Paulo Assis Barbosa

Masterização e dispositivos portáteis: reflexões sobre
a otimização de áudio para a escuta móvel privada

Dissertação apresentada ao Programa de
Pós-Graduação em Música da Escola de
Comunicações e Artes da Universidade
de São Paulo para obtenção do título de
Mestre em Música.

Área de concentração: processos de
criação musical

Orientador: Fernando Henrique de
Oliveira Iazzetta

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

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

Barbosa, Paulo Assis

Masterização e dispositivos portáteis:: reflexões sobre a
otimização de áudio para a escuta móvel privada / Paulo Assis
Barbosa ; orientador, Fernando Henrique de Oliveira
Iazzetta. -- São Paulo, 2019.
140 p.: il. + DVD de dados.

Dissertação (Mestrado) - Programa de Pós-Graduação em Música
- Escola de Comunicações e Artes / Universidade de São Paulo.
Bibliografia
Versão corrigida

1. Escuta 2. Música 3. Mediação 4. Masterização de áudio I.
de Oliveira Iazzetta, Fernando Henrique II. Título.

CDD 21.ed. - 780

Elaborado por Alessandra Vieira Canholi Maldonado - CRB-8/6194

BARBOSA, P. A. **Masterização e dispositivos portáteis:** reflexões sobre a otimização de áudio para a escuta móvel privada. 2019. Dissertação (Mestrado em Processos de Criação Musical) - Escola de Comunicação de Artes, Universidade de São Paulo, São Paulo, 2019.

Aprovado em:

Banca Examinadora

Prof. Dr. _____
Instituição: _____
Julgamento: _____

Prof. Dr. _____
Instituição: _____
Julgamento: _____

Prof. Dr. _____
Instituição: _____
Julgamento: _____

À Laura, que me faz querer ser melhor.

Agradecimentos

À Laura, a quem eu nunca consigo encontrar palavras suficientes, e mesmo assim apoia minhas loucuras e segura minha mão quando preciso. Ou seja, o tempo todo.

Ao meu pai Ricardo, que me transmitiu a paixão por equipamentos e tecnologias, e à minha mãe Rosângela, que me ensinou a ouvir. Ao Roberto, de quem eu aprendi a questionar tudo (será?), e à Marilena, de quem eu roubei o hábito de tomar notas. À minha irmã Déborah e aos meus cunhados e cunhadas queridos, e suas alegrias em formato de sobrinhos.

Ao Fernando Iazzetta, que acreditou em minhas ideias iniciais e, com muita paciência, soube orientar a pesquisa e me tirar dos bicos em que eu invariavelmente me colocava.

Aos colegas que me incentivaram desde a criação do projeto de pesquisa, André Martins, Micael Antunes da Silva, Luzilei Aliel e Yuri Baer, essa enrascada só começou graças à ajuda de vocês. Aos queridos Julián Jaramillo e Esteban Astorga, que escreveram comigo um artigo essencial para a dissertação.

Aos amigos de tantos outros lugares, que dividiram comigo as dúvidas e preocupações com as distrações necessárias para continuar. Clara, Marcel, Tati, Felipe, Joana, Roberta, Diego, Maria Mercedes, Felipe, Luiz, Lilian, Anne: agora posso voltar a falar de outros assuntos!

À banca de qualificação, José Augusto Mannis e Julián Jaramillo, pelos caminhos apontados que direcionaram partes da pesquisa para o rumo correto.

À banca examinadora, Marcelo Wanderley, José Augusto Mannis e Fernando Iazzetta, que aportaram observações e ajustes a essa versão corrigida.

Aos professores e amigos que fiz durante o mestrado, Rogério Costa, Sílvio Ferraz, Régis Rossi, Heloísa Valente, José Augusto Mannis e todos os colegas do NuSom, da Sonora, do GPI, do NEAC e do AMU, que eu não vou citar nominalmente com medo de ser injusto. Vocês sabem que são queridos!

À CAPES, pelo apoio à pesquisa. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Esse período foi incrível, mesmo se, por vezes, estressante. Obrigado por me aturarem.

Resumo

BARBOSA, P. A. **Masterização e dispositivos portáteis:** reflexões sobre a otimização de áudio para a escuta móvel privada. 2019. Dissertação (Mestrado em Processos de Criação Musical) - Escola de Comunicação de Artes, Universidade de São Paulo, São Paulo, 2019.

