

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Escola De Música
Programa de Pós-Graduação em Música - Sonologia

Thiago de Almeida Magalhães Campolina

**Caracterização de Ataques na Clarineta via Modelagem e
Decomposição Espectral**

Belo Horizonte
2019

Thiago de Almeida Magalhães Campolina

Caracterização de Ataques na Clarineta via Modelagem e Decomposição Espectral

Versão Final

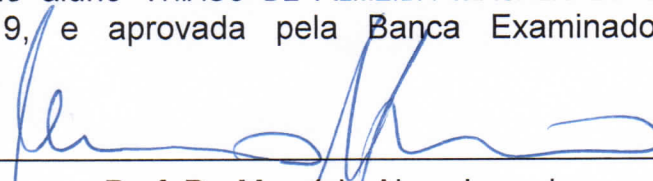
Tese apresentada ao Programa de Pós -
Graduação em Música da Escola de Música da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Doutor em Música.

Orientador: Maurício Alves Loureiro
Coorientador: Hani Camille Yehia

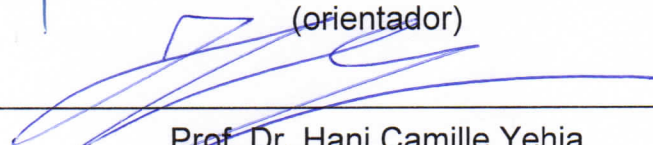
Belo Horizonte
2019

C198c	Campolina, Thiago de Almeida Magalhães. Caracterização de ataques na clarineta via modelagem e decomposição espectral [manuscrito] / Thiago de Almeida Magalhães Campolina. - 2019. 106 f., enc.; il. Orientador: Maurício Alves Loureiro. Coorientador: Hani Camille Yehia Linha de pesquisa: Sonologia. Tese (doutorado) - Universidade Federal de Minas Gerais, Escola de Música. Inclui bibliografia. 1. Música - Teses. 2. Clarineta. 3. Processamento do som por computador. I. Loureiro, Maurício Alves. II. Yehia, Hani Camille. III. Universidade Federal de Minas Gerais. Escola de Música. IV. Título. CDD: 789.96
-------	---

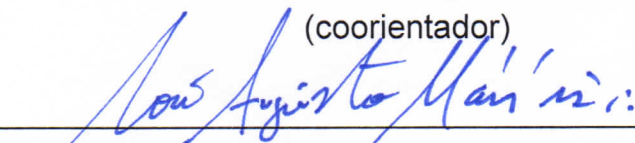
Tese defendida pelo aluno THIAGO DE ALMEIDA MAGALHÃES CAMPOLINA, em 19 de novembro de 2019, e aprovada pela Banca Examinadora constituída pelos Professores:



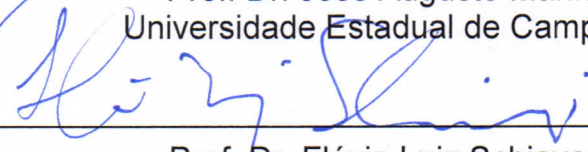
Prof. Dr. Mauricio Alves Loureiro
Universidade Federal de Minas Gerais
(orientador)



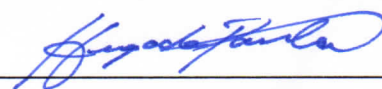
Prof. Dr. Hani Camille Yehia
Universidade Federal de Minas Gerais
(Escola de Engenharia)
(coorientador)



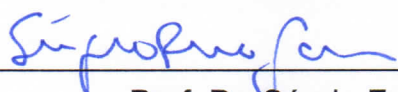
Prof. Dr. José Augusto Mannis
Universidade Estadual de Campinas



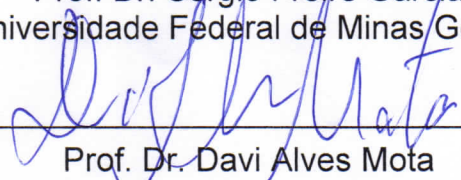
Prof. Dr. Flávio Luiz Schiavoni
Universidade Federal de São João del-Rei



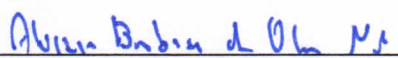
Prof. Dr. Hugo Bastos de Paula
Pontifícia Universidade Católica de Minas Gerais



Prof. Dr. Sérgio Freire Garcia
Universidade Federal de Minas Gerais



Prof. Dr. Davi Alves Mota
(Programa de Pós-Doutorado com Experiência no Exterior/Capes)
Universidade Federal de Minas Gerais



Prof. Dr. Aluizio Barbosa de Oliveira Neto
(Músico)

*Dedico esta Tese a todas as pessoas que fazem parte da minha vida de forma positiva,
sendo amadas por mim e me amando.*

Agradecimentos

Agradeço aos meus orientadores Prof. Maurício Alves Loureiro e Prof. Hani Camille Yehia, e ao Prof. Sérgio Freire, pelos valiosos ensinamentos, conselhos, motivações e apoio. Aos membros da Banca Examinadora pelas críticas e contribuições. Agradeço imensamente aos meus pais e irmãos pelo apoio e carinho em todos os momentos, me motivando a sempre seguir o meu caminho. À minha companheira Thatiana, por ser inspiração de alegria e amor nos momentos felizes, e pelo suporte nos momentos de stress. Aos meus amigos, que se preocupam comigo e estão sempre dispostos a me ajudar no que for preciso, e a dividir momentos de felicidade. Aos colegas do CEGeME¹ (Centro de Estudos do Gesto Musical e Expressão), CEFALA² (Centro de Estudos da Fala, Acústica, Linguagem e Música), e do PPGMUS (Programa de Pós-Graduação em Música) pela amizade, contribuições e troca de conhecimentos. Ao PPGMUS (Programa de Pós-Graduação em Música), à FAPEMIG (Fundação de Amparo à Pesquisa do Estado de Minas Gerais), ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), e à CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), pelo suporte e contribuições diretas ou indiretas à minha pesquisa.

¹ <www.musica.ufmg.br/cegeme/>

² <www.cefala.org/>

“A mente que se abre a uma nova ideia jamais voltará ao seu tamanho original.”
(Albert Einstein)

Resumo

O ataque das notas pode assumir papéis importantes em performances musicais. No ataque, músicos são capazes de controlar timbre, articulação e ritmo, manipulações essenciais para transmitir suas intenções expressivas e tornar a música mais apreciável. Desde as primeiras investigações objetivas sobre performances, feitas no início do século XX, modelos para a expressividade musical são propostos e aprimorados. A modelagem busca um melhor entendimento sobre como são manipulados os parâmetros relacionados à expressividade. Na literatura, ao mesmo tempo que a importância dos ataques para a percepção e construção da performance é evidenciada, são expostas também as dificuldades de analisá-los. Eles são eventos acústicos de curta duração, sujeitos a interações com a reverberação do ambiente, e com características transientes e não lineares. Estudos sobre ataques geralmente criam condições distantes da performance, como análise de notas isoladas e gravações em câmaras anecoicas, visando isolar variáveis. Eles, por um lado, trazem resultados essenciais para a construção do conhecimento. Por outro lado, esse distanciamento desconsidera a interação entre músico e ambiente, sendo que adaptar a performance à acústica específica de uma sala de concertos faz parte da rotina dos músicos profissionais. Estudos relativamente recentes revelaram resultados que corroboram com o fato de que músicos profissionais ajustam conscientemente seu estilo de tocar sob diferentes condições acústicas, com investigações objetivas dos ajustes das performances sob diferentes condições acústicas da sala. Além disso, a maioria dos estudos voltados para análise de ataques e transições entre notas, encontrados na revisão bibliográfica realizada, utiliza a energia total do sinal, propondo descritores como Índice de Legato, Inclinação e Log do Tempo de Ataque. Poucos estudos foram encontrados utilizando como base a análise do conteúdo espectral.

Motivada por essas limitações apontadas e com o objetivo de superá-las, a pesquisa apresentada nesta tese propõe e explora o uso de modelagem e decomposição espectral do sinal acústico, como uma via para descrever e analisar ataques de notas em gravações de clarineta em condições de prática cotidiana. Obtendo por essa via a modelagem e a decomposição dos harmônicos no espectro de frequências, são propostas três representações para a região do ataque nesse instrumento: (i) a proporção da energia harmônica, isolada do resíduo, no trecho da transição entre notas no qual o ataque está contido (ii) considerando de forma independente o conteúdo harmônico da nota que está sendo tocada, da reverberação da nota anterior e dos resíduos que compõem o ataque mas não fazem parte das estruturas harmônicas; (iii) considerando o perfil temporal dos inícios dos harmônicos modelados. Na primeira representação, o cálculo de um índice de legato já proposto na literatura é adaptado para utilizar a energia da componente harmônica da transição, em vez da energia total, nomeado de Índice de Legato Harmônico (**ILh**). Na segunda, foram confrontadas em duas dimensões (2D), as três proporções de energia (em relação ao sinal original) das componentes extraídas no ataque. Nessa representação 2D (nomeada de plano **H2R2**), foi possível

verificar visualmente e fazer inferências sobre as proporções das componentes harmônicas, não-harmônicas e interações com a acústica da sala, visando enriquecer o estudo das performances musicais na condição de prática cotidiana, onde ambientes reverberantes estão sempre presentes. Na terceira frente, foi proposto um descritor tridimensional, nomeado de Perfil Harmônico do Ataque (**PHA**), envolvendo tempo, frequência e magnitude espectral, com o objetivo de caracterizar a estrutura harmônica de ataques na clarineta. O **PHA** foi definido como a sequência de inícios dos harmônico, começando pela frequência fundamental, seguindo a ordem harmônica crescente. O início de cada harmônico foi considerado como 50 % da faixa de variação de sua magnitudes em dB no trecho do ataque.

O **ILh**, o plano **H2R2** e o **PHA** foram testados usando gravações de um trecho de uma peça para clarineta do repertório clássico tradicional, tocadas por seis músicos profissionais. Para as análises das proporções de energia **H2** e **R2** e do **ILh**, foram selecionados dois grupos de notas do trecho. Testes estatísticos **ANOVA** (*Analysis of Variance*) indicaram diferenças significativas ($p < 0,05$), considerando “músico” e “nota” como fator. O fator “condição” (com e sem metrônomo) não obteve diferenças significativas ($p > 0,05$) nos testes t para nenhuma variável. Para o plano **H2R2** e o **PHA**, foram examinados ataques de notas em transições com diferentes demandas técnicas de execução, em condição de *legato* e articuladas. Na análise do plano **H2R2**, testes estatísticos para comparação entre médias bidimensionais (**MANOVA** - *Multivariate Analysis of Variance*) indicaram diferenças significativas ($p < 0,05$), considerando “músico” como fator. No caso do **PHA**, foi avaliada sua capacidade de classificação por músicos e, sendo este um perfil tridimensional, a verificação foi feita nas projeções Harmônico-Magnitude e Harmônico-Tempo, separadamente. A capacidade de classificação foi avaliada através de comparações utilizando Coeficiente de Correlação de *Pearson* entre pares de perfis em duas situações: *Intra* e *Inter* classes. Os grupos *Intra* classes e as médias *Intra* e *Inter* classes foram comparados utilizando testes estatísticos **ANOVA** e teste t. Os testes apontaram separabilidade ($p < 0,05$) *Intra* classes em todos os perfis de início dos harmônicos (**PHA**) dos ataques das sete notas avaliadas. Diferenças significativas ($p < 0,05$) entre as médias de correlações *Intra* e *Inter* ocorreram em 4 notas, com 3 delas para a projeção **HxT** em notas ligadas, sugerindo um indicativo de que a informação sobre o tempo de inícios dos harmônicos são mais relevantes em notas ligadas. Valores altos da média *Inter* para as projeções **HxM** (o mesmo não acontece com **HxT**) sugere a existência de uma estrutura de magnitudes dos harmônicos que compõem o **PHA**, comum aos músicos, que pode estar relacionada às proporções entre energias dos harmônicos que define o timbre do instrumento.

Palavras-chaves: Ataque de notas, Clarineta, Timbre, Expressividade Musical.

Abstract

Note attacks have an important role in musical performances. By manipulating note attack quality, musicians are able to control timbre, articulation, rhythm, that are essential parameters for conveying their expressive intentions making music more appreciable. Models for musical expressiveness are proposed and improved since the first objective investigations on performances at the beginning of the last century. Modeling aims at a better understanding of how these parameters related to expressiveness are manipulated. In the literature, while the attack's importance for the perception and the performance construction is pointed, the analysis difficulties are also exposed. They are short-duration acoustic events, subjected to interactions with environmental reverberation, with transient and nonlinear characteristics. Studies about note attacks often create conditions that are far from performance, analyzing isolated notes and recording in anechoic chambers to isolate variables. They bring essential results for the construction of knowledge. However, this distance disregards the interaction between the musician and environment, while adapting the performance to the specific acoustics of a concert hall is part of the routine of professional musicians. Relatively recent studies investigating performance adjustments under different room acoustic conditions have revealed results that corroborate that professional musicians consciously adjust their style under different concert halls. In addition, most of the studies found in the bibliographic review carried out that analyze attacks and transitions between notes, use the total energy of the signal, proposing descriptors such as Legato Index, Inclination, and Attack Time Log. Few studies have been found based on spectral content analysis.

Motivated by these limitations and aiming to overcome them, the research presented in this thesis proposes and explores the use of modeling and spectral decomposition of the acoustic signal, as a way to describe and analyze note attacks in clarinet recordings under daily practice conditions. By obtaining the modeling and decomposition of the harmonics in the frequency spectrum, three representations are proposed for the attack region in this instrument: (i) the proportion of the harmonic energy, isolated from the residue, during the transition between notes in which the attack is contained (ii) considering independently the harmonic content of the note being played, the reverberation of the previous note and the residues that are part of the attack but not from the harmonic structures; (iii) considering the temporal profile of the modeled harmonics onsets. In the first representation, a legato index already proposed in the literature is adapted to use the energy of the harmonic component of the transition, instead of the total energy, named Harmonic Legato Index (ILh). In the second representation, the energy proportions of the three components (in relation to the original signal) extracted during the attack were confronted each other in only two dimensions (2D). In this 2D representation (called H2R2 plane), it was possible to visualize and make inferences about the proportions of the harmonic, non-harmonic

components and interactions with the acoustics of the room, aiming to enrich the study of musical performances in everyday practice condition, where reverberating environments are always present. In the third front, it is proposed a three-dimensional descriptor, named Harmonic Attack Profile (PHA), involving time, frequency, and spectral magnitudes, in order to characterize the harmonic structure of note attacks on the clarinet. PHA was defined as the sequence of the harmonic onsets, starting with the fundamental frequency, in the increasing harmonic order. Each harmonic onset was considered as 50% of its magnitudes range in dB during the attack.

The ILh, the H2R2 plane, and the PHA were tested using recordings of a clarinet piece from the traditional classical repertoire, played by six professional musicians. For the analysis of the H2, R2, and ILh energy proportions, two groups of notes were selected. statistical tests ANOVA (Analysis of Variance) indicated significant differences ($p < 0.05$), considering “musician” and “note” as factors. The “condition” factor (with and without a metronome) did not present significant differences ($p > 0.05$) in the t-tests for any variable. For the H2R2 plane and for the PHA, note attacks were examined in transitions with different technical execution demands, in *legato* and articulated conditions. In the H2R2 plane analysis, statistical tests for comparison between two-dimensional means (MANOVA - Multivariate Analysis of Variance) indicated significant differences ($p < 0.05$), considering “musician” as a factor. For the PHA, we evaluated its capacity for classification by musicians, and as it is a three-dimensional profile, the verification was made in the Harmonic-Magnitude and Harmonic-Time projections, separately. The classification capacity was evaluated through comparisons using Pearson’s Correlation Coefficient between pairs of profiles in two situations: Intra and Inter classes. The Intra class groups and the Intra and Inter class means were compared using ANOVA and t-test. The tests showed Intra class differences ($p < 0.05$) in all profiles of harmonic onsets (PHA) of the seven-note attacks evaluated. Significant differences ($p < 0.05$) between Intra and Inter correlation means occurred in four notes, with three of them to the HxT projection in *legato* notes, suggesting an indication that the time information of harmonic onsets is more relevant in *legato* notes. High values of the Inter means for HxM projections (the same does not happen with HxT) suggest the existence of a harmonic magnitude structure composing the PHA, common to musicians, which may be related to the harmonic energy proportions that define the timbre of the instrument.

Key-words: Note Attacks, Clarinet, Timbre, Musical Expressiveness.

Lista de Figuras

Figura 1 – Representação da anatomia do aparelho auditivo humano. Adaptado de (HUANG et al., 2001, p.29).	26
Figura 2 – Curvas isofônicas médias: determinadas experimentalmente com volumes fixados, quando variada a intensidade sonora (Nível de Intensidade Sonora) e frequência, tomando como referência $0dB$ a $1KHz$. Adaptado de (FLETCHER; MUNSON, 1933).	28
Figura 3 – Curva com as frequências centrais das 24 bandas críticas da escala de Bark. Fica evidenciando o caráter não linear em relação à frequência. Adaptado de (HUANG et al., 2001, p.33).	28
Figura 4 – Mascaramento em Frequência. Influência do Limite de Mascaramento de uma senoide a $1kHz$ e um nível de pressão sonora igual a $69dB$ na curva de Limiar Absoluto de Audição. Adaptado de (HUANG et al., 2001, p.35).	29
Figura 5 – Limite de mascaramento temporal. Som mascarador uniforme entre $0ms$ e $200ms$: comportamento diferente de pré e pós-mascaramento. Adaptado de (HUANG et al., 2001, p.36).	29
Figura 6 – Espectrograma de um fragmento extraído de uma gravação utilizada nas análises desta tese. Três elementos encontrados nos ataques das notas são destacados.	35
Figura 7 – Áreas usadas para o cálculo do Índice de Legato.	37
Figura 8 – Processo de segmentação utilizado. Detecção dos instantes delimitadores: Início de Nota (<i>Onsets</i>), Final de Nota (<i>Offsets</i>), Início de Decaimento, e Final de Ataque.	44
Figura 9 – Decomposição espectral do sinal: cálculo das curvas de energia RMS da componente harmônica e do sinal original da transição; Definição das proporções de energia H1 , R1 , H2 , R2 das componentes extraídas do ataque.	45
Figura 10 – Diagrama representando o Índice de Legato calculado a partir da curva de energia RMS da parte harmônica (ILh), e o cálculo do Índice de Legato tradicional (IL).	48
Figura 11 – Confronto entre as curvas de energia calculada utilizando apenas a componente harmônica do sinal (RMS_h), utilizando o sinal completo (RMS) e o complemento de um do Fluxo Espectral. As setas apontam para as regiões de transição de seis notas de uma das gravações utilizadas nos resultados (detalhes das gravações no Capítulo 6). Os saltos do eixo tempo são de $5,81ms$	49

Figura 12 – Plano H2 x R2 : 100 pontos simulando o limite superior, com R2(dB) assumindo valores entre $0dB$ (igual a soma total S_E) e $-20dB$, variando de acordo com: $R2(dB) = S_E - H2(dB)$, e a componente H1 sem energia.	51
Figura 13 – Dois exemplos de PHA : harmônicos na região do ataque de duas notas com seus PHAs representados por linhas tracejadas (metade superior da figura), e seus respectivos perfis isolados (metade inferior da figura).	52
Figura 14 – Exemplos de PHA : dois casos de ataques com harmônicos que decrescem antes de terem suas energias aumentadas	52
Figura 15 – Excerto do primeiro movimento da Sonata para Clarineta e Piano em Fá menor op. 120 No. 1 de Johannes Brahms, utilizado nos testes experimentais (figura adaptada de Teixeira et al. (2015)).	55
Figura 16 – Mapa de recorrência de movimento da campana da clarineta, retirado de (TEIXEIRA et al., 2015), que contribuiu para a escolha dos grupos de notas utilizadas para análise. Grupo 1: Compassos 5 a 10 e Grupo 2: compassos 17 e 18.	55
Figura 17 – <i>Boxplots</i> das proporções H2 para as notas do Grupo 1.	58
Figura 18 – <i>Boxplots</i> das proporções H2 para as notas do Grupo 2.	58
Figura 19 – Notas selecionadas para a ANOVA : 37 a 42 do Grupo 2.	59
Figura 20 – Resíduos para análise e verificação das premissas da ANOVA . Na metade esquerda, o ajuste à distribuição normal. Na metade direita, a semelhança das variâncias dos resíduos dos ajustes.	59
Figura 21 – Resultado da comparação múltipla H2 x Nota	60
Figura 22 – Boxplots (à esquerda) e médias com intervalos de confiança a 95 % (à direita) para o fator Clarinetista utilizado na ANOVA	61
Figura 23 – Resíduos para análise e verificação das premissas da ANOVA para o teste do fator Clarinetista. Na metade esquerda, o ajuste à distribuição normal. Na metade direita, a semelhança das variâncias dos resíduos dos ajustes.	62
Figura 24 – Resultado da comparação múltipla para o fator Clarinetista na variável H2	62
Figura 25 – Boxplots (à esquerda) e médias com intervalos de confiança a 95 % (à direita) para o fator Condição utilizado no teste t	63
Figura 26 – Resíduos para análise e verificação das premissas do teste t pareado para o fator Condição. Na metade direita, o ajuste à distribuição normal. Na esquerda, o boxplot dos resíduos.	64
Figura 27 – <i>Boxplots</i> das proporções R2 para as notas do Grupo 1.	65
Figura 28 – <i>Boxplots</i> das proporções R2 para as notas do Grupo 2.	66

Figura 29 – Notas selecionadas para a ANOVA no teste do fator Notas para R2: 37 a 42 do Grupo 2.	66
Figura 30 – Resíduos para análise e verificação das premissas da ANOVA . Na metade esquerda, o ajuste à distribuição normal. Na metade direita, a semelhança das variâncias dos resíduos dos ajustes.	67
Figura 31 – Resultado da comparação múltipla R2 x Nota	67
Figura 32 – Boxplots (à esquerda) e médias com intervalos de confiança a 95 % (à direita) para o fator Clarinetista utilizados na ANOVA	68
Figura 33 – Resíduos para análise e verificação das premissas da ANOVA para o teste da influência do fator Clarinetista na variável R2. Na metade esquerda, o ajuste à distribuição normal. Na metade direita, a semelhança das variâncias dos resíduos dos ajustes.	69
Figura 34 – Resultado da comparação múltipla para o fator Clarinetista na variável R2.	70
Figura 35 – Boxplots (à esquerda) e médias com intervalos de confiança a 95 % (à direita) para o fator Condição utilizados no teste t	70
Figura 36 – Resíduos para análise e verificação das premissas do teste t pareado para o fator Condição com a variável R2 . Na metade direita, o ajuste à distribuição normal. Na esquerda, o boxplot dos resíduos.	71
Figura 37 – <i>Boxplots</i> de IL para as notas do Grupo 1.	72
Figura 38 – <i>Boxplots</i> de IL para as notas do Grupo 2.	73
Figura 39 – <i>Boxplots</i> de IL para as notas ligadas do Grupo 1.	74
Figura 40 – Notas selecionadas para a IL x Nota : 1 a 13 do Grupo 1.	74
Figura 41 – Resíduos para análise e verificação das premissas da ANOVA . Na metade esquerda, o ajuste à distribuição normal. Na metade direita, a semelhança das variâncias dos resíduos dos ajustes.	75
Figura 42 – Resultado da comparação múltipla IL x Nota	75
Figura 43 – Boxplots (à esquerda) e médias com intervalos de confiança a 95 % (à direita) para o fator Clarinetista utilizado na ANOVA	76
Figura 44 – Resíduos para análise e verificação das premissas da ANOVA para o teste da influência do fator Clarinetista na variável IL . Na metade esquerda, o ajuste à distribuição normal. Na metade direita, a semelhança das variâncias dos resíduos dos ajustes.	77
Figura 45 – Resultado da comparação múltipla para o fator Clarinetista na variável IL	77
Figura 46 – <i>Boxplots</i> por nota e médias com intervalos de confiança a 95 % para as diferenças IL - ILh avaliadas no Grupo 1: Compassos 5 a 10, e no Grupo 2: compassos 17 e 18.	78

Figura 47 – Resíduos para análise e verificação das premissas do teste t pareado IL x ILh . Na metade direita, o ajuste à distribuição normal. Na esquerda, o boxplot dos resíduos.	79
Figura 48 – (a) Plano H2R2 para os ataques de todas as execuções da nota 4 (nota com transição em condição de <i>legato</i> , partindo da nota 3). E1 e E2 indicam duas execuções em destaque de um mesmo clarinetista. Em (b) e (c) estão mostradas as formas de onda, espectrogramas e fluxo espectral para E2 e E1	81
Figura 49 – Planos H2R2 para os ataques de todas as execuções das notas 10, 12, 14 e 22, cujas transições são em condição de <i>legato</i> partindo da nota anterior.	82
Figura 50 – Planos H2R2 para os ataques de todas as execuções das notas 9 e 13 que, nesses dois casos, apresentam transições articuladas (não ligadas) partindo da nota anterior.	82
Figura 51 – Notas excluídas da análise das projeções do PHA , utilizando um limiar de corte de -3 dB, em ambas as variáveis do plano H2R2 , para identificar ataques mal sucedidos. Os ataques excluídos estão listados no canto esquerdo inferior de cada plano H2R2 da figura, com a simbologia de “C” para clarinetista e “E” para execução.	85
Figura 52 – Estatísticas das comparações <i>Intra</i> e <i>Inter</i> grupos (de tamanho N) utilizando coeficientes de correlação de <i>Pearson</i> entre as projeções do PHA - Harmônico por Tempo (HxT) e por Magnitude (HxM) - clarinetistas C1 a C6 executando a nota 4. Médias das correlações com Intervalos de Confiança (I.C.) a 95% nas colunas à direita. Média geral <i>Intra</i> e o valor p da ANOVA ao centro.	86
Figura 53 – Projeções do PHA , nota 4, clarinetistas C1 a C6: Harmônico por Tempo (HxT) em (1a) até (1f) e Harmônico por Magnitude (HxM) em (2a) até (2f).	87
Figura 54 – Estatísticas das comparações <i>Intra</i> e <i>Inter</i> grupos (de tamanho N) utilizando coeficientes de correlação de <i>Pearson</i> entre as projeções do PHA - Harmônico por Tempo (HxT) e por Magnitude (HxM) - clarinetistas C1 a C6 executando a nota 9. Médias das correlações com Intervalos de Confiança (I.C.) a 95% nas colunas à direita. Média geral <i>Intra</i> e o valor p da ANOVA ao centro.	88
Figura 55 – Projeções do PHA , nota 9, clarinetistas C1 a C6: Harmônico por Tempo (HxT) em (1a) até (1f) e Harmônico por Magnitude (HxM) em (2a) até (2f).	89

Figura 56 – Estatísticas das comparações <i>Intra</i> e <i>Inter</i> grupos (de tamanho N) utilizando coeficientes de correlação de <i>Pearson</i> entre as projeções do PHA - Harmônico por Tempo (HxT) e por Magnitude (HxM) - clarinetistas C1 a C6 executando a nota 10. Médias das correlações com Intervalos de Confiança (I.C.) a 95% nas colunas à direita. Média geral <i>Intra</i> e o valor p da ANOVA ao centro.	90
Figura 57 – Projeções do PHA , nota 10, clarinetistas C1 a C6: Harmônico por Tempo (HxT) em (1a) até (1f) e Harmônico por Magnitude (HxM) em (2a) até (2f).	91
Figura 58 – Estatísticas das comparações <i>Intra</i> e <i>Inter</i> grupos (de tamanho N) utilizando coeficientes de correlação de <i>Pearson</i> entre as projeções do PHA - Harmônico por Tempo (HxT) e por Magnitude (HxM) - clarinetistas C1 a C6 executando a nota 12. Médias das correlações com Intervalos de Confiança (I.C.) a 95% nas colunas à direita. Média geral <i>Intra</i> e o valor p da ANOVA ao centro.	92
Figura 59 – Projeções do PHA , nota 12, clarinetistas C1 a C6: Harmônico por Tempo (HxT) em (1a) até (1f) e Harmônico por Magnitude (HxM) em (2a) até (2f).	93
Figura 60 – Estatísticas das comparações <i>Intra</i> e <i>Inter</i> grupos (de tamanho N) utilizando coeficientes de correlação de <i>Pearson</i> entre as projeções do PHA - Harmônico por Tempo (HxT) e por Magnitude (HxM) - clarinetistas C1 a C6 executando a nota 13. Médias das correlações com Intervalos de Confiança (I.C.) a 95% nas colunas à direita. Média geral <i>Intra</i> e o valor p da ANOVA ao centro.	94
Figura 61 – Projeções do PHA , nota 13, clarinetistas C1 a C6: Harmônico por Tempo (HxT) em (1a) até (1f) e Harmônico por Magnitude (HxM) em (2a) até (2f).	95
Figura 62 – Estatísticas das comparações <i>Intra</i> e <i>Inter</i> grupos (de tamanho N) utilizando coeficientes de correlação de <i>Pearson</i> entre as projeções do PHA - Harmônico por Tempo (HxT) e por Magnitude (HxM) - clarinetistas C1 a C6 executando a nota 14. Médias das correlações com Intervalos de Confiança (I.C.) a 95% nas colunas à direita. Média geral <i>Intra</i> e o valor p da ANOVA ao centro.	96
Figura 63 – Projeções do PHA , nota 14, clarinetistas C1 a C6: Harmônico por Tempo (HxT) em (1a) até (1f) e Harmônico por Magnitude (HxM) em (2a) até (2f).	97

- Figura 64 – Estatísticas das comparações *Intra* e *Inter* grupos (de tamanho N) utilizando coeficientes de correlação de *Pearson* entre as projeções do **PHA** - Harmônico por Tempo (**HxT**) e por Magnitude (**HxM**) - clarinetistas C1 a C6 executando a nota 22. Médias das correlações com Intervalos de Confiança (I.C.) a 95% nas colunas à direita. Média geral *Intra* e o valor p da **ANOVA** ao centro. 98
- Figura 65 – Projeções do **PHA**, nota 22, clarinetistas C1 a C6: Harmônico por Tempo (**HxT**) em (1a) até (1f) e Harmônico por Magnitude (**HxM**) em (2a) até (2f). 99
- Figura 66 – Comparações entre as médias dos coeficientes de correlação para todas as notas. Projeção HxT em (a) e HxM em (b). Valores das médias com Intervalos de Confiança 95 % e valores p dos testes t na parte inferior dos gráficos. 100

Lista de tabelas

Tabela 1 – Relação entre a designação dos atributos perceptivos e físicos dos sons.	27
Tabela 2 – Estatísticas da ANOVA para o teste H2 x Nota .	59
Tabela 3 – Estatísticas da ANOVA para o teste H2 x Clarinetista .	61
Tabela 4 – Estatísticas do teste t pareado para H2 x Clarinetista .	63
Tabela 5 – Estatísticas da ANOVA para o teste R2 x Nota .	65
Tabela 6 – Estatísticas da ANOVA para o teste R2 x Clarinetista .	68
Tabela 7 – Estatísticas do teste t pareado para R2 x Condição .	71
Tabela 8 – Estatísticas da ANOVA para o teste IL x Nota .	73
Tabela 9 – Estatísticas da ANOVA para o teste IL x Clarinetista .	76
Tabela 10 – Estatísticas do teste t pareado para IL x ILh .	79
Tabela 11 – Resultados do teste MANOVA (<i>Multidimensional Analysis of Variance</i>) considerando “clarinetista” (C1 a C6) como fator. São apresentadas as variâncias das variáveis <i>R2</i> e <i>H2</i> , o par de clarinetistas comparados com Contraste Máximo e, finalmente, os valores-p, para cada uma das notas selecionadas.	83
Tabela 12 – Resumo dos resultados da análise de correlações realizada nas duas projeções do PHA , Harmônico x Tempo (HxT) e Harmônico x Magnitude (HxM): Testes ANOVA (comparação entre as médias dos coeficientes de correlação dos grupos <i>Intra</i>) e testes t (comparação entre as médias <i>Intra</i> e <i>Inter</i> grupos).	101

Lista de abreviaturas e siglas

RMS	<i>Root Mean Square.</i>
RMS_h	RMS da componente harmônica do sinal.
SMS	<i>Spectral Modeling Synthesis.</i>
PHA	<i>Perfil Harmônico do Ataque.</i>
IL	Índice de Legato.
Log	Logaritmo.
ABX	Teste para detectar diferenças entre duas opções de estímulos sensoriais.
MOS	<i>Mean Opinion Score.</i>
dB	<i>Decibel.</i>
FFT	<i>Fast Fourier Transform.</i>
STFT	<i>Short Time Fourier Transform.</i>
IFFT	<i>Inverse Fast Fourier Transform.</i>
TWM	<i>Two Way Mismatch.</i>
IL_h	Índice de Legato da componente harmônica do sinal.
H2R2	Plano que confronta as componentes H2 e R2 .
H1R1	Plano que confronta as componentes H1 e R1 .
H1	Componente espectral harmônica do ataque.
H2	Componente espectral harmônica da nota anterior presente no ataque.
R1	Componente espectral do resíduo de H1 .
R2	Componente espectral do resíduo de H1 e H2 .
LPC	<i>Linear Predictive Coding.</i>
FM	<i>Frequency-Modulation Synthesis</i> (síntese FM).
OLA	<i>Overlap Add.</i>

ANOVA	<i>Analysis of Variance.</i>
MANOVA	<i>Multivariate Analysis of Variance.</i>
2D	Duas dimensões.
RTF	<i>Radiodiffusion Télévision Française.</i>
EXPAN	Software desenvolvido no CEGeME .
CEGeME	Centro de Estudos do Gesto Musical e Expressão.
NPS	Nível de Pressão Sonora.
NIS	Nível de Intensidade Sonora.
F0	Frequência Fundamental.
G1	Grupo de notas 1.
G2	Grupo de notas 2. Df Sum Sq Mean Sq F value Pr(>F)
diff	Diferenças entre as notas comparadas em dB.
lwr	Valor inferior do intervalo de confiança;
upr	Valor superior do intervalo de confiança;
padj	Valor p ajustado pela comparação múltipla.
Min	Valor mínimo das diferenças;
Med	Mediana das diferenças;
Max	Valor máximo das diferenças.
W	Estatística do teste <i>Shapiro-Wilk</i> .
Df	Grau de liberdade.
Sum Sq	Soma dos quadrados.
Mean Sq	Soma quadrática média.
F value	Valor da estatística F.
Pr(>F)	Valor p para a estatística F.
IC	Intervalo de Confiança.
L	Condição Livre.

M	Condição Metrônomo.
C1,...,C6	Clarinetista 1, ..., Clarinetista 6.
E1,...,E6	Execução 1, ..., Execução 6.
HxM	Projeção Harmônico-Magnitude.
HxT	Projeção Harmônico-Tempo.
Intra	Grupo Intra-classe.
Inter	Grupo Inter-classe.

.

.

Sumário

1	INTRODUÇÃO	23
2	ASPECTOS PERCEPTIVOS E MOTIVAÇÕES	26
2.1	Aparelho Auditivo	26
2.2	Atributos e Fenômenos Psicoacústicos	27
2.2.1	Volume	27
2.2.2	Banda Crítica	27
2.2.3	Mascaramento em Frequência	29
2.2.4	Mascaramento Temporal	29
2.3	Percepção de Ataques em Objetos Sonoros	30
2.3.1	Limiares de Audição	30
2.3.2	Altura (<i>Pitch</i>)	31
2.3.3	Intensidade	31
2.3.4	Duração	32
2.3.5	Timbre	33
2.4	Motivações	34
3	DESCRIÇÃO DE CONTEÚDO ACÚSTICO EM PERFORMANCES	36
3.1	Deteccção dos Ataques e Decaimentos	36
3.2	Índice de Legato	37
3.3	Fluxo Espectral	38
4	MODELAGEM E DECOMPOSIÇÃO ESPECTRAL	39
4.1	Spectral Modeling Synthesis (SMS)	39
5	METODOLOGIA PROPOSTA	43
5.1	Segmentação	43
5.2	Deteccção dos Ataques e Decaimentos	44
5.3	Decomposição Espectral	45
5.4	Modelagens e Parâmetros de Ajuste	46
5.5	Índice de Legato	48
5.6	Plano H2R2	50
5.7	Perfil Harmônico do Ataque	51
6	RESULTADOS E DISCUSSÕES	54
6.1	Material	54
6.2	Variáveis de interesse	55

6.3	Fatores de Interesse	56
6.4	Testes Realizados	56
6.5	Verificação de Efeito nas variáveis H2 e R2	57
6.6	Transição entre Notas	71
6.7	Caracterização dos Ataques	79
7	CONCLUSÃO	103
	REFERÊNCIAS	105

1 Introdução

Em uma performance musical, o intérprete atua sobre elementos fundamentais da música para se expressar: altura, timbre, intensidade e duração. Nesse processo, estão envolvidos tanto aspectos pessoais quanto interpretativos da obra interpretada. Desde o início do século XX, modelos para representar a expressividade musical vêm sendo propostos e aprimorados. A modelagem busca um melhor entendimento sobre como são manipulados os elementos expressivos que tornam a música mais apreciável. Nesse contexto, o ataque das notas pode assumir papéis importantes.

Na literatura, ao mesmo tempo que a importância dos ataques para a percepção e construção da performance é evidenciada, são expostas também as dificuldades de analisá-los. Os ataques são eventos acústicos de curta duração, sujeitos a interações com a reverberação do ambiente, e com características transientes e não lineares. Estudos relacionados à região de ataque em sinais musicais vêm sendo feitos desde a década de 1960. Sua importância para a percepção é demonstrada em vários trabalhos (LUCE, 1963; RISSET, 1965; GREY, 1977). Luce (1963), por exemplo, mostrou que a identificação do instrumento musical é possível com apenas 60 ms do ataque, enquanto que com a sustentação são necessários 150 ms. Risset (1965) analisou sons de trompete e, após a obtenção de curvas individuais de evolução temporal das amplitudes e frequências para cada harmônico, conseguiu sintetizar sons de trompete a partir de aproximações, por segmentos lineares, das curvas obtidas na análise. Mostrou também que há diferença no tempo de início de cada harmônico no ataque. Grey (1977) mostrou que durante o ataque ocorre uma mudança no equilíbrio entre os harmônicos devido a diferentes taxas de crescimento e decrescimento dos mesmos.

Esses estudos sobre ataques apontados no parágrafo anterior criam condições distantes da performance, como análise de notas isoladas e gravações em câmaras anecoicas, visando isolar variáveis. Eles trazem resultados essenciais para a construção do conhecimento. Por outro lado, esse distanciamento desconsidera a interação entre músico e ambiente, pois adaptar a performance à acústica específica de uma sala de concertos faz parte da rotina dos músicos profissionais. Em Kato, Ueno e Kawai (2008), um experimento com músicos profissionais revelou que eles ajustam conscientemente seu estilo de tocar sob diferentes condições acústicas da sala. Uma entrevista com os músicos indicou que cada participante ajustou conscientemente o estilo de atuação de acordo com alguns dos seguintes atributos: tempo, vibrato, harmônicos, qualidade do som, articulação, agógica e dinâmica. Além de entrevistas, esse e outros estudos se concentraram também na investigação objetiva dos ajustes das performances feitas por músicos profissionais sob diferentes condições acústicas da sala (KATO; UENO; KAWAI, 2008; KALKANDJIEV;

WEINZIERL, 2015; BOLZINGER; WARUSFEL; KAHLE, 1994; FISCHINGER; FRIELER; LOUHIVUORI, 2015). Em Kalkandjiev e Weinzierl (2015), por exemplo, modelos computacionais de 14 salas foram testados por síntese binaural dinâmica com artistas tocando fisicamente em uma câmara anecoica, enquanto usavam fones de ouvido simulando ambientes acústicos virtuais correspondentes. Este experimento revelou conceitos distintos de ajuste às condições acústicas da sala, bem como individualidade significativa com relação à interação dos músicos com a acústica da sala.

Com isso, a motivação desta tese é contribuir com métodos para tornar as análises objetivas dos ataques mais próximas à condição de performance. Nessa direção, em Loureiro et al. (2009) foram estudadas as transições entre notas musicais em trechos de repertório gravados em ambiente com reverberação, evidenciando diferentes características da qualidade das transições, sugerindo que tais características são determinantes na construção de uma interpretação musical. Nas situações em que notas devem ser ligadas (condição de *Legato*), por exemplo, a manipulação dos ataques assume um papel importante. A qualidade desse tipo de transição demanda habilidades para sincronizar manipulações de energia, digitação e pressão, lidando com as características acústicas do instrumento, além do ajuste à resposta acústica do ambiente. A análise objetiva do *Legato* na literatura se concentra na análise da energia total como o Índice de Legato proposto em Maestre e Gómez (2005), sendo pouco explorada no âmbito de detalhamento espectral. Portanto, pelo fato de as transições em *Legato* estarem diretamente envolvidas com ataques, o Índice de Legato é também tema desta tese.

