativação cortical dos dois grupos de voluntários, a partir do sinal médio dos voluntários músicos e não-músicos, comparando visualmente os sinais de EEG resultantes da região central do escalpo e os mapas topográficos dos dois grupos para as músicas apresentadas nos experimentos, verificando a existência dos componentes de ERP N100 e P200 e as diferenças entre grupos.

3.6 CLASSIFICAÇÃO

Para a formação do vetor de entrada do classificador são levados em consideração dois cenários, que variam de acordo com a latência observadas, utilizando os valores registrados em cada canal de eletrodos nas épocas médias de cada característica acústica, para cada voluntário.

No primeiro cenário são utilizados os valores do potencial elétrico medidos em todos os eletrodos nos instantes 100 ms e 200 ms pós-estímulo para cada característica acústica.

No segundo cenário, é observado na região central do escalpo (média entre os sinais de EEG registrados em C3 e C4) os instantes em que os pontos mínimos ocorrem para cada voluntário dentro de uma janela de tempo de 80-150 ms pós-estímulo, referentes a componente N100, e os instantes em que os pontos máximos ocorrem para cada voluntário entre 150-275 ms pós-estímulo, referentes a componente P200 (VUSST et al., 2014). Após serem determinados estes instantes, são realizados os registros dos valores do potencial elétrico medidos em todos os eletrodos para cada característica acústica. Desta forma, cada voluntário será representado

pelos valores do potencial elétrico registrado nos eletrodos em instantes no tempo que não são, necessariamente, iguais entre si.

Portanto, cada voluntário será representado por um vetor de entrada composto pela concatenação dos valores registrados para todos os eletrodos Ch, para cada característica acústica selecionada fp, para os dois instantes analisados. Portanto, a dimensionalidade deste vetor é dada por $d = fp \times Ch \times 2$.

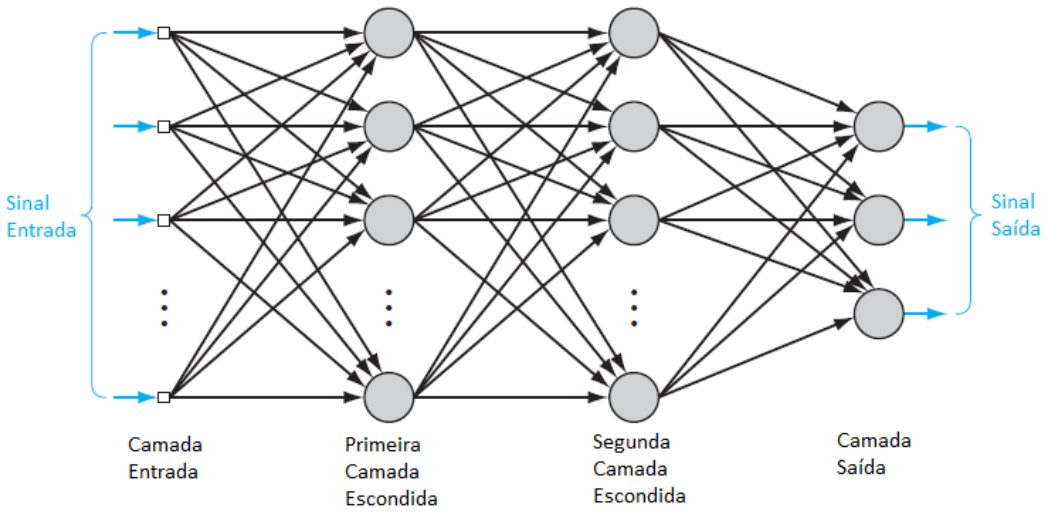
Este vetor de entrada será utilizado para a classificação dos dados obtidos em Músicos e Não-músicos por meio das respostas neurais induzidas pelas músicas apresentadas aos voluntários. Os dados obtidos para cada voluntário são agrupados em uma matriz $\mathbf{S_{dx26}}$ onde as 26 colunas se referem ao número de voluntários, ou seja:

$$\mathbf{S} = \begin{bmatrix} \mathbf{S_1} & \mathbf{S_2} & \dots & \mathbf{S_{26}} \end{bmatrix} = \begin{bmatrix} S_{1,1} & S_{1,2} & \dots & S_{1,26} \\ S_{2,1} & S_{2,2} & \dots & S_{2,26} \\ \vdots & \vdots & \vdots & \vdots \\ S_{d,1} & S_{d,2} & \dots & S_{d,26} \end{bmatrix} \tag{20}$$

Os métodos de classificação utilizados possuem aprendizado supervisionado, sendo necessário selecionar parte dos dados para o treinamento e parte para teste e validação dos dados, utilizando o X-*fold Cross-Validation*. Os métodos de classificação utilizados são o k-NN (k Nearest Neighbors) e a Rede Neural *MultiLayer Perceptron* (MLP).

As redes neurais são modelos computacionais inspirados no funcionamento do sistema nervoso. A rede neural MLP é uma rede que possui uma ou mais camadas escondidas entre as camadas de entrada e saída, contendo um alto nível de conectividade, relacionado com os pesos sinápticos da rede, e possui uma função de ativação não-linear.

Figura 24 – Arquitetura da rede *MultiLayer Perceptron* com 2 camadas escondidas



Fonte: Adaptado de Haykin, 2001, p.225.

A Figura 24 apresenta a ilustração de uma arquitetura gráfica da rede MLP com 2 camadas escondidas. Cada nó da rede corresponde a um neurônio e os neurônios de cada camada estão conectados com todos os neurônios da camada posterior.

O sinal de entrada apresentado é propagado pela rede neurônio por neurônio. O sinal que chega a cada neurônio da rede é calculado como uma função das entradas associadas aos pesos sinápticos. Os neurônios de saída correspondem à camada de saída da rede. Entre a camada de entrada e saída estão as camadas escondidas, podendo haver uma ou mais camadas escondidas, que atuam como detectores de características (HAYKIN, 2001).

O k-NN é uma técnica de classificação não-paramétrica, onde não se considera a informação da distribuição da população, mas os próprios dados definem sua estrutura e são capazes de encontrar estas estruturas explícita ou implicitamente (LIMA, 2012). Neste método define-se um conjunto de treinamento, onde se sabe a classificação dos dados, e então é definida uma métrica para calcular a distância entre os dados do conjunto de treinamento e os dados a serem classificados, sendo a Distância Euclidiana a métrica mais comum aplicada.

É definido o valor de k (número de vizinhos mais próximos), onde os dados são classificados de acordo com o voto majoritário dos k vizinhos mais próximos. O processo é repetido, variando o número de k, sendo selecionado como modelo final o número de vizinho que apresentam menor erro na classificação (LIMA, 2012).

Os dois classificadores são utilizados para cada cenário proposto, realizando uma análise e comparação entre os resultados obtidos pelos dois.

4 EXPERIMENTOS E RESULTADOS

Este capítulo descreve os experimentos e resultados realizados neste trabalho. Basicamente, as quatro primeiras seções se referem aos experimentos e três últimas aos resultados obtidos.

4.1 AQUISIÇÃO DOS SINAIS DE EEG

Para realizar a aquisição dos sinais de EEG, foi utilizada a plataforma OpenBCI, possuindo uma taxa de aquisição de 250Hz para 8 eletrodos e 125Hz para 16 eletrodos, com resolução de 32bits por canal, utilizando eletrodos a seco de Cloreto de Prata (AgCl), posicionados no escalpo de acordo com o sistema 10-20, utilizando a placa Cyton (8 canais de eletrodos) e o módulo Daisy (8 canais de eletrodos sobressalentes) para realizar o registro e o módulo USB para comunicação com o computador, conforme a Figura 25 (OPENBCI, 2016).

Figura 25 – Placa Cyton e módulo Daisy conectados e módulo USB

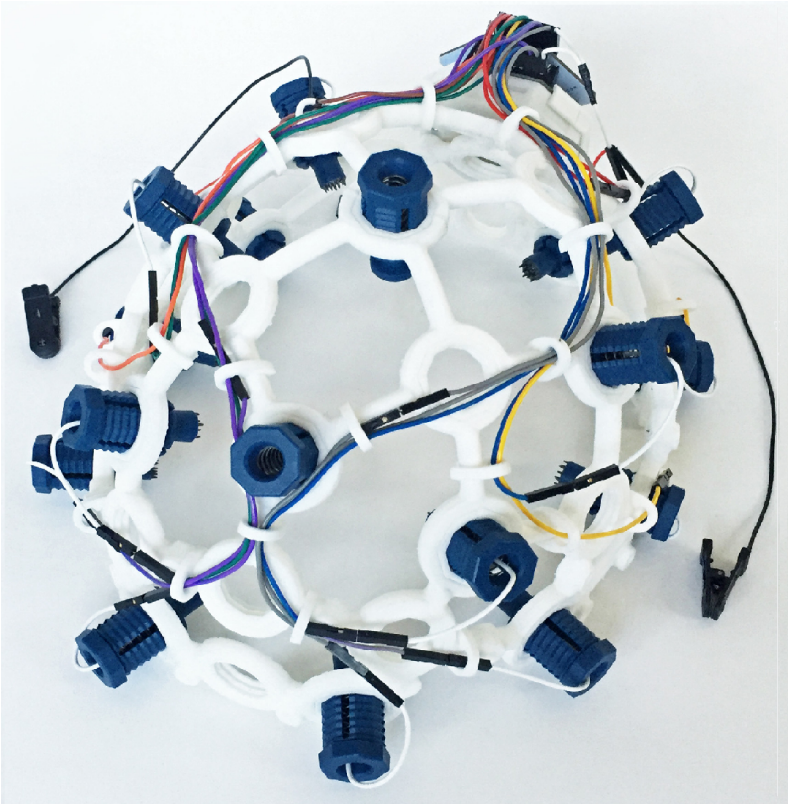
Fonte: Adaptado de OpenBCI, 2017.

Os eletrodos são conectados na placa Cyton e no módulo Daisy de acordo com o código de cores proposto pelo fabricante, conforme a Figura 26. Os eletrodos de referência são conectados nos pinos BIAS e RSB da placa Cyton, sendo necessário conectar em paralelo os pinos RSB da placa Cyton e do módulo Daisy para iniciar o registro dos sinais de EEG.

Figura 26 – Conexão dos eletrodos na Placa Cyton e módulo Daisy

Fonte: Autor.

Figura 27 – Ultra Cortex IV com Placa Cyton e módulo Daisy conectados

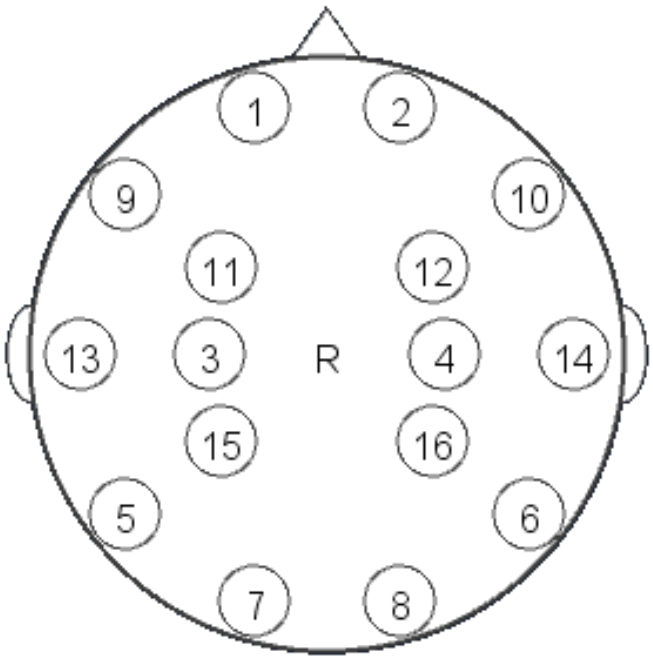


Fonte: OpenBCI, 2017.

Os eletrodos ficam fixados no Cortex IV, conforme a Figura 27, onde a placa Cyton e o módulo Daisy se comunicam com o módulo USB via Bluetooth, evitando fios pendurados,

trazendo maior praticidade e conforto para a realização dos experimentos, utilizando eletrodos a seco para a captura do sinal de EEG.

Figura 28 – Posicionamento dos eletrodos de EEG de acordo com a plataforma OpenBCI para 16 eletrodos



Fonte: OpenBCI, 2017.

Sendo assim, os 16 eletrodos utilizados neste experimento são dispostos no escalpo de acordo com a Figura 28.

4.2 PARTICIPANTES

Para a realização dos experimentos foram utilizados participantes voluntários, que não possuíam histórico de desordens neurológicas ou psicológicas. As pessoas que não possuíam conhecimento musical, ou seja, que não tiveram nenhum treinamento musical formal e nunca aprenderam a tocar nenhum instrumento, foram categorizados como *Não-Músicos*. As pessoas com conhecimento musical, que possuíam no mínimo 2 anos de treinamento musical formal, foram categorizadas como *Músicos*.

Foram utilizados 30 voluntários para os experimentos, sendo metade músicos e a outra metade não-músicos. Destes 30 voluntários, 4 foram eliminados da análise por problemas durante a aquisição de dados, tais como, o voluntário apresentar sonolência durante o experimento, ou o sinal não satisfazer os critérios de processamento descritos no capítulo anterior.

