UNIVERSIDADE DE SÃO PAULO
ESCOLA DE COMUNICAÇÕES E ARTES

Pedro Paulo Köhler Bondesan dos Santos

PERCEPÇÃO MÉTRICA:
ESTUDANDO A PERCEPÇÃO DO RITMO MUSICAL
ATRAVÉS DE EXPERIMENTOS PSICOFÍSICOS

São Paulo
2017

Pedro Paulo Köhler Bondesan dos Santos

PERCEPÇÃO MÉTRICA: ESTUDANDO A PERCEPÇÃO DO RITMO MUSICAL ATRAVÉS DE EXPERIMENTOS PSICOFÍSICOS

Tese apresentada ao Programa de Pós-Graduação em Música, da Universidade de São Paulo, para obtenção do título de Doutor em Música. (versão corrigida)

Orientadora: Profa. Dra. Adriana Lopes da Cunha Moreira.

Co-orientador: Prof. Dr. Ronald Dennis P. K. C. Ranvaud

São Paulo
2017

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Catálogo na Publicação
Serviço de Biblioteca e Documentação
Escola de Comunicações e Artes da Universidade de São Paulo
Dados fornecidos pelo(a) autor(a)

Bondesan dos Santos, Pedro Paulo Köhler
Percepção métrica: Estudando a percepção do ritmo musical
através de experimentos psicofísicos / Pedro Paulo Köhler
Bondesan dos Santos. -- São Paulo: P. P. K. Bondesan dos
Santos, 2017.
100 p.: il.

Tese (Doutorado) - Programa de Pós-Graduação em Música -
Escola de Comunicações e Artes / Universidade de São Paulo.
Orientadora: Adriana Lopes da Cunha Moreira
Coorientador: Ronald Dennis Paul Kenneth Clive Ranvaud
Bibliografia
1. ritmo musical 2. entrainment 3. Ambiguidade rítmica 4.
percepção de música 5. experimentos psicofísicos I. Moreira,
Adriana Lopes da Cunha II.
Título.

CDD 21.ed. - 780

DEDICATÓRIA

Aos meus filhos,

Julia Mendonça Bondesan dos Santos e

Lorenzo Mendonça Bondesan dos Santos

AGRADECIMENTOS

Profa. Dra. Adriana Lopes Moreira da Cunha, Prof. Dr. Ronald Ranvaud (ICB/USP), Prof. Dr. Flo Menezes (Estúdio PANaroma – UNESP), Shiro Takashima (programação do script Psychtoolbox-3), Prof. Dr. Pedro Paulo Salles (USP), Prof. Dr. Fernando Iazzetta (USP), Prof. Dr. Marcelo Queiroz (IME-USP), Profa. Dra. Patricia Vanzela (Neurociências – UFABC), Profa. Dra. Graziela Bortz (UNESP), Daniel Avilez (Estúdio PANaroma – UNESP)

Pelo incentivo e pela força, Valeria Pinheiro Mendonça, Ricardo Barini, Vane Sanches Barini, Ana Sanches Barini, Caio Sanches Barini, Cecilia Köhler, Ciro Bondesan dos Santos, Gustavo Köhler Bondesan dos Santos, Wildicelia Melhem e João Catarino.

Pelas conversas inspiradoras e conselhos, Dr. Paulo Gonçalves Silva Filho, Dr. Antônio Miguel Vieira Monteiro (INPE), Me. Carlos Eduardo Pedrasse e Luciana Rathsam.

apoio:

Studio PANaroma
de Música Eletroacústica da Unesp

Diretor Artístico: Flo Menezes

Epígrafe

*“.... Tempo rei, ó tempo rei, ó tempo rei
Transformai as velhas formas do viver
Ensinai-me, ó Pai, o que eu ainda não sei
Mãe Senhora do Perpétuo socorrei...”*

Tempo Rei, Gilberto Gil (1984)

Resumo

BONDESAN DOS SANTOS, Pedro Paulo Köhler. Percepção Métrica: Estudando a percepção do ritmo musical através de experimentos psicofísicos. 2017. 108 f. Tese (Doutorado)– Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo, 2017.

Nesta tese de doutorado abordamos modelos cognitivos de percepção da métrica musical e *entrainment* a partir de questões musicológicas, como a ambiguidade métrica decorrente de exemplos da literatura. Partindo de uma verificação do estado da arte em pesquisas rítmicas que envolvem o estudo de anacruses, realizamos um percurso experimental que investiga a efetividade da chamada *percepção da acentuação subjetiva* revelada por Povel e Okkerman (1981), por acreditarmos que o fenômeno da acentuação subjetiva esteja envolvido na desambiguação da percepção de referenciais métricos dúbios. Para tanto, desenvolvemos uma metodologia de quantificação das similaridades entre os padrões de acentuação coletados em grupo universitário da cidade de São Paulo e os padrões de referência da literatura, sobretudo de Povel e Essens (1985). Estes experimentos revelaram que há uma tendência significativa à percepção da acentuação subjetiva predominantemente em grupo sem estudo formal de música. Por outro lado, os estudantes de música revelaram uma tendência de acentuação mais relacionada à pulsação musical.

Palavras-chave: Ritmo musical. *Entrainment*. Ambiguidade rítmica. Percepção da música. Experimentos psicofísicos.

Abstract

BONDESAN DOS SANTOS, Pedro Paulo Köhler. Beat Perception: Studying the musical rhythm perception through psychophysical experiments. 2017.108 f. Thesis (Ph.D. in Music) – Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo, 2017.

In this doctoral thesis we address cognitive models of perception of musical meter and entrainment from musicological issues, such as the metric ambiguity arising from examples of literature. Based on a state-of-the-art check on rhythmic researches involving the study of anacrusis, we conducted an experimental study that investigates the effectiveness of the so-called subjective accent revealed by Povel and Okkerman (1981), because we believe that the phenomenon of subjective accent is involved in the disambiguation of dubious beat references perception. Therefore, we developed a methodology to quantify the similarities between the accentuation patterns collected in university group of São Paulo and the literature reference standards, especially Povel and Essens (1985). These experiments revealed that there is a significant tendency to subjective perception of accent predominantly in people without formal music study. On the other hand, the students of music revealed a tendency of accentuation more related to the musical beat.

Keywords: Musical rhythm. Entrainment. Rhythmic ambiguity. Perception of music. Psychophysical experiments.

Lista de Figuras

- Figura 1 - Beethoven, Sinfonia n. 5, I (comp. 1-5).
- Figura 2 - Modelo representativo da ideia de ressonância neural, onde um estímulo externo provocaria a emergência de osciladores internos em proporções de harmônicos, perfazendo analogamente os níveis métricos musicais, e promovendo o que Jones (2009, p. 84) denomina vinculação métrica (metric binding)
- Figura 3 - Um oscilador sincronizando com um ritmo isócrono. Os componentes básicos são o período, o pulso e a fase. À medida que o evento confirma a expectativa ou não, o foco atencional se estreita ou se alarga. (JONES, 2009, p. 85)
- Figura 4 - Observa-se a mesma melodia com duas interpretações e níveis métricos distintos: a de cima, com interpretação anacrúsica (acento métrico fora de fase em relação ao início da melodia) e a de baixo com interpretação tética (em fase com a melodia).
- Figura 5 - Imagem representando um Potencial relacionado a evento genérico.
- Figura 6 - Modelo de etapas Hinkelman e Kempthorne à esquerda e o modelo de Gil à direita
- Figura 7 - Conversão das acentuações em sequências de “zeros e uns” coletadas dos participantes através do computador e processada em planilha Excel, com a fórmula de similaridade citada
- Figura 8 - Imagens correspondentes às telas do experimento da Fase 1 com grupos de sons isócronos equitonais de quatro elementos e nomeados SEQ4. A tela da esquerda mostra as quatro barras, representando cada elemento do grupo e na imagem da direita as barras preenchidas parcialmente em branco, com o nível de acentuação relativa assinalada por um participante, no caso específico com os dois primeiros sons não acentuados e os dois últimos igualmente acentuados.
- Figura 9 - Comparação dos resultados obtidos entre dois participantes, R9 e R11, respectivamente, mostram sua avaliação do nível da acentuação relativa percebida (no eixo das ordenadas y), referente a cada um dos 4 elementos do grupo (no eixo x), representados pelos números de 1 a 4.
- Figura 10 - Respostas dos *trials* do participante 9 representadas pelas linhas finas e a linha grossa mostrando a média aritmética do conjunto.
- Figura 11 - Resultados (catch trial) nos retângulos da terceira e da sexta linha da tabela acima indicam grande foco atencional, próximo de 100%.
- Figura 12 - Resultado de experimento com estímulos de 4 elementos isócronos equitonais agrupados, mostrando a acentuação percebida pelo participante.
- Figura 13 - Resultado de experimento com estímulos de 3 elementos isócronos equitonais agrupados.
- Figura 14 - Exemplo da tela da interface linear circular contendo o estímulo E8, correspondente a representação linear da Figura 16
- Figura 15 - Estímulos-exemplos das categorias retirados do artigo de Povel e Essens p.420 e utilizados no nosso experimento.
- Figura 16 - Estímulo-exemplo retirado do artigo de Povel e Essens p.416 e utilizado no nosso experimento com o nome E8.

- Figura 17 - Resultados dos índices de similaridade (eixo y) do estímulo E1, obtidos individualmente por cada Participante (legendas coloridas) relativos a cada tentativa T1 a T4 (eixo x)
- Figura 18 - Resultados dos índices de similaridade (eixo y) do grupo todo acima e dos sub grupos dos alunos de audiovisual (sem nenhum estudo de música no gráfico do meio, e com estudo informal de música no gráfico abaixo), obtidos individualmente por cada Participante (legendas coloridas).
- Figura 19 - Tela demonstrativa da interface linear da Fase 2 do experimento, onde se pode ver a representação visual do estímulo correspondente àquele da Figura 17, nomeado E8 no nosso experimento.
- Figura 20 - Gráfico comparativo das médias de similaridade (eixo y) por Participantes estudantes de música Fase 3 (legendas coloridas) em relação aos estímulos (eixo x).
- Figura 21 - Gráfico comparativo das médias de similaridade (eixo y) por Participantes do Curso de Audiovisual da Fase 2 (legendas coloridas) em relação aos estímulos (eixo x).
- Figura 22 - Gráfico comparativo das médias de similaridade (eixo y) por Participantes do Curso de Audiovisual Fase 3 (legendas coloridas) em relação aos estímulos (eixo x).
- Figura 23 - As barras coloridas significam os índices de percepção de acentuação (eixo y) calculados para cada grupo (legendas coloridas) e referentes a cada evento (eixo x) dos agrupamentos das sequências. O asterisco ao lado do elemento da sequência significa a coincidência do elemento com uma pulsação.
- Figura 24 - As barras coloridas significam os índices de percepção de acentuação dos estímulos E5, E6 e E8 (eixo y) calculados para cada grupo (legendas coloridas) e referentes a cada evento (eixo x) dos agrupamentos das sequências. O asterisco ao lado do elemento da sequência significa a coincidência do elemento com uma pulsação.
- Figura 25 - Resultados dos índices de similaridade (eixo y) dos sub grupos dos alunos de audiovisual (sem nenhum estudo de música no gráfico de cima, e com estudo informal de música no gráfico abaixo), obtidos individualmente por cada Participante (legendas coloridas).
- Figura 26 - Comparativo entre os padrões de acentuação dos participantes de P1 a P4 e de P5 a P9, mostrando a inversão de tendências de acentuação para o estímulo E4, tanto no primeiro grupo D1* D2, quanto no segundo grupo. Neles, a tendência para acentuação em D2, se inverte para acentuação maior em D1.
- Figura 27 - Comparativo entre os padrões de acentuação dos participantes de P1 a P4 e de P5 a P9, mostrando a inversão de tendências de acentuação para o estímulo E8, grupo D1 D2 ao final da sequência . Nele, a tendência para acentuação em D2, se inverte para acentuação maior em D1.
- Figura 28 - Comparativo entre os padrões de acentuação dos participantes de P1 a P4 e de P5 a P9, mostrando a inversão de tendências de acentuação para o estímulo E9, no grupo D1 D2. Nele, a tendência para acentuação em D2, se inverte para acentuação maior em D1.

Figura 29 - Resultados do Participante R9 – Fase 1. Os retângulos representam cada condição de Intervalo Inter Onset - IOIs de 200ms a 800ms, com grupos de 4 elementos, onde cada numero inteiro de 1 a 4 do eixo y representa respectivamente cada elemento do grupo, e no eixo x, temos o índice de acentuação relativa marcado pelo participante. (catch trial) nos retângulos da terceira linha da tabela acima indicam grande foco atencional, próximo de 100%.

Figura 30 - Resultados do Participante R11 – Fase 1. Os retângulos representam cada condição de Intervalo Inter Onset - IOIs de 200ms a 800ms, com grupos de 4 elementos, onde cada numero inteiro de 1 a 4 do eixo y representa respectivamente cada elemento do grupo, e no eixo x, temos o índice de acentuação relativa marcado pelo participante. (catch trial) nos retângulos da terceira linha da tabela acima indicam grande foco atencional, próximo de 100%.

Figura 31 - Resultados dos índices de similaridade (eixo y) dos estímulos de 1 a 4, obtidos individualmente por cada Participante (legendas coloridas) relativos a cada tentativa T1 a T4 (eixo x)

Figura 32 - Resultados dos índices de similaridade (eixo y) dos estímulos de 5 a 9, obtidos individualmente por cada Participante (legendas coloridas) relativos a cada tentativa T1 a T4 (eixo x)

Figura 33 - Gráfico comparativo das médias de similaridade (eixo y) por Participantes da Fase 3 (legendas coloridas) em relação aos estímulos (eixo x). Acima os nove Participantes estudantes do Curso de Audiovisual, e abaixo os seis Participantes do Curso de Música a UNESP.

Figura 34 - As barras coloridas significam os índices de percepção de acentuação dos estímulos E1, E2 (eixo y) calculados para cada grupo (legendas coloridas) e referentes a cada evento (eixo x) dos agrupamentos das sequências. O asterisco ao lado do elemento da sequência significa a coincidência do elemento com uma pulsação.

Figura 35 - As barras coloridas significam os índices de percepção de acentuação dos estímulos E3 a E6 (eixo y) calculados para cada grupo (legendas coloridas) e referentes a cada evento (eixo x) dos agrupamentos das sequências. O asterisco ao lado do elemento da sequência significa a coincidência do elemento com uma pulsação.

Figura 36 - As barras coloridas significam os índices de percepção de acentuação dos estímulos E7 a E9 (eixo y) calculados para cada grupo (legendas coloridas) e referentes a cada evento (eixo x) dos agrupamentos das sequências. O asterisco ao lado do elemento da sequência significa a coincidência do elemento com uma pulsação.

Lista de Tabelas

Tabela 1 - Médias dos índices de percepção dos participantes não músicos da Fase 3 de P1 a P9, obtidos em cada estímulos de E1 a E9, e referente ao gráfico da Figura 19. Em vermelho os índices acima de 50%.

Tabela 2 - Percentual acentuações no último elemento do grupo, sendo que a Fase 3 neste ponto, tinha 4 participantes (P1 a P4) .

Tabela 3 - Percentual acentuações no último elemento do grupo, contabilizando a inclusão de cinco participantes não músicos, na Fase 3.

Tabela 4 - Resultados para os padrões de acentuações mais percebidas por cada grupo. Em vermelho os resultados iguais ou maiores a 0,50.

Lista de abreviaturas

E1, E2 E9 – Estímulos utilizados nos experimentos.

EP – inicialmente nomeada *Evoked Potentials* (EP) ou Potenciais Evocados, os Potenciais Relacionadas a Eventos ou *Event Related Potentials* (ERP) porque são potenciais elétricos evocados por estímulos (em oposição aos ritmos espontâneos EEG). (LUCK, 2005, p. 6)

ERF (*Event-Related Field*) - campos magnéticos relacionados a eventos, é o equivalente aos ERP coletados por de Magnetoencefalografia

ERP (*Event-Related Potentials*) – o mesmo que Potenciais Evocados

EEG – Eletroencefalografia

fMRI –(*Funcional Magnetic Ressonance Imaging*) Ressonância Magnética Funcional

IOI - (*Inter-Onset-Intervals*) intervalo entre os beeps de uma sequência

P1, P2 P9 – Participantes dos experimentos

PET – (*Positron Emission Tomography*), Tomografia por emissão de pósitrons

PF - Ponto de fixação visual:

RP- Regras de Preferencia

SEQ3 - sequência de 3 beeps.

SEQ4 - sequência de 4 beeps.

SS-EP – (*Steady-State Evoked Potential*) ou Potenciais Evocados de Estado Estável

SUMÁRIO

Dedicatória	3
Agradecimentos	4
Epígrafe	5
Resumo.....	6
Abstract.....	7
Lista de Figuras.....	8
Lista de abreviaturas.....	11
Sumário	12
Introdução	14
Capítulo 1	25
Percepção de regularidades temporais, Sincronia adaptativa (<i>entrainment</i>) e Modelos cognitivos de percepção métrica	25
1.1 Percepção de regularidades temporais	25
1.1.1 Acentuação.....	25
1.2 <i>Entrainment</i> e percepção métrica.....	29
1.2.1 Histórico das pesquisas sobre sincronização (<i>entrainment</i>)	33
1.3 O Modelo de Regras de Preferência (<i>Preference Rule-Based Model</i> , PRM) ou Teoria da Codificação (<i>Encoding Theory</i>).....	36
1.4 Teoria da Assistência Dinâmica, DAT.....	38
1.5 Teoria de ressonância neural da métrica musical	42
Capítulo 2	45
O estado da arte para a nossa pesquisa.....	45
2.1 Duas abordagens de pesquisa experimental em percepção rítmica	46
2.1.1 Infraestrutura experimental	53
Capítulo 3	58
Percurso experimental	58
3.1 <i>Design</i> experimental	58
3.2 Metodologia de medição das respostas coletadas nos Experimentos	63
3.3 Experimentos comportamentais.....	68
3.3.1 Fase 1: Experimento exploratório 1, com estruturas de 4 elementos	69
3.3.2 Fase 1: Experimento exploratório 2, com estruturas de 4 elementos e de 3 elementos	75

Fase 1: Análise dos resultados dos experimentos exploratórios.....	79
3.3.3 Fase 2: Experimento com interface circular.....	80
3.3.3.1 Materiais e Métodos	81
3.3.3.1.1 Procedimento.....	81
3.3.3.1.2 Estímulos.....	82
3.3.3.1.3 Medidas.....	84
3.3.3.2 Resultados	84
3.3.3.2.1 Discussão.....	84
3.3.3.2.2 Conclusão	87
3.3.4 Fase 3: Experimento com interface linear	87
3.3.4.1 Materiais e Métodos.....	88
3.3.4.1.1 Procedimento.....	88
3.3.4.1.2 Estímulos.....	90
3.3.4.1.3 Medidas.....	90
3.3.4.1.4 Discussão.....	90
3.3.4.1.5 Conclusão	93
Conclusão	100
Bibliografia Completa	108
ANEXO I – Gráficos dos resultados obtidos.....	111
Glossário.....	120

Introdução

O caminho natural desta pesquisa, originado de nossa dissertação de mestrado sobre a percepção do ritmo musical e seus respectivos modelos cognitivos aplicados à percepção de alguns exemplos da literatura da música erudita e popular, passa pelo experimentalismo científico com suas codificações, seu vocabulário diferenciado, bem como suas metodologias próprias e estabelecidas.

Isto por si representa um grande desafio para este trabalho. Construir vínculos consistentes entre a ciência em sua busca pela medição dos fenômenos psicofísicos no caso desta pesquisa, e a arte, ao incluir um estudo voltado para a percepção do ritmo no fenômeno musical e seus recortes no interior de toda complexidade das eventuais ambiguidades presentes nas obras dos grandes criadores e que lhes confere muitas vezes valor de grande arte.

Em outras palavras, esta tese tem como um dos objetivos gerais aproximar a observação científica da compreensão do fenômeno artístico.

A experimentação no campo da ciência visa muitas vezes confirmar hipóteses, sugerir e testar modelos e quando os modelos são aplicáveis universalmente, ganham o status de leis científicas. Já a experimentação no campo artístico não busca a confirmação do esperado; pelo contrário, a busca artística tem como meta muitas vezes descobrir o inusitado, muitas vezes também as ambiguidades residentes nas fronteiras da percepção e dos significados, ou das interpretações possíveis da percepção.

A dificuldade reside nisso, e apesar de tudo o que estamos tentando fazer para entender um único aspecto da rítmica e da métrica musical, esses aspectos ainda estão dissociados da arte musical em si. E aqui cabe a metáfora da medicina que estuda partes do corpo dissecando cadáveres, em contraste com outra medicina que estuda o corpo em pleno funcionamento.

A tecnologia atual, já se aproxima da possibilidade de observar em tempo real o pleno funcionamento do cérebro e suas reações frente a estímulos propostos. As técnicas de EEG (Eletroencefalografia) apresentam hoje a rapidez suficiente para se observar, com precisão, respostas da ordem de milissegundos, em 128 canais simultâneos. Esse avanço permite mapear respostas relativas a estudos de comportamentos ligados a fenômenos temporais, como o ritmo musical. Mesmo assim a “precariedade” do experimento científico atual ainda não consegue estudar a assim chamada música ecologicamente válida¹ de uma maneira mais holística, embora mesmo que os ganhos de conhecimento no campo musical ainda sejam restritos, os ganhos para a neurociência e a compreensão do funcionamento do cérebro são inegáveis e importantes.

Nosso problema central reside em investigar possíveis fatores envolvidos na percepção de ambiguidades métricas encontradas na literatura musical, e analisadas do ponto de vista dos mecanismos e modelos da cognição e da neurociência em nossa dissertação de mestrado (BONDESAN DOS SANTOS, 2012), retornaremos a esse assunto mais adiante. Entretanto, a investigação experimental utilizando estímulos musicais, no sentido de estímulos ecologicamente válidos, o que quer dizer utilizar gravações de música de verdade, ainda é tão complexo como a compreensão dos processos cerebrais em situações reais. Os equipamentos de estudo ainda requerem ambientes e situações altamente artificiais para o registro do funcionamento do cérebro em tempo real.

Estes obstáculos citados, juntamente com a aquisição da expertise dos protocolos científicos para a realização da pesquisa experimental, apresentaram-se de modo que precisamos tomar como ponto de partida um percurso que alia os ganhos de aprendizado experimental à geração de conhecimento consistente para as questões levantadas. No caso das ambiguidades, sejam visuais, sejam acústicas, os processos de desambiguação, por assim dizer, passam inevitavelmente pelo estabelecimento de um referencial subjetivo. Nesse sentido, tratamos a investigação sobre o efeito da

¹ A expressão “Música ecologicamente válida” refere-se a utilização de estímulos efetivamente musicais. Em condições experimentais atualmente estudamos apenas um aspecto, o ritmo no nosso caso, utilizando estímulos que estão muito distanciados da realidade musical.

acentuação subjetiva como uma possível fonte de conhecimento a respeito do referencial subjetivo adotado.

As regras de preferência na acentuação de agrupamentos rítmicos, lastreadas pelo efeito da acentuação subjetiva (POVEL; OKKERMAN, 1981), não funcionam como lei universal ainda, pois não existe pesquisa confirmatória desses efeitos para a grande maioria dos indivíduos e em culturas diferentes.