Motivada pelas limitações nas metodologias de análise do ataque apontadas, a pesquisa apresentada nesta tese propõe e explora o uso de modelagem e decomposição espectral como uma via para descrever e analisar ataques e transições entre notas na clarineta em condições próximas às de performance. Para as transições, o Índice de Legato é comparado com uma versão contendo apenas a parte harmônica. Para a região do ataque, são propostas duas representações: (i) considerando de forma independente, o conteúdo harmônico da nota que está sendo tocada (denominado **H1**), da reverberação da nota anterior (denominado **H2**), e também o conteúdo dos resíduos que compõem o ataque mas não fazem parte das estruturas harmônicas (denominado **R2**); (ii) considerando o perfil temporal dos harmônicos modelados. Na primeira, o objetivo é confrontar em duas dimensões as proporções das componentes extraídas em relação ao sinal original. Nesse plano, é possível verificar visualmente e fazer inferências sobre as proporções das componentes harmônicas, não-harmônicas e interações com a acústica da sala, visando enriquecer o estudo das performances musicais na condição de prática cotidiana, onde ambientes reverberantes estão sempre presentes. Na segunda frente, foi proposto um descritor tridimensional, Perfil Harmônico do Ataque (**PHA**), envolvendo tempo, frequência (harmônicos) e magnitude. Partindo da fundamental, o **PHA** cruza na sequência, o ponto central da excursão de energia em dB de cada harmônico no trecho do ataque. Foram avaliadas as projeções do

PHA nos planos magnitude-harmônico e tempo-harmônico. A abordagem de análise foi testada usando gravações de um trecho de uma peça para clarineta do repertório clássico tradicional, tocadas por seis músicos profissionais. Foram examinadas notas em condição de *legato*, bem como notas articuladas, apresentando diferentes demandas técnicas de execução.

Em princípio, os trechos selecionados para análise são segmentados em notas e, em seguida, detectada a região de transição e de ataques. Nas regiões detectadas, é aplicada a técnica de modelagem espectral *Spectral Modeling Synthesis* (**SMS**), capaz de decompor o sinal em uma parte harmônica e outra parte residual, resultante da subtração espectral da primeira. Na parte residual dos ataques (com outras notas anteriores próximas), a segunda componente harmônica referente à reverberação é também modelada pelo **SMS**. Realizada a decomposição espectral dos trechos, primeiramente a proporção de energia de cada componente é avaliada individualmente por meio de testes estatísticos **ANOVA** - *Analysis of Variance*) e teste t para comparação entre as médias de grupos e verificação de efeitos dos seguintes fatores envolvidos na gravação: músico executante (seis músicos), nota executada (26 notas) e duas situações de execução (livre e com metrônomo). Verificadas as relevâncias estatísticas para os fatores **Músico** e **Notas**, o próximo passo é a avaliação da representação proposta pelo plano formado pelas variáveis **H2** e **R2**, ambas normalizadas em relação à energia total em dB. Por último, é mostrada a avaliação do Perfil Harmônico do Ataque (**PHA**) através das projeções magnitude-harmônico e tempo-harmônico. A avaliação é feita através da comparação dos coeficientes de correlação de *Pearson* entre perfis da mesma nota e músico para gravações diferentes (foram gravadas 6 execuções para cada músico), com os coeficientes de correlação entre perfis da mesma nota e gravações para músicos diferentes.

O próximo capítulo apresenta alguns aspectos perceptivos e as motivações que levaram ao desenvolvimento da pesquisa apresentada nesta tese. Na sequência, os capítulos seguintes tratam da descrição de conteúdo acústico em performances de clarinetistas (Capítulo 3), da modelagem e decomposição espectral utilizada (Capítulo 4), da metodologia proposta (Capítulo 5), dos resultados e discussões (Capítulo 6), e da conclusão (Capítulo 7).

2 Aspectos Perceptivos e Motivações

Neste capítulo são apresentadas as motivações para o desenvolvimento da pesquisa apresentada nesta tese. Inicialmente, as principais questões físicas, fisiológicas e perceptivas envolvidas na escuta dos ataques de notas de instrumentos musicais são abordadas e, no final do capítulo, estão presentes os detalhes das motivações para esta pesquisa.

2.1 Aparelho Auditivo

A Figura 1 contém uma representação da anatomia do aparelho auditivo humano. A estrutura é dividida em três seções: A Orelha Externa, a Orelha Média e a Orelha Interna. A Orelha Externa tem cerca de 2,5 *cm* de comprimento e é coberta pelo Tímpano na extremidade. O Tímpano transmite as vibrações aos ossos ligados a ele no seu lado oposto. Este lado oposto é a Orelha Média, uma cavidade cheia de ar com cerca de 1,3 *cm* de diâmetro e cerca de 6 *cm*³ de volume. A Janela Oval é uma pequena membrana na Cóclea que opera como uma interface para a Orelha Interna. Dentro da Cóclea, região preenchida por um líquido chamado Perilinf, tem seu conteúdo agitado pelos movimentos da Janela Oval e, com isso, milhares de células ciliadas presentes no Órgão de Corti são colocadas em movimento. Na Cóclea, faixas de frequência são detectadas por diferentes regiões. Ligado à Cóclea, encontra-se o Nervo Auditivo que transmite a informação para o cérebro. A atividade neuronal realiza uma integração temporal, fazendo com que o sistema interprete as variações instantâneas indiretamente.

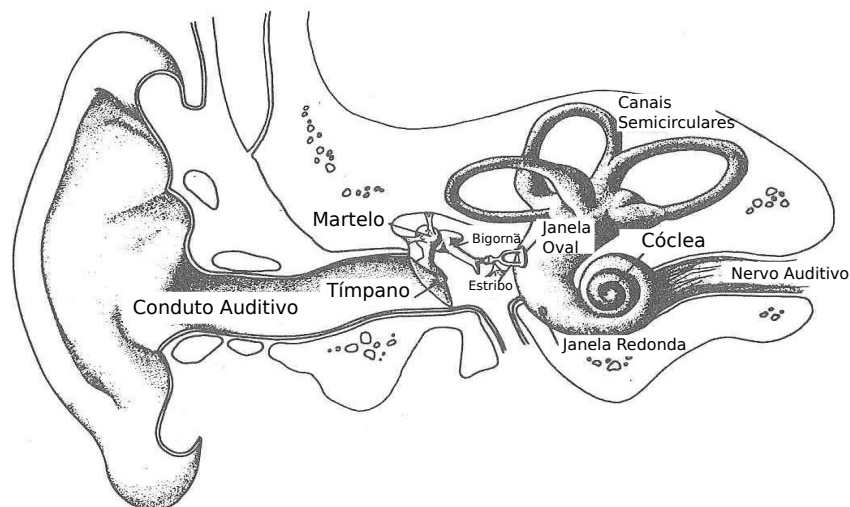


Figura 1 – Representação da anatomia do aparelho auditivo humano. Adaptado de (HUNG et al., 2001, p.29).

2.2 Atributos e Fenômenos Psicoacústicos

O som é uma onda longitudinal de pressão e pode ser medida objetivamente. Por outro lado, a percepção do som (fenômeno estudado pela psicoacústica) é uma sensação causada pelo envio para o cérebro de impulsos elétricos gerados pelo mecanismo auditivo em resposta a uma variação de pressão. Em contraste aos atributos físicos, os atributos subjetivos são difíceis de serem estimados.

É necessário diferenciar as nomenclaturas dos atributos perceptivos e físicos do som na psicoacústica. Cada um dos atributos perceptivos, conforme listado na Tabela 1, tem uma forte conexão com uma propriedade física principal, mas a relação é complexa. Um dos fatos mais importantes é a percepção em escala logarítmica da intensidade sonora e da frequência.

Tabela 1 – Relação entre a designação dos atributos perceptivos e físicos dos sons.

<i>Quantidade Física</i>	<i>Atributo Perceptivo</i>
Pressão Sonora	Volume
Intensidade Sonora	Volume
Frequência Fundamental	Altura
Forma Espectral	Timbre
Duração	Duração

2.2.1 Volume

Volume é a resposta do mecanismo de audição ao parâmetro físico Pressão Sonora. A medida mais comum para variação de pressão sonora é o Nível de Pressão Sonora (NPS), definido como: $20\log_{10}(\frac{P}{P_{ref}})$, com unidade em decibel, onde P é a pressão medida e $P_{ref} = 2 \times 10^{-5} \frac{N}{m^2}$ definido como o Limiar de audição. O NPS é numericamente igual ao Nível de Intensidade sonora (NIS) definido como: $10\log_{10}(\frac{I}{I_{ref}})$, com unidade em decibel, onde I é a intensidade medida e $I_{ref} = 10^{-12} \frac{W}{m^2}$ definido como o Limiar de audição. O volume é uma função dependente da frequência e do nível de pressão sonora. Assim, faixas distintas de frequência têm percepção de intensidade diferentes para um mesmo nível de pressão sonora. Na Figura 2 podem ser vistas curvas isofônicas médias (curvas de *Fletcher-Munson* apresentadas em [Fletcher e Munson \(1933\)](#)), determinadas experimentalmente com volumes fixados, quando variadas intensidade sonora (nível de pressão sonora) e frequência, tomando como referência $0dB$ a $1KHz$ para tons puros.

2.2.2 Banda Crítica

O sistema auditivo realiza análise de frequência de sons. A estrutura da cóclea comporta-se como um banco de filtros com sobreposição entre eles e largura de banda

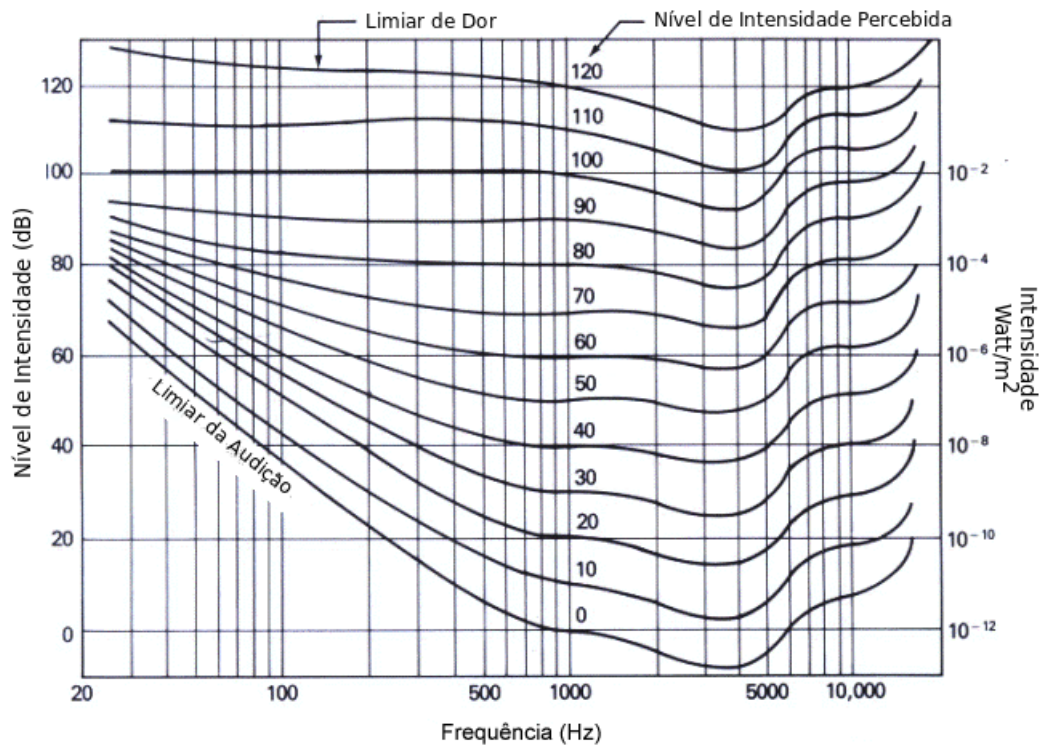


Figura 2 – Curvas isofônicas médias: determinadas experimentalmente com volumes fixados, quando variada a intensidade sonora (Nível de Intensidade Sonora) e frequência, tomando como referência 0dB a 1KHz . Adaptado de (FLETCHER; MUNSON, 1933).

definida como banda crítica. Este fenômeno é importante na compreensão de processos auditivos como volume, altura e timbre. Uma das representações da escala de frequência da cóclea é a escala de Bark, que varia de 1 a 24, correspondendo a 24 bandas críticas de audição. Na Figura 3 pode ser observada uma curva com as frequências centrais das 24 bandas críticas, evidenciando seu caráter não linear em relação à frequência.

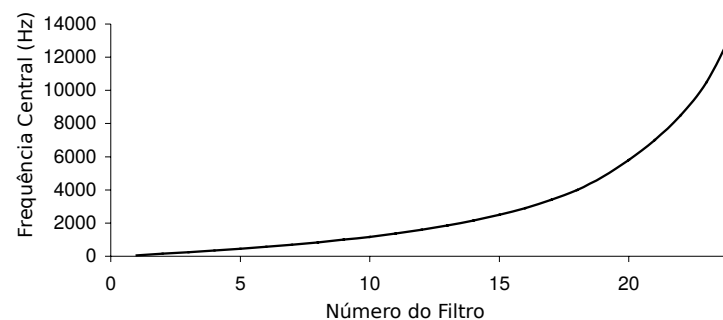


Figura 3 – Curva com as frequências centrais das 24 bandas críticas da escala de Bark. Fica evidenciando o caráter não linear em relação à frequência. Adaptado de (HUANG et al., 2001, p.33).

2.2.3 Mascaramento em Frequência

O fenômeno Mascaramento em Frequência ocorre quando um som mascara toda uma região de outros sons com frequência próxima e energia abaixo de um limiar (determinado empiricamente). No lado esquerdo da Figura 4 pode ser observado o Limiar Absoluto de Audição. No lado direito da mesma figura, o mesmo limiar pode ser observado, porém agora com a influência do Limite de Mascaramento de uma onda senoidal com frequência igual a $1kHz$ e um nível de pressão sonora igual a $69dB$.

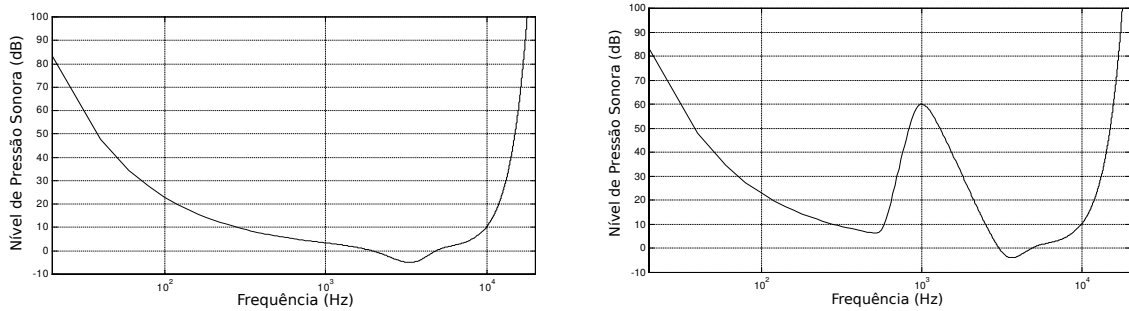


Figura 4 – Mascaramento em Frequência. Influência do Limite de Mascaramento de uma senoide a $1kHz$ e um nível de pressão sonora igual a $69dB$ na curva de Limiar Absoluto de Audição. Adaptado de (HUANG et al., 2001, p.35).

2.2.4 Mascaramento Temporal

O fenômeno chamado de Mascaramento Temporal ocorre quando um som não pode ser percebido por estar muito próximo temporalmente de outro som. Há pré e pós-mascaramento. Para uma mesma condição de nível de pressão sonora e duração de um som mascarador, enquanto o pré-mascaramento pode ocorrer em cerca de $5ms$, o pós-mascaramento pode durar de 50 a $300ms$. Na Figura 5 pode ser observado um exemplo de um limite de mascaramento temporal com mascarador uniforme começando em $0ms$ e duração de $200ms$, mostrando comportamento diferente no pré e pós-mascaramento.

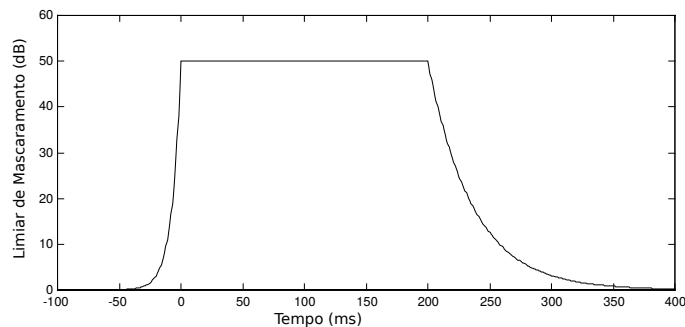


Figura 5 – Limite de mascaramento temporal. Som mascarador uniforme entre $0ms$ e $200ms$: comportamento diferente de pré e pós-mascaramento. Adaptado de (HUANG et al., 2001, p.36).

2.3 Percepção de Ataques em Objetos Sonoros

De uma forma resumida, objeto sonoro é o som em si, dissociado da fonte de produção e de contexto previamente envolvidos na produção e escuta dos sons. Segundo Schaeffer no **Solfejo dos Objetos Sonoros** (SCHAEFFER, 1967), o ataque é um dos eventos mais familiares na música. Schaeffer relata ainda que o conhecimento adquirido com a música concreta e técnicas disponíveis na época levaram às seguintes conclusões: O ataque é muito importante na criação de objetos musicais e varia com a natureza do corpo em vibração e com a forma em que foi posto a vibrar. Varia também com a dinâmica do início da nota, qualitativamente classificada como percussiva, explosiva, etc. A complexidade harmônica emitida pelo corpo ressonante é também uma variável importante.

O **Solfejo dos Objetos Sonoros** é uma sequência de áudios, na qual são apresentados vários experimentos perceptivos. Guiados pela narração do autor, os áudios dos experimentos possibilitam aos ouvintes a verificação das conclusões com o próprio sistema auditivo. Pierre Henri Marie Schaeffer, nasceu na França em 1910 e foi um importante compositor, engenheiro de telecomunicações e musicólogo de sua época. Desenvolveu estudos importantes sobre percepção auditiva, conduziu trabalhos pela *Radiodiffusion Télévision Française* (RTF) em Paris, e foi precursor da Música Concreta. As considerações relativas ao ataque em objetos sonoros e limiares de audição, baseadas no **Solfejo dos Objetos Sonoros**, são apresentadas nesta seção.

2.3.1 Limiares de Audição

No **Solfejo dos Objetos Sonoros**, são realizados experimentos com pulsos sintetizados. Até 24 pulsos por segundo, ouvem-se pulsos distintos. A partir de 29 pulsos por segundo, uma sensação de altura é percebida, e o som passa a ter uma granulosidade. À medida que a frequência dos pulsos é aumentada, uma sensação de rugosidade aparece. O mesmo foi feito a partir de um grão de uma nota grave de fagote. A sequência de sensações foi descrita como: inicialmente, choques regulares, seguidos de vestígios rítmicos, chamados de grão, paralelamente a um crescente efeito de altura e, por último, a emergência de uma textura, colorindo a altura.

Outro experimento mostra que existe um limite para a acumulação temporal de objetos. De certa forma a música estabeleceu este limite como sendo semifusas a 60 batidas por minuto, o que corresponde a 62,5 ms ou $\frac{1}{16}$ s. Uma escala descendente de piano foi executada com durações iguais a semifusas ($\frac{1}{16}$ s) e, como esperado, é possível perceber a separação das notas. Não podendo o pianista executar mais rápido, foi usada a função acelerando do gravador e, com isso, as durações passaram a ser de $\frac{1}{32}$ s cada uma. Neste ponto, passa-se a não haver mais uma distinção clara das notas. O mesmo

experimento foi realizado para palavras. Com sílabas de duração média de 40 ms, não foi mais possível reconhecer o que estava sendo dito. Ao aumentar a duração média para 80 ms, a inteligibilidade foi restaurada. A conclusão foi de que a barreira de separação é de 50 ms.

A constante de tempo, considerada o menor intervalo de tempo abaixo do qual o ouvido é indiferente a natureza dos estímulos foi de 5 ms. Qualquer aparição entre 0 e 5 ms é percebido como ruído parasita, ou seja, uma pequena explosão.

Um novo experimento foi realizado com o intuito de medir o menor intervalo de tempo no qual o aparelho auditivo consegue reconhecer timbre e altura. Quatro séries de durações iguais a 3, 5, 10, 25, 50, e 250 ms, duas para clarineta, e duas para trompete foram executadas. Em aproximadamente 10 ms, começa a surgir a percepção de altura. A partir daí, começa a surgir a sensação de cor, timbre e, posteriormente, de reconhecimento de instrumento. Foi também demonstrado que, com durações abaixo de 6 ms, fragmentos de alturas diferentes concatenados fundem-se, criando uma melodia subliminar. A partir de 10 ms de duração, passaram a formar uma estrutura melódica.

2.3.2 Altura (*Pitch*)

Schaeffer realizou uma filtragem em uma nota grave de piano, deixando apenas os três primeiros harmônicos. O timbre é afetado consideravelmente, porém, o *pitch* e a intensidade não se alteram. Retirando apenas a fundamental, nenhuma diferença muito aparente foi constatada. Contudo, cortando-se a fundamental de uma nota de altura média, o timbre foi gravemente alterado. E por último, repetindo o procedimento para uma nota aguda, o observado é o oitavamento da nota. Portanto, a altura não está somente associada com a frequência fundamental.

Neste mesmo tema, Schaeffer realiza outro conjunto de experimentos. Filtragens com diferentes frequências centrais foram feitas no ruído branco, sucessivamente, resultando em uma percepção de sequência melódica. A mesma filtragem foi feita em um som complexo estruturado. Neste caso o timbre é alterado mas o som não evolui em tessitura.

2.3.3 Intensidade

Para sons sustentados, nos quais o conteúdo harmônico não varia consideravelmente no decorrer da nota, o ouvido é mais sensível à variação de energia no tempo. Em sons percussivos seguidos de ressonância, o ouvido é mais sensível a como a energia desaparece do que como ela aparece. Outra consideração é que, sons do tipo percussão-ressonância, têm a caracterização do timbre no momento do ataque. Já em sons sustentados, o ataque tem um papel secundário na caracterização do timbre. Nestes casos, o timbre é o resultado da combinação da percepção do ataque com a percepção do restante da nota.

Schaeffer mostra que suprimindo-se até um segundo do início de uma nota grave de piano, a nota permanece quase sem alteração no timbre. Paradoxalmente, quando o ataque de um som de sino soando é recortado, a alteração na percepção é bastante alterada. A supressão do ataque em alguns casos altera muito a percepção, e em outros casos não. Uma explicação é o fato da nota grave do piano ser formada por um único som, e uma nota de vibrafone ser formada por dois sons: um choque metálico muito breve, e uma ressonância que depende da construção do instrumento. Portanto, para o caso de um som único, a supressão do ataque não altera muito a percepção, enquanto que, para sons duplos, a retirada do ataque eliminaria uma das componentes do som. Outro experimento descarta a existência de uma correlação na forma de onda entre ataques parecidos. Oito staccatos de trompete foram gravados. Suas formas de onda foram comparadas com o intuito de achar algum padrão, sem nenhum sucesso. Porém, a diferença na forma da curva de dinâmica é um fator importante.

Os primeiros 50 ms foram cortados de um **réb** de flauta e o observado foi uma diferença sutil para o original. A diferença segundo Schaeffer, é que os primeiros 50 ms fornecem uma espécie de ruído causado pelo sopro no instrumento. Ao realizar a mesma experiência para o trompete, esta diferença sutil não foi encontrada. A explicação fornecida pelo autor é de que, no trompete, o ataque é principalmente influenciado pelo formato da curva dinâmica. Portanto, para reforçar seu argumento, Shaeffer produziu um ataque artificial em uma nota sustentada de trompete, cortando a fita magnética obliquamente, e o resultado foi bastante similar à nota original. O mesmo para o violino pôde ser constatado. Um ataque artificial foi realizado primeiramente com um corte reto e posteriormente com um corte oblíquo, e, ao serem comparados com o original, o corte oblíquo se mostrou bastante similar.

Finalmente, o autor apresenta dois experimentos, nos quais notas de instrumentos diferentes, piano e flauta, são transmutadas entre si apenas pela alteração da curva de dinâmica. No primeiro caso, um **mi4** de piano teve seu ataque recortado na fita magnética e posto a soar simultaneamente com a nota **mi4** da flauta. As duas notas se mostraram bem parecidas. Outra manipulação foi feita, agora com um **fá4** de flauta, cuja forma dinâmica foi modelada por um modulador de envoltória de forma a ficar com a dinâmica semelhante à de um **fá4** de piano. Mais uma vez, as notas foram comparadas e se mostraram semelhantes.

2.3.4 Duração

Shaeffer realizou um experimento em que sons complexos com ataques curtos e ressonâncias longas foram executados. Os mesmos sons foram executados de trás para frente e, curiosamente, na execução invertida, os ataques pareceram bem menores. O trajeto da escuta não é mais realizado nem com a mesma velocidade nem da mesma

maneira. Os ataques foram cortados e executados separadamente, ficando claro que parecem muito menores isolados do que quando tocados no conjunto da nota. A conclusão deste conjunto de experimentos é que existe uma anamorfose tempo-duração. Há uma distorção entre a duração percebida e o tempo físico medido. Os elementos da forma perturbam consideravelmente os valores métricos.

2.3.5 Timbre

Schaeffer apresenta um conjunto de experimentos com o intuito de mostrar que o timbre não é dependente nem somente do espectro harmônico, e nem somente da dinâmica, e sim, de uma associação das duas coisas. Inicialmente oito sons provenientes de flauta, fagote, clarineta, síntese, oboé, trompete, e síntese, tiveram suas curvas dinâmicas alteradas e, com isso, o reconhecimento da fonte produtora do som foi bastante dificultado, isto mostra que a curva de dinâmica possui uma grande importância na composição do timbre. Para demonstrar a importância da composição harmônica, um tom puro (**sol3**), foi modulado com a mesma curva de dinâmica de um **sol3** de piano, e o resultado foi uma nota parecida mas com uma cor diferente. Fazendo a mesma coisa mas agora modulando uma nota de flauta, que possui espectro harmônico bem mais próximo ao do piano, o resultado foi bem mais próximo ao original.

Schaeffer coloca a questão de como reconhecemos um timbre característico de um instrumento, como o piano por exemplo, se cada uma de suas notas possui timbre próprio. Foi constatado que a dinâmica das notas do piano se torna cada vez mais acentuada à medida que aumenta-se a altura. Schaeffer gravou 22 notas de piano sobre as sete oitavas consecutivamente, e podendo perceber claramente a mudança de dinâmica. Para constatar a mudança harmônica, foram gravadas notas graves transpostas pelo gravador para duas oitavas acima, e notas agudas transpostas pelo gravador para duas oitavas abaixo. As notas transpostas foram comparadas com as notas originais. Percebendo-se claramente que notas mais graves possuem um espectro harmônico muito mais rico que notas mais agudas. Schaeffer define então o que ele chama de lei do piano. A lei do piano foi definida como uma lei de compensação entre timbre dinâmico e timbre harmônico. Em uma progressão do grave para o agudo, a inclinação da curva dinâmica cresce constantemente, enquanto que o timbre harmônico decresce proporcionalmente.

A causalidade é colocada como forte influente no discernimento das fontes emissoras. Um experimento foi feito com uma nota de trompete soando com um incidente no início. O incidente causa certa estranheza que o autor chamou de excesso de timbre. Em outro experimento, duas ressonâncias provenientes de uma excitação de chapa de metal e outra proveniente de uma simulação do piano foram comparadas. Quando apenas a parte de ressonância é executada, é muito difícil dizer o que gerou ambos os sons. No momento em que o som por inteiro é tocado, o gerado pelo piano é imediatamente reconhecido.

Segundo Schaeffer, quando o contexto de causalidade intervém, é inserido um novo fator: a psicologia da audição propriamente dita.

2.4 Motivações

Neste capítulo, até aqui, foram elencadas questões físicas, fisiológicas e perceptivas úteis para ajudar a inferir sobre os processos auditivos envolvido na escuta dos ataques e para definição de termos que serão abordados no decorrer da tese. O entendimento de fenômenos psicoacústicos, como o Mascaramento Temporal por exemplo, ajuda na compreensão de fatos perceptivos como a influência do ataque no evento temporalmente seguinte, a sustentação, alterando o timbre da nota, como descrito no **Solfejo dos Objetos Sonoros** citado nos parágrafos anteriores.

A Figura 6 ilustra a principal motivação deste estudo. O espectrograma de um fragmento extraído de uma gravação utilizada neste estudo, delimitado pelo retângulo na partitura do fragmento, resume os três elementos acústicos que observamos na região de ataque: (1) o conteúdo harmônico da reverberação da nota anterior (destaque pelo círculo); (2) o comportamento específico do aumento da magnitude dos harmônicos e das defasagens de início em notas atacadas individuais (destacadas pela elipse); (3) um elemento que não faz parte da estrutura harmônica da nota, nem da reverberação, geralmente tratada como ruído, que pode se originar de teclas de instrumentos, sopro de ar ou qualquer turbulência (destacada por um quadrado). Esses três elementos acústicos apontados influenciam diretamente na percepção auditiva dos ataques, podendo influenciar a escuta das notas e da peça musical como um todo. Um espectrograma simples, como o mostrado na Figura 6, possibilita uma análise qualitativa ou quantitativa de forma manual. A detecção, decomposição e modelagem espectral desses três elementos permite superar dificuldades encontradas em análises quantitativas de gravações de áudio de clarineta, podendo enriquecer o estudo de performances musicais na condição de prática cotidiana, onde ambientes reverberantes estão sempre presentes.

A modelagem e quantificação desses componentes acústicos do ataque das notas de forma isolada permite, também, inferências sobre as proporções das componentes harmônicas, não-harmônicas e interações do músico com a acústica da sala, inferências essas, encontradas ainda bem incipientes na literatura. Como mostrado na seção de introdução desta tese, estudos sobre ataques geralmente criam condições distantes da performance, como análise de notas isoladas e gravações em câmaras anecoicas, visando isolar variáveis. Eles trazem resultados essenciais para a construção do conhecimento. Por outro lado, esse distanciamento desconsidera a interação entre músico e ambiente, pois adaptar a performance à acústica específica de uma sala de concertos faz parte da rotina dos músicos profissionais. Estudos relativamente recentes revelaram resultados

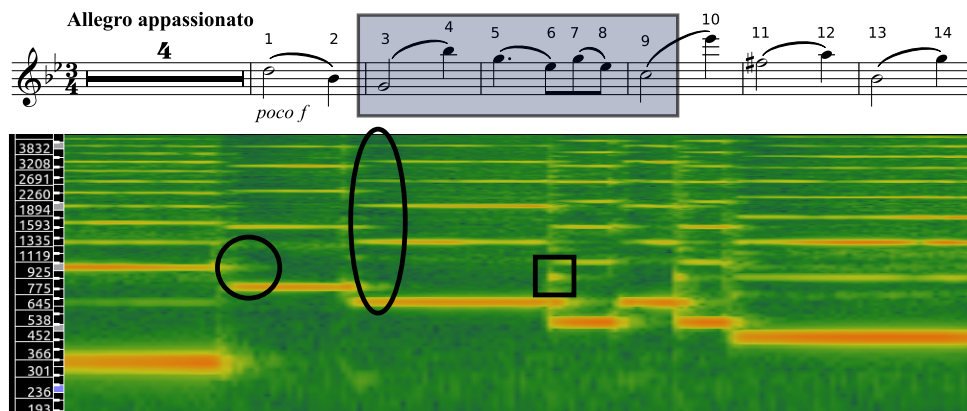


Figura 6 – Espectrograma de um fragmento extraído de uma gravação utilizada nas análises desta tese. Três elementos encontrados nos ataques das notas são destacados.

que corroboram com o fato de que músicos profissionais ajustam conscientemente seu estilo de tocar sob diferentes condições acústicas, com investigações objetivas dos ajustes das performances sob diferentes condições acústicas da sala (KATO; UENO; KAWAI, 2008; KALKANDJIEV; WEINZIERL, 2015; BOLZINGER; WARUSFEL; KAHLE, 1994; FISCHINGER; FRIELER; LOUHIVUORI, 2015). Em Kalkandjiev e Weinzierl (2015), por exemplo, modelos computacionais de 14 salas foram testados por síntese binaural dinâmica com artistas tocando fisicamente em uma câmara anecoica, enquanto usavam fones de ouvido simulando ambientes acústicos virtuais correspondentes. Este experimento revelou conceitos distintos de ajuste às condições acústicas, bem como individualidade significativa com relação à interação dos músicos com a acústica da sala.

Com isso, a motivação principal desta tese pode ser resumida na busca de métodos para tornar as análises objetivas dos ataques mais próximas à condição real de performance.

3 Descrição de Conteúdo Acústico em Performances

No capítulo anterior, foram apresentados conceitos básicos sobre percepção auditiva, importantes para as discussões levantadas. O presente capítulo é sobre extração de informação do material acústico de performance de clarinetistas.

A Expressão de músicos em uma performance engloba aspectos pessoais e interpretativos da obra musical, envolvendo trabalho gestual tanto no âmbito acústico quanto de movimentação corporal. O foco aqui é no conteúdo acústico da informação musical em performances, mais especificamente no conteúdo espectral dos ataques das notas em instrumentos monofônicos. Diversos estudos mostraram a importância do ataque de notas para a percepção, ao mesmo tempo que evidenciam as dificuldades de análise desses eventos acústicos, por apresentarem curta duração temporal, além de características transientes e não lineares. Existe uma vasta quantidade de descritores acústicos na literatura que podem ser usados para descrição e caracterização de conteúdo musical. Na sequência deste capítulo, serão apresentados conceitos e métodos para a descrição de conteúdo acústico em performances musicais envolvidos nesta pesquisa.

3.1 Detecção dos Ataques e Decaimentos

Os métodos de detecção de regiões de transição das notas (Ataques e Decaimentos) utilizam um dos quatro princípios: variações na energia do sinal, variações na magnitude do espectro de frequência, variações na fase do espectro e detecção por modelagem (DAUDET, 2006). São destacados aqui dois métodos: o primeiro baseado em variação de energia do sinal, e o segundo por estabilização do Fluxo Espectral.

No primeiro método, a detecção dos instantes delimitadores Início de Decaimento e Final de Ataque pode ser feita a partir da estimação dos máximos das taxas de variação de energia dentro da nota. O método é baseado na análise do contorno do envelope de energia da nota (MAESTRE; GÓMEZ, 2005). Considerando o envelope de energia como uma função diferenciável contínua no tempo, os pontos de máxima curvatura são considerados variações máximas locais da derivada primeira do sinal, e consequentemente, máximos e mínimos da derivada segunda. Os pontos correspondentes aos valores mínimos da derivada segunda determinam os candidatos para os instantes procurados na curva de energia **RMS**. O candidato escolhido para o Início de Decaimento será aquele que apresentar maior inclinação negativa entre ele e o final da nota, enquanto o candidato escolhido para Final de Ataque será aquele que apresentar maior inclinação positiva entre ele e o início da nota.

No segundo método, é utilizado o Fluxo Espectral para a detecção da região de transição das notas. O Final de Ataque é detectado por limiar de estabilização do Fluxo Espectral a partir do *Onset* (do *Onset* para frente no tempo). O Início de Decaimento é detectado por limiar de estabilização do Fluxo Espectral a partir do *Offset* (do *Offset* para trás no tempo).

3.2 Índice de Legato

Notas ligadas na clarineta são produzidas por meio de um único sopro, sem interromper o fluxo de ar durante a transição. Uma forma de investigar como esse tipo de transição é executada é através do cálculo do Índice de Legato (IL), apresentado em [Maestre e Gómez \(2005\)](#). Esse Índice está relacionado com a manipulação da energia do som durante a transição. Se baseia na comparação entre a envoltória de energia do sinal analisado, e uma linha reta traçada entre o início do decaimento e o final do ataque da nota seguinte. É calculado como descrito na equação 3.1.

$$IL = 1 - \frac{A1}{A1 + A2} = \frac{A2}{A1 + A2} \quad (3.1)$$

Os termos A1 e A2 presentes na equação 3.1, são áreas definidas como demonstrado na Figura 7. A interpretação para seus valores é feita da seguinte forma: Valores iguais a zero, significaria uma transição em que a energia do som estaria o mais longe possível da curva $r(t)$ que representa um legato de referência sem queda nenhuma na energia. Valor igual a um representa energia na transição igual à da referência $r(t)$.

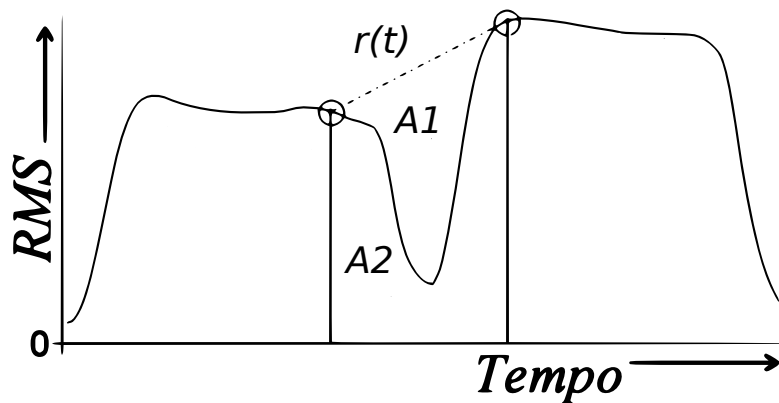


Figura 7 – Áreas usadas para o cálculo do Índice de Legato.

O Índice de Legato descrito acima é uma ferramenta para caracterizar a energia do sinal na transição. Com as técnicas de decomposição de sinais em partes determinística, estocástica e transiente, é possível estender a avaliação do Índice de Legato investigando

a contribuição individual de cada uma destas componentes em relação à energia total da transição. A motivação para avaliar cada componente separadamente vem do fato de serem sinais de comportamento físico distintos e, principalmente, por serem percebidos de forma diferente. Exemplos de sinais com essas diferenças são artefatos como barulho de chave, ruídos de aspiração e ruído de unha na corda.

3.3 Fluxo Espectral

O Fluxo Espectral é definido pela correlação da magnitude do espectro de frequência entre quadros consecutivos (MCADAMS, 1999), conforme Equação 3.2. Seu valor tende a aumentar em regiões de pouca variação na evolução temporal do espectro, e estabilizar-se somente em regiões de sustentação das notas. Sua variação está, portanto, geralmente associada à mudança de nota, o que permite utilizá-lo para identificação das regiões de transição entre elas. Nesta pesquisa, a região do ataque é considerada como o trecho do sinal contido entre o início da nota e o ponto onde o complemento de um (subtraído da unidade) do Fluxo Espectral deixa de ser superior à média do fluxo ao longo de toda a nota.

O Fluxo Espectral é calculado como descrito na Equação 3.2:

$$F(q) = \frac{1}{M} \sum_{p=1}^M |r(X(p)_q, X(p)_{q-1})| \quad , \quad (3.2)$$

e seu complemento como na Equação 3.3:

$$\overline{F(q)} = 1 - \frac{1}{M} \sum_{p=1}^M |r(X(p)_q, X(p)_{q-1})| \quad , \quad (3.3)$$

onde $F(q)$ é o valor do Fluxo Espectral no quadro q , M o tamanho do quadro, r o coeficiente de correlação, e X o Espectro de frequência do quadro em questão.

O valor do Fluxo Espectral tende a aumentar em regiões com pouca variação na evolução temporal do espectro. Portanto, sua variação está geralmente associada à mudança de nota. Assim, a estabilização do Fluxo Espectral acontece somente em regiões de sustentação das notas. Isso permite a utilização do Fluxo Espectral para detecção de região de transição de notas.

4 Modelagem e Decomposição Espectral

Existem muitos métodos para modelagem e decomposição espectral de sinais de áudio. Eles geralmente são inspirados no processo auditivo, podem ser baseados em *Short-Time Fourier Transform* (**STFT**) como *Phase Vocoder* e *Spectral Modeling Synthesis* (**SMS**), ou em outras formas de representação como o *Linear Predictive Coding* (**LPC**), *Frequency-Modulation Synthesis* (síntese **FM**), e Síntese Aditiva ([SMITH, 2011](#)). Nesta tese, foi utilizado o **SMS**. A comparação entre os diferentes métodos para esse fim foge do escopo desta pesquisa. A seguir serão descritas as etapas de modelagem, síntese e decomposição do SMS.

4.1 Spectral Modeling Synthesis (SMS)

O *Spectral Modeling Synthesis* (**SMS**) é uma abordagem de modelagem espectral. O princípio básico do **SMS** é a representação do sinal através de uma soma de senoides quase-estacionárias, modelando assim uma componente determinística ([SERRA; SMITH, 1990](#), p.60). Considerando um sinal $s(t)$, sua representação é, portanto, da forma (Equação 4.1):

$$s(t) = \sum_{n=1}^N A_n(t) \cos[\theta_n(t)] + e(t) \quad (4.1)$$

onde N é o número de componentes senoidais, $e(t)$ é o resíduo (diferença entre o sinal e a componente determinística), $A_n(t)$ e $\theta_n(t)$ são a amplitude instantânea e a fase instantânea das componentes senoidais. A fase instantânea é definida como (Equação 4.2):

$$\theta_n(t) = \int_0^t \omega_n(\tau) d\tau + \theta_n(0) + \phi_n \quad (4.2)$$

onde $\omega_n(t)$ é a frequência instantânea, $\theta_n(0)$ o valor inicial da fase e ϕ_n é um deslocamento fixo de fase.