Tabela 3 – Relação de Voluntários

ID	Sexo	Idade	Destro/Canhoto	Instrumento	Anos de prática	Horas de prática semanal
NM1	Masculino	27 anos	Destro	-	-	-
NM2	Masculino	46 anos	Canhoto	-	-	-
NM3	Masculino	40 anos	Canhoto	-	-	-
NM4	Masculino	30 anos	Destro	-	-	-
NM5	Masculino	25 anos	Destro	-	-	-
NM6	Feminino	25 anos	Destra	-	-	-
NM7	Masculino	42 anos	Canhoto	-	-	-
NM8	Masculino	24 anos	Destro	-	-	-
NM9	Masculino	28 anos	Destro	-	-	-
NM10	Masculino	25 anos	Canhoto	-	-	-
NM11	Masculino	25 anos	Destro	-	-	-
NM12	Feminino	35 anos	Destra	-	-	-
NM13	Feminino	34 anos	Destra	-	-	-
M1	Masculino	26 anos	Destro	Bateria	10 anos	-
M2	Masculino	18 anos	Destro	Flauta transversal	7 anos	4hrs
M3	Feminino	23 anos	Destro	Contrabaixo acústico	13 anos	4hrs
M4	Feminino	45 anos	Destro	Saxofone	28 anos	4hrs
M5	Feminino	25 anos	Destro	Violino	14 anos	4hrs
M6	Masculino	26 anos	Destro	Bateria	20 anos	8hrs
M7	Masculino	25 anos	Destro	Piano	21 anos	35hrs
M8	Masculino	31 anos	Destro	Violão	17 anos	-
M9	Masculino	26 anos	Destro	Guitarra	10 anos	-
M10	Feminino	22 anos	Destro	Violão	7 anos	8hrs
M11	Feminino	28 anos	Destro	Canto/Violino	13 anos	-
M12	Feminino	28 anos	Destro	Violino	18 anos	4hrs
M13	Masculino	26 anos	Destro	Saxofone	13 anos	4hrs

Fonte: Autor.

Portanto, para o experimento foram considerados então 26 voluntário (17 homens e 9 mulheres), 13 não-músicos (31,2±7,4 anos) e 13 músicos (26,8±6,3 anos) que praticam seus instrumentos há, em média, 14 anos, conforme descrito na Tabela 3.

Como é possível perceber, todos os voluntários músicos, com exceção do M7, são amadores, apresentando poucas horas de prática semanal, apesar dos muitos anos de prática do instrumento, enquanto que o voluntário M7 tem por principal instrumento o piano, mas toca outros instrumentos também e pratica seu instrumento diariamente.

Todos os participantes assinaram Termo de Consentimento (Apêndice A) para a realização dos experimentos e preencheram um questionário sobre as informações descritas na Tabela 3.

4.3 ESTÍMULOS

Duas peças musicais instrumentais foram utilizadas como estímulo: (1) a música “Dança Húngara No.5”, de Johannes Brahms (201,4182 s de duração), apresentada pela *Orquestra Sinfônica Húngara Budapeste*, e (2) a música “O Barbeiro de Sevilha - Abertura”, de Gioachino Rossini (420,6116 s de duração), apresentada pela *Orquestra Filarmônica de São Petersburgo*.

Os dois áudios foram tirados de uma gravação ao vivo, possuem som estéreo e uma taxa de amostragem de 44,1 kHz. Estas peças musicais foram selecionadas devido ao alto nível de variação das suas características musicais como dinâmica, timbre, tonalidade e ritmo.

4.4 PROCEDIMENTOS

O experimento foi desenvolvido em linguagem *C#* através da plataforma *Microsoft Visual Studio*. Essencialmente, o programa desenvolvido realiza a sincronização entre o início da reprodução dos áudios e o início da gravação dos dados de EEG pela plataforma OpenBCI.

Antes de iniciar os experimentos, um áudio irrelevante para a análise foi apresentado ao voluntário para determinar o volume do áudio, sendo adequado de acordo com a preferência do voluntário.

Todos os experimentos foram realizados em uma sala fechada, não havendo ruídos externos ou interrupções, onde o voluntário esteve sentado em uma cadeira confortável, permanecendo o mais imóvel possível e de olhos abertos, sendo instruído a prestar o máximo de atenção à música. As músicas foram apresentadas aos voluntários por meio de fones de ouvido intra-auriculares.

Durante os experimentos, as músicas foram apresentadas em sequência, havendo uma breve conversa entre o pesquisador e o voluntário entre cada uma delas. Sendo assim, as músicas foram apresentadas a todos os voluntários na seguinte ordem:

a) O Barbeiro de Sevilha – Abertura, de Gioachino Rossini;

b) Dança Húngara No. 5, de Johannes Brahms.

4.5 EXTRAÇÃO E SELEÇÃO DAS CARACTERÍSTICAS ACÚSTICAS

Conforme detalhado no capítulo anterior, foram extraídas 12 características acústicas ($c = 12$) de baixo nível, utilizando um janelamento de 50 ms ($l = 0{,}05s$) com fator de sobreposição $h = 50\%$. Tais parâmetros resultaram em uma matriz $\mathbf{M}$ de dimensão 24880×12, concatenando $w_1 = 8056$, referente a música "Dança Húngara No. 5", de Johannes Brahms, e $w_2 = 16824$ referente a música "O Barbeiro de Sevilha - Abertura", de Gioachino Rossini.

O PCA foi aplicado sobre a matriz de dados $\mathbf{M}$ normalizada, onde foram selecionados os 3 primeiros componentes principais, de acordo com os critérios propostos, que explicam

Tabela 4 – Componentes principais, autovalores, porcentagem da variância explicada e proporção acumulada pelos componentes

Componente Principal	Autovalores	Proporção	Proporção Acumulada
CP1	6.0852	0.5071	0.5071
CP2	2.2937	0.19114	0.69825
CP3	1.9916	0.16596	0.86421
CP4	0.66901	0.055751	0.91996
CP5	0.33262	0.027719	0.94768
CP6	0.26804	0.022337	0.97002
CP7	0.133	0.011083	0.9811
CP8	0.08443	0.0070358	0.98814
CP9	0.069542	0.0057952	0.99393
CP10	0.052664	0.0043886	0.99832
CP11	0.014469	0.0012057	0.99953
CP12	0.0056935	0.00047446	1

Fonte: Autor.

Tabela 5 – Carregamento das variáveis nos fatores

N°	Características	F1	F2	F3
1	RMS	0.15783	**0.9666**	0.068971
2	ZCR	0.77885	0.14297	-0.25826
3	S. Rolloff	0.89205	0.18486	-0.34445
4	Roughness	0.17045	0.79403	-0.1078
5	Brightness	0.86602	0.14573	-0.29105
6	S. Entropy	0.83221	0.20548	0.26655
7	S. Flatness	0.7226	0.13985	0.053021
8	S. Skewness	-0.021481	-0.025767	**0.99704**
9	S. Kurtosis	-0.15715	-0.069974	0.95244
10	S. Centroid	**0.94926**	0.15338	-0.26571
11	S. Spread	0.76615	0.20286	0.23778
12	S. Flux	0.24738	0.90153	-0.072014

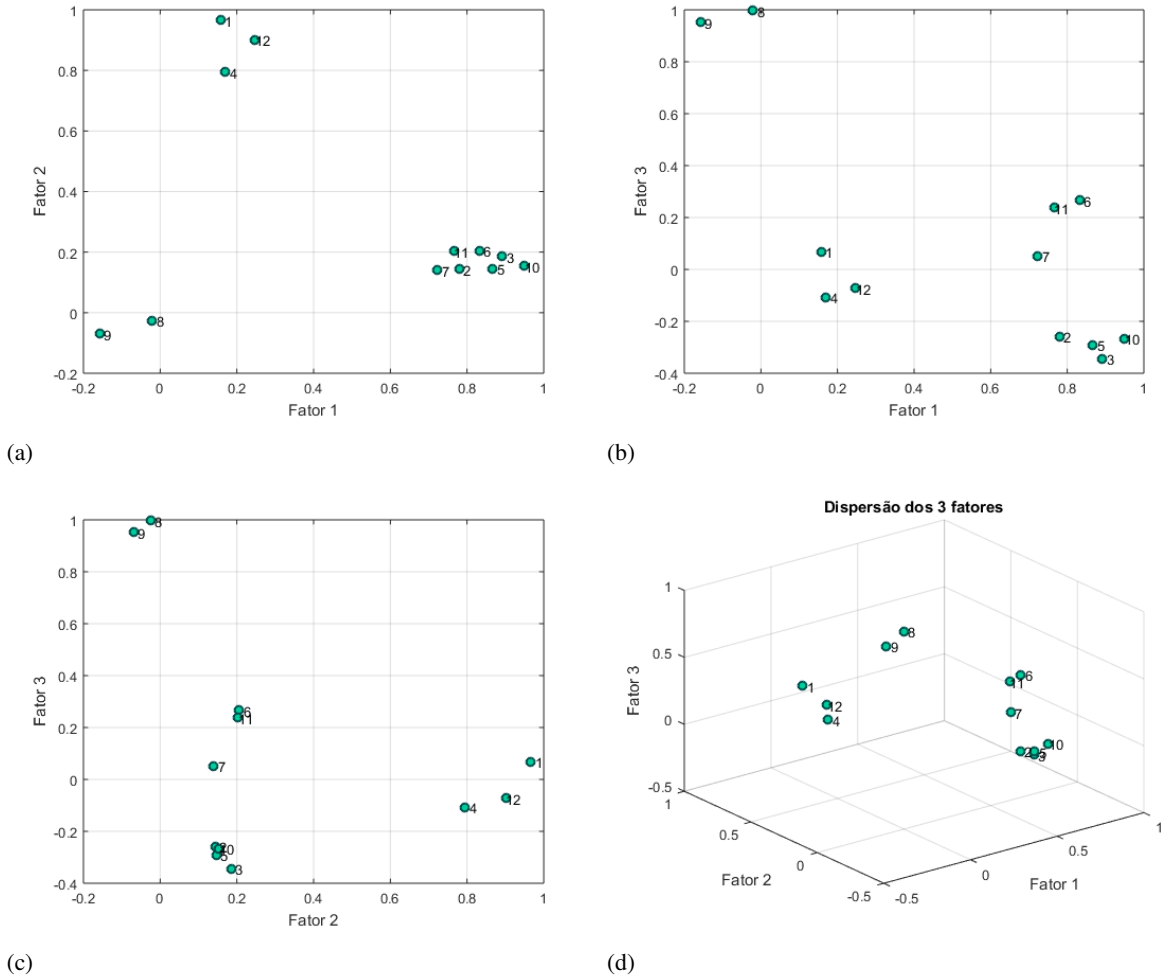
Fonte: Autor.

mais de 85% da variância dos dados, conforme apresentado na Tabela 4, sendo suficientes para representar os áudios selecionados.

Desta forma, foi realizado o FA, utilizando 3 fatores ($p = 3$). A Tabela 5 apresenta a carga fatorial de cada variável em cada fator, destacando as características acústicas que apresentam maior carga fatorial, indicando o número de correspondência entre a variável e o fator.

A Figura 29 mostra o gráfico de dispersão dos 3 fatores avaliados gerados a partir da carga fatorial de cada característica acústica em cada fator. É possível verificar que existem agrupamentos entre as características acústicas extraídas dos áudios. A Figura 29d apresenta de forma ainda mais clara a existência destes agrupamentos. Portanto, o resultado encontrado no FA revela que existe redundância entre as características acústicas extraídas dos áudios e estas características se agrupam de acordo com a similaridade da informação extraída por elas.

Figura 29 – Gráfico de dispersão das características acústicas nos fatores da Análise Fatorial



(a) (b)

(c) (d)

Fonte: Autor.

Aplicando os critérios propostos e descritos no capítulo anterior, as características acústicas *Spectral Centroid*, *Root Mean Square* e *Spectral Skewness* deveriam ser selecionadas para a análise. Porém, tanto no Fator 1 quanto no Fator 3, as características acústicas que possuíam maior carga fatorial foram reprovadas, pois as características *Spectral Centroid* e *Spectral Skewness* não apresentavam o número mínimo ($tr \geq 4$) de *triggers* estabelecidos.

Apesar de terem sido reprovadas, os *triggers* encontrados nas características *Spectral Centroid* e *Spectral Skewness* estão contidos nas características que os substituíram, sustentando a afirmação de que existe redundância nas informações contidas nas características acústicas extraídas dos áudios.

Sendo assim, as características acústicas que foram selecionadas para a análise são:

a) *S. Rolloff*: 89% de carga fatorial no F1;

b) *RMS*: 96% de carga fatorial no F2;

c) *S. Kurtosis*: 95% de carga fatorial no F3.

Para cada característica acústica selecionada foram encontrados os *triggers* conforme mostra a Tabela 6, onde foram utilizados os limiares superior (V_{p+}) e inferior (V_{p-}) como sendo +/-20% da média de cada característica acústica.

