

Distribuição espectral da energia sonora na Produção Musical: desafios, estilos e concepções

Fernando Braga Campos¹

Universidade Federal de Minas Gerais/Programa de Pós Graduação em Música

SIMPOM: *Sonologia*

fernandobragacampos@gmail.com

Resumo: O equilíbrio na distribuição espectral da energia sonora e o máximo aproveitamento da faixa dinâmica disponível são grandes desafios da produção musical. A presente comunicação pretende trazer à superfície as variáveis envolvidas na questão, elencando o referencial teórico necessário à interpretação de dados objetivos e subjetivos e analisando as soluções utilizadas pelos profissionais. Os resultados demonstram que questões de ordem estética, como estilo e concepção sonora, são tão relevantes ao trabalho do profissional de áudio quanto as de ordem técnica.

Palavras-chave: Produção musical; Sons graves; Áudio; Gravação; *Mixagem*.

Spectral Distribution of Sound Energy in Music Production

Abstract: The balance in the spectral distribution of the sound energy and maximum utilization of the available dynamic range are major challenges of music production. This communication aims at bringing to the surface the variables involved in the issue, listing the theoretical framework necessary to the interpretation of objective and subjective data and analyzing the solutions used by professionals. The results demonstrate that aesthetic issues, such as style and sound conception, are as relevant to the work of audio professionals as the ones of technical order.

Keywords: Music Production; Low-End Sounds; Audio; Recording; Mixing.

Introdução

Uma das principais questões enfrentadas pela produção musical é o aproveitamento de toda a faixa dinâmica disponível com uma correta distribuição espectral do material sonoro. Sabe-se que nossa percepção não é tão apurada nos registros extremos grave e agudo quanto na porção média do espectro². Especial atenção deve ser dada ao registro

¹ Sérgio Freire Garcia (Orientador).

² ROBINSON, D.W.; DADSON, R.S.: 1956.

grave, que normalmente necessita de amplitudes comparativamente maiores para que se alcance um equilíbrio auditivo.

O texto que se segue é um recorte de uma dissertação defendida em dezembro de 2013 junto ao Programa de Pós-Graduação em Música da UFMG, linha de pesquisa Sonologia. Sua primeira seção trata dos conceitos e ferramentas utilizados na análise de trechos gravados. A seção seguinte analisa alguns excertos de Stravinsky e Ravel, e também discute um breve estudo de caso realizado na pesquisa. Essas breves análises sugerem que, lado a lado com os desafios técnicos postos pelo registro grave, há que se considerar as abordagens estilísticas a respeito do material musical e as concepções estéticas dos profissionais que executam tarefas consideradas, equivocadamente, como técnicas.

1. Conceitos, definições e ferramentas utilizadas

Alguns conceitos matemáticos, como a função logarítmica e os decibéis, são importantes para se compreender desde o mecanismo da transdução auditiva até o funcionamento de equipamentos como a mesa de som, pois as principais grandezas utilizadas - intensidade e *pitch* - são expressas ou se relacionam diretamente com a escala logarítmica. O estudo da psicoacústica se mostra também extremamente relevante para a adequada interpretação dos dados numéricos objetivos.

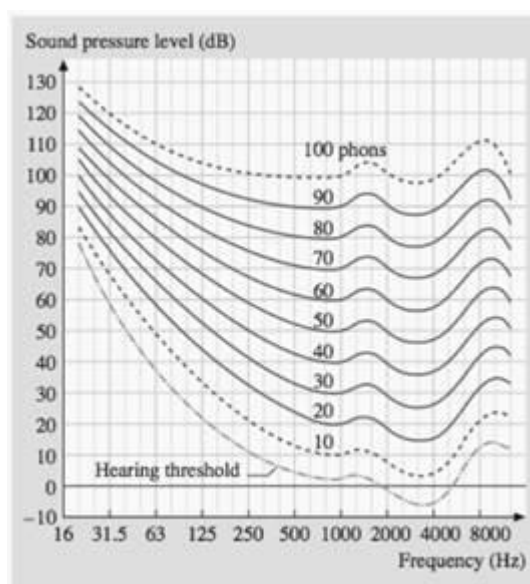


Figura 1: Curvas de mesma intensidade percebida para intensidades de 10 a 100 phons para sons apresentados de forma binaural na direção frontal. Curva de limiar auditivo (*hearing threshold*) é também mostrada. As curvas de 10 e 100 phons estão pontilhadas por serem calculadas por interpolação e extrapolação, respectivamente. Fonte: Rossing (2007, p.468) Trad. do autor.