Ambiguidades métricas e as questões da pesquisa

Nossa motivação inicial para o desenvolvimento desta pesquisa foi o questionamento relativo às possibilidades de interpretação da percepção da métrica no motivo inicial da *Sinfonia n. 5* de Beethoven (Figura 1). Havia indícios para indagar sobre a razão da aparente ambiguidade surgida da comparação entre o modo proposto pelo compositor e o modo como percebemos o ritmo dos compassos iniciais do primeiro movimento da maneira utilizada pelo compositor na estruturação de seu pensamento musical.

Beethoven propõe uma aparente dúvida sobre a métrica do motivo inicial nos primeiros cinco compassos (é um metro binário ou ternário?) e o recontextualiza nos compassos seguintes propondo a partir dali a percepção do motivo claramente em métrica binária.

Figura 1 - Beethoven, Sinfonia n. 5, I (comp. 1-5).

Nesta abordagem, realizada na nossa dissertação de mestrado (BONDESAN DOS SANTOS, 2012), lançamos mão de conceitos que serão explicitados no capítulo 1, envolvendo a utilização das Regras de Preferência e do fenômeno da

acentuação subjetiva em sequências equitonais, na tentativa de explicar razões da possibilidade da percepção da métrica ternária sobre um motivo grafado de modo binário pelo compositor.

Analizamos diversas e conhecidas gravações da obra, sobre as quais já havia questionamentos sobre a dificuldade da correta interpretação do ritmo da escrita do motivo principal nos seus primeiros compassos por parte de diversos maestros (SCHULLER, 1997, p. 122). Este era um indicativo de que havia problemas relacionando a interpretação da partitura à percepção do seu resultado sonoro. Neste sentido, seria possível pensar na existência de algum aspecto determinante na maneira de processar a métrica musical no início das obras musicais. Poderia ser algo já contido no sinal acústico, ou inferido e desambiguado pelo processamento cognitivo em situações onde o sinal acústico não oferece claramente as referências métricas.

Este questionamento nos levou à realização de experimentos que avaliassem a percepção do referido motivo. Esses primeiros experimentos vieram a constituir a Fase 1 dos nossos experimentos, a qual será apresentada mais adiante no presente trabalho. Em consequência dos resultados obtidos e da discussão gerada pelos mesmos, outros questionamentos se seguiram.

O primeiro, de natureza mais conceitual, refere-se ao nível epistemológico da questão da ambiguidade. Em princípio, a representação escrita da música é altamente simbólica, portanto, considerada de alto nível, ou processo *top-down*, uma vez que tem como referência o processamento cognitivo da informação ali contida. Os processos *top-down* operam junto à informação sensorial, transformando-a em simbólica, assim como ocorre com os conhecimentos e as expectativas, que podem vir a influenciar nossa percepção.

Já a percepção e o processamento dos sinais acústicos são considerados de baixo nível, ou processos *bottom-up*. A informação de baixo nível é usada para construir e estabelecer informações de alto nível.

Considerando os dois estilos de processamento, *top-down* e *bottom-up*, da

maneira como foi exposto, por se manifestar apenas no momento em que o evento acústico ocorre, ou seja como um aspecto diretamente ligado à percepção, poderíamos pensar na possibilidade da acentuação subjetiva ser um processamento inicial do tipo *bottom-up*, e o processamento *top-down* como advindo de um treinamento do tipo percepção do *beat* por músicos com estudo formal de música.

Por outro lado, observa-se que a intencionalidade da interpretação da pulsação modifica os sinais psicofísicos, como demonstrado no experimento de Nozaradan (Exploring how musical rhythm entrains brain activity with electroencephalogram frequency-tagging, 2014, p. 4), indicando a evidência, coletada em EEG², da interação entre processamentos de alto nível alterando a interpretação sobre outros de baixo nível.

Sendo assim, seria possível a modificação da percepção da acentuação de acordo com a intencionalidade aprendida no treinamento musical? Em outras palavras, seria possível haver essa acentuação através do condicionamento do músico ao tempo métrica musical? O condicionamento levaria a uma busca preferencial pelo referencial do *beat*?

Esta possibilidade já havia sido aventada nos experimentos de Povel e Essens (1985), onde questionavam a atuação de acentuações locais (resultantes das regras de preferência) em concorrência com acentuações globais ligadas a métrica e ao *beat*, que eles chamaram de “indução de um relógio interno”. Para eles, haveria a tendência de participantes não qualificados (musicalmente) confundirem acentos locais com aqueles produzidos pela indução da pulsação., resultando em uma reprodução degradada do estímulo proposto (POVEL; ESSENS, 1985, p. 429-431). Este experimento não se propunha a questionar a percepção dos acentos relacionados aos estímulos, mas sim medir a quantidade de acertos na reprodução dos estímulos pelos participantes, instados a ouvir os estímulos e reproduzi-los o mais fielmente possível, acionando um botão que registrava performance rítmica.

Ainda sobre a questão epistemológica da ambiguidade, temos a consideração

² EEG, Eletroencefalografia.

de que o mapeamento de eventos rítmicos numa hierarquia é ambíguo e dependente de contexto prévio, sendo que uma determinada estrutura rítmica, precedida de uma métrica binária ou ternária pode ser categorizada de diferentes maneiras (THOMPSON-SCHILL; HAGOORT, *et al.*, 2013, p. 294), em processo *top-down* influenciado também por expectativas construídas pela exposição anterior à música, além do contexto prévio já citado. Existem evidências experimentais que suportam estas idéias, porém o objeto da nossa investigação difere das mesmas, porque nossa abordagem trata de questões referentes ao *início de obras musicais*, onde portanto, não há contexto musical prévio ainda, uma vez que o ouvinte conta apenas com as influências de expectativas construídas pela sua vivência anterior com a música.

Trataremos, a seguir, de duas visões sobre os níveis de processamento cognitivo da música que são importantes para podermos localizar adequadamente nosso objeto.

A primeira está relacionada ao ciclo percepção-ação na perspectiva de Seifert et al. (2013, p. 211-212), que cita os níveis relacionados por Scherer (2009) para tratar do processamento cognitivo das emoções:

1. Um baixo nível sensório-motor com um mecanismo de avaliação da correspondência de padrões, que é em grande parte determinada geneticamente, utilizando critérios que consistem em modelos apropriados.
2. Um nível esquemático, baseado em traços de memória de processos individuais de aprendizagem, sob a forma de modelos uniformes e holísticos e ocorrendo de forma bastante automática e inconsciente.
3. Um nível de associação, baseado em associações multiformes espalhando e envolvendo várias áreas de associação cortical. Isso pode ocorrer automática e inconscientemente ou de forma deliberada e consciente, dependendo do investimento atencional.
4. Um nível conceitual, envolvendo conhecimento proposicional e sistemas de significado representacional culturalmente moldados, provavelmente exigindo consciência e cálculos de esforço em áreas corticais pré-frontais.

Este modelo trata de processos subjacentes que, de acordo com Seifert et al. (2013, p. 211-212)

[...] ocorram muito provavelmente em diferentes níveis de processamento da informação (por exemplo, sensório-motor, esquemático, associativo, conceitual) e que alguns critérios de avaliação, principalmente

impulsionados por características de estímulo intrínseco, podem ser processadas mais facilmente e mais rapidamente em níveis mais baixos, enquanto que as verificações envolvendo inferência externa requerem processamento em níveis mais elevados.³

O processamento do ritmo parece ter em comum, principalmente, os níveis 1 e 3 de Seifert como o mecanismo sensório-motor de avaliação de correspondência de padrões em baixo nível e um nível associativo, para aqueles que não têm uma formação musical específica. Já o nível 4, conceitual, culturalmente moldado por uma educação formal, deve ser bem comum naqueles cuja formação musical criou raízes.

Na segunda visão, Koelsch (Neural Correlates of Music Perception, 2013, p. 141-142) sugere que os (sub)processos cognitivos envolvidos no processamento das diversas categorias da sintaxe da música tonal que estamos estudando são seis:

1. Extração de elementos
2. *Knowledge-free structuring* (algo como Estruturação sem necessidade de conhecimento prévio)
3. Formação de expectativas musicais
4. *Structure building* (Construção de estruturas)
5. Re-análise e revisão estrutural
6. Integração sintática

³ “[...] occur on different levels of information processing (e.g., sensorimotor, schematic, associative, conceptual) and that some appraisal criteria, mostly driven by intrinsic stimulus characteristics, might be processed more easily and more rapidly at lower levels, whereas checks involving external inference require processing at higher levels” (SEIFERT; VERSCHURE, *et al.*, 2013, p. 211-212).

Nossa pesquisa, está focada principalmente no nível da estruturação sintática, particularmente na extração de elementos do fluxo contínuo da informação acústica. Em se tratando da estrutura temporal e de suas possíveis ambiguidades, referimo-nos à extração da pulsação musical, ou seja, à interpretação da ambiguidade temporal contida no estabelecimento do referencial da pulsação musical. Um fator importante no reconhecimento deste referencial é a percepção de acentos no sinal acústico, a Gestalt envolvida no processo de agrupamento e sequenciamento de eventos temporais.

Estes aspectos, envolvem em nossa opinião, prioritariamente o processamento informação de baixo nível, considerando os subprocessos supracitados.

Ainda no nível da integração sintática, Koelsch fornece um exemplo onde a recontextualização métrica de uma cadência musical poderia ser percebida como errada, se o acorde da tônica final não coincidisse com uma da pulsação musical forte e conclusiva (2013, p. 145). Argumentamos uma vez mais que nosso foco está localizado exatamente nas situações da literatura musical onde as informações importantes para o estabelecimento do referencial métrico ou estão degradadas, ou não são evidentes a ponto de serem percebidas como referência, gerando incerteza ou a possibilidade de dupla interpretação no ouvinte.

Nesse sentido, a importância dos mecanismos de percepção das acentuações num sinal acústico se torna fundamental para a definição do referencial temporal e a contextualização dos elementos do discurso dentro da integração sintática da cognição musical.

Na linguagem falada isso ocorre também com algumas palavras. Por exemplo, se tomarmos as palavras “aconteCEram” e “aconteceRÃO” apenas pela sonoridade e não pela representação escrita, veremos que a acentuação diferente muda o som da palavra e também seu significado, apesar dos fonemas serem muito semelhantes. Em um contexto de sonoridade dúbia em relação à sílaba tônica, o

ouvinte precisará de outros referenciais para promover a desambiguação. Um deles seria o contexto em que a palavra foi empregada, em que o processamento cognitivo seria mais de alto nível e haveria um grande investimento de energia da atenção. Embora com o processamento de ritmos musicais seja possível haver este recurso de alto nível para desambiguar uma situação semelhante, existem situações onde sincronizamos com músicas ou eventos sonoros recorrentes sem nem mesmo estarmos com a atenção focada especialmente neles.

O etnomusicólogo Simha Arom também faz uma possível referência ao fenômeno, no livro *African Polyphony and Polyrhythm*:

[...] O entrelaçamento de acentos e cores sonoras, combinados à falta de uma acentuação regular criam no ouvinte o sentimento de incerteza, uma impressão de ambiguidade relativa à articulação do período. Este sentimento pode ser comparado ao que se pode experimentar em um trem, quando percebemos a acentuação regular no som das rodas nos trilhos e repentinamente temos a impressão de que o período mudou: o que parecia um “tempo forte”, marcando o retorno do ciclo temporal, se torna um “tempo fraco”, e vice-versa. O mesmo fenômeno ocorre se ouvimos por algum tempo ao binário tic-tac do funcionamento de um relógio. O acento um, originalmente atribuído ao “tic” (TIC-tac) parece repentinamente mudar para o “tac”, e se pode ouvir TAC-tic. Maior ambiguidade pode resultar até de um período ternário. [...] (AROM, 2004, p. 230).

É possível que o tipo de ambiguidade referido por Arom e aquela exemplificada no início deste capítulo sejam de natureza semelhante e estejam associadas a mecanismos cognitivos mais relacionados ao processamento das informações de baixo nível, que provocam choque com expectativas e conceitos gerados em processamentos de alto nível, originando sensações de estranhamento no momento em que somos expostos a estas situações sonoras específicas.

Ao optarmos por investigar a efetividade do fenômeno da acentuação subjetiva, entendemos que este seja um bom ponto de partida para o aprendizado da pesquisa científica que emprega experimentos psicofísicos, além de acreditarmos que o fenômeno exerce importante papel no processamento cognitivo dos referenciais

métricos musicais.

Nesse sentido, nosso experimento parece ser original, ao registrar a percepção dos acentos em relação a estímulos métricos, diferentemente de Povel e Okkerman (1981).

O percurso experimental

Nosso percurso experimental foi iniciado após o estudo de matérias ligadas ao conhecimento do funcionamento do sistema nervoso, dos neurônios e da codificação dos estímulos realizada pelo cérebro, dentre as quais destacamos o curso de “Neurociência Cognitiva: Percepção e Ação”. Durante esse período inicial, este curso acabou sendo fundamental para a pesquisa, pois um de seus professores foi o Prof. Dr. Ronald Ranvaud, do Departamento de Fisiologia e Biofísica do Instituto de Ciências Biomédicas da Universidade de São Paulo. Juntamente com a revisão da literatura de referência, realizamos o primeiro esboço de experimento comportamental com a orientação do Prof. Ranvaud e a colaboração de membros de seu grupo de pesquisa, sobretudo Jorge Shiro Takashima, com a programação dos scripts do Psychtoolbox-3.

Nossa primeira abordagem, nomeada Fase 1, foi dedicada a investigar a acentuação em estruturas de 4 elementos equitoniais isolados, como no motivo de Beethoven. Sendo seguida por experimentos avaliando estruturas de 3 e 4 elementos isolados. Cogitamos que a duração destes estímulos não fosse suficiente para que a indução do efeito da acentuação subjetiva se manifestasse de maneira clara, em decorrência dos resultados obtidos.

Em uma segunda abordagem nomeada Fase 2, buscamos estímulos com duração maior e encontramos aqueles elaborados por Povel e Essens para investigar a existência de um “relógio interno” induzido pela exposição padrões rítmicos com pulsação isócrona, e que equivaleria ao que em música é chamado de *tactus* ou *beat*. A vantagem de utilizar estes estímulos decorria de sua classificação em categorias, conforme as supostas acentuações subjetivas coincidiam mais vezes às pulsações recorrentes, diminuindo até coincidir apenas duas vezes em diferentes configurações

de agrupamentos de elementos sonoros equitoniais. No desenho da interface de coleta de dados e apresentação do experimento, resolvemos adotar uma representação circular dos estímulos como metáfora do relógio, imaginando ser esta uma representação mais intuitiva que a representação linear da escrita musical, e supondo uma abordagem mais amigável para os participantes sem formação musical.

Na terceira abordagem, denominada Fase 3, alteramos a interface circular de anotação das acentuações percebidas dos estímulos para uma anotação linear, para que a ordem nas sequências de grupos se apresentasse na maneira mais clara possível.

Desenvolvemos uma metodologia de medição dos resultados obtidos, considerando o grau de similaridade entre as respostas dos participantes, e o Padrão esperado das Regras de Preferência, que passaremos a nomear Padrão-RP. Consideramos também a hipótese de que a nosso grupo esteja sujeito a um padrão diferente do esperado Padrão-RP e desenvolvemos metodologia de cálculo para esta possível situação.

Finalmente concluímos com as análises dos resultados obtidos abrindo também para novas questões surgidas no decorrer do nosso percurso experimental.

Capítulo 1

Percepção de regularidades temporais, Sincronia adaptativa (*entrainment*) e Modelos cognitivos de percepção métrica

1.1 Percepção de regularidades temporais

Perceber padrões, quaisquer que sejam, visuais, auditivos, táteis, exige habilidades de diferenciação entre elementos. Na percepção de padrões temporais, a diferenciação de elementos ocorre pela variação do intervalo entre os seus inícios, em que sucessões de intervalos semelhantes, por exemplo, provocarão o efeito de agrupamento dos mesmos. Agrupamentos podem ocorrer, ainda, através da acentuação de alguns elementos em detrimento de outros.

Em nossa pesquisa tratamos, fundamentalmente, das implicações da acentuação na percepção de regularidades ou padrões temporais. Nosso interesse está vinculado a conhecer os fatores que levam sujeitos a adotar suas referências temporais na detecção de padrões, de recorrências e de pulsações temporais.

Mas o que vem a ser acentuação? Todos conhecemos seus efeitos através da prática da nossa língua.

1.1.1 Acentuação

Acento significa, em fonética ou em música, a ênfase relativa dada a um elemento determinado dentro de um contexto. Na fonética, o acento refere-se à ênfase numa sílaba dentro de uma palavra; em música, significa a ênfase dada a

algumas notas ou sons específicos dentro de uma estrutura musical. Tanto na música como na fonética, esta ênfase pode ocorrer na intensidade sonora ou na duração da nota, som ou da sílaba.

A acentuação (ou *stress*) é uma das características dentro da prosódia e, bem como o ritmo e a entonação, vem sendo considerada como sendo muito relevante para a caracterização da identidade da língua falada:

Estudos de percepção sobre a identificação da linguagem humana confirmam que a informação prosódica, especificamente a frequência e a intensidade, é utilizada para a identificação do idioma sob condições em que a acústica de unidades sonoras e fonotáticas está degradada (Mori et al, 1999; Komatsu et al, 2001). Um estudo utilizando ressíntese revelou a importância do ritmo e da entonação para a discriminação linguística (Ramus e Mehler, 1999) (MARY; YEGNANARAYANA, 2008, p. 782-783)⁴

Musicalmente, esta ênfase dedicada a alguns elementos dentro de uma frase musical exerce influência sobre a compreensão da mesma. Se, acentuarmos de duas maneiras diferentes algumas das notas de uma mesma escala musical, como por exemplo:

DO, ré, **MI**, fá, **SOL**

e

dó, **RÉ**, mi, **FÁ**, sol⁵,

notamos que o simples deslocamento da acentuação muda o sentido da frase musical para a percepção. Trata-se de uma mudança qualitativa, que remete a interpretações

⁴ “Prosodic characteristics such as rhythm, stress and intonation in speech conveys some important information regarding the identity of the spoken language. Results of perception studies on human language identification, confirm that prosodic information, specifically pitch and intensity, are used for language identification under conditions where the acoustics of sound units and phonotactics are degraded (Mori et al., 1999; Komatsu et al., 2001). A study using resynthesis has revealed the importance of rhythm and intonation for language discrimination (Ramus and Mehler, 1999).” (MARY; YEGNANARAYANA, 2008, p. 782-783).

⁵ As duas escalas possuem exatamente as mesmas notas, porém a primeira contém acentos nas notas escritas em letra maiúscula (**DO**, **MI** e **SOL**) e na segunda a acentuação ocorre com o mesmo padrão intercalado, mas deslocando o acento inicial para a segunda nota da escala (**RÉ** e **FÁ**).

distintas, tanto harmônicas e melódicas, como de condução do movimento de projeção musical.

Lerdahl e Jackendoff (1983, p. 17), importando conceitos da linguística para analisar a organização estrutural da música, consideram a distinção de três tipos de ênfase ou de acentos: **fenomenal, estrutural e métrico**: “Por acento fenomenal, entendemos qualquer evento, na superfície musical, que dê ênfase ou estresse a momentos no fluxo musical”⁶. É importante dizer que a palavra **fenomenal**, traduzida do inglês (*phenomenal*) significa fenômeno e sua tradução do grego (*φαινόμενον*) significa coisa observável, perceptível.

Então, **acentos fenomenal**⁷ representa a ênfase percebida no contexto musical e se relaciona com diferenças entre aspectos físicos do som, como volume e duração relativos, saltos de frequência e separação temporal entre as notas.

O **acentos estrutural**⁸ se define pela “acentuação ou ênfase causada por pontos de gravidade harmônica/melódica numa frase ou seção, especialmente pela cadência, considerada o alvo principal do movimento tonal.” (LERDAHL; JACKENDOFF, 1983, p. 17)

Já o **acentos métrico** “é um constructo mental, inferido mas não idêntico aos padrões de acentuação na superfície musical”⁹ (LERDAHL; JACKENDOFF, 1983, p. 18), portanto abstrato. É altamente vinculado à noção de hierarquia métrica, na qual os eventos temporais são agrupados em unidades recorrentes, que contêm subunidades e superunidades, e ocorrem em níveis de pulsações justapostas que se desdobram na proporção de números inteiros e seu grau de acentuação é definido pela coincidência entre pulsos de dois ou mais níveis.

Repp afirma que:

⁶ (...) By *phenomenal accent* we mean any event at the musical surface that gives emphasis or stress to a moment in the musical flow. (...) (p. 17)

⁷ Phenomenal accents are conveyed by aspects of the physical sound structure, such as differences in loudness or duration, leaps in pitch, and temporal separation. Structural accents are less relevant to the present study and can be left aside.

⁸ (...) By *structural accent* we mean an accent caused by the melodic/harmonic points of gravity in a phrase or section-especially by the cadence, the goal of tonal motion. (...) (p.18)

⁹ (...) Metrical accents, by contrast, are “a mental construct, inferred from but not identical to the patterns of accentuation at the musical surface” (...) (p. 18).

[...] acentos fenomenais são importantes, mas existem também determinantes endógenos que podem persistir em face de um sinal de entrada conflitante. Se o padrão de acentos fenomenais está empobrecido ou ambíguo, o mesmo trecho musical pode dar origem a diferentes estruturas métricas percebidas.¹⁰ (REPP, 2010, p. 1390).

A afirmação acima mostra que existem outros aspectos envolvidos na percepção de acentos. A diferença na percepção das acentuações seria decorrente de estarem presentes no sinal físico e exógeno (de origem externa), ou se teriam origem no processamento endógeno (originado internamente) a partir do sinal de entrada acústica? Como inferimos a métrica, ou seja, como esse constructo mental se relaciona com a acentuação de superfície em uma música?

Refletindo sobre essa questão fulcral, suponho que os fatores endógenos atuariam na possível desambiguação de um sinal de entrada ambíguo, ou com fraca ênfase nas acentuações. Considero a chamada acentuação subjetiva, se não a única, pelo menos a mais estudada evidência deste fator endógeno proposto por Repp.

Assim, a **acentuação subjetiva** pode ser considerada um fenômeno que media o processo perceptivo dos diversos tipos de acentuação elencados acima, e que exerce função desambiguadora atuante nas ocasiões em que o sinal exógeno (de origem externa) se encontra degradado ou sua natureza é propriamente ambígua.

Dentro da categoria dos fatores endógenos existe, ainda, a questão da percepção modulada pela familiaridade, por fatores culturais ou pela interação entre o repertório cultural e a predisposição genética evolutiva do indivíduo a tal modo de percepção. Ontogenia versus filogenia. Embora reconheçamos a pertinência, tais questões não serão abordadas por nós neste trabalho. Dentre os fatores endógenos está envolvida a expectativa, sobre a qual trataremos mais adiante, sob a perspectiva de David Huron.

Em uma segunda reflexão, conjecturamos que as nuances da acentuação

¹⁰ “[...] Phenomenal accents, although important, are not the only determinants of perceived meter. There are endogenous determinants as well, as is already suggested by the fact that an induced metrical structure can persist in the face of conflicting input. If the pattern of phenomenal accents is impoverished or ambiguous, the same musical passage can give rise to different perceived metrical structures. [...]” (REPP, 2010, p. 1390).

subjetiva se estendem desde passagens onde há a presença de um acento muito claro, marcante e objetivo até as muito ambíguas. Em outras palavras, existe um núcleo da percepção da acentuação que seria mais ligado ao fenomenal e que, hipoteticamente, seria perceptível a todos, de maneiras muito semelhantes. Por outro lado, a experiência individual altera sutilmente mesmo as escutas mais determinadas por fatores objetivos. Em todos os casos, os acentos são as entidades aglutinadoras, formando os agrupamentos necessários à compreensão da sintaxe musical, que ocorre através da interpretação da informação contida no sinal acústico em processos *bottom-up* e sua contextualização em processos *top-down*.