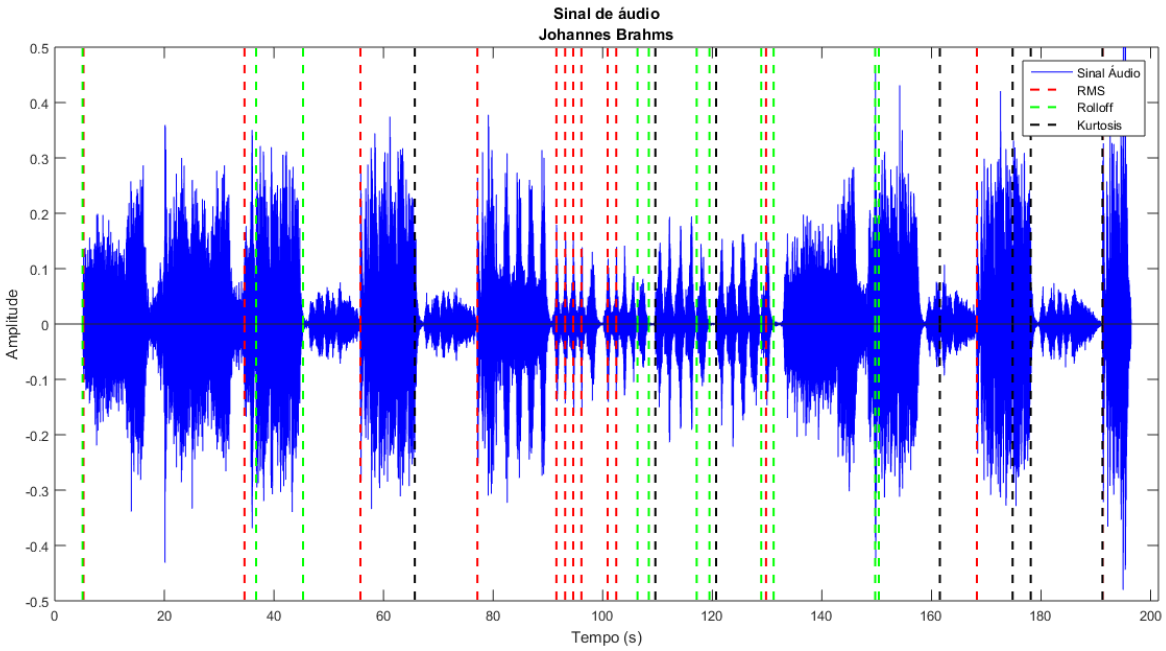
Tabela 6 – Apresentação dos parâmetros utilizados para determinação dos *triggers* de cada característica e o número de *triggers* encontrados

Características	Média	PLFP(ms)	*triggers*
Johannes Brahms:			
RMS	0,0330	750	13
Rolloff	3192,6	400	11
Kurtosis	20,0857	500	7
Gioachino Rossini:			
RMS	0,0278	500	14
Rolloff	2306,8	500	16
Kurtosis	27,9389	400	15

Fonte: Autor.

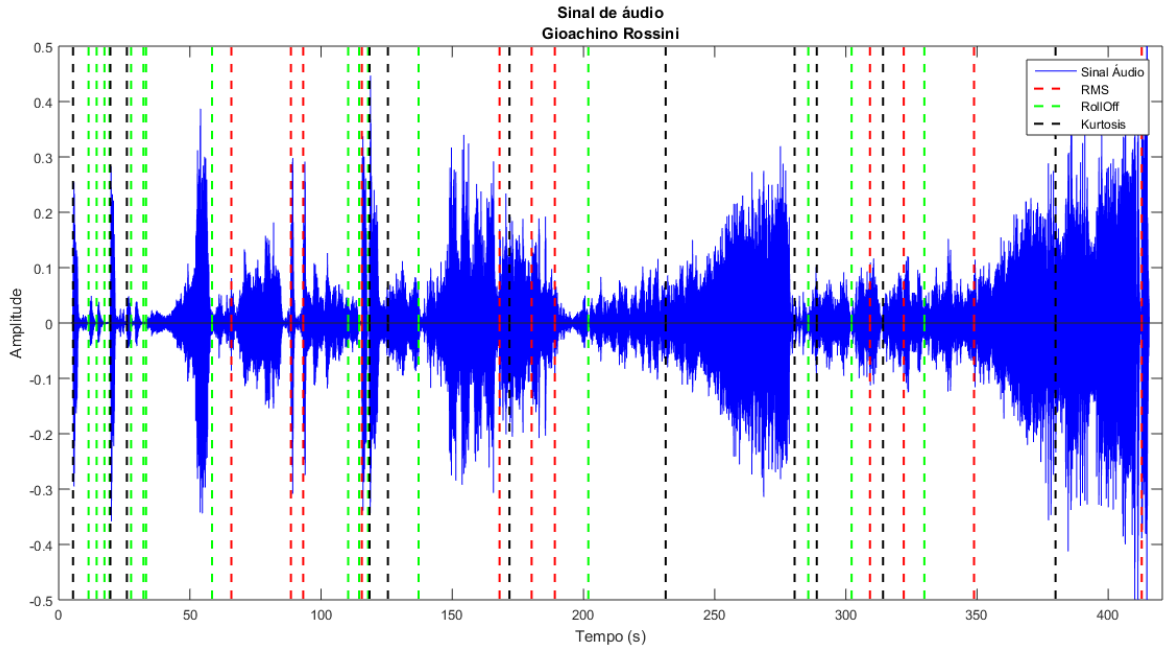
Apesar da música "Dança Húngara No.5", de Johannes Brahms ter aproximadamente a metade da duração apresentada na música "O Barbeiro de Sevilha - Abertura", de Gioachino Rossini, o número de *triggers* encontrado para ela apenas um pouco menor.

Figura 30 – Forma de onda da música “Dança Húngara No.5” de Johannes Brahms com os *triggers* encontrados



Fonte: Autor.

Figura 31 – Forma de onda da música “O Barbeiro de Sevilha - Abertura” de Gioachino Rossini com os *triggers* encontrados



Fonte: Autor.

As Figuras 30 e 31 apresentam as formas de onda dos dois áudios, indicando os instantes que ocorreram os *triggers* para cada uma das p característica acústica selecionada ao longo do tempo, onde é possível ver que boa parte dos *triggers* ocorreram nos momentos em que há grande contraste na amplitude do sinal.

Este conjunto de *triggers* selecionados correspondem a instantes na duração das músicas apresentadas aos voluntários, onde serão analisadas, no sinal de EEG, quais são as respostas neurais decorrentes da exposição deles aos voluntários.

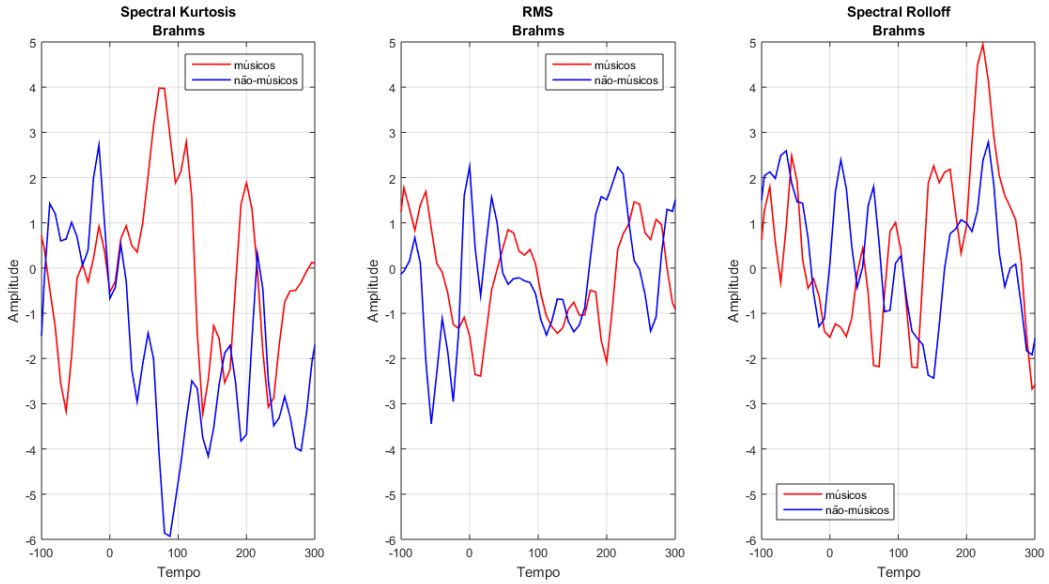
4.6 PROCESSAMENTO E ANÁLISE DO SINAL DE EEG

Os sinais de EEG foram filtrados, utilizando um filtro Butterworth passa faixa (1-30Hz), de ordem 10, pós-processando o sinal conforme dito no capítulo anterior onde o sinal foi janelado em épocas de 200ms pré-estímulo e 800ms pós-estímulo para cada *trigger*, resultando em um único sinal médio para cada característica acústica, contendo a resposta neural resultante da exposição delas ao voluntário no decorrer das músicas.

Para cada voluntário, portanto, são obtidos 3 sinais de 1000ms de duração, resultante da média das épocas de cada característica acústica. Estes sinais são diferentes para cada voluntário, porém, espera-se que haja similaridade entre os sinais em cada grupo.

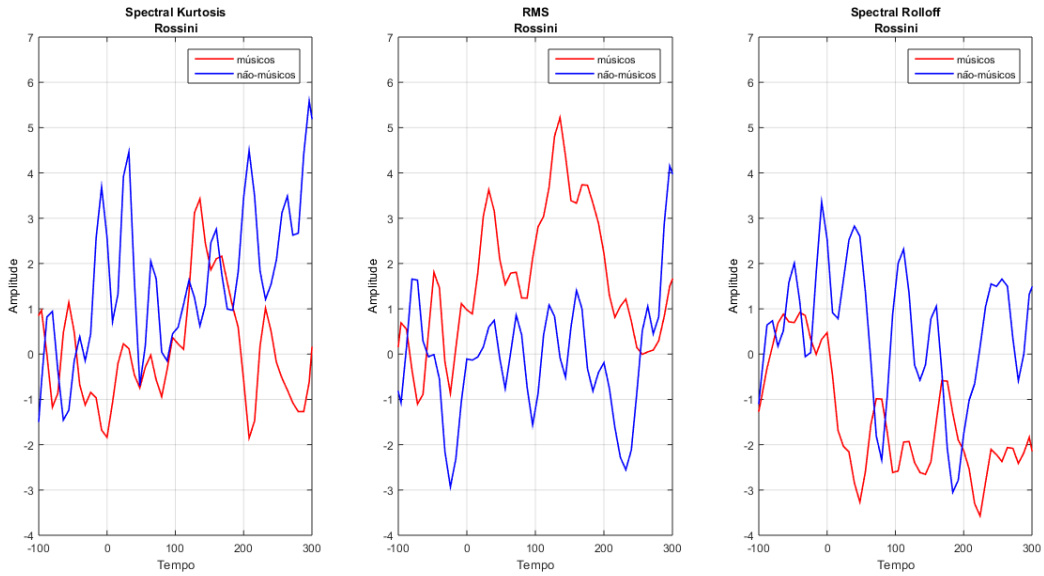
Portanto, levando em consideração a classificação prévia dos voluntários, foi realizada uma média geral destes sinais dentro dos dois grupos de voluntários (músicos *versus* não-músicos).

Figura 32 – Sinal médio da região central de músicos e não-músicos para a música: Dança Húngara No.5, de Johannes Brahms



Fonte: Autor.

Figura 33 – Sinal médio da região central de músicos e não-músicos para a música: O Barbeiro de Sevilha, de Gioachino Rossini



Fonte: Autor.

Para analisar a ocorrência dos componentes N100, caracterizado por um vale entre 80-150ms pós-estímulo, e P200, caracterizado por um pico entre 150-275ms pós-estímulo, foram observados os sinais resultante da região central do córtex (média entre os eletrodos C3 e C4), conforme apresentam as Figuras 32 e 33.

Em ambas as músicas, os sinais resultante dos dois grupos são diferentes entre si, porém não é possível tirar muitas conclusões destes sinais, pois eles se apresentam muito ruidosos, onde a expectativa era que estas curvas se apresentassem mais suavizadas, apresentando os componentes de ERP de forma clara no sinal. Isto evidencia que possivelmente não ocorreram os componentes de ERP para todos os voluntários para as características acústicas observadas.

Realizando a média dos *trigger* de cada característica é possível atenuar estes ruídos para cada voluntário, evidenciando o comportamento que é comum em cada época, filtrando o sinal pela média (LUCK, 2014). Porém, com a remoção de artefatos, alguns destes sinais podem ser rejeitados, devido a inúmeros fatores que podem causá-los, evidenciando assim a necessidade de um número mínimo de *triggers* para cada característica, como proposto nos critérios de seleção das características acústicas.

De fato, durante a remoção de ruídos, houve voluntários cujas características acústicas foram representadas apenas por uma única época e, desta forma, durante a realização do filtro pela média (média entre *triggers* para cada característica acústica) o sinal resultante para aquele voluntário não teve os seus ruídos atenuados, impactando no resultado do grupo.

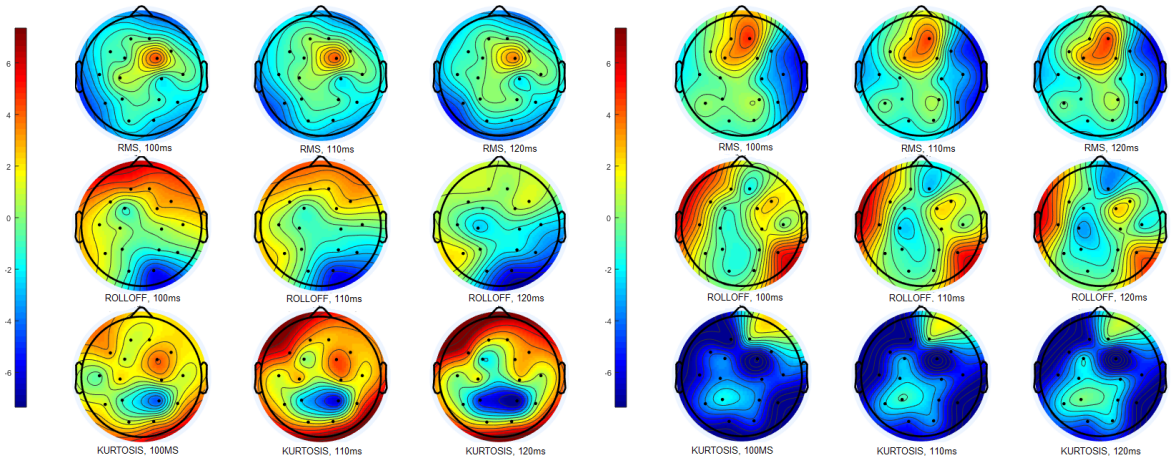
Este é um fator importante a se levar em consideração. Quanto menor a duração de uma música, menores são as chances de ocorrência de *triggers* para as características acústicas selecionadas, e os resultados obtidos sugerem que ao ser representado apenas por um pequeno número de *triggers*, a característica acústica pode falhar em apresentar os componentes de ERP desejados.