Para a detecção dos valores dos parâmetros das componentes senoidais, são necessárias três etapas. A primeira é a análise em frequência dos quadros de curta duração do sinal, ou *Short Time Fourier Transform* (**STFT**). A segunda etapa é a detecção de picos do espectro, de acordo com algum critério. E, por último, a síntese utilizando os parâmetros extraídos.

O **SMS** modela a parte determinística dos sinais, e retorna como resultado uma saída do tipo Modelo + Resíduo. Este resíduo contém componentes de ruídos, transientes e erro de modelagem somados. O SMS modela também a parte estocástica (ruído), extraindo

a envoltória do espectro do resíduo por decimação. Porém, como descrito em [Serra e Smith \(1990, p.59\)](#), a modelagem estocástica do resíduo não é eficiente para transientes.

Representação em Frequência

Na etapa de análise do **SMS**, o sinal de entrada é dividido em quadros de análise de curta duração. Com isso, a Transformada de Fourier de cada quadro possibilita uma representação da evolução do comportamento do espectro de frequências ao longo do tempo. Essa técnica é chamada de *Short Term Fourier Transform*. Cada quadro é então multiplicado por uma janela de análise espectral, para melhorar a discriminação das frequências no espectro.

Detecção de Picos

A detecção de picos no espectro é feita a partir da comparação de cada *bin* com os respectivos anteriores e posteriores. Se esse *bin* for maior do que seus dois adjacentes, é considerado um máximo local e, conseqüentemente um pico no espectro.

Interpolação de Picos

No processo de análise do **SMS**, a magnitude do espectro de frequências é representado na escala de decibels. Nesse caso, por ser uma escala logarítmica, os lóbulos centrais das janelas de análises podem ser representados por parábolas, visto que na escala linear, os lóbulos centrais podem ser aproximados por gaussianas, que se tornam parábolas em escalas logarítmicas. Visando um refinamento na detecção de picos no espectro, os valores encontrados são interpolados junto aos seus *bins* adjacentes, encontrando com isso, o valor calculado de máximo da parábola e o respectivo valor interpolado de fase.

Detecção de Frequência Fundamental

O SMS utiliza o método *Two Way Mismatch* - **TWM** para detecção da Frequência Fundamental ([MAHER; BEAUCHAMP, 1994](#)). O **TWM** foi desenvolvido para sons monofônicos e tem o intuito de ser robusto a ruídos e reverberações. Para tanto, é aplicada a **STFT** utilizando uma janela de tamanho fixo. A cada salto, normalmente com sobreposição da janela, obtém-se as magnitudes dos picos espectrais (harmônicos candidatos). Para estimar a F_0 , é feita a comparação entre os candidatos a harmônicos obtidos com a aplicação da transformada de Fourier, e as sequências previstas de harmônicos desses candidatos. A partir dessa comparação, é calculado o valor de discrepância entre os harmônicos parciais e os harmônicos previstos, utilizando o método **TWM**.

Peak Continuation

No sistema desenvolvido por Serra ([SERRA; SMITH, 1990](#)), o processo definido como *peak continuation* no caso de sons em geral, define guias e trajetórias a partir dos picos encontrados no espectro de frequência do sinal. Guias e trajetórias são duas definições criadas para o processo de busca dos caminhos das parciais (componentes determinísticas senoidais não necessariamente harmônicas) do som ao longo do tempo no espectro de frequências. Enquanto as guias são sugestões iniciais do local do espectro onde essas parciais estão passando (usando informações iniciais a respeito dos sons), as trajetórias são o resultado final do processo de busca de picos no espectro que mais se aproximam dessas guias. No caso da definição a priori de sons quasi-harmônicos, as guias são definidas através da detecção da frequência fundamental usando o método *Two Way Mismatch* ([MAHER; BEAUCHAMP, 1994](#)), descrito na seção anterior 4.1. Para cada quadro de análise, a frequência fundamental é encontrada e os múltiplos harmônicos são definidos como as guias. A partir dessas guias, picos no espectro são rastreados de acordo com parâmetro de restrições passados pelos usuários, delineando assim, as trajetórias das parciais harmônicas (ou quase-harmônicas usando maior rigor).

Subtração Espectral

Depois de estimados os máximos das parábolas referentes a cada parcial do espectro, o lóbulo central da janela de análise utilizada é calculado e subtraído do espectro analisado. Com a subtração, tem-se portanto dois sinais: a componente harmônica formada pelos lóbulos centrais das parciais do espectro, e o resíduo resultante da subtração.

Análise e Representação Estocástica (Ruído)

A componente estocástica (ruído) do resíduo é modelada com a extração da envoltória do espectro do resíduo por decimação. Com isso, a componente estocástica é ressintetizada pela aplicação dessa envoltória a um ruído branco sintético. A presença de componentes harmônicos nessa etapa de modelagem atrapalha a estimação da envoltória. Assim, é recomendável aplicar esta etapa apenas quando o sinal estiver livre das componentes harmônicas. Os transientes não são bem modelados por esta etapa, pois esses eventos se espalham no espectro por sua característica impulsiva no tempo, se misturando à componente estocástica.

Ressíntese: Soma Sobreposta

A ressíntese da componente harmônica e do resíduo é feita utilizando o método de superposição de quadros *Overlap Add* (**OLA**). O espectro reconstruído a partir da

resposta espectral da janela de análise utilizada é transformado para o tempo (**IFFT**) e, em seguida, aplicada a superposição de quadros ou Overlap Add (OLA).

5 Metodologia Proposta

Neste capítulo é apresentada a metodologia de análise proposta nesta tese. Lembrando que o objetivo é analisar os ataques e as transições entre notas de gravações de áudio em situações próximas à de concerto, com execução de trechos de peças de repertório em ambiente com reverberação presente e controlada.

Em uma visão geral, pode-se dividir a metodologia proposta em três etapas: Segmentação do áudio, modelagem e decomposição espectral e, por último, caracterização das transições entre notas e dos ataques. A última etapa, por sua vez, possui três ferramentas para caracterização: (i) o Índice de Legato, avaliando-se a proporção da componente harmônica nas transições, chamado de Índice de Legato Harmônico (**ILh**), (ii) o plano **H2R2** para os ataques e (iii) o Perfil Harmônico do Ataque (**PHA**).

Em princípio, os trechos selecionados para análise são segmentados em notas e, em seguida, detectadas as regiões de transição. Com a transição detectada, é aplicada nessa região a técnica de modelagem espectral *Spectral Modeling Synthesis* (SMS), capaz de decompor o sinal em uma parte harmônica e outra parte residual, resultante da subtração espectral da primeira. Na parte residual dos ataques com outras notas anteriores próximas, uma segunda componente harmônica referente à reverberação é também modelada pelo **SMS**. Realizada a decomposição espectral do sinal, a proporção de energia de cada componente é então avaliada por meio de testes estatísticos para verificação de efeitos significantes dos fatores de interesse envolvidos. A diferença entre o Índice de Legato (**IL**) e o Índice de Legato Harmônico (**ILh**) é também testada. Além da caracterização dos ataques pela avaliação da proporção de energia de suas componentes, é proposta nesta tese uma curva nomeada de **Perfil Harmônico do Ataque (PHA)**, buscando uma forma de parametrização do comportamento harmônico da região, mesmo com a presença de reverberação.

5.1 Segmentação

O processo de detecção do ataque e da região de transição entre as notas depende da segmentação do sinal. A segmentação é a detecção dos instantes delimitadores de eventos de interesse que, no caso desta pesquisa são: Início de Nota (*Onsets*), Final de Nota (*Offsets*), Início de Decaimento, e Final de Ataque. A segmentação dos áudios gravados foi realizada com o auxílio do software EXPAN ([CAMPOLINA; LOUREIRO; MOTA, 2009](#)), desenvolvido pelo CEGeME, mesmo grupo de pesquisa onde foi realizada a coleta dos dados.

O diagrama da Figura 8 ilustra o processo de segmentação utilizado. O processo se inicia pelo janelamento do sinal, e a partir daí são calculadas as curvas de energia **RMS** (*Root Mean Square*), Altura, e Fluxo Espectral (as duas últimas depois de aplicada a **FFT** em cada janela de análise). A partir das curvas de **RMS** e de Altura, são detectados os *Onsets* e os *Offsets*. O Final de Ataque e Início de Decaimento são detectados por limiar de estabilização do Fluxo Espectral dentro de uma mesma nota.

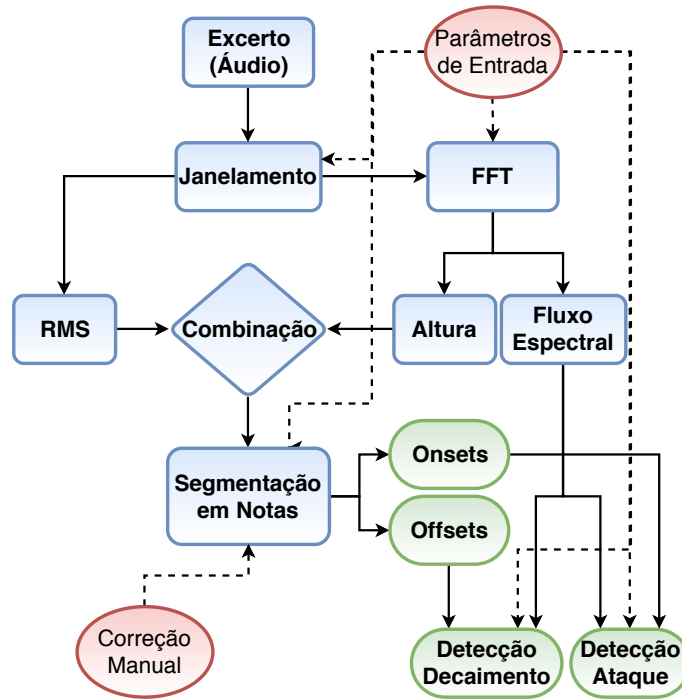


Figura 8 – Processo de segmentação utilizado. Detecção dos instantes delimitadores: Início de Nota (*Onsets*), Final de Nota (*Offsets*), Início de Decaimento, e Final de Ataque.

5.2 Detecção dos Ataques e Decaimentos

Foi utilizado nesta tese o método de detecção da região de transição das notas pelo Fluxo Espectral (detalhado na Seção 3.3). Dentro de uma nota (entre *Onset* e *Offset*), o Final de Ataque e Início de Decaimento são detectados por estabilização do Fluxo Espectral. Essa estabilização ocorre quando o complemento do Fluxo Espectral encontra-se abaixo de um limiar pré-definido. O Final de Ataque é então estabelecido como sendo o primeiro instante dessa região de estabilização, enquanto o Início de Decaimento é definido como o último instante da mesma.

5.3 Decomposição Espectral

Detectadas as regiões de transição e os ataques, o espectro dessas regiões é analisado a partir da modelagem *SMS*. Para a análise, o sinal é decomposto de acordo com o diagrama da Figura 9.

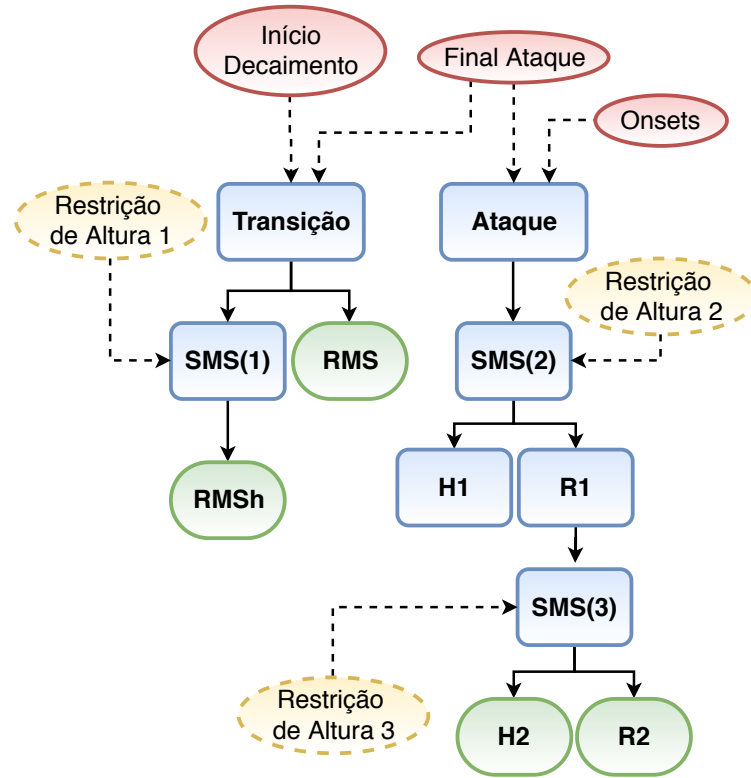


Figura 9 – Decomposição espectral do sinal: cálculo das curvas de energia **RMS** da componente harmônica e do sinal original da transição; Definição das proporções de energia **H1**, **R1**, **H2**, **R2** das componentes extraídas do ataque.

A sigla **RMSH** representa a curva de energia **RMS** calculada utilizando apenas a componente harmônica do sinal (extraída com **SMS(1)**) da transição entre as notas. Por outro lado, a sigla **RMS** representa a curva de energia do sinal original. Ainda no diagrama da Figura 9, **H1** e **R1** representam a proporção de energia em dB da componente harmônica e do resíduo da subtração espectral de uma primeira modelagem do ataque (**SMS(2)**), respectivamente, em relação ao sinal original. A partir de **R1**, é realizada uma nova modelagem (**SMS(3)**) resultando em **H2** e **R2**. Por sua vez, **H2** representa a proporção de energia da componente harmônica relativa à reverberação da nota anterior, depois de extraída a componente harmônica da nota em questão. Assim, **R2** representa um segundo resíduo, depois de extraídos **H1** e **H2**. As proporções **H2** e **R2** são também relativas ao sinal original, em dB. As modelagens em **SMS(1)**, **SMS(2)** e **SMS(3)** se diferenciam pelo parâmetro de entrada que define a restrição de F0 procurada. Enquanto **SMS(1)** se restringe a variação de procura de F0 na faixa das duas notas envolvidas na transição, **SMS(2)** se restringe a variação de F0 da nota específica em que o ataque está

inserido, com uma tolerância de 10%. Por fim, **SMS(3)** se restringe a variação de procura de F0 da nota anterior em que o ataque está inserido. Na seção seguinte, as modelagens e os parâmetros de ajuste serão detalhados.

5.4 Modelagens e Parâmetros de Ajuste

Nesta seção, serão listadas as modelagens e parâmetros de ajuste utilizados, para elucidação da metodologia de análise proposta.

Siglas do Diagrama da Figura 8:

Janelamento - O janelamento no processo de segmentação é aplicado no sinal de áudio para determinação das curvas de **RMS**, **Altura** e **Fluxo Espectral** (as duas últimas depois de aplicada a FFT). **Janela utilizada:** de *hamming*. **Tamanho da Janela:** 1024 amostras (23,2 ms a 44,1 kHz). **Salto:** 256 amostras (5,8 ms a 44,1 kHz).

FFT - Transformada de Fourier aplicada no sinal janelado. **Tamanho:** 2^{12} .

Altura - A curva de altura é calculada a partir do janelamento definido acima. Para cada janela de análise, a F0 é estimada pela escolha do máximo local do espectro que apresenta maior somatório de energia em sua série harmônica. **Número de Candidatos:** 7. **Tamanho da Série Harmônica por candidato:** 13. **Tamanho do Filtro de Moda aplicado:** 5 quadros.

Fluxo Espectral - Curva definida pelo cálculo do coeficiente de correlação do espectro de frequências do quadro atual com o espectro do quadro imediatamente anterior.

RMS - Curva de energia quadrática média calculada a partir do janelamento definido acima. Para cada janela de análise, é calculada a raiz quadrada das médias dos quadrados.

Combinação - Essa etapa no diagrama representa a combinação das curvas de **RMS** e de **Altura** para a segmentação das notas. Notas consecutivas de alturas iguais são detectadas por variação na curva **RMS**, e notas consecutivas de alturas diferentes são detectadas por variação na curva de altura.

Segmentação em Notas - A detecção de *Onsets* por energia é feita pelo cruzamento da curva **RMS** por uma curva média do próprio **RMS**. **Tamanho da Janela RMS Médio:** 1 s. A detecção de *Onsets* por frequência é feita pela variação na curva de altura. **Limiar de Variação de Altura:** 6 %.

Detecção Decaimento - Pontos correspondentes aos valores mínimos da derivada segunda do **RMS**, com maior inclinação negativa em relação ao *Offset*.

Detecção Ataque - O Final de Ataque é detectado por limiar de estabilização do Fluxo Espectral a partir do *Onset* (do *Onset* para frente no tempo). **Limiar de estabilização:** 10 % da média da nota.

Correção Manual - A correção manual é feita utilizando as curvas de **RMS**, **Altura** e **Fluxo Espectral**.

Siglas do Diagrama da Figura 9:

SMS(1) - Modelagem **SMS** aplicada na região de transição entre duas notas: nota atual e nota seguinte. **Instantes delimitadores:** Início de Decaimento da nota atual e Final de Ataque da nota seguinte. **Janela de Análise:** **blackmanharris** de tamanho 1024. **Resolução da FFT:** $N = 2^{12}$. **Limiar de detecção de energia:** -160 dB. **Número de Harmônicos:** 40. **Erro máximo para detecção de fundamental (algoritmo TWM):** 10. **Desvio máximo de harmônicos:** 6%. **F0 mínimo:** F0 da nota atual, estimado pelo EXPAN, com tolerância de 5% abaixo. **F0 máximo:** F0 da nota seguinte, estimado pelo EXPAN, com tolerância de 5% acima. **F0 mínimo** e **F0 máximo** para este caso, compõem a **Restrição de Altura 1** do diagrama da Figura 9.

SMS(2) - Modelagem **SMS** aplicada na região de ataque de uma nota. **Instantes delimitadores:** Início da nota e Final de Ataque da mesma nota. **F0 mínimo:** F0 da nota atual, estimado pelo EXPAN, com tolerância de 5% abaixo. **F0 máximo:** F0 da nota atual, estimado pelo EXPAN, com tolerância de 5% acima. O restante dos parâmetros é o mesmo da modelagem **SMS(1)**. **F0 mínimo** e **F0 máximo** para este caso, compõem a **Restrição de Altura 2** do diagrama da Figura 9.

SMS(3) - Modelagem **SMS** aplicada no resíduo de **SMS(2)**, ajustada para procurar F0 da nota anterior. O objetivo dessa etapa é modelar a parcela harmônica da reverberação da nota anterior contida no ataque de notas ligadas. **Instantes delimitadores:** Início da nota e Final de Ataque da mesma nota. **F0 mínimo:** F0 da nota anterior, estimado pelo EXPAN, com tolerância de 5% abaixo. **F0 máximo:** F0 da nota anterior, estimado pelo EXPAN, com tolerância de 5% acima. O restante dos parâmetros é o mesmo da modelagem **SMS(1)**. **F0 mínimo** e **F0 máximo** para este caso, compõem a **Restrição de Altura 3** do diagrama da Figura 9.

RMS - Curva de energia **RMS** (*Root Mean Square*). Mesmos parâmetros de janelamento da curva de **RMS** do diagrama da Figura 8. **Instantes delimitadores:** Início de Decaimento da nota atual e Final de Ataque da nota seguinte.

RMS_h - Curva de energia **RMS** (*Root Mean Square*), calculada a partir da componente harmônica retornada por **SMS(1)**. Mesmos parâmetros de janelamento da curva de

RMS do diagrama da Figura 8. **Instantes delimitadores:** Início de Decaimento da nota atual e Final de Ataque da nota seguinte.

- H1** - Representa a proporção de energia, em dB, da componente harmônica de uma primeira modelagem do ataque (**SMS(2)**), em relação ao sinal original como referência.
- H2** - Representa a proporção de energia, em dB, da componente harmônica da segunda modelagem do ataque (**SMS(3)**), em relação ao sinal original como referência.
- Restrição de Altura 2** ajustada para procurar F0 da nota anterior, visando modelar a componente harmônica da reverberação da nota anterior.
- R1** - Representa a proporção de energia, em dB, do resíduo da subtração espectral da primeira modelagem harmônica do ataque (**SMS(2)**), em relação ao sinal original como referência.
- R2** - Representa a proporção de energia, em dB, do resíduo da subtração espectral da segunda modelagem harmônica do ataque (**SMS(3)**), em relação ao sinal original como referência.

5.5 Índice de Legato

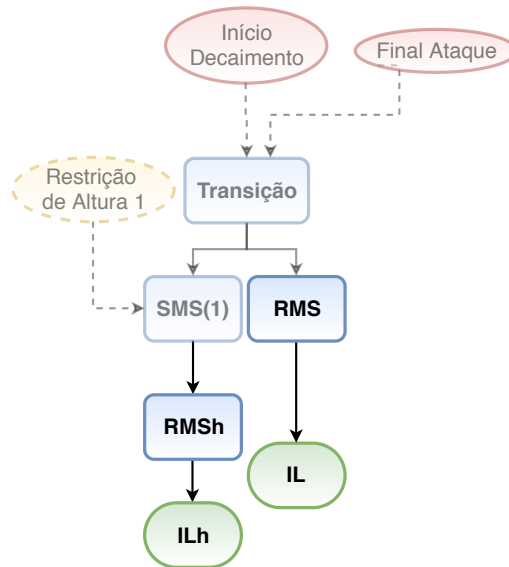


Figura 10 – Diagrama representando o Índice de Legato calculado a partir da curva de energia **RMS** da parte harmônica (**ILh**), e o cálculo do Índice de Legato tradicional (**IL**).

Nesta seção, é definida a primeira das três ferramentas para caracterização utilizada nesta tese. Neste caso, o alvo é as transições em condição de *Legato*. A sigla **IL** no diagrama da Figura 10 representa o Índice de Legato como descrito na Seção 3.2 (calculado a partir da curva de energia **RMS** do sinal original). Por outro lado, a sigla **ILh** representa o

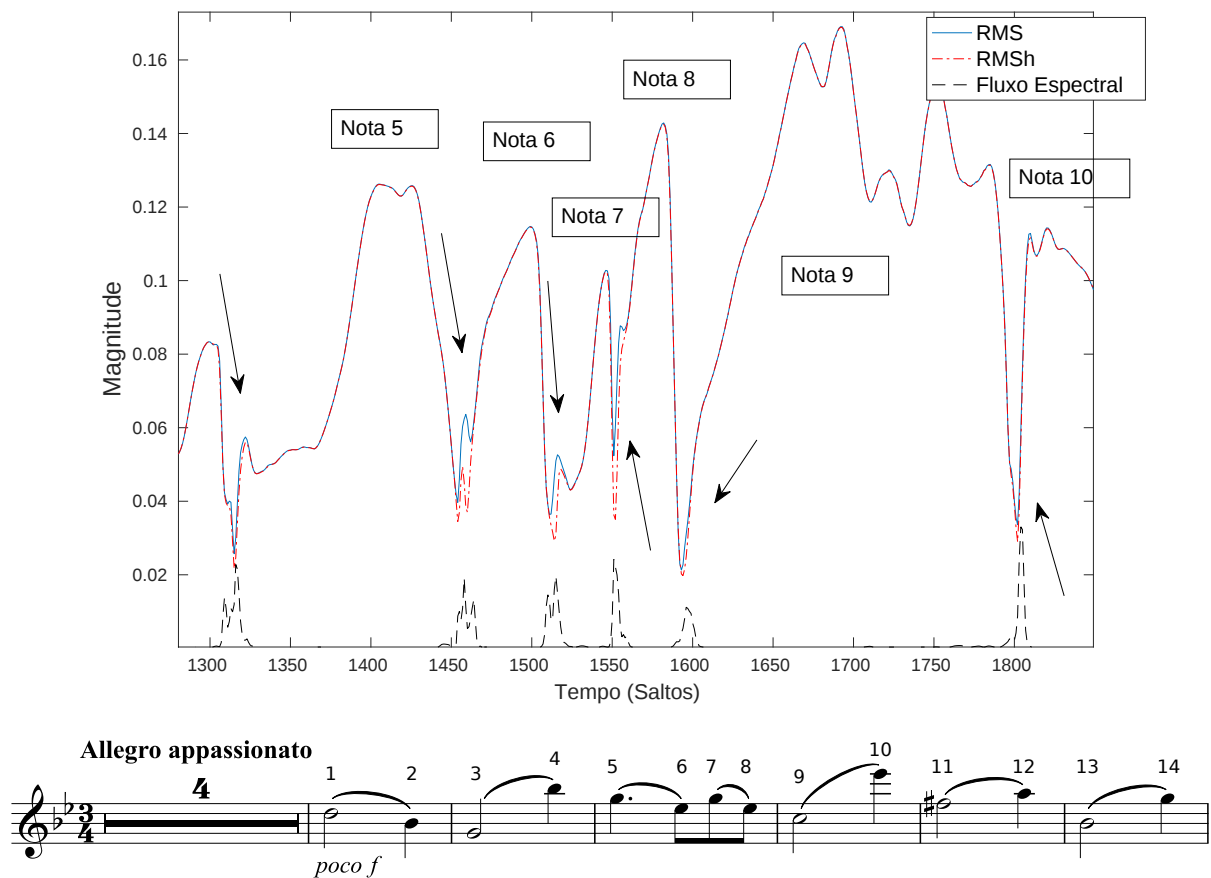


Figura 11 – Confronto entre as curvas de energia calculada utilizando apenas a componente harmônica do sinal (**RMSh**), utilizando o sinal completo (**RMS**) e o complemento de um do Fluxo Espectral. As setas apontam para as regiões de transição de seis notas de uma das gravações utilizadas nos resultados (detalhes das gravações no Capítulo 6). Os saltos do eixo tempo são de 5,81 ms.

Índice de Legato calculado a partir da curva de energia apenas para a parte harmônica, representada por **RMSh**. Os resultados da comparação entre esses dois índices são mostrados no Capítulo 6. Na Figura 11, estão confrontadas as curvas de energia **RMS** e **RMSh** sobrepostas, além do Fluxo Espectral. Nessa mesma figura, as setas apontam para as regiões de transição de seis notas de uma das gravações utilizadas nos resultados (detalhes das gravações, do trecho e obra utilizada estão descritos no Capítulo 6). Ainda sobre a Figura 11, as transições coincidem com os picos do Fluxo Espectral e, como pode ser verificado visualmente, nessas regiões são encontradas as maiores diferenças entre as duas curvas de envoltória de energia.

5.6 Plano H2R2

A segunda ferramenta para caracterização proposta é descrita nesta seção e, neste caso, o objetivo é descrever o conteúdo dos ataques monofônicos, levando em conta a reverberação da nota anterior como uma variável de interesse. A ideia é confrontar em duas dimensões as componentes **H2** e **R2** para caracterizar os ataques, visando facilitar a comparação entre o conteúdo de energia devido à reverberação da nota anterior com a energia harmônica da nota que está sendo atacada. Considerando a energia total do sinal no ataque (S_E) como a soma das componentes:

$$S_E = H1 + R1, \quad (5.1)$$

e

$$R1 = H2 + R2, \quad (5.2)$$

onde $H1$, $H2$ e $R1$ e $R2$ são as proporções de energia definidas na Seção 5.4. Substituindo a equação 5.2 em 5.1 temos:

$$S_E = H1 + H2 + R2. \quad (5.3)$$

O interesse é definir um plano confrontando **H2** (Equação 5.4) x **R2** (Equação 5.5), com as duas componentes normalizadas em relação à energia total (S_E) e representadas em decibel. Podemos escrever:

$$H2(dB) = 10 \log_{10}\left(\frac{H2}{S_E}\right), \quad (5.4)$$

$$R2(dB) = 10 \log_{10}\left(\frac{R2}{S_E}\right). \quad (5.5)$$

Com essa representação, temos valores negativos para as variáveis **R2(dB)** e **H2(dB)** limitados em $0dB$ (zero decibel). Valores altos (próximos de $0dB$) para **R2(dB)** e/ou **H2(dB)**, correspondem a um valor baixo de energia harmônica no ataque da nota atual **H1**. A Figura 12 contém um exemplo do plano proposto, onde estão mostrados 100 pontos simulando o limite superior para as duas variáveis confrontadas, com **R2(dB)** variando entre $0dB$ (igual a soma total S_E) e $-20dB$ alternando os valores de acordo com: **R2(dB)** = $S_E - H2(dB)$, e a componente **H1** sem energia.

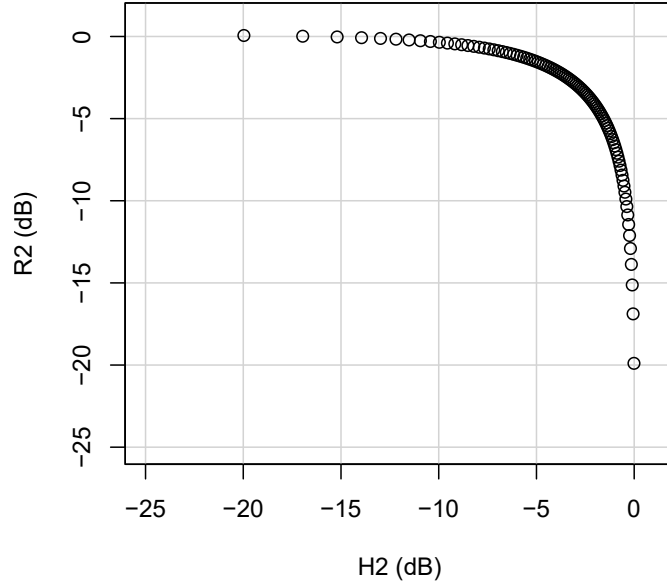


Figura 12 – Plano **H2** x **R2**: 100 pontos simulando o limite superior, com **R2(dB)** assumindo valores entre $0dB$ (igual a soma total S_E) e $-20dB$, variando de acordo com: $\mathbf{R2(dB)} = S_E - \mathbf{H2(dB)}$, e a componente **H1** sem energia.

5.7 Perfil Harmônico do Ataque

O final do ataque é o instante no qual os harmônicos se estabilizam após o início da nota. Com o intuito de caracterizar o comportamento harmônico do ataque das notas, é proposta uma heurística para traçar um perfil harmônico desses eventos delimitados pelo início da nota e o final do ataque, chamado de **Perfil Harmônico do Ataque (PHA)**. O traçado de um perfil estabelece uma parametrização e possibilita a análise a partir da comparação entre os fatores de interesse.

O **PHA** é definido, portanto, tomando o final do ataque como referência. A partir desse ponto de estabilização, para cada harmônico é encontrado um instante de início e, exatamente por esses inícios, é por onde passa o perfil. O **Início do Harmônico** foi definido como o último instante cuja sua magnitude supera 50% da faixa de variação em dB, faixa esta avaliada entre o final do ataque e o mínimo de magnitude do harmônico em questão. O ponto inicial do perfil foi definido como o primeiro instante da $F0$, ou seja, do *Onset* da nota tomado na $F0$. A sequência do **PHA** passa então por cada instante de Início do Harmônico, na sequência crescente, a partir de $F0$. A Figura 13 ilustra dois casos, onde harmônicos na região do ataque de duas notas com seus **PHAs** representados por linhas tracejadas (metade superior da figura), e seus respectivos perfis isolados (metade inferior da figura).

A heurística para determinação do **PHA** foi pensada de maneira a evitar a influência da reverberação da nota anterior. Essa influência fica mais relevante quando acontecem casos de notas com harmônicos coincidentes. Para ilustrar a situação, a Figura 14 contém dois

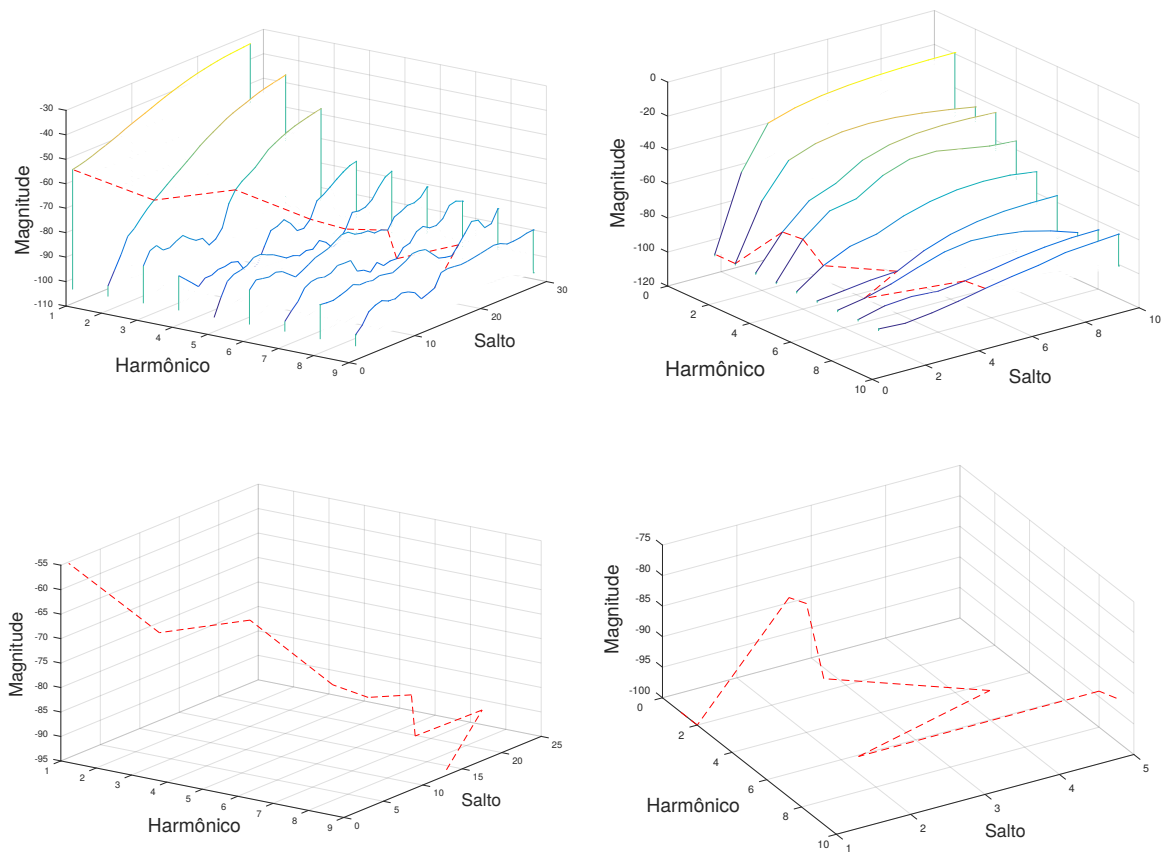


Figura 13 – Dois exemplos de **PHA**: harmônicos na região do ataque de duas notas com seus **PHAs** representados por linhas tracejadas (metade superior da figura), e seus respectivos perfis isolados (metade inferior da figura).

casos de ataques com harmônicos que decrescem antes de terem suas energias aumentadas. Assim, o **PHA** define uma trajetória tridimensional que envolve três variáveis: tempo (instante de início ou inflexão dos harmônicos), magnitude (intensidade de cada início ou inflexão) e número do harmônico.

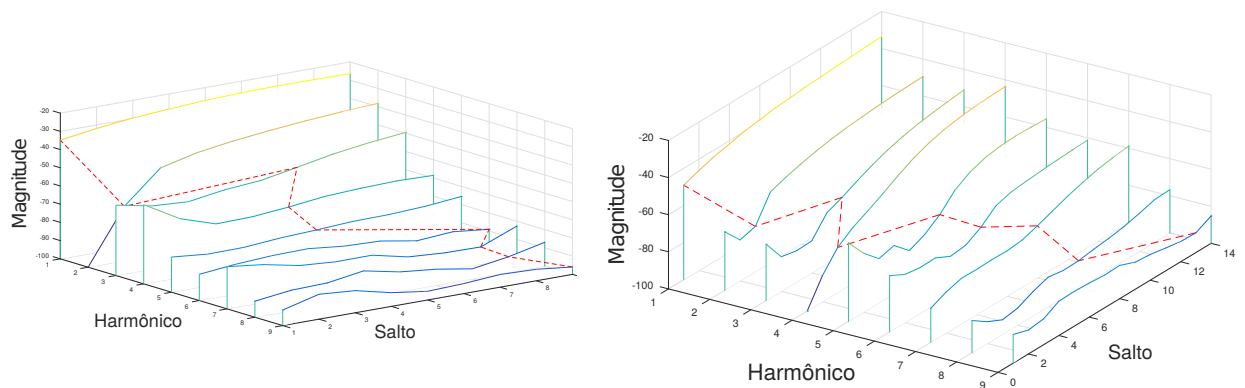


Figura 14 – Exemplos de **PHA**: dois casos de ataques com harmônicos que decrescem antes de terem suas energias aumentadas

Para a verificação da relevância **PHA**, foi avaliada sua capacidade de classificação por músicos e, sendo este um perfil tridimensional, a verificação foi feita nas projeções Harmônico-Magnitude e Harmônico-Tempo, separadamente. A capacidade de classificação foi avaliada através de comparações utilizando Coeficiente de Correlação de *Pearson* entre pares de perfis em duas situações: *Intra* e *Inter* classes.

O Coeficiente (ρ) de Correlação de *Pearson* (PEARSON, 1895) entre duas variáveis X e Y é definido conforme a equação:

$$\rho = \frac{cov(X, Y)}{\sqrt{var(X).var(Y)}} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (5.6)$$

onde $cov(X, Y)$ é a covariância entre X e Y, $var(X)$ e $var(Y)$ são as variâncias de X e Y, x_i e y_i são os i – ésimos elementos de X e Y, sendo $\sum_{i=1}^n$ o símbolo para somatória dos n elementos. Este coeficiente, assume apenas valores entre -1 e 1, sendo que:

- $\rho = 1$ significa uma correlação positiva (ou direta) perfeita entre as duas variáveis X e Y. Isto é, se uma aumenta, a outra sempre aumenta e, se uma diminui, a outra sempre diminui.
- $\rho = -1$ significa uma correlação negativa (ou inversa) perfeita entre as duas variáveis. Isto é, se uma aumenta, a outra sempre diminui e, se uma diminui, a outra sempre aumenta.
- $\rho = 0$ significa que as duas variáveis não dependem linearmente uma da outra. No entanto, pode existir uma dependência não linear. Assim, o resultado $\rho = 0$ deve ser investigado por outros meios.

Um valor de ρ igual ou maior a 0,9, positivo ou negativo, indica correlação muito forte (direta ou inversa). Valores entre 0,7 e 0,9, positivos ou negativos, indicam correlação forte. Valores entre 0,5 e 0,7, positivos ou negativos, indicam correlação moderada. Valores entre 0,3 e 0,5, positivos ou negativos, indicam correlação fraca. Por fim, valores entre 0 a 0,3 positivo ou negativo indica uma correlação baixa.

Além do Coeficiente de Correlação, foi verificada a existência de diferença significativa entre as médias dos grupos *Intra* e *Inter* classes através do teste estatístico (**ANOVA** - *Analysis of Variance*).

6 Resultados e Discussões

6.1 Material

Foram feitos testes com um conjunto de gravações utilizadas no estudo [Teixeira et al. \(2015\)](#), realizadas no Centro de Estudos do Gesto Musical e Expressão (CEGeME) da Escola de Música da UFMG. Na ocasião do estudo citado, foram apontadas e analisadas as recorrências dos padrões de movimento dos músicos, além dos desvios de andamento ocorrido em cada compasso. O excerto escolhido é um trecho do primeiro movimento da Sonata para Clarineta e Piano em Fá menor op. 120 No. 1 de Johannes Brahms (Figura 15). A base de dados utilizada nesta tese foi composta por seis clarinetistas do conjunto de gravações referido no início deste parágrafo. O áudio foi gravado com taxa de amostragem de 44.1KHz, e foi capturado por um microfone condensador posicionado em um pedestal a um metro de distância dos clarinetistas. Cada um dos seis clarinetistas tocou o trecho 6 vezes em duas condições: três vezes em condição livre, e outras três seguindo um metrônomo ajustado no andamento (manualmente medido) da última execução livre. Com isso, o material permite a avaliação de efeitos nos seguintes fatores envolvidos: músico executante (seis músicos), nota executada (26 notas na soma dos dois grupos), duas situações de execução (livre e com metrônomo), e por último, tipo de transição (*Legato* e não *Legato*).

Dois grupos de notas foram selecionados para análise dos primeiros testes desta tese. Grupo 1: Compassos 5 a 10 e Grupo 2: compassos 17 e 18. O Grupo 1, composto pelas 14 primeiras notas do excerto (compasso 5 ao 10), foi selecionado pela característica de alternância entre situação de *Legato* e não *Legato*, com diferentes intervalos e níveis de dificuldade de execução. E o Grupo 2, composto pelas notas 31 a 42 do excerto (compassos 17 e 18), com duas sequências de 6 colcheias ligadas. A Figura 16 (figura retirada e adaptada de [Teixeira et al. \(2015\)](#)) contém o mapa de recorrência de movimento da campana da clarineta, reforçando a escolha desses dois grupos por estarem em regiões de alta recorrência. Na sequência, são descritos os testes estatísticos realizados com os dados experimentais descritos acima, assim como as variáveis e fatores envolvidos.