1.1 As curvas de intensidade subjetiva (*equal loudness contours*)

O contorno sinuoso das curvas de *loudness* demonstra claramente a variação da sensibilidade do sistema auditivo às frequências e intensidades. Cada uma das curvas representa a percepção de uma mesma intensidade (*loudness*, na unidade phons) para sinais de frequências distintas. Assim, para que dois sons de frequências diferentes sejam percebidos como equivalentes em intensidade, busca-se, ao longo da curva, a pressão sonora necessária (expressa no eixo Y). Todos os pontos ao longo das curvas apresentadas produzem a mesma sensação de intensidade no ouvinte. Vê-se claramente o quanto nossa audição privilegia a percepção da região média (entre 250Hz e 5000Hz, especialmente) em detrimento das regiões do extremo grave e do extremo agudo.

1.2 O valor RMS³ de um sinal e sua expressão em decibéis

Para discorrer sobre a distribuição da energia espectral nos diferentes estilos musicais, importa esclarecer a forma de se calcular a energia aproximada de um sinal sonoro. Ela é normalmente representada pela envoltória de amplitude do sinal, construída a partir de valores RMS das amplitudes instantâneas. Para seu cálculo, convertem-se os diversos valores instantâneos dentro de determinado intervalo temporal (tipicamente entre 5 e 25 ms) em um único valor eficaz, ou RMS, que indica a amplitude de um sinal contínuo que conteria a mesma energia do sinal ondulatório. A figura abaixo ilustra diferentes modos de se calcular a amplitude de um sinal.

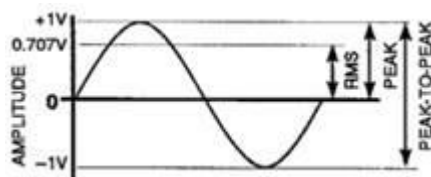


Figura 2: Valores RMS; de pico, e de pico-a-pico de um sinal de 1V de amplitude. Fonte: Davis & Jones (1990, p.27).

A razão pela qual o valor RMS é o mais adequado para as análises de distribuição espectral é dada por Davis e Jones:

O valor médio de um sinal (energia) não se altera muito, mesmo que haja transientes com picos altos. Reciprocamente, o nível dos picos pode variar drasticamente sem grande efeito nos níveis médios. Quando desejamos avaliar a intensidade percebida de um som, o valor RMS relaciona mais diretamente a percepção de nossos ouvidos à energia do áudio. (DAVIS & JONES, 1990, p. 27).

³ RMS, na sigla em inglês, *Root Mean Square*, representa a média quadrática.

A percepção de intensidade não está ligada somente ao valor RMS do sinal. Ela depende também da duração dos sons e de outros fatores como distribuição espectral e mascaramento, por exemplo. “O ouvido integra a energia com o passar do tempo dentro de uma janela temporal de aproximadamente 200 ms” (GELFAND, 2009, p. 170).

As escalas em decibéis (dB) são largamente utilizadas pelos profissionais de áudio em sua rotina de trabalho, mas sua interpretação nem sempre é fácil. A dificuldade se dá, principalmente, em função de sua natureza comparativa. É necessário estabelecer valores de referência. No universo do áudio, as referências utilizadas são bastante variadas (dBV, dBu, dBm, dBW, dB^{SPL}, dBFS etc). Para auxiliar a discussão das análises realizadas neste estudo, criamos duas tabelas comparativas: a primeira lida com a soma de dois sinais de mesmo conteúdo espectral e fase, comparando seus níveis e influência na *mixagem*. A segunda tabela (pensada em dBFS - *full scale*, com amplitude máxima igual a 1,0) indica o valor RMS de um sinal digital e seu correspondente *headroom*.

Soma de dois sinais* de áudio		
Quando a diferença é de (dB):	Soma-se ao mix (dB):	menor/maior (%)
0	6.02	100.0
-0.5	5.77	94.4
-3	4.65	70.8
-6	3.53	50.1
-9.5	2.51	33.5
-12	1.95	25.1
-15	1.42	17.8
-25	0.48	5.6
-30	0.27	3.2
-40	0.09	1.0

Tabela 1: Soma de dois sinais de áudio. *Sinais de mesmo conteúdo espectral e em fase. Fonte: Construção dos autores.

Headroom digital em relação a um único sinal		
Sinal solo (dBFS)	Ainda cabe um sinal de (dBFS)	Ainda cabe (%)
0	-	0.0
-0.5	-25.0	5.6
-3	-10.7	29.2
-6	-6.0	49.9
-10	-3.3	68.4
-12	-2.5	74.9
-15	-1.7	82.2
-20	-0.9	90.0
-30	-0.3	96.8
-40	-0.1	99.0

Tabela 2: Headroom⁴ digital em relação a um único sinal. Fonte: Construção dos autores.