Em contrapartida, músicas muito longas podem levar o voluntário a um estado de relaxamento e, algumas vezes, de sonolência, afetando os resultados do experimento. E as respostas neurais podem variar de uma música para outra, como é possível perceber nas escolhidas para este experimento, onde a música de Johannes Brahms possui contrastes mais marcantes em comparação com a de Gioachino Rossini.

Diferentemente dos experimentos tradicionais de ERP com sons artificiais controlados, que podem ser repetidos diversas vezes, em uma música, apesar dos *triggers* selecionados estarem relacionados a um mesmo aspecto acústico, ou a uma mesma característica acústica, eles são compostos por acontecimentos diferentes ao longo da música que podem ou não se repetir, limitando a sua representabilidade dentro deste aspecto que está sendo observado.

Quanto maior o contraste na característica acústica analisada, maior a chance de ocorrência de um potencial evocado. Ao determinar os parâmetros utilizados para a seleção de *triggers*, quanto mais rígido eles forem, maior o contraste acústico por eles representado. Porém, quanto mais rígido estes parâmetros forem, menor o número de *triggers* extraídos nas músicas.

Figura 34 – Mapa topográfico cortical da música Dança Húngara No.5, de Johannes Brahms, para a componente N100



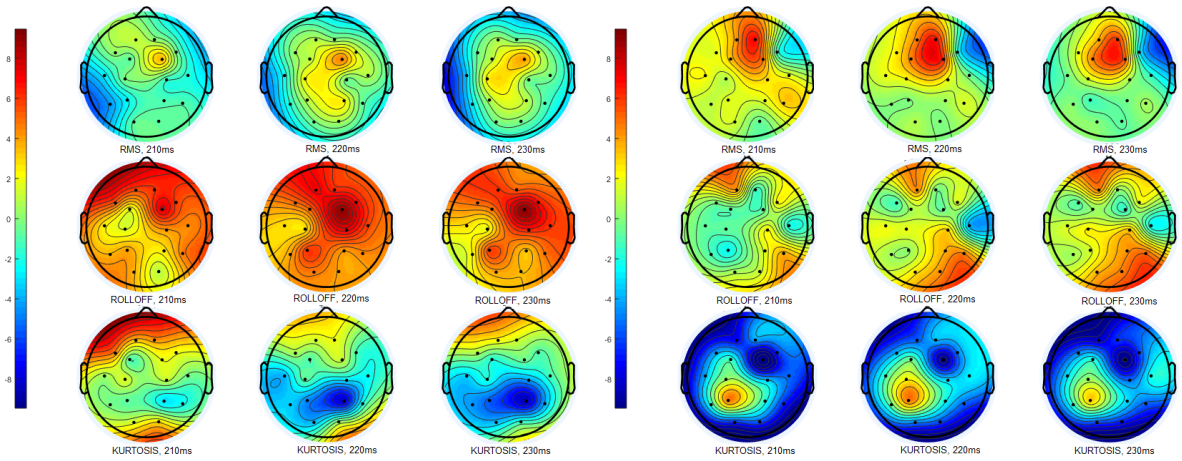
(a) (b)

Fonte: Autor.
Legenda: Média do sinal de (a) músicos e (b) não-músicos para as características RMS, *Spectral Rolloff* e *Spectral Kurtosis* entre 100 ms e 120 ms pós-estímulo.

Um outro tipo de abordagem que pode ser feito é a análise dos mapas topográficos corticais, que representam, de forma visual, a relação do potencial elétrico pelo escalpo como um todo, não olhando apenas para o sinal presente em um eletrodo, mas apresentando uma visão global da atividade elétrica.

A Figura 34 apresenta o mapa topográfico referente ao componente N100 entre os instantes 100-120 ms pós-estímulo para músicos e não-músicos, da música de Johannes Brahms.

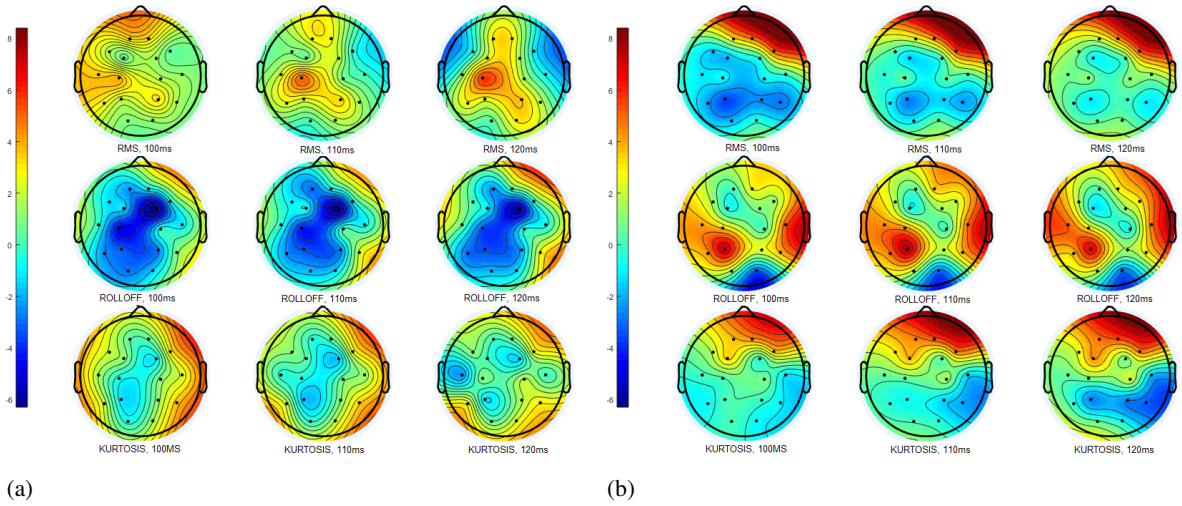
Figura 35 – Mapa topográfico cortical da Dança Húngara No. 5, de Johannes Brahms, para a componente P200



(a) (b)

Fonte: Autor.
Legenda: Média do sinal de (a) músicos e (b) não-músicos para as características RMS, *Spectral Rolloff* e *Spectral Kurtosis* entre 210 ms e 230 ms pós-estímulo.

Figura 36 – Mapa topográfico cortical da música O Barbeiro de Sevilha - Abertura, de Gioachino Rossini, para a componente N100

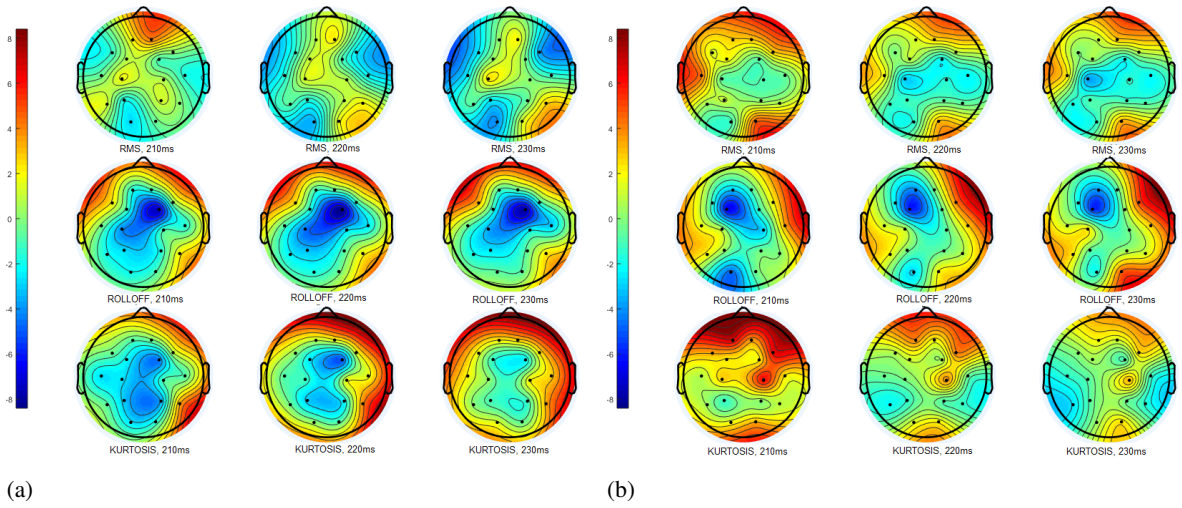


Fonte: Autor.
Legenda: Média do sinal de (a) músicos e (b) não-músicos para as características RMS, *Spectral Rolloff* e *Spectral Kurtosis* entre 100 ms e 120 ms pós-estímulo.

Na Figura 34a, que apresenta o mapa topográfico de músicos, é possível observar, na característica *Spectral Kurtosis*, a geração de uma área negativa na região parietal, que se inicia em 100 ms pós-estímulo e se propaga ao longo do tempo, como esperado para a componente N100, diferente das outras características acústicas, onde este comportamento não ocorreu.

Em contrapartida, a Figura 34b, que apresenta o mapa topográfico de não-músicos, não é possível observar a geração de uma área negativa distinta como no outro grupo, pelo contrário,

Figura 37 – Mapa topográfico cortical da música O Barbeiro de Sevilha - Abertura, de Gioachino Rossini, para a componente P200



Fonte: Autor.
Legenda: Média do sinal de (a) músicos e (b) não-músicos, para as características RMS, *Spectral Rolloff* e *Spectral Kurtosis* entre 210 ms e 230 ms pós-estímulo.

a característica *Spectral Kurtosis* apresentava valores muito baixos, não sendo possível verificar o mesmo comportamento característico da componente N100 como no grupo de músicos.

A Figura 35 apresenta o mapa topográfico referente ao componente P200 entre os instantes 210 ms e 230 ms pós-estímulo para músicos e não-músicos, da música de Johannes Brahms. Na Figura 35a é possível observar, para a característica *Spectral Rolloff*, a geração de uma área positiva na região fronto-central, como esperado da componente P200. Além disso, com menor intensidade, é possível observar o mesmo comportamento para a característica RMS na região frontal. Porém, para a característica *Spectral Kurtosis* isso não ocorre.

Na Figura 35b, a característica RMS apresenta uma área positiva na região frontal e a característica *Spectral Kurtosis*, com menor intensidade, apresenta a geração de uma área positiva na região parietal à esquerda. A característica *Spectral Rolloff* não apresenta o mesmo comportamento característico.

A Figura 36 presenta o mapa topográfico entre os instantes 100-120 ms, referente ao componente N100, pra músicos e não-músicos, da música de Gioachino Rossini. Na Figura 36a é possível observar, na característica Spectral Rolloff, a geração de uma área negativa em toda a região fronto-centro-parietal do escalpo, característica da componente N100. As outras características não apresentaram essa área negativa.

Em contrapartida, a Figura 36b não é possível observar esse comportamento, com exceção da característica RMS que apresenta uma região negativa em 100ms pós-estimulo com intensidade menor que a encontrada no grupo de músicos.

A Figura 37 presenta o mapa topográfico entre os instantes 210-230ms, referente ao componente P200, pra músicos e não-músicos, da música de Gioachino Rossini. Enquanto que, para a música de Johannes Brahms, ocorreram áreas positivas distintas para algumas características acústicas em ambos os grupos, na música de Gioachino Rossini, este comportamento não foi observado, com exceção da característica *Spectral Kurtosis*, na Figura 37b, na região central.

Como é possível perceber, ao longo do tempo ocorrem mudanças no potencial elétrico do córtex resultantes das informações novas (estímulos) que estão sendo apresentadas aos voluntários e processadas pelo cérebro.

Dentro dos grupos, as mudanças que ocorrem devido a apresentação destes estímulos são distintas ao longo do tempo, destacando a diferença entre músicos e não-músicos, especialmente na música de Johannes Brahms, onde estas diferenças aparentam ser mais claras.

Através dos mapas topográficos é possível perceber, portanto, a distinção entre os grupos de músicos e não-músicos e a ocorrência dos potenciais evocados para as características acústicas selecionadas, tanto para músicos quanto para não-músicos e confirma-se que não foram todas as características acústicas que geraram os componentes de ERP tradicionais neste experimento.

4.7 CLASSIFICAÇÃO

Conforme proposto na metodologia, para a classificação foram analisados 2 cenários, onde, cada voluntário é representado por um vetor composto pela concatenação dos valores registrados para todos os eletrodos ($Ch = 16$), para cada característica acústica selecionada ($fp = 3$), para os dois instantes analisados, resultando em um vetor de dimensionalidade $d = 96$. Sendo assim, os dados obtidos de cada voluntário foram agrupados em uma matriz $\mathbf{S_{96x26}}$, onde as 26 colunas se referem ao número de voluntários para cada cenário, sendo eles:

a) Cenário 1: Conjunto de dados constituídos do potencial elétrico registrado nos eletrodos para cada característica acústica selecionada, nos instantes 100 ms e 200 ms pós-estímulo;

b) Cenário 2: Conjunto de dados constituídos do potencial elétrico registrado nos eletrodos para cada característica acústica selecionada, nos instantes em que ocorrem os pontos mínimos e máximos pós-estímulo correspondentes às componentes N100 e P200, na região central do escalpo (média entre os eletrodos C3 e C4);

A partir dos cenários propostos foi possível comparar entre classificar os voluntários por meio de instantes iguais (determinístico) entre eles, ou seja, as latências 100 ms e 200 ms pós-estímulo, que não leva em consideração o que está acontecendo com o sinal de EEG ao longo do tempo dentro das épocas analisadas, ou classificar os voluntários por meio das respostas esperadas para ocorrência dos potenciais evocados, utilizando como valores os pontos mínimos entre 80 ms e 150 ms (N100) pós-estímulo e os pontos máximos entre 150 ms e 275 ms (P200) pós-estímulo, observando instantes diferentes entre os voluntários.