Esta pesquisa trata das relações entre a masterização do áudio, última etapa de produção de um fonograma antes de sua replicação ou distribuição, e a escuta móvel privada, ou seja, a escuta por meio de fones de ouvido em dispositivos portáteis. O primeiro capítulo apresenta as origens da masterização e os problemas e desafios surgidos com a digitalização da música, resultando na “guerra de volumes”, e com a virtualização do fonograma, mudando paradigmas da masterização, com a compactação do áudio e os sistemas de normalização de intensidade. Um estudo de caso de uma música masterizada diversas vezes ilustra as tendências técnicas e estéticas da masterização ao longo de cinco décadas. O Capítulo 2 discorre sobre o discurso da alta-fidelidade e a proposta divergente dos dispositivos portáteis. Uma análise da cultura de escuta por meio de fones de ouvido localiza esse dispositivo como peça determinante na escuta privada. Também é realizada uma reflexão sobre as situações de escuta baseadas no uso dos dispositivos portáteis, e sobre o encontro das ideias de alta-fidelidade e alta-conveniência, propondo-se como o conceito de masterização contextual pode contribuir com a personalização do áudio para o usuário. O Capítulo 3 consiste na apresentação de exercícios de masterização, ilustrando os processamentos de áudio e situações de escuta discutidas nos capítulos anteriores.

Palavras-chave: Escuta. Música. Mediação. Masterização de áudio.

Abstract

BARBOSA, P. A. **Audio mastering and portable devices:** thoughts on audio optimization for private mobile listening. 2019. Dissertação (Mestrado em Processos de Criação Musical) - Escola de Comunicação de Artes, Universidade de São Paulo, São Paulo, 2019.

This research addresses the relationship between audio mastering, the last step in audio production process before replication or distribution, and private mobile listening, that is, music listening through headphones in portable devices. The first chapter introduces the origins of audio mastering and the problems and challenges that have arisen with the digitization of music, resulting in the “loudness wars”, and phonogram virtualization, changing audio mastering paradigms with audio compression and loudness normalization systems. A case study of music mastered several times illustrates the aesthetic trends of audio mastering over five decades. Chapter 2 discusses the discourse of high-fidelity and the divergent proposal of portable devices. An analysis of the listening culture through headphones finds the device as a key player in the listening privatization, followed by an analysis of listening situations based on the use of portable devices, and the meeting of high-fidelity and high-convenience ideas, in which contextual mastering can contribute to the customization of audio for the user. Chapter 3 consists of mastering exercises, illustrating the audio processing and listening situations discussed in previous chapters.

Key Words: Listening. Music. Mediation. Audio mastering.

Sumário

Introdução	15
1. Masterização	19
1.1. A fita magnética	20
1.2. Uma cadeia de produção	21
1.3. Da transferência para manipulação	22
1.4. Masterização digital e a guerra de volumes	23
1.5. <i>Streaming</i> e o fim da guerra	25
1.6. A masterização como processo artístico	28
1.7. Aspectos do áudio na masterização	29
1.7.1. Medição de intensidade	31
1.7.2. Dinâmica do áudio	32
1.7.3. Frequências	35
1.8. Estudo de caso: Tales from Topographic Oceans	36
1.8.1. Quatro versões e duas curiosidades	38
1.8.2. Equilíbrio tonal	42
1.8.3. Dinâmica	46
1.8.4. Conclusões do estudo	52
2. Dois caminhos da escuta	55
2.1. O discurso da alta-fidelidade: fiel em relação a quê?	56
2.2. A ideia de alta-conveniência: portabilidade e praticidade	57
2.3. A cultura de escuta por meio de fones de ouvido	59
2.3.1. Uma <i>commodity</i> tecnológica	60
2.3.2. Equipamentos de escuta	62
2.3.3. O (pseudo)cientificismo mercantil	63
2.3.4. Uma estratégia urbana	65
2.3.4.1. Os efeitos do Walkman	66
2.3.4.2. Os fones de ouvido na era do computador	68
2.3.4.3. A transformação em item de ostentação	70
2.3.4.4. Áudio localizado	71
2.3.5. Conclusões sobre os fones de ouvido	72
2.4. Da escuta privada à personalizada	74
2.4.1. Manuseio e práticas de escuta	75

2.4.2. Virtualização do fonograma	76
2.4.3. <i>Singles</i> , álbuns e <i>playlists</i>	77
2.4.4. iPod	79
2.4.5. <i>Smartphones</i>	81
2.4.5.1. Armazenamento	81
2.4.5.2. Conexão	82
2.4.5.3. Processamento	84
2.4.6. A escuta personalizada	85
2.4.7. Situações de escuta	87
2.4.7.1. A escuta que mascara	87
2.4.7.2. A escuta de fruição	88
2.4.7.3. A escuta distraída	89
2.5. Os desafios da masterização hoje	89
2.5.1. Hi-fi no <i>smartphone</i> ?	92
2.5.2. Problematizações sobre os fonogramas virtualizados	97
2.5.3. Por uma masterização contextual	102
3. Exercícios de masterização	107
3.1. Faixas utilizadas nos exercícios	109
3.2. Preparação e <i>softwares</i>	112
3.3. Normalização de intensidade	113
3.4. Alteração de extensão dinâmica	115
3.5. Mascaramento do ambiente	118
3.6. Equilíbrio tonal	120
3.7. Espacialização	122
3.8. Compactação de arquivo	123
3.9. Masterizações contextuais	125
3.9.1. A escuta que mascara	125
3.9.2. A escuta de fruição	127
3.9.3. A escuta distraída	128
3.10. Conclusões dos exercícios	129
Considerações finais	131
Referências	133
Anexo I - faixas do DVD	139