⁴ Headroom é o termo técnico utilizado para denominar o que resta de faixa dinâmica disponível a partir de um dado ponto na escala.

Os dados das tabelas acima são fruto de um exercício de cálculo, no sentido de que pressupõe-se uma situação extrema: a relação entre dois sinais cuja única diferença é a amplitude. Assim as colunas centrais devem ser lidas como: "soma-se ao mix no máximo (dB)", e "ainda cabe um sinal de pelo menos (dBFS)", respectivamente. Por exemplo, se adicionarmos a um sinal pré-existente outro com mesma intensidade, espera-se que a *mixagem* tenha um ganho de no máximo 6 dB (na primeira tabela), e que o *headroom* seja da ordem de 6 dB (na segunda tabela).

1.3 Filtragem em três regiões

Com o intuito de analisar comparativamente a intensidade dos graves, médios e agudos, fizemos a divisão dos áudios originais em faixas de frequência. Utilizamos para isso uma filtragem baseada na transformada de Fourier, separando o espectro sonoro em três regiões principais: grave (primeiras três oitavas - 20Hz a 160Hz), médio (4 oitavas intermediárias - 160Hz a 2,56kHz) e agudo (3 oitavas mais agudas - 2,56kHz a 20kHz). A utilização da FFT⁵ possibilita que a separação seja feita sem vazamento entre as regiões. Cada uma dessas regiões foi direcionada a arquivos de áudio distintos, que foram então analisados no software Praat. A implementação da separação em faixas foi realizada no *software* Max, a partir de um filtro espectral de 1/4 de oitava, com funcionamento em tempo real, desenvolvido por Padovani e Freire (2006). As envoltórias de amplitudes utilizadas nas análises a seguir utilizaram intervalos temporais diferentes para o cálculo dos valores RMS, variando entre 5 ms (para a região aguda) e 25 ms (para a região grave). O tamanho variável da janela se justifica pela resolução necessária a cada faixa de frequências analisada. Embora a divisão em um número maior de regiões não ofereça nenhum problema técnico, consideramos que essas três faixas básicas são suficientes para uma análise quanti-qualitativa da distribuição espectral em graves, médios e agudos. Devemos lembrar que boa parte dos sons musicais ocupará mais de uma dessas faixas.

2. Análise de excertos musicais

Para ilustrar as diferenças da distribuição espectral em diferentes estilos, comparamos inicialmente dois trechos de peças escritas para orquestra sinfônica. Em seguida analisamos as diferentes versões de *mixagem* de uma peça composta para o teste que está descrito no item 3.3.3 da dissertação, da qual a presente comunicação é parte.

⁵ *Fast Fourier Transform* (Transformada Rápida de Fourier).

2.1 Trechos sinfônicos

Um dos principais desafios na gravação de uma orquestra sinfônica é a grande faixa dinâmica explorada por boa parte do repertório, acentuada por uma variação constante do conteúdo espectral. A título de exemplo, observamos as seções iniciais e finais de gravações de duas obras amplamente conhecidas (*Bolero*⁶, de Ravel; e *Sagração da Primavera*⁷, de Stravinsky), ambas iniciando em baixa intensidade, no registro médio, e terminando com um grande *tutti* orquestral.

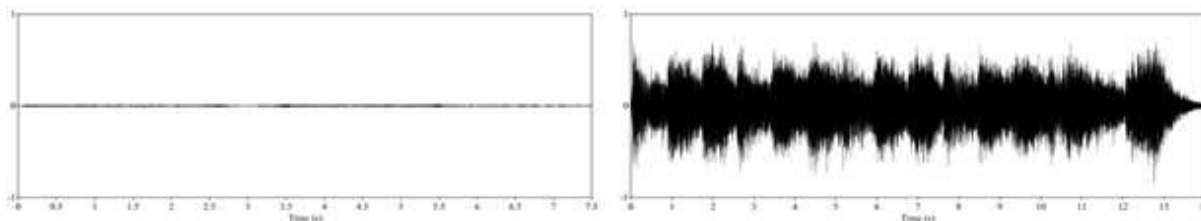


Figura 3: Forma de onda nos compassos iniciais (esq.) e finais (dir.) do *Bolero*, de Ravel.