Em outra reflexão possível, consideramos que as acentuações fenomenal e estrutural têm caráter preferencialmente exógeno, ou seja, em princípio estão contidas fisicamente no sinal acústico. Já o acento métrico, definido como um “constructo mental”, teria um caráter mais propriamente endógeno, assim como ocorre com a acentuação subjetiva. Nesse sentido, penso eu, a acentuação subjetiva poderia ser nomeada **acentuação endógena**, uma vez que a idéia de subjetividade se contrapõe à de objetividade; no entanto, mesmo as interpretações mais objetivas envolvem certa dose de subjetividade, posto que são interpretações.

A diferença na percepção das acentuações seria, efetivamente, em que medida ela decorre de sua presença no sinal físico e exógeno, e em que medida está presente no processamento endógeno do sinal de entrada acústica.

1.2 *Entrainment* e percepção métrica

A percepção da regularidade de eventos temporais sonoros faz parte do processo conhecido como *entrainment*, “um processo biológico que adapta sincronicamente oscilações internas da atenção a um evento externo” (JONES, 2009,

p. 83)¹¹. Este processo de percepção de regularidades temporais em eventos externos é o responsável, por exemplo, pela capacidade humana de sincronizar os movimentos corporais dos músicos aos do maestro em uma orquestra, ou os dos pares ao dançar, seguindo o tempo da música em um baile.

Biologicamente falando, a música e a dança são manifestações de comportamentos rítmicos e, nesse sentido, podem ser didaticamente divididas em três componentes (FITCH, 2011, p. 75-76): extração de pulso, *entrainment* (ou processo de sincronização) e geração de padrão motor. Segundo Fitch, “o primeiro envolve apenas um processamento auditivo, o segundo envolve a integração inter-modal (do sistema auditivo com o motor) e o último, um comportamento motor sensorialmente guiado”¹².

Certamente, a resposta comportamental rítmica expressa enquanto música e dança tem mais nuances do que essa divisão didática em componentes, que não se constitui enquanto um modelo científico, posto que o amplo processamento da percepção temporal ainda se encontra em estado de mapeamento, não tendo sido desvendado por completo.

No entanto, esta divisão didática torna mais clara a visão de que o *entrainment* não é um fenômeno que atua na percepção apenas de modo intelectual e analítico, e também não é fruto somente da cultura humana, mas tem uma importância evolutiva, uma vez que integra nossos sistemas sensório-motores e nos prepara para uma ação adequada. No reino animal, é o *entrainment* quem prepara a presa para a fuga a partir da percepção sonora dos passos rápidos de um predador que se aproxima, permitindo que ela planeje estratégias de escape, assim como um jogador de futebol dribla os seus oponentes.

Entendido mais amplamente como o acoplamento de osciladores temporais

¹¹ “[...] Entrainment is a biological process that realizes adaptive synchrony of *internal* attending oscillations with an *external* event. [...]” (JONES, 2009, p. 83).

¹² “[...] I hypothesize three components as necessary for ensemble playing and dancing: pulse extraction, beat entrainment, and motor pattern generation. [...] The first involves auditory processing only, the second cross-modal (auditory to motor) integration, and the last sensorily-guided motor behavior. [...]” (FITCH, 2011, p. 75-76).

dos relógios biológicos com estímulos cíclicos do ambiente, o *entrainment* é ubíquo e bastante comum em sistemas biológicos, conforme observa Fitch. Por outro lado, a habilidade e propensão humana, de sincronizar ações motoras com estímulos auditivos parece ser rara. Assim, forma-se um aparente “paradoxo do ritmo”:

Periodicidade e *entrainment* parecem estar entre as características mais básicas de seres vivos, mas a capacidade humana (e a inclinação) para sincronizar a sua saída motora a estímulos auditivos parece ser muito rara: não há evidência crível de sincronismo de pulsação entre outros primatas ou mamíferos, mesmo que simples (Merker, 1999, 2000) [...]” (FITCH, 2011, p. 78)¹³.

A inferência de regularidades temporais em estímulos musicais, também conhecida como métrica¹⁴, representa um refinamento do *entrainment* enquanto processo biológico operado pela cultura e manifesto na música e na dança. A métrica define-se pela percepção da recorrência de padrões de acentuações relativas, contidas no estímulo físico ou formadas pelo ouvinte com o auxílio do fenômeno da chamada acentuação subjetiva (POVEL; OKKERMAN, 1981). O fenômeno da métrica, “ao proporcionar modelo temporal recorrente, tem uma função preditiva na codificação da ocorrência de eventos” (HURON, 2006, p. 198)¹⁵.

A representação métrica temporal da música ocidental, como a conhecemos, apresenta níveis hierárquicos relacionados a acentuações recorrentes e um nível de referência denominado *beat* na língua inglesa. Em português esse termo não tem

¹³ “[...] Periodicity and entrainment seem to be among the most basic features of living things, yet the human ability (and proclivity) to entrain our motor output to auditory stimuli appears to be very rare: there is no credible evidence of even simple beat entrainment among other primates or mammals (Merker, 1999, 2000). [...]” (FITCH, 2011, p. 78)

¹⁴ A acentuação que costumamos perceber na música ocidental é aquela em que o pulso da música pode apresentar divisões simples de duas, três ou quatro subunidades, como podemos ouvir na música *pop*, na qual a bateria toca as unidades maiores com tambores mais graves e médios, e as subunidades são tocadas nos pratos metálicos. Numa valsa, que tem pulsação ternária, costumamos mover os pés sempre quando se conta mentalmente o 1 da sequência 1-2-3, 1-2-3, indicando que associamos o acento métrico ao primeiro tempo da sequência da dança.

¹⁵ “[...] the phenomenon of meter, we find the meter provides a recurring temporal template for coding and predicting event onsets. [...]” (HURON, 2006, p. 198)

tradução adequada e a literatura musical na nossa língua traduz o *beat* como *tactus*¹⁶ ou como tempo do compasso, definidos aqui como unidade de referência temporal utilizada pelo ouvinte, possivelmente para guiar o processo de sincronização adaptativa (*entrainment*), reunindo pulsações isócronas em grupos recorrentes de subdivisões em números inteiros, e usado para coordenar passos de dança da maneira mais natural. “Em suma, o *tactus* é um nível perceptivelmente proeminente da estrutura métrica, que as regras não conseguiram designar de maneira mais específica”¹⁷ (LERDAHL; JACKENDOFF, 1983, p. 71).

Como vimos, a dificuldade de determinar categoricamente essa unidade se deve, provavelmente, ao fato de compreendermos ainda pouco de como ocorre esse processo do ponto de vista neuronal.

Dada a inexistência de uma tradução adequada para o *beat* ou o *tactus* em português, passaremos a adotar a palavra “pulsação” para designar o conceito.

Existem alguns modelos que explicam a formação da representação métrica, como o modelo de Regras de Preferência, apoiado também no fenômeno da acentuação subjetiva, a teoria da assistência dinâmica e, mais modernamente, a teoria da ressonância neural. Trataremos desses modelos após uma breve contextualização histórica do estudo dos comportamentos rítmicos.

¹⁶ *Tactus* é um termo do século 15 e 16 usado para designar o que conhecemos por *beat*, ou unidade referencial de tempo. Quando se acompanha uma música batendo os pés ou batendo palmas em geral é o equivalente ao *beat*.

¹⁷ “[...] Yet metrical intuitions about music clearly include at least one special designated metrical level, which we are calling the *tactus*. This is the level of beats that is conducted and with which one most naturally coordinates foot-tapping and dance steps. When one wonders whether to “feel” a piece “in 4” or “in 2”, the issue is which metrical level is the *tactus*. In short, the *tactus* is a perceptually prominent level of metrical structure that the rules so far fail to designate as any way special. We can incorporate this notion into the formal theory by designating a particular level in a metrical structure as the *tactus*. [...]” (LERDAHL; JACKENDOFF, 1983, p. 71).

1.2.1 Histórico das pesquisas sobre sincronização (*entrainment*)

As pesquisas mais antigas envolvendo sincronização datam do final do século XIX e início do século XX e eram basicamente comportamentais no que se poderia chamar de “estudos de sincronização” (*tapping studies*). Essas pesquisas consistiam basicamente em experimentos com voluntários que simplesmente acompanhavam metrônomo com batidas na proporção de 1:1 em vários andamentos. Esse tipo de estudo “envolve não só a atenção, mas também ação, um comportamento (bater em sincronia) que depende da atenção ao ritmo” (LONDON, 2012, p. 25) ¹⁸.

Nos anos 1970 e 80, as pesquisas sobre a percepção temporal, revelaram importantes conhecimentos no sentido de estabelecer os limites psicofísicos, como relata Patel:

Do ponto de vista do ouvinte, a percepção do *tactus* normalmente é relacionada ao movimento enquanto sincronização com o tempo da música. Para muitas pessoas, essa sincronização é parte natural da experiência musical e não requer um esforço especial. Pode ser uma surpresa, então, que os humanos sejam a única espécie que sincroniza espontaneamente com o *tactus* musical. Mesmo que a sincronia seja conhecida em outras partes do reino animal, como os coros dos sapos e os chamados sincronizados dos insetos (Gerhardt & Huber 2002, Ch. 8; Strogatz, 2003), a sincronização humana com o *tactus* é singular em muitos aspectos.

Obviamente, a percepção do *tactus* não causa automaticamente movimento (alguém pode sentar e ouvir), mas a exclusividade da percepção na sincronização do *tactus* sugere que a percepção do tempo merece investigação psicológica.

Primeiro, existe uma banda de tempo preferida para percepção do *tactus*. As pessoas têm dificuldade de seguir um *tactus* mais rápido que 200ms e mais lento que 1.2s. Dentro desta banda, existe a preferência por *tactus* que ocorram entre 500-700ms (Parncutt, 1994; van Noorden e Moelants, 1999) (PATEL, 2008, p. 100) ¹⁹.

¹⁸“Tapping studies involve not only attention but also action, a behavior (tapping in synchrony) that depends on rhythmic attending. Studies of tapping go back to the 19th century and early 20th centuries (e.g., Dunlap 1910; Stevens 1886 Woodrow 1932)” (PATEL, 2008, p. 96-97).

¹⁹ “[...] From a listener's perspective, perception of a beat is often linked to movement in the form of synchronization to the beat. For many people, this synchronization is a natural part of musical experience requiring no special effort. It may come as a surprise, then, that humans are the only

Outros trabalhos se encaminharam no sentido de comparar o ritmo musical com o ritmo na fala, já que ambos são caracterizados pela sistematização temporal de acentos e padrões de fraseamento. O domínio e a fluência nativa no aprendizado de uma língua também “requer mais do que o domínio dos fonemas, do vocabulário e da gramática. É preciso dominar os padrões de tempo e acentuação que caracterizam o fluxo das sílabas nas sentenças” (PATEL, 2008, p. 96-97).

Patel completa essa informação sobre essa questão do ritmo da fala:

[...] cada língua tem um ritmo que faz parte de sua estrutura sonora, e um conhecimento implícito deste ritmo faz parte da competência de um orador em sua língua. A incapacidade de adquirir ritmo nativo é um fator importante na criação de um sotaque estrangeiro na fala (Taylor, 1981; Faber, 1986; Chela-Flores, 1994). [...] (PATEL, 2008, p. 96-97)²⁰.

A pesquisa contemporânea sobre o tempo musical foi iniciada pelo clássico trabalho de Fraisse²¹ de 1963, que tratou mais especificamente dos agrupamentos rítmicos de forma serial, nos quais teriam prevalência relações de Gestalt como proximidade e similaridade. Ainda segundo Jones (2009, p. 82), aspectos métricos entraram em cena quando Povel (1981) indicou a formação da percepção de hierarquias temporais. Neste artigo, Povel (1981, p. 13) verifica que, na proporção preferencial de 3:1, por exemplo, a duração menor é percebida como subdivisão da maior, sendo esta última percebida também como acentuada em relação à menor,

species to spontaneously synchronize to the beat of music. Although synchrony is known from other parts of the animal kingdom, such as the chorusing of frogs or the synchronized calls of insects (Gerhardt & Huber 2002, Ch. 8; Strogatz, 2003), human synchronization with a beat is singular in a number of respects (see Chapter 7, section 7.5.3, for further discussion of this point). Of course, beat perception does not automatically cause movement (one can always sit still), but the human uniqueness of beat synchronization suggests that beat perception merits psychological investigation. Research in music cognition has revealed several interesting facts about beat perception. [...]” (PATEL, 2008, p. 100)

²⁰ “[...] Speaking a language with native fluency requires more than mastering its phonemes, vocabulary, and grammar. One must also master the patterns of timing and accentuation that characterize the flow of syllables in sentences. That is, each language has a rhythm that is part of its sonic structure, and an implicit knowledge of this rhythm is part of a speaker's competence in their language. A failure to acquire native rhythm is an important factor in creating a foreign accent in speech (Taylor, 1981; Faber, 1986; Chela-Flores, 1994). [...]” (PATEL, 2008, p. 96-97)

²¹ FRAISSE, P. *The psychology of time*. New York, NY: Harper and Row, 1963.

denotando a percepção de uma estrutura temporal subjacente aos padrões rítmicos usados em seu experimento e estabelecendo a proposição de um modelo denominado *beat-based model* (“modelo baseado na pulsação musical”, numa tradução livre) (POVEL, 1981, p. 13). A estrutura subjacente nesse experimento é o percepto²² equivalente ao **compasso musical**.

Assim, Jones considera que começaram a se estabelecer duas correntes psicológicas teóricas da percepção métrica: a Teoria da codificação métrica (*metric encoding*) e a Teoria da Assistência Dinâmica, discutidas individualmente mais abaixo.

Como pudemos notar, desde 2009 Jones refere-se a esses temas usando a designação “Teorias”. Tendo em vista o amplo emprego da terminologia modelo e teoria, cabe aqui uma explanação e uma diferenciação desses conceitos.

No presente trabalho, iremos considerar a concepção de modelo como sendo uma forma estruturada que possibilita uma compreensão do mundo em sua complexidade. Ou seja, constitui-se enquanto uma representação esquemática simplificada de algum aspecto da realidade do mundo, permitindo previsões de resultados ou comportamentos. Já a teoria ocupa um degrau acima do modelo, pois além de permitir previsões acerca de um conhecimento, também permite identificar eventuais ações de controle sobre os elementos.

Os referenciais citados no presente trabalho empregam ambos os termos, indiscriminadamente, havendo uma tendência ao uso do termo modelo nos textos mais antigos e do termo teoria nos mais recentes. Mantivemos as designações originais, porém acrescidas desta ressalva, para evitar uma leitura confusa dos mesmos. E observamos que o fato do corpus das ideias a respeito do modelo de regras de preferência receber a designação de Teoria da Codificação, reflete o reconhecimento da formação de um conjunto de conhecimentos mais robusto sobre o tema.

²² O termo percepto refere-se ao resultado da percepção, àquilo que foi percebido, seja ele um aspecto musical (como o ritmo ou a altura), uma imagem visual, um cheiro.

1.3 O Modelo de Regras de Preferência (*Preference Rule-Based Model*, PRM) ou Teoria da Codificação (*Encoding Theory*)²³

No texto *Parallelism as a Factor in Metrical Analysis*, Temperley e Bartlett (2002) apresentam uma síntese da abordagem denominada Modelo de Regras de Preferência, PRM. O Modelo de Regras de Preferência deriva do Teorema de Bayes (1763), aplicado a processos de aprendizagem em que probabilidades iniciais são revistas à presença de nova informação, analogamente ao que ocorre com o método científico. Conforme suas concepções, a percepção humana também reflete processos dinâmicos de aprendizagem. Existe evidência empírica de que o funcionamento do cérebro humano e a percepção da realidade seguem a lógica bayesiana, como já observava a psicologia experimental em *A Bayesian formulation of visual perception* (KNILL, KERSTEN e YUILLE, 1996). Ao comparar a percepção visual à percepção das estruturas musicais (mesmo sem manifestação direta na citação), Lerdahl e Jackendoff descrevem o princípio elementar da lei de Bayes na seguinte passagem:

[...] o campo visual em geral é altamente ambíguo. Portanto, a função de mapeamento terá que incluir um sistema de preferências que escolha a análise ou as análises mais altamente favorecidas para o campo visual apresentado. As leis de organização da Gestalt podem ser construídas como declarações informais de vários aspectos deste sistema de preferências” (LERDAHL; JACKENDOFF, 1983, p. 306).

Segundo a abordagem de Temperley e Bartlett (2002, p. 119-119), a percepção do ritmo seria determinada de acordo com algumas tendências a atribuir acentuações a eventos articulados ritmicamente, de acordo com o seu posicionamento dentro da estrutura apresentada. Ou seja, a preferência pela

²³ Este tópico apresenta-se enquanto uma ampliação de passagens que integram nossa dissertação de mestrado (BONDESAN DOS SANTOS, 2012) e o artigo publicado nos anais da Associação Nacional de Pesquisa e Pós-Graduação em Música (BONDESAN DOS SANTOS, 2016).

atribuição de acento subjetivo de um evento em meio a outros de igual origem seria determinada por sua relação com outros elementos da sequência. Neurocientistas de diversas tendências da área chegaram a um consenso sobre cinco fatores, que orientam o processo de indução da pulsação musical (*beat induction*). Estes fatores atendem ao modelo que segue: (1) preferência por pulsações (*beats*) que coincidem com ataques de notas; (2) preferência por alinhar tempos métricos com as notas mais longas (para evidência psicológica dessa afirmação, ver os supracitados POVEL e ESSENS, 1985); (3) preferência pela regularidade dos tempos métricos; (4) formação de “agrupamento”, uma tendência a se considerar o tempo métrico forte no início de uma série de notas que formam um grupo ou uma frase (POVEL; ESSENS, 1985); (5) formação de progressões harmônicas, uma tendência a associar tempos métricos fortes com os pontos de mudança na harmonia (TEMPERLEY, 2001, p. 51).

As regras de preferência se referem, basicamente, a regras de agrupamento temporal nas quais um sujeito tende a atribuir acentuação relativa a alguns elementos em uma sequência de sons acusticamente idênticos. Um aspecto importante no PRM é o fenômeno da acentuação subjetiva. Uma acentuação subjetiva é aquela atribuída “subjetivamente” pelo ouvinte a sequências de estímulos acústicos equitonais, com pulsação isócrona e destituídas de qualquer acentuação física (POVEL; OKKERMAN, 1981).

Segundo Jones (2009), a Teoria da Codificação (*Encoding Theory*) prestou continuidade aos estudos iniciados através do PRM, através das pesquisas de sincronização de batidas (*synchronized tapping*), que conduzem para a suposição da existência de um relógio interno induzido através da percepção e/ou reprodução de estímulos externos ritmados. O ponto de discussão entre os pesquisadores é se a indução do relógio depende mais das batidas não acentuadas, das acentuadas, ou de um modelo híbrido que considera os dois casos. Nas pesquisas realizadas por Povel o modelo das batidas não acentuadas previu melhor performance dos não músicos, enquanto o modelo híbrido previu melhor a performance dos músicos.

Experimentos com batidas fixas em 800ms, variando em subdivisões isócronas e não isócronas revelam que as subdivisões isócronas são melhor

reproduzidas que as não isócronas (PATEL, 2005), reforçando a ideia de que os ouvintes dependem de relações de proporção entre múltiplos níveis métricos.

O fenômeno da acentuação subjetiva, que presta suporte à Teoria da Codificação, parece ser um efeito que tem conexão importante com o processo de sincronização adaptativa. Este último leva à percepção ou ao estabelecimento de uma unidade temporal (pulsção.ou *beat*), que torna possível o comportamento rítmico sincronizado dos seres humanos. Nossa intenção é iniciar este trabalho investigativo com experimentos que testem o efeito da acentuação subjetiva em grupos universitários sem estudo musical e com estudo musical.

1.4 Teoria da Assistência Dinâmica, DAT

A Teoria da Assistência Dinâmica (*Dynamic Attending Theory*, DAT) (JONES; BOLTZ, 1989 apud FUJIOKA; FIDALI; ROSS, 2014, p. 2) propõe que a sincronia adaptativa de oscilações internas da atenção²⁴ com um evento externo (*entrainment*) é auxiliada pela percepção de hierarquias métricas temporais, as quais constituem um processamento dos níveis temporais subjacentes ao estímulo, os quais, por sua vez, induzem oscilações neurais correlatas.

Assim, as expectativas temporais seriam a manifestação da persistência endógena da periodicidade, induzida por cada nível métrico contido no estímulo. Sua adaptabilidade se dá pelo ajuste da fase em resposta à *violação das expectativas*, em razão da sintonização recorrente dos intervalos de tempo de um dado nível (JONES, 2009, p. 83).

É importante notar ainda que, segundo Jones, haveria a influência de fatores de treinamento e aculturação contribuindo para a familiarização junto aos níveis métricos e promovendo o que ela chama **vinculações métricas** (*metric binding*).

²⁴ A **Atenção** aqui, refere-se a uma função cognitiva, definida pela seleção e manutenção de um foco, para a atenção seletiva, em um estímulo ou informação proveniente dos nossos sentidos, das memórias ou de outros processos cognitivos. Na atenção dividida o indivíduo coloca o foco em mais de um estímulo ou processo ao mesmo tempo.

As vinculações métricas resumem-se à sincronização interna entre as oscilações neurais referentes a cada nível métrico, produzindo ressonância entre eles e formando um bloco (*cluster*) (JONES, 2009, p. 84), como podemos observar na figura 2.

Figura 2 – Modelo representativo da ideia de ressonância neural, onde um estímulo externo provocaria a emergência de osciladores internos em proporções de harmônicos, perfazendo analogamente os níveis métricos musicais, e promovendo o que Jones (2009, p. 84) denomina vinculação métrica (metric binding)

Outro conceito também importante para uma melhor compreensão dos fenômenos da percepção é o chamado **foco atencional** (*attentional focusing*), que se trata, basicamente da “habilidade de focar a atenção em sinais relevantes do ambiente para uma determinada tarefa; esta habilidade pode também incluir a supressão de estímulos distratores (COGNITIVE ATLAS, [s.n.]). Existem muitas tentativas de explicação do fenômeno da **atenção auditiva**:

[...] extensivos estudos em humanos usando técnicas de ERP (potenciais relacionados a eventos) e ERFs (campos magnéticos relacionados a eventos)[...] Analisando alterações de ERPs e ERFs em função do direcionamento da atenção, é possível fazer inferências sobre o momento [timing], nível de processamento, e localização anatômica de processos de seleção de estímulos no cérebro. [...] (WOLDORFF, 2015)²⁵.