Introdução

Até o final do século XIX, a música era um evento efêmero, ocorrido apenas na presença de musicistas – amadores ou profissionais. Com o surgimento da fonografia, passa a ser possível captar, armazenar e reproduzir as vibrações do ar que percebemos como sons. O registro sonoro torna-se, ao longo do século XX, muito mais cotidiano do que a música executada ao vivo (IAZZETTA, 2012, p. 19).

Uma das primeiras tecnologias de fixação de áudio utilizava cones acústicos para captar as vibrações do ar e gravar mecanicamente cilindros de cera, um a um. As tecnologias subsequentes foram aumentando a capacidade de replicação e a qualidade de reprodução, como a prensagem de discos a partir de matrizes, a eletrificação do processo de captura e a adoção de materiais mais adequados para suporte, como o vinil.

No final dos anos 1940, a introdução da fita magnética como suporte de fixação intermediário permite a separação entre os momentos de captura e de criação do disco matriz de replicação (CUMMINGS, 2013). A partir dessa ruptura, são ampliadas as etapas de produção do registro sonoro, com captações dos sinais sonoros em camadas, e posterior redução do número de canais no processo de mixagem. A masterização surge como a última etapa criativa no processo de produção de áudio (KATZ, 2015), sua manipulação final antes da replicação em determinado suporte ou, mais recentemente, antes da distribuição eletrônica.

Os gramofones, rádios, toca-discos e *home theaters*, que substituíram a presença do piano na sala de visitas das famílias, criam novos modos de escuta, dissociando, no tempo e no espaço, os momentos da performance e da escuta (IAZZETTA, 2009). A ideia de *alta-fidelidade* (*hi-fi*) surge a partir da necessidade da indústria do áudio de convencer os consumidores sobre a equivalência entre a reprodução sonora de seus equipamentos com a experiência da música ao vivo, criando uma cultura de escuta tecnológica (TAYLOR, 2001).

Paralelamente a essa cultura *hi-fi*, surgem equipamentos voltados para a portabilidade, como vitrolas portáteis e rádios de pilha. A grande popularização do Walkman no início da década de 1980 gera uma outra cultura de escuta, móvel e privada, a partir do uso de fones de ouvido no ambiente urbano (HOSOKAWA, 2012). Essa cultura de escuta de *alta-conveniência* expande-se menos de duas décadas depois, com o surgimento dos arquivos digitais de áudio compactado sendo distribuídos na internet e reproduzidos em dispositivos portáteis como iPods e smartphones (CUMMINGS, 2013, p. 220).

O onipresença de smartphones e a substituição das vendas físicas de álbuns em *compact-disc* pela distribuição *on-line* de *singles* e serviços de *streaming* alteraram os hábitos de escuta. Assim, cada vez mais, a fruição musical se dá como uma atividade secundária, paralela a outras (STERNE, 2012, p. 225). Ela acontece, por exemplo, nos trajetos urbanos, onde o usuário de um dispositivo portátil de reprodução sonora com fones de ouvido encontra novos usos para suas músicas, utilizando sua *playlist* como trilha sonora para sua escuta privada (BULL, 2012). Isso cria situações de escuta diferentes do cenário de fruição musical construído pelo discurso *hi-fi*.

Na finalização do áudio para mídias físicas, parece lógico adequá-lo às limitações de seu suporte, buscando preservar a qualidade de reprodução, sem considerar as possíveis situações de escuta. Os áudios distribuídos eletronicamente, por outro lado, abrem outros desafios para os engenheiros de masterização, que são os responsáveis por adaptar cada áudio para a sua replicação ou distribuição. Assim, eles passam a ter que considerar se priorizarão - e como - as novas situações de escuta em seu trabalho.