Para realizar a classificação, é necessário separar os dados em um conjunto de treinamento e um de testes. Para avaliar a generalização dos métodos propostos foi utilizado o X-*Fold Cross-validation*, onde o conjunto de dados é separado aleatoriamente em X partes. A cada iteração, uma das X partes é selecionada como teste e o restante é selecionado como treinamento, sendo que este processo é repetido X vezes, utilizando todas as partes por ele selecionadas, ou seja, se $X = 5$, a classificação será realizada 5 vezes, utilizando 5 conjuntos de treinamento diferentes. Após ser realizada a classificação para cada conjunto de treinamento, é realizada a média de todas as taxas de acerto na classificação encontrada para cada X conjunto de treinamento.

Para garantir que no conjunto de teste e treinamento houvessem um número proporcional de músicos e não-músicos, o *cross-validation* foi realizado separadamente para cada grupo primeiramente. Após serem determinados os conjuntos de teste e treinamento para os dois grupos eles foram agrupados novamente.

Além disso, para verificar a influência do tamanho do conjunto de treinamento nos classificadores, foram utilizados 2 valores para o X-fold, sendo eles o 5-fold, onde o conjunto de dados é dividido em 5 partes iguais, apresentando 80% dos dados como conjunto de treina-

mento e 20% para teste, e o 10-fold, onde o conjunto de dados é dividido em 10 partes iguais, apresentando 90% dos dados como conjunto de treinamento e 10% para teste.

Para o classificador k-NN, os dados foram divididos em duas partes, o conjunto de treinamento e o de teste. Foi realizado o PCA sobre a matriz $\mathbf{S_{96x26}}$, projetando os dados nos autovetores, variando o número de componentes principais entre 5 e 25, e variando o número de k vizinhos mais próximos entre 1 e 3, selecionando como resultado aquele que apresentava melhor taxa de acerto na classificação.

Para a rede neural MLP, o conjunto de dados foi dividido em três partes, o conjunto de treinamento, validação e teste. O conjunto de validação corresponde a uma parte (*fold*) e o conjunto de testes corresponde a uma parte (*fold*) diferente do de validação, enquanto que o restante dos dados foi utilizado para treinamento. A classificação foi feita utilizando como treinamento o algoritmo Gradiente Descendente, utilizando como critério de parada o número máximo de 1000 épocas (iterações) e como função de ativação dos neurônios uma função não-linear do tipo sigmoide.

Foi utilizado apenas 1 camada escondida, tendo o número H de neurônios variando entre 5 e 115, selecionando aquele H que apresentava melhor taxa de acerto na classificação como resultado. Além disso, a camada de saída possui dois neurônios, onde (1,0) corresponde ao grupo de músicos e (0,1) corresponde ao grupo de não-músicos.

Portanto, para cada música foram obtidos dois valores, resultantes do *cross-validation* 5-fold e 10-fold, aplicados para cada um dos cenários propostos, variando, portanto, o tamanho do conjunto de treinamento para cada análise.

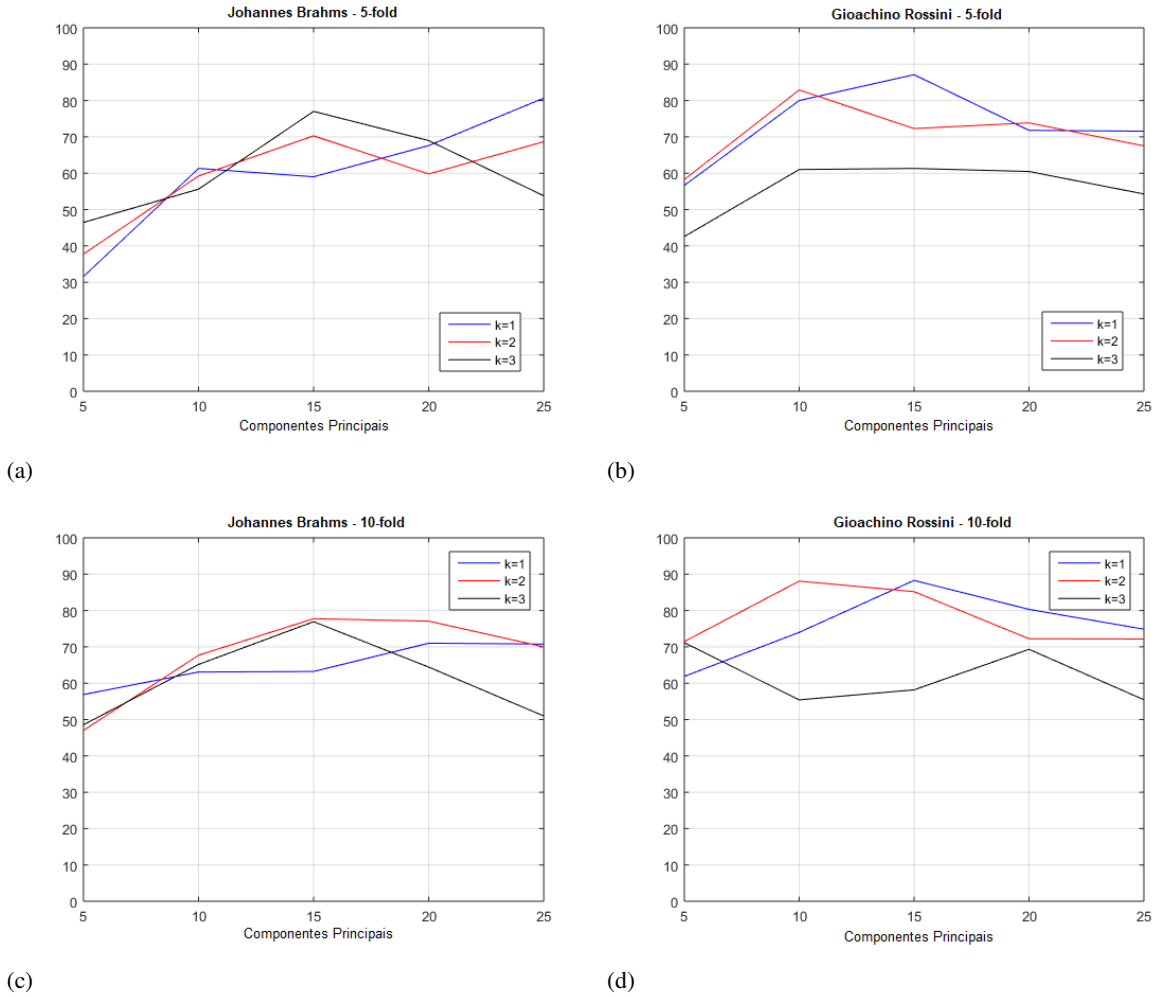
4.7.1 Classificação de acordo com Cenário 1

Para a classificação no cenário 1 a matriz S de entrada de dados apresenta os valores do potencial elétrico nos instantes 100 ms e 200 ms pós-estímulo, para cada característica acústica selecionada, analisando o cérebro como um todo, porém sem levar em consideração a variação do sinal de EEG ao longo do tempo, observando instantes determinados de forma igual para cada voluntário.

A Figura 38 apresenta os resultados encontrados pelo k-NN variando o número de componentes principais, para $k = 3$ vizinhos mais próximos, onde a música de Gioachino Rossini apresentou a melhor classificação, com 88,3% de acerto na classificação utilizando um conjunto de treinamento maior (10-fold).

É possível observar que ao aumentar o tamanho do conjunto de treinamento, de 5-fold para 10-fold, o classificador teve uma leve melhora. Em todas as condições, aumentando-se o número de componentes principais melhorou a classificação, até atingir 15 componentes principais, quando a taxa de acerto na classificação parece se manter estável ou declinar um pouco.

Figura 38 – Taxa de acerto na classificação do Cenário 1 para o classificador k-NN



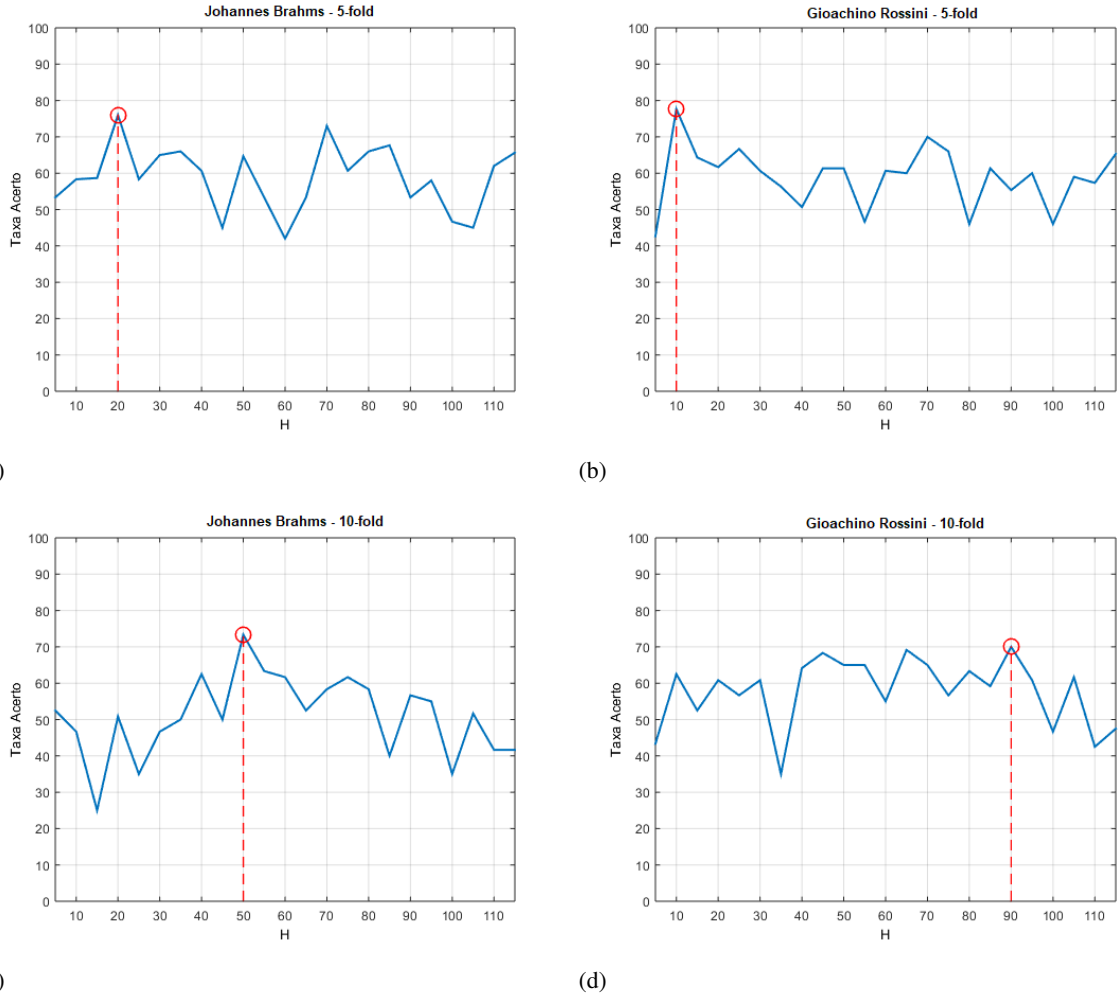
Fonte: Autor.
Legenda: (a) 5-fold para Johannes Brahms, (b) 5-fold para Gioachino Rossini, (c)10-fold para Johannes Brahms e (d) 10-fold para Gioachino Rossini.

Para a música de Johannes Brahms as 10 primeiras componentes principais representam 78,3% da variância dos dados, alcançando 95% da variância dados na componente 18. Já para a música de Gioachino Rossini, as 10 primeiras componentes principais representam 81,8% da variância dos dados, alcançando 95% da variância na componente 16.

A Rede Neural foi realizada apresentando uma camada de entrada, uma camada escondida e uma camada de saída, sendo 96, H e 2 o número de neurônios para cada camada, respectivamente, onde os valores de H variavam entre 5 e 115. As duas saídas desejadas da rede neural correspondem ao grupo de músicos (1,0) e ao grupo de não-músicos (0,1).

A Figura 39 apresenta a taxa de acerto para cada H onde é possível perceber uma instabilidade na rede conforme varia-se o número de neurônio na camada escondida, atingindo a melhor classificação para a música de Gioachino Rossini com 77,6% de acerto na classificação utilizado o menor conjunto de treinamento (5-fold).

Figura 39 – Taxa de acerto na classificação do Cenário 1 para a rede neural MLP



Fonte: Autor.
Legenda: (a) 5-fold para Johannes Brahms, (b) 5-fold para Gioachino Rossini, (c)10-fold para Johannes Brahms e (d) 10-fold para Gioachino Rossini.

Na Tabela 7 são apresentados os valores do número de k vizinhos mais próximos que apresentaram melhor taxa de acerto na classificação das duas músicas no k-NN com o valor da taxa de acerto e o desvio padrão, e o número H de neurônios na camada escondida da Rede Neural que apresentaram melhor taxa de acerto na classificação, apresentando a taxa de acerto e o desvio padrão. A Figura 40 apresenta um gráfico de barras com os valores de classificação para os dois classificadores.

Neste cenário, para as duas músicas, o classificador que obteve a maior taxa de acerto na classificação foi o k-NN, que é um classificador que se baseia na distância entre os vizinhos, enquanto que a rede neural, que foi projetada para encontrar uma relação não-linear entre os dados, apresenta um problema de *overfitting*, devido ao baixo número de amostras.