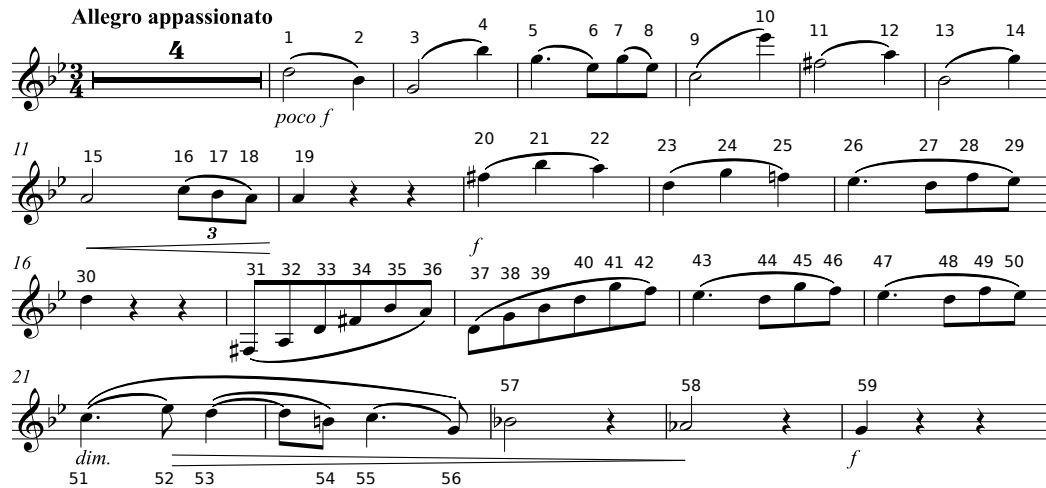


Figura 15 – Excerto do primeiro movimento da Sonata para Clarineta e Piano em Fá menor op. 120 No. 1 de Johannes Brahms, utilizado nos testes experimentais (figura adaptada de [Teixeira et al. \(2015\)](#)).

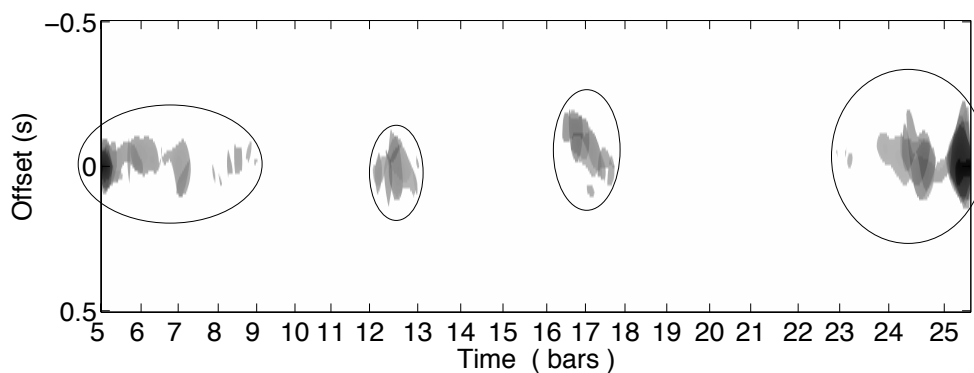


Figura 16 – Mapa de recorrência de movimento da campana da clarineta, retirado de ([TEIXEIRA et al., 2015](#)), que contribuiu para a escolha dos grupos de notas utilizadas para análise. Grupo 1: Compassos 5 a 10 e Grupo 2: compassos 17 e 18.

6.2 Variáveis de interesse

As variáveis de interesse utilizadas nos testes são as seguintes:

- H2** - Proporção de energia da componente harmônica presente na nota em questão, relativa à reverberação da nota anterior.
- R2** - Proporção de energia do segundo resíduos do SMS, depois de ter sido subtraída a componente H2.
- ILh** - Índice de Legato calculado utilizando a curva de **RMS** apenas da parte harmônica.

IL - Índice de Legato tradicional.

As variáveis espúrias identificadas (influentes mas sem interesse) são listadas a seguir:

Instrumento Musical - O Instrumento Musical faz parte da identidade do timbre do intérprete, e por isso não foi padronizado. A acústica de cada instrumento influencia de maneira própria no som produzido, interferindo nas variáveis de interesse.

Palheta Utilizada - Assim como o instrumento, a palheta faz parte da identidade do timbre do intérprete, e por isso não foi padronizado. Esta variável também interfere nas variáveis de interesse.

Harmônicos coincidentes - Em notas seguidas, principalmente ligadas, existe a possibilidade de harmônicos da nota anterior coincidirem com harmônicos da nota em questão, influenciando nas variáveis de interesse.

Identificadas as variáveis espúrias, o ideal é fazer com que elas se distribuam de maneira equivalente em todas as outras de interesse, de forma aleatorizada.

6.3 Fatores de Interesse

Os fatores de interesse são:

Clarinetista - 6 clarinetistas;

Nota - Variam de 1 a 14 no Grupo 1 (**G1**), de 31 a 42 no Grupo 2 (**G2**) nas análises das variáveis **H2**, **R2**, **IL** e **ILh** isoladamente. Nas análises do plano **H2R2** e do **PHA**, as notas selecionadas foram: 4, 10, 12, 14 e 22 (*legatos*) e 9 e 13 (articuladas);

Condição - São duas condições de execução impostas: uma livre, e outra com metrônomo.

Tipo de Nota - Tem-se dois tipos de notas: Notas ligadas e Notas não ligadas.

6.4 Testes Realizados

Estão listados a seguir os testes estatísticos realizados:

Verificação de Efeito nas variáveis H2 e R2:

H2 x Nota - teste **ANOVA** + teste de *Tukey* (comparação múltipla a posteriori);

H2 x Clarinetista - teste **ANOVA** + teste de *Tukey* (comparação múltipla a posteriori);

H2 x Condição - teste **t** pareado;

R2 x Nota - teste **ANOVA** + teste de *Tukey* (comparação múltipla a posteriori);

R2 x Clarinetista - teste **ANOVA** + teste de *Tukey* (comparação múltipla a posteriori);

R2 x Condição - teste **t** pareado;

Transição entre Notas:

IL x Nota - teste **ANOVA** + teste de *Tukey* (comparação múltipla a posteriori);

IL x Clarinetista - teste **ANOVA** + teste de *Tukey* (comparação múltipla a posteriori);

IL x ILh - teste **t**.

Caracterização dos Ataques:

Plano H2-R2 x Clarinetista - teste **MANOVA**

Perfil Harmônico do Ataque (PHA) x Clarinetista - teste **ANOVA** dos Coeficientes de Correlação de Pearson.

6.5 Verificação de Efeito nas variáveis H2 e R2

H2 x Nota

Começando os testes com a verificação de efeito nas variáveis **H2** e **R2**, são apresentados nesta seção, efeitos da variável **H2** para o fator “Nota”. A princípio, são mostradas os *boxplots* das proporções **H2** para as notas do Grupo 1 (Figura 17) e do Grupo 2 (Figura 18).

Na sequência, são apresentados os resultados do teste **H2 x Nota**: efeito do fator Nota na variável **H2**, a partir do teste **ANOVA**. O teste foi aplicado utilizando as notas 37 a 42 selecionadas do Grupo 2 (notas selecionadas mostradas na Figura 19). A escolha das notas analisadas foi motivada pelo cumprimento das premissas de simetria e de homoscedasticidade (igualdade de variância) da distribuição dos resíduos do teste **ANOVA**. As estatísticas do teste estão na Tabela 2, onde *Df* é o grau de liberdade, *Sum Sq* é a soma dos quadrados, *Mean Sq* é a soma quadrática média, *F value* é o valor da estatística F, e *Pr(>F)* é o valor p.

A Figura 20 contém a apresentação dos resíduos para análise e verificação das premissas do teste. Na metade esquerda da figura, por inspeção visual pode ser verificado o ajuste à distribuição normal, enquanto na metade da direita, pode ser verificada a semelhança das variâncias dos resíduos dos ajustes. Os resíduos foram avaliados pelo teste

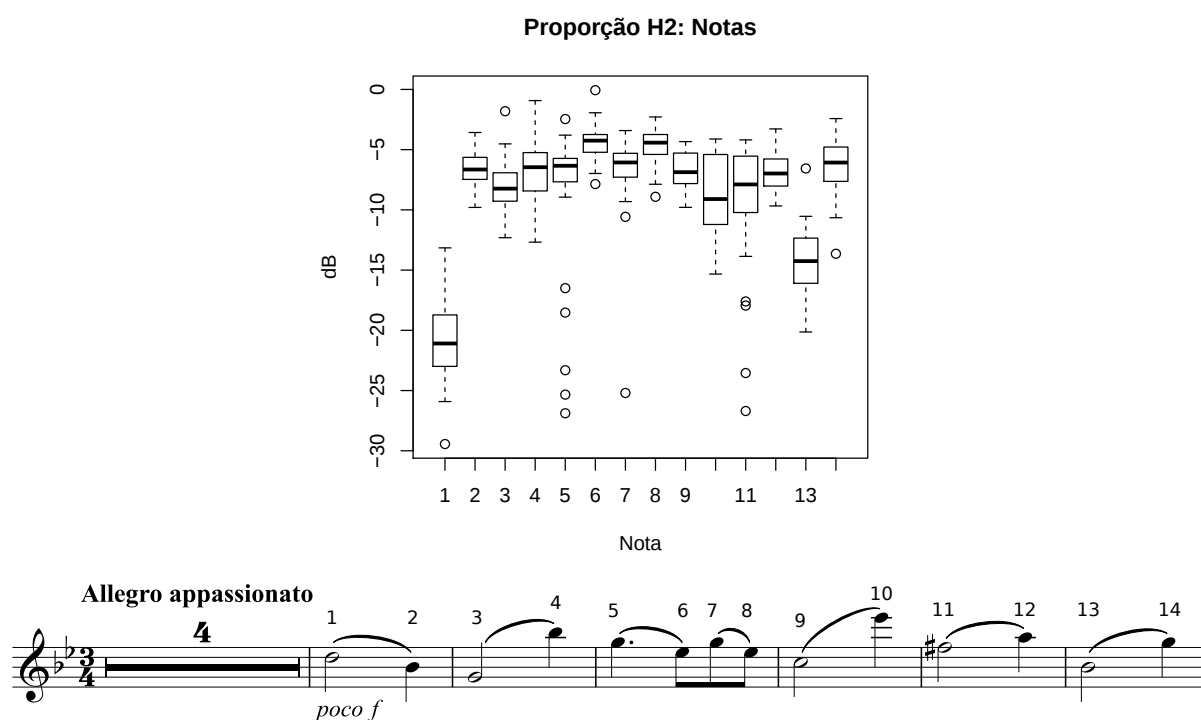


Figura 17 – *Boxplots* das proporções **H2** para as notas do Grupo 1.

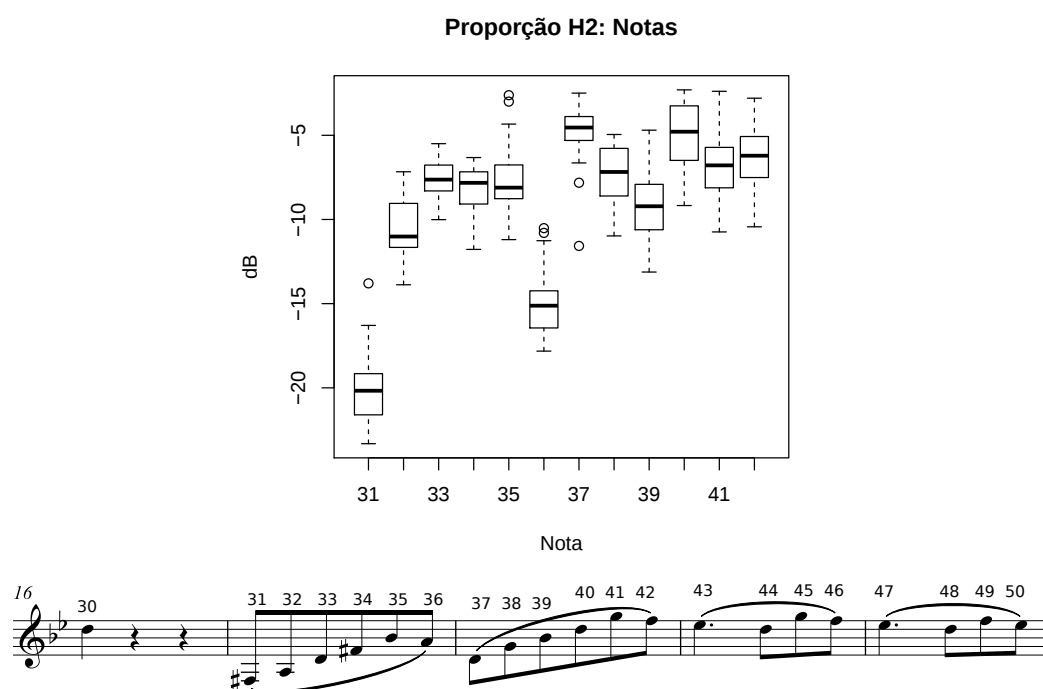


Figura 18 – *Boxplots* das proporções **H2** para as notas do Grupo 2.

de normalidade *Shapiro-Wilk*, no qual não foi identificado um comportamento não normal ($W = 0,989$ e valor $p = 0,119$).

Na Figura 21 podem ser vistos os dados da comparação múltipla a posteriori (teste

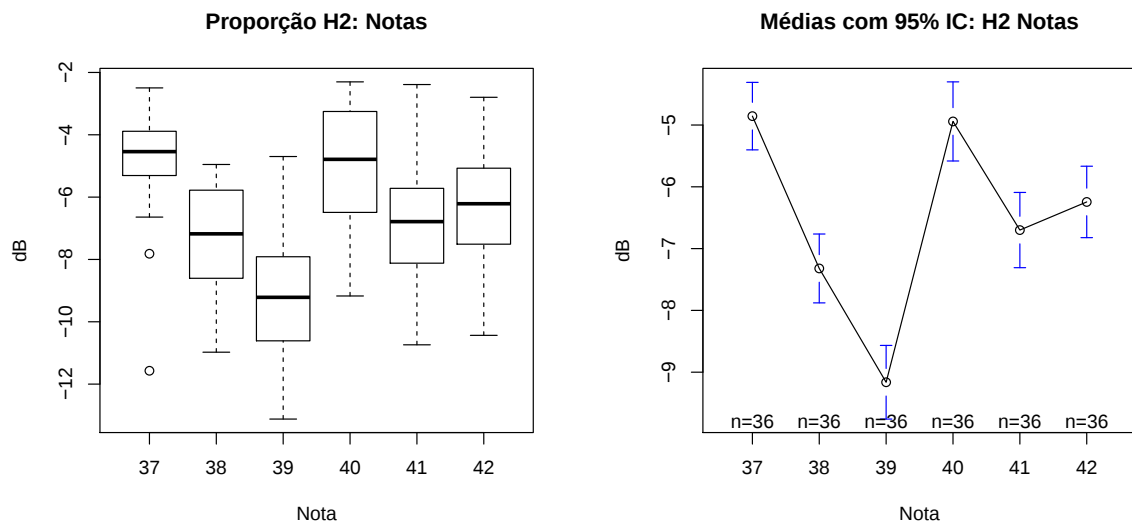


Figura 19 – Notas selecionadas para a **ANOVA**: 37 a 42 do Grupo 2.

Tabela 2 – Estatísticas da **ANOVA** para o teste **H2 x Nota**.

Fator	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Nota	5	467.75	93.55	30.91	<0.0001
Residuals	210	635.48	3.03		

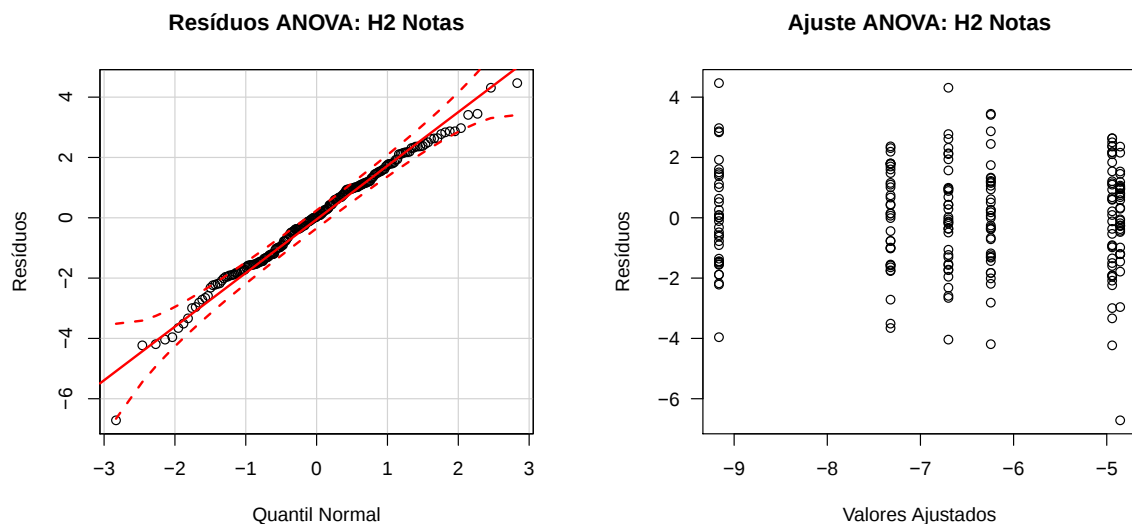


Figura 20 – Resíduos para análise e verificação das premissas da **ANOVA**. Na metade esquerda, o ajuste à distribuição normal. Na metade direita, a semelhança das variâncias dos resíduos dos ajustes.

de *Tukey*) para o teste **H2 x Nota**. Na figura, a metade esquerda contém um gráfico com as diferenças para cada comparação, com intervalos de confiança a 95 %. A metade direita da figura contém os valores das estatísticas das comparações. Ainda na tabela da metade direita, as abreviações são explicadas a seguir: diff - diferenças entre as notas comparadas

em dB; lwr - valor inferior do intervalo de confiança; upr - valor superior do intervalo de confiança; padj - valor p ajustado pela comparação múltipla. Min - valor mínimo das diferenças; Med - mediana das diferenças; Max - valor máximo das diferenças.

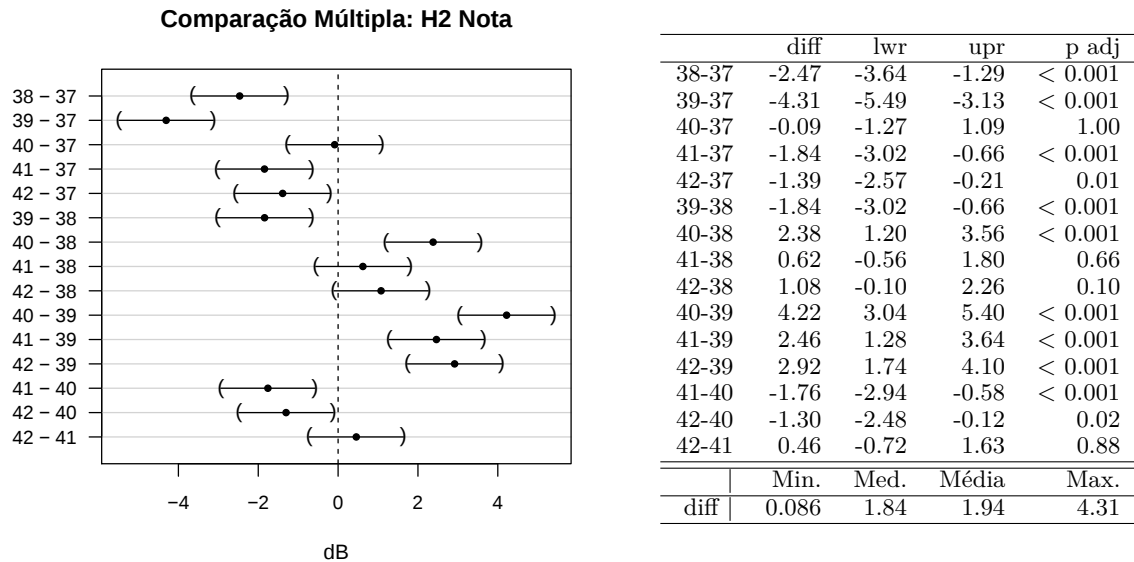


Figura 21 – Resultado da comparação múltipla **H2 x Nota**

As notas 1 e 31 apresentam medianas de **H2** abaixo de $-20dB$. Por não apresentarem notas anteriores, a restrição para **SMS(3)** foi ajustada, nesse caso, para a F0 da própria nota. Portanto, esse valor pode ser considerados um valor de fundo para a modelagem. Em algumas notas como a 5, e 11, ocorreram **outliers** em torno de $-20dB$, podendo significar execuções mal sucedidas. Além disso, a nota 6 apresenta alguns **outliers** em $0dB$, o que significa que a modelagem, para esse caso, encontrou exatamente a energia da nota original, indicando algum problema novamente. Para esses casos, deve ser realizada a verificação auditiva. No conjunto de notas avaliadas pela **ANOVA**, a hipótese nula das médias serem iguais foi refutada (valor $p < 0,0001$) com as premissas do teste sendo bem atendidas. A comparação múltipla (teste de *Tukey*) apontou a menor diferença significativa de **H2** entre as notas como sendo 1,76 dB e a máxima de 4,31 dB.

H2 x Clarinetista

Na sequência, são apresentados os resultados do teste **H2 x Clarinetista**: efeito do fator Clarinetista na variável **H2**, a partir do teste **ANOVA**. O teste foi aplicado utilizando as notas 37 a 42 selecionadas do Grupo 2 (as mesmas notas do teste **H2 x Nota**). Os **boxplots** e médias com intervalos de confiança a 95 % para o fator Clarinetista estão na Figura 22. As estatísticas do teste estão na Tabela 3.

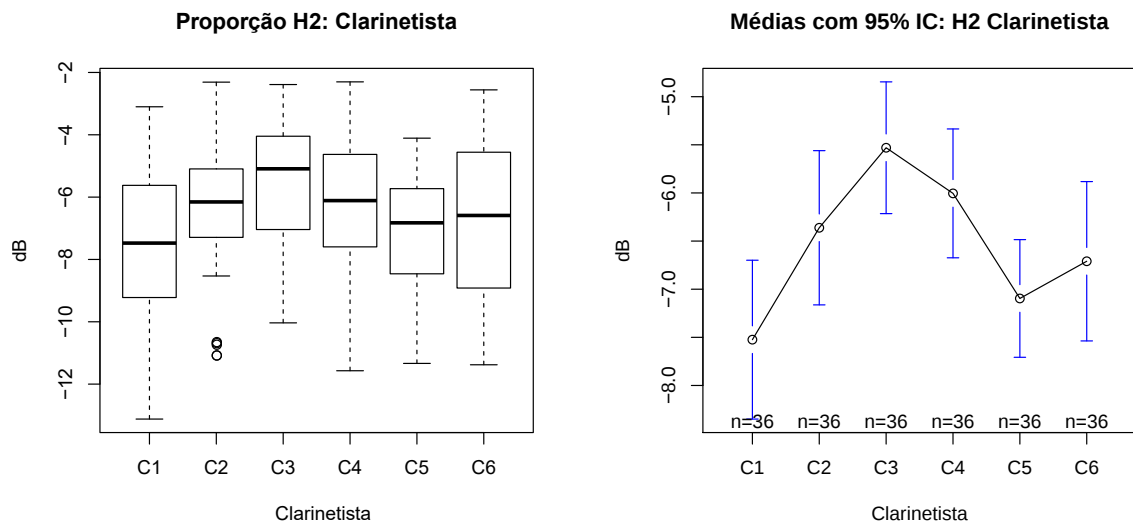


Figura 22 – **Boxplots** (à esquerda) e médias com intervalos de confiança a 95 % (à direita) para o fator Clarinetista utilizado na **ANOVA**.

Tabela 3 – Estatísticas da **ANOVA** para o teste **H2 x Clarinetista**.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Clarinetista	5	95.19	19.04	3.97	0.0019
Residuals	210	1008.05	4.80		

A Figura 23 contém a apresentação dos resíduos para análise e verificação das premissas do teste. Na metade esquerda da figura, por inspeção visual pode ser verificado o ajuste à distribuição normal, enquanto na metade da direita, pode ser verificada a semelhança das variâncias dos resíduos dos ajustes. Os resíduos foram avaliados pelo teste de normalidade *Shapiro-Wilk*, no qual foi identificado um comportamento não normal somente ao nível de significância 0,04 ($W = 0,987$ e valor $p = 0,042$).

Na Figura 24 podem ser vistos os dados da comparação múltipla a posteriori (teste de *Tukey*) para o teste **H2 x Clarinetista**. Na figura, a metade esquerda contém um gráfico com as diferenças para cada comparação, com intervalos de confiança a 95 %. A metade direita da figura contém os valores das estatísticas das comparações. Ainda na tabela da metade direita, as abreviações são explicadas a seguir: diff - diferenças entre as notas comparadas em dB; lwr - valor inferior do intervalo de confiança; upr - valor superior do intervalo de confiança; padj - valor p ajustado pela comparação múltipla. Min - valor mínimo das diferenças; Med - mediana das diferenças; Max - valor máximo das diferenças.

Portanto, o teste **ANOVA** obteve a hipótese nula das médias serem iguais refutada (valor $p = <0,0019$) com as premissas do teste sendo bem atendidas. A comparação múltipla (teste de *Tukey*) apontou a menor diferença significativa de **H2** entre Clarinetistas como sendo 1,57dB e a máxima de 1,99dB. Os efeitos para o fator Clarinetista foram menores do que para o fator Nota, apesar de ainda serem significantes. A premissa de

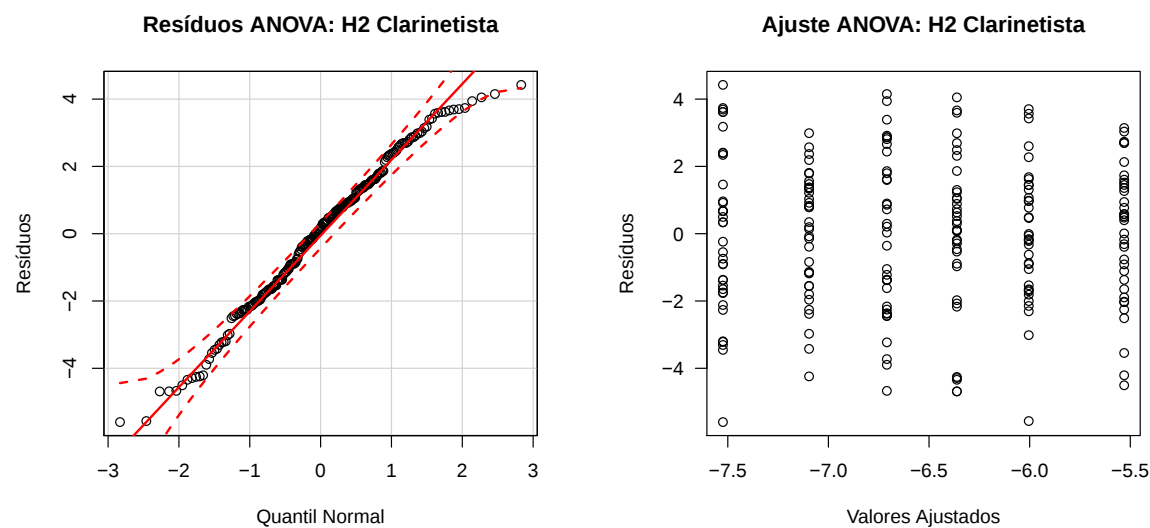


Figura 23 – Resíduos para análise e verificação das premissas da ANOVA para o teste do fator Clarinetista. Na metade esquerda, o ajuste à distribuição normal. Na metade direita, a semelhança das variâncias dos resíduos dos ajustes.

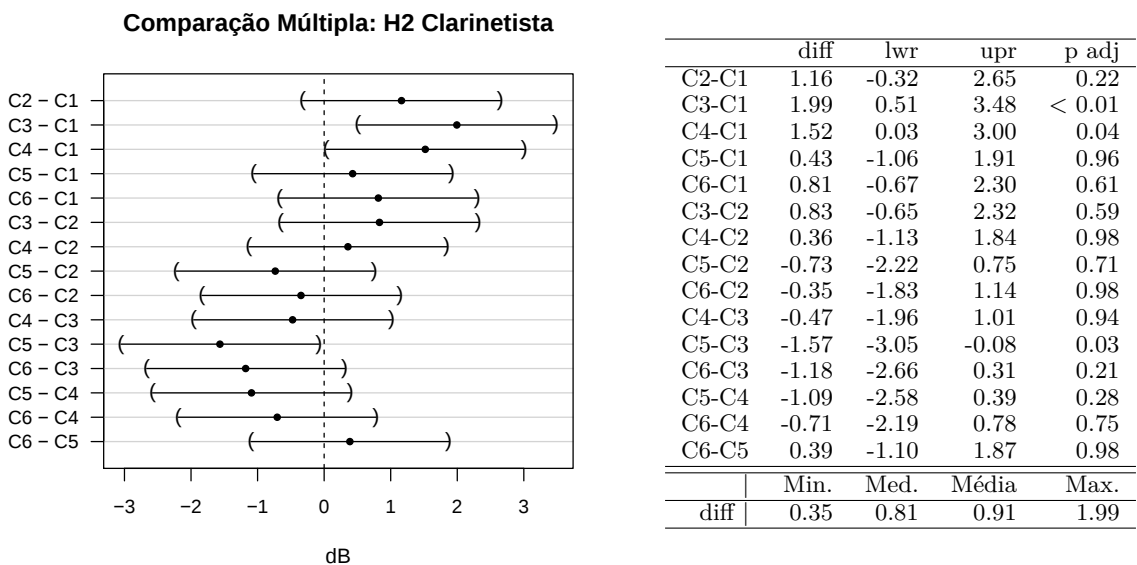


Figura 24 – Resultado da comparação múltipla para o fator Clarinetista na variável **H2**.

normalidade dos resíduos não foi numericamente atendida segundo o teste *Shapiro-Wilk* a 95 % de nível de confiança (valor $p = 0,042$), ficando porém, muito próximo de não refutar a hipótese de normalidade. Como visualmente os resíduos se mantiveram dentro dos limites no gráfico Quantil Normal com poucas exceções, não foi considerada a validade do teste.

H2 x Condição

Na sequência, são apresentados os resultados do teste **H2 x Condição**: efeito do fator Clarinetista na variável **H2**, a partir de teste **t** pareado. O teste foi aplicado utilizando as notas 37 a 42 selecionadas do Grupo 2 (as mesmas notas do teste **H2 x Condição**). Os **boxplots** e médias com intervalos de confiança a 95 % para o fator Condição estão na Figura 25.

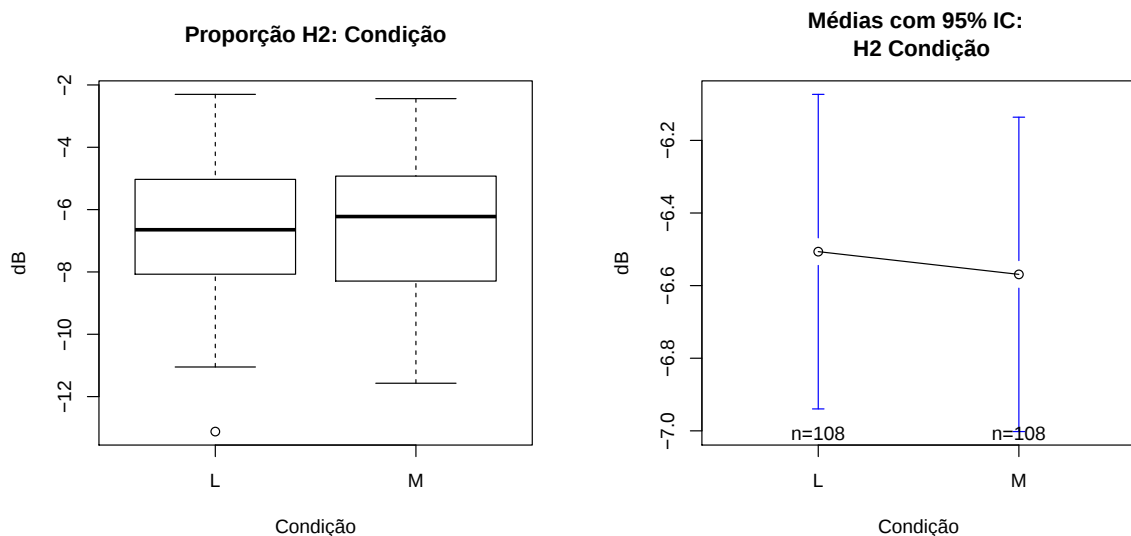


Figura 25 – **Boxplots** (à esquerda) e médias com intervalos de confiança a 95 % (à direita) para o fator Condição utilizado no teste **t**.

As estatísticas do teste estão na Tabela 4. Como a hipótese nula não foi rejeitada ao nível de confiança 95 %, foi testada a potência do teste. Considerando um tamanho de efeito mínimo de 1 dB (ordem de grandeza da mediana dos efeitos das notas), a potência foi de 0,99. Para uma potência de 0,8, o tamanho de efeito mínimo é de 0,55 dB.

Tabela 4 – Estatísticas do teste **t** pareado para **H2 x Clarinetista**.

	Dif.	Df.	t-value	p-value	IC(95%) lwr	upr
Condição	-0.063	107	0.325	0.746	-0.320	0.445

A Figura 26 contém a apresentação dos resíduos para análise e verificação das premissas do teste. Na metade esquerda da figura, por inspeção visual pode ser verificada a simetria da distribuição pelo **boxplot** dos resíduos, enquanto na metade direita, o ajuste à distribuição normal. Os resíduos foram avaliados pelo teste de normalidade *Shapiro-Wilk*, no qual não foi identificado um comportamento não normal ($W = 0,987$ e valor $p = 0,388$).

Com isso, não foram obtidas diferenças significativas (valor $p = 0,445$) para o fator “Condição” na variável **H2**, com diferença igual a apenas 0,06dB. O tamanho mínimo de efeito seria de 0,55 dB, considerando uma potência de 0,8. O teste de normalidade não

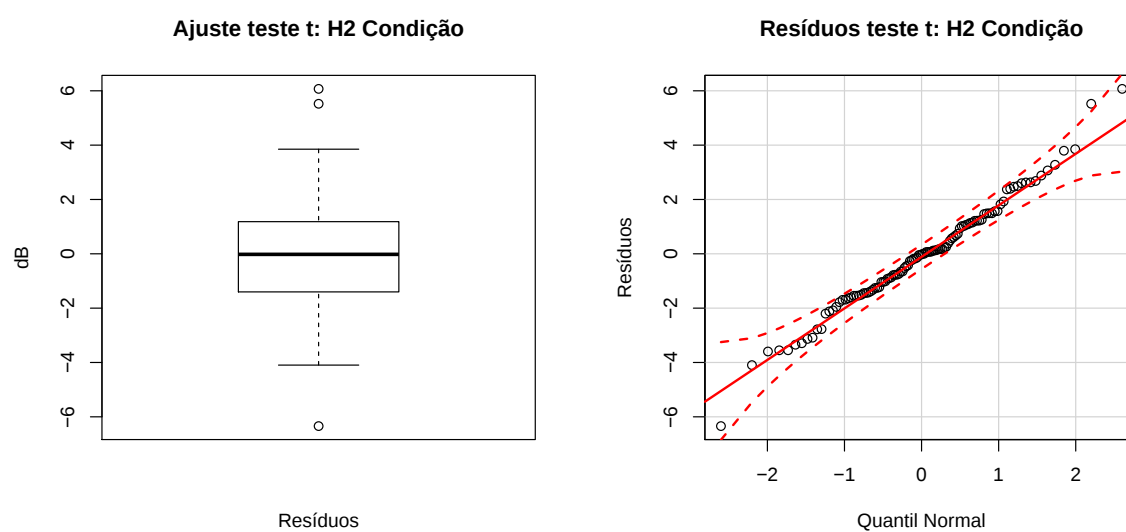


Figura 26 – Resíduos para análise e verificação das premissas do **teste t** pareado para o fator Condição. Na metade direita, o ajuste à distribuição normal. Na esquerda, o **boxplot** dos resíduos.

rejeitou a hipótese nula de os resíduos serem normais a 95 % de confiança, (valor $p = 0,388$). Assim, as premissas foram atendidas.

R2 x Nota

Mudando para a variável **R2**, teste **R2 x Nota**, no conjunto de notas avaliadas pela **ANOVA** (mesmas notas para **H2**), a hipótese nula das médias serem iguais foi também refutada (valor $p = <0,0001$). A premissa de normalidade dos resíduos não foi numericamente atendida pelo teste *Shapiro-Wilk* a 95% de nível de confiança (valor $p = 0,01$), ficando porém não muito distante disso. Já o gráfico de resíduos por valores ajustados apontou uma ligeira diferença nas variâncias. Apesar dos pequenos desvios nas premissas para os resíduos, o descarte do teste não foi considerado necessário. A comparação múltipla (teste de *Tukey*) apontou a menor diferença significativa de **R2** entre as notas como sendo 1,96 dB e a máxima de 3,97 dB.

A princípio, são mostradas os *boxplots* das proporções **R2** para as notas do Grupo 1 (Figura 27) e do Grupo 2 (Figura 28). As estatísticas do teste estão na Tabela 5.

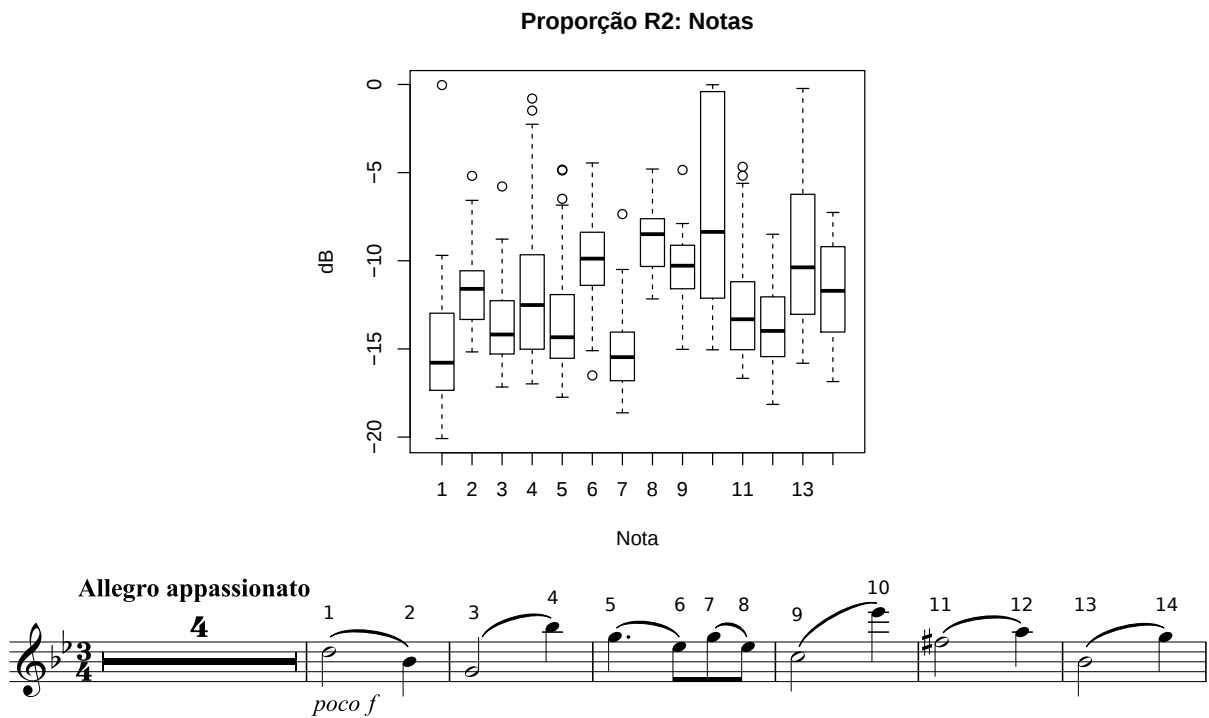


Figura 27 – *Boxplots* das proporções **R2** para as notas do Grupo 1.