Como engenheiro de masterização e usuário de dispositivos portáteis de reprodução sonora, várias vezes escutei as mesmas músicas em ambiente acústico de estúdio e em fones de ouvido nos trajetos pela cidade. Muitas masterizações não me parecem compatíveis com algumas das situações de escuta: grandes variações de intensidade entre passagens musicais, por exemplo, fazem uma música ficar ou alta demais, ou submergir no barulho da cidade. Outras masterizações funcionam bem em fones de ouvido, mas sua apreciação em estúdio mostra que, para obter qualidade na escuta portátil, às vezes é necessário sacrificar a qualidade da reprodução em equipamentos mais sofisticados. Como o uso de dispositivos portáteis não parece encontrar lugar no discurso *hi-fi*, a conveniência da escuta móvel privada e a ideia da alta-fidelidade se intercalam em minha vida, entre a rua e o estúdio.

Esta pesquisa parte dos questionamentos a respeito da adequação da masterização nas situações de escuta atuais, criadas a partir do uso de dispositivos portáteis. Como a masterização se adapta à substituição do suporte físico pelos arquivos digitais de áudio? Como ocorreu essa separação entre fidelidade e conveniência? Quais são os problemas enfrentados pelos engenheiros de masterização nas novas configurações de fruição musical? Quais soluções podem ser vislumbradas para responder às necessidades de uma escuta móvel privada?

A popularização dos *smartphones* e o número expressivo de vendas de *singles* em

formato eletrônico (RIAA, 2019) indicam uma tendência de forma de escuta que merecia ser estudada. Assim, o recorte escolhido para esta pesquisa foi refletir sobre as relações entre a masterização e a escuta móvel privada de músicas, ou seja, como a finalização do áudio se articula com a fruição musical, por meio de fones de ouvido, em ambientes urbanos.

O objetivo da dissertação foi analisar o problema da cisão entre os equipamentos *hi-fi* e os dispositivos portáteis, como ela cria novas situações de escuta e como a masterização lida com os problemas ocasionados por essa multiplicação de situações.

A escuta mediada, primeira etapa para compreender o surgimento da fonografia e suas consequências, é analisada por autores como Timothy Taylor (2001), Jonathan Sterne (2003), Fernando Iazzetta (2009) e Mark Katz (2010). Os dispositivos e a escuta portáteis são estudados por acadêmicos como Shuhei Hosokawa (2012) e Michael Bull (2012). O diálogo sobre masterização, assunto pouco tratado na academia, foi estabelecido nesta pesquisa com dois profissionais reconhecidos na área, Bob Katz (2015) e Bobby Owsinski (2017), e alguns pesquisadores do tema, como Robert Taylor (2018) e Rudolf Ortner (2012). Para complementar a pouca bibliografia acadêmica especializada, além de comentadores do meio musical, como David Byrne (2012), foram utilizadas outras fontes de escrita informal, como entrevistas e matérias de revistas e *sites* das áreas de áudio profissional e de equipamentos *hi-fi*.

Para realizar o percurso da pesquisa, foi necessário estudar a história da masterização e as ferramentas básicas de manipulação do áudio, fazendo uma análise de algumas masterizações diferentes de um mesmo material, o que contribuiu para construir um olhar sobre as variações estéticas ocorridas ao longo dos anos e dos suportes (Capítulo 1).

As mudanças ocorridas na escuta mediada são contempladas no Capítulo 2, que aprofunda um assunto negligenciado academicamente, o fone de ouvido, responsável direto pela privatização da escuta. A trajetória do fonograma digital para dentro dos dispositivos portáteis também é analisada, para chegar ao encontro do discurso da alta-fidelidade com os equipamentos de alta-conveniência, e procurar o lugar da masterização nas formas de escuta do século XXI.

A relevância do que analisamos ao longo da pesquisa é ilustrada a partir de exercícios de masterização (Capítulo 3), que materializam a discussão teórica, revelando alguns dos problemas e desafios da masterização, e testando caminhos para um eventual encontro da qualidade de reprodução sonora com a qualidade de escuta privada.

1. Masterização

As primeiras tecnologias de fixação de som não permitiam edição. Uma delas, o fononautógrafo de Leon Scott, nem sequer conseguia converter novamente o que captava em som, apenas fixando-o em uma representação gráfica (STERNE, 2003, p. 35). Quando surgiram os primeiros dispositivos comerciais de captura e reprodução de som, o processo era inteiramente mecânico: o fonógrafo de Edison, comercializado a partir de 1877, e o gramofone de Emile Berliner, patenteado em 1887, entre outros inventos semelhantes do período, reproduziam som a partir da amplificação mecânica das vibrações de uma agulha de leitura passando pelos sulcos de um cilindro de cera ou de um disco de goma-laca.