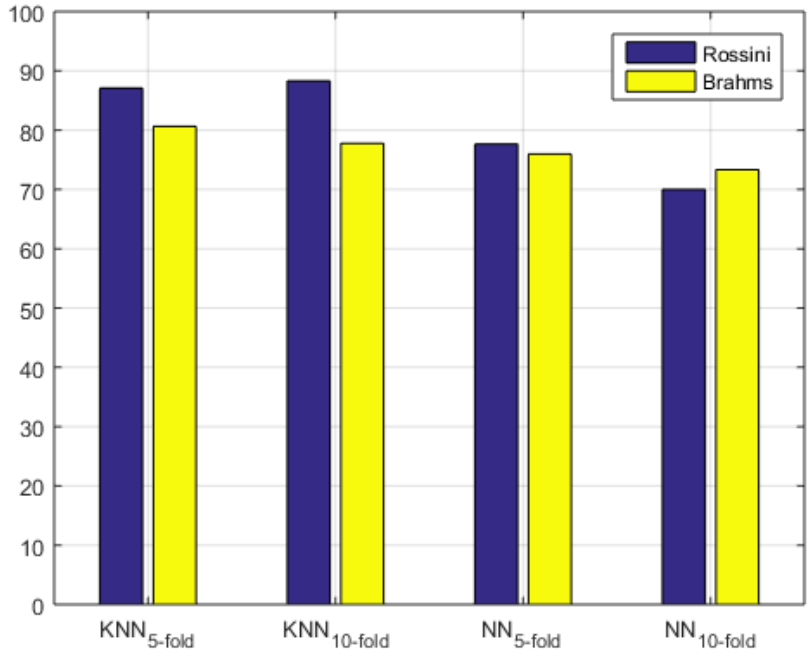
Neste caso, aumentar o tamanho do conjunto de treinamento (X-fold maior) não resultou em nenhuma mudança significativa no resultado dos classificadores.

Tabela 7 – Resultado da classificação dos voluntários para o Cenário 1 para os classificadores k-NN e a Rede Neural MLP

		k-NN			Rede Neural		
X-**fold**	**música**	k	Média	Desvio Padrão (%)	H	Média	Desvio Padrão (%)
5-fold	*Rossini*	1	87,1%	8,20	10	77,6%	8,9
	Brahms	1	80,6%	8,40	20	76%	18
10-fold	*Rossini*	1	88,3%	8,90	90	70%	11
	Brahms	2	77,7%	8,91	50	73,3%	32

Fonte: Autor.

Figura 40 – Gráfico de Barras, apresentando a classificação dos voluntários para o Cenário 1



Fonte: Autor.

Ainda assim, os dois classificadores encontraram uma taxa de acerto na classificação superior a 70% para as duas música, enquanto que a música O Barbeiro de Sevilha - Abertura, de Gioachino Rossini, apresentou uma melhor classificação.

4.7.2 Classificação de acordo com Cenário 2

A matriz S utilizada para a classificação no cenário 2 apresenta os valores do potencial elétrico determinado de acordo com os pontos mínimos (80 ms a 150 ms pós-estímulo) e máximos (150 ms a 275 ms pós-estímulo) nas regiões centrais do escalpo.

Tabela 8 – Latência média dos grupos de voluntários para cada característica acústica apresentada.

Classe	Característica Acústica	Latência N100	Latência P200
	Gioachino Rossini		
Músicos	Spectral Kurtosis	107 ms	206 ms
	RMS	109 ms	199 ms
	Spectral Rolloff	129 ms	203 ms
Não-músicos	Spectral Kurtosis	115 ms	217 ms
	RMS	115 ms	218 ms
	Spectral Rolloff	108 ms	214 ms
	Johannes Brahms		
Músicos	Spectral Kurtosis	135 ms	220 ms
	RMS	121 ms	237 ms
	Spectral Rolloff	104 ms	219 ms
Não-Músicos	Spectral Kurtosis	122 ms	236 ms
	RMS	126 ms	224 ms
	Spectral Rolloff	115 ms	227 ms

Fonte: Autor.

Portanto, para cada voluntário será registrado o valor do potencial elétrico de acordo com a latência (instante) em que ocorrem os pontos mínimos e máximos dentro das janelas observadas, podendo ser observados instantes diferente entre cada voluntário, considerando o cérebro como um todo e levando em consideração as alterações do sinal de EEG ao longo do tempo.

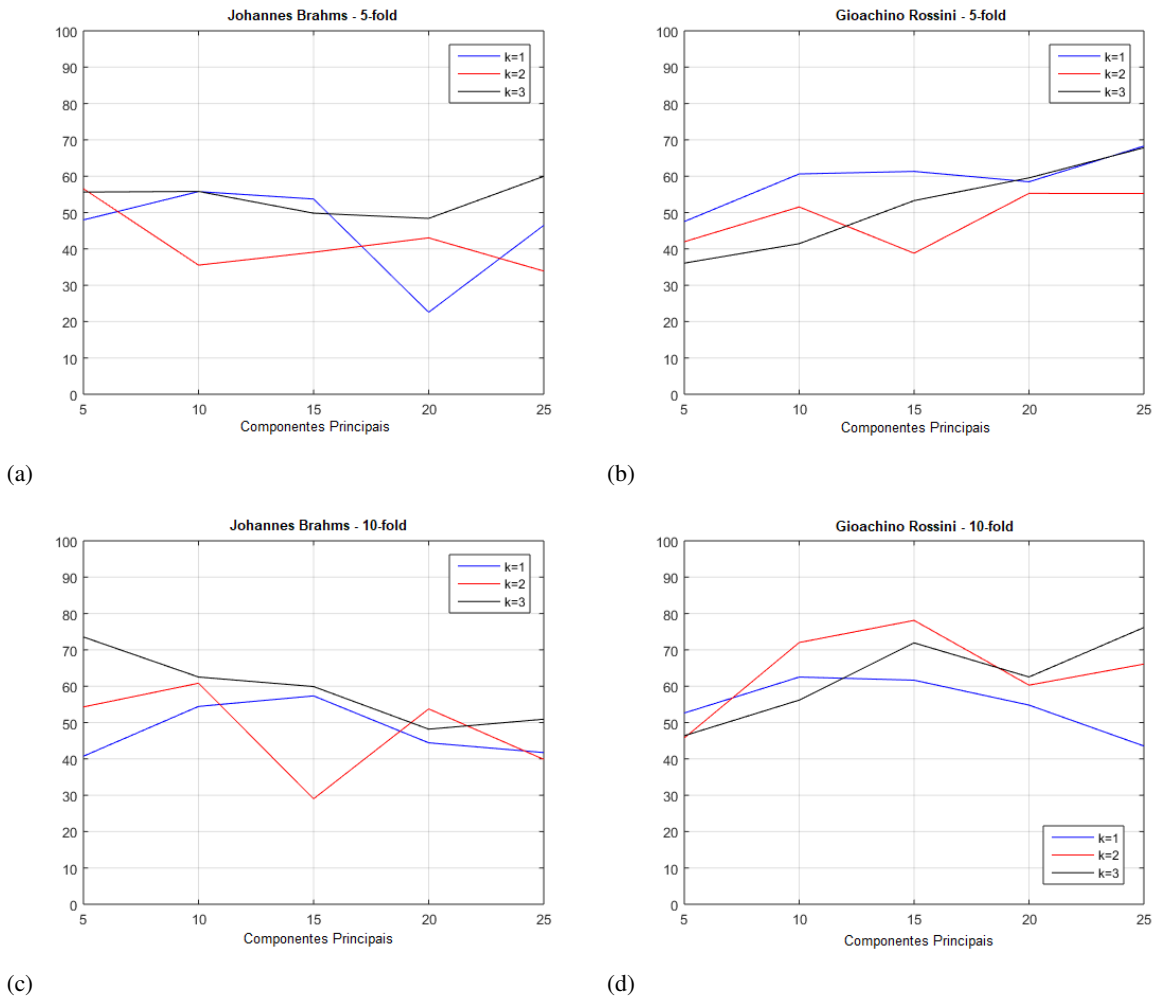
A Tabela 8 apresenta as latências médias para cada característica acústica resultante da média do grupo de voluntários autodenominados músicos e não-músicos, onde é possível observar a diferença nas latências para cada grupo, variando também entre as características acústicas.

Para a música de Gioachino Rossini, os pontos de mínimo foram apresentados em uma latência maior para não-músicos, com exceção apenas característica *Spectral Rolloff*, enquanto que, para a música de Johannes Brahms, essa distinção não é tão clara, portanto não é possível generalizar o período de ocorrência dos pontos mínimos e máximos para cada grupo.

A Figura 41 apresenta os resultados encontrados pelo k-NN variando o número de componentes principais, para $k = 3$ vizinhos mais próximos determinados, onde música de Gioachino Rossini apresentou a melhor taxa de acerto na classificação, de 78,1%.

É possível observar que independente do tamanho do conjunto de treinamento, a classificação para a música de Johannes Brahms (Figuras 41a e 41c) se mantém abaixo de 70% e o aumento do número de componentes principais não parece influenciar na classificação. Em contrapartida, para as músicas de Gioachino Rossini (Figuras 41b e 41d), conforme aumenta-se o número de componentes principais, melhora-se a classificação.

Figura 41 – Taxa de acerto na classificação do Cenário 2 para o classificador k-NN



Fonte: Autor.
Legenda: (a) 5-fold para Johannes Brahms, (b) 5-fold para Gioachino Rossini, (c)10-fold para Johannes Brahms e (d) 10-fold para Gioachino Rossini.

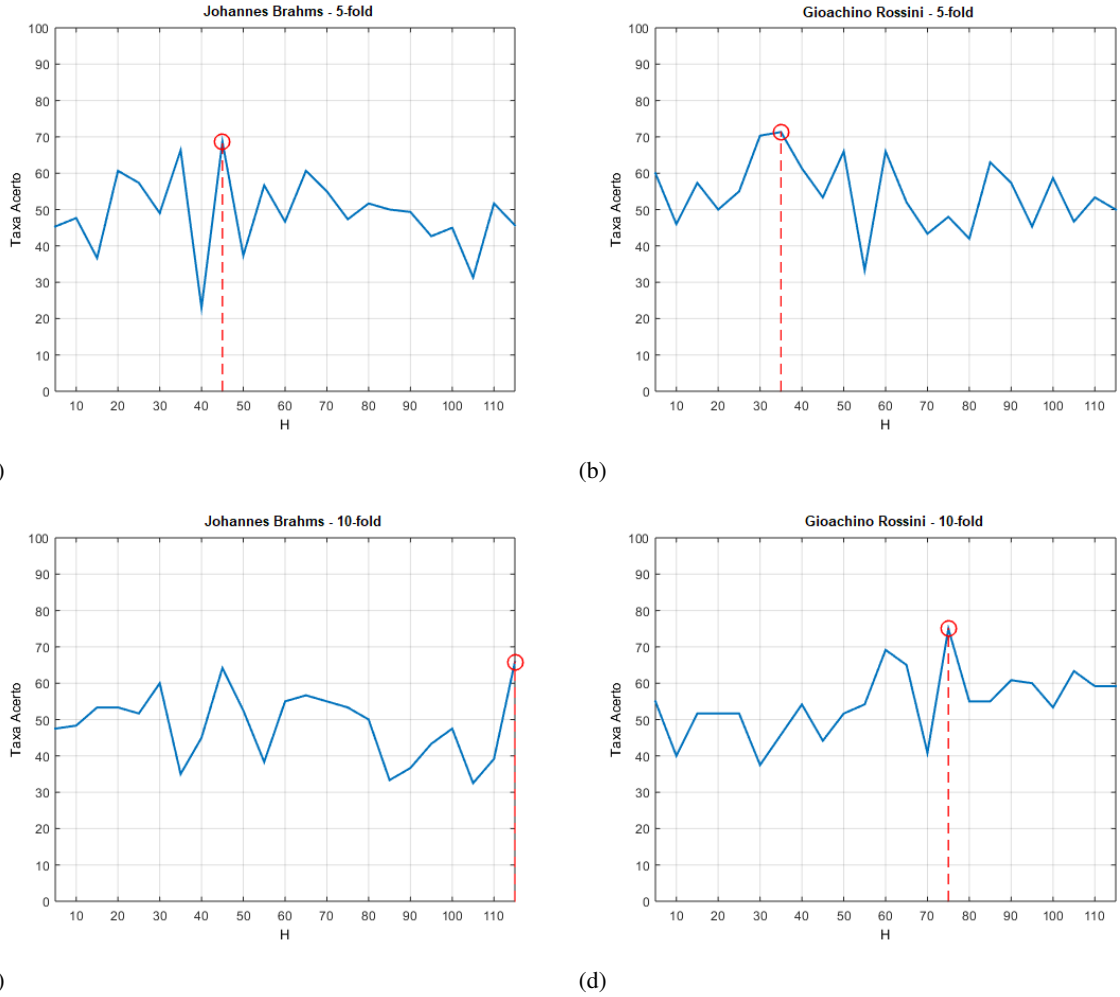
Para a música de Johannes Brahms as 10 primeiras componentes principais representam 81,2% da variância dos dados, alcançando 95% da variância dados na componente 18. Já para a música de Gioachino Rossini, as 10 primeiras componentes principais representam 84,1% da variância dos dados, alcançando 95% da variância na componente 16.

Os parâmetros da rede neural foram determinados de forma semelhante ao Cenário 1, variando o número de H neurônios na camada escondida, tendo as duas saídas correspondendo ao grupo de músicos (1,0) e não-músicos (0,1).

A Figura 42 apresenta a taxa de acerto para cada H onde é possível perceber uma instabilidade na rede conforme varia-se o número de neurônio na camada escondida, atingindo a melhor classificação para a música de Gioachino Rossini, com uma taxa de acerto de 75%.