Tabela 5 – Estatísticas da **ANOVA** para o teste **R2 x Nota**.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Nota	5	314.25	62.85	14.68	2.44e-12
Resíduos	210	899.34	4.28		

Na Figura 30 estão os resíduos para análise e verificação das premissas do teste. Na metade esquerda da figura, por inspeção visual pode ser verificado o ajuste à distribuição normal, enquanto na metade da direita, pode ser verificada a semelhança das variâncias dos

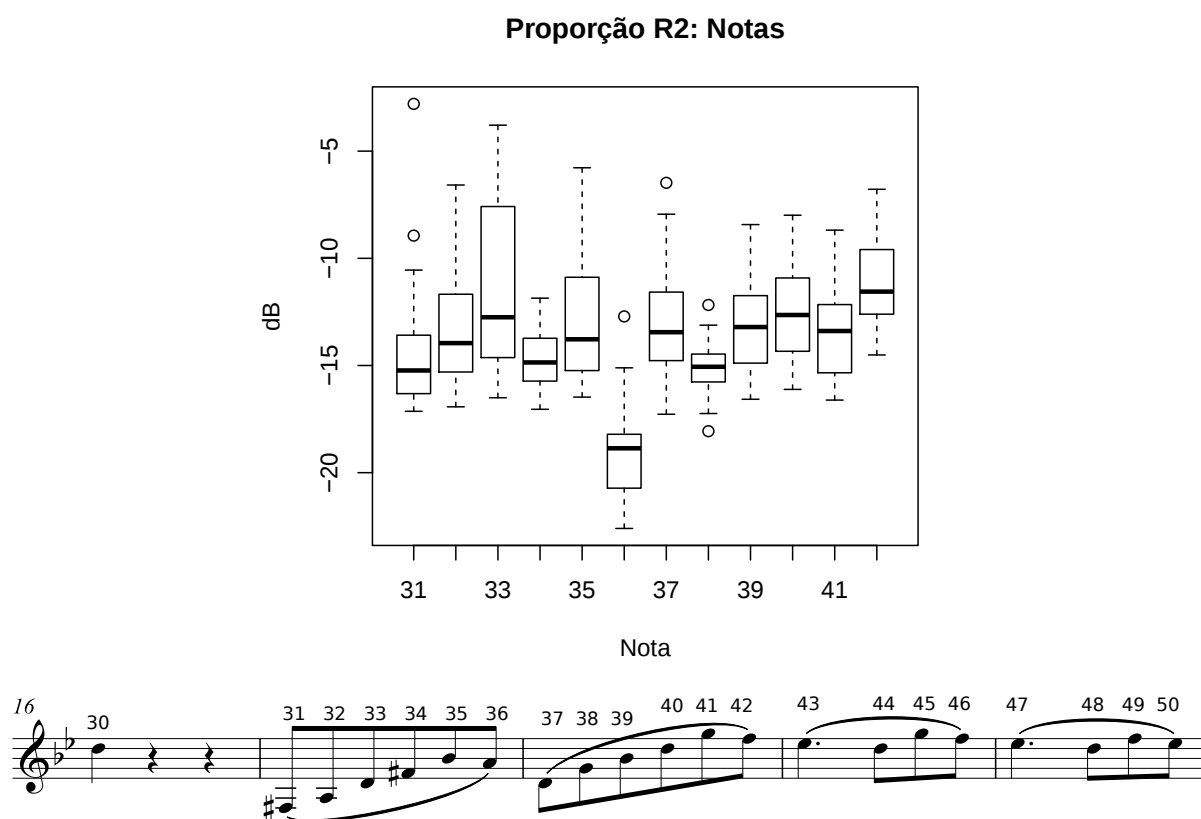


Figura 28 – *Boxplots* das proporções **R2** para as notas do Grupo 2.

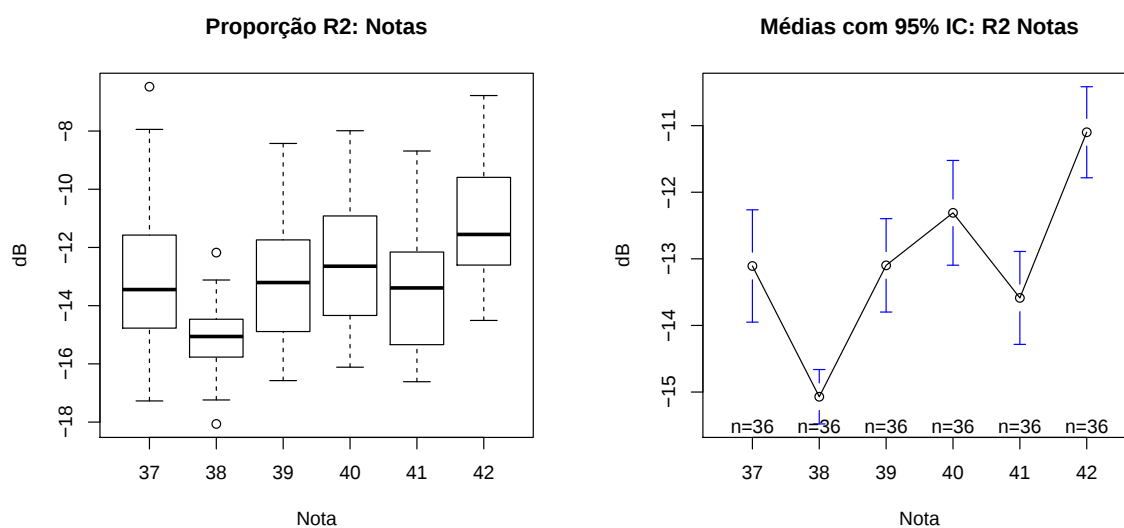


Figura 29 – Notas selecionadas para a **ANOVA** no teste do fator Notas para R2: 37 a 42 do Grupo 2.

resíduos dos ajustes. Os resíduos foram avaliados pelo teste de normalidade *Shapiro-Wilk*, no qual foi identificado um comportamento não normal ao nível de confiança 95 % ($W = 0,983$, valor $p = 0,010$).

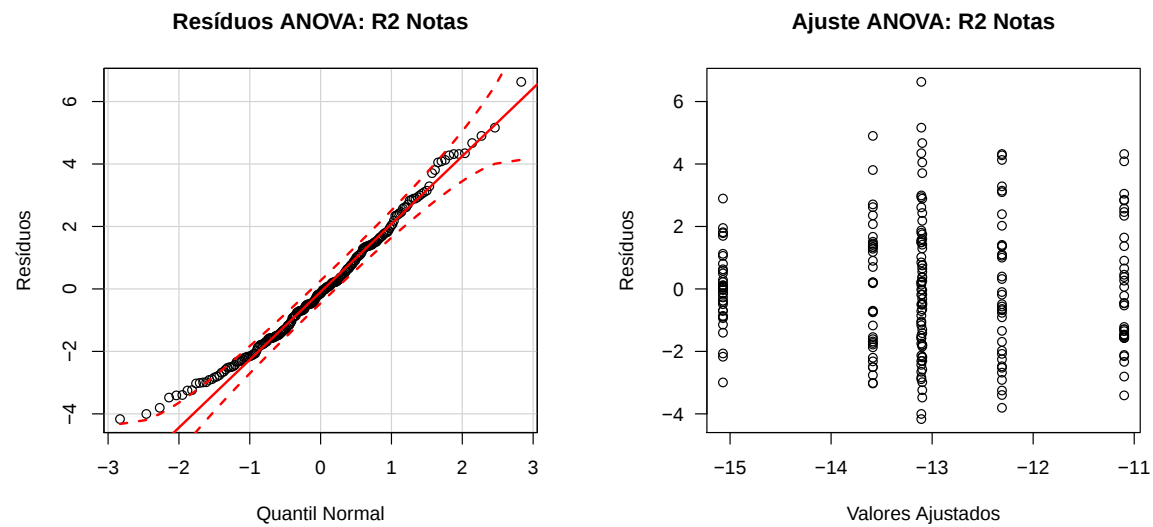


Figura 30 – Resíduos para análise e verificação das premissas da **ANOVA**. Na metade esquerda, o ajuste à distribuição normal. Na metade direita, a semelhança das variâncias dos resíduos dos ajustes.

Na Figura 31 podem ser vistos os dados da comparação múltipla a posteriori (teste de *Tukey*) para o teste **R2 x Nota**. Na figura, a metade esquerda contém um gráfico com as diferenças para cada comparação, com intervalos de confiança a 95 %. A metade direita da figura contém os valores das estatísticas das comparações. Ainda na tabela da metade direita, as abreviações são explicadas a seguir: diff - diferenças entre as notas comparadas em dB; lwr - valor inferior do intervalo de confiança; upr - valor superior do intervalo de confiança; padj - valor p ajustado pela comparação múltipla. Min - valor mínimo das diferenças; Med - mediana das diferenças; Max - valor máximo das diferenças.

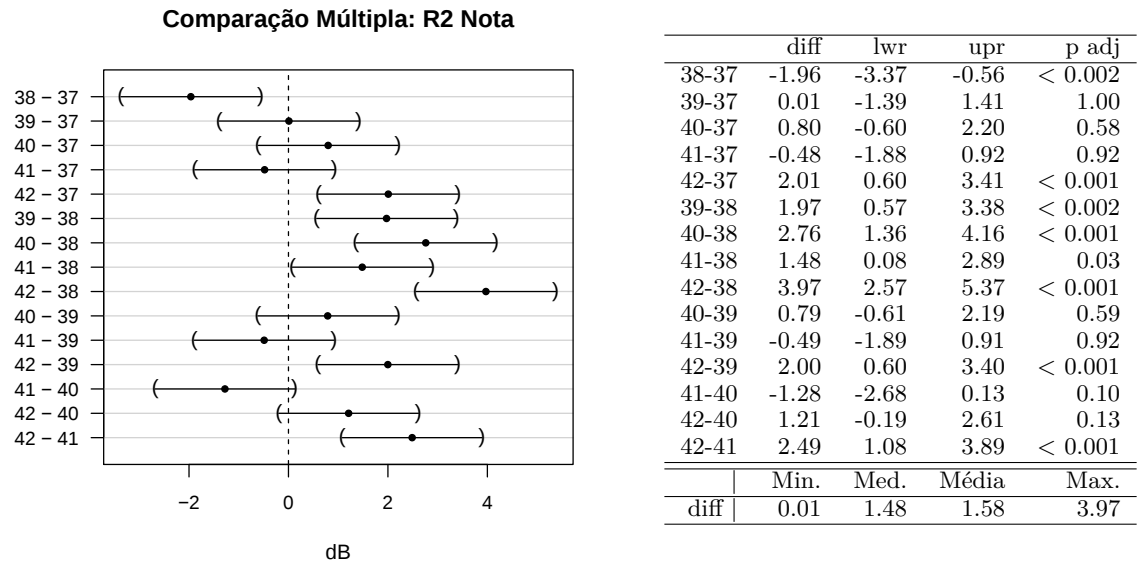


Figura 31 – Resultado da comparação múltipla **R2 x Nota**.

R2 x Clarinetista

Avaliando agora os resultados do teste **R2 x Clarinetista**, o teste **ANOVA** obteve também a hipótese nula refutada das médias serem iguais (valor $p = <0,0009$). A premissa de normalidade dos resíduos não foi numericamente atendida segundo o teste *Shapiro-Wilk* a 95% de nível de confiança (valor $p = 0,006$), ficando porém não muito distante disso. A análise visual do gráfico de Quantil Normal para os resíduos aponta que o teste está no limite da aceitabilidade, com muitos valores próximos ao limite de confiança a 95%. A comparação múltipla (teste de *Tukey*) apontou a menor diferença significativa de **R2** entre Clarinetistas como sendo 1,56 dB e a máxima de 1,73 dB.

O teste foi também aplicado utilizando as notas 37 a 42 selecionadas do Grupo 2 (as mesmas notas do teste **R2 x Nota**). Os **boxplots** e médias com intervalos de confiança a 95 % para o fator Clarinetista estão na Figura 32. As estatísticas do teste estão na Tabela 6.

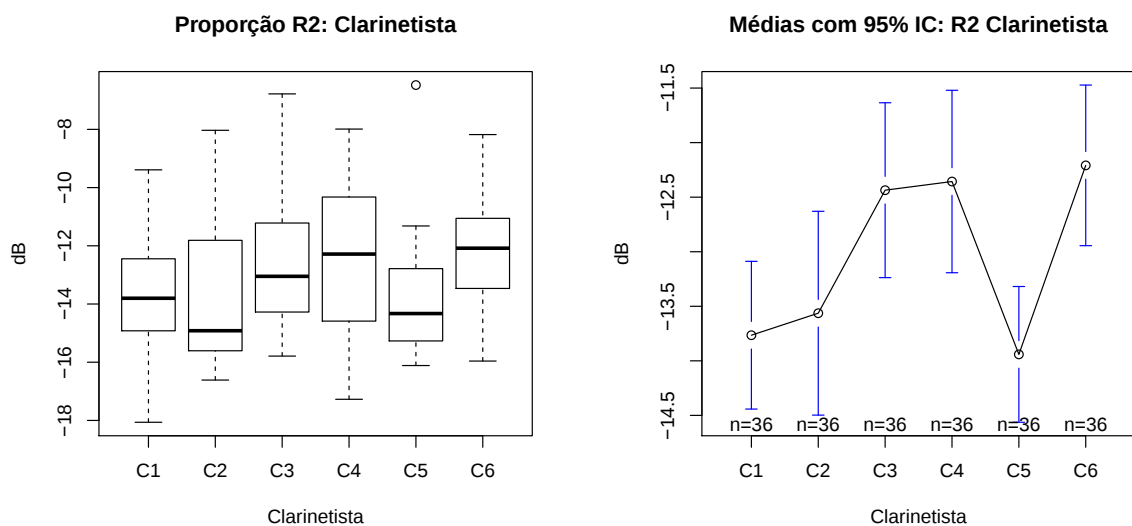


Figura 32 – **Boxplots** (à esquerda) e médias com intervalos de confiança a 95 % (à direita) para o fator Clarinetista utilizados na **ANOVA**.

Tabela 6 – Estatísticas da **ANOVA** para o teste **R2 x Clarinetista**.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Clarinetista	5	112.84	22.57	4.31	0.09e-2
Resíduos	210	1100.75	5.24		

A Figura 33 contém os resíduos para análise e verificação das premissas do teste. Na metade esquerda da figura, por inspeção visual pode ser verificado o ajuste à distribuição normal, enquanto na metade da direita, pode ser verificada a semelhança das variâncias dos resíduos dos ajustes. Os resíduos foram avaliados pelo teste de normalidade *Shapiro-Wilk*,

no qual foi identificado um comportamento não normal ao nível de confiança 95% ($W = 0,982$, valor $p = 0,006$).

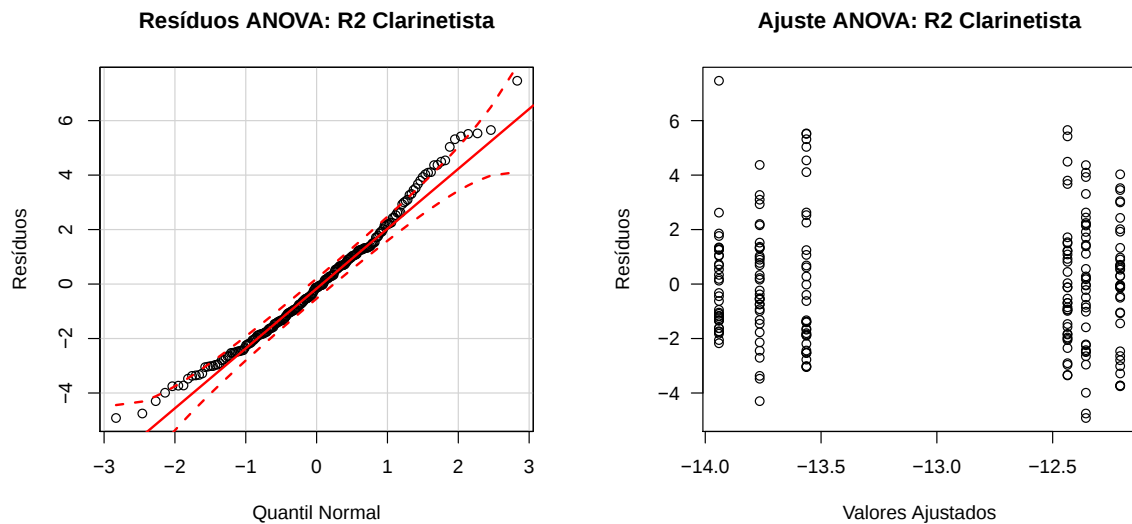


Figura 33 – Resíduos para análise e verificação das premissas da **ANOVA** para o teste da influência do fator Clarinetista na variável R2. Na metade esquerda, o ajuste à distribuição normal. Na metade direita, a semelhança das variâncias dos resíduos dos ajustes.

Na Figura 34 podem ser vistos os dados da comparação múltipla a posteriori (teste de *Tukey*) para o teste **R2 x Clarinetista**. Na figura, a metade esquerda contém um gráfico com as diferenças para cada comparação, com intervalos de confiança a 95 %. A metade direita da figura contém os valores das estatísticas das comparações. Ainda na tabela da metade direita, as abreviações são explicadas a seguir: diff - diferenças entre as notas comparadas em dB; lwr - valor inferior do intervalo de confiança; upr - valor superior do intervalo de confiança; padj - valor p ajustado pela comparação múltipla. Min - valor mínimo das diferenças; Med - mediana das diferenças; Max - valor máximo das diferenças.

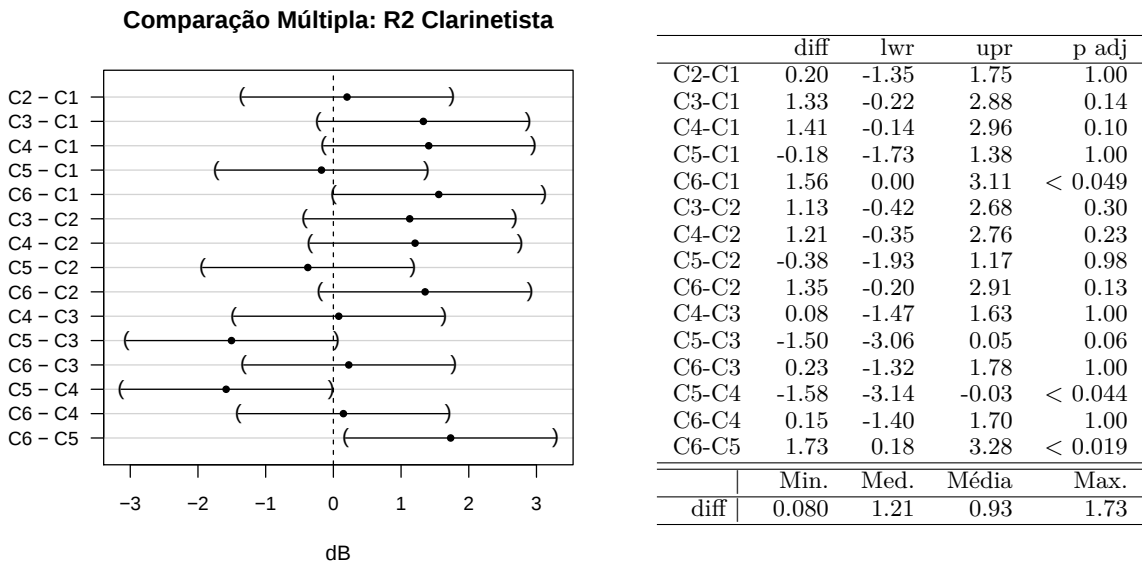


Figura 34 – Resultado da comparação múltipla para o fator Clarinetista na variável R2.

R2 x Condição

Passando para o fator Condição (**R2 x Condição**), ainda sobre a variável **R2**, o fator não obteve diferenças significativas em seu efeito (valor $p = 0,647$), com diferença igual a apenas 0,26 dB. O tamanho mínimo de efeito seria de 0,56 dB, considerando uma potência de 0,8. As premissas do teste foram atendidas. O teste foi também aplicado utilizando as notas 37 a 42 selecionadas do Grupo 2. Os **boxplots** e médias com intervalos de confiança a 95 % para o fator Condição estão na Figura 35.

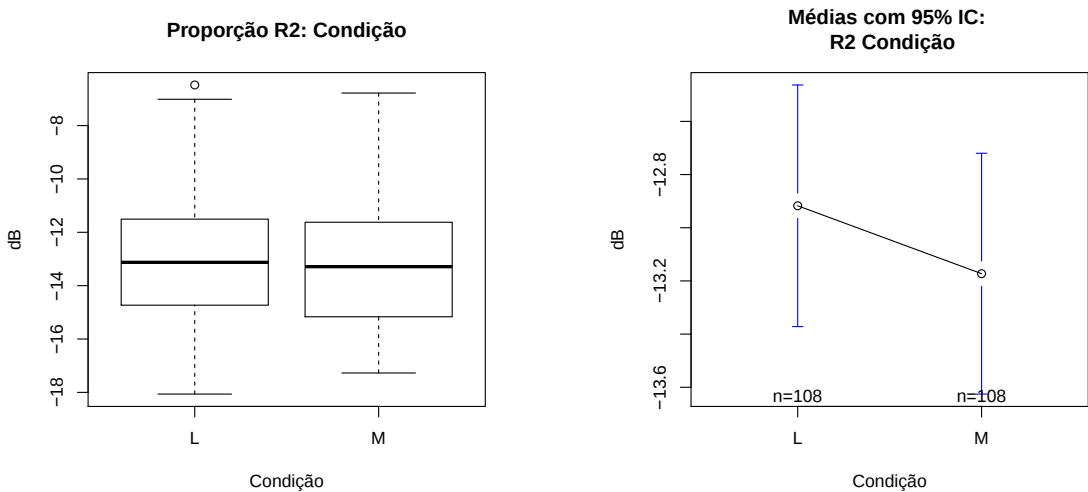


Figura 35 – **Boxplots** (à esquerda) e médias com intervalos de confiança a 95 % (à direita) para o fator Condição utilizados no teste **t**.

As estatísticas do teste estão na Tabela 7. Como a hipótese nula não foi rejeitada

ao nível de confiança 95%, foi testada a potência do teste. Considerando um tamanho de efeito mínimo de 1,00 dB (ordem de grandeza da mediana dos efeitos das notas), a potência foi de 0,9. Para uma potência de 0,8, o tamanho de efeito mínimo é de 0,56 dB.

Tabela 7 – Estatísticas do teste **t** pareado para **R2 x Condição**.

	Dif.	Df.	t-value	p-value	IC(95%) lwr	upr
Condição	0.255	107	1.293	0.199	-0.136	0.647

A Figura 36 contém a apresentação dos resíduos para análise e verificação das premissas do teste (variável **R2**). Na metade esquerda da figura, por inspeção visual pode ser verificada a simetria da distribuição pelo **boxplot** dos resíduos, enquanto na metade direita, o ajuste à distribuição normal. Os resíduos foram avaliados pelo teste de normalidade *Shapiro-Wilk*, no qual não foi identificado um comportamento não normal ao nível de confiança 95 % ($W = 0,992$, valor $p = 0,783$).

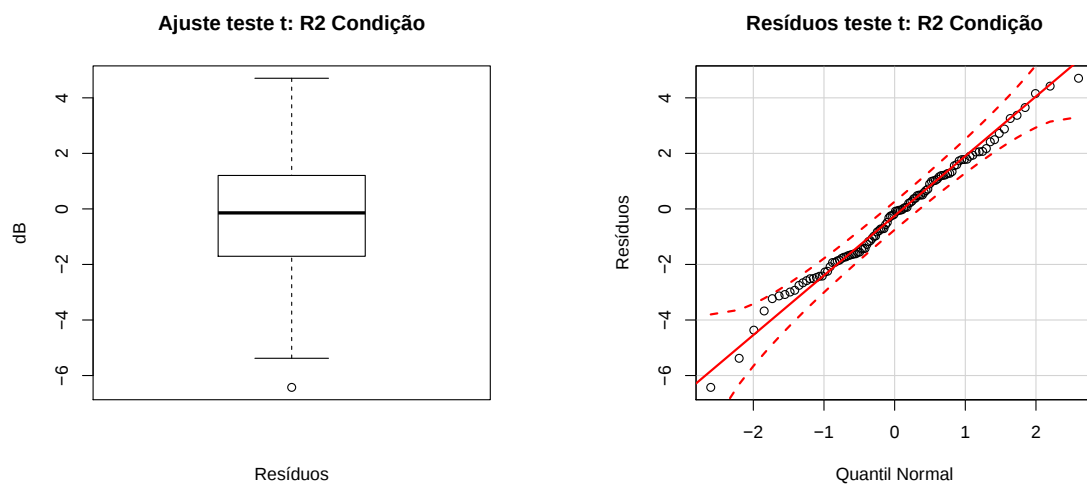


Figura 36 – Resíduos para análise e verificação das premissas do teste **t** pareado para o fator Condição com a variável **R2**. Na metade direita, o ajuste à distribuição normal. Na esquerda, o **boxplot** dos resíduos.

6.6 Transição entre Notas

Mudando agora para os testes com o Índice de Legato (**IL**), uma consideração inicial deve ser feita. Como se trata de um índice para avaliar transições entre notas, um valor para o índice envolve duas notas consecutivas. Para facilitar a representação numérica, os valores de Índice de Legato foram associados ao número da transição. Associar ao número da transição equivale numericamente a se associar o valor do **IL** para a primeira das duas notas envolvidas. Exemplo: para 3 notas temos duas transições; com isso, teríamos **IL** para as transições 1 e 2, que numericamente podem ser associadas com as notas 1 e 2. Assim,

visando manter o nome para o fator *Nota* nos testes do **IL**, foi adotado esse procedimento de representação.

IL x Nota

Com o Índice de Legato, começando com **IL x Nota**, o teste **ANOVA** refutou a hipótese das médias serem iguais (valor $p < 0,0001$) para as notas ligadas do grupo 1. Houve um pequeno desvio de normalidade segundo o teste *Shapiro-Wilk* a 95 % de nível de confiança (valor $p = 0,228e-3$), porém, no gráfico de resíduos por quantil normal, os resíduos não apresentam um desvio visualmente impeditivos. A comparação múltipla (teste de *Tukey*) apontou a menor diferença significativa de **IL** entre Notas como sendo 0,16 e a máxima de 0,26 dB.

Na sequência, são mostradas os *boxplots* dos índices para as notas do Grupo 1 (Figura 37) e do Grupo 2 (Figura 38), além dos *boxplots* dos índices para as notas do Grupo 1 (Figura 39) apenas para as notas ligadas.

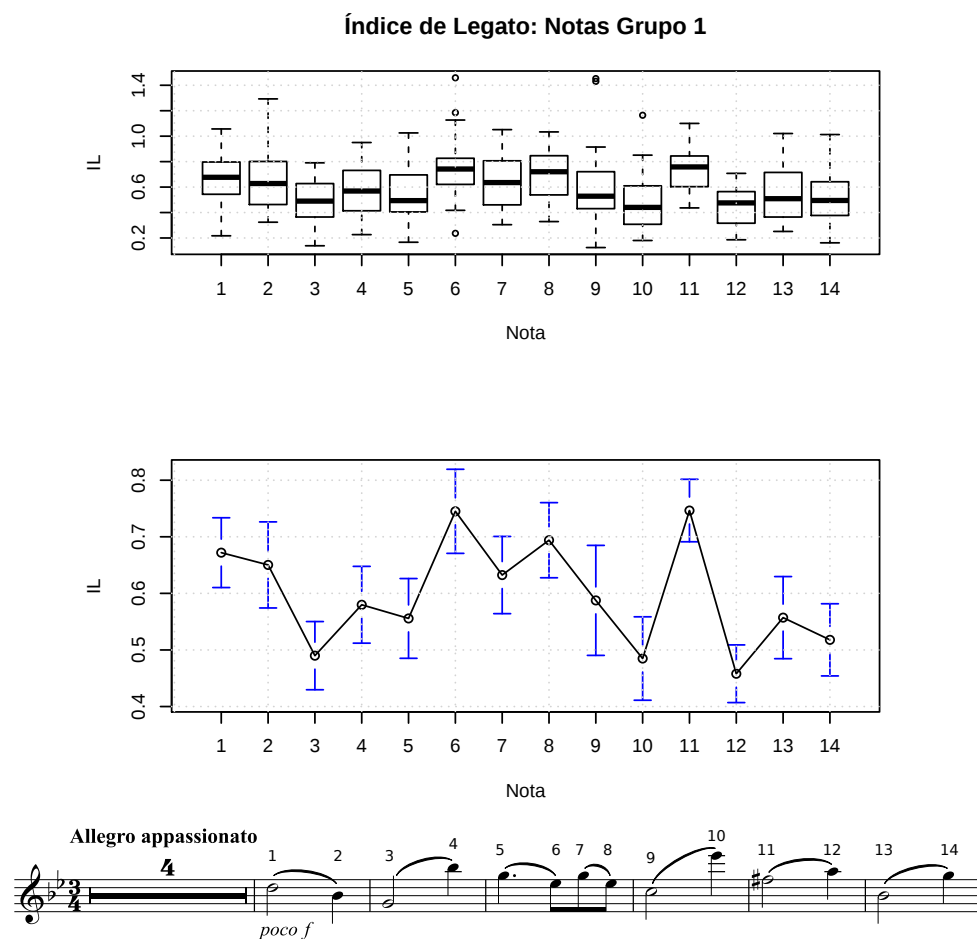


Figura 37 – *Boxplots* de **IL** para as notas do Grupo 1.

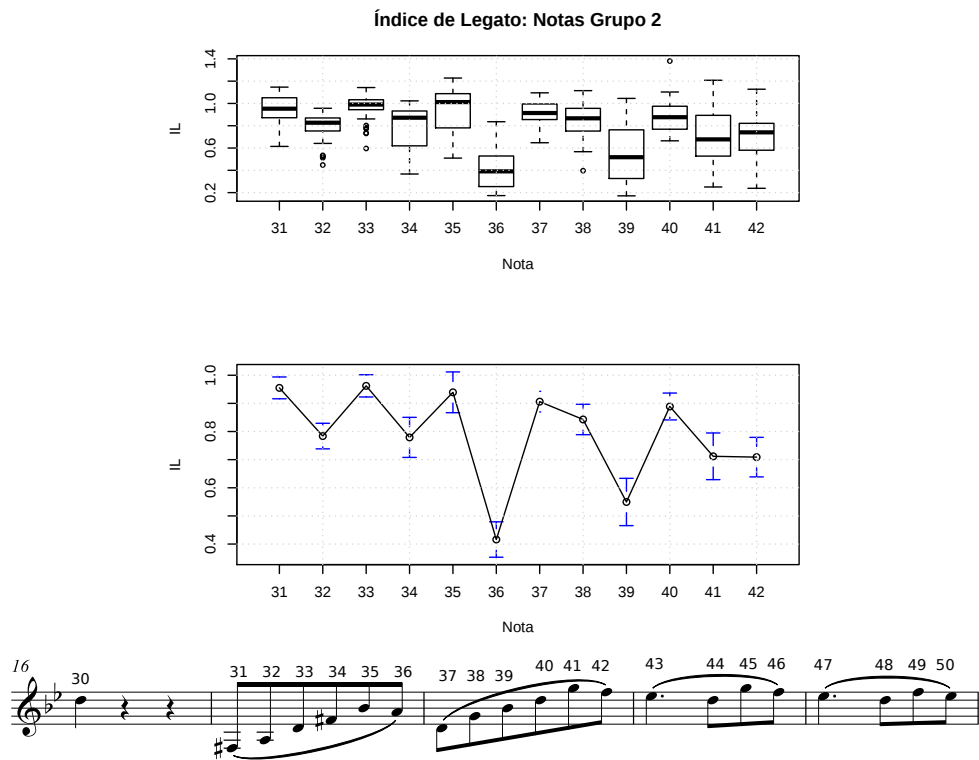


Figura 38 – *Boxplots* de **IL** para as notas do Grupo 2.

Em seguida, são apresentados os resultados do teste **IL x Nota**: efeito do fator Nota na variável **IL**, a partir do teste **ANOVA**. O teste foi aplicado utilizando as notas 1 a 13 selecionadas do Grupo 1 (notas selecionadas mostradas na Figura 40). As estatísticas do teste estão na Tabela 8.

Tabela 8 – Estatísticas da **ANOVA** para o teste **IL x Nota**.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Nota	6	1.57	0.26	6.01	7.05e-06
Residuals	245	10.64	0.04		

A Figura 41 contém os resíduos para análise e verificação das premissas do teste. Na metade esquerda da figura, por inspeção visual pode ser verificado o ajuste à distribuição normal, enquanto na metade da direita, pode ser verificada a semelhança das variâncias dos resíduos dos ajustes. Foi identificada a presença de dois **outliers**, porém muito próximos ao limiar de 95% de nível de confiança para normalidade dos resíduos (linha tracejada da Figura 41). Os resíduos foram avaliados pelo teste de normalidade *Shapiro-Wilk*, no qual foi identificado um comportamento não normal ao nível de confiança 95% ($W = 0,975$, valor $p = 0,228e-3$).

Na Figura 42 podem ser vistos os dados da comparação múltipla a posteriori (teste de *Tukey*) para o teste **IL x Nota**. Na figura, a metade esquerda contém um gráfico com as diferenças para cada comparação, com intervalos de confiança a 95 %. A metade direita

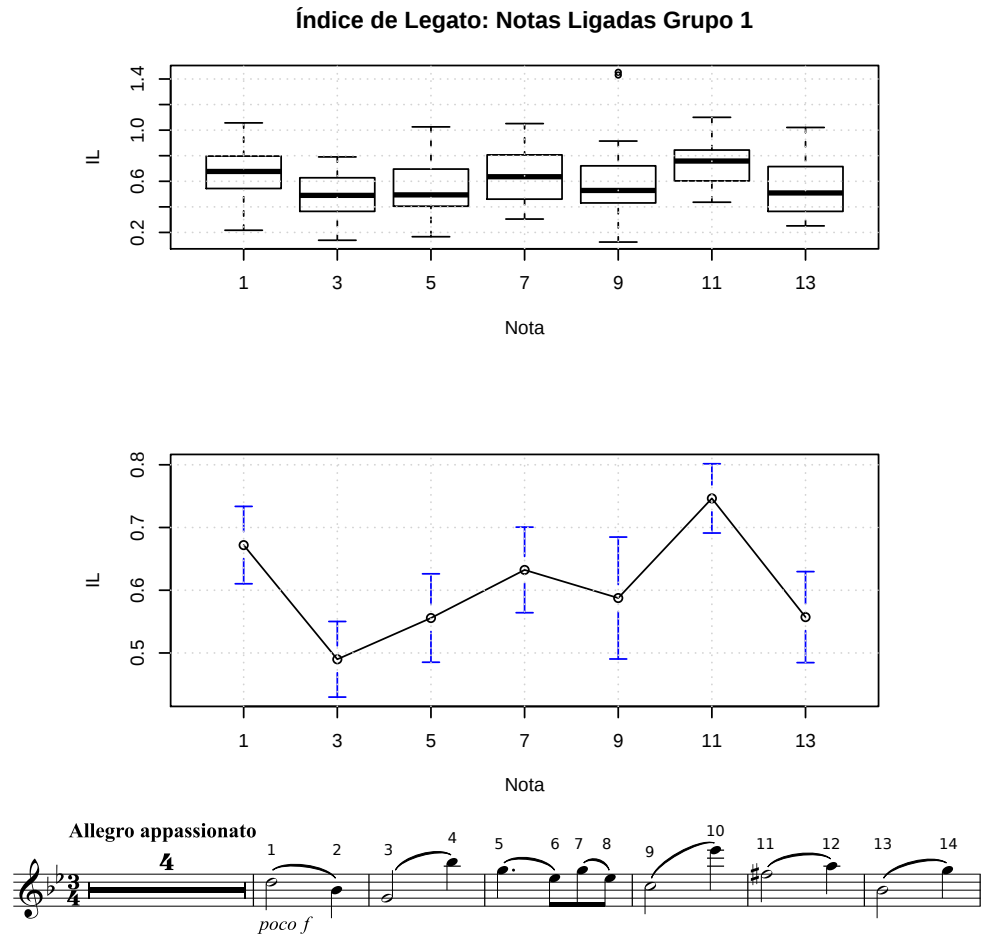


Figura 39 – *Boxplots* de **IL** para as notas ligadas do Grupo 1.

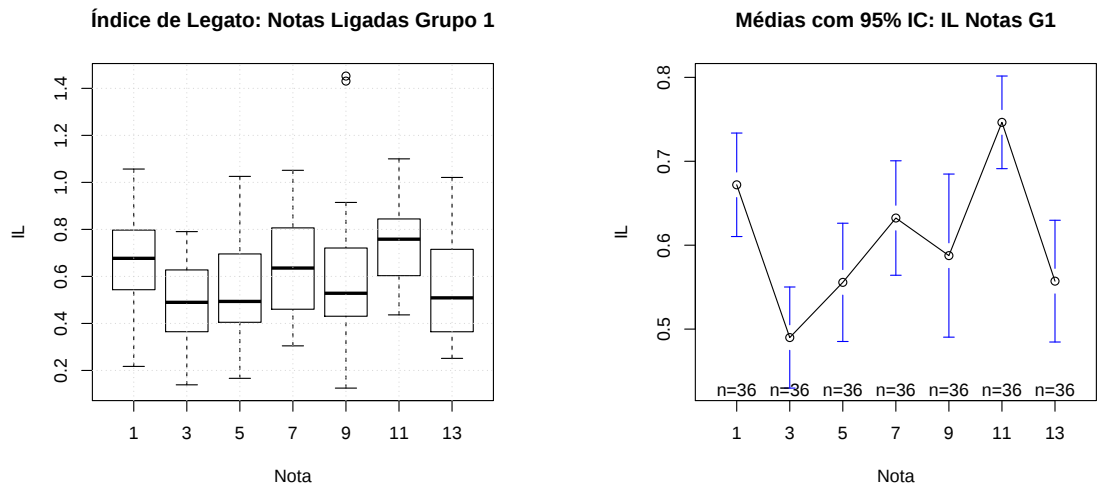


Figura 40 – Notas selecionadas para a **IL x Nota**: 1 a 13 do Grupo 1.

da figura contém os valores das estatísticas das comparações. Ainda na tabela da metade direita, as abreviações são explicadas a seguir: diff - diferenças entre as notas comparadas em dB; lwr - valor inferior do intervalo de confiança; upr - valor superior do intervalo

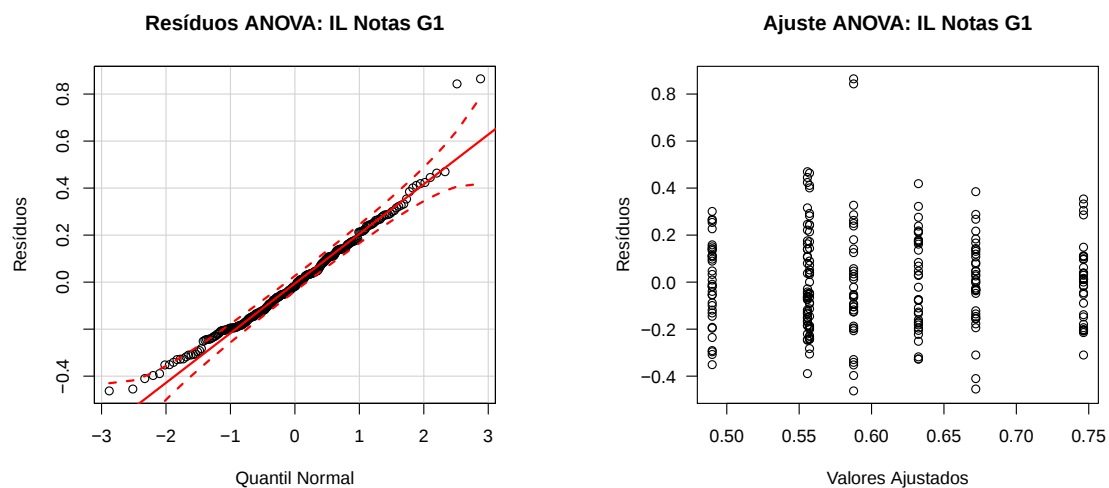


Figura 41 – Resíduos para análise e verificação das premissas da **ANOVA**. Na metade esquerda, o ajuste à distribuição normal. Na metade direita, a semelhança das variâncias dos resíduos dos ajustes.

de confiança; padj - valor p ajustado pela comparação múltipla. Min - valor mínimo das diferenças; Med - mediana das diferenças; Max - valor máximo das diferenças.

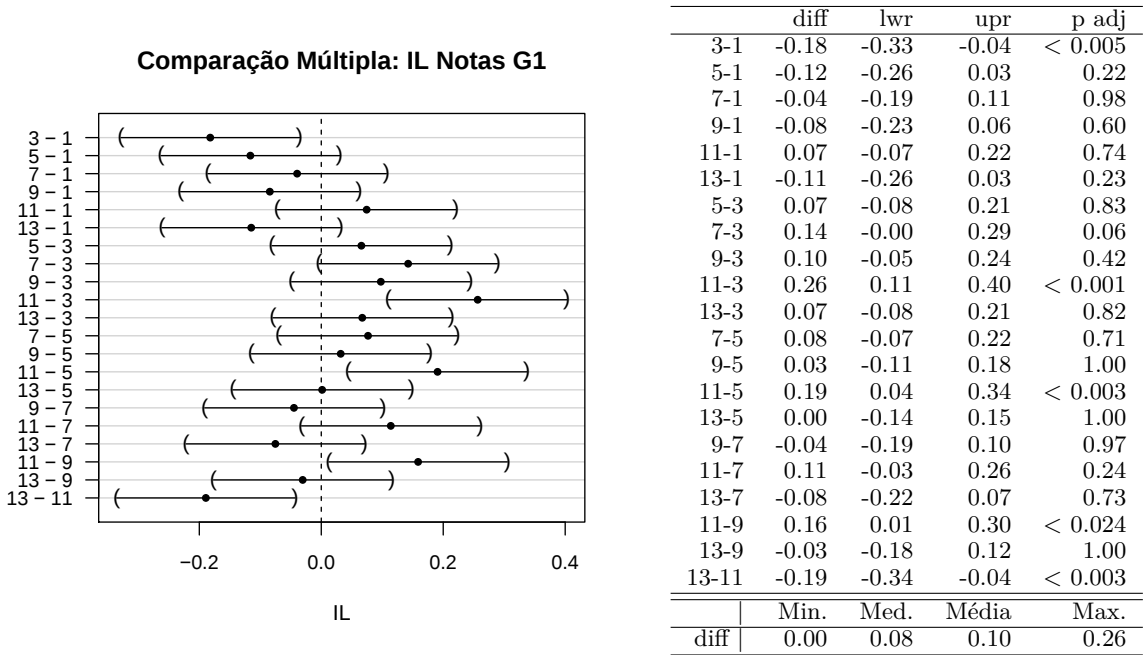


Figura 42 – Resultado da comparação múltipla **IL x Nota**.