A gravação desses sulcos, por sua vez, dava-se de forma inversa: um cone concentrava as vibrações do ar para a ponta de uma agulha de gravação, que riscava o material do cilindro ou do disco. Essa transdução mecânica de som em sulcos não podia ser editada - no caso dos cilindros de cera, nem mesmo replicada industrialmente, já que cada cilindro era o produto de uma performance. Nos modelos mais eficientes de produção de cilindros, vários gravadores eram dispostos no ambiente da performance, captando, de diferentes posições (mesmo que lado a lado), o material sonoro executado (KATZ, 2010, p. 43).

Como aponta David Byrne (2012, p. 99), ao final da segunda guerra mundial, algumas tecnologias alemãs, capturadas pelas tropas norte-americanas, foram estudadas e lançadas comercialmente nos Estados Unidos. Entre elas estava a gravação de áudio em fita magnética. Já havia alguns experimentos nesse sentido, mas a tecnologia alemã era muito superior, criando a dúvida nos soldados americanos sobre se o que ouviam nas rádios nazistas eram gravações ou orquestras tocando ao vivo nas estações de rádio, em meio à guerra.

O gravador de rolo de fita magnética foi comercializado a partir de 1948, mudando o paradigma de qualidade de captação e reprodução. As emissoras de rádio norte-americanas logo adotaram esse equipamento: a qualidade sonora da gravação, agora, era comparável à qualidade da transmissão de uma performance ao vivo (CUMMINGS, 2013, p. 71). No mesmo ano, os discos de vinil começam a ser vendidos, e a combinação dessas duas tecnologias - o disco de vinil como equipamento caseiro de reprodução musical e a fita magnética como equipamento de gravação e edição em estúdio - tem como resultado, entre outras consequências, a inserção de novas etapas na cadeia produtiva da indústria fonográfica¹ (BARBOSA, 2016, p. 42).

1 Desde que os registros sonoros começaram a ser comercializados - ou seja, a partir da venda de cilindros pré-gravados de Edison para reprodução caseira - uma indústria foi se estabelecendo para

1.1. A fita magnética

Com a disseminação das redes de energia elétrica a partir das duas últimas décadas do século XIX, surgem várias tecnologias, como a lâmpada e o elevador. As redes comerciais de transmissão de rádio, implantadas durante a década de 1920, utilizam algumas dessas novas tecnologias: microfones condensadores e processadores de áudio como amplificadores e compressores.

A indústria fonográfica, criada junto com os cilindros e discos de cera, passa a utilizar equipamentos elétricos a partir da década de 1920 (CUMMINGS, 2013, pg. 47), com microfones e amplificadores otimizando a captação de som antes de sua gravação mecânica em uma matriz de replicação (o molde utilizado na fabricação dos discos).

Essas tecnologias elétricas são utilizadas, no final dos anos 1940, junto a duas novidades do pós-guerra que mudam a qualidade de reprodução e o custo de replicação: a fita magnética como meio intermediário de fixação, durante a produção; e o vinil, material mais barato que a goma-laca, para a replicação em larga escala dos discos. Isso amplia a popularização e o consumo de música gravada (CUMMINGS, 2013).

A fita magnética viabiliza a separação dos momentos de performance e de criação da matriz de replicação. Não cabe aqui o levantamento de todas as consequências dessa ruptura temporal entre a performance dos músicos e a fixação em um suporte, mas citamos, como exemplos, a gravação em camadas não simultâneas (*overdub*) e todo tipo de manipulação de fita magnética, incluindo colagens e efeitos como *delays* e *loops*. Antes da fita magnética, apenas um engenheiro de áudio era responsável pela transdução de som para sinal elétrico, tratamento de sinal e transferência para o disco matriz, o que ocorria simultaneamente. A partir da introdução da fita magnética no processo de gravação de discos, surgem as figuras do engenheiro de gravação e do engenheiro de transferência (*cutter*), esse último, futuramente chamado de engenheiro de masterização (OWSINSKI, 2017). A separação de funções fica ainda mais evidente quando, em 1955, a empresa Ampex inicia a venda comercial de gravadores de fita magnética de vários canais, ampliando as possibilidades de manipulação pós-captação e fazendo surgir a figura do engenheiro de mixagem.

produzir, comercializar e lucrar com fonogramas (aqui entendido como o registro sonoro em um suporte físico que permita sua reprodução). Ela inclui: aquisição de direitos autorais, contratos jurídicos com intérpretes e campanhas de marketing, além de distribuição, comercialização e produção de fonogramas. Nosso foco nesta pesquisa, no entanto, é a masterização e as situações de escuta dos fonogramas, nos interessando apenas analisar sua produção e fruição, não cabendo aprofundamentos sobre outras etapas dessa cadeia de produção.

1.2. Uma cadeia de produção

A separação entre a performance e a fixação definitiva cria novas etapas na cadeia de produção da indústria fonográfica. Como já mencionado, o que antes era um complexo processo de captar e simultaneamente gravar uma mídia matriz, é desmembrado em diversas etapas, não simultâneas, que vamos aqui dividir em três momentos distintos: captação, mixagem e masterização.