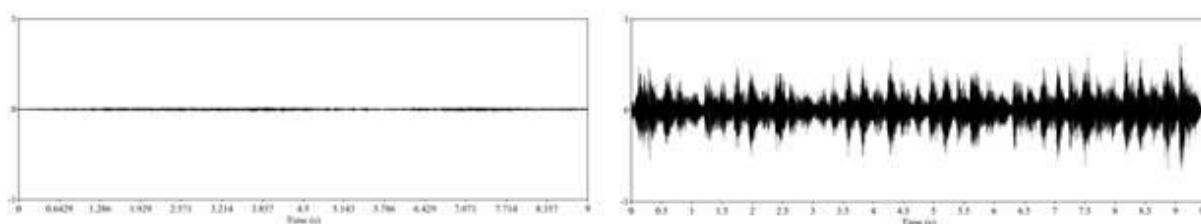


Figura 4: Forma de onda nos compassos iniciais (esq.) e finais (dir.) da *Sagração da Primavera*, de Stravinsky. Foi mantida a mesma escala de amplitude da figura anterior.

Os gráficos que se seguem foram feitos utilizando o software de análise sonora Praat, que atribui uma intensidade máxima de 90 dB para um sinal (teórico) de 0 dBFS (RMS). As formas de onda dos trechos foram colocadas acima (Figs. 3 e 4) para melhor visualização das variações energéticas, uma vez que é o gráfico mais usual.

No caso do Bolero, analisando em conjunto as Figs. 5 e 6 (abaixo), podemos verificar que há uma variação dinâmica de aproximadamente 40dB entre os extremos da dinâmica musical. Esses valores são RMS, o que significa que há valores de pico ainda mais extremos. Isso dificulta bastante a captação, posto que tais picos, embora não contribuam consideravelmente para o valor RMS, podem causar saturação na entrada e na saída do conversor analógico-digital-analógico (A/D/A), resultando em distorção.

⁶ RAVEL, Maurice: *Bolero*.

⁷ STRAVINSKY, Igor: *Sagração da Primavera*.

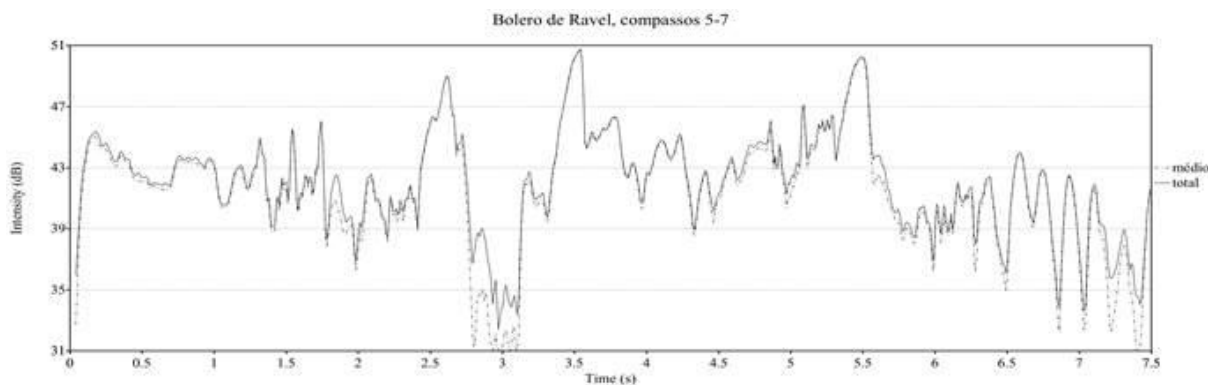


Figura 5: Envelopes dinâmicos total e das oitavas médias nos compassos 5-12 do Bolero de Ravel.

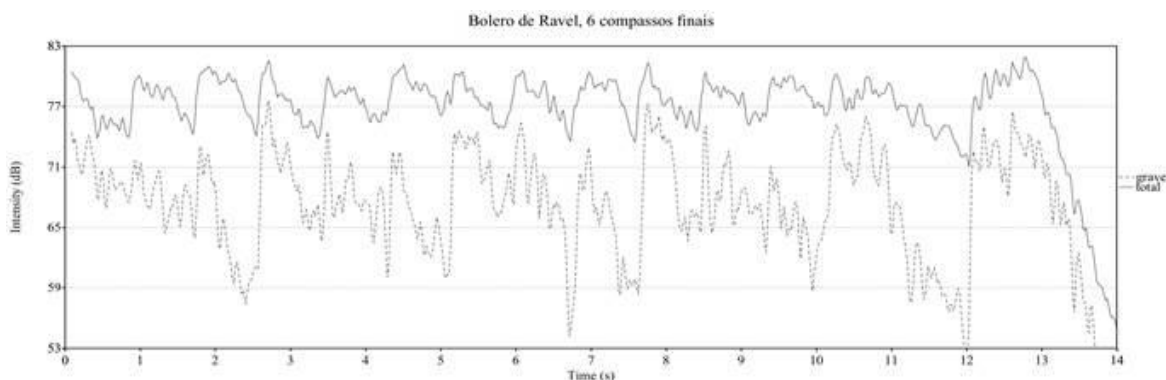


Figura 6: Envelopes dinâmicos total e das oitavas graves nos 6 compassos finais do Bolero, de Ravel.