IL x Clarinetista

Ainda com **IL** agora para o fator Clarinetista, a **ANOVA** refutou a hipótese das médias serem iguais (valor $p < 0,0292$), com as premissas atendidas de maneira fraca.

A comparação múltipla a posteriori apontou a única diferença significativa de **IL** entre Clarinetista como sendo 0,15.

O teste foi aplicado utilizando as notas 1 a 13 selecionadas do Grupo 1 (as mesmas notas do teste **IL x Nota**). Os **boxplots** e médias com intervalos de confiança a 95% para o fator Clarinetista estão na Figura 43. As estatísticas do teste estão na Tabela 9.

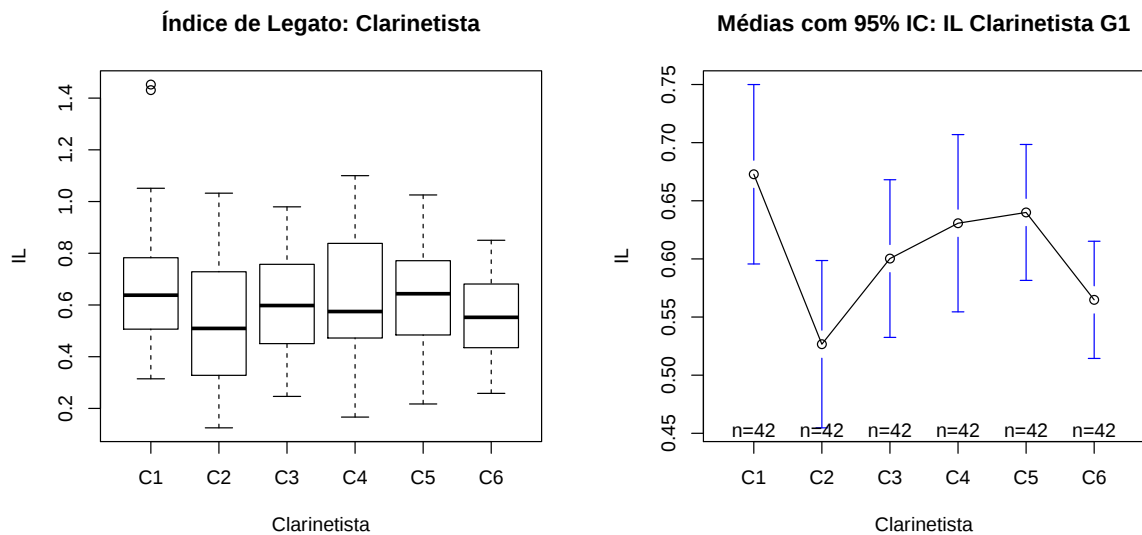


Figura 43 – **Boxplots** (à esquerda) e médias com intervalos de confiança a 95 % (à direita) para o fator Clarinetista utilizado na **ANOVA**.

Tabela 9 – Estatísticas da **ANOVA** para o teste **IL x Clarinetista**.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Clarinetista	5	0.60	0.12	2.54	0.0292
Residuals	246	11.61	0.05		

A Figura 44 contém a apresentação dos resíduos para análise e verificação das premissas do teste. Na metade esquerda da figura, por inspeção visual pode ser verificado o ajuste à distribuição normal, enquanto na metade da direita, pode ser verificada a semelhança das variâncias dos resíduos dos ajustes. Os resíduos foram avaliados pelo teste de normalidade *Shapiro-Wilk*, no qual foi identificado um comportamento não normal ao nível de confiança 95% ($W = 0,985$, valor $p = 0,008$).

Na Figura 45 podem ser vistos os dados da comparação múltipla a posteriori (teste de *Tukey*) para o teste **IL x Clarinetista**. Na figura, a metade esquerda contém um gráfico com as diferenças para cada comparação, com intervalos de confiança a 95%. A metade direita da figura contém os valores das estatísticas das comparações. Ainda na tabela da metade direita, as abreviações são explicadas a seguir: diff - diferenças entre as notas comparadas em dB; lwr - valor inferior do intervalo de confiança; upr - valor superior

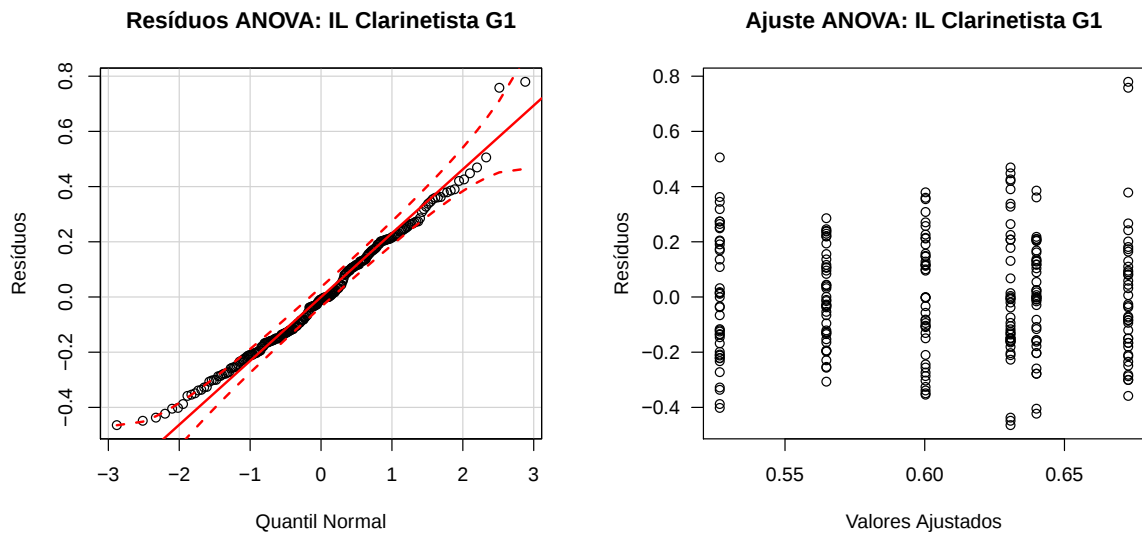


Figura 44 – Resíduos para análise e verificação das premissas da **ANOVA** para o teste da influência do fator Clarinetista na variável **IL**. Na metade esquerda, o ajuste à distribuição normal. Na metade direita, a semelhança das variâncias dos resíduos dos ajustes.

do intervalo de confiança; padj - valor p ajustado pela comparação múltipla. Min - valor mínimo das diferenças; Med - mediana das diferenças; Max - valor máximo das diferenças.

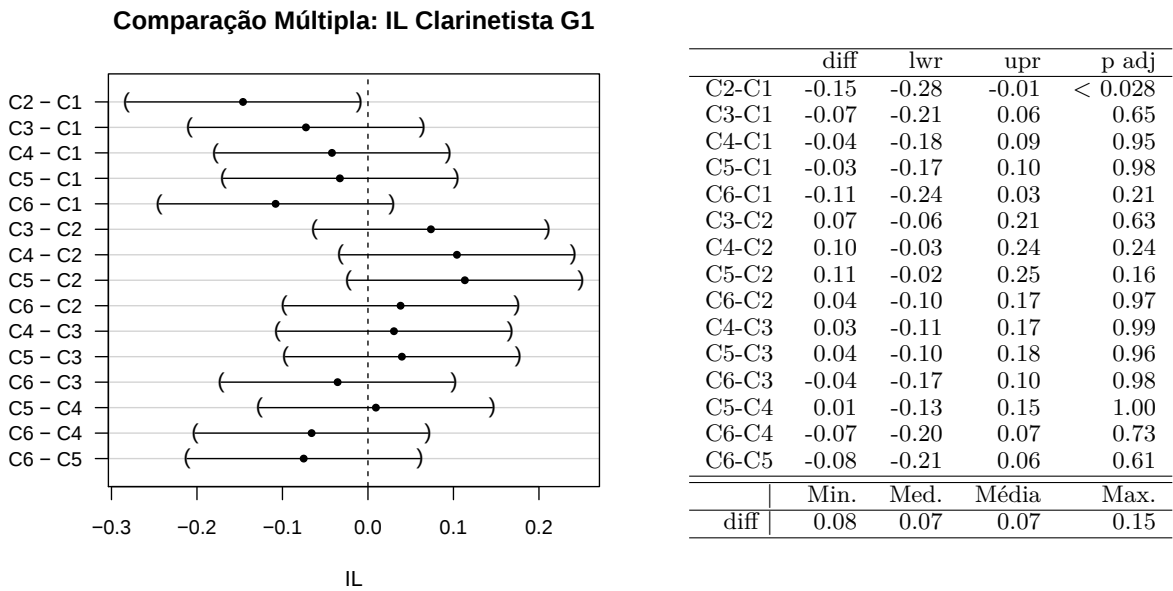


Figura 45 – Resultado da comparação múltipla para o fator Clarinetista na variável **IL**.

IL x ILh

Por fim, o teste t **IL x ILh**. O teste foi aplicado utilizando todas as notas ligadas dos dois grupos. Os **boxplots** por nota e médias com intervalos de confiança a 95% para as diferenças estão na Figura 46.

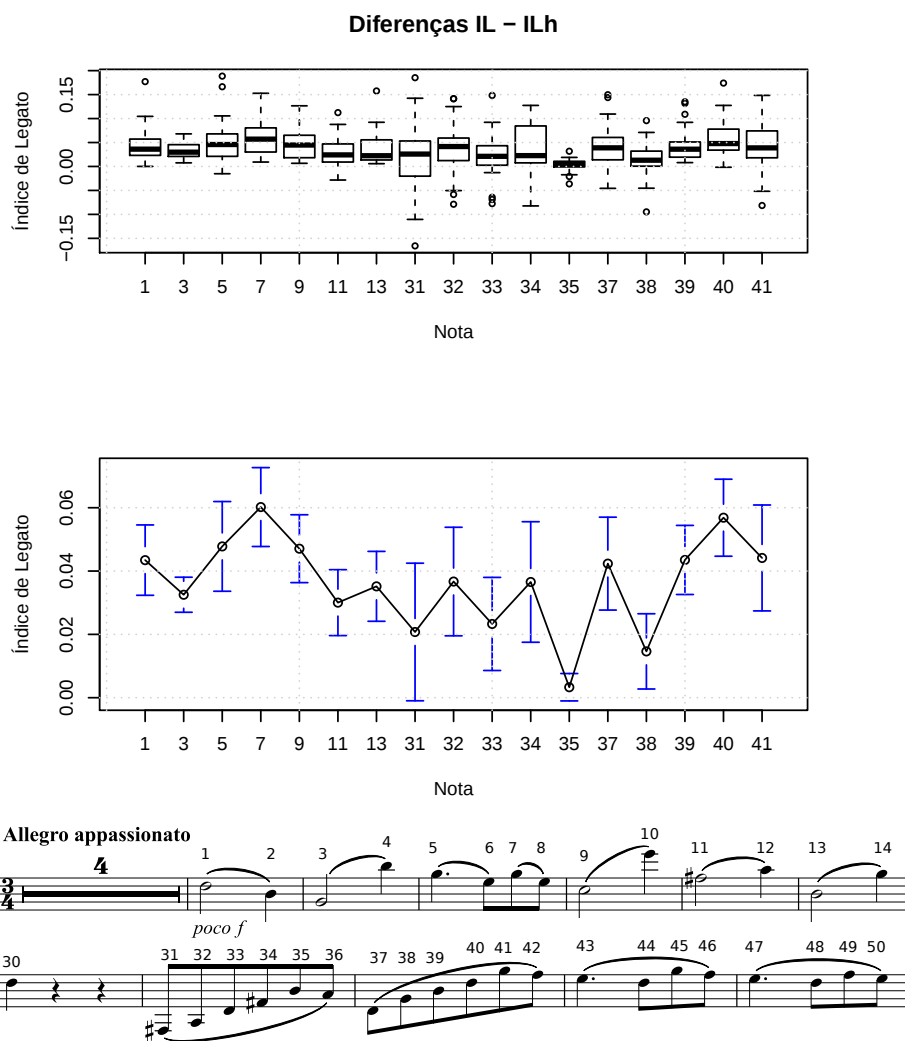


Figura 46 – *Boxplots* por nota e médias com intervalos de confiança a 95 % para as diferenças **IL - ILh** avaliadas no Grupo 1: Compassos 5 a 10, e no Grupo 2: compassos 17 e 18.

Houve diferença significativa entre **IL** e **ILh** (valor $p < 0,0001$), com diferença igual a 0,04dB. O teste *Shapiro-Wilk* refutou fortemente a hipótese dos resíduos serem normais ao nível de confiança de 95 % (valor $p = 1,059e-11$). Em casos em que o teste de normalidade dos resíduos é refutado, outras questões devem ser avaliada antes de se invalidar o resultado por não atender a premissa. Isso devido ao fato de o teste t ser robusto à não normalidade desde que o tamanho amostral seja grande e que haja simetria da distribuição dos resíduos (MONTGOMERY; RUNGER, 2009, pg. 177). Como o número

de amostras neste caso é grande ($n = 612$) e o *boxplot* dos resíduos é simétrico como pode ser verificado visualmente na Figura 47, as premissas do teste foram consideradas aceitas.

Tabela 10 – Estatísticas do teste t pareado para **IL x ILh**.

	Dif.	Df.	t-value	p-value	IC(95%) lwr	upr
Condição	0.036	611	21.41	2.2e-16	0.033	0.040

As estatísticas do teste estão na Tabela 10. A Figura 47 contém a apresentação dos resíduos para análise e verificação das premissas do teste **IL x ILh**. Na metade esquerda da figura, por inspeção visual pode ser verificada a simetria da distribuição pelo **boxplot** dos resíduos, enquanto na metade direita, o ajuste à distribuição normal. Os resíduos foram avaliados pelo teste de normalidade *Shapiro-Wilk*, no qual a hipótese nula dos resíduos serem normais foi refutada ao nível de confiança 95 % ($W = 0,961$, valor $p = 1,059e-11$).

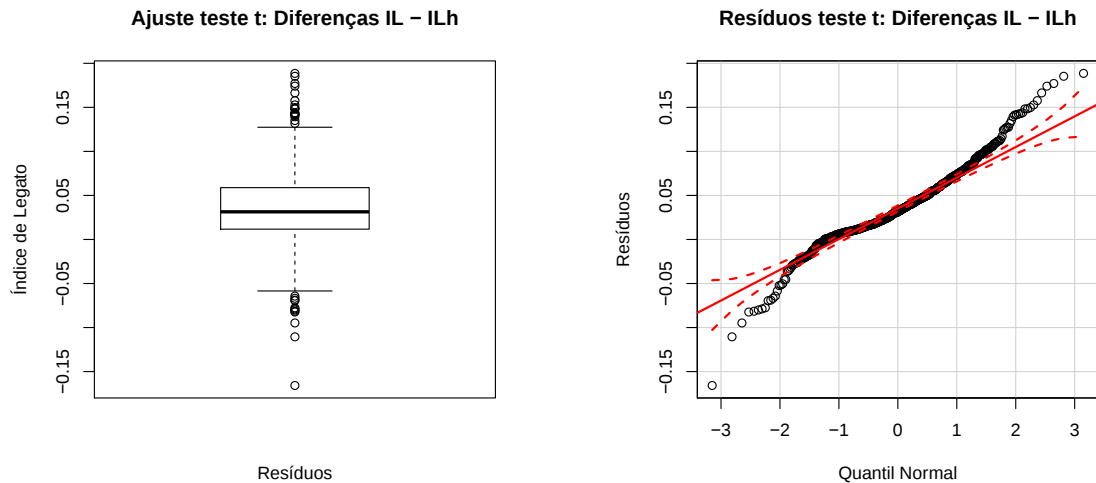


Figura 47 – Resíduos para análise e verificação das premissas do teste t pareado **IL x ILh**. Na metade direita, o ajuste à distribuição normal. Na esquerda, o **boxplot** dos resíduos.

6.7 Caracterização dos Ataques

Foram selecionados sete casos de transição entre notas consecutivas para aplicação das representações propostas para caracterizar ataques de notas. A seleção foi baseada em níveis de exigência técnica de cada transição, tipo de articulação e intervalo entre as notas. Cinco delas são transições *legato* (conectadas por um arco na partitura): notas 3 a 4, 9 a 10, 11 a 12, 13 a 14 e 21 a 22. Duas são transições articuladas (destacadas): notas 8 a 9 e 12 a 13.

Plano H2R2

Na Figura 48 (a) está mostrado o plano **H2R2** para os ataques de todas as execuções da nota 4 (nota com transição em condição de *legato*, partindo da nota 3). No plano referido, **E1** e **E2** indicam duas execuções em destaque de um mesmo clarinetista. Na Figura 48 (b) e (c) estão mostradas as formas de onda, espectrogramas e fluxo espectral para as execuções indicadas por **E2** e **E1**, onde pode ser visualmente confirmada a distância entre as duas execuções na representação do plano **H2R2**. A execução **E2** exibe conteúdo harmônico (**H1**) mais baixo do que a execução **E1**. As formas de onda e espectrogramas mostrados nessa figura foram extraídos do software Sonic Visualizer (CANNAM; LANDONE; SANDLER, 2010), enquanto o Fluxo Espectral (linha branca contínua) foi extraído usando o plugin MIR.EDU (SALAMON; GÓMEZ, 2014). Nesta seção, as elipses dos gráficos são delimitadores de concentração de dados estimadas como o contorno de probabilidade normal bivariada com nível de confiança igual a 0,8. Os pontos sólidos associados nessas figuras representam o centro das elipses.

Na figura seguinte (Figura 49), estão mostrados os planos **H2R2** para os ataques das notas 10, 12, 14 e 22, cujas transições são também em condição de *legato* partindo da nota anterior. Na Figura 50, por sua vez, estão mostrados os planos **H2R2** para os ataques das notas 9 e 13 que, nesses dois casos, apresentam transições articuladas partindo da nota anterior.

A tabela 11 mostra os resultados do teste **MANOVA** considerando os clarinetistas (C1 a C6) como fator de comparação. São apresentadas as variâncias das variáveis *R2* e *H2*, o par de clarinetistas comparados com Contraste Máximo e, finalmente, os valores-p, para cada uma das notas selecionadas (4, 9, 10, 12, 13, 14 e 22).

Transições de notas em *Legato* na clarineta, como nas notas 4, 10, 12, 14 e 22 (que devem ser executadas ligadas à nota anterior), são alcançadas mantendo a pressão de sopro ao longo da transição, ou seja, a vibração da palheta não deve ser interrompida nem por qualquer diminuição da pressão de sopro nem tocando a palheta com a língua. A qualidade das transições *legato* entre notas consecutivas depende bastante da capacidade de manter a estabilidade da pressão de sopro. Além disso, à medida que o intervalo entre as notas aumenta, a execução da transição pode impor demandas técnicas adicionais que podem variar de acordo com o tamanho do salto. Pode exigir precisão nos movimentos de sincronização que envolvem mais dedos, às vezes de mãos diferentes. Além disso, alguns saltos também podem apresentar habilidades extras para sincronizar o movimento dos dedos, com a adaptação da pressão bucal à alteração da vibração do tubo do instrumento da próxima nota, especialmente quando a transição envolve a alteração do registro.

A transição que abrange o ataque da nota 22 (Figura 49 a) envolve um semi-tom descendente que requer apenas uma chave para fechar um furo, acionada por uma mola de

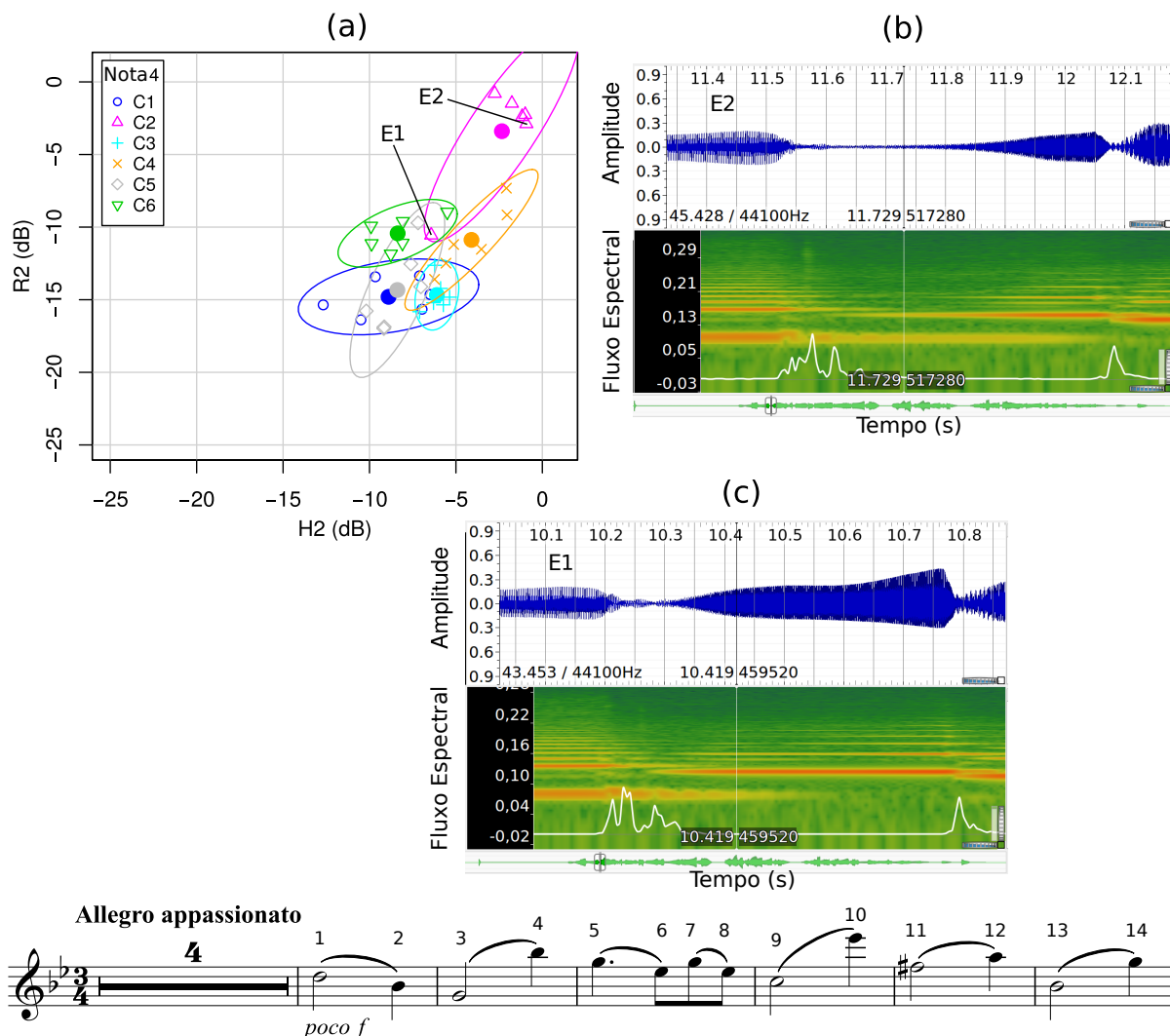


Figura 48 – (a) Plano **H2R2** para os ataques de todas as execuções da nota 4 (nota com transição em condição de *legato*, partindo da nota 3). **E1** e **E2** indicam duas execuções em destaque de um mesmo clarinetista. Em (b) e (c) estão mostradas as formas de onda, espectrogramas e fluxo espectral para **E2** e **E1**.

aço, enquanto a transição que abrange o ataque da nota 12 (Figura 49 b) requer levantar um dedo da mão esquerda simultaneamente com um dedo da mão direita, exigindo precisão na sincronização do movimento de ambos os dedos. Os gráficos mostram um conteúdo harmônico mais alto no ataque da nota 22, portanto, menos reverberação ($H2$) e menos ruído ($R2$), o que pode indicar uma qualidade *legato* mais alta nessa transição. Também foi observada menor variação de ambas as dimensões no ataque da nota 22, tanto em relação a todas as execuções, quanto as execuções dentro dos grupos de cada participante. Isso sugere que os clarinetistas são capazes de obter maior sucesso na execução dessa transição.

Os *legatos* que abrangem o ataque da nota 14 (Figura 49 c) e o da nota 4 (Figura 48 a) envolvem um intervalo de sexta maior ascendente e uma décima menor, respectivamente, mudando do registro grave ao registro alto do instrumento. A dificuldade nessas transições consiste não apenas na ação múltipla simultânea do dedo, mas também em manter estável

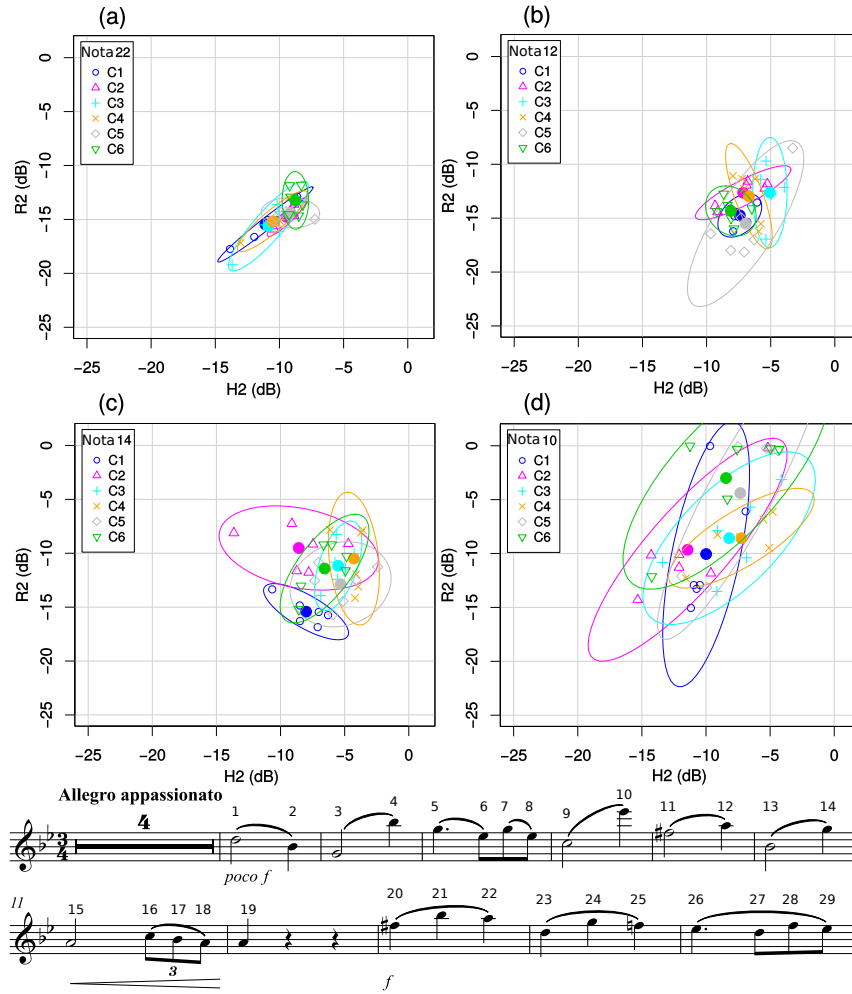


Figura 49 – Planos **H2R2** para os ataques de todas as execuções das notas 10, 12, 14 e 22, cujas transições são em condição de *legato* partindo da nota anterior.

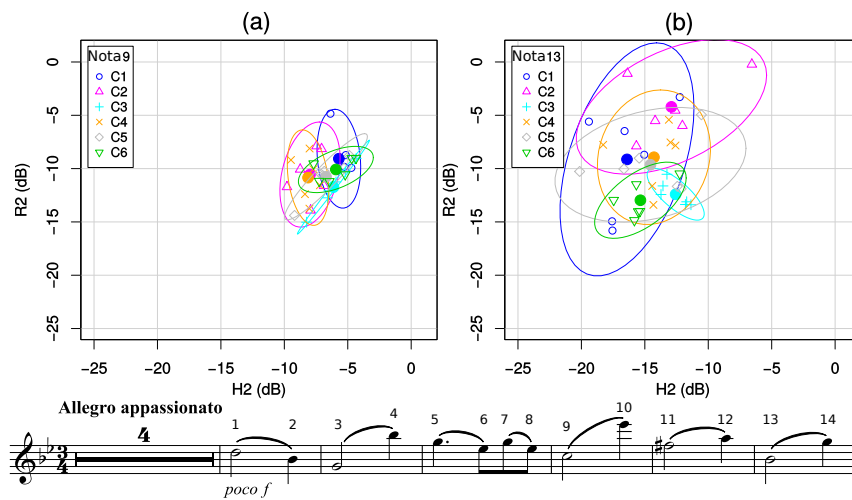


Figura 50 – Planos **H2R2** para os ataques de todas as execuções das notas 9 e 13 que, nesses dois casos, apresentam transições articuladas (não ligadas) partindo da nota anterior.

Tabela 11 – Resultados do teste **MANOVA** (*Multidimensional Analysis of Variance*) considerando “clarinetista” (C1 a C6) como fator. São apresentadas as variâncias das variáveis *R2* e *H2*, o par de clarinetistas comparados com Contraste Máximo e, finalmente, os valores-p, para cada uma das notas selecionadas.

	Variância						Contraste Max. in dB (par)	MANOVA P valor
	C1	C2	C3	C4	C5	C6		
Nota 04							4.4 (C4-C5)	< 0.001*
R2	1.5	12.9	1.3	5.2	8.0	1.2	-	-
H2	6.0	4.5	0.4	3.3	1.7	2.6	-	-
Nota 09							2.1 (C1-C4)	0.007
R2	4.8	5.3	4.3	4.4	3.6	1.0	-	-
H2	0.7	1.3	1.8	0.6	2.4	1.9	-	-
Nota 10							4.8 (C1-C2)	0.015
R2	33.6	23.7	14.2	4.7	40.9	23.6	-	-
H2	2.5	13.3	10.0	7.1	9.2	14.3	-	-
Nota 12							2.4 (C3-C5)	0.013
R2	0.9	1.4	5.8	5.3	13.1	1.2	-	-
H2	0.7	3.1	0.4	0.7	4.6	0.9	-	-
Nota 13							5.6 (C2-C6)	< 0.001
R2	26.5	8.8	1.3	8.8	6.4	2.8	-	-
H2	6.2	12.5	1.12	4.4	12.6	2.9	-	-
Nota 14							4.3 (C4-C5)	< 0.001
R2	1.5	3.4	3.8	8.4	3.5	5.6	-	-
H2	2.4	8.6	0.8	0.9	3.4	2.7	-	-
Nota 22							2.3 (C2-C3)	0.027
R2	2.7	0.7	3.8	1.8	0.4	1.6	-	-
H2	3.1	0.5	2.1	1.7	1.2	0.2	-	-

* O grupo do clarinetista C2 foi excluído do teste **MANOVA** para a nota 4, devido à forte rejeição no teste de normalidade.

a pressão de sopro, bem como em sincronizar o acoplamento da pressão da boca com o movimento do dedo. Comparando as execuções de ambas as transições no plano **H2R2**, podemos observar uma variação maior de ambas as dimensões no ataque da nota 4, o que pode sugerir que os clarinetistas tenham mais dificuldade em obter um *legato* de melhor qualidade para essa transição. De fato, além de envolver um intervalo maior, a transição do G4 para o Bb5 (nota 4) requer ações complexas de fechamento simultâneo de orifícios e pressionamento de chaves com os dedos das duas mãos.

A mais exigente delas é a transição que abrange o ataque da nota 10 (Figura 49 d). Como na nota 4, a nota 10 é alcançada com um intervalo de décima menor ascendente, mas passando de uma nota baixa do registro alto para o registro *altissimo*. Essa passagem impõe uma dificuldade técnica adicional devido à instabilidade acústica do instrumento no registro *altissimo*, onde os clarinetistas devem sempre usar vários dedilhados alternativos, a fim de corrigir a afinação, o que impõe uma demanda adicional na sincronização do acoplamento da pressão bucal com o dedo do movimento. É comum que os clarinetistas não consigam executar corretamente essa passagem, independentemente do seu nível de experiência, como algumas das execuções observadas.

Comparando as execuções dos ataques de ambas as transições articuladas (notas 9 e 13) da Figura 50 (a) e (b), respectivamente, podemos observar que o ataque da nota 9 exibe um conteúdo harmônico de reverberação da nota de origem (**H2**) maior no geral do que o

da nota 13, exceto por uma exceção. Isso pode ser explicado por uma articulação maior imposta pela alteração do registro (registro alto para baixo) e maior intervalo envolvido na transição que abrange o ataque da nota 13. Como dito anteriormente, as alterações no registro exigem mais sincronização das ações dos dedos e adaptação da pressão da boca, o que pode resultar em um tempo de articulação mais longo entre as notas. Isso também pode explicar a maior variação das execuções da nota 13, provavelmente refletindo as tentativas dos clarinetistas de executar a articulação com menos interrupção do som “buraco”) entre as notas.

Perfil Harmônico do Ataque

Na sequência, são apresentados os resultados para o **Perfil Harmônico do Ataque (PHA)**, proposto nesta tese para caracterização da estrutura harmônica de ataques na clarineta. Foram utilizados os ataques das mesmas sete notas selecionadas para análise no plano **H2R2** (4, 9, 10, 12, 13, 14 e 22). Lembrando que, cinco delas são transições *legato* (ligadas à nota anterior, com indicação de um arco na partitura): transições 3 a 4, 9 a 10, 11 a 12, 13 a 14 e 21 a 22. Duas são transições articuladas: notas 8 a 9 e 12 a 13.

Para a análise dos **PHAs**, foi avaliada a capacidade de discriminação de músicos utilizando comparações entre perfis e, como os perfis são tridimensionais, as comparações foram feitas através de suas projeções em duas dimensões. Foram calculados Coeficientes de Correlação de Pearson entre perfis de diferentes execuções através das projeções no plano Harmônico-Magnitude (**HxM**) e no plano Harmônico-Tempo (**HxT**). Para cada uma das duas projeções, foram calculados coeficientes de correlação entre perfis de execuções diferentes do mesmo músico, gerando assim, seis grupos chamados de *Intra* com 15 elementos cada, resultantes da combinação 2 a 2 das 6 execuções de cada músico. Além da comparação entre os 6 grupos *Intras*, foi definido também um grupo chamado de *Inter*, contendo os coeficientes de correlação entre as projeções dos **PHAs** de cada clarinetista com os outros 5 colegas. Como são 6 execuções, para cada nota, em cada uma das duas projeções **HxT** e **HxM**, é estabelecido um grupo *Inter* de 30 elementos (6 execuções vezes 5 comparações). Para cada nota e projeção do **PHA**, as médias dos coeficientes de correlação desses 6 grupos *Intras* de 15 elementos e o grupo *Inter* de 30 elementos, foram comparadas usando o teste **ANOVA** (*Analysis of Variance*). As médias com os respectivos intervalos de confiança a 95% para os 7 grupos, a média geral e o valor p estão nas Figuras 52, 54, 56, 58, 60, 62 e 64.

Usando os resultados dos planos **H2R2**, foi possível identificar ataques mal sucedidos. Foi verificado que esses mal sucedidos atrapalham as comparações por correlações entre projeções do **PHA**. Por esse motivo, foi definida uma heurística para detectar esses ataques, estabelecendo um limiar de corte de -3 dB em ambas as variáveis **H2(dB)** e

R2(dB), para excluir execuções com pelo menos uma variável acima desse limiar. Na Figura 51, são mostradas as quatro notas nas quais pelo menos um ataque mal sucedido foi detectado e excluído das comparações entre projeções do **PHA**. Os ataques excluídos estão listados no canto esquerdo inferior de cada plano **H2R2** da figura, com a simbologia de “C” para clarinetista e “E” para execução.

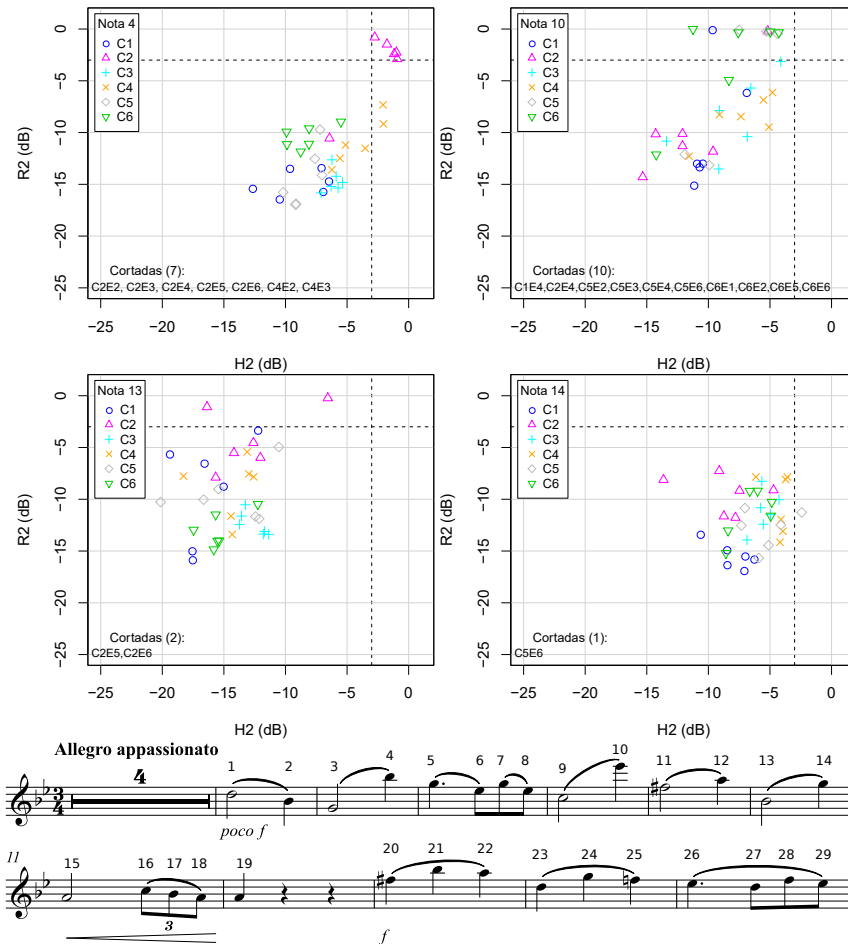


Figura 51 – Notas excluídas da análise das projeções do **PHA**, utilizando um limiar de corte de -3 dB, em ambas as variáveis do plano **H2R2**, para identificar ataques mal sucedidos. Os ataques excluídos estão listados no canto esquerdo inferior de cada plano **H2R2** da figura, com a simbologia de “C” para clarinetista e “E” para execução.

Nota 4

Na Figura 52, estão mostradas as estatísticas das comparações *Intra* e *Inter* grupos para a nota 4, executada pelos seis clarinetistas (C1 a C6), utilizando coeficientes de correlação de *Pearson* entre as projeções do Perfil Harmônico do Ataque (**PHA**). Lembrando que as projeções são: Harmônico por Tempo (**HxT**) e Harmônico por Magnitude (**HxM**). Os valores das colunas N (mostrados na figura) indicam quantos coeficientes de correlação foram calculados para cada grupo (combinações 2 a 2), e podem variar caso

haja corte de execuções devido à ataques mal sucedidos. No caso desta nota 4, houve 5 cortes de execução no grupo C2, e 2 cortes no grupo C4. Com os cortes, restou apenas uma execução no grupo C2, resultando em 0 comparações, e 4 execuções no grupo C4, resultando em 6 comparações. Os valores das médias das correlações estão nas colunas à direita, com Intervalos de Confiança (I.C.) a 95% entre parênteses. Nas colunas ao centro, para as duas projeções, encontram-se a média geral *Intra* e o valor p da **ANOVA** (teste para verificação de existência ou não de diferença entre as médias dos grupos).