²⁵ “[...] Auditory attention has been investigated extensively in humans using event-related potentials (ERPs) and event-related magnetic fields (ERFs). These recordings can noninvasively track with high temporal resolution the brain activity associated with different types of stimulus events. By analyzing changes in the ERPs or ERFs as a function of the direction of attention, one can make inferences

Faremos aqui uma breve interrupção para explicitar que a atenção funciona de forma análoga à transdução de potenciais de ação entre os neurônios, privilegiando os estímulos do ambiente, relativos a uma tarefa específica, e suprimindo os estímulos distratores. Como ocorre normalmente nos estudos do cérebro, as questões frequentes relativas a qualquer aspecto de processamento referem-se a quais sistemas estão envolvidos naquele processamento, em que grau, se há uma ordem no caminho do processamento e se a avaliação se faz no âmbito químico, celular ou no âmbito do processamento dos potenciais de ação (que são parte dos potenciais relacionados a eventos²⁶). Estes últimos resultam da inversão da polarização elétrica da membrana dos neurônios e são responsáveis pelo efeito de transdução das informações no sistema nervoso. Para haver a transdução de um potencial de ação, é preciso que um estímulo produza a despolarização acima do limiar do neurônio. Somente assim a célula neuronal dispara o potencial de ação transduzido a outro neurônio e assim por diante. A transdução dos potenciais de ação ocorre quando os neurônios liberam substâncias denominadas neurotransmissores, que podem ser excitatórios ou inibitórios, produzindo ou não a transdução da ação correspondente à uma via neural. Assim, quando focamos a atenção em algo, inibimos os elementos periféricos a este.

Se, no âmbito da percepção, o conhecimento acumulado ainda não é definitivo em relação à atenção, no caso do processamento auditivo, o percurso para explicar seu funcionamento carece ainda de muita pesquisa.

Algumas evidências apontam para o aprimoramento das respostas aos estímulos selecionados no córtex auditivo primário de macacos para tarefas atencionais seletivas de escuta dicótica (BENSON; HEINZ, 1978 apud

about the timing, level of processing, and anatomical location of stimulus selection processes in the brain. [...]” (WOLDORFF, 2015).

²⁶ Mais à frente, esse trabalho traz explicações mais específicas sobre os potenciais relacionados a eventos.

WOLDORFF, 2015)²⁷. Outras investigações operam no âmbito das proteínas reguladas pela estimulação sonora em pássaros canoros (PINAUD; OSORIO, *et al.*, 2008). A última edição deste último trabalho traz uma revisão sobre o assunto e afirma existirem evidências que apontam para uma função primordial da transmissão de neurônios GABAérgicos de suprimir as respostas neuronais a frequências que ladeiam a frequência preferida de neurônios envolvidos no processamento auditivo:

Os benefícios desta forma de inibição no processamento sensorial incluem a intensificação dos campos receptivos e uma eficiência melhor na codificação de sinais sensoriais. Pesquisas posteriores revelaram, especificamente, a importância da inibição GABAérgica para determinar a localização do som, bem como de outros processamentos mais complexos como a sensibilidade, a harmonia, ou os movimentos direcionalmente restritos na frequência [de pássaros canoros]. (TREMERE; KOVALESKI; PINAUD, 2013, p. 321)

Estas evidências observadas em pássaros canoros demonstram um mecanismo neural atuando através do neurotransmissor GABA, cuja função é inibir a transdução sináptica de um determinado neurônio. Assim, temos uma possível associação entre o foco da atenção do pássaro e a inibição de sinais menos importantes, tendo por consequência uma intensificação das informações relevantes para o sujeito.

Voltando a Jones, seus experimentos trazem evidências rítmicas da atenção, indicando que a previsibilidade temporal não é o único determinante importante na segregação de fluxo. O tempo de exposição aos estímulos, sua estrutura e a atenção fazem diferença na percepção de eventos (JONES; KIDD; WETZEL, 1981, p. 1070-1071). A figura 3 apresenta um modelo simplificado de funcionamento da energia atencional, atuando com um oscilador focado na violação ou confirmação da expectativa gerada por um ritmo isócrono.

²⁷ Benson and Heinz (1978), studying single cells in monkey primary auditory cortex during a selective attention task (dichotic listening), reported relative enhancement of the responses to attended stimuli.

Figura 3 – Um oscilador sincronizando com um ritmo isócrono. Os componentes básicos são o período, o pulso e a fase. À medida que o evento confirma a expectativa ou não, o foco atencional se estreita ou se alarga. (JONES, 2009, p. 85)

Tendo sido inicialmente uma teoria evidenciada por experimentos basicamente comportamentais, desde os anos 1990 a DAT vem sendo fortalecida pela obtenção de evidências experimentais no campo da neurociência, através de pesquisas com ondas de Eletroencefalograma (EEG), entre outros métodos de medição de atividade eletrofisiológica.

Este percurso experimental agrega novos conhecimentos advindos da neurociência e sua consequência direta é o que vem sendo designado por **Teoria da ressonância neural**.

1.5 Teoria de ressonância neural da métrica musical

A ideia de um **acoplamento neural** com estímulos sonoros vem da observação da correlação entre a sincronização comportamental de indivíduos a estímulos musicais e a respectiva ocorrência de atividade neuronal em frequência equivalente. Esta observação foi documentada através da técnica de Potenciais

Relacionados a Eventos (ERP, *Event-Related Potentials*).

Sobre estes últimos, Luck esclarece:

[...]Os Potenciais Relacionados a Eventos foram inicialmente chamados Potenciais Evocados (EP de *Evoked Potentials*) porque tratavam de potenciais elétricos evocados por estímulos (em oposição às ondas EEG espontâneas). [...] A expressão ‘Potenciais Relacionados a Eventos’ é proposta para designar a classe geral dos potenciais que exibem relações de tempo estável para um evento de referência definível[...] (VAUGHAN, 1969, p. 46 apud LUCK, 2005)²⁸.

Em outras palavras, esses potenciais evocados referem-se a respostas de populações de neurônios a estímulos externos. No nosso caso, eles são estímulos acústicos não necessariamente periódicos, mas que contêm no mínimo uma pulsação. Esta pulsação, de algum modo, é percebida pelo sistema nervoso que responderia adaptativamente, procurando projetar expectativas de continuidade desta mesma pulsação. A ressonância de sub-harmônicos ou harmônicos da frequência dessa pulsação produziria subunidades ou superunidades temporais, estabelecendo níveis a que musicalmente chamamos de níveis métricos. Nozaradan e Peretz resumem a Teoria da Ressonância neural da seguinte maneira:

[...]A teoria de ressonância para o *tactus* e a percepção métrica (Large, 2008; Large e Kolen, 1994; van Noorden e Moelants, 1999) propõe que a percepção do *tactus* emerge do processo de sincronização adaptativa [*entrainment*] de populações neuronais ressoando na frequência do *tactus*, em si dando origem a ressonância em sub-harmônicos desta frequência de ordem mais elevada (na hierarquia métrica), correspondente ao compasso[...] (NOZARADAN; PERETZ, 2013, p. 104).²⁹

A primeira evidência direta desta hipótese, um grande avanço na área, veio

²⁸ “[...] ERPs were originally called evoked potentials (EPs) because they were electrical potentials that were evoked by stimuli (as opposed to the spontaneous EEG rhythms). The term ‘event related potentials’ (ERP) is proposed to designate the general class of potentials that display stable time relationships to a definable reference event. [...]” (Vaughan, 1969, p. 46 apud LUCK, 2005).

²⁹ “[...] The resonance theory for beat and meter perception (Large, 2008; Large and Kolen, 1994; van Noorden and Moelants, 1999), proposes that beat perception emerges from the entrainment of neuronal populations resonating at the beat frequency, itself giving rise to higher order resonance at subharmonics of beat frequency, corresponding to the meter [...]” (NOZARADAN; PERETZ, 2013, p. 104).

através da medição da frequência de respostas de populações de neurônios em um experimento em que os participantes ouviam uma pulsação com acentuações irregulares, e imaginavam métricas binárias e ternárias. Os resultados mostraram diferenças entre as respostas para cada situação, indicando que a intencionalidade da interpretação de um mesmo estímulo físico pode ser observada diretamente na resposta neuronal pela abordagem chamada *frequency tagging* (mapeamento de frequência) (NOZARADAN; PERETZ et al., 2011), que consiste em observar os picos de frequência registrados no EEG após o início do estímulo.

Uma interessante linha de investigação derivada desta abordagem, seria observar que tipos de respostas podem ser elicitadas em sujeitos ouvindo estruturas metricamente ambíguas. Nossa pesquisa atual ganharia muito com este tipo de abordagem.

Apesar dos modelos supraexpostos apresentarem diferenças de abordagem do fenômeno, o Modelo de Regras de Preferência, a Teoria da Assistência Dinâmica e a abordagem da Ressonância Neural não são excludentes e podem ser complementares.

O fenômeno da acentuação subjetiva, um dos pilares do Modelo de Regras de Preferência, apresentou indícios de correlação entre potencial evocado e acentuação subjetiva, reforçando a hipótese que vem sendo explorada por diversos pesquisadores, como os do grupo de Potter e Fenwick (*Perceiving rhythm where none exists: event-related potential (ERP) correlates of subjective accenting.*, 2009) , que vincula potencial evocado e percepção métrica.

Os próximos desenvolvimentos da pesquisa geral apontam para a investigação da influência do sistema motor na percepção do ritmo (IVERSEN; REPP; PATEL, 2009; PATEL; IVERSEN, 2014), com a hipótese de que o sistema motor e o sistema auditivo podem se afetar reciprocamente.

Outro aspecto importante levantado pelas pesquisas sobre comportamentos rítmicos é a mudança de conceito a respeito do senso comum de que a percepção do ritmo é um processo ativo, e não apenas passivo analítico, pois envolve a geração de expectativas e a consequente preparação do indivíduo para o movimento.

Capítulo 2

O estado da arte para a nossa pesquisa³⁰

Embora muito se tenha estudado sobre o ritmo musical e o *entrainment* nos últimos anos, os questionamentos apontados nesta tese têm um caráter um pouco diferente daqueles objetos de maior dedicação das pesquisas.

A investigação a respeito dos mecanismos de sincronização que atuam na percepção de ouvintes no início de peças musicais é relativamente pequena, e este é um momento importante onde se forma o referencial primeiro da métrica que aquela obra vai propor ao ouvinte, é onde ocorre a indução da métrica temporal da música. Assim, uma investigação relacionada a percepção de *anacruses* toca diretamente nossos questionamentos. Igualmente importante é o estudo do *mecanismo de indução* da pulsação e seus possíveis correlatos neurais. Para tanto, o recurso de medição mais comumente utilizado tem sido o eletroencefalograma (EEG).

Eletroencefalograma (EEG).

Quando se fala na técnica de EEG, voltada para a possibilidade de coleta de potenciais evocados relacionados a eventos (ERP), normalmente se pensa que os sinais coletados através da entrada sensorial acústica meramente refletem a característica do sinal físico. Porém, o experimento do grupo de Nozaradan (NOZARADAN; PERETZ, *et al.*, 2011) mostra, pela primeira vez, uma correlação maior entre o sinal *interpretado* e a estrutura métrica subjacente intencionalmente proposta, ao invés de haver uma correlação entre o sinal *físico* de entrada e uma mera reprodução endógena deste mesmo sinal.

Como foi citado anteriormente, o percurso lógico para a realização dos

³⁰ O corpo do texto que integra esta seção foi publicado na forma do artigo *Duas abordagens de pesquisa experimental em percepção rítmica*, na revista *Percepta* (v. 1, n. 2, p.103-112, 2014), da Associação Brasileira de Cognição e Artes Musicais. Disponível em: <<http://www.abccogmus.org/journals>>, com acesso em 28 jun. 2016.

experimentos medindo respostas eletrofisiológicas em EEG passa pela confirmação comportamental do fenômeno que buscamos medir. Para tanto, escolhemos duas abordagens experimentais como paradigmas da nossa pesquisa: Uma comportamental, que configura-se como uma entrevista seguida por uma coleta de dados comportamentais: outra abordagem é aquela que representa uma continuação natural da nossa investigação, a experimentação com EEG – que, apesar de não fazer parte dessa tese de doutorado, fornece a evidencia de que a intencionalidade modifica a interpretação da informação acústica contida no sinal físico e inicialmente codificada pelo córtex auditivo.

A abordagem com EEG é uma possível linha de pesquisa posterior, na medida em que tenhamos a confirmação do efeito comportamental da acentuação subjetiva.

Em seguida, à exposição dos experimentos tomados como modelo, trataremos do nosso próprio percurso experimental.

2.1 Duas abordagens de pesquisa experimental em percepção rítmica³¹

A pesquisa experimental em percepção do ritmo musical tem se beneficiado muito do desenvolvimento das neurociências e da neuropsicologia. Em virtude do avanço tecnológico, tanto o computacional como a modernização e a ampliação ao acesso de equipamentos de diagnóstico por imagens do cérebro, esse campo de estudos atualmente tem gerado conhecimento para a área de música, assim como ela própria tem sido usada como paradigma para se estudar a cognição humana.

Nesse sentido, o mecanismo da percepção do ritmo musical, apesar de apresentar-se ainda com muitos aspectos a serem desvendados, já conta com um

³¹Esta seção constitui uma ampliação do artigo publicado na revista de cognição musical *Percepta*, disponível em <http://www.abcogmus.org/journals/index.php/percepta/article/view/25> e nos Anais do X Simpósio de Cognição e Artes Musicais, 2014.

corpus de conhecimento considerável, tendo sido acumulado de forma mais sistematizada a partir da década de 70.

Apresentamos, a seguir, duas abordagens de pesquisa experimental em música, visando uma reflexão sobre suas possibilidades de aplicação. A primeira trata de uma aproximação mais tradicional e inclui a apresentação de exemplos musicais para apreciação dos sujeitos e uma avaliação estatística dos resultados obtidos. A singularidade desta abordagem refere-se ao estudo de fatores estruturais (por exemplo, intervalos duracionais) que influem na percepção de anacruses, estudando, em princípio, a variação de uma estrutura básica CLL (curto-longo-longo) e LCC (longo-curto-curto).

A segunda abordagem trata das evidências a respeito da possibilidade de captação direta, em eletroencefalograma (EEG), de sincronismo neuronal (*neuronal entrainment*) associado à pulsação e à métrica musical. Esta captação é obtida através da técnica de Potenciais Relacionados a Eventos (*Event Related Potential*, ERP) e faz uso de estímulos ambíguos que não tendem para nenhuma interpretação métrica específica.

Esta técnica EEG de Potenciais Evocados de Estado Estável (*Steady State Evoked Potential*) no estudo do sincronismo neuronal para pulso e métrica musical tem indicado uma maneira segura e não invasiva de se observar o comportamento da percepção frente a estímulos musicais estruturados a partir da perspectiva do ritmo mais tradicional, presente na música popular e na música de tradição europeia ocidental, na qual persiste uma hierarquia temporal com agrupamentos de unidades e subunidades em sua maioria binária ou ternária.

Estratégia 1: Coleta de dados comportamentais.

Essa estratégia consiste na apresentação de exemplos sonoros a sujeitos e na coleta da apreciação deles em relação aos estímulos. Focaremos essa estratégia apoiados em um experimento realizado sobre a percepção de anacruses (LONDON; CROSS; HIMBERG, 2009). Esse estudo é intitulado *The effect structural and performance factor in the perception of anacruses* e contém dois experimentos, para

os quais foram elaborados 36 estímulos, com exemplos musicais em que a estrutura rítmica consistia em duas sequências básicas de três intervalos temporais: CCL e LCC (com proporções de 1:1:2 e 2:1:1, respectivamente). O objetivo desse estudo foi avaliar a influência da interação entre fatores estruturais (ritmo e alturas) e fatores de performance (articulação e andamento) na percepção de anacruses. A hipótese era de que a percepção anacrúsica recairia principalmente sobre a estrutura CCL, onde se presume que o acento tético seja predominantemente percebido na nota longa, ou então nos sons inicial e final desta sequência, conforme as regras de preferência, onde a tendência é acentuar o primeiro e o último sons de um grupo de três ou mais sons (POVEL; OKKERMAN, 1981).

Estes dois experimentos contaram com a participação de sujeitos em três níveis de expertise musical (menos de cinco anos de estudo, de 5 a 10 anos de estudo e mais de 10 anos). O primeiro experimento consistiu em verificar as interações das duas estruturas rítmicas sem melodia e com vários tipos de estruturas melódicas (ascendentes, descendentes, por graus conjuntos, com saltos intervalares) e em andamentos distintos. O segundo experimento aproveitou os resultados do primeiro para refinar a avaliação do efeito do aumento do andamento na percepção das anacruses.

Os sujeitos foram convidados a ouvir os estímulos por três a cinco tempos antes de percutir com palmas o tempo acentuado percebido ao longo das melodias, o que permitiu a observação do tipo de alinhamento das frases com o respectivo acento forte. Através da progressão retrógrada do acento declarado, obtém-se a indicação sobre a percepção anacrúsica ou não anacrúsica do estímulo, como se pode observar nos níveis métricos pontilhados da Figura 4.

Figura 4 - Observa-se a mesma melodia com duas interpretações e níveis métricos distintos: a de cima, com interpretação anacrúsica (acento métrico fora de fase em relação ao início da melodia) e a de baixo com interpretação tética (em fase com a melodia).

No experimento 1, com estímulos isotônicos, foram encontrados indícios de que os padrões CCL são mais sensíveis a mudanças de interpretação na interação com diferentes andamentos, havendo grande impacto na mudança de sua percepção de tético para anacrúsico, quando adicionada qualquer estrutura melódica (LONDON; CROSS; HIMBERG, 2009, p. 110).

No experimento 2 verificou-se haver uma mudança sistemática na percepção da anacruse à medida que o andamento aumenta (LONDON; CROSS; HIMBERG, 2009, p. 111).

No artigo supracitado, London, Cross e Himberg lembram de duas observações feitas em experimentos anteriores, de 2004: na percepção de anacruses, algumas melodias apresentam o que eles chamam de “maleabilidade métrica”. Esta seria uma espécie de ambiguidade métrica. Também observam o comportamento diferente da percepção de estruturas rítmicas frente à alteração da velocidade no andamento. Assim, variando-se o andamento de uma estrutura rítmica de três durações, por exemplo, curto-curto-longo (1:1:2), a partir de uma determinada velocidade na articulação, a interpretação de um ouvinte tenderia a deslocar o acento métrico de uma duração para outra, tornando a estrutura mais anacrúsica ou mais tética, conforme o caso.

A correlação entre a variação de andamentos e o deslocamento da percepção métrica provocado nos ouvintes, apresentada por alguns tipos de estruturas rítmicas, foi um dos fatores importantes que guiaram a pesquisa dos três autores.

Estratégia 2: Mapeamento da sincronização adaptativa neural (*neural entrainment*) através de EEG.

O conceito de *entrainment* seria “um processo biológico que promove a sincronia adaptativa de oscilações internas da atenção com um evento externo”

(JONES, 2009, p. 83)³², ao que nos referiremos por *sincronização adaptativa*. Consequentemente, a sincronização adaptativa neural (*neural entrainment*) decorre de uma teoria aceita atualmente, segundo a qual os neurônios sincronizam com estímulos cíclicos externos, oscilando na mesma frequência. Partindo desses pressupostos, o experimento descrito no artigo *Tagging the neuronal entrainment to beat and meter* (NOZARADAN; PERETZ, et al., 2011), que poderia ser traduzido como algo próximo de “mapeando³³ a sincronização neuronal para o acento métrico e a métrica”, descreve a observação da sincronização neural através do EEG e também acaba por mapear as características do seu ciclo como binário ou ternário.

No experimento, a estratégia de utilização do EEG foi eficaz, na medida em que apresenta um estímulo acústico com pulsação de 2,4 Hz, ou 144 bpm (batidas por minuto), e fisicamente neutro, ou seja, não privilegia acentuações ciclicamente binárias ou ternárias. No entanto, pede para os sujeitos imaginarem acentuações indutoras de métricas com base primeiramente em dois e, depois, em três. O artifício utilizado foi medir o comportamento dos impulsos gerados no cérebro por estímulos auditivos externos, chamados Potenciais Evocados (*Evoked Potentials*, EP), mais precisamente com a técnica de Potenciais de Eventos Relacionados (*Event Related Potentials*, ERP). Estes são observáveis através de eletroencefalograma (EEG) e da ressonância magnética (MEG), as “duas principais técnicas de estudo da dinâmica do processamento auditivo no cérebro humano” (LARGE, 2008, p. 118)³⁴.

Os resultados obtidos nas imagens geradas pelo EEG demonstram representações cerebrais diferentes para cada uma das métricas imaginadas. Assim, a leitura possível é que as diferentes imagens da codificação cerebral, obtidas para cada um dos ciclos harmônicos do estímulo original (2,4 Hz ou 144 bpm), seja binário (1,2 Hz ou 72 bpm) ou ternário (0,8 Hz ou 48 bpm), indicam a compatibilidade com um mapeamento da qualidade de sincronização realizada pelos

³² “[...] Entrainment is a biological process that realizes adaptive synchrony of internal attending oscillations with an external event [...]” (JONES, 2009, p. 83).

³³ O termo mapear é utilizado no sentido de vasculhar e descrever, e não no sentido matemático de transformar.

³⁴ “[...] Electroencephalography (EEG) and magnetoencephalography (MEG) are the two main techniques for studying the temporal dynamics of auditory processing in the human brain [...]” (LARGE, 2008, p. 118).

sujeitos (NOZARADAN; PERETZ et al., 2011, p. 10236).

Para este experimento, a variante da técnica de ERP, denominada Potenciais Evocados de Estado Estável (*Steady State Evoked Potential*) no estudo do sincronismo neuronal para pulso e métrica musical, aponta para uma maneira eficiente de se observar o comportamento da percepção, frente a estímulos musicais estruturados, a partir da perspectiva do ritmo mais tradicional presente na música popular e na música tonal de tradição europeia ocidental, na qual persiste uma hierarquia temporal com agrupamentos de unidades e subunidades, em sua grande maioria binária ou ternária.

Potenciais relacionados a eventos.

A técnica Potenciais de Eventos Relacionados (*Event Related Potentials*, ERP) foi inicialmente nomeada *Evoked Potentials* (EP) ou Potenciais Evocados, por terem sido potenciais elétricos evocados por estímulos (em oposição aos ritmos espontâneos do EEG) (LUCK, 2005, p. 6). Posteriormente, com o desenvolvimento da neurociência cognitiva e conjuntamente com o incremento na capacidade de processamento computacional, essa técnica ganhou maior reconhecimento e sua importância está ligada à alta resolução temporal em relação às duas outras técnicas mais modernas, que são a Ressonância Magnética Funcional (*Functional Magnetic Resonance Imaging*, fMRI) e a Tomografia por emissão de pósitrons (*Positron Emission Tomography*, PET), sendo considerada um complemento a estas últimas (LUCK, 2005, p. 6).

Atualmente, para pesquisa, que envolve o processamento de eventos onde a precisão temporal é tão importante quanto no estudo sobre processamento rítmico, ainda tem sido usada a técnica de ERP, como também técnicas desenvolvidas a partir dela como podemos observar na abordagem de *frequency tagging* (NOZARADAN, 2014) explicitada no capítulo 1 e na avaliação uma das estratégias que segue.

Em resumo, entre os sinais pós-sinápticos captados dos disparos das redes de neurônios cerebrais, os potenciais elétricos evocados por estímulos (auditivos, visuais, entre outros) contêm uma forma característica de onda análoga, cujas

componentes são classificadas de acordo com a latência e a natureza dos picos (positivos ou negativos), em que são registradas, após a ocorrência do estímulo original. Por exemplo, o pico inicial nos primeiros 100 ms da onda evocada representa a resposta neuronal para o sinal físico provocado pelo estímulo, chamado também de P1 ou P100 (P para positivo) é de componente “exógeno”, ou seja, que depende mais das características de fatores externos. Já para aqueles picos que ocorrem mais tardiamente, por volta dos 300 ms (P3 ou P300), por exemplo, a componente da resposta sensorial é considerada “endógena”, pois depende mais de fatores internos (LUCK, 2005, p. 11).