Nas Figs. 7 e 8, para fins de comparação, estão duas versões de gravações contendo análises dos compassos 198 a 200 da *Sagração da Primavera*. Auditivamente, resta clara a diferença de concepção, já que na segunda versão (Fig. 8) há uma presença do bumbo orquestral um tanto mais acentuada.

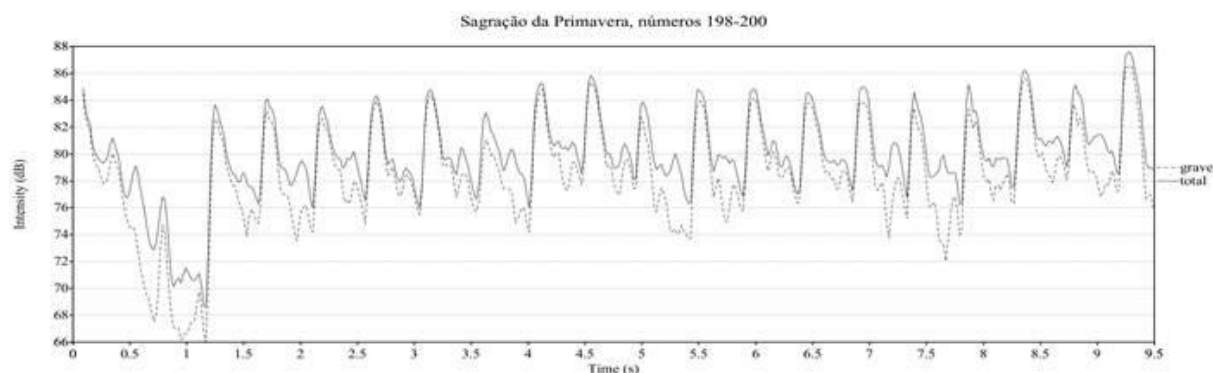


Figura 7: Envelopes dinâmicos total e das oitavas graves nos compassos 198-200 da Sagração da Primavera (Versão A).

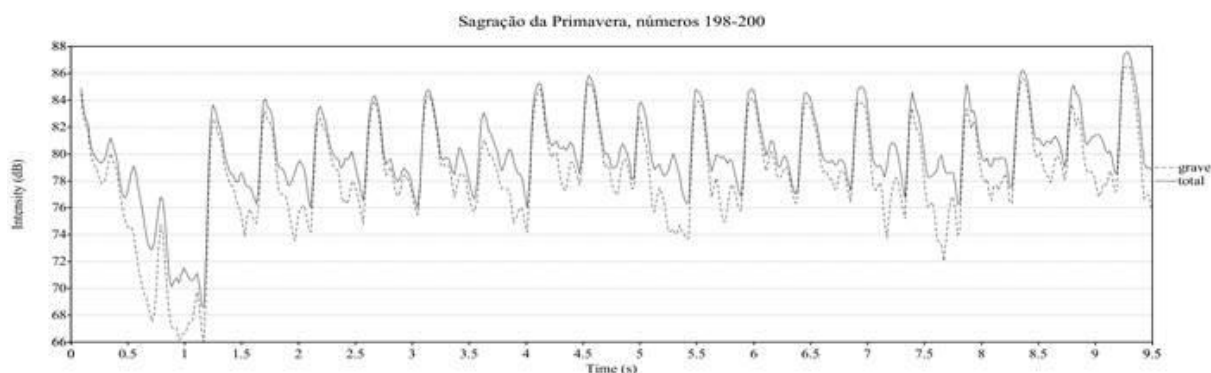


Figura 8: Envelopes dinâmicos total e das oitavas graves nos números 198-200 da Sagração da Primavera (Versão B).

Interessante observar que, no trecho onde o bumbo aparece com maior predominância, os graves acabam por dar forma à curva energética total, aproximando-se principalmente em seus valores máximos, onde chegam a atingir 0,5 dB em relação à curva de energia total. De acordo com a tabela 1, significa dizer que o sinal da porção grave ocupa grande parte de toda a gama dinâmica disponível. Uma análise conjunta entre a percepção auditiva e os dados objetivos retirados dos gráficos nos dão uma amostra de quão definidores do contorno energético total podem ser os sons graves, especialmente nas passagens fortes. Nas passagens com pouca energia total, portanto, seria correto presumir que a presença de componentes graves é mínima, e que o contorno energético acaba sendo determinado pela região média. Tal observação corrobora a análise feita dos dados apresentados na Fig.1, uma vez que é na região média que o ouvido humano tem maior sensibilidade, não sendo necessária grande quantidade de energia para que o som seja percebido.