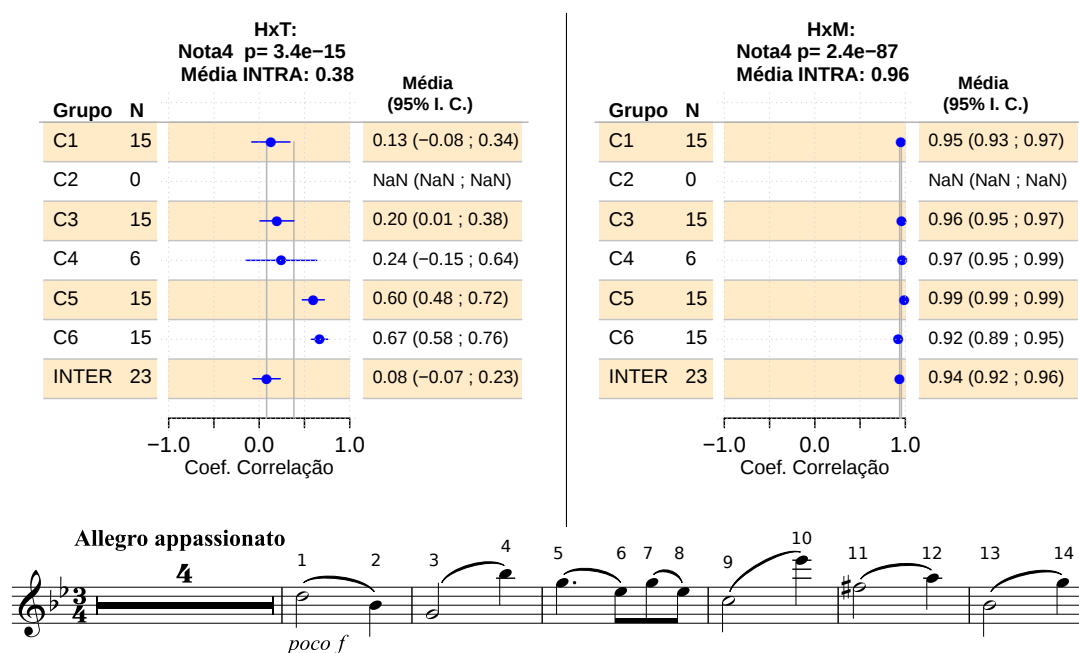


Figura 52 – Estatísticas das comparações *Intra* e *Inter* grupos (de tamanho N) utilizando coeficientes de correlação de *Pearson* entre as projeções do **PHA** - Harmônico por Tempo (**HxT**) e por Magnitude (**HxM**) - clarinetistas C1 a C6 executando a nota 4. Médias das correlações com Intervalos de Confiança (I.C.) a 95% nas colunas à direita. Média geral *Intra* e o valor p da **ANOVA** ao centro.

Os resultados indicaram, para ambas as projeções **HxT** e **HxM** da nota 4, diferença significativa ao nível de confiança 95% (valor $p < 0,05$) entre as médias dos grupos de correlações. As diferenças entre as médias *Intra* e *Inter* foram de 0.3 (**HxT**) e 0.02 (**HxM**). A média *Intra* e a média *Inter* dos coeficientes de correlação para a projeção **HxT** (0,38 e 0,08) foram bem menores do que as mesmas da projeção **HxM** (0,96 e 0,94). Portanto, as médias *Intra* e *Inter* para **HxT** indicam correlação fraca e baixa, com dois clarinetistas (C5 e C6) apresentando médias de correlações moderadas (entre 0,5 e 0,7). No caso de **HxM**, assim como as médias *Intra* e *Inter*, todos os grupos apresentaram médias de correlações muito fortes (acima de 0,9). A transição para a nota 4 é em condição de *Legato* em um intervalo ascendente de décima menor, saltando do registro baixo do instrumento para o alto. Na Figura 53 estão mostradas as projeções **HxT** (1a até 1f) e **HxM** (2a até 2f) de todas as 6 execuções realizadas pelos 6 clarinetistas.

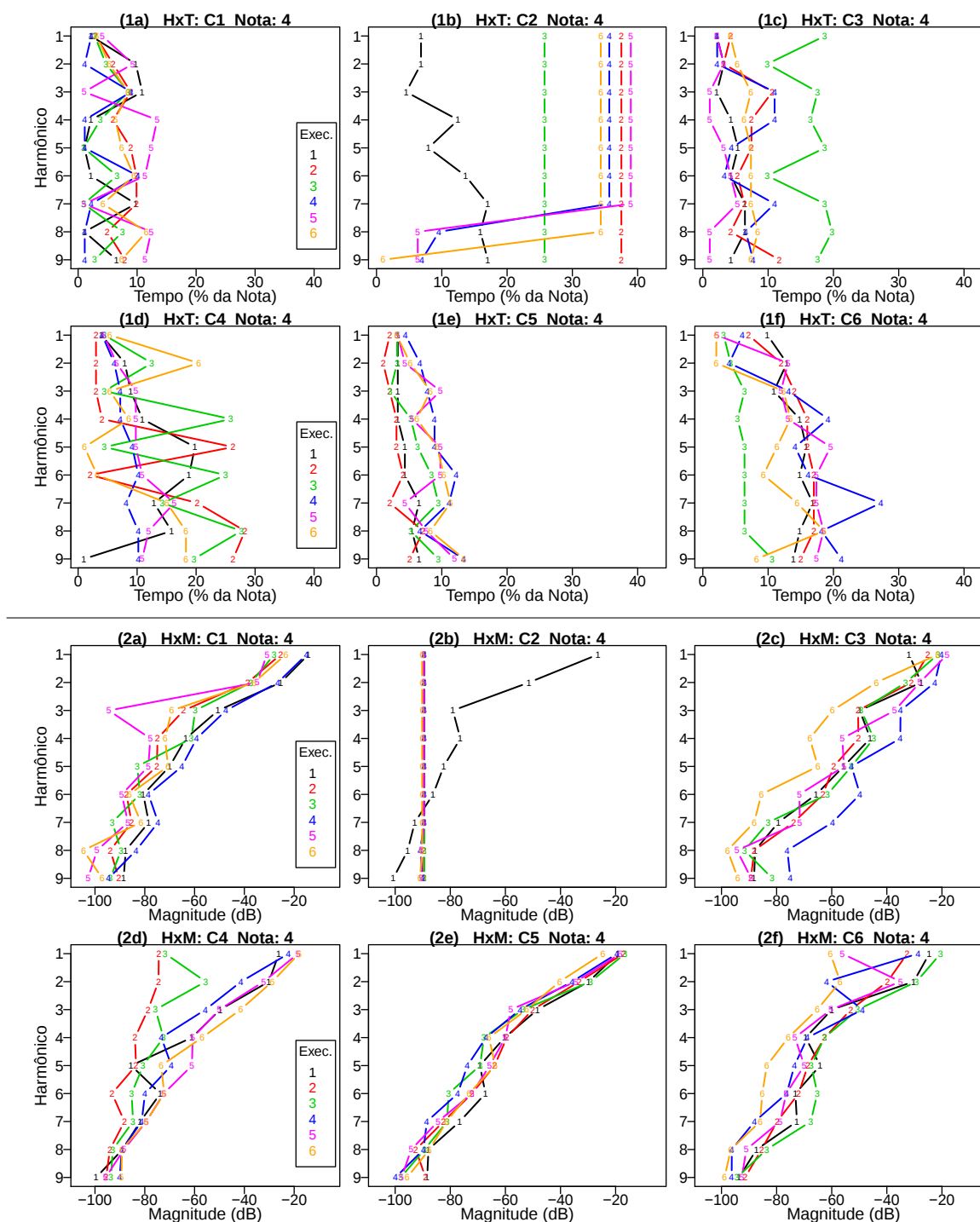


Figura 53 – Projeções do PHA, nota 4, clarinetistas C1 a C6: Harmônico por Tempo (HxT) em (1a) até (1f) e Harmônico por Magnitude (HxM) em (2a) até (2f).

Nota 9

Na Figura 54, estão mostradas as estatísticas das comparações *Intra* e *Inter* grupos para a nota 9, executada pelos seis clarinetistas (C1 a C6), utilizando coeficientes de correlação de *Pearson* entre as projeções HxT e HxM do PHA. Os valores das colunas N (mostrados na figura) indicam quantos coeficientes de correlação foram calculados para

cada grupo, e podem variar caso haja corte de execuções devido à ataques mal sucedidos. No caso desta nota 9, não houve nenhum corte de execução nos 6 grupos de clarinetistas, resultando em 15 comparações (combinações 2 a 2 das 6 execuções) em cada grupo *Intra*, e 30 comparações no grupo *Inter* (5 comparações de cada tomada vezes 6 tomadas). Os valores das médias das correlações estão nas colunas à direita, com Intervalos de Confiança (I.C.) a 95% entre parênteses. Nas colunas ao centro, para as duas projeções, encontram-se a média geral *Intra* e o valor p da **ANOVA** (teste para verificação de existência ou não de diferença entre as médias dos grupos).

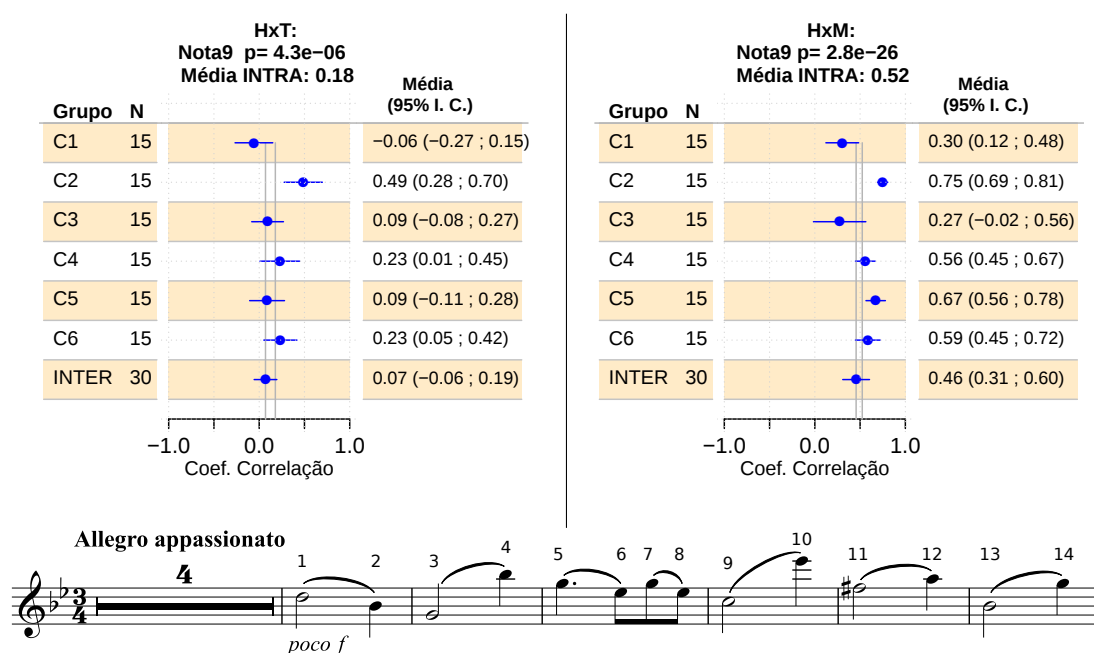


Figura 54 – Estatísticas das comparações *Intra* e *Inter* grupos (de tamanho N) utilizando coeficientes de correlação de *Pearson* entre as projeções do **PHA** - Harmônico por Tempo (**HxT**) e por Magnitude (**HxM**) - clarinetistas C1 a C6 executando a nota 9. Médias das correlações com Intervalos de Confiança (I.C.) a 95% nas colunas à direita. Média geral *Intra* e o valor p da **ANOVA** ao centro.

Os resultados indicam, para ambas as projeções **HxT** e **HxM** da nota 9, diferença significativa ao nível de confiança 95% (valor $p < 0,05$) entre as médias dos grupos de correlações. As diferenças entre as médias *Intra* e *Inter* foram de 0.11 (**HxT**) e 0.06 (**HxM**). A média *Intra* e a média *Inter* dos coeficientes de correlação para a projeção **HxT** (0,18 e 0,07) foram menores do que as mesmas da projeção **HxM** (0,52 e 0,46). Portanto, as médias *Intra* e *Inter* para **HxT** indicam correlação baixa, com apenas o clarinetista C2 apresentando médias de correlações próxima de moderada (entre 0,5 e 0,7). No caso de **HxM**, as médias *Intra* e *Inter* ficaram muito próximas a 0,5 (moderada), com dois grupos (C1 e C3) apresentando médias de correlações baixa (próximas de 0,3) e C2 com média de correlações forte (0,75). A transição para a nota 9 é articulada em um intervalo descendente, dentro do mesmo registro do instrumento. Na Figura 55 estão mostradas as

projeções **HxT** (1a até 1f) e **HxM** (2a até 2f) de todas as 6 execuções realizadas pelos 6 clarinetistas.

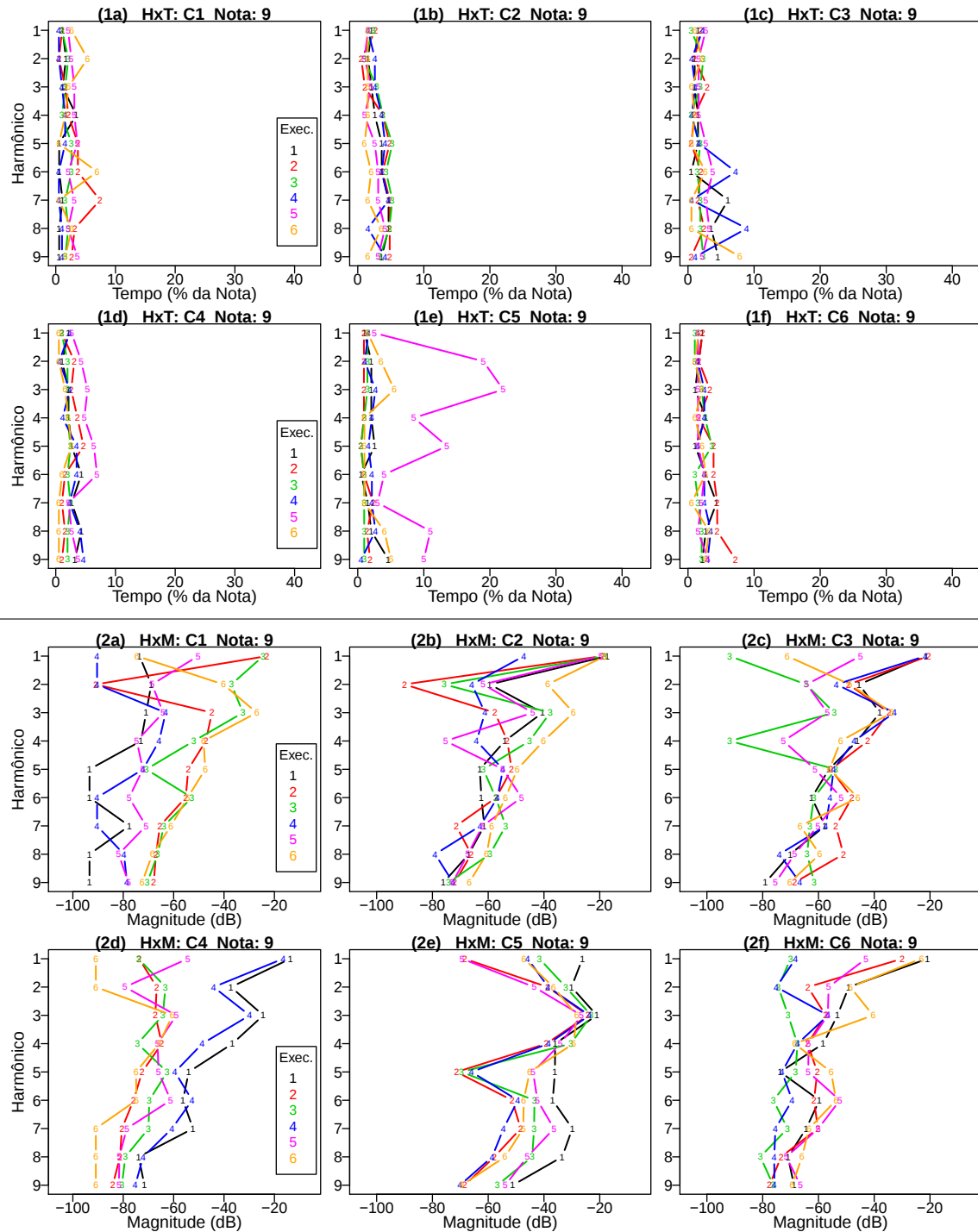


Figura 55 – Projeções do **PHA**, nota 9, clarinetistas C1 a C6: Harmônico por Tempo (**HxT**) em (1a) até (1f) e Harmônico por Magnitude (**HxM**) em (2a) até (2f).

Nota 10

Na Figura 56, estão mostradas as estatísticas das comparações *Intra* e *Inter* grupos para a nota 10, executada pelos seis clarinetistas (C1 a C6), utilizando coeficientes de correlação de *Pearson* entre as projeções **HxT** e **HxM** do **PHA**. Os valores das colunas N (mostrados na figura) indicam quantos coeficientes de correlação foram calculados para cada grupo, e podem variar caso haja corte de execuções devido à ataques mal sucedidos. No caso desta nota 10, houve um corte de execução nos grupos C1 e C2, e 4 cortes em C5 e C6, resultando em 10 comparações em C1 e C2, e apenas 1 comparação em C5 e C6. Os valores das médias das correlações estão nas colunas à direita, com Intervalos de Confiança (I.C.) a 95% entre parênteses. Nas colunas ao centro, para as duas projeções, encontram-se a média geral *Intra* e o valor p da **ANOVA** (teste para verificação de existência ou não de diferença entre as médias dos grupos).

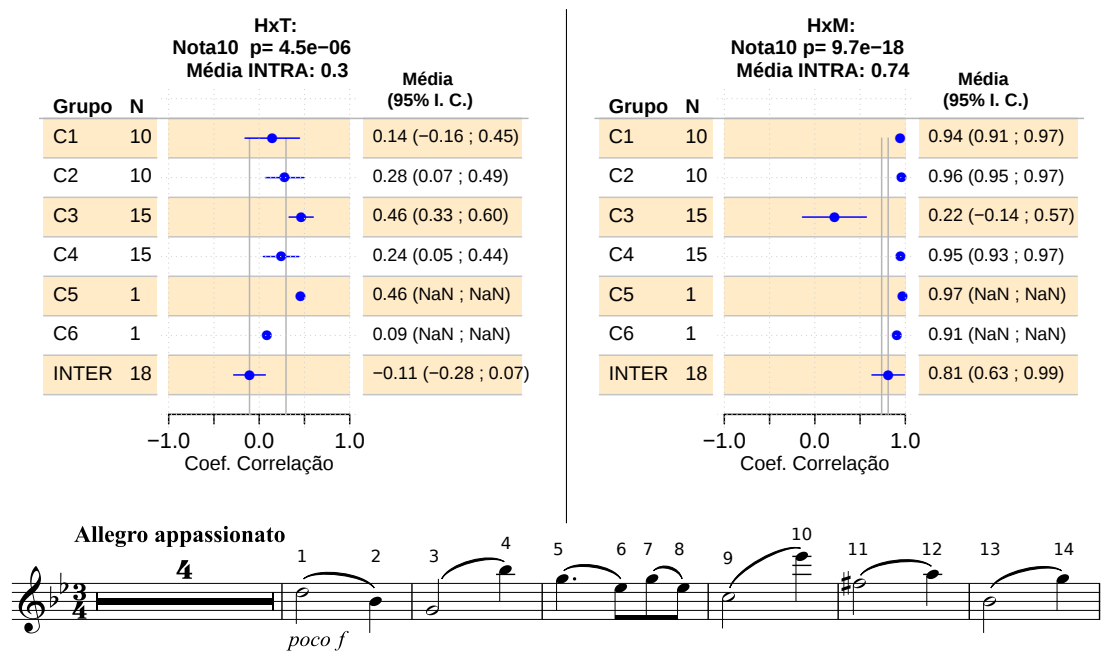


Figura 56 – Estatísticas das comparações *Intra* e *Inter* grupos (de tamanho N) utilizando coeficientes de correlação de *Pearson* entre as projeções do **PHA** - Harmônico por Tempo (**HxT**) e por Magnitude (**HxM**) - clarinetistas C1 a C6 executando a nota 10. Médias das correlações com Intervalos de Confiança (I.C.) a 95% nas colunas à direita. Média geral *Intra* e o valor p da **ANOVA** ao centro.

Os resultados para **HxT** e **HxM** da nota 10 indicam diferença significativa ao nível de confiança 95% ($p < 0,05$) entre as médias dos grupos de correlações. As diferenças entre as médias *Intra* e *Inter* foram de 0.41 (**HxT**) e -0.07 (**HxM**). A média *Intra* e a média *Inter* dos coeficientes de correlação para a projeção **HxT** (0,3 e -0,11) foram bem menores do que as mesmas da projeção **HxM** (0,74 e 0,81). Portanto, as médias *Intra* e *Inter* para **HxT** indicam correlação baixa, com médias de C3 e C4 pouco maiores (fracas, entre 0,3 e 0,5). No caso de **HxM**, as médias *Intra* e *Inter* foram fortes (entre 0,7 e 0,9), com apenas

C3 apresentando média baixa ($< 0,3$), puxando a média *Intra* para baixo, enquanto todos os outros grupos de clarinetistas obtiveram médias de correlações muito fortes ($> 0,9$). A transição para a nota 10 é em condição de *Legato*, em um intervalo ascendente de décima menor, variando do registro alto do instrumento para o altíssimo. Esta é a transição de maior dificuldade de execução dentre todas as avaliadas. Na Figura 57 estão mostradas as projeções **HxT** (1a até 1f) e **HxM** (2a até 2f) de todas as 6 execuções dos 6 clarinetistas.

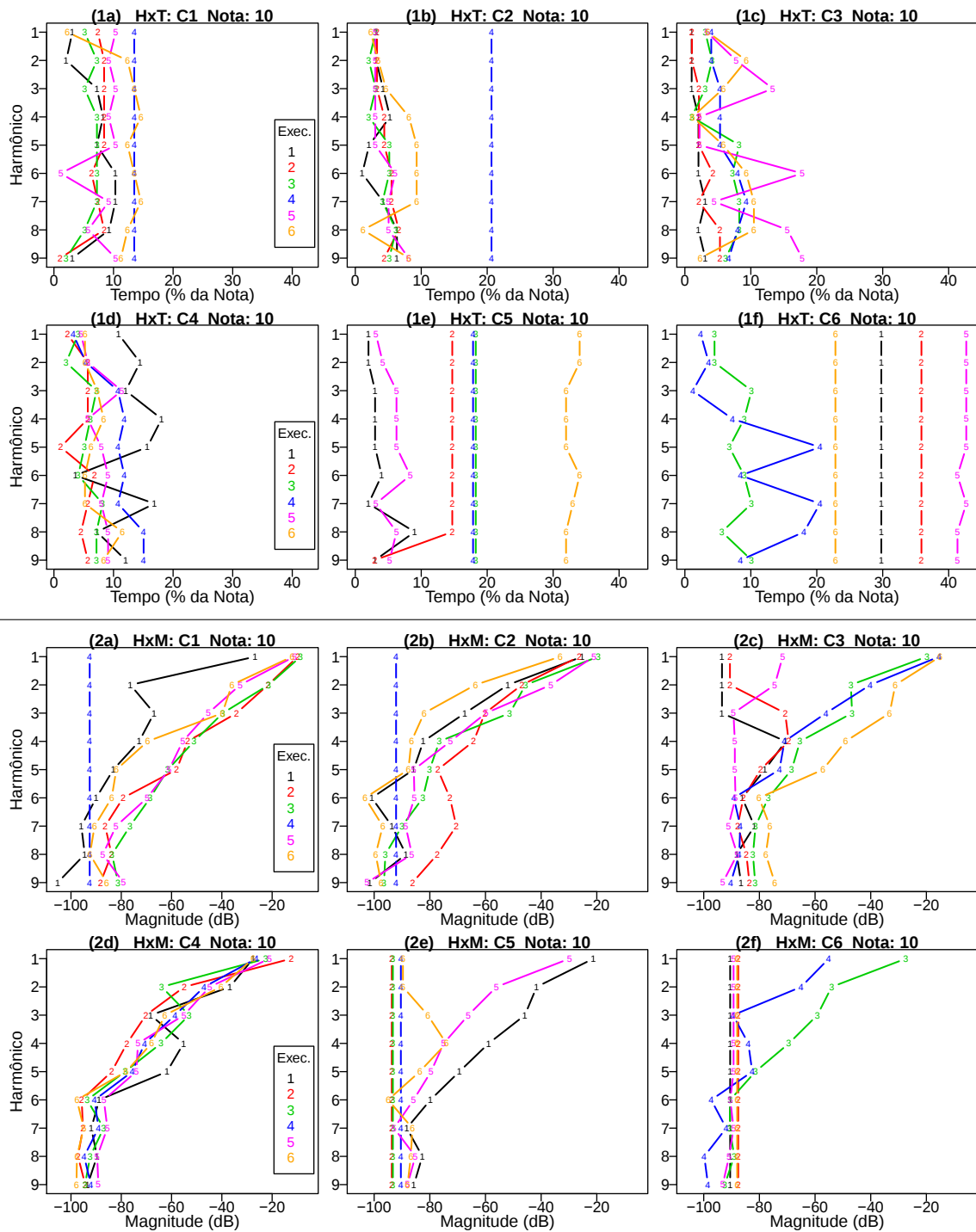


Figura 57 – Projeções do **PHA**, nota 10, clarinetistas C1 a C6: Harmônico por Tempo (**HxT**) em (1a) até (1f) e Harmônico por Magnitude (**HxM**) em (2a) até (2f).

Nota 12

Na Figura 58, estão mostradas as estatísticas das comparações *Intra* e *Inter* grupos para a nota 12, executada pelos seis clarinetistas (C1 a C6), utilizando coeficientes de correlação de *Pearson* entre as projeções **HxT** e **HxM** do **PHA**. Os valores das colunas N (mostrados na figura) indicam quantos coeficientes de correlação foram calculados para cada grupo, e podem variar caso haja corte de execuções devido à ataques mal sucedidos. No caso desta nota 12, não houve nenhum corte de execução nos 6 grupos de clarinetistas, resultando em 15 comparações (combinações 2 a 2 das 6 execuções) em cada grupo *Intra*, e 30 comparações no grupo *Inter* (5 comparações de cada tomada vezes 6 tomadas). Os valores das médias das correlações estão nas colunas à direita, com Intervalos de Confiança (I.C.) a 95% entre parênteses. Nas colunas ao centro, para as duas projeções, encontram-se a média geral *Intra* e o valor p da **ANOVA** (teste para verificação de existência ou não de diferença entre as médias dos grupos).

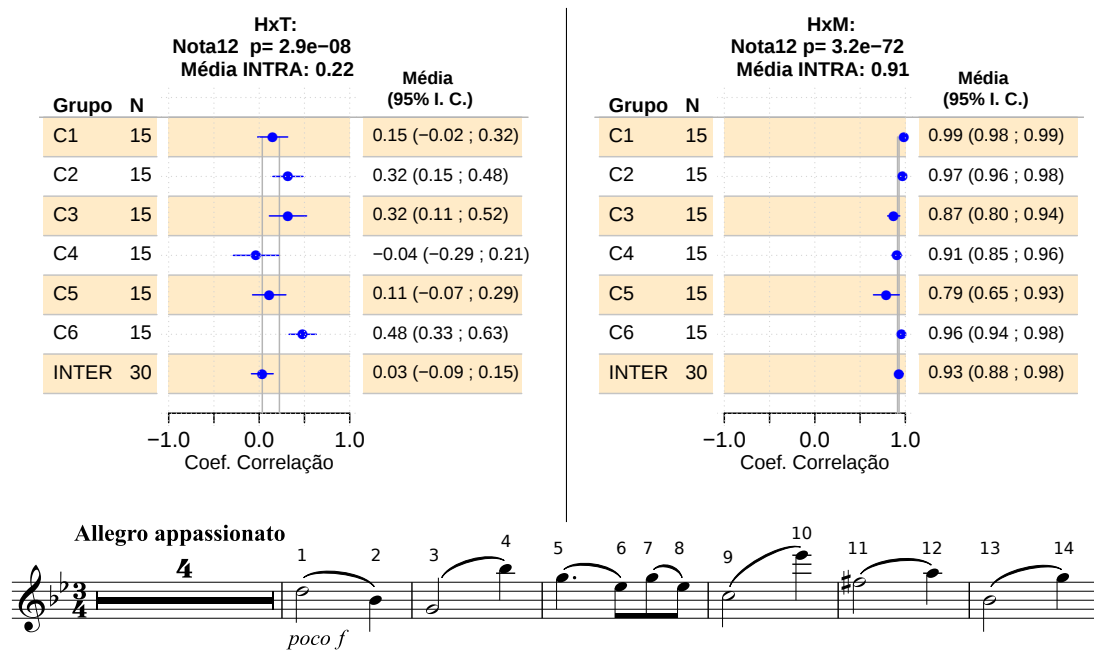


Figura 58 – Estatísticas das comparações *Intra* e *Inter* grupos (de tamanho N) utilizando coeficientes de correlação de *Pearson* entre as projeções do **PHA** - Harmônico por Tempo (**HxT**) e por Magnitude (**HxM**) - clarinetistas C1 a C6 executando a nota 12. Médias das correlações com Intervalos de Confiança (I.C.) a 95% nas colunas à direita. Média geral *Intra* e o valor p da **ANOVA** ao centro.

Os resultados indicam, para ambas as projeções **HxT** e **HxM** da nota 12, diferença significativa ao nível de confiança 95% (valor $p < 0,05$) entre as médias dos grupos de correlações. As diferenças entre as médias *Intra* e *Inter* foram de 0.19 (**HxT**) e 0.02 (**HxM**). A média *Intra* e a média *Inter* dos coeficientes de correlação para a projeção **HxT** (0,22 e 0,03) foram bem menores do que as mesmas da projeção **HxM** (0,91 e 0,93). Portanto, as médias *Intra* e *Inter* para **HxT** indicam correlação baixa, com apenas o

clarinetista C6 apresentando média de correlações próxima de moderada (entre 0,5 e 0,7). No caso de **HxM**, as médias *Intra* e *Inter* ficaram muito próximas a 0,9 (correlações muito fortes), com apenas C3 e C5 apresentando médias de correlações fortes (entre 0,7 e 0,9). A transição para a nota 12 é em condição de *Legato* em um intervalo ascendente dentro do mesmo registro do instrumento. Na Figura 59 estão mostradas as projeções **HxT** (1a até 1f) e **HxM** (2a até 2f) de todas as 6 execuções realizadas pelos 6 clarinetistas.

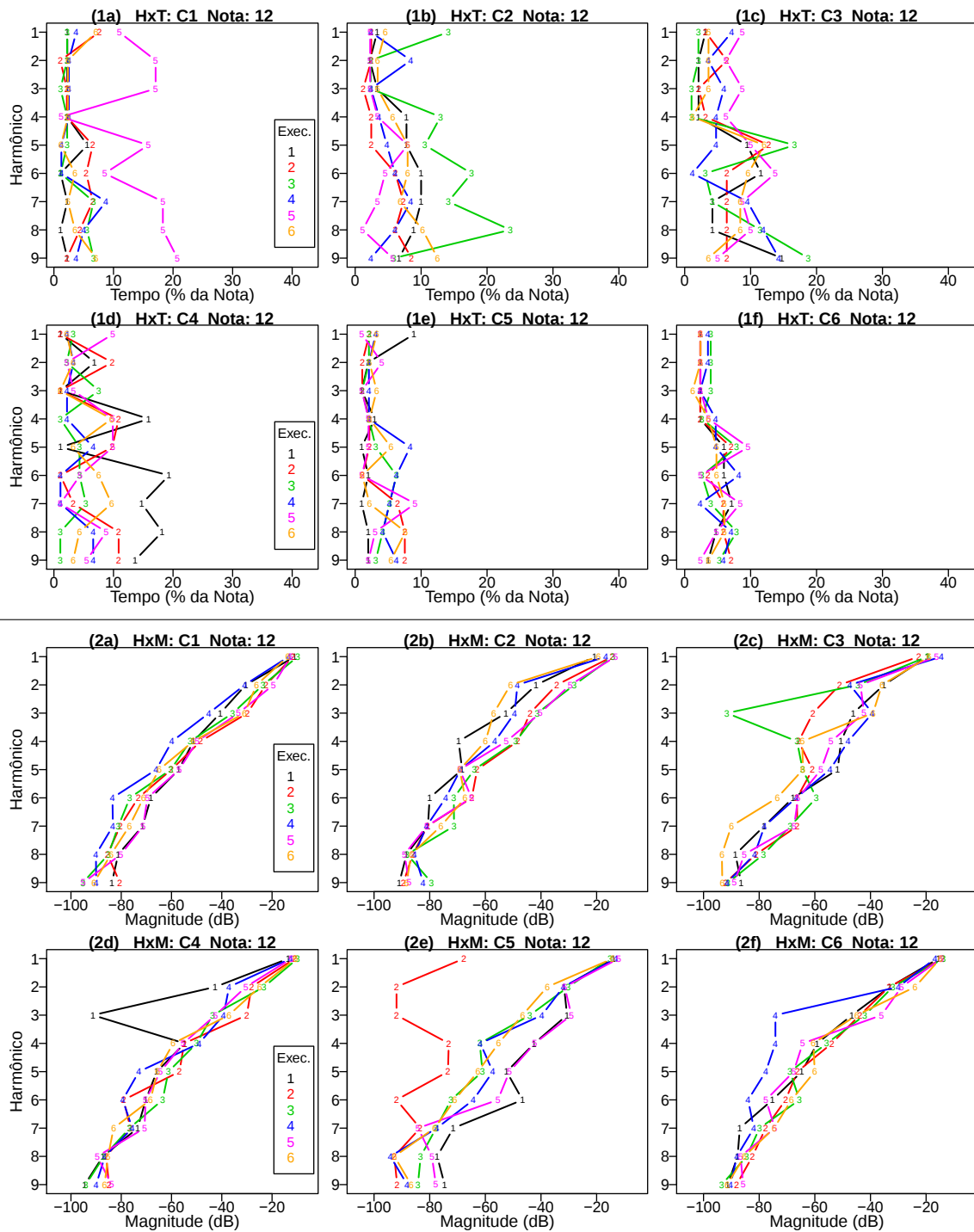


Figura 59 – Projeções do **PHA**, nota 12, clarinetistas C1 a C6: Harmônico por Tempo (**HxT**) em (1a) até (1f) e Harmônico por Magnitude (**HxM**) em (2a) até (2f).

Nota 13

Na Figura 60, estão mostradas as estatísticas das comparações *Intra* e *Inter* grupos para a nota 13, executada pelos seis clarinetistas (C1 a C6), utilizando coeficientes de correlação de *Pearson* entre as projeções **HxT** e **HxM** do **PHA**. Os valores das colunas N (mostrados na figura) indicam quantos coeficientes de correlação foram calculados para cada grupo, e podem variar caso haja corte de execuções devido à ataques mal sucedidos. No caso desta nota 13, houve dois cortes de execução no grupos de clarinetista C2, resultando em 6 comparações (combinações 2 a 2 de 4 execuções) em cada grupo *Intra*, e 28 comparações no grupo *Inter*. Os valores das médias das correlações estão nas colunas à direita, com Intervalos de Confiança (I.C.) a 95% entre parênteses. Nas colunas ao centro, para as duas projeções, encontram-se a média geral *Intra* e o valor p da **ANOVA** (teste para verificação de existência ou não de diferença entre as médias dos grupos).

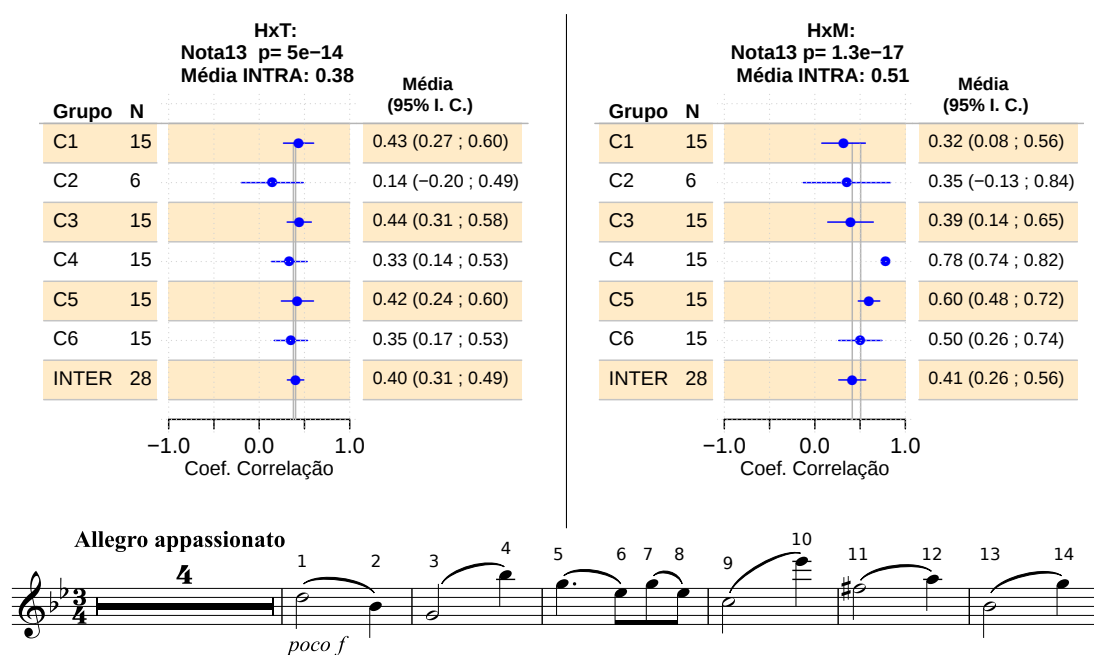


Figura 60 – Estatísticas das comparações *Intra* e *Inter* grupos (de tamanho N) utilizando coeficientes de correlação de *Pearson* entre as projeções do **PHA** - Harmônico por Tempo (**HxT**) e por Magnitude (**HxM**) - clarinetistas C1 a C6 executando a nota 13. Médias das correlações com Intervalos de Confiança (I.C.) a 95% nas colunas à direita. Média geral *Intra* e o valor p da **ANOVA** ao centro.

Os resultados indicam, para ambas as projeções **HxT** e **HxM** da nota 13, diferença significativa ao nível de confiança 95% (valor $p < 0,05$) entre as médias dos grupos de correlações. As diferenças entre as médias *Intra* e *Inter* foram de -0.02 (**HxT**) e 0.10 (**HxM**). A média *Intra* e a média *Inter* dos coeficientes de correlação para a projeção **HxT** (0,38 e 0,40) foram pouco menores do que as mesmas da projeção **HxM** (0,51 e 0,41). Portanto, as médias *Intra* e *Inter* para **HxT** indicam correlações baixas, com apenas C2 apresentando média de correlações baixa ($< 0,3$). No caso de **HxM**, as médias *Intra* e

Inter se mostraram fraca e moderada, com o C4 se destacando com média de correlações forte (entre 0,7 e 0,9). A transição para a nota 13 é articulada em um intervalo descendente, ocorrendo a mudança de registro do instrumento. Na Figura 61 estão mostradas as projeções **HxT** (1a até 1f) e **HxM** (2a até 2f) de todas as 6 execuções realizadas pelos 6 clarinetistas.

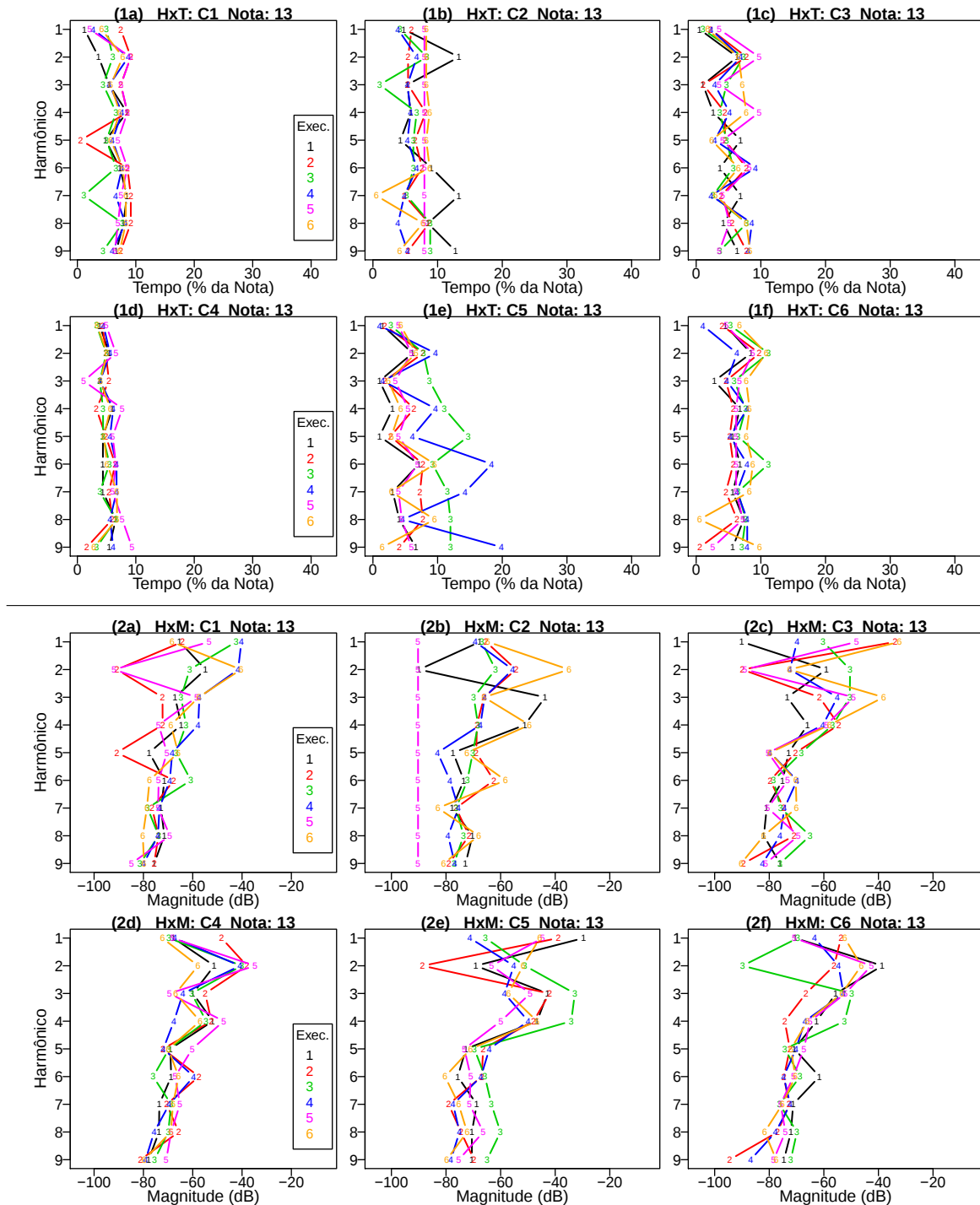


Figura 61 – Projeções do **PHA**, nota 13, clarinetistas C1 a C6: Harmônico por Tempo (**HxT**) em (1a) até (1f) e Harmônico por Magnitude (**HxM**) em (2a) até (2f).