Na captação, são fixados em suporte intermediário diferentes sinais sonoros, gerando fonogramas independentes. Os sinais podem ser captados simultaneamente de uma mesma fonte ou recolhidas em tempos e locais distintos, e podendo ser utilizados equipamentos e processamentos como microfones, pré-amplificadores e mesmo alguns efeitos - dependendo da época e dos recursos disponíveis da gravação. Na década de 1950, esse suporte de fixação intermediário era a fita magnética e, apesar de ter sido amplamente substituída por tecnologias digitais décadas depois, ela é encontrada até hoje em alguns estúdios, para clientes que buscam sonoridades analógicas em suas produções (OWSINSKI, 2008, p. 26).

Na mixagem, esse conjunto de áudios captados é organizado, processado e convertido em um fonograma com o número de canais do suporte final de reprodução – por exemplo, dois para um registro sonoro estéreo ou oito para uma obra octafônica. Podem ser utilizados diversos aparatos tecnológicos de colagem, edição e alteração do sinal original, sendo essa uma etapa de criação em diversos gêneros e escolas musicais. A mixagem preocupa-se com as relações entre as fontes sonoras captadas, modificando suas características individuais para enfatizar aspectos de composição e performance desejados. As captações ganham processamentos como equalização, reverberação e compressão, os níveis de intensidade são ajustados e o palco sonoro é distribuído de acordo com a intenção estética dos artistas envolvidos.

Para Bob Katz (2015, p. 1), “masterização é a última etapa criativa no processo de produção de áudio, a ponte entre a mixagem e a replicação ou distribuição²”. Já Bobby Owsinski, apresenta duas definições para a masterização: na primeira, masterização é “[...] o processo de transformar uma coleção de músicas em um álbum, fazendo elas parecerem um conjunto em tom, volume e espaçamento entre elas” (OWSINSKI, 2008, p. 3). Sua segunda definição, atualizada para uma época com mais *singles* vendidos pela internet e menos álbuns vendidos fisicamente, afirma que “a masterização é o processo de ajuste fino

2 Todas as traduções de citações desta pesquisa são nossas.

do nível, do equilíbrio tonal e da *metadata* de uma faixa no preparo para sua distribuição” (OWSINSKI, 2017, não paginado).

Bobby Owsinski (2017) destaca uma significativa diferença no posicionamento da masterização nas cadeias produtivas dos Estados Unidos e do Reino Unido: enquanto a indústria norte-americana coloca a masterização como a última etapa de criação de um álbum, os britânicos localizam-na como a primeira etapa do processo de fabricação do suporte de fixação do áudio (mídia). Isso daria aos engenheiros norte-americanos, locados ainda na fase criativa da cadeia de produção, muito mais liberdade artística do que é dado aos colegas europeus. Já enxergar a masterização como um processo ligado diretamente a um suporte de áudio, como fazem os britânicos, ajuda a perceber sua relação com as contingências da reprodução sonora - os limites físicos de determinado suporte de áudio, as diferentes capacidades dos dispositivos de reprodução e as possíveis situações de escuta.

Na masterização, a ênfase está nas manipulações do sinal voltadas à transferência do áudio para determinado suporte, e no diálogo desse registro com as referências sonoras do ouvinte, afetando diretamente a percepção da escuta mediada por alto-falantes e fones de ouvido. Ela tenta trazer o registro sonoro de um plano idealizado de composição e sonoridade para uma realidade de escuta que, como analisaremos, pode, atualmente, estar também em um plano idealizado.

1.3. Da transferência para manipulação

A função inicial do engenheiro de transferência, ou *cutter*, era apenas passar o conteúdo de uma fita magnética para o disco matriz, que seria usado de base para a fabricação dos discos de vinil. Um álbum de grande tiragem, ou com replicação em diversos países, possuía muitas transferências para discos matriz. De acordo com Matt Shelvock (2012, p. 9), a utilização do estéreo nos discos de vinil ampliou o trabalho dos *cutters*, responsáveis agora pela melhora de uma “fidelidade” do som - algo relacionado com a espacialidade trazida pelo uso de dois canais de áudio. Na década de 1950, surgem equipamentos dedicados ao processamento para a transferência, e os engenheiros de transferência passam a ter maior controle na manipulação do som:

Tornou-se crucial para os engenheiros de transferência tomar decisões de natureza tanto técnica quanto estética, e a denominação “engenheiro de transferência” foi substituída por “engenheiro de masterização”. Se uma fita tivesse um conteúdo excessivo de frequências baixas, por exemplo, os graves teriam que ser atenuados para garantir que a agulha do toca-discos do usuário não saltasse do sulco (SHELVOCK, 2012, p. 9).