2.2 Breve estudo de caso

A fim de observar as opções de distribuição espectral dentro de um determinado estilo, foi criado um estudo de caso - pequeno trecho instrumental com bateria, contrabaixo e guitarras, aproximando-se em estilo do *funk* norte-americano - que, durante a coleta de dados na pesquisa empírica, foi mixado por quatro profissionais de áudio de Belo Horizonte/MG. Uma breve descrição dos procedimentos de captação utilizados:

Foi montada uma barreira lateral em forma de semicírculo ao redor da bateria para evitar reflexões, notadamente excessivas no auditório. As peças de maior interesse foram microfônicas separadamente. O interesse foi determinado pela predominância dos graves no espectro sonoro emitido de cada peça. O trecho foi inicialmente tocado de forma convencional, todas as peças da bateria num mesmo passo. Em seguida, ouvindo o trecho gravado, o músico executou cada peça em separado, de modo que não há vazamento do som

de cada peça nos demais microfones. O resultado da *mixagem* é, portanto, a soma dos canais que contém o material das peças gravadas separadamente. O sinal gravado pelos microfones *Over* representa os tom-tons e os pratos de ataque e condução simultaneamente. O chimbau foi gravado em microfone individual e separadamente. Todas as gravações e *mixagens* foram feitas em 24 bits de faixa dinâmica e 48 kHz de taxa de amostragem e foram renderizadas em mono para maior precisão das análises de intensidades e de espectro sonoro. Os níveis de cada microfone foram ajustados de modo a aproveitar toda a faixa dinâmica disponível sem picos de saturação durante todo o teste. O trecho escolhido para análise contém as variações dinâmicas intrínsecas ao material musical. As guitarras e o baixo foram gravados conectados por cabos diretamente às entradas de instrumento da interface de áudio. A descrição detalhada deste teste encontra-se na referida dissertação.

Os materiais manipulados por cada um dos sujeitos estão indicados pela letra S seguida de um número, no topo do gráfico⁸. Os gráficos das análises de intensidade de cada faixa de frequências e também do contorno energético total estão nas Figs. 9 a 12. A partir da leitura destes gráficos e das impressões auditivas pode-se perceber que o tratamento dado aos graves dá a estes um papel de destaque sobre as demais regiões em todas as mixagens do teste. É possível perceber ainda que a proporção entre os graves e as demais faixas varia consideravelmente, embora o contorno das curvas graves não seja tão alterado quanto o das demais. Sobre esses contornos, vale notar o quanto são distintas as abordagens sobre os agudos. Nesse sentido nota-se uma similaridade entre as mixagens das Figs. 9 e 11 e entre aquelas das Figs. 10 e 12. Nas primeiras as curvas dos agudos se encontram mais "achatadas", possivelmente devido a uma maior compressão aplicada ao material, enquanto nas demais o que se percebe é um contorno mais oscilante. O resultado sonoro dessa última abordagem pode ser descrito como mais "escuro", com "menos brilho". Da comparação entre as mixagens pode-se extrair que, apesar de tratarem de um mesmo material gravado original, de um mesmo estilo musical e de uma mesma instrumentação, o resultado sonoro é único para cada uma das abordagens. Tal constatação nos leva a inferir que, para além do conceito de estilo, que é fruto de um consenso coletivo, a concepção estética do profissional influencia diretamente o resultado sonoro da *mixagem* por ele produzida.

⁸ Todas as gravações analisadas no artigo estão disponíveis para *download* em:
<https://www.dropbox.com/1/C93nZ9q6dkfUyEjfHMOcZp>