Na Figura 5 podemos ver uma típica forma de onda captada em EEG, com seus respectivos picos P100 e P300, mencionados acima:

Figura 5 - Imagem representando um Potencial relacionado a evento genérico.

No gráfico (Figura 5) podemos notar que os picos considerados “positivos” estão voltados para o lado de baixo. Isso se deve ao fato de que o eixo da amplitude em μV (micro Volts) está invertido e a voltagem positiva está para baixo. Isso não é incomum de se encontrar em textos científicos, porém também é comum encontrar este eixo vertical na forma cartesiana, com a parte positiva voltada para cima.

A variante da técnica de Potenciais Evocados utilizada no experimento de EEG, denominada Potenciais Evocados de Estado Estável (*Steady-State Evoked*

Potential, SS-EP) tira proveito da periodicidade do sinal de entrada determinado experimentalmente para identificar a resposta neural no domínio da frequência, utilizando o paradigma do *imageamento mental métrico* (NOZARADAN; PERETZ, 2013, p. 103). Assim, através desta técnica se obtém as frequências predominantes de oscilação neural evocadas pela construção endógena da métrica, realizada pelos voluntários sobre a pulsação sugerida pelo estímulo externo do experimento.

2.1.1 Infraestrutura experimental

Materiais e métodos dos experimentos. Os participantes dos experimentos de percepção de anacruses foram 18, com experiência musical dividida em três categorias: menos de cinco anos de estudo com instrumento ou canto (seis voluntários), de 5 a 10 anos de estudo (quatro voluntários) e mais de 10 anos de estudo (oito voluntários).

Os participantes do experimento de EEG foram oito voluntários (sendo três mulheres e cinco homens, dentre todos sete destros) e a experiência musical abrangeu desde aqueles que tinham boa experiência musical como *performers* (três deles) até aqueles que tinham experiência como ouvintes amadores ou como dançarinos.

Experimento com anacruses. Foram desenvolvidos estímulos que abrangeram as combinações entre figuras rítmicas (CCL *versus* LCC), saltos melódicos (ascendentes *versus* descendentes), intervalos melódicos (saltos maiores do que terça menor (3m), *versus* saltos até 3m), harmonia implícita do quinto para o primeiro grau e do primeiro para o primeiro grau (V-I *versus* I-I) e variação de tempo entre os ataques das notas (*inter-onset-intervals*, IOI; 80 bpm/750 ms IOI, 105 bpm/570 ms IOI, e 140 bpm/428 ms IOI). O termo *inter-onset-intervals* (IOI) poderia ser traduzido por “intervalos-entre-inícios”, em que os intervalos de tempo entre os respectivos ataques das notas constituem a estrutura rítmica num

agrupamento temporal de eventos musicais dos estímulos estudados por London, Cross e Himberg (2009).

Experimento com EEG. Estímulos com duração de 33 segundos, compostos por sons puros (senoidais) de 333,3 Hz (o que corresponde, aproximadamente, à nota Mi3 ou E3), ciclicamente repetidos numa frequência de 2,4 Hz, ou seja, 144 bpm. Para gerar acentuações com o intuito de se produzir uma estrutura pseudoperiódica que não induzisse a percepção dos voluntários nem a uma interpretação binária e nem ternária, a solução encontrada foi bem engenhosa.

Para não ter uma frequência com razão inteira de 2,4 Hz, na modulação da amplitude foi utilizada uma função sinusoidal de 11 Hz (número primo), variando entre 0,3 e 1, para modular a amplitude das batidas em 144 bpm. Isto fez com que estímulos gerados com esta técnica, chamada de “convolução” de sinais, não crie nenhuma indução binária ou ternária nos ouvintes.

A frequência de 2,4 Hz foi escolhida por ser confortável para o ouvinte produzir mentalmente a frequência harmônica binária de 1,2 Hz e a ternária de 0,8 Hz; além disso, a mesma também se encontra dentro da banda intervalar considerada confortável para a percepção e a produção temporal (DRAKE; BOTTE, 1993).

Coleta comportamental. A estratégia da coleta de dados comportamentais inclui uma entrevista e a aquisição de dados através de uma interface computacional. A entrevista para a avaliação da verbalização da percepção de estímulos auditivos constitui a base de qualquer avaliação subjetiva³⁵ de um estímulo físico. Ela revela o tipo e o nível de interpretação que o participante faz do sinal físico que lhe é apresentado. Porém, no caso de estímulos musicais, participantes sem uma formação musical podem não distinguir declaradamente a diferença de um ciclo binário de

³⁵ Naturalmente, aos fatores subjetivos iniciais de qualquer percepção (p.ex., da métrica binária de uma marcha), agregam-se fatores interpretativos ligados ao conhecimento acumulado ao longo do tempo pelo participante, que ampliam a subjetividade inicial (em nosso exemplo, aspectos inicialmente ignorados da marcha são agora percebidos por ter havido uma reinterpretação retrospectiva do sujeito à luz de sua experiência acumulada).

outro ternário, e isso nos leva a um questionamento: a diferenciação entre um metro ternário e um binário, contida em estímulos externos, pode ser ouvida e processada mesmo que não seja conscientemente notada? Ou significa apenas que o participante musicalmente leigo não distingue nominalmente as diferenças (apesar de “percebê-las”, de alguma maneira) e com algum treinamento pode ser capaz de nomear adequadamente um ou outro? Como se dá, neurofisiologicamente falando, a interpretação da métrica musical e a consequente sincronização do sujeito com o estímulo acústico externo? A capacidade de perceber ou distinguir uma pulsação cíclica estaria vinculada diretamente a um nível mais alto de processamento da informação no cérebro? Neste ponto, uma questão interessante se coloca: até onde o processamento de um estímulo físico sonoro cíclico é resultado da ontogenia? Em outras palavras, o processamento de um sinal acústico em ciclo binário, como tal, é inato ou aprendido? Se for um pouco de cada, quanto ou que parte é resultado de sua interação com o meio?

Nesse sentido, a estratégia de avaliação através da técnica ERP, pode ser a mais eficiente em revelar diferenças na codificação de sinais binários e ternários de estímulos acústicos. E isto decorre da sua maior precisão temporal em relação ao PET e ao fMRI, que demandam muito mais processamento computacional no registro em tempo real dos dados.

O tipo de experimento com ERP (NOZARADAN et al., 2011) abre caminho para interessantes desdobramentos, indicando a possibilidade de se observar a interpretação realizada por um sujeito e dizer como o seu cérebro codificou este estímulo e em qual categoria classificou o sinal (no caso, binário ou ternário).

As consequências desse tipo de desenvolvimento para o campo da música são promissoras, abrindo oportunidades para se calibrar melhor a interpretação musical. A ampliação do conhecimento acerca da recepção de performances musicais por parte do público conduz a um aprimoramento do planejamento da performance dos executantes, que podem optar por diminuir ambiguidades rítmicas na execução de obras musicais, conduzindo a uma escuta de senso mais comum, ou por provocar intencionalmente um aumento no grau de ambiguidades métricas, se seu intuito for

provocar uma escuta mais ativa e participativa por parte do público ouvinte.

Por outro lado, mesmo com o avanço da tecnologia e a consequente queda de preços nos equipamentos, o EEG ainda não é uma ferramenta prática e portátil. Apesar do surgimento de sensores EEG para uso doméstico, inclusive em jogos e treinamentos de resposta neural (*neuro feedback*), por exemplo, as pesquisas ainda demandam instalações (*setups*) relativamente complexas. Isso, muitas vezes, dificulta medições com um número maior de voluntários. Nesse sentido, a coleta comportamental, embora seja uma estratégia menos precisa, possibilita a avaliação de um número maior de voluntários em cada realização de experimento.

Assim, uma pesquisa que contemple as duas abordagens tem maior probabilidade de se mostrar completa, hajam vistas que estas estratégias não são antagônicas, mas se complementam na tarefa de compreensão do fenômeno estudado.

Notamos também que a percepção de algumas estruturas rítmicas comporta-se de maneira diferente do esperado, conforme estudos anteriores indicaram. Tivemos a oportunidade de apreciar exemplos disso em trechos de composições de Beethoven - seja no motivo principal do início do primeiro movimento de sua *Sinfonia* n. 5 op. 67 ou no início da *Sonata* para piano op. 14, n. 2 (BONDESAN DOS SANTOS, 2012). Estes estudos envolvem questionamentos que nos levaram ao percurso experimental que veremos no capítulo seguinte.

Considerando os trechos musicais acima referidos, podemos observar que, para o ouvinte, a acentuação métrica percebida apresenta a tendência de ficar *fora de fase* com a intenção expressa na partitura pelo compositor, ou seja, o ouvinte percebe a métrica deslocada em relação àquela do compasso representado na partitura. Curiosamente, os inícios das melodias ou de peças musicais são os momentos em que identificamos a maior ocorrência de ambiguidades ou “maleabilidades” métricas. Estudos com esse tipo de abordagem foram levados a termo anteriormente por Sloboda (*The communication of musical metre in piano performance*, 1983), onde o foco temático era a performance anacrúsica e não anacrúsica de pianistas sobre uma

mesma melodia com característica métrica “maleável” ou ambígua.

Imaginamos que o início de uma peça musical é um momento crucial, com grande potencial de ambiguidade métrica, por ser um momento onde ocorre a indução da métrica musical no ouvinte ou o “treinamento” deste em relação à experiência temporal proporcionada pela música. Consequentemente, é um bom indicativo de que este aspecto configura um grande momento para estudos sobre a formação das estruturas e a melhor compreensão e parametrização do estabelecimento das expectativas temporais propostas aos ouvintes por uma determinada composição e/ou performance musical.

Capítulo 3

Percurso experimental

3.1 *Design experimental*

O processo de realização de experimentos requer uma organização metodológica de seus eventos, o que tem sido denominado *design experimental* por alguns pesquisadores da literatura científica de língua inglesa (GIL, 2002; HINKELMAN; KEMPTHORNE, 1994).

No percurso para a definição do *design experimental*, é possível classificar metodológica e didaticamente as pesquisas em três grandes grupos sequenciais: exploratórias, descritivas e explicativas, sendo que as “pesquisas exploratórias e descritivas constituem etapa prévia indispensável para que se possa obter explicações científicas” (GIL, 2002, p. 41-42). Dentro desta sequenciação, também se pode dizer que as pesquisas explicativas podem ser consideradas pesquisas causais, ou seja, que buscam demonstrar hipóteses com relação de causa e efeito entre variáveis e um determinado comportamento. Na visão de Gil, esta divisão (exploratórias, descritivas e explicativas) classifica as pesquisas com base nos objetivos gerais das pesquisas e ainda afirma sua importância como ferramenta:

[...]é muito útil para o estabelecimento de seu marco teórico, ou seja, para possibilitar uma aproximação conceitual. Todavia, para analisar os fatos do ponto de vista empírico, para confrontar a visão teórica com os dados da realidade, torna-se necessário traçar um modelo conceitual e operativo da pesquisa [...] (GIL, 2002, p. 43).

Este autor prossegue associando o **modelo conceitual** citado com o *design experimental*, como é costumeiro entre os autores ingleses desse segmento. É importante observar que, apesar de Gil tratar basicamente de experimentação científica mensurável, suas ideias incluem a pesquisa em ciências humanas. Já as ideias de Hinkelman e Kempthorne vão mais no sentido da observação do que, de algum modo, possa ser medido - ou seja, das ciências exatas, biomédicas e da psicologia. Nesse contexto, a psicologia parece ocupar a fronteira da ciência onde nem tudo é mensurável, ainda que alguma porção dos seus estudos se utilizem de

grandezas quantificáveis, sobretudo através dos estudos experimentais atuais em neurociência.

Como estamos lidando basicamente com a psicologia experimental na fronteira entre Música e Neurociência, a dificuldade principal é encontrar um modo de medir aspectos que envolvem grandezas relativas. Por exemplo, nossa pesquisa procura confirmar percepções psicofísicas de padrões de acentuação em grupamentos de eventos sonoros equitoniais. A percepção da acentuação sonora tem caráter relativo, sua natureza requer a comparação com o que se apresenta na sua vizinhança temporal, os vizinhos antecedentes e os posteriores. Outra característica sobre este tipo de percepção é sua variabilidade numa população, supondo que existam graus de variabilidade na percepção das intensidades sonoras relativas entre indivíduos. Mais à frente trataremos de propostas para solução desta questão.

Tanto Gil como Hinkelman e Kempthorne subdividem em ações o design de experimentos a saber (Figura 6):

Figura 6 - Modelo de etapas Hinkelman e Kempthorne à esquerda e o modelo de Gil à direita

Como podemos notar na Figura 6, o modelo de Hinkelman e Kempthorne é mais sucinto que o de Gil, pois o primeiro agrupa várias ações do segundo numa só instância. Enquanto o primeiro fala em formulação do problema, o segundo o subdivide em duas ações (formulação do problema e construção de hipóteses). A ação seguinte de Hinkelman, escolha do design experimental, é dividida por Gil em 5 ações (determinação do plano, operacionalização das variáveis, elaboração dos instrumentos de coleta de dados, pré-teste dos instrumentos e seleção da amostra). E as outras ações seguem iguais num modelo e no outro (execução do experimento, análise dos dados e interpretação dos dados).

Na fase exploratória inicial do nosso experimento, buscamos explorar que tipos de agrupamentos isócronos equitoniais provocam o comportamento de acentuação subjetiva de acordo com as regras de preferência descritas por Povel e Essens e resumidas por Temperley e Bartlette (POVEL; ESSENS, 1985; POVEL; OKKERMAN, 1981; TEMPERLEY; BARTLETTE, 2002). Ademais, buscamos nesses experimentos exploratórios possíveis correlações entre a variação dos intervalos temporais que separam os inícios dos IOI e a percepção da acentuação subjetiva.

Pudemos observar que, na prática, as etapas do *design* experimental propostas por Gil (2002) não são estanques, uma vez que no início de nossa elaboração não tínhamos conhecimento formalizado destas etapas, e procedemos mais de modo intuitivo do que estes modelos demonstram. O contato com estas informações e com este modelo organizado de explicitar as etapas ocorreu quando os experimentos já estavam em andamento. Por essa razão, ao início desta seção, classificamos a proposta de Gil (2002) como didática, além de ser metodológica. No entanto, conseguir nomear as fases do processo ajuda na identificação de soluções para as questões que vão surgindo no percurso experimental, e isto é de extrema importância nos momentos em que é preciso tomar decisões rápidas e acertadas.

No plano de formalização do nosso *design* experimental, fez-se necessária a confecção de um protocolo experimental que, grosso modo, constitui-se na descrição detalhada de todas as etapas e materiais utilizados no processo investigativo, sendo

apresentado como uma documentação da proposta para efeito de avaliação da ação por comitês de ética, cuja aprovação valida institucionalmente os posteriores resultados de pesquisas experimentais. No nosso caso, como o experimento é direcionado a seres humanos, foi inscrito digitalmente na Plataforma Brasil³⁶ e, simultaneamente, em papel, junto à Comissão de Ética em Pesquisa Envolvendo Seres Humanos (CEPSH) do Instituto de Ciências Biomédicas da USP.

Soma-se a isso o fato da elaboração do protocolo experimental ser imprescindível à auto-organização do pesquisador para a realização do experimento.

“A maioria dos experimentos são de natureza exploratória, em razão do interesse em descobrir quais fatores influenciam no resultado de um determinado processo”³⁷ (HINKELMAN; KEMPTHORNE, 1994, p. 30-31). Em nosso caso, não é diferente, uma vez que nossa investigação também foca os fatores que têm influência no processo de percepção de ritmos musicais. Hinkelman e Kempthorn especificam:

[...] A experimentação é essencialmente um processo sequencial. Um experimento conduz a outro, na medida em que alguns *insights* são obtidos a partir do processo e novas perguntas vão sendo feitas. Um experimento exploratório como descrito acima, pode ser seguido por aquilo que podemos chamar de um experimento de confirmação[...] (HINKELMAN; KEMPTHORNE, 1994, p. 30-31)³⁸.

No caso do presente trabalho, a etapa exploratória consistiu na avaliação de grupamentos simples de quatro elementos isócronos equitoniais apresentados para os participantes. Procuramos estabelecer se haveria uma relação de causa e efeito envolvendo a variação do intervalo temporal entre os elementos e a percepção de uma acentuação endógena (não presente no estímulo acústico). A hipótese era de que

³⁶ Disponível em: <<http://aplicacao.saude.gov.br/plataformabrasil/login.jsf>> Acesso em: nov. 2015

³⁷ “[...] Most experiments are of an exploratory nature in the following sense. The investigator is interested in finding out what factors have an influence on the outcome of a certain process. [...]” (HINKELMAN; KEMPTHORNE, 1994, p. 30-31)

³⁸ “[...] Experimentation is essentially a sequential process. One experiment leads to another as some insight is gained from a process and new questions are being asked. An exploratory experiment as described above, may be followed by what we may call a confirmatory experiment. [...]” (HINKELMAN; KEMPTHORNE, 1994, p. 30-31)

quanto menor o intervalo temporal entre os elementos, portanto o grupo com pulsação mais rápida, o efeito endógeno se faria mais evidente. As observações se mostraram inconclusivas e a possível causa, creditamos ao fato de que grupos de duração tão pequena (de 800ms a 3,2s) não seriam suficientes para provocar a indução do efeito de acentuação endógena.

O passo seguinte dentro deste processo, partiu de duas premissas básicas:

- necessidade de aumentar a duração dos estímulos para testar o surgimento do efeito endógeno;
- confirmação do efeito endógeno e suas regras de acentuação, reproduzindo os padrões utilizados nos experimentos de acentuação subjetiva (POVEL; ESSENS, 1985).

A solução encontrada para aumentar a duração dos estímulos foi buscar estímulos mais complexos e, se possível, já utilizados em outros experimentos do tipo.

Neste sentido, o próprio experimento de Povel e Essens (1985) já continha uma lista de estímulos usados para avaliar a possível indução da formação de “relógios internos” em sujeitos. Os estímulos estão classificados por categorias que seguem gradativamente da categoria 1, de melhor indução (em que os elementos acentuados coincidem com os tiques do relógio) para a categoria 9, indicando uma indução mais fraca do relógio interno (estímulos que apresentam pelo menos dois silêncios nas posições onde ocorreriam os tiques do relógio) (POVEL; ESSENS, 1985, p. 419).

Uma idéia interessante que nos ocorreu foi a elaboração de uma interface visual com a representação dos estímulos de forma circular, o que em princípio consideremos ser mais intuitiva do que a representação linear, das partituras musicais.

A segunda fase também se mostrou um tanto inconsistente e, a partir da

supervisão do Prof. Ranvaud, vimos que a aplicação de experimentos psicofísicos necessita seguir alguns requisitos, para a obtenção de resultados confiáveis:

1. É preciso ter certeza de que o participante entendeu a questão;
2. É preciso se assegurar de que ela diga o que entendeu da proposta da tarefa;
3. É preciso assegurar que ela diga o que de fato pretendia.

3.2 Metodologia de medição das respostas coletadas nos Experimentos

Estabelecer uma escala de medição da acentuação nos estímulos escolhidos exigiu desenvolver um modo de comparar a percepção dos participantes com as regras de acentuação subjetiva (que se constitui como uma referência de nosso texto teórico). Assim, buscamos³⁹ modos de se estabelecer algum grau de comparação entre as acentuações esperadas, segundo as regras de preferência para os estímulos dados pela literatura (POVEL; ESSENS, 1985) e as acentuações atribuídas aos mesmos pelos participantes dos grupos escolhidos.

A coleta dos dados gera uma sequência de **zeros** e **uns**, conforme o Participante (P_x) percebe elementos acentuados (1) e não-acentuados (0), para cada posição na sequência, que chamaremos de vetor. No nosso caso, as sequências utilizadas como estímulos são compostas, predominantemente, de oito elementos agrupados em diversas configurações.

Consequentemente, temos oito vetores para cada sequência, onde cada vetor representa a percepção de um acento (p) numa posição (i) dentro da sequência ($p[i]$), em que i corresponde à posição do vetor dentro da sequência (no nosso caso, com i

³⁹ As soluções matemáticas surgiram de conversa junto ao Prof. Dr. Marcelo Queiroz (IME-USP), a quem agradecemos imensamente a generosidade e disposição em ouvir o problema e pensar em soluções relativamente simples e muito eficazes no estabelecimento de uma metodologia para quantificar as respostas obtidas nos nossos experimentos.

variando de 1 a 8).

As sequências padrão de Povel e Essens, podem ser igualmente codificadas de acordo com a acentuação esperada segundo as regras de preferência.

Assim, por exemplo, podemos comparar o grau de similaridade entre a acentuação percebida na posição 1 da sequência (**p[1]**) e seu correspondente esperado na posição 1 da sequência padrão, que chamaremos de (**q[1]**), os quais são melhor visualizados na Figura 7.

Figura 7- Conversão das acentuações em sequências de “zeros e uns” coletada dos participantes através do computador e processada em planilha Excel, com a fórmula de similaridade citada.

A similaridade entre as acentuações esperadas nos estímulos usados por Povel e Essens, e aquelas percebidas pelos nossos participantes, pareceu ser uma boa direção a seguir. Consequentemente, o Prof. Queiroz sugeriu uma fórmula apoiada no **delta de Kronecker**, cuja notação é definida por:

$$\delta_{p[i] q[i]} = \begin{cases} 1, & \text{se } p[i] = q[i] \\ 0, & \text{se } p[i] \neq q[i] \end{cases}$$

Assim, a fórmula sugerida para similaridade entre um estímulo padrão (q) com oito elementos (vetores) em relação a outro percebido (p), onde "i" significa a posição do vetor na sequência, fica dessa maneira:

$$S = \frac{\sum_{i=1}^8 \delta(p[i], q[i])}{8}$$

O oito do divisor, nesse caso, é o número total de vetores da sequência. Se, porventura, cada posição da sequência tiver a acentuação exatamente igual, considerados o padrão e a acentuação percebida pelo participante, a soma será 8 dividido pelo número de vetores da sequência (8) e o resultado será igual a 1, indicando similaridade total entre padrão e percepção do participante naquela tentativa (T_x).

Outra fórmula mais simples de compreender, para calcular o que chamaremos de **grau de similaridade**, é 1 menos o módulo da “distância” entre o par vetor padrão-vetor participante; ou seja, se forem iguais, significa que a distância entre um e outro é **zero** (ou seja, similaridade máxima) e, se forem diferentes, a distância entre um e outro será 1 ou -1 (não similaridade). Então, para a somatória da diferença de cada par de vetores, teremos uma medida de aproximação entre estes e, quando os dividimos pelo número total de vetores, chegamos a um grau de similaridades entre as sequências.

Segue a fórmula derivada daquela com o **delta de Kronecker**:

$$S = \frac{\sum_{i=1}^8 (1 - |q[i] - p[i]|)}{8}$$

A fórmula acima determina a similaridade através da medição da “distância” entre os dois vetores, pois, se forem iguais, o módulo da subtração de $q[i]$ por $p[i]$ resulta em 1, então $(1 - 1)$ resulta em 0, ou seja, estes vetores são similares.

Existe a probabilidade também de que nossa população manifeste a percepção de acentuação subjetiva seguindo padrão diverso daquele manifestado nas Regras de Preferência (POVEL; ESSENS, 1985).

O modo que encontramos de tentar quantificar a possibilidade deste padrão construído a partir da média das respostas dos participantes foi elaborar um índice para as acentuações mais percebidas em cada posição do vetor para cada população.