Nota 14

Na Figura 62, estão mostradas as estatísticas das comparações *Intra* e *Inter* grupos para a nota 14, executada pelos seis clarinetistas (C1 a C6), utilizando coeficientes de correlação de *Pearson* entre as projeções **HxT** e **HxM** do **PHA**. Os valores das colunas N (mostrados na figura) indicam quantos coeficientes de correlação foram calculados para cada grupo, e podem variar caso haja corte de execuções devido à ataques mal sucedidos. No caso desta nota 14, houve apenas um corte de execução em C5, resultando em 10 comparações nesse grupo, e 29 comparações no grupo *Inter*. Os valores das médias das correlações estão nas colunas à direita, com Intervalos de Confiança (I.C.) a 95% entre parênteses. Nas colunas ao centro, para as duas projeções, encontram-se a média geral *Intra* e o valor p da **ANOVA** (teste para verificação de existência ou não de diferença entre as médias dos grupos).

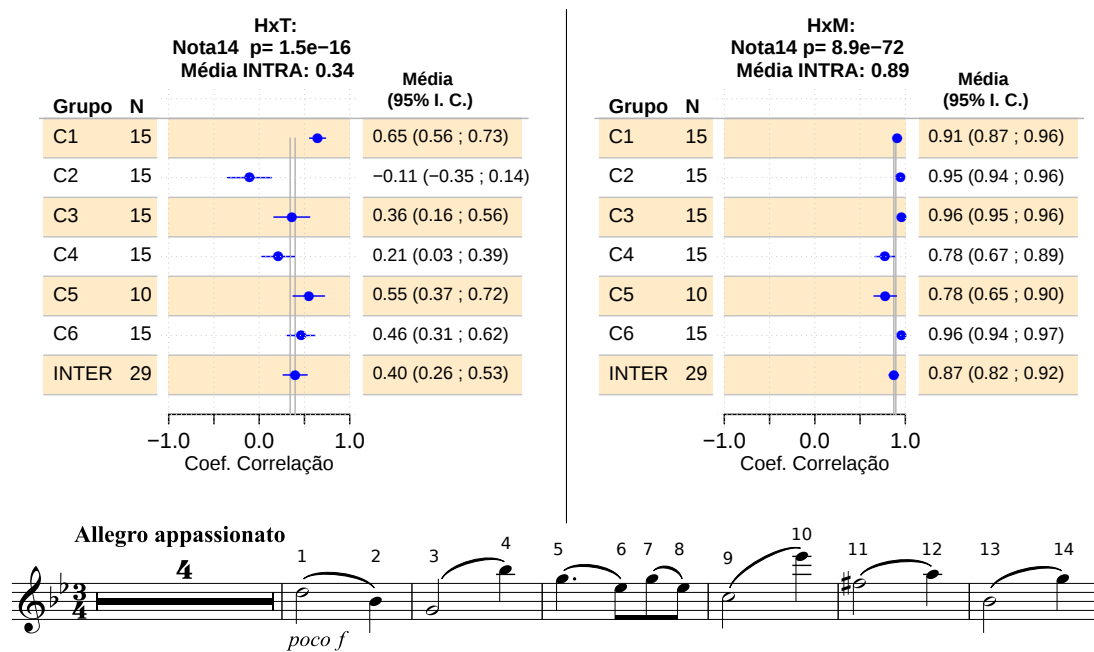


Figura 62 – Estatísticas das comparações *Intra* e *Inter* grupos (de tamanho N) utilizando coeficientes de correlação de *Pearson* entre as projeções do **PHA** - Harmônico por Tempo (**HxT**) e por Magnitude (**HxM**) - clarinetistas C1 a C6 executando a nota 14. Médias das correlações com Intervalos de Confiança (I.C.) a 95% nas colunas à direita. Média geral *Intra* e o valor p da **ANOVA** ao centro.

Os resultados para **HxT** e **HxM** da nota 14 indicam diferença significativa ao nível de confiança 95% ($p < 0,05$) entre as médias dos grupos de correlações. As diferenças entre as médias *Intra* e *Inter* foram de -0.06 (**HxT**) e 0.02 (**HxM**). A média *Intra* e a média *Inter* dos coeficientes de correlação para a projeção **HxT** (0,34 e 0,40) foram bem menores do que as mesmas da projeção **HxM** (0,89 e 0,87). Portanto, as médias *Intra* e *Inter* para **HxT** indicam correlação moderada, com a média de C2 baixa, indicando inclusive, correlação negativa (-0.11). No caso de **HxM**, as médias *Intra* e *Inter* foram fortes (entre

0,7 e 0,9), com C1, C2, C3 e C6 apresentando médias de correlações muito fortes ($> 0,9$). A transição para a nota 14 é em condição de *Legato*, em um intervalo ascendente de sexta maior, subindo um registro do instrumento. Na Figura 63 estão mostradas as projeções **HxT** (1a até 1f) e **HxM** (2a até 2f) de todas as 6 execuções dos 6 clarinetistas.

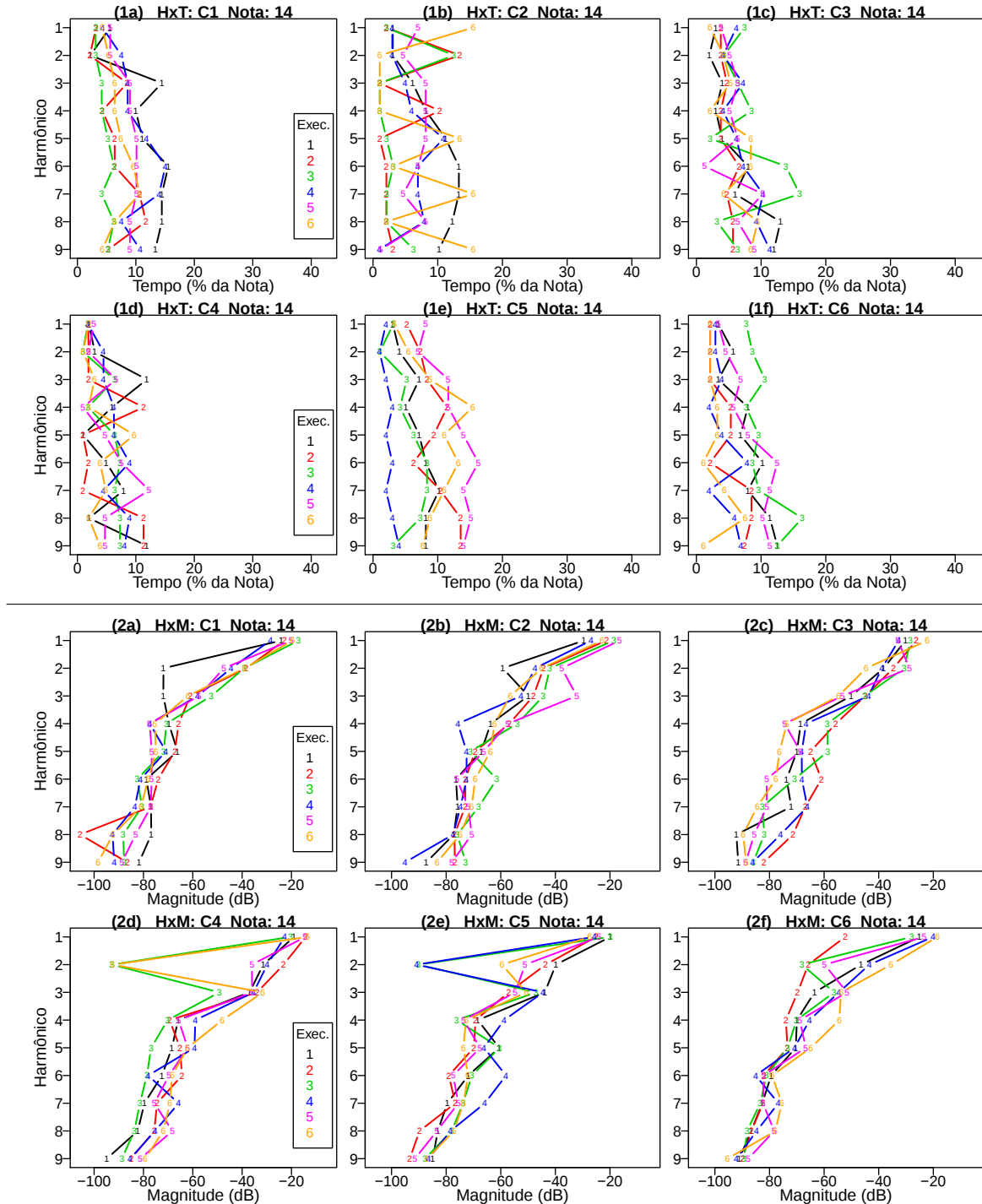


Figura 63 – Projeções do **PHA**, nota 14, clarinetistas C1 a C6: Harmônico por Tempo (**HxT**) em (1a) até (1f) e Harmônico por Magnitude (**HxM**) em (2a) até (2f).

Nota 22

Na Figura 64, estão mostradas as estatísticas das comparações *Intra* e *Inter* grupos para a nota 22, executada pelos seis clarinetistas (C1 a C6), utilizando coeficientes de correlação de *Pearson* entre as projeções **HxT** e **HxM** do **PHA**. Os valores das colunas N (mostrados na figura) indicam quantos coeficientes de correlação foram calculados para cada grupo, e podem variar caso haja corte de execuções devido à ataques mal sucedidos. No caso desta nota 22, não houve nenhum corte de execução nos 6 grupos de clarinetistas. Os valores das médias das correlações estão nas colunas à direita, com Intervalos de Confiança (I.C.) a 95% entre parênteses. Nas colunas ao centro, para as duas projeções, encontram-se a média geral *Intra* e o valor p do teste **ANOVA**.

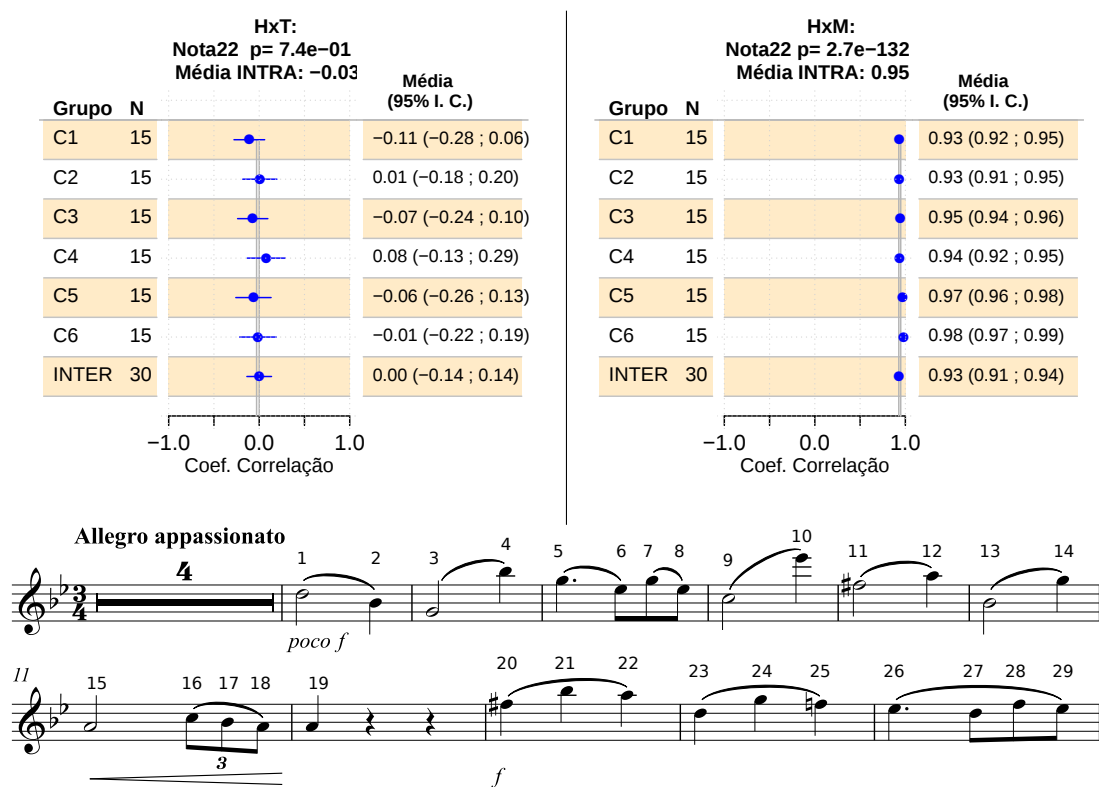


Figura 64 – Estatísticas das comparações *Intra* e *Inter* grupos (de tamanho N) utilizando coeficientes de correlação de *Pearson* entre as projeções do **PHA** - Harmônico por Tempo (**HxT**) e por Magnitude (**HxM**) - clarinetistas C1 a C6 executando a nota 22. Médias das correlações com Intervalos de Confiança (I.C.) a 95% nas colunas à direita. Média geral *Intra* e o valor p da **ANOVA** ao centro.

Os resultados da nota 22 indicam diferença significativa entre as médias dos grupos de correlações ($p < 0,05$), apenas para **HxM**. A projeção **HxT**, por outro lado, não obteve diferença significativa ($p = 0,74$). As diferenças entre as médias *Intra* e *Inter* foram de -0.03 (**HxT**) e 0.02 (**HxM**). As médias *Intra* e *Inter* dos coeficientes de correlação para **HxT** (-0,03 e 0,00) foram bem menores do que as mesmas de **HxM** (0,95 e 0,93). Portanto, as médias *Intra* e *Inter* para **HxT** indicam correlação muito baixa, com nenhum grupo

de clarinetista apresentando média de correlações maiores que 0,08. No caso de **HxM**, as médias *Intra* e *Inter* ficaram acima de 0,9 (correlações muito fortes), comportamento igual para todos os grupos *Intra*. A transição para a nota 22 é em condição de *Legato*, com intervalo descendente de apenas um semitom dentro do mesmo registro do instrumento, realizada com apenas a liberação de uma chave. Na Figura 65 estão mostradas as projeções **HxT** (1a até 1f) e **HxM** (2a até 2f) de todas as 6 execuções realizadas pelos 6 clarinetistas.

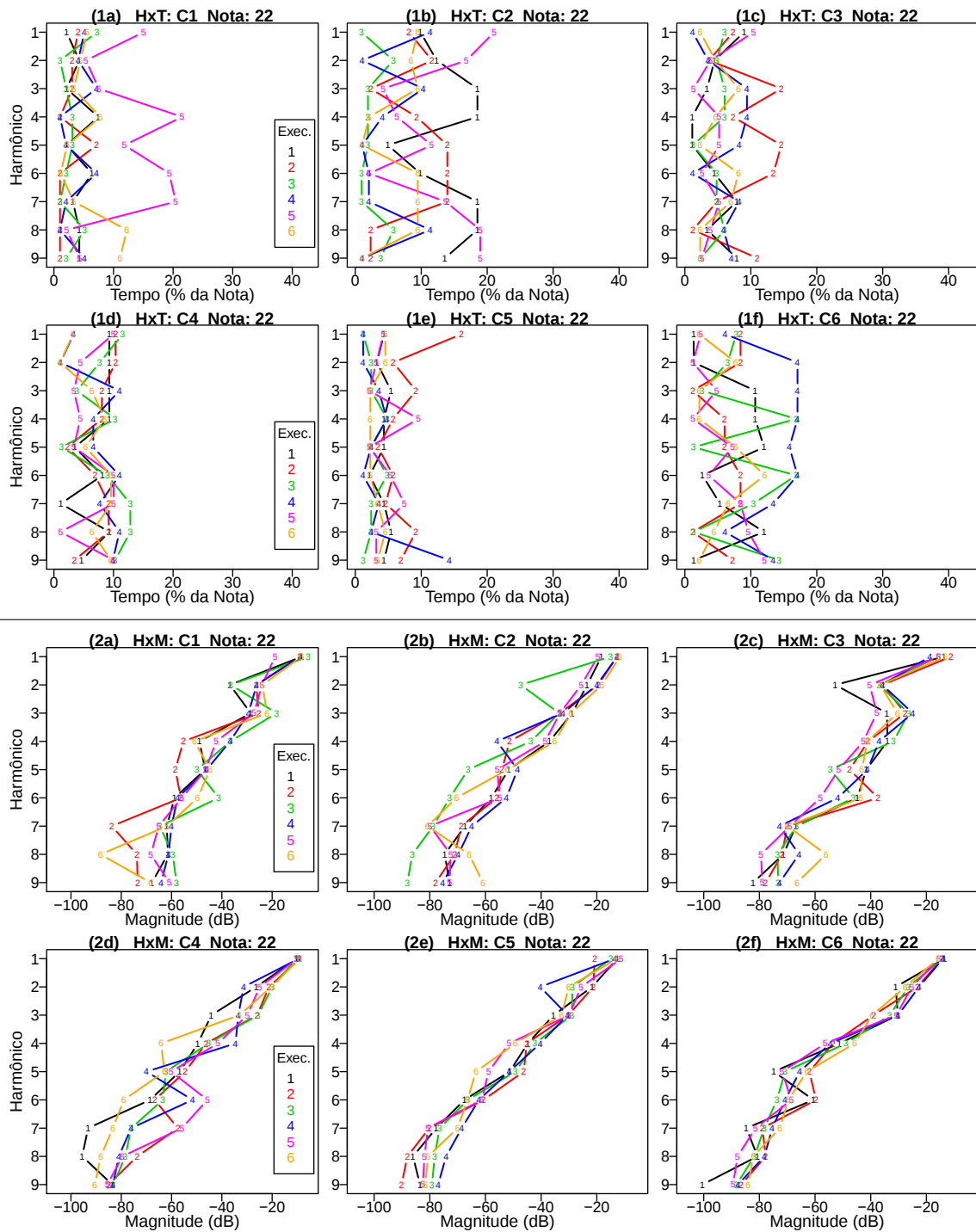


Figura 65 – Projeções do **PHA**, nota 22, clarinetistas C1 a C6: Harmônico por Tempo (**HxT**) em (1a) até (1f) e Harmônico por Magnitude (**HxM**) em (2a) até (2f).

Visão Geral dos Coeficientes de Correlação

Nesta seção, é apresentada uma visão geral dos dados dos coeficientes de correlação entre execuções *Intra* (mesmo clarinetista) e *Inter* grupo (clarinetistas diferentes) nas projeções do **PHA** (Harmônico por Tempo - HxT e Harmônico por Magnitude - HxM), para todas as notas. Inicialmente, é feita a comparação entre as médias dos coeficientes de correlação dos grupos *Intra* e *Inter* através de teste t, para as duas projeções HxT e HxM em cada uma das notas consideradas, cujos resultados estão presentes na Figura 66.

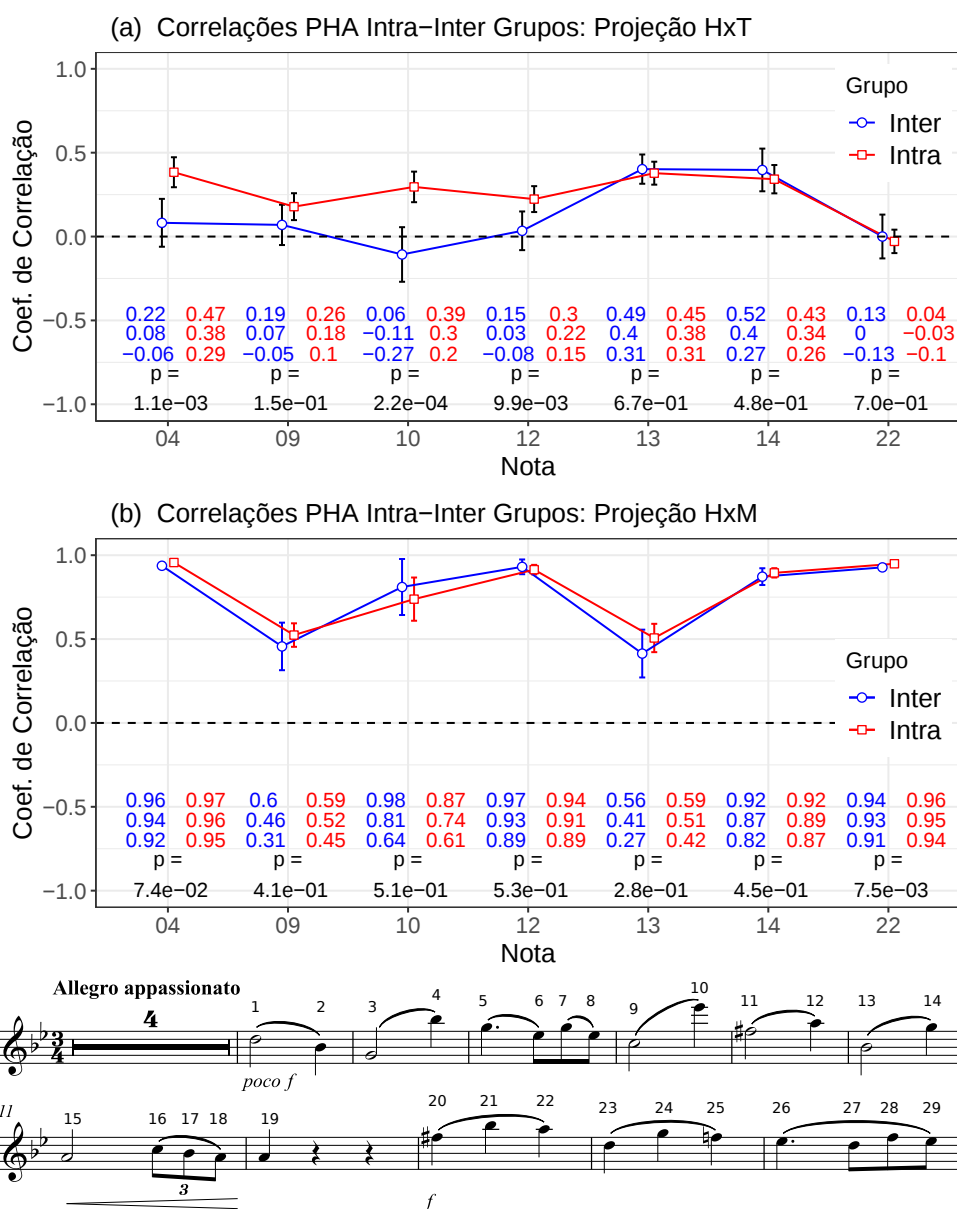


Figura 66 – Comparações entre as médias dos coeficientes de correlação para todas as notas. Projeção HxT em (a) e HxM em (b). Valores das médias com Intervalos de Confiança 95 % e valores p dos testes t na parte inferior dos gráficos.

Lembrando que o Coeficiente de Correlação de *Pearson* é uma medida do grau de dependência entre as variáveis comparadas. Portanto, médias *Intra* próximas de 1, indicam

que existe um alto grau de coerência nas execuções de um músico. Se o Intervalo de Confiança das correlações *Intra* é estreito, significa alto grau de consistência. Porém, não basta avaliar apenas a média *Intra*. Deve-se também avaliar a média *Inter*, pois esta indica se existem outros fatores, além dos músicos, aumentando o grau de dependência entre os perfis. Neste ponto, visando facilitar a análise geral, foram reunidos os resultados da análise de correlações realizadas nas projeções do **PHA** na Tabela 12.

Tabela 12 – Resumo dos resultados da análise de correlações realizada nas duas projeções do **PHA**, Harmônico x Tempo (HxT) e Harmônico x Magnitude (HxM): Testes ANOVA (comparação entre as médias dos coeficientes de correlação dos grupos *Intra*) e testes t (comparação entre as médias *Intra* e *Inter* grupos).

Projeção H x T					N o t a	Projeção H x M				
Médias			Valor p	Valor p		Médias			Valor p	Valor p
<i>Intra</i> (μ_1)	<i>Inter</i> (μ_2)	$\mu_1 - \mu_2$	ANOVA <i>Intra</i>	Teste t $\mu_1 - \mu_2$		<i>Intra</i> (μ_1)	<i>Inter</i> (μ_2)	$\mu_1 - \mu_2$	ANOVA <i>Intra</i>	Teste t $\mu_1 - \mu_2$
0,38	0,08	0,30	<0,01*	<0,01*	04	0,96	0,94	0,02	<0,01*	0,07
0,18	0,07	0,11	<0,01*	0,15	09	0,52	0,46	0,06	<0,01*	0,41
0,30	-0,11	0,41	<0,01*	<0,01*	10	0,74	0,81	-0,07	<0,01*	0,51
0,22	0,03	0,19	<0,01*	<0,01*	12	0,91	0,93	-0,02	<0,01*	0,53
0,38	0,40	-0,02	<0,01*	0,67	13	0,51	0,41	0,10	<0,01*	0,28
0,34	0,40	-0,06	<0,01*	0,48	14	0,89	0,87	0,02	<0,01*	0,45
-0,03	0,00	-0,03	0,74	0,70	22	0,95	0,93	0,02	<0,01*	<0,01*

(*) Valores p com significância estatística ao nível de confiança 95%.

Inicialmente, pode ser observado na Tabela 12 que, para os ataques de todas as sete notas avaliadas, o perfil tridimensional de início dos harmônicos no ataque (**PHA**) se mostrou capaz de diferenciar os grupos de clarinetistas analisados. Isso porque, com exceção da projeção **HxT** da nota 22, todas as outras projeções, incluindo **HxM** da nota 22, apresentaram diferenças significativas ($p < 0,05$) entre as médias de correlações *Intra* grupos.

Podemos observar também, as médias de correlações *Intra* (de 0,52 a 0,96) e *Inter* (de 0,46 a 0,97), para a projeção **HxM**, com valores bem mais elevados do que as mesmas (de -0,03 a 0,38) e (de -0,11 e 0,40) para a projeção **HxT**. O valor mais alto da média *Inter* da projeção **HxM** quando comparado com a mesma média da projeção **HxT**, sugere a existência de uma estrutura de magnitudes dos harmônicos que compõem o **PHA**, comum aos músicos. Essa estrutura comum pode estar relacionada às proporções entre energias dos harmônicos que define o timbre do instrumento. A mesma inferência não pôde ser feita em relação ao tempo de início dos harmônicos, pois as projeções **HxT** apresentaram valores baixos de médias de correlações *Inter*.

Nas projeções **HxT**, das sete transições analisadas, apenas para a nota 22 não houve diferenças significativas, tanto entre as médias de correlações *Intra*, quanto entre as médias de correlações *Intra* e *Inter*, apresentando ambas as médias de correlações próximas

de zero. Isso reflete o fato de essa transição simples de semitom descendente ser realizada apenas pelo acionamento de uma chave com mola, deixando pouca margem de influência do músico na produção do ataque desta nota. Esse fato pode explicar também a maior consistência das magnitudes tanto *Intra* quanto *Inter* grupo, quando comparadas com os tempos de inícios dos harmônicos do **PHA**. Somado à explicação acima, o fato de os clarinetistas terem usados instrumentos diferentes no experimento, pode explicar as diferenças significativas entre as médias de correlações, tanto *Intra* quanto *Intra-Inter* para a projeção **HxM** da nota 22.

Ainda nas projeções **HxT**, os saltos para as notas 4, 10 e 12 apresentaram diferença significativa no teste t entre as médias *Intra* e *Inter*. As notas 10 e 4 são as duas de maior dificuldade de execução, e foram as duas com maiores diferenças entre médias *Intra* e *Inter* (0,41 e 0,30).

Nas projeções **HxM**, com exceção para a nota 22, nenhuma outra apresentou diferença significativa no teste t entre as médias *Intra* e *Inter*. Uma explicação para isso pode ser o fato da possível estrutura de magnitudes dos harmônicos que define o timbre do instrumento (inferida pelas altas médias *Inter* classes), comum entre os grupos de clarinetistas, estar interferindo nas correlações, mascarando as possíveis diferenças entre músicos.

As notas articuladas (notas 9 e 13), na projeção **HxM**, apresentam médias *Intra* (de 0,51 a 0,52) e *Inter* (de 0,41 a 0,46) bem menores do que as mesmas das notas ligadas, que variam de 0,74 a 0,96 nas médias *Intra*, e 0,81 a 0,94 nas médias *Inter*. Essas diferenças são estatisticamente significativas, pois os limites superiores dos intervalos de confiança (ICs) das notas 9 e 13 são menores do que os limites inferiores dos ICs das outras notas (que são ligadas), como pode ser verificado na Figura 66. Uma possível explicação para esse fato é que, como em notas articuladas há a interrupção do fluxo de ar produzido pelo clarinetista, consequentemente há uma diminuição de energia antes do ataque. Isso faz com que os harmônicos de ataques em notas articuladas variem mais em magnitude do que os mesmos em notas ligadas. Essa maior variação amplia a margem para atuação do músico nas magnitudes dos harmônicos, diminuindo com isso a proporção da estrutura fixa de balanceamento das energias dos harmônicos que define o timbre do instrumento em regiões de sustentação de energia. Um estudo realizado por Grey (1977) mostrou que durante o ataque de vários instrumentos, inclusive de clarineta, ocorre uma mudança no equilíbrio entre os harmônicos devido a diferentes taxas de crescimento e decrescimento dos mesmos.

Com essas últimas análises, encerram-se as discussões dos resultados. Na próxima seção, será feita a conclusão desta tese, com um compilado de toda a metodologia utilizada na pesquisa realizada, incluindo os principais resultados alcançados. Na sequência da mesma seção da conclusão, serão apontados os trabalhos futuros.

7 Conclusão

Foi explorada nesta tese a ideia de descrever o conteúdo de ataques de notas de clarineta por meio de modelagem e decomposição espectral. Mais especificamente, comparando as proporções de energia das componentes harmônicas e residuais de forma independente e através da análise do perfil temporal dos harmônicos modelados. Para a caracterização do trecho da transição entre notas no qual o ataque está contido, foi proposta a adaptação do cálculo de um índice de legato, já apresentado na literatura, para utilizar a energia da componente harmônica da transição, em vez da energia total. O índice proposto foi chamado de Índice de Legato Harmônico (**ILh**). Para caracterizar ataques de notas, foi proposto um plano para confrontar duas componentes: **H2** - o conteúdo de energia devido à reverberação harmônica da nota anterior e **R2** - a componente residual subtraindo as componentes harmônicas da nota anterior (**H2**) e a nota sendo atacada (**H1**) a partir do sinal original. O confronto dessas duas componentes em termos de proporções de energia em relação ao sinal original permitiu inferir sobre três componentes (**H1**, **H2** e **R2**) em uma visualização em **2D** chamada de plano **H2R2**. Essa abordagem fornece descrição mais detalhada sobre o conteúdo espectral da região de ataque e permite inferências como interações com a acústica da sala, enriquecendo assim, o estudo de apresentações musicais nas condições de prática cotidiana, onde ambientes reverberantes estão sempre presentes.

As propostas foram testadas usando um conjunto de gravações de um trecho de uma peça de clarineta do repertório clássico tradicional, o primeiro movimento da Sonata de Brahms para Clarineta em Fá menor, op. 120 n° 1, tocado por seis clarinetistas profissionais. Testes **ANOVA** indicaram diferenças significativas ($p < 0,05$), considerando “músico” e “nota” como fator para **H2** e **R2** e do **ILh**. Os testes **MANOVA** para as duas variáveis (**R2** e **H2**) indicaram diferenças ($p < 0,05$) ao considerar “músico” como fator para todos os ataques das transições entre notas avaliadas. Foram examinados diferentes *legatos*, bem como ataques de notas em transições articuladas, apresentando diferentes demandas técnicas de realização. Observou-se maior variância de ambas as dimensões para notas com maior demanda técnica. Essa abordagem de análise também foi capaz de revelar ataques de notas que não foram executados com sucesso, conforme observado nas execuções da nota 4 do clarinetista C2. Investigações futuras podem ser feitas tentando encontrar relações entre as proporções das componentes e outros fatores, como estimativas de impedância acústica da nota individual ou variações na acústica da sala. Em outra frente, foi proposto um descritor tridimensional, Perfil Harmônico do Ataque (**PHA**), envolvendo tempo, frequência (harmônicos) e magnitude. Partindo da fundamental, o **PHA** cruza, na sequência, o ponto de 50% da fixa de variação das magnitudes em dB de cada harmônico no trecho do ataque. Foram avaliadas as projeções do **PHA** nos planos

harmônico-magnitude e harmônico-tempo para as mesmas notas das gravações usadas para os testes do plano **H2R2**. Para cada uma das duas projeções, foram calculados os coeficientes de correlação de *Pearson* entre perfis de execuções diferentes do mesmo músico (chamados de *Intra*), e outro chamado de *Inter*, contendo os coeficientes de correlação entre as projeções dos **PHAs** de cada clarinetista com os outros colegas. Os testes apontaram separabilidade ($p < 0,05$) *Intra* classes em todos os perfis de início dos harmônicos (**PHA**) dos ataques das sete notas avaliadas. Valores altos da média *Inter* para as projeções **HxM** (o mesmo não acontece com **HxT**) sugere a existência de uma estrutura de magnitudes dos harmônicos que compõem o **PHA**, comum aos músicos, que pode estar relacionada às proporções entre energias dos harmônicos que define o timbre do instrumento. Diferenças significativas ($p < 0,05$) entre as médias de correlações *Intra* e *Inter* ocorreram em 4 notas, com 3 delas para a projeção **HxT** em notas ligadas, sugerindo um indicativo de que a informação sobre o tempo de inícios dos harmônicos são mais relevantes em notas ligadas.

Trabalhos Futuros

Investigações futuras em relação ao plano **H2R2** podem ser feitas na tentativa de encontrar relações entre as proporções das componentes e outros fatores, como estimativas de impedância acústica da nota individual ou variações na acústica da sala. Também pode ser aplicada a abordagem em registros diferentes e outros instrumentos, buscando definir regiões características no plano.

Um outro trabalho futuro é uma melhoria para o **PHA**. A melhoria se trata de uma interpolação das curvas dos harmônicos extraídos no ataque. Essa interpolação aumenta a resolução temporal, que levaria ao aumento da precisão na detecção do **PHA**.

Outra questão importante para validação da metodologia proposta são os testes subjetivos, tanto para a verificação da modelagem, quanto para a verificação de relevância perceptiva das variáveis avaliadas. A seguir são descritos os testes subjetivos propostos para trabalhos futuros:

Avaliação da Modelagem - Testar sinal original contra **H1+H2+R2**, utilizando teste ABX e MOS, visando avaliar a qualidade da modelagem;

Relevância de H2 - Testar **H1+H2+R2** contra **H1+R2**, utilizando teste ABX e MOS, visando avaliar se a componente harmônica relativa à reverberação da nota anterior é perceptivamente relevante na composição do ataque;

Relevância de R2 - Testar **H1+H2+R2** contra **H1+H2**, visando avaliar se a componente residual (sinal original subtraído das componentes harmônicas) é perceptivamente relevante na composição do ataque.

Referências

- BOLZINGER, S.; WARUSFEL, O.; KAHLE, E. A study of the influence of room acoustics on piano performance. *Le Journal de Physique IV*, EDP sciences, v. 4, n. C5, p. C5–617, 1994. Citado 3 vezes nas páginas 23, 24 e 35.
- CAMPOLINA, T.; LOUREIRO, M.; MOTA, D. Expan: a tool for musical expressiveness analysis. In: *Proceedings of the 2nd International Conference of Students of Systematic Musicology*. [S.l.: s.n.], 2009. p. 24–27. Citado na página 43.
- CANNAM, C.; LANDONE, C.; SANDLER, M. Sonic visualiser: An open source application for viewing, analysing, and annotating music audio files. In: *Proceedings of the ACM Multimedia 2010 International Conference*. Firenze, Italy: [s.n.], 2010. p. 1467–1468. Citado na página 80.
- DAUDET, L. A review on techniques for the extraction of transients in musical signals. *Computer Music Modeling and Retrieval*, Springer, p. 219–232, 2006. Citado na página 36.
- FISCHINGER, T.; FRIELER, K.; LOUHIVUORI, J. Influence of virtual room acoustics on choir singing. *Psychomusicology: Music, Mind, and Brain*, Educational Publishing Foundation, v. 25, n. 3, p. 208, 2015. Citado 3 vezes nas páginas 23, 24 e 35.
- FLETCHER, H.; MUNSON, W. A. Loudness, its definition, measurement and calculation. *Bell System Technical Journal*, Wiley Online Library, v. 12, n. 4, p. 377–430, 1933. Citado 3 vezes nas páginas 11, 27 e 28.
- GREY, J. Multidimensional perceptual scaling of musical timbres. *J. Acoust. Soc. Am*, v. 61, n. 5, p. 1270–1277, 1977. Citado 2 vezes nas páginas 23 e 102.
- HUANG, X. et al. *Spoken language processing: A guide to theory, algorithm, and system development*. [S.l.]: Prentice Hall PTR, 2001. Citado 4 vezes nas páginas 11, 26, 28 e 29.
- KALKANDJIEV, Z. S.; WEINZIERL, S. The influence of room acoustics on solo music performance: An experimental study. *Psychomusicology: Music, Mind, and Brain*, Educational Publishing Foundation, v. 25, n. 3, p. 195, 2015. Citado 3 vezes nas páginas 23, 24 e 35.
- KATO, K.; UENO, K.; KAWAI, K. Musicians’ adjustment of performance to room acoustics, part iii: Understanding the variations in musical expressions. *Journal of the Acoustical Society of America*, [New York: Acoustical Society of America], v. 123, n. 5, p. 3610, 2008. Citado 3 vezes nas páginas 23, 24 e 35.
- LOUREIRO, M. A. et al. Content analysis of note transitions in music performance. In: INESC PORTO. *Proceedings of the 6th Sound and Music Computing Conference (SMC 2009)*, Porto, Portugal. [S.l.], 2009. p. 355–359. Citado na página 24.
- LUCE, D. *Physical correlates of nonpercussive musical instrument tones*. Tese (Doutorado) — MIT, 1963. Citado na página 23.

- MAESTRE, E.; GÓMEZ, E. Automatic characterization of dynamics and articulation of expressive monophonic recordings. In: CITESEER. *Proceedings of the 118th Audio Engineering Society Convention*. [S.l.], 2005. Citado 3 vezes nas páginas 24, 36 e 37.
- MAHER, R. C.; BEAUCHAMP, J. W. Fundamental frequency estimation of musical signals using a two-way mismatch procedure. *The Journal of the Acoustical Society of America*, ASA, v. 95, n. 4, p. 2254–2263, 1994. Citado 2 vezes nas páginas 40 e 41.
- MCADAMS, S. Perspectives on the contribution of timbre to musical structure. *Computer Music Journal*, MIT Press, v. 23, n. 3, p. 85–102, 1999. Citado na página 38.
- MONTGOMERY, D. C.; RUNGER, G. C. *Estatística Aplicada e Probabilidade para Engenheiros*, 4ª. [S.l.]: Ed. Rio de Janeiro: Editora LTC, 2009. Citado na página 78.
- PEARSON, K. *Notes on Regression and Inheritance in the Case of Two Parents* *Proceedings of the Royal Society of London*, 58, 240-242. [S.l.]: ed. Royal Society (Great Britain), 1895. Citado na página 53.
- RISSET, J. Computer study of trumpet tones. *The Journal of the Acoustical Society of America*, v. 38, p. 912, 1965. Citado na página 23.
- SALAMON, J.; GÓMEZ, E. Mir. edu: An open-source library for teaching sound and music description. *Proceedings of the 15th International Society for Music Information Retrieval (ISMIR)*, Tapei, Taiwan, 2014. Citado na página 80.
- SCHAEFFER, P. *Solfège de l'objet Sonore*. INA GRM. [S.l.]: Seuil e GRM, 1967. Citado na página 30.
- SERRA, X.; SMITH, J. Spectral modeling synthesis: A sound analysis/synthesis system based on a deterministic plus stochastic decomposition. *Computer Music Journal*, JSTOR, v. 14, n. 4, p. 12–24, 1990. Citado 3 vezes nas páginas 39, 40 e 41.
- SMITH, J. O. *Spectral audio signal processing*. [S.l.]: W3K publishing, online book., 2011. Citado na página 39.
- TEIXEIRA, E. C. et al. Motion analysis of clarinet performers. *Journal of New Music Research*, Taylor & Francis, v. 44, n. 2, p. 97–111, 2015. Citado 3 vezes nas páginas 12, 54 e 55.