Na Tabela 9 são apresentados os valores do número de k vizinhos mais próximos que apresentaram melhor taxa de acerto na classificação das duas músicas no k-NN com o valor da taxa de acerto e desvio padrão, e o número H de neurônios na camada escondida da Rede

Figura 42 – Taxa de acerto na classificação do Cenário 2 para a rede neural MLP



Fonte: Autor.
Legenda: (a) 5-fold para Johannes Brahms, (b) 5-fold para Gioachino Rossini, (c)10-fold para Johannes Brahms e (d) 10-fold para Gioachino Rossini.

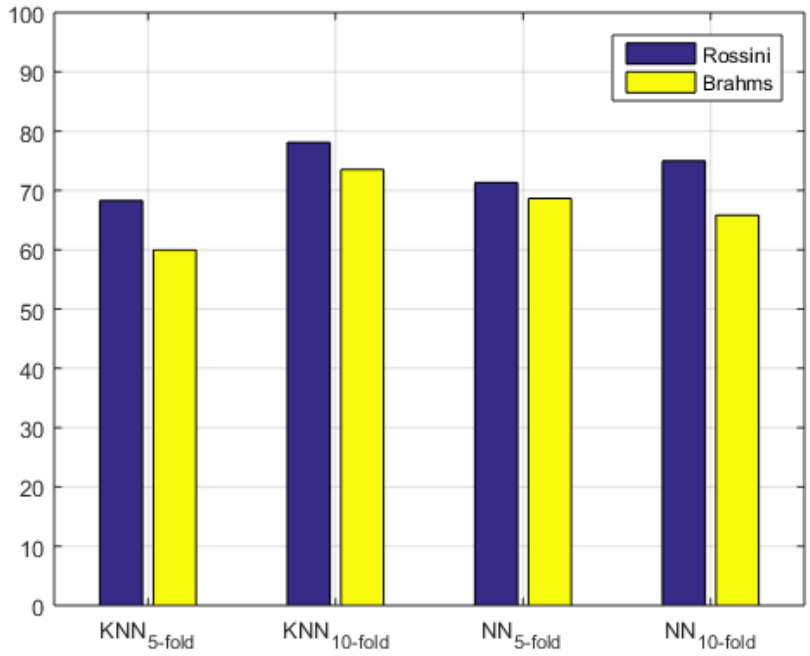
Neural que apresentam melhor taxa de acerto na classificação, apresentando a taxa de acerto e desvio padrão. A Figura 43 apresenta um gráfico de barras com os valores de classificação para os dois classificadores.

Tabela 9 – Resultado da classificação dos voluntários para o Cenário 2 para os classificadores k-NN e a Rede Neural MLP.

		k-NN			Rede Neural		
X-**fold**	**música**	k	Média	Desvio Padrão (%)	H	Média	Desvio Padrão (%)
5-fold	*Rossini*	1	68,2%	8,12	35	71,3%	19,3
	Brahms	3	59,9%	3,10	45	68,7%	22,1
10-fold	*Rossini*	2	78,1%	9,61	75	75%	9,4
	Brahms	3	73,5%	4,09	115	65,8%	22,5

Fonte: Autor.

Figura 43 – Gráfico de Barras, apresentando a classificação dos voluntários para o Cenário 2



Fonte: Autor.

Como é possível observar, o k-NN apresentou, não apenas uma maior taxa de acerto na classificação, como ele se apresentou mais estável do que a rede neural, semelhante ao cenário 1, com o problema de *overfitting* devido ao baixo número de voluntários para comporem a amostra. Em ambos os casos, a música O Barbeiro de Sevilha - Abertura, de Gioachino Rossini, apresentou uma melhor classificação.

Neste caso, o aumento do conjunto de treinamento (X-fold maior) resultou em uma melhora na classificação para o k-NN, porém não apresentou nenhuma mudança significativa no resultado da rede neural.

Portanto, em comparação ao cenário anterior, este segundo cenário apresentou uma classificação inferior, onde não apenas a taxa de acerto na classificação foi inferior, mas o desvio padrão da rede neural foi maior em comparação ao cenário 1.

Ao observar o instante de ocorrência do potencial evocado N100 e P200, assume-se que eles de fato estejam ocorrendo para cada voluntário, porém isso pode não estar ocorrendo devido a uma série de limitações do experimento, como a presença de artefatos durante a ocorrência de triggers, conforme dito anteriormente. Inclusive, na análise anterior dos grupos de voluntários foi revelado que para algumas características acústicas não ocorreram os potenciais evocados, o que influencia na classificação deste segundo cenário. Portanto, para realizar a classificação a partir deste segundo cenário é necessário que ocorram, de fato, os componentes de ERP esperados para todos os voluntários.

Além disso, ao serem observados instantes diferentes para cada voluntário, é possível que estejam sendo analisados processamentos musicais distintos, podendo este ser mais um fator de confusão para o classificador.

Desta forma, para o segundo cenário, música O Barbeiro de Sevilha - Abertura, de Gioachino Rossini apresentou a melhor taxa de acerto na classificação, igual a 78,1%, para o classificador k-NN.

5 CONCLUSÃO

Músicas possuem um conjunto de características que podem ser extraídas e que revelam suas qualidades como o timbre, ritmo, tonalidade, entre outras. Neste trabalho foram analisadas as respostas neurais de um grupo de voluntários ao ouvirem duas músicas (Dança Húngara No 5, de Johannes Brahms, e O Barbeiro de Sevilha - Abertura, de Gioachino Rossini), buscando classificá-los por instantes específicos dentro de suas épocas (Cenário 1), ou por meio dos instantes em que ocorriam os componentes N100 e P200 (Cenário 2), verificados através do sinal resultante da região central do escalpo.

Foram extraídas 12 características acústicas. Esperava-se que houvesse redundância nos vetores de características extraídos e ao realizar a Análise Fatorial chegou-se a conclusão de que apenas 3 características acústicas eram suficientes estatisticamente para descrever as músicas escolhidas. A partir destas características foram definidos os *triggers*, para processamento, análise e classificação do sinal de EEG.

Os voluntários músicos utilizados neste trabalho são amadores, com exceção de um, praticando seus instrumentos por poucas horas durante a semana, mas, apesar disso, ainda assim foi possível diferenciá-los dos não-músicos, com uma taxa de acerto superior a 80% para o classificador k-NN em ambas as músicas.

Classificar os dados de acordo com os pontos mínimos e máximos da região central tem por intuito utilizar o resultado do potencial evocado para diferenciar o grupo de músicos dos não-músicos. Porém, os resultados mostraram que para o conjunto de voluntários selecionados, esta forma de classificação é menos assertiva em comparação ao uso de valores do potencial elétrico obtidos no mesmo instante para todos, de forma determinística.

O classificador k-NN se mostrou melhor que a Rede Neural, por encontrar uma relação entre a distância de cada voluntário no espaço projetado, enquanto que a rede neural apresenta uma outra abordagem, buscando encontrar uma relação não-linear entre os dados. Acredita-se que este comportamento inferior da Rede Neural seja devido ao baixo número de amostras analisadas, que implica no problema de *overfitting* típico deste método nestas situações.

Faz-se importante também discutir algumas limitações quanto ao trabalho apresentado. São elas:

a) O uso da touca de EEG com eletrodos a seco traz maior comodidade ao voluntário e maior flexibilidade na realização dos experimentos. Porém, poucos movimentos do voluntário podem gerar grandes ruídos no sinal e parte dos voluntários tiveram dificuldade em se manter parados;

b) Como as músicas utilizadas foram apenas instrumentais (musicas clássicas), dependendo do gosto do voluntário, ouvi-las se tornou uma tarefa tediosa. Somando isso ao estado físico e mental do voluntário, o experimento pode levá-lo a um estado de relaxamento, sendo necessário removê-lo da análise de dados;

c) Os instantes marcados como *triggers* constituem em estímulos diferentes ao longo da música, mesmo que estejam correspondendo a características iguais. Isso significa que, uma vez tendo passado um determinado *trigger*, ele não necessariamente se repetirá. Portanto, se durante o experimento houver um artefato naquele *trigger* marcado, aquele sinal se perderá na análise, revelando a necessidade de um maior número de *triggers* na análise;

d) Deve ser feito uma compensação entre o número de *triggers* extraídos das músicas e a rigidez na determinação dos parâmetros utilizados para selecioná-los. Quanto mais rígido os parâmetros forem, maior o contraste acústico gerado por ele, porém, menor é o número de *triggers* extraídos;

e) Durante a remoção de artefatos, o número de *triggers* que representa a resposta cortical para a característica acústica observada pode reduzir, e apenas um *trigger* não é suficiente para compor a análise adequadamente, apresentando um sinal ruidoso. Caso uma média global entre os voluntários seja realizada, ela pode ser afetada por este sinal;

f) O número de eletrodos utilizados foram capazes de obter uma boa classificação, porém, utilizando um número maior de eletrodos, espera-se obter uma melhor classificação por haver mais informação sobre a atividade cortical.

A partir desta dissertação, algumas possibilidades podem ser executadas como trabalhos futuros, tais como: Diversificar os estilos das músicas apresentadas como estímulo, aumentar o número de eletrodos para a análise dos sinais cerebrais e aumentar o número de voluntários na análise. Por fim, espera-se que utilizando músicos profissionais a classificação se torne mais assertiva, havendo maior distinção entre os grupos.

REFERÊNCIAS

ABRAMS, Daniel A. et al. Inter-subject synchronization of brain response during natural music listening. **European Journal of Neuroscience**, v. 37, n. 9, p. 1458–1469, maio 2013. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/23578016>. Acesso em: 12 out. 2016.

ALLURI, Vinoo; TOIVIAINEN, Petri. Exploring perceptual and acoustic correlates of polyphonic timbre. **Music Perception**, v. 27, n. 3, p. 223–241, 2010. Disponível em: <https://www.researchgate.net/publication/232767615_Exploring_perceptual_and_acoustic_correlates_of_polyphonic_timbre>. Acesso em: 15 fev. 2017.

ALLURI, Vinoo et al. From Vivaldi to Beatles and back: predicting lateralized brain responses to music. **NeuroImage**, v. 83, p. 627–636, 2013. Disponível em: <https://doi.org/10.1016/j.neuroimage.2013.06.064>. Acesso em 3 jun. 2016.

ALLURI, Vinoo et al. Large-scale brain networks emerge from dynamics processing of musical timbre, key and rhythm. **NeuroImage**, v. 59, p. 3677–3689, 2012. Disponível em: <https://doi.org/10.1016/j.neuroimage.2011.11.019>. Acesso em: 3 jun. 2016.

ANGHINAH, Renato et al. Artefatos biológicos no EEG quantitativo. **Arq. Neuropsiquiatr**, São Paulo, v. 64, n. 2, p. 264–268, 2006.

ANGULO-PERKINS, Arafat et al. Music listening engages specific cortical regions within the temporal lobes: differences between musicians and non-musicians. **Cortex**, v. 59, p. 126–137, 2014. Disponível em: <http://dx.doi.org/10.1016/j.cortex.2014.07.013>. Acesso em: 7 abr. 2016.

ARGOUD, Fernanda Isabel Marques. **Contribuição à automatização da detecção e análise de eventos epileptiformes em eletroencefalograma**. 2001. 185f. TESE (Pós-graduação em Engenharia Elétrica) - Universidade Federal de Santa Catarina, Florianópolis, 2001. Disponível em: <http://repositorio.ufsc.br/xmlui/handle/123456789/82114>. Acesso em: 22 abr. 2016.

BEAR, Mark F.; CONNORS, Barry W.; PARADISO, Michael A. **Neuroscience:** exploring the brain. 3rd. Baltimore: Lippincott Williams e Wilkins, 2006.

BENNET, Alex; BENNET, David. The human knowledge system: music and brain coherence. **Information and knowledge management systems**, Frost, West Virginia, v. 38, n. 3, 2008.

BESSON, Mireille; FAÏTA, Frédérique. An event-related potential (ERP) study of musical expectancy: comparison of musicians with nonmusicians. **Journal of experimental psychology human perception and performance**, v. 21, n. 6, p. 1278–1296, 1995. Disponível em: <http://psycnet.apa.org/record/1996-16299-001>. Acesso em: 3 jun. 2016.

BROSSI, Adriana Bortoleto et al. Verificação das respostas do mismatch negativity (MMN) em sujeitos adultos normais. **Rev. Bras. Otorrinolaringol.**, São Paulo, v. 6, n. 6, p. 793–802, 2007.

BUZSÁKI, György. **Rhythms of the brain**. 1st ed. New York: Oxford University, 2006.

BUZSÁKI, György; ANASTASSIOU, Costas A.; KOCH, Christof. The origin of extracellular field and currents: EEG, ECoG, LFP and spikes. **Nature Review**, v. 13, p. 407–420, 2012. Disponível em: <http://www.nature.com/nrn/journal/v13/n6/full/nrn3241.html?foxtrotcallback=true>. Acesso em: 12 abr. 2016.

COLLURA, Thomas F. The measurement, interpretation and use of EEG frequency bands. **Brain Master Technologies**, 1997. Disponível em: <http://www.stresstherapysolutions.com/pubs/qeeg/TFC+BMrPublications/535-005Measurement,InterpretationandUseofEEGFrequency.pdf>. Acesso em: 22 abr. 2016.

DAYAN, Peter; ABBOTT, L. F. **Theoretical neuroscience:** computational and mathematical modeling of neural systems. Revised ed. Cambridge: The MIT, 2005.

DUFFY, F. H. **Topographic mapping of brain electrical activity: clinical applications and issues**. In: Maurer K. (eds) Topographic Brain Mapping of EEG and Evoked Potentials. Springer, Berlin, Heidelberg, p. 19-52. 1989.