Como a média populacional para cada posição de vetor não pode ser associada diretamente, sempre há uma lógica binária ou há um valor absoluto (0 para vetor sem acento e 1 para o com acento), é possível a aplicação do que em matemática é chamado de lógica difusa ou logica “fuzzy”.

Neste tipo de lógica, as proposições podem ter valores intermediários entre o “verdadeiro” e o “falso” – no nosso caso, valores intermediários entre “acentuado” e “não acentuado”, supondo variabilidade de percepções numa população determinada. Esta lógica tem o objetivo de modelar raciocínios imprecisos e acreditamos poder tomar estas percepções subjetivas como um tipo de “raciocínio” com um grau de imprecisão razoável, em que as respostas podem ser relativas. A percepção da acentuação sonora não é um valor absoluto, posto que somente existe em relação à sua vizinhança, como acontece com outras modalidades de percepção psicofísica, como por exemplo, a percepção de “quente” pode ser classificada em relação à percepção do “frio” e do “morno”.

Segundo neste raciocínio, a construção de um padrão de acentuação subjetiva (q) para uma determinada população pode ser traduzida matematicamente como a somatória das percepções de todas as respostas de todos os participantes P (de 1 a N) para cada posição do vetor (i), onde i vai de 1 até o último vetor da sequência.

$$q[i] = \frac{\sum_{j=1}^N P_j[i]}{N}, \quad \forall i \in \{1, 8\}$$

Consequentemente, é possível construir um padrão de acentuações como no exemplo a seguir:

$$\begin{aligned} \mathbf{i} &= \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{matrix} \\ \mathbf{q} &= \begin{matrix} \frac{40}{40} & \frac{10}{40} & \frac{15}{40} & \frac{38}{40} & \frac{30}{40} & \frac{3}{40} & \frac{12}{40} & \frac{31}{40} \end{matrix} \\ \mathbf{p} &= \begin{matrix} 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \end{matrix} \end{aligned}$$

Supondo que o número total de participantes seja 40, para cada vetor $\mathbf{i}$, teríamos que o padrão $\mathbf{q}[\mathbf{i}]$ será a razão entre a somatória das percepções dos acentos de todos os participantes $\mathbf{P}[\mathbf{i}]$ para cada vetor e o número total dos participantes $N=40$. Os resultados para cada posição seriam números entre 0 e 1. Então, para definir um padrão de acentuações em cada posição da sequência dada no exemplo, é preciso estabelecer um “gatilho” para considerar determinada posição como acentuada ou não.

Se considerarmos para a sequência do exemplo um gatilho de 50%, ou seja para se para a população estudada, a média das percepções acentuadas daquela posição do vetor for maior ou igual a 0,5, teremos que, acima da razão 20/40, a posição é considerada acentuada, abaixo desta razão 20/40, a posição é considerada não acentuada.

3.3 Experimentos comportamentais

A confirmação da efetividade de um efeito psicofísico como esse que estamos estudando, relaciona-se com o comportamento esperado e elencado nos experimentos de Povel e Essens (1985) onde a acentuação subjetiva, se expressa na forma do Modelo de Regras de Preferência exposto anteriormente no capítulo 1.

Assim sendo, nosso percurso experimental compreendeu três fases, inicialmente com um experimento exploratório chamamos de Fase 1, onde testamos a emergência do efeito em estímulos com apenas um grupo de sons equitonais e isócronos de quatro e três elementos. A Fase 2 representou um aprimoramento do ganho de conhecimento da Fase 1, através da mudança da abordagem dos estímulos, tomados da literatura (POVEL; ESSENS, 1985) onde não refazemos exatamente aquele experimento. Enquanto Povel e Essens consideravam as Regras de Preferência como um determinante para o estabelecimento da percepção da pulsação métrica, e avaliavam se padrões rítmicos com as acentuações das Regras de Preferência, concomitantes com a pulsação, seriam mais facilmente reproduzidos que outros padrões cujas acentuações fossem deslocadas para não coincidir com a pulsação esperada.

Nossa abordagem a partir da Fase 2, procura exatamente avaliar se uma pequena amostra populacional de dois cursos universitários da cidade de São Paulo, com estudantes de música como um grupo controle, e outros estudantes sem imersão na formação musical ou sem vivência musical nenhuma, confirmam o comportamento da percepção de acentos expresso nas Regras de Preferência, frente aos mesmos estímulos usados por Povel e Essens (1985).

3.3.1 Fase 1: Experimento exploratório 1, com estruturas de 4 elementos

Esta foi a fase exploratória do nosso experimento. Significou, basicamente, uma primeira abordagem das perguntas iniciais, acerca da percepção de acentos relativos ao grupo de 4 elementos que equivale aos 4 elementos do motivo principal da Sinfonia n. 5 de Beethoven.

Ao proceder o experimento exploratório com estruturas equitonais de 3 e 4 elementos, vimos que não era suficiente o tempo de duração para haver a indução do efeito de acentuação subjetiva. Consequentemente, nosso foco de investigação foi direcionado para os limites da indução da acentuação subjetiva. Assim, passamos para um desenvolvimento ainda exploratório que constituiu a Fase 2 dos experimentos.

Precisávamos testar um tamanho mínimo para o surgimento do fenômeno da percepção da acentuação subjetiva. Recorremos, então, a padrões rítmicos já utilizados em experimentos relacionados ao fenômeno (POVEL; ESSENS, 1985).

3.3.1.1 Método

3.3.1.1.1 Procedimento

Neste experimento, testamos as respostas comportamentais de dois participantes com estímulos contendo grupos de 4 elementos isócronos e equitonais que passaremos a designar SEQ4.

Objetivo. Verificar padrão de acentuação subjetiva em grupos de 4 *beeps* isocrônicos, de mesma intensidade, em função de diferentes IOIs (intervalo entre os *beeps* de uma sequência).

Hipótese 1. Haverá padrões de acentuação subjetiva em grupos de 4 elementos?

Hipótese 2. Havendo padrões de acentuação subjetiva nestes grupos de 4 elementos, eles sofreriam influência da variação de IOIs?

Participantes. Dois participantes saudáveis (2 homens, destros; faixa etária, 49-56 anos; média de idade, 52,5) tomaram parte deste exploratório. Eles têm experiência musical em performance (35,5 anos de prática, em média) e formação superior em música (1 na Graduação e no 1 Mestrado). Não possuem histórico de problemas neurológicos ou psiquiátricos, porém um deles relatou ter episódios de zumbido no ouvido e 5% de perda de audição. Não estavam tomando nenhum tipo de remédio no período da realização do experimento.

Estímulo audiovisual. O PF (ponto focal)⁴⁰ aparece. Logo em seguida, ouve-se no fone de ouvido a SEQ4 sendo tocada 4 vezes (as quatro SEQ4 são idênticas), com intervalos entre as SEQ4 de 3 a 4 segundos. Em seguida na Figura 8, há a apresentação visual de 4 colunas (retângulos verticais lado a lado).

Figura 8 – Imagens correspondentes às telas do experimento da Fase 1 com grupos de sons isócronos equitoniais de quatro elementos e nomeados SEQ4. A tela da esquerda mostra as quatro barras, representando cada elemento do grupo e na imagem da direita as barras preenchidas parcialmente em branco, com o nível de acentuação relativa assinalada por um participante, no caso específico com os dois primeiros sons não acentuados e os dois últimos igualmente acentuados.

⁴⁰ Como o experimento foi concebido para ser aplicado juntamente com EEG, previmos a utilização do ponto focal, que significa um ponto colocado na tela do computador, onde o participante é instado a olhar fixamente. O Ponto Focal (PF) é utilizado em experimentos de EEG para reduzir os “ruídos” eletrofisiológicos provocados pelo movimento dos olhos, melhorando a qualidade dos dados coletados.

Este estímulo é apresentado em 4 tentativas (*trials*) para 4 IOIs determinados em 200ms, 400ms, 600ms e 800ms, sorteados em ordem aleatória.

Assim, temos os parâmetros principais:

nrep = 4; #número de respostas p cada condição é 4

IOIs = 0.2 0.4 0.6 0.8 ms ; # as condições são os intervalos entre beeps

ntrial = 4; # número de apresentações de grupos em cada condição 4

nbeep = 4; # número de beeps 4

Resultados.

O gráfico da Figura 9 mostra as respostas comportamentais da acentuação relativa percebida pelos participantes 9 e 11, de cada um dos grupos de 4 elementos apresentados nas 4 condições de IOI, ou seja, com intervalos de 200ms, 400ms, 600ms e 800ms, onde a variação desses intervalos corresponde a pulsos musicais que vão do mais rápido – 200ms equivalentes a 300BPM (batidas por minuto) – ao mais lento – 800ms (75BPM), passando por 400ms (150BPM) e 600ms (100BPM).

No eixo vertical (y), temos o nível de acentuação relativa, porque a acentuação percebida, além de não ter um valor absoluto, é comparada aos elementos circunvizinhos.

No eixo horizontal (x), vemos a representação da posição de cada evento no grupo ouvido, ordenado de 1 a 4. Acima de cada quadro, temos o valor do IOI respectivo (200ms, 400ms, 600ms e 800ms).

Figura 9 - Comparação dos resultados obtidos entre dois participantes R9 e R11, respectivamente, mostram sua avaliação do nível da acentuação relativa percebida (no eixo das ordenadas y), referente a cada um dos 4 elementos do grupo (no eixo x), representados pelos números de 1 a 4.

Cabe ainda, esclarecer que cada uma das linhas mais finas representa o nível das acentuações percebidas em uma apresentação, sendo que a ampliação em detalhe de uma das respostas na Figura 10 mostra 4 linhas finas correspondendo a 4 respostas de 4 grupos, apresentados e uma quinta linha mais grossa, representando a média aritmética das respostas.

Figura 10 – Respostas dos *trials* do participante 9 representadas pelas linhas finas e a linha grossa mostrando a média aritmética do conjunto.

De acordo com as regras de preferência, portanto, com a tendência a acentuar subjetivamente o primeiro e o último elemento num grupo com 3 ou mais elementos isócronos e equitoniais, os resultados obtidos não refletem claramente as expectativas como mostram os gráficos da Figura 9.

Revelam apenas um indicio de maior acentuação relativa entre o primeiro e o último elemento na faixa de 200ms, no participante 9 (nomeado R9 no gráfico da Figura 9) e um pouco menor no participante 11 (R11) . Nas faixas mais lentas, a partir de 400ms, a diferença de acentuação entre o primeiro e o último elemento ou diminui ou fica praticamente irrelevante, indicando que praticamente não há percepção de acentuação subjetiva dessa faixa em diante, ou seja, 400ms ou maior.

Nas respostas do participante 11, a partir de 400ms os resultados indicam a

ocorrência de um padrão binário com efeito “tic-tac”, como se fosse uma acentuação percebida no movimento do ponteiro de um relógio.

Com o objetivo de testar o grau de atenção dos participantes, propusemos estímulos com acentuação real no segundo e no terceiro elementos da SEQ4. Conforme a literatura indica (POVEL; OKKERMAN, 1981, p. 6), para ser percebida com eficácia dentro de um agrupamento, o elemento acentuado precisa de um acréscimo de 4dB ⁴¹ para compensar os acentos circunvizinhos. Estas duas configurações propostas contradizem a tendência de acentuação subjetiva de se acentuar o primeiro e o último elemento da sequência. Denominamos este artifício de *catch trial* (“pegadinha”), e o comportamento esperado é que o sujeito seja capaz de perceber a acentuação real.

Os resultados obtidos através das *catch trials* (Figura 11) revelam que os dois participantes demonstraram grande foco atencional, ao acertar quase na totalidade onde se encontravam os acentos reais.

Figura 11 - Resultados (catch trial) nos retângulos da terceira e da sexta linha da tabela acima indicam grande foco atencional, próximo de 100%.

⁴¹ “[...] The data indicate, first, that the interval-produced accent is quite a robust phenomenon. An increase of about 4 dB in intensity is needed to balance it. The intensity difference limen for 1.000-Hz tones at 50 dB is less than .7 dB (Riesz, 1928). A tone with a 2-dB higher intensity than the surrounding tones will invariably be perceived as accented (Thomassen, Note 2) [...]” (POVEL; OKKERMAN, 1981, p.570)

Os *catch trials* denominados *catch2* e *catch3* referem-se respectivamente a acentuações reais no segundo e terceiro elementos das sequências e como se pode observar nos respectivos resultados, a acentuação percebida nos mesmos apresenta coerência com as acentuações reais propostas.

3.3.2 Fase 1: Experimento exploratório 2, com estruturas de 4 elementos e de 3 elementos

No segundo experimento exploratório, testamos as respostas comportamentais de uma voluntária com estímulos contendo grupos de 4 e 3 elementos isocrônicos e equitoniais, que passaremos a designar SEQ4 e SEQ3, respectivamente.

A proposta do experimento é a mesma do anterior mas, devido à manifestação de um dos participantes do experimento anterior de que o teste foi longo e cansativo, tendo sido de 19 minutos para o participante 11 (R11) e de 24 minutos para participante 9 (R9, que não o considerou cansativo), resolvemos alterar seu desenho. No primeiro, diminuímos de 4 para 2 o número de respostas para cada condição com SEQ4 e diminuímos, para SEQ3, tanto o número de apresentações, de 4 para 3, como o número de respostas (*trials*), de 4 para 2.

Por outro lado, aumentamos o número de IOIs testados, adicionando 300ms e 700ms aos IOIs testados anteriormente. Assim, reduzimos para 11 minutos o exploratório SEQ4 e para 14 minutos o SEQ3. Com o tempo menor, pudemos realizar os dois testes numa mesma seção.

Neste último, apesar das sequências serem menores, o tempo de duração foi maior, devido ao tempo maior dedicado pela voluntária a refletir sobre as respostas.

Houve o relato de que o ponto focal (PF) “divide a atenção” e ela preferia fechar o olho para ouvir melhor.

Objetivo. Verificar padrão de acentuação subjetiva em grupos de 4 e 3 *beeps* isocrônicos, de mesma intensidade, em função de diferentes IOIs (intervalo entre os *beeps* de uma sequência).

Hipótese 1. Haverá padrões de acentuação subjetiva em grupos de 4 elementos?

Hipótese 2. Havendo padrões de acentuação subjetiva nestes grupos de 4 elementos, eles sofreriam influência da variação de IOIs?

Participantes. Um participante (1 mulher, destra; idade, 25) tomou parte no experimento. Ela tem experiência musical (19 anos de prática) e estudo superior em música (Graduação em andamento). Não possui histórico de problemas auditivos, neurológicos ou psiquiátricos. Não estava tomando nenhum tipo de remédio no período da realização do experimento.

Estímulo audiovisual. O PF (ponto focal) aparece. Logo em seguida, ouve-se no fone de ouvido a SEQ4 sendo tocada 4 vezes (as quatro SEQ4 são idênticas), com intervalos entre as SEQ4 de 3 a 4 segundos. Em seguida, há a apresentação visual de 4 colunas (retângulos verticais lado a lado). Para as SEQ3 ocorre idêntico ao SEQ4, com a diferença que após a apresentação dos estímulos sonoros, a tela do computador apresenta aos participantes uma representação visual dos estímulos contendo três retângulos verticais lado a lado, onde a tarefa é registrar o acento percebido correspondente.

Este estímulo é apresentado em 2 tentativas (*trials*) para cada uma das 6 condições de IOIs, determinados em 200ms, 300ms, 400ms, 600ms, 700ms e 800ms, sorteados em ordem aleatória.

Para o grupo de 4 elementos (SEQ4), os parâmetros principais são:

nrep = 4; #número de repetições do grupo p cada condição (IOI)
IOIs = 0.2 0.3 0.4 0.6 0.7 0.8 ms ;# intervalo entre beeps
ntrial = 2; # número de trials 2
nbeep = 4; # número de beeps 4

Os resultados apresentados para SEQ4 foram os seguintes (Figura 12):

Figura 12 - Resultado de experimento com estímulos de 4 elementos isócronos equitoniais agrupados, mostra acentuação percebida pelo participante.

Para o grupo de 3 elementos (SEQ3), os parâmetros principais são:

nrep = 3; #número de repetições do grupo p cada condição (IOI)
IOIs = 0.2 0.3 0.4 0.6 0.7 0.8 ms ;# intervalo entre beeps
ntrial = 2; # número de trials 2

nbeep = 3; # número de beeps 3

Os resultados apresentados para SEQ3 foram os seguintes (Figura 13):

Figura 13 - Resultado de experimento com estímulos de 3 elementos isócronos equitoniais agrupados.

Resultados. Nas SEQ4 com andamento a partir de 300ms do experimento exploratório 2 houve grande compatibilidade entre os resultados obtidos e os esperados, em referência às regras de preferência. Falha ocorrida na coleta da segunda resposta em 300ms não comprometeu tais resultados e ocorreu por erro do sistema. Já os *catch trials* indicam que a atenção não estava com foco em grau elevado, tendo havido uma diminuição no número de acertos para o patamar de 50% para *catch2* e de 60% para *catch3*.

Nas SEQ3, há a tendência de acentuação para o primeiro e/ou terceiro elemento, exceto na condição de IOI de 300ms. O *catch trial* teve índice de acerto de quase 100%, havendo apenas uma resposta sem correspondência com o acento real.

Fase 1: Análise dos resultados dos experimentos exploratórios

Considerando os três grupos de respostas para SEQ4, sendo duas do primeiro exploratório e uma do segundo, podemos dizer que os indícios apontam para inconsistência do efeito da acentuação subjetiva nos dois primeiros, e um pouco mais de consistência no segundo exploratório, com exceção da faixa de 300ms, onde a resposta foi contrária à regra da acentuação subjetiva. A acentuação subjetiva da regra de preferência ou não foi confirmada, ou recaiu sobre o segundo elemento.

O fato de termos resultados em discordância com as expectativas das regras de preferência, leva a duas reflexões:

1. Os agrupamentos de 3 e 4 elementos equitoniais isócronos são breves demais para induzir alguma métrica e, conseqüentemente, o efeito de acentuação subjetiva.
2. Se é verdade que esses agrupamentos não induzem nenhum tipo de métrica, pode-se dizer que eles tendem a uma métrica ambígua, gerando em cada ouvinte uma interpretação individualizada.

O item 2 fortalece a hipótese da ambiguidade métrica no motivo principal do início do primeiro movimento da Sinfonia n. 5 de Beethoven, anteriormente questionada (BONDESAN DOS SANTOS, 2012), em razão da estrutura rítmica daquele motivo ser semelhante à estrutura testada nos experimentos SEQ4.

3.3.3 Fase 2: Experimento com interface circular

Na fase 2, as mudanças envolveram a forma de apresentação do experimento, e a natureza dos estímulos, que passaram a ser mais complexos, compostos por grupos de sons equitoniais de 1 a cinco elementos, visando testar a resposta comportamental da acentuação subjetiva nos participantes em relação às Regras de Preferência observadas na literatura.

Estes estímulos, além de serem constituídos por diferentes grupos de eventos equitoniais separados por pulsações ou grupos de pulsações isócronas, em sua representação visual para os participantes passaram a ter uma interface circular, como podemos ver no exemplo da Figura 14,, onde o estímulo representado corresponde aquele da Figura 17 extraído do artigo de Povel e Essens (1985, p. 416) e nomeado daqui para frente de Estímulo 8 (E8).

Figura 14 - Exemplo da tela da interface linear circular contendo o estímulo E8, correspondente a representação linear da Figura 16

Em princípio a circularidade da representação dos padrões rítmicos nos pareceu intuitiva e apropriada, considerando a leitura por parte de participantes sem familiaridade com a representação temporal linear como a da escrita musical. Também levamos em conta que noção de ciclo rítmico possivelmente também seria favorecida na representação circular.

3.3.3.1 Materiais e Métodos

Nossa infra estrutura experimental foi composta dos seguintes materiais:

- Fone de ouvido Beyerdynamic DT 990 PRO (250ohm)
- Computador MacBook Pro, Processador 2.5GHz Intel Core i5, 4GB RAM, Sistema operacional, Versão OS X 10.9.5
- Psychtoolbox-3 - Psychophysics Toolbox Version 3 (PTB-3)⁴² e Software GNU Octave⁴³, versão 3.8.0 - Copyright (C) 2013, configurado para "x86_64-apple-darwin13.0.0".
- Mouse USB Microsoft.

3.3.3.1.1 Procedimento

Cada participante primeiro assistiu um vídeo demonstrativo, para que se familiarizasse com o processo. Eles foram posicionados à frente de um computador para visualizar a representação gráfica das sequências de cada estímulo. Cada sequência foi apresentada através de fones de ouvido em que os estímulos correspondentes são apresentados quatro vezes para os participantes, sendo que na primeira apresentação (nomeada T1 para tentativa 1) os participantes ouvem uma vez a sequência, na segunda apresentação (T2), duas vezes a sequência, na terceira (T3), três vezes e na quarta (T4), quatro vezes.

Ao término de cada tentativa, o participante foi instruído a marcar na

⁴² Disponível em: < <http://psychtoolbox.org/> >. Acesso em: out. 2015.

⁴³ Disponível em: < <http://www.gnu.org/software/octave/> >. Acesso em: out. 2015.

representação gráfica da sequência aqueles elementos que considerou acentuados e a apertar o botão verde do canto esquerdo da tela do computador, acionado por mouse USB, para passar para a próxima tentativa.

O grupo de participantes foi formado por oito alunos do curso de audiovisual da Escola de comunicações e Artes da Universidade de São Paulo, sendo cinco homens e três mulheres, todos destros, com idades que variaram de 17 a 36 anos; a média de idade foi de 22 anos ($DP = 6,11$)⁴⁵. Quatro deles possuem de 4 a 6 anos de prática de instrumento musical (violão, piano e percussão) equivalente a um estudo informal de música e todos participaram do experimento individualmente.

É preciso esclarecer que entendemos estudo informal de música pelo aprendizado musical que leva em conta preferencialmente a transmissão oral do conhecimento, aquele que se relaciona mais com a memorização auditiva do aprendiz, que com a decodificação da escrita musical tradicional. Ao aprendizado desta última no Brasil, consideramos que existe um grande peso do treinamento da percepção do ritmo no sentido da acentuação ser fortemente relacionada à estrutura do compasso musical. Este tema apresenta muitas nuances e merece ser estudado com mais profundidade em trabalhos futuros.

3.3.3.1.2 Estímulos

Os estímulos são compostos por nove sequências equitonais de Povel e Essens (Perception of Temporal Patterns, 1985), sendo 8 dos exemplos retirados do experimento, p. 420 (figura 15) e um da p.416 (figura 16) com *beeps* de 660Hz, onda senoidal e duração de 50m, IOI de 183ms. Os oito estímulos representados na figura 15, contem oito sons equitonais - posteriormente considerados vetores de acentuação - para efeito de coleta de dados e comparação entre as acentuações percebidas pelos participantes e aquela esperada de acordo com as regras de preferência RP. O estímulo da figura 15 contém apenas sete sons equitonais.

⁴⁵ DP significa Desvio Padrão Amostral e serve para indicar o quanto a amostra populacional se dispersa em torno da média da amostra. Um baixo valor (nesse caso 6,11) indica que a idade da maioria dos indivíduos da nossa amostra tem idade próxima da média de 22 anos.

Cada estímulo foi nomeado de acordo com a categoria a que pertence, sendo a categoria 1, usada para classificar muitos eventos coincidindo com a pulsação esperada, gradativamente até a categoria 9, para menos eventos coincidindo com a pulsação esperada. Assim o estímulo da categoria 1 foi chamado E1, o da categoria 2 foi chamado E2, e assim por diante. Como no artigo não há exemplo da categoria 8, nomeamos E8 o estímulo (figura 16) usado como exemplo na p. 416 do referido artigo.