Com o aperfeiçoamento dos materiais e técnicas de transferência, vão ficando evidentes algumas vantagens na manipulação do áudio nessa etapa: não comprometido com a coesão sonora dos instrumentos e o tratamento de cada captação individualmente - tarefas dos engenheiros de mixagem -, o engenheiro de masterização pode focar na melhoria de outros aspectos estéticos, como a extensão dinâmica na faixa, e aspectos técnicos específicos, como ampliar a distância entre o conteúdo sonoro e o ruído de fundo do suporte utilizado. Bobby Owsinski cita também uma consequência direta dessa função da masterização:

Engenheiros de masterização [...] encontraram maneiras de aumentar o volume dos discos (e assim deixá-los com menos ruído), aplicando equalização e compressão. Produtores e artistas perceberam, então, que algumas gravações soavam mais altas no rádio e, por estarem mais altas, pareciam melhores aos usuários (apesar de que eles estavam especulando, e não usando dados científicos), e talvez por isso vendessem mais (OWSINSKI, 2008, pg. 4).

Podemos perceber nessa tentativa de aumento das vendas a chamada “guerra de volumes”, que fica mais evidente e causa distorções significativas no som quando, como analisaremos em seguida, os processamentos deixam de ser feitos por equipamentos analógicos e passam para o domínio digital.

1.4. Masterização digital e a guerra de volumes

Em qualquer mídia física utilizada como suporte de fixação de áudio há um ruído de fundo intrínseco, e quanto maior a relação sinal-ruído, maior a *extensão dinâmica* que pode ser registrada nesse suporte, ou seja, maior pode ser a diferença entre os momentos mais fortes e os mais suaves fixados nele.

Bob Katz (2015) sugere que os primeiros sinais de uma “guerra de volumes” (*loudness war*) são encontrados ainda no final do século XIX, quando as gravações nos cilindros de cera de Edison, muito barulhentos, deviam ser realizadas o mais forte possível para que o material captado não ficasse muito próximo ou mais baixo que o ruído de fundo do suporte e os próprios ruídos mecânicos do reproduzidor.

O advento da gravação por meios elétricos, na década de 1920, e o surgimento do disco de vinil, em 1948, aumentaram a extensão dinâmica disponível. A introdução do *compact-disc*, em 1982, forneceu uma *faixa de extensão dinâmica*³ (*dynamic range of an*

3 A faixa de extensão dinâmica refere-se à capacidade de dinâmica de um determinado suporte, medida em dB. A extensão dinâmica de um fonograma, por sua vez, refere-se à relação entre os níveis máximo e mínimo do sinal de um áudio (TAYLOR, 2018, p. 115).

audio system) ainda maior para mídias de consumo final (KATZ, 2015, p. 241), o que permitiu gravações com maior diferença de intensidade, tanto em sons rápidos - como os picos resultantes do ataque de um instrumento percussivo - quanto entre as músicas e entre suas partes. Além dessa faixa ter sido ampliada, o *compact-disc* não possui algumas limitações físicas do disco de vinil, que impediam o abuso nas intensidades do áudio fixado nesse suporte: sinais muito intensos geravam sulcos demasiado bruscos, ejetando a agulha de leitura de seu curso, e interrompendo a leitura (TAYLOR, 2018, p. 94).

Ao passar para o domínio digital, os métodos de medição dos níveis de intensidade do áudio também mudaram: a medição analógica marcava algo próximo de um nível de referência, enquanto a digital mostra cada ponto da resolução existente, explicitando os limites máximos para os picos do sinal. Bob Katz explora essa ideia em seu livro, mostrando que a medição de nível médio (*average-reading VU meters*) do domínio analógico foi substituída por uma medição de picos máximos das ondas (*peak metering meters*), com precisão matemática. Enquanto no domínio digital não há representação possível para valores acima de zero decibels relativos à escala inteira⁴ (dBFS, por sua sigla em inglês), no domínio analógico o *headroom* (faixa entre o nível nominal e o limite do sistema) é bastante flexível: a distorção analógica, ocasionada no limite dessa faixa de *headroom*, é resultado da resposta não-linear da compressão de componentes eletrônicos e da fita magnética, um efeito frequentemente utilizado pelos engenheiros de áudio como recurso estético (KATZ, 2015, p. 99).