GEORGE, Elyse M.; DONNA, Coch. Music training and working memory: an ERP study. **Neuropsychology**, p. 1083–1094, 2011. Disponível em: <https://doi.org/10.1016/j.neuropsychologia.2011.02.001>. Acesso em: 3 jun. 2016.

HAIR, Joseph F. Jr. et al. **Multivariate data analysis**. 7th ed. Harlow: Pearson, 2014.

HAYKIN, Simon. **Redes neurais:** princípios e prática. 2. ed. Porto Alegre: ARTMED, 2001.

HINRICHS, H.; SIMON, G.; KÜNKEL, H. Methods of spatial interpolation applied to EEG mapping. **Psychiatry Research**, v. 29, p. 431–432, 1989. Disponível em: <http://dx.doi.org/10.1016/0165-1781(89)90113-3>. Acesso em: 2 maio 2017.

ILARI, Beatriz Senoi. **Em busca da mente musical:** ensaios sobre os processos cognitivos em música. 1. ed. Curitiba: UFPR, 2006.

JOHNSON, Richard A.; WICHERN, Dean W. **Applied multivariate statistical analysis**. 6th ed. New Jersey: Pearson, 2007.

KALAN, A. P. J. Abdul. **A. P. P. Abdul Kalan Quotes**. 2015. Disponível em: <http://www.noellesdefiningmoments.com/2015/12/never-stop-fighting.html>. Acesso em: 19 set. 2017.

KAMEL, Nidal; MALIK, Aamir Saeed. **EEG/ERP analysis:** methods and applications. 1st ed. New York: CRC, 2014.

KANDEL, Eric R. et al. **Principles of neural science**. 5th ed. New York: McGraw-Hill, 2013.

KNESS, Peter; SCHEDL, Markus. **Music similarity and retrieval:** an introduction to audio and web based strategies. 1st ed. Heidelberg: Springer, 2016.

KOELSCH, Sterfan et al. Brain indices of music processing: nonmusicians are musical. **Journal of Cognitive Neuroscience**, Massachusetts, v. 12, n. 3, p. 520–541, 2000.

KOLES, Zoly J.; PARANJAPE, R. B. Topographic mapping of the EEG: an examination of accuracy and precision. **Brain Topography**, Alberta, v. 1, n. 2, p. 87–95, 1988.

KREBS, Claudia; WEIBERG, Joanne; AKESSON, Elizabeth. **Lippincotts illustrated reviews: neuroscience**. 1st ed. Baltimore: Lippincott Williams e Wilkins, 2012.

LARTILLOT, Oliver. **MIRtoolbox 1.6.1 users manual**. Aalborg: Department of architecture, design e media technology, 2014. Disponível em: <https: //www.jyu.fi/hytk/fi/laitokset/mutku/en/research/materials/mirtoolbox/MIRtoolbox1.6guide>. Acesso em: 17 nov. 2016.

LEHMANN, Adreas C.; SLOBODA, John A.; WOODY, Robert H. **Psychology for musicians:** understanding and acquiring the skills. 1st ed. New York: Oxford University, 2007.

LEHMANN, D. **From mapping to the analysis and interpretation of EEG/EP maps**. In: Maurer K. (eds) Topographic Brain Mapping of EEG and Evoked Potentials. Springer, Berlin, Heidelberg, p. 53-75. 1989.

LENT, Roberto. **Cem bilhões de neurônios:** conceitos fundamentais de neurociência. 2. ed. Rio de Janeiro: Ateneu, 2001.

LERCH, Alexander. **An introduction to audio content analysis:** applications in signal processing and music informatics. 1st ed. New Jersey: IEEE, 2012.

LIMA, Edirlei Soares. **K-Nearest neighbor**. 2012. Disponível em: <http://edirlei.3dgb.com.br/aulas/ia_2012_1/IA_Aula_16_KNN.pdf>. Acesso em: 3 jul. 2017.

LUCK, Steven J. **An introduction to the event-related potential technique**. 2nd ed. Massachusetts: The MIT, 2014.

MALMIVUO, Jaakko; PLONSEY, Robert. **Bioelectromagnetism:** principles and applications of bioelectric and biomagnetic fields. 1st ed. New York: Oxford University, 1995.

MARCUSE, Lara V.; FIELDS, Madeline C.; YOO, Jiyeoun. **Rowan's primer of EEG**. 2nd ed. New York: Elsevier, 2016.

MONTEIRO, Adriano; MANZOLLI, Jonatas. **Análise de áudio e recuperação da informação musical em um ambiente computacional voltado a improvisação**. 2011. Disponível em: <https://monteiroadriano.wordpress.com/writings/>. Acesso em: 20 out. 2016.

MOURÃO, Carlos Alberto Junior; MELO, Lucilene Bandeira Rodrigues. Integração de três conceitos: função executiva, memória de trabalho e aprendizado. **Psicologia teoria e pesquisa**, Juiz de Fora, v. 23, n. 3, p. 309–314, 2011. Disponível em: <http://www.scielo.br/pdf/ptp/v27n3/06.pdf>. Acesso em: 23 agos. 2017.

MULERT, Christoph; LEMIEUX, Louis. **EEG-fMRI: physiological basis, technique and applications**. 1st ed. Heidelberg: Springer, 2010.

MURRAY, Micah M.; BRUNET, Denis; MICHEL, Christoph M. Topographic ERP analysis: A step-by-step tutorial review. **Brain Topography**, v. 20, p. 249–264, 2008.

MUSZKAT, Mauro; CORREIA, Cleo M. F.; CAMPOS, Sandra M. Música e neurociência. **Neurociências**, v. 8, n. 2, p. 70–75, 2000.

NUNEZ, Paul L.; SRINIVASAN, Ramesh. **Electric fields of the brain:** the neurophysis of EEG. 2nd ed. New York: Oxford University, 2006.

OLEJNICZAK, Piotr. Neurophysiologic basis of EEG. **Journal of Clinical Neurophysiology**, New Orleans, v. 23, n. 3, p. 186–189, 2006.

OLIVEIRA, Marcelo. **Neurônios**. 2006. Disponível em: <http://www.infoescola.com/sistema-nervoso/neuronios/>. Acesso em: 11 jun. 2017.

OPENBCI. **OpenBCI: open source biosensing tools**. 2016. Disponível em: <http://openbci.com>. Acesso em: 2 nov. 2016.

PERETZ, Isabelle; ZATORRE, Robert J. Brain organization for music processing. **Annual Reviews Psychology**, v. 56, p. 89–114, 2004. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/15709930>. Acesso em: 8 jun. 2016.

PERRIN, F. et al. Mapping of scalp potentials by surface spline interpolation. **Eletroencephalography and clinical neurophysiology**, v. 66, p. 75–81, 1987. Disponível em: <https://doi.org/10.1016/0013-4694(87)90141-6>. Acesso em: 2 maio 2017.

PIGEON, Ir. S. **White noise Co.** 2013. Disponível em: <https://mynoise.net/NoiseMachines/whiteNoiseGenerator.php>. Acesso em: 3 jul. 2017.

PLOMP, R.; LEVELT, W. J. M. Exploring perceptual and acoustic correlates of polyphonic timbre. **The journal of the acoustical society of america**, Soesterberg, v. 38, p. 549–560, 1965.

POIKONEN, H. et al. Event-related brain responses while listening to entire pieces of music. **Neuroscience**, v. 312, p. 58–73, 2016. Disponível em: <https://doi.org/10.1016/j.neuroscience.2015.10.061>. Acesso em: 3 jun. 2016.

SARNTHEIN, Johannes et al. Persistent patterns of brain activity: an EEG coherence study of the positive effect of music on spatial-temporal reasoning. **Neurological Research**, v. 19, p. 107–116, 1997. Disponível em: <https: //www.researchgate.net/publication/14043917_Persistent_patterns_of_brain_activity_An_ EEG_coherence_study_of_the_positive_effect_of_music_on_spatial-temporal_reasoning>. Acesso em: 7 abr. 2016.

SETHARES, William A. **Tuning, timbre, spectrum, scale**. 2nd ed. Madison: Springer, 2005.

SHENG-FU, Liang et al. Classification of EEG signals from musicians and non-musicians by neural networks. **World Congress on Intelligent Control and Automation, IEEE**, v. 8, p. 865–869, 2011. Disponível em: <http://ieeexplore.ieee.org/document/5970639/?reload=true>. Acesso em: 21 jul. 2017.

SIVANANDAN, N. Quantitative EEG analysis technique for computerized digital brain signals. **International Journal of Scientific and Research Publications**, Coimbarore, v. 3, 2013. Dispoível em: <http://www.ijsrp.org/research-paper-0613/ijsrp-p18117.pdf>. Acesso em 1 ago. 2017.

SQUIRE, Larry R. et al. **Fundamental neuroscience**. 3rd ed. Massachusetts: Elsevier, 2013.

SUBRAMANIAN, Hariharan; RAO, Preeti; ROY, Sumantra D. **Audio signal classification**. 2004. Disponível em: <https://www.ee.iitb.ac.in/~esgroup/es_mtech04_sem/es_sem04_paper_04307909.pdf>. Acesso em: 1 ago. 2017.

SUR, Shravani; SINHA, V. K. Event-related potential: an overview. **Industrial Psychiatry**, v. 18, n. 1, 2009. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3016705/>. Acesso em: 2 abr. 2017.

USAKLI, Ali Bulent. Improvement of EEG signal acquisition: an electrical aspect for state of the art of front-end. **Computational Intelligence and Neuroscience**, p. 264–268, 2010. Disponível em: <https://www.researchgate.net/publication/285574770_Improving_Measurement_Performance_of_EEG_Signal_Acquisition_An_Electrical_Aspect_for_Front-end>. Acesso em: 22 abr. 2016.

VIRTALA, P. et al. Musicianship facilitates the processing of Western music chords: an ERP and behavioral study. **Neuropsychologia**, v. 61, p. 247–258, 2014. Disponível em: <https://doi.org/10.1016/j.neuropsychologia.2014.06.028>. Acesso em: 7 abr. 2016.

VUSST, Peter et al. The sound of music: differentiating musicians using a fast, musical multi-feature mismatch negativity paradigm. **Neuropsychologia**, v. 61, p. 1432–1443, 2014. Disponível em: <https://doi.org/10.1016/j.neuropsychologia.2012.02.028>. Acesso em: 7 abr. 2016.

WARREN, Jason D. How does the brain process music? **Clinical Medicine**, London, v. 8, n. 1, p. 32–36, 2008.

WECKER, Jonas Edison; SOARES, Marcelo Marques; NEMOS, Douglas Fabiano Lenz. **Aula de anatomia**. 2001. Disponível em: <http://www.auladeanatomia.com/novosite/sistemas/sistema-nervoso/>. Acesso em: 10 dez. 2016.

WOODMAN, Geoffrey F. **A brief introduction to the use of event-related potentials (ERPs) in studies of perception and attention**. 2010. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3816929/>. Acesso em: 2 abr. 2017.

ZULL, James E. The art of changing the brain. **Teaching for meaning**, v. 62, n. 1, p. 68–72, 2004. Disponível em: <http://www.ascd.org/ASCD/pdf/journals/ed_lead/el200409_zull.pdf>. Acesso em: 7 abr. 2016.

APÊNDICE A – TERMO DE CONSENTIMENTO LIVRE ESCLARECIDO

LABORATÓRIO DE PROCESSAMENTO DE IMAGENS
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
CENTRO UNIVERSITÁRIO DA FEI

Pesquisador Responsável: Carlos Eduardo Thomaz
Endereço: Avenida Humberto de Alencar Castelo Branco, 3972, Sala K5-01
CEP: 09.850-901 – São Bernardo do Campo – SP
Fone: (11) 4353-2910 ramal: 2183 - E-mail: cet@fei.edu.br

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Projeto de pesquisa: "Predição de musicalidade por meio da análise de sinais de EEG". O objetivo do presente projeto é classificar os voluntários em músicos e não-músicos por meio das respostas neurais induzidas por músicas, a partir de características acústicas extraídas de cada áudio.

Você está sendo convidado(a) a participar do projeto de pesquisa citado. Para este estudo, ouvirá algumas músicas enquanto o seu encefalograma é registrado.

O eletroencefalograma é um exame não invasivo e que não apresenta risco para o indivíduo. É colocada uma touca com eletrodos a seco, distribuídos em várias regiões da cabeça, para realizar o registro da atividade cerebral.

Nesta pesquisa a aquisição do eletroencefalograma não tem objetivo diagnóstico. Seu nome ou o material que indique a sua pessoal participação não serão liberados a terceiros. O(A) Sr(a). não será identificado(a) em nenhuma publicação que possa resultar deste estudo.

Para participar deste estudo você não terá nenhum custo, nem receberá qualquer vantagem financeira. Você será instruído(a) sobre o estudo em qualquer aspecto que desejar e estará livre para participar ou recusar-se a participar. Poderá retirar seu consentimento ou interromper a participação a qualquer momento.

Este termo de consentimento encontra-se impresso em duas vias, sendo que uma cópia será arquivada pelo pesquisador responsável, no Laboratório de Processamento de Imagens do Departamento de Engenharia Elétrica do Centro Universitário da FEI em São Bernardo do Campo, e a outra será fornecida a você.

Eu, _____, portador do documento de Identidade_____, concordo de livre e espontânea vontade em participar como voluntário(a) do projeto supra-citado. Declaro que fui informado(a) dos objetivos do estudo de maneira clara e detalhada e esclareci minhas dúvidas. Recebi uma cópia deste termo de consentimento livre e esclarecido e me foi dada a oportunidade de ler e esclarecer as minhas dúvidas.

São Bernardo do Campo, _____ de _____ de ____.

_____ _____

Assinatura participante Assinatura pesquisador