Os sons equitonais por sua vez constituem um timbre senoidal de 660Hz e duração total de 50ms.

Figura 15 - Estímulos-exemplos das categorias retirados do artigo de Povel e Essens p.420 e utilizados no nosso experimento.

Fig. 2. A temporal pattern as presented (a), with the accent rule applied (b) indicating the perceived accents.

Figura 16 - Estímulo-exemplo retirado do artigo de Povel e Essens p.416 e utilizado no nosso experimento com o nome E8.

3.3.3.1.3 Medidas

Os participantes ouvem os estímulos e assinalam as acentuações nos elementos das sequências. O interesse aqui é medir o grau de similaridade das acentuações percebidas pelos participantes músicos e não músicos com o padrão das regras de preferência, e também verificar se há a formação de padrão próprio da população em relação ao da literatura, através do cálculo proposto na metodologia do item 3.2.

3.3.3.2 Resultados

Os resultados totais, encontram-se com a totalização das médias das quatro tentativas de cada participante (Figura 18 da pg. 95) e no ANEXO I de forma individualizada por tentativas e por participantes (Figura 31 e Figura 32, pgs. 121 e 122).

Adotaremos a estratégia de apresentar nas discussões apenas os gráficos demonstrativos das observações em questão.

3.3.3.2.1 Discussão

No Estimulo 1 observamos o índice de similaridade permanecer alto, ou aumentar a medida em que as repetições aumentam (de T1 para T4), com os

participantes P6, P7 e P8. P1 e P5 apresentam similaridade baixa, já na Tentativa 1 (T1) e se mantém abaixo ou na faixa dos 37,5% de similaridade. P2 e P3, oscilam e caem na última tentativa T4, ficando igual ou abaixo de 37,5% de similaridade como se pode ver na Figura 17.

No geral tem um grupo que aumenta o índice da percepção similar ao padrão RP e outro grupo diminui o índice à medida que é exposto a mais repetições do padrão contido no estímulo. Estes resultados podem ser vistos nos gráficos da Figura 18 e indicam a tendência observada em todos os experimentos que realizamos, não há correlação entre o número de vezes que um estímulo é repetido, e a emergência do efeito da percepção da acentuação do padrão RP.

Figura 17 - Resultados dos índices de similaridade (eixo y) do estímulo E1, obtidos individualmente por cada Participante (legendas coloridas) relativos a cada tentativa T1 a T4 (eixo x)

Observamos tendências que dividem os alunos do audiovisual em dois grupos: um que nunca teve nenhum tipo de estudo de música e outro com algum estudo informal de música.

Dos participantes com estudo informal de musica tivemos três que estudaram

violão e um piano, uma media de 3,63 anos (DP = 2,28).

Os dois subgrupos separados nos gráficos (Figura 18) mostram um razoável grau de coerência no sentido de dispersar a maior parte das medias dentro de uma faixa de 30% nos resultados, ou seja, dos estudantes de audiovisual que nunca estudaram música a diferença entre os maiores resultados e a maioria dos menores, ocorreu abaixo de 80% e acima ou igual a 50%, na maior parte dos estímulos.

Já os estudantes de audiovisual com estudo informal de música têm a dispersão media dos seus resultados na faixa entre 70% e 40% de similaridade dos resultados. Como podemos notar no espaço compreendido entre as linha vermelhas nos dois gráficos abaixo.

Figura 18 - Resultados dos índices de similaridade (eixo y) do grupo todo acima e dos sub grupos dos alunos de audiovisual (sem nenhum estudo de música no gráfico do meio, e com estudo informal de música no gráfico abaixo), obtidos individualmente por cada Participante (legendas coloridas).

Estes resultados apontam para a tendência de participantes sem estudo musical apresentarem maior compatibilidade com a percepção da acentuação subjetiva de acordo com o padrão RP e os participantes com estudo informal de música apresentam menor compatibilidade uma vez que suas medias encontram-se numa faixa entre o que consideramos um bom índice de percepção 0,70 (ou 70%) e outro em 0,4 (ou 40%) consideramos abaixo da media de percepção de 50%. Significando que menos da metade das exposições aos estímulos equitoniais as percepções apresentam resultados similares ao padrão RP.

3.3.3.2.2 Conclusão

Aqueles participantes sem formação musical tiveram os maiores índices (P3 a P6) 50% com índices de presença de acentuação subjetiva em 56% dos Estímulos e três deles com medias de similaridade entre 0,63 e 0,66 (as mais altas do Experimento).

3.3.4 Fase 3: Experimento com interface linear

Esta fase testou a acentuação subjetiva com uma interface que propõe a visualização gráfica dos estímulos de forma linear, conforme mostra a Figura 19, de

modo a eliminar a probabilidade de indução a erro detectada na fase anterior, com a possível inversão da leitura na ordem dos elementos dos grupos entre a parte superior da circunferência e a parte inferior da mesma.

Figura 19 - Tela demonstrativa da interface linear da Fase 2 do experimento, onde se pode ver a representação visual do estímulo correspondente àquele da Figura 16, nomeado E8 no nosso experimento.

3.3.4.1 Materiais e Métodos

Nossa infraestrutura experimental foi composta dos mesmos materiais da Fase 2: Mesmo fone de ouvido, mesmo computador e mesmo software.

As diferenças deram-se em alterações no procedimento em relação à Fase 2.

3.3.4.1.1 Procedimento

Quinze participantes, sendo seis da graduação do Departamento de Música da UNESP e nove não estudantes de música - da graduação do Departamento de Audiovisual da Universidade de São Paulo, interagiram com o experimento individualmente.

Dos não estudantes de música foram cinco homens e quatro mulheres, um canhoto e os outros destros, com idades que variaram de 21 a 24 anos; a média de

idade foi de 22 anos ($DP = 1,36$). Cinco deles possuem de 8 meses a 6 anos de prática de instrumento musical (violão, piano e percussão) equivalentes a um estudo informal de música e todos participaram do experimento individualmente.

Dos estudantes de música foram cinco homens e quatro mulheres, um canhoto e os outros destros, com idades que variaram de 18 a 37 anos; a média de idade foi de 23 anos ($DP = 7,41$). O desvio padrão do grupo todo foi de 4,55 indicando que a maioria dos participantes tem idades próximas da media de 22 anos.

Os primeiros 4 (P1 a P4) participantes não músicos inicialmente assistiram a um vídeo demonstrativo, para que se familiarizassem com o processo. Eles foram posicionados à frente de um computador para visualizar a representação gráfica das sequências. Cada sequência foi apresentada através de fones de ouvido em quatro tentativas: a primeira tocou a sequência apenas uma vez, na segunda, duas vezes, na terceira, três vezes e na quarta tentativa, quatro.

Ao término de cada tentativa, o participante foi instruído a marcar na representação gráfica da sequência aqueles elementos que considerou acentuados e a clicar no botão verde usando o mouse do canto esquerdo da tela do computador para passar para a próxima tentativa.

Os seguintes 5 participantes não músicos (P5 a P9) tiveram uma pequena alteração no protocolo: ao invés de assistirem um vídeo mostrando o experimento, eles passaram por uma simulação do experimento, com acentuação real, marcando os acentos reais no intuito de verificarmos se o conceito de acentuação foi bem compreendido pelos participantes. O restante do experimento com esses participantes correu de maneira igual aos outros.

3.3.4.1.2 Estímulos

Os estímulos são os mesmos da Fase 2, composto por nove sequências equitonais de Povel e Essens (Perception of Temporal Patterns, 1985), sendo 8 dos exemplos retirados do experimento, p. 420 (figura 15) e um da p.416 (figura 16) com *beeps* de 660Hz, onda senoidal e duração de 50ms, IOI de 216ms.

3.3.4.1.3 Medidas

Os participantes ouvem os estímulos e assinalam as acentuações nos elementos das sequências. O interesse aqui é medir o grau de similaridade das acentuações percebidas pelos participantes músicos e não músicos com o padrão das regras de preferência, e também verificar se há a formação de padrão próprio do grupo em relação ao da literatura, através do cálculo proposto na metodologia do item 3.2.

3.3.4.1.4 Discussão

Entre os participantes, havia estudantes de música e estudantes de audiovisual (não músicos) e o objetivo foi comparar se há alguma diferença significativa entre a percepção de estudantes de música treinados para perceber acentuações com mais acuidade e aqueles sem um estudo específico de música. *Conjeturamos que se, por um lado, os estudantes de música tendem a detectar mais facilmente a acentuação que os outros, paradoxalmente seu treinamento pode induzir a percepções padronizadas pela maior imersão no condicionamento da métrica musical, produzidas pela busca do referencial da pulsação da música e da estrutura da organização rítmica; diferentemente dos outros, cujo treinamento está direcionado para outros parâmetros do áudio mais relacionados à estética, à produção de emoções específicas ligadas aos objetivos e intenções do produto audiovisual.*

É importante lembrar que, entre a coleta de dados dos estudantes do Curso de Música e daqueles do Curso Audiovisual, procedemos uma pequena mudança no protocolo de aplicação do experimento. Inserimos dois estímulos a mais com acentuação física em alguns elementos, para nos assegurarmos que os participantes

tivessem a escuta efetiva das acentuações realmente contidas nos estímulos e que as percebiam corretamente conforme as diretrizes de experimentos psicofísicos sugeridas pelo Prof. Ranvaud.

Com relação às medições da similaridade dos estímulos, procedemos também a contabilização das médias dos índices de similaridade obtidos nas 4 tentativas de avaliação da percepção de cada participante por estímulo, para proceder uma comparação de seus respectivos índices médios perceptuais de acentuação em relação ao Padrão-RP (Figuras 20, 21 e 22).

Esta estratégia supostamente torna possível a visualização de estímulos mais refratários à percepção similar ao Padrão-RP e aqueles onde o efeito manifestado é maior.

Figura 20 - Gráfico comparativo das médias de similaridade (eixo y) por Participantes estudantes de música Fase 3 (legendas coloridas) em relação aos estímulos (eixo x).

Embora as médias obtidas contenham um grau razoável de dispersão (Figura 20), é possível notar claras tendências localizadas, como nas médias dos estudantes de Música onde o grau de similaridade para quatro dos participantes (P1, P2, P3 e P6) decresce em conjunto nos estímulos E3 e E4, atingindo grau acima de 70% em E5 (agora junto com P5) e caindo conjuntamente no estímulo E6, voltando à dispersão em E7 e dividindo-se em dois grupos de 3 participantes em E8 (um grupo com índice acima de 70% e outro abaixo de 50%).

Figura 21 - Gráfico comparativo das médias de similaridade (eixo y) por Participantes estudantes de Audiovisual da Fase 2 (legendas coloridas) em relação aos estímulos (eixo x).

No gráfico das médias de similaridade dos estudantes do Audiovisual - Fase 2 (Figura 21), podemos notar a predominância da dispersão dos resultados, acima da faixa dos 50%, revelando bons indícios da similaridade da percepção dos participantes com o padrão-RP. Interessante ainda neste gráfico é que a curva do participante P1 apresenta baixa similaridade com o padrão e se destaca em relação às demais, sendo a maior diferença entre o perfil deste participante e dos outros é que sua alfabetização não foi em português. Este indício mereceria a investigação sobre uma possível correlação entre a percepção da acentuação subjetiva e língua materna. Outra alternativa de análise dos seus resultados, seria a sensação de dúvida em relação às acentuações reportadas por P1, o que nos leva a questionar se fomos claros na apresentação da tarefa e na explicitação do conceito de acentuação para P1.

Figura 22 - Gráfico comparativo das médias de similaridade (eixo y) por Participantes do Curso de Audiovisual Fase 3 (legendas coloridas) em relação aos estímulos (eixo x).

O gráfico das médias dos estudantes do Audiovisual da Fase 3 (Figura 22), apresenta muito mais ocorrências de médias abaixo dos 50% que o da Fase 2 (Figura 21), o que pode ser decorrência da introdução dos estímulos com acentos reais na etapa de apresentação do protocolo. Esse aspecto indica a possível emergência de um efeito conhecido como *viés induzido* (cognitive bias) das respostas. Este viés consiste na possibilidade indução de algumas percepções de acentuação, introduzida pelos estímulos com acentuação, que dificilmente poderiam ser relacionadas à uma memória consciente e que aparentemente invertem a tendência esperada de resposta observada no grupo que não foi exposto ao estímulo com acentos reais.

3.3.4.1.5 Conclusão

Dos nove participantes não músicos da fase 3, P1 foi o que teve a percepção menos similar ao Padrão-RP (Regras de Preferência). É possível que não estivesse em bom momento para se concentrar na tarefa proposta, suas respostas apresentam baixa similaridade em geral, somente chegando em 50% no estímulo 9, além de não apresentar um padrão próprio consistente com alguma regra alternativa.

Dentro do grupo da Fase 3, o estímulo 8 apresenta os maiores índices de similaridade com o padrão das regras de preferência para três dos nove participantes, em outras palavras, é percebido como aquele em que o fenômeno da acentuação subjetiva se mostra bem forte para mais participantes ao mesmo tempo. Três dos nove participantes demonstraram índices médios acima de 70%, sendo que um deles tem experiência musical autodidata. Ao mesmo tempo os outros cinco participantes tiveram índices de similaridade abaixo dos 50%.

Os estímulos E5 e E6, na média, apresentam os maiores índices de similaridade com o padrão das regras de preferência, sendo percebidos como aqueles em que o fenômeno da acentuação subjetiva se mostra mais forte para a maioria dos participantes não músicos. Oito dos nove participantes não músicos demonstraram índices médios acima de 50%, no estímulo E6, sendo cinco deles na faixa entre 60% e 80%, e um participante com 91%. No estímulo E5, sete dos nove participantes tiveram índices médios acima dos 50%. Como se pode conferir no gráfico da Figura 19 e na Tabela 1 respectiva que segue.

	E1	E2	E3	E4	E5	E6	E7	E8	E9
P1	0,41	0,31	0,34	0,44	0,31	0,34	0,28	0,32	0,50
P2	0,78	0,47	0,69	0,53	0,53	0,69	0,00	0,79	0,53
P3	0,56	0,56	0,34	0,47	0,53	0,66	0,22	0,29	0,41
P4	0,47	0,56	0,56	0,81	0,63	0,72	0,00	0,71	0,75
P5	0,41	0,41	0,28	0,31	0,50	0,91	0,31	0,46	0,34
P6	0,34	0,38	0,50	0,44	0,28	0,53	0,44	0,46	0,47
P7	0,66	0,59	0,56	0,28	0,84	0,75	0,69	0,43	0,69
P8	0,38	0,78	0,84	0,78	0,78	0,56	0,84	0,71	0,78
P9	0,50	0,63	0,72	0,41	0,72	0,63	0,59	0,64	0,66

Tabela 1 – Médias dos índices de percepção dos participantes não músicos da Fase 3 de P1 a P9, obtidos em cada estímulos de E1 a E9, e referente ao gráfico da Figura 19. Em vermelho os índices acima de 50%.

Curiosamente, o Estímulo 7, foi aquele onde dois participantes não responderam e os outros dois responderam praticamente o oposto do esperado para o Padrão-RP. Os participantes estudantes de música demonstraram grande dispersão na percepção desse estímulo.

Comparação entre os resultados das Fases 2 e 3 para padrões próprios de cada expressos pelos grupos analisados.

Os resultados dos padrões de percepção dos participantes para cada estímulo E, de 1 a 9 (**E1...E9**), sendo que: **F3** corresponde à fase 3, com a interface de representação linear; **F2**, à fase 2, com interface circular; **AV**, aos estudantes de Artes Visuais, que interagiram com todas as interfaces; **Músicos**, aos estudantes de Música, que realizaram o experimento apenas com a interface linear. Estes resultados devem ajudar a revelar possíveis regras alternativas àquelas do Padrão-RP para a percepção de acentos subjetivos e podem ser visualizados detalhadamente no ANEXO I (Figuras 34, 35 e 36 e Tabela 4)

As designações utilizadas nos gráficos abaixo (Figuras 23 e 24) são as que seguem:

- **E1-F3-AV** (Estímulo 1, experimento na Fase 3 e interface de representação linear, com barras azuis).
- **E1-F2-AV** (Estímulo 1, experimento na Fase 2 e interface de representação circular, com barras vermelhas).
- **E1-F3-Músicos** (Estímulo 1, experimento Fase 3 e a interface de representação linear realizado com estudantes de música, com barras verdes).
- No eixo X, temos a representação dos grupos de elementos equitoniais.
- **U** significa grupo formado por um elemento.
- **D** significa grupo formado por dois elementos, sendo D1 o primeiro e D2 o segundo.
- **T** significa grupo formado por três elementos, sendo T1 o primeiro e T2 o segundo.
- **Q** significa grupo formado por quatro elementos.
- **C** significa grupo formado por cinco elementos.
- No eixo Y, temos a representação do padrão de acentuação subjetiva (**q**) de cada população estudada (com alunos de Música em verde, e alunos do Audiovisual em azul e vermelho).

Figura 23 - As barras coloridas significam os índices de percepção de acentuação (eixo y) calculados para cada grupo (legendas coloridas) e referentes a cada evento (eixo x) dos agrupamentos das sequências. O asterisco ao lado do elemento da sequência significa a coincidência do elemento com uma pulsação musical.

Figura 24 - As barras coloridas significam os índices de percepção de acentuação dos estímulos E5, E6 e E8 (eixo y) calculados para cada grupo (legendas coloridas) e referentes a cada evento (eixo x) dos agrupamentos das sequências. O asterisco ao lado do elemento da sequência significa a coincidência do elemento com uma pulsação musical.

Os resultados gráficos podem ser melhor conferidos na Tabela 4 no ANEXO I, que mostra em vermelho os resultados iguais ou maiores que 50% obtidos pela percepção da acentuação no vetor de cada posição. As legendas ladeadas por (*) referem-se aos pontos de ocorrência das pulsações induzidas no experimento de Povel e Essens e concomitantes com o elemento do grupo na sequência.

Apesar das taxas de similaridade entre o padrão das regras de preferência e os resultados dos participantes participantes não se mostrarem altos o suficiente em nenhum grupo estudado até o momento, o cálculo das acentuações mais percebidas em cada posição de vetor para cada grupo indica uma tendência maior para acentuação apenas no último elemento de grupos de três ou mais elementos para os participantes da Fase 2 (50% para grupos de três e quatro elementos), e para os participantes da Fase 3 (66% para grupos de quatro elementos).

O grupo com maior tendência de percepção apenas no último elemento, seja de dois, três ou quatro, é aquela da Fase 2, com resultado de 50% para grupos de três elementos e maior que 50% nos grupos de 2 e de 4 elementos (Tabela 3).

Ocorrências 2 Elementos	Em 10 grupos		Ocorrências 3 Elementos	Em 6 grupos		Ocorrências 4 Elementos	Em 3 grupos	
Respostas-D-F3- P1 a P4	1	10%	Respostas-T-F3	3	50%	Respostas-Q-F3	1	33%
Respostas-D-F2	8	80%	Respostas-T-F2	3	50%	Respostas-Q-F2	2	66%
Respostas-D-Mus	6	60%	Respostas-T-Mus	1	17%	Resposta-Q-Mus	0	0%

Tabela 2 - Percentual acentuações no último elemento do grupo, sendo que a Fase 3 neste ponto, tinha 4 participantes (P1 a P4) .

Ocorrências 2 Elementos	Em 10 grupos		Ocorrências 3 Elementos	Em 6 grupos		Ocorrências 4 Elementos	Em 3 grupos	
Respostas-D-F3 - Todos	2	20%	Respostas-T-F3	4	67%	Respostas-Q-F3	1	33%
Respostas-D-F2	8	80%	Respostas-T-F2	3	50%	Respostas-Q-F2	2	66%
Respostas-D-Mus	6	60%	Respostas-T-Mus	1	17%	Resposta-Q-Mus	0	0%

Tabela 3 - Percentual acentuações no último elemento do grupo, contabilizando a inclusão de cinco participantes não músicos, na Fase 3.

O grupo dos estudantes de música apresentou um padrão de acentuações mais compatível com as posições vetoriais que coincidem com pulsações, como se pode observar nos gráficos dos estímulos E3, E4, E5 e E6 (Figuras 21 e 22), onde os elementos com asterisco “*” referem-se àqueles posicionados de acordo com os pulsos esperados para cada Estímulo segundo o experimento de Povel e Essens

(1985). Seus índices de acentuação acima de 50% nas posições referidas resultam em grande diferenciação dos padrões do grupo dos estudantes de audiovisual, que apresentam tendência a modificar parcialmente a regra de preferência, uma vez que acentuam apenas o último elemento do grupo de dois ou mais elementos (ao invés de acentuarem o primeiro e o último).

Conclusão

Os experimentos de Povel e Essens (1985) testavam a indução da pulsação, ou do relógio interno através da eficácia da reprodução de estímulos e adotando a Regra da Acentuação Subjetiva como válida para os mesmos.

E iniciamos nosso percurso experimental procurando testar a hipótese com agrupamentos isolados, primeiramente imitando o motivo inicial da *Sinfonia n. 5* de Beethoven com uma estrutura de quatro elementos, logo em seguida testando o efeito com três elementos também e obtivemos resultados inconclusivos. Este fato levou-nos à hipótese de que os elementos isolados não tinham duração suficiente para a emergência do fenômeno da acentuação subjetiva. Buscando elaborar estímulos estruturados com maior duração, resolvemos testar a emergência do fenômeno com os próprios estímulos de Povel e Essens (1985), afinal estes nunca haviam sido testados avaliando a força do fenômeno efetivamente, da maneira como o fizemos. Testando inclusive, grupos de universitários de música e universitários sem estudo formal de música.

Ao adotarmos seus estímulos como referência para nossos experimentos, com a intenção de testar a força dos mesmos em grupo de estudantes universitários paulistanos, supúnhamos que o efeito emergisse para todos os participantes com grande intensidade.

Desenvolvemos uma metodologia matemática para a medição da similaridade entre o Padrão-RP da literatura e as respostas dos nossos participantes. Também foi desenvolvida uma metodologia para o estabelecimento de um padrão alternativo de cada grupo avaliado, buscando possíveis regras alternativas de acentuação, caso fosse necessário.

Uma das fragilidades encontradas no cálculo da similaridade, reside na situação onde o participante acentua todos os elementos de um grupo. Isso equivale não ter acento nenhum naquele grupo, mas por outro lado ele pode estar percebendo aquele grupo como mais acentuado que os vizinhos, com uma visão mais global do padrão. Nesse caso, como o número de acentos esperados para o grupo de três elementos é ter acentuado o primeiro e último, o índice de similaridade se torna maior, por coincidir dois acentos em três possibilidades. A fragilidade do cálculo também se mostra no outro extremo da situação, que é o caso do participante que não

ouve acentuação em nenhum elemento. Para esse caso o resultado da similaridade não será zero porque os zeros de sua resposta somarão na similaridade com aqueles zeros do Padrão de referência. Para estas situações ainda não chegamos a nenhuma solução.

Os resultados de nossos experimentos já numa segunda fase, contando apenas com universitários sem estudo formal de música, apresentaram dispersão média de aproximadamente 30%, considerando o espalhamento dos resultados entre o maior índice médio e o menor (Figura 25), exceção apenas para os estímulos E1 e E4.

Figura 25 - Resultados dos índices de similaridade (eixo y) dos sub grupos dos alunos de audiovisual (sem nenhum estudo de música no gráfico de cima, e com estudo informal de música no gráfico abaixo), obtidos individualmente por cada Participante (legendas coloridas).