Considerações finais

As diferenças de conteúdo espectral e de contorno dinâmico entre a peça composta para o teste e as peças orquestrais analisadas são claras aos ouvidos. É notório que trata-se de materiais musicais muito distintos. A presença da região grave é mais marcante nas *mixagens* do teste do que nas peças orquestrais utilizadas. Há, no conjunto dos materiais *mixados* para o teste, uma variação tão grande de escolhas estéticas que somos levados a levantar questões a respeito da diversidade de estilos de *mixagem* existentes e da interferência destes nas características sonoras dos materiais trabalhados. Uma mesma peça, como visto, pode ter versões com sonoridades muito diferentes. As diferenças podem ser resultado tanto da variação entre os arranjos musicais como do fato de serem fonogramas produzidos por profissionais cujas concepções estéticas sejam muito próprias. De uma forma ou de outra podemos afirmar que, para além de escolhas técnicas, o profissional envolvido na produção musical está sempre em busca de um resultado estético, que é fruto de uma determinada concepção de sonoridade. Porém, ainda que existam infinitas possíveis concepções, as limitações e características do aparelho auditivo irão determinar boa parte das escolhas durante o processo de produção. As demais escolhas seriam pautadas pelo diálogo entre o imaginário sonoro do ouvinte ao qual aquela produção se destina e as novidades que o material produzido possam trazer a ele.

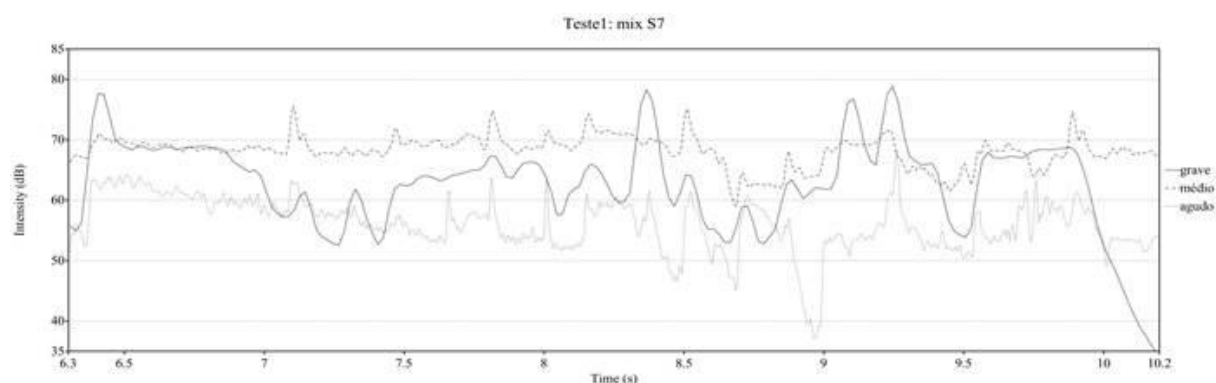


Figura 9: *Mix* teste 1 Sujeito 7.

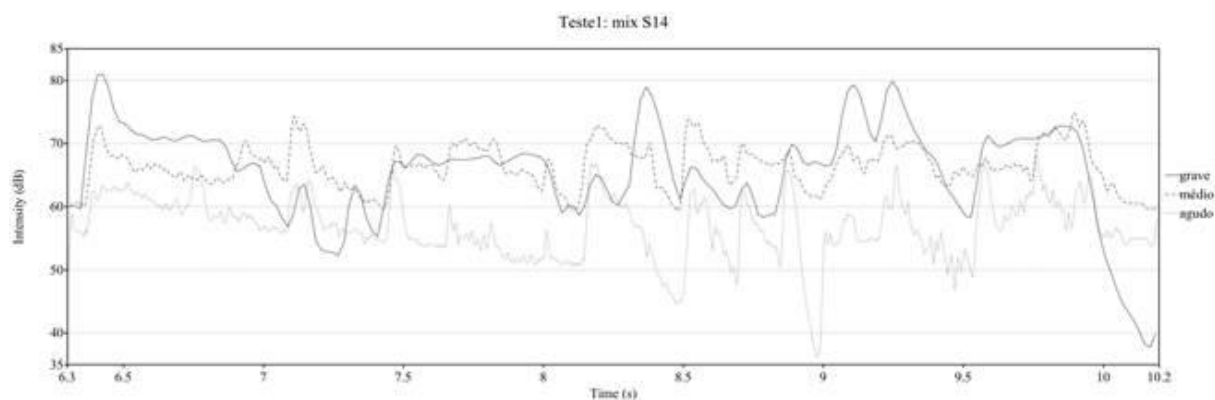


Figura 10: Mix teste 1 Sujeito 12.