Porém, olhando mais detalhadamente e dividindo o grupo dos alunos de áudio visual em subgrupos considerando a presença ou ausência de estudo informal de música, a faixa media de 30% na dispersão dos resultados apresenta uma queda

(Figura 24) indicando possível correlação entre presença de estudo musical e percepção similar ao padrão RP. Ou seja, *quando não há treinamento musical a percepção de acentos é mais próxima do padrão RP.*

Este indício é corroborado pelas médias dos estudantes de música da Fase 3, cuja percepção somente alcança alta similaridade com o padrão RP nos estímulos em que as acentuações relacionadas à pulsação coincidem com mais acentos do padrão RP, permanecendo baixa nos outros estímulos onde a pulsação não coincide com o padrão RP, reafirmando assim a tendência de que a formação musical pode alterar a percepção da acentuação subjetiva, expressando-se em baixos índices médios de similaridade com o padrão RP.

Ainda com esta dispersão dos resultados, os gráficos de similaridade da Fase 2, no geral (tanto das médias de todos os participantes, quanto os outros, que se individualizam por sub grupos), ajudaram-nos a visualizar que esta dispersão se deu predominantemente acima da faixa dos 50%, sugerindo que o Padrão-RP é significativo para os grupos pesquisados, apesar de não haverem indicado tendências fortes para similaridade entre todos os participantes dos experimentos das Fases 2 e 3.

Já os gráficos de padrão alternativo revelam a tendência a um padrão derivado do Padrão-RP (visto na Tabela 3), que incorpora uma de suas regras, acentuando preferencialmente os últimos eventos de grupos entre 2 e 4 elementos para os estudantes de audiovisual na Fase 2 (de 80 a 50% das ocorrências), para os estudantes de audiovisual na Fase 3 (de 20 a 67% das ocorrências) com acentuação no último evento de grupos de 3 e 4 elementos e para os estudantes de música na Fase 3 (60% das ocorrências) no último elemento de grupos de 2 elementos apenas. Há que se lembrar que o número de participantes do audiovisual na Fase 3 é ainda muito pequeno (9 participantes) e pouco significativo, se o considerarmos isoladamente. Porém, se forem adicionados aos estudantes de audiovisual da Fase 2, a tendência ganha em significação. (para detalhamento ver Figuras 32 a 34 e Tabela 4 no ANEXO I)

Entretanto, os estudantes de música demonstram uma tendência em perceber acentuações diversas do Padrão-RP nos momentos concomitantes com a regularidade das pulsações. e estes dados indicam a tendência de o treinamento musical influenciar o efeito do Padrão-RP. Neste ponto, podemos vislumbrar que haveria uma influência da cultura musical sobre a percepção “inata” de recorrências temporais, se algum dia for comprovado que o Padrão-RP pode ser tomado efeito Gestáltico⁴⁶ e presente nos seres humanos desde o seu nascimento, ou desde a fase pré-verbal.

Estas considerações relativas à comparação entre os resultados da Fase 2 e da Fase 3 levam, também, a um questionamento a respeito da eficácia de cada interface de representação de cada experimento. Se conseguirmos realizá-los com mais participantes, além dos ganhos estatísticos de informação sobre a percepção de acentuações, poderíamos avaliar também o grau de influência da representação em relação aos resultados obtidos.

Neste ponto, uma questão interessante se coloca: até onde o processamento de um estímulo físico sonoro cíclico é resultado da ontogenia? Em outras palavras, o processamento de um sinal acústico em ciclo binário, como tal, é inato ou aprendido? Se for um pouco de cada um deles, quanto ou que parte é resultado de sua interação com o meio?

Considerando os dois estilos de processamento, *top-down* e *bottom-up*, da maneira como foi exposto no corpo do presente trabalho, poderíamos pensar na possibilidade da acentuação subjetiva ser um processamento inicial do tipo *bottom-up*, enquanto o processamento *top-down* como sendo advindo de um treinamento do tipo que desenvolve a percepção da pulsação por músicos através do estudo formal de música.

⁴⁶ Entendemos por Gestalt a um dinamismo auto regulador do sistema nervoso central, que ao buscar sua própria estabilidade tende a organizar a percepção de acordo com princípios estruturantes de semelhança, proximidade, unidade, segregação e continuidade. Esse pensamento tem como principais representantes, Max Wertheimer (1880-1943), Wolfgang Köhler (1887-1967) e Kurt Koffka (1886-1940).

Sendo assim, repetimos o questionamento que fizemos na Introdução deste trabalho: seria possível a modificação da percepção da acentuação de acordo com a intencionalidade aprendida no treinamento musical?

Na fase 3, a última fase dos experimentos, identificamos duas forças que possivelmente atuam em concorrência na determinação dos acentos subjetivos percebidos pelos participantes. Uma delas seria as regras de preferência; a outra, a própria força da métrica do compasso musical.

Questões advindas do participante P-6 e possíveis conclusões: O fato do participante ter declarado dificuldade para perceber acentuações ao final do experimento constitui motivo para um estudo de caso buscando configurar melhor o que vem a ser essa percepção específica.

Os grupos de dois elementos não apresentam muitas ocorrências na fase 3, isso pode ser efeito do viés cognitivo induzido?

O tempo de resposta do participante, pode ser um indício de nível de processamento cognitivo? O tempo maior para as respostas dos participantes poderia indicar uma reinterpretação da percepção dos estímulos, ou seja a aplicação de um processo *top-down*, resignificando a primeira percepção dada pelo processo *bottom-up*?

Acreditamos que a inclusão da coleta deste tipo de informação no script do experimento pode fornecer uma possível correlação entre o tempo de resposta e a qualidade das respostas obtidas em relação à similaridade com o Padrão-RP, ou com o padrão próprio do grupo.

Nas médias de similaridade para os estudantes de audiovisual de P5 a P9, somente P5 e P6 ficaram abaixo da média dos 50%, na maioria dos estímulos apresentados. Os outros P7, P8 e P9 apresentaram índices de similaridade acima dos 50% na maior parte dos estímulos, sendo que os maiores índices foram obtidos nos estímulos E3, E5 e E6.

As diferenças no estabelecimento do padrão próprio entre os participantes da fase dois e os da fase três podem ser consequência: da mudança da interface circular para a linear, ou da aplicação dos estímulos com acentuação efetiva.

A introdução dos estímulos com acentos reais para verificação da escuta efetiva de acentuações, mostrou a influência de grupos de acentuação dos estímulos com acentos reais nas respostas dos participantes, indicando a possibilidade de um "viés induzido" nas respostas respectivas, também conhecido como "cognitive bias". Este fato nos leva a uma discussão interessante, porque dificilmente a memória consciente dos participantes expostos aos estímulos com acentos reais, atuaria no sentido de relacionar grupos equivalentes em número de elementos acentuados aos outros sem acentos reais.

Esta possibilidade pode ser explicada através da heurística aplicada às diferenças de julgamento e tomada de decisões humanas. "Heurísticas envolvem atalhos mentais que fornecem estimativas rápidas sobre a possibilidade de ocorrências incertas" (Baumeister & Bushman, 2010, p.141). Seu processamento é simples para o cérebro computar, mas às vezes as heurísticas introduzem "erros severos e sistemáticos" (Tversky & Kahneman, 1974, p.1115).

A hipótese do viés induzido pode ser observada na comparação entre as tendências observadas antes da introdução da apresentação de estímulos com acentos reais, aplicada aos participantes de P1 a P4 e depois da mesma, com os participantes de P5 a P9, na comparação entre os gráficos que seguem, Figuras 26,27 e 28:

Figura 26 - Comparativo entre os padrões de acentuação dos participantes de P1 a P4 e de P5 a P9, mostrando a inversão de tendências de acentuação para o estímulo E4, tanto no primeiro grupo D1* D2, quanto no segundo grupo. Neles, a tendência para acentuação em D2, se inverte para acentuação maior em D1.

Figura 27 - Comparativo entre os padrões de acentuação dos participantes de P1 a P4 e de P5 a P9, mostrando a inversão de tendências de acentuação para o estímulo E8, grupo D1 D2 ao final da sequência. Nele, a tendência para acentuação em D2, se inverte para acentuação maior em D1.

Figura 28 - Comparativo entre os padrões de acentuação dos participantes de P1 a P4 e de P5 a P9, mostrando a inversão de tendências de acentuação para o estímulo E9, no grupo D1 D2. Nele, a tendência para acentuação em D2, se inverte para acentuação maior em D1.

Como conclusão final temos que até onde pesquisamos, a acentuação subjetiva é um efeito bem sutil, que pode ser influenciado por aspectos exógenos, como na hipótese do viés induzido, e endógenos como a influencia do treinamento musical nos resultados obtidos sobre a percepção de acentos, haja visto que nos grupos pesquisados aparentemente este efeito não emerge exatamente como relatado para os grupos europeus.

Estes resultados, trazem novas questões para serem trabalhadas como continuação deste experimento (aumentar o numero de participantes por exemplo) e em próximos experimentos que abrirão perspectivas de aprimoramento desta pesquisa.

Bibliografia Completa

- AROM, S. **African Polyphony and Polyrhythm: Musical Structure and Methodology**. NY: Cambridge University Press, 2004.
- BONDESAN DOS SANTOS, P. P. **Ambiguidade rítmica: estudo do ritmo musical sob a perspectiva de modelos atuais de percepção e cognição**. *Dissertação (Mestrado em Processos de Criação Musical)*. São Paulo: Escola de Comunicações e Artes, Universidade de São Paulo, 2012.
- COGNITIVE ATLAS. Projeto coordenado por Russell Poldrack, Professor da Stanford University em colaboração com o UCLA Center for Computational Biology e o UCLA Consortium for Neuropsychiatric Phonemics. [s.n.]: [s.n.]. Disponível em: <http://www.cognitiveatlas.org/concept/attentional_focusing>. Acesso em: 20 set. 2016.
- FITCH, W. T. The biology and evolution of rhythm: unravelling a paradox. In: REBUSCHAT, P., et al. **Language and Music as Cognitive Systems**. New York : OUP Oxford, 2011. p. 356. ISBN 9780199553426.
- FUJIOKA, T.; FIDALI, B. C.; ROSS, B. Neural correlates of intentional switching from ternary to binary meter in a musical hemiola pattern. **Frontiers in Psychology**, 2014. doi:10.3389/fpsyg.2014.01257.
- GIL, A. C. **Como elaborar projetos de pesquisa**. 4. ed. ed. São Paulo: Atlas, 2002.
- HINKELMAN, K.; KEMPTHORNE, O. Introduction to experimental design. New York: John Wiley and Sons, Inc. 1994.
- HURON, D. **Sweet anticipation: music and the psychology of expectation**. Cambridge: MIT Press, 2006.
- IVERSEN, J. R.; REPP, B. H.; PATEL, A. **Top-Down Control of Rhythm Perception Modulates Early Auditory Responses**. Brain Mechanisms of Metrical Interpretation. [S.l.]: Annals of the New York Academy of Sciences. 2009. p. 58-73. doi: 10.1111/j.1749-6632.2009.04579.x.
- JONES, M. R. Musical Time. In: HALLAM, S.; CROSS, I.; THAUT, M. **Oxford Handbook of Music Psychology (Oxford Library of Psychology)**. [S.l.]: OXFORD UNIVERSITY PRESS, 2009. ISBN-10: 0199298459.
- JONES, M. R.; BOLTZ, M. Dynamic attending and responses to time. **Psychological Review**, v. 96, n. 3, p. 459-491, 1989.
- JONES, M. R.; G. KIDD; R. WETZEL. Evidence for rhythmic attention. **Journal of Experimental Psychology. Human Perception and Performance**. 7(5), 1981, p.1059–1073.

- KNILL, D. C.; KERSTEN, D.; YUILLE, A. A Bayesian formulation of visual perception. In: KNILL, D. C.; RICHARDS, **Perception as Bayesian Inference**. [S.l.]: Cambridge University Press, 1996.
- KOELSCH, S. Neural Correlates of Music Perception. In: **Language, music, and the brain: A mysterious relationship**, Cambridge: MIT Press, 2013, p.141-172.
- LARGE, W. Resonating to musical rhythm: theory and experiment. In: _____ **Psychology of time**. [S.l.]: [s.n.], 2008. p. 189-232.
- LERDAHL, F.; JACKENDOFF, R. **A generative theory of tonal music**. Cambridge: MIT Press, 1983.
- LONDON, J. **Hearing in Time: Psychological Aspects of Musical Meter**. New York: Oxford University Press, 2012.
- LONDON, J.; CROSS, I.; HIMBERG, T. The Effect Structural and Performance Factor in the Perception of Anacruses. **Music Perception: An Interdisciplinary Journal**, v. 27, n. 2, p. 103-120, December 2009. <http://www.jstor.org/stable/10.1525/mp.2009.27.2.103>.
- LUCK, S. J. **An introduction to the event-related potential technique**. [S.l.]: MIT Press, 2005. 374 p.
- MARY, L.; YEGNANARAYANA, B. Extraction and representation of prosodic features for language and speaker recognition. **Speech Communication**. **50**(10), 2008, p.782–796.
- NOZARADAN, S. Exploring how musical rhythm entrains brain activity with electroencephalogram frequency-tagging. **Philosophical Transactions of the Royal Society. Series B, Biological Sciences**, London, 369, n. 1658, 2014. Acesso em: 20 nov. 2014. doi:10.1098/rstb.2013.0393.
- NOZARADAN, S. et al. Tagging the Neuronal Entrainment to Beat and Meter. **The Journal of Neuroscience**, v. 31, n. 28, p. 10234-10240, 13 July 2011. DOI:10.1523/JNEUROSCI.0411-11.2011.
- NOZARADAN, S.; PERETZ, I. **Exploring the Neural Entrainment to Musical Rhythms and Meter: a Steady-State Evoked Potential Approach**. PhD Dissertation. ed. Montreal: [s.n.], 2013. Tese de Doutorado.
- PATEL, A. D. **Language, music, and the brain**. [S.l.]: OUP, 2008.
- PATEL, A. D.; IVERSEN, J. R. The evolutionary neuroscience of musical beat perception: the action simulation for auditory prediction (ASAP) hypothesis.. **Front. Psychol.**, v. 8, n. 57, 2014. (doi:10.3389/fnsys. 2014.00057).
- PINAUD, R. et al. Profiling of experience-regulated proteins in the songbird auditory forebrain using quantitative proteomics. **European Journal of Neuroscience**. **27**(6), 2008, p.1409–1422.

- POTTER, D. D. et al. Perceiving rhythm where none exists: event-related potential (ERP) correlates of subjective accenting. **Cortex; a Journal Devoted to the Study of the Nervous System and Behavior**, v. 45 , n. 1, p. 103–9, 2009. doi:10.1016/j.cortex.2008.01.004.
- POVEL, ; ESSENS, P. Perception of Temporal Patterns. **Music Perception**, v. 2, n. 4, p. 411-440, 1985.
- POVEL, D.J. Internal representation of simple temporal patterns. **Journal of experimental psychology. Human perception and performance.**, v. 7, n. 1, p. 3-18, Fev 1981.
- POVEL, D.J.; OKKERMAN, M. Accents in equitone sequences. **Perception & Psychophysics**, 30, n. 6, 1981. 565-572. Copyright 1982.
- REPP, B. H. Do metrical accents create illusory phenomenal accents? **Attention, Perception & Psychophysics**. 72(5), 2010, p.1390–1403.
- SCHERER, K. R. Emotions are emergent processes: they require a dynamic computational architecture. **Philosophical Transactions of the Royal Society B: Biological Sciences**. 364(1535), 2009, p.3459–3474.
- SCHULLER, Gunther. **The Compleat Conductor**. Oxford: Oxford University Press, 1997.
- SEIFERT, U., P. F. M. J. VERSCHURE, M. A. ARBIB et al. Semantics of internal and external worlds. *In: Language, music, and the brain: A mysterious relationship*, Cambridge: MIT Press, 2013, p.203-232.
- TEMPERLEY, D. **The Cognition of Basic Musical Structures**. Cambridge: [s.n.], 2001.
- TEMPERLEY, D.; BARTLETTE, C. Parallelism as a Factor in Metrical Analysis. **Music Perception**, v. 20, n. 2, p. 117-149, Winter 2002.
- WOLDORFF, M. (21 de 08 de 2015). **MIT Encyclopedia of Cognitive Science**. (MIT) Acesso em 27 de 09 de 2016, disponível em Auditory Attention: <http://ato.ms/MITECS/Entry/woldorff.html>

ANEXO I – Gráficos dos resultados obtidos

Figura 29 - Resultados do Participante R9 – Fase 1. Os retângulos representam cada condição de Intervalo Inter Onset - IOIs de 200ms a 800ms, com grupos de 4 elementos, onde cada numero inteiro de 1 a 4 do eixo y representa respectivamente cada elemento do grupo, e no eixo x, temos o índice de acentuação relativa marcado pelo participante. (catch trial) nos retângulos da terceira linha da tabela acima indicam grande foco atencional, próximo de 100%.

Figura 30 - Resultados do Participante R11 – Fase 1. Os retângulos representam cada condição de Intervalo Inter Onset - IOIs de 200ms a 800ms, com grupos de 4 elementos, onde cada numero inteiro de 1 a 4 do eixo y representa respectivamente cada elemento do grupo, e no eixo x, temos o índice de acentuação relativa marcado pelo participante. (catch trial) nos retângulos da terceira linha da tabela acima indicam grande foco atencional, próximo de 100%.

Figura 31 - Resultados dos índices de similaridade (eixo y) dos estímulos de 1 a 4, obtidos individualmente por cada Participante (legendas coloridas) relativos a cada tentativa T1 a T4 (eixo x)

Figura 32 - Resultados dos índices de similaridade (eixo y) dos estímulos de 5 a 9, obtidos individualmente por cada Participante (legendas coloridas) relativos a cada tentativa T1 a T4 (eixo x).

Figura 33 - Gráfico comparativo das médias de similaridade (eixo y) por Participantes da Fase 3 (legendas coloridas) em relação aos estímulos (eixo x). Acima os nove Participantes estudantes do Curso de Audiovisual, e abaixo os seis Participantes do Curso de Música a UNESP.

	T1*	T2	T3		U*			D1	D2*			D1	D2*	
E1-F3	0,22	0,22	0,44		0,50			0,36	0,44			0,33	0,31	
E1-F2	0,28	0,13	0,31		0,53			0,41	0,44			0,44	0,47	
E1-Mus.	0,17	0,25	0,33		0,38			0,21	0,63			0,21	0,63	
	U*		Q1	Q2	Q3*	Q4			U*			D1	D2*	
E2-F3	0,56		0,19	0,22	0,33	0,47			0,50			0,39	0,39	
E2-F2	0,50		0,13	0,19	0,13	0,50			0,50			0,28	0,72	
E2-Mus.	0,83		0,04	0,04	0,29	0,21			0,58			0,00	0,58	
	D1*	D2		T1	T2*	T3			U*			D1	D2*	
E3-F3	0,33	0,53		0,31	0,28	0,56			0,53			0,39	0,39	
E3-F2	0,16	0,50		0,25	0,31	0,50			0,38			0,31	0,75	
E3-Mus.	0,63	0,25		0,04	0,50	0,25			0,79			0,13	0,71	
	D1*	D2		T1	T2*	T3				D1*	D2		U*	
E4-F3	0,56	0,28		0,28	0,22	0,69				0,39	0,36		0,53	
E4-F2	0,19	0,59		0,31	0,44	0,72				0,38	0,59		0,53	
E4-Mus.	0,54	0,29		0,00	0,42	0,38				0,63	0,38		0,71	
	C1*	C2	C3	C4	C5*		U			U			U*	
E5-F3	0,44	0,33	0,22	0,22	0,44		0,47			0,44			0,53	
E5-F2	0,25	0,31	0,31	0,22	0,56		0,34			0,53			0,75	
E5-Mus.	0,67	0,04	0,04	0,04	0,50		0,54			0,54			0,63	
	U*			T1	T2*	T3		U			T1	T2	T3*	
E6-F3	0,75			0,47	0,03	0,44		0,64			0,33	0,17	0,69	
E6-F2	0,41			0,28	0,31	0,66		0,63			0,16	0,28	0,66	
E6-Mus.	0,75			0,00	0,71	0,04		0,46			0,08	0,21	0,58	
	Q1*	Q2	Q3	Q4		D1	D2			U			U*	
E7-F3	0,32	0,18	0,39	0,32		0,46	0,36			0,46			0,43	
E7-F2	0,19	0,19	0,19	0,63		0,28	0,75			0,38			0,72	
E7-Mus.	0,54	0,08	0,33	0,21		0,13	0,79			0,46			0,63	
	T1*	T2	T3				U			U			D1	D2
E8-F3	0,31	0,28	0,72				0,44			0,53			0,39	0,42
E8-F2	0,31	0,28	0,84				0,31			0,25			0,22	0,84
E8-Mus.	0,46	0,04	0,58				0,71			0,54			0,50	0,54
	Q1*	Q2	Q3	Q4		D1	D2			U			U*	
E9-F3	0,56	0,06	0,44	0,39			0,47	0,33		0,61			0,61	
E9-F2	0,13	0,28	0,28	0,81			0,19	0,63		0,31			0,75	
E9-Mus.	0,42	0,00	0,42	0,08			0,54	0,28		0,46			0,75	

Tabela 4 - Resultados para os padrões de acentuações mais percebidas por cada grupo. Em vermelho os resultados iguais ou maiores a 0,50.

As designações utilizadas nos gráficos abaixo (Figuras 32 a 34) são as que seguem:

- **E1-F3-AV** (Estímulo 1, experimento na Fase 3 e interface de representação linear, com barras azuis).
- **E1-F2-AV** (Estímulo 1, experimento na Fase 2 e interface de representação circular, com barras vermelhas).
- **E1-F3-Músicos** (Estímulo 1, experimento Fase 3 e a interface de representação linear realizado com estudantes de música, com barras verdes).
- No eixo X, temos a representação dos grupos de elementos equitoniais.
- **U** significa grupo formado por um elemento.
- **D** significa grupo formado por dois elementos, sendo D1 o primeiro e D2 o segundo.
- **T** significa grupo formado por três elementos, sendo T1 o primeiro e T2 o segundo.
- **Q** significa grupo formado por quatro elementos.
- **C** significa grupo formado por cinco elementos.
- No eixo Y, temos a representação do padrão de acentuação subjetiva (**q**) de cada grupo estudado (com alunos de Música em verde, e alunos do Audiovisual em azul e vermelho).

Figura 34 - As barras coloridas significam os índices de percepção de acentuação dos estímulos E1, E2 (eixo y) calculados para cada grupo (legendas coloridas) e referentes a cada evento (eixo x) dos

agrupamentos das seqüências. O asterisco ao lado do elemento da seqüência significa a coincidência do elemento com uma pulsação.

Figura 35 - As barras coloridas significam os índices de percepção de acentuação dos estímulos E3 a E6 (eixo y) calculados para cada grupo (legendas coloridas) e referentes a cada evento (eixo x) dos

agrupamentos das seqüências. O asterisco ao lado do elemento da seqüência significa a coincidência do elemento com uma pulsação.

Figura 36 - As barras coloridas significam os índices de percepção de acentuação dos estímulos E7 a E9 (eixo y) calculados para cada grupo (legendas coloridas) e referentes a cada evento (eixo x) dos agrupamentos das seqüências. O asterisco ao lado do elemento da seqüência significa a coincidência do elemento com uma pulsação.

Glossário

Beep - som senoidal com duração de 60 ms, incluindo *fade in/out* de 10 ms no início e fim.