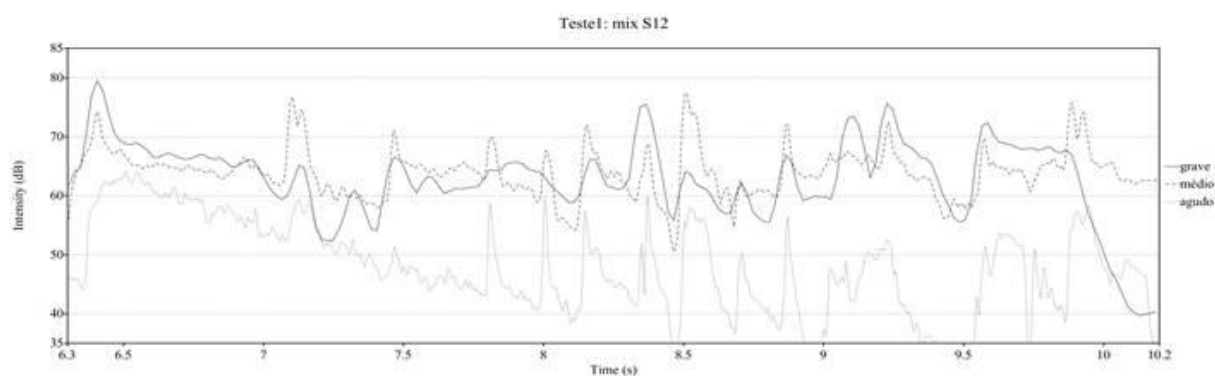


Figura 11: Mix teste 1 Sujeito 14.

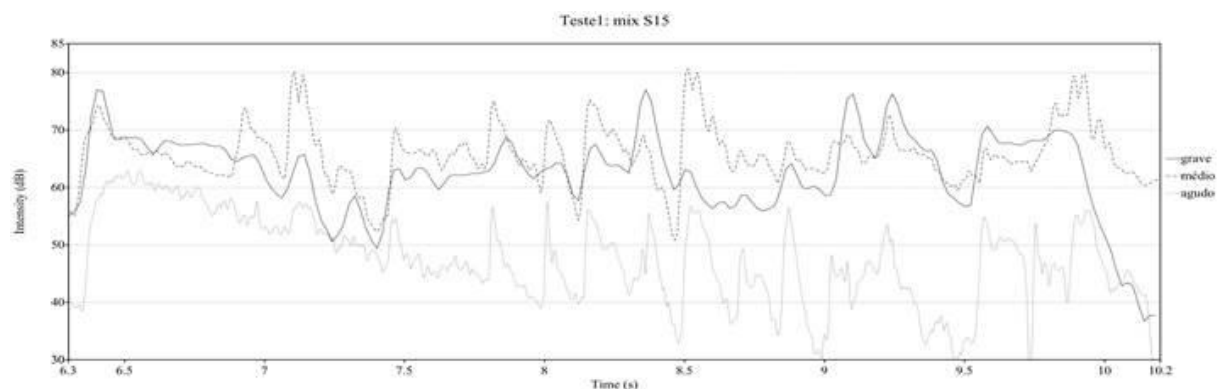


Figura 12: Mix teste 1 Sujeito 15.

Referências

BOLERO. Maurice Ravel (Compositor). The Boston Pops Orquestra (Intérprete). Jerome Rosen (Regente). Boston: Universal Music, 2007. [Compact Disc]

CAMPOS, Fernando. *A Região Grave Do Espectro Sonoro: aspectos teóricos e práticos na produção musical*. Belo Horizonte, 2013. 146f. Dissertação (Mestrado em Música-Sonologia). Universidade Federal de Minas Gerais, Belo Horizonte, 2013.

DAVIS, Gary; JONES Ralph. *Sound Reinforcement Handbook*. 2nd edition. Buena Park, CA: Hal Leonard, 1990.

FLETCHER, H.; MUNSON, W.A. Loudness, its definition, measurement and calculation. *Journal of the Acoustic Society of America*, 5v., 82-108, 1933.

GELFAND, Stanley. *Hearing: an introduction to psychological and physiological acoustics*. 5th edition. New York: Informa Healthcare, 2009.

LE SACRE DU PRINTEMPS. Igor Stravinsky (Compositor). Cleveland Orchestra (Intérprete). Lorin Maazel (Regente). Cleveland: Telarc, 1983. [Compact Disc]

LE SACRE DU PRINTEMPS. Igor Stravinsky (Compositor). London Symphony Orchestra (Intérprete). Claudio Abbado (Regente). Berlim: Deutsche Grammophon, 1986. [Compact Disc]

PADOVANI, José; FREIRE, Sérgio. Explorando envoltórias espectrais em sistemas musicais interativos. In: XVI CONGRESSO DA ANPPOM, 16., 2006, Brasília. *Anais....* Brasília: p. 280-287, 2006.

ROBINSON, D.W.; DADSON, R.S. A re-determination of the equal-loudness relations for pure tones. *British Journal of Applied Physics*, 7v., 5n., 166-181, 1956.

ROSSING, Thomas. *Springer Handbook of Acoustics*. New York, NY: Springer, 2007.