Segundo Katz, essa diferença nas medições teria ocasionado uma mudança na maneira como os engenheiros de masterização percebiam o quanto ainda poderiam comprimir e aumentar os volumes dos materiais trabalhados: se antes havia uma referência média e picos soltos no *headroom*, agora, devido à facilidade no domínio digital de normalizar pelo pico do sinal (ou seja, levar a posição mais alta da onda para zero dBFS) a referência transforma-se no teto máximo, fazendo os engenheiros achatarem os picos cada vez mais. Como ilustra Katz (2015, p. 243), “estava aberta a caixa de Pandora” da “guerra de volumes” (*loudness war*). Robert Taylor (2018, p. 98) elenca vários estudos que apontam o ano de 2004 como o auge do excesso de compressão na música comercial.

4 Decibels relativos à escala inteira (dBFS) é uma escala logarítmica comumente utilizada para descrever áudio no domínio digital. Ela possui como valor máximo zero dBFS, e é definida pelos padrões AES 17-1998 e ITU-T P.38x. Valores superiores obtidos de algum processamento, como aumento de ganho excessivo, geram um máximo na onda em zero dBFS, chamado de *clipping*.

David Byrne (2012, p. 123) aponta a liberdade técnica do *compact-disc*, sem as limitações físicas do vinil, como responsável pela escalada do volume, com a diminuição da extensão dinâmica das músicas. Bob Katz, talvez com maior precisão técnica, afirma que, graças ao avanço das técnicas de processamento de áudio, as dinâmicas podem ser manipuladas cada vez com mais precisão, o que favorece a percepção de intensidade da música, mas diminui sua extensão dinâmica. O resultado, segundo Bob Katz, é que “o *compact-disc* trouxe consigo uma acelerada nova “guerra de volumes” e, aproximadamente em 1995, a volta estava completa. Estávamos fazendo gravações de música *pop* com extensões dinâmicas menores do que a do cilindro de Edison de 1909!” (KATZ, 2015, p. 241).

Por controlar essa última etapa do processo de manipulação do áudio, claramente os engenheiros de masterização são, em parte, responsáveis pelas consequências sonoras das músicas “vitimadas” pela “guerra de volumes”. Importante lembrar, porém, que essa demanda por fonogramas cada vez mais achatados (com o intuito de ficarem mais altos do que a concorrência) não vem dos engenheiros de masterização, mas de toda a indústria fonográfica, que tem dificuldade em perceber que a “guerra de volumes” é, em grande parte, também uma guerra contra a qualidade sonora, já que essa manipulação acaba criando distorções e perdas de dinâmica (KATZ, 2015, p. 244). Robert Taylor (2018, p. 354) nota que, apesar dos engenheiros de masterização terem o controle da dinâmica final, muitas mixagens já possuem compressão em excesso, deixando o engenheiro de masterização sem muitas opções. Ele também aponta a pressão dos artistas e dos executivos das gravadoras em todo o processo de gravação como responsável por essa escalada da compressão nas músicas comerciais.

1.5. *Streaming* e o fim da guerra

A “guerra de volumes” chega a um “empate” quando a maioria das músicas comerciais é lançada distorcida e com redução sensível de extensão dinâmica. Não há mais como destacar uma produção com aumento da intensidade das músicas, e a guerra tem seu auge durante o domínio dos *compact-discs* como suporte principal na comercialização de fonogramas (KATZ, 2015, p. 242).

Como veremos em maior detalhe no segundo capítulo, a virada do século XX para o XXI testemunha a virtualização do fonograma, ou seja, a passagem do fonograma de um suporte físico para um formato eletrônico como o *mp3*. Essa mudança impacta nos hábitos de fruição musical e vem acompanhada de outras novidades tecnológicas. A internet faci-

lita as trocas virtuais e os fonogramas, compactados em arquivos digitais de áudio *lossy*⁵, passam a ser trocados e ouvidos no computador e, logo depois, em dispositivos móveis como iPods e *smartphones*.

Um álbum em *compact-disc* ou em disco de vinil é concebido para ser ouvido como um conjunto. Arquivos de *mp3* pirateados na internet, extraídos e baixados individualmente, permitem ao usuário sua seleção e ordem de reprodução, independente dos álbuns de onde vieram. Essa liberdade, indutora da cultura das *playlists*, acaba expondo as diferenças de finalização em faixas de distintas origens, principalmente a disparidade de intensidade entre elas - saltar da reprodução de uma gravação de música clássica, cheia de sutilezas, para um intenso heavy metal pode ser, além de assustador, realmente perigoso para a saúde auricular. Assim, os fabricantes de dispositivos reprodutores procuraram estratégias para mitigar essa disparidade de volume e diminuir a inconveniência da alteração manual do volume a cada mudança de música.

As redes de televisão norte-americanas e europeias procuram, a partir dos anos 1980, implementar padrões para os níveis de intensidade de áudio entre a programação e as propagandas que, por razões semelhantes, passaram por um excesso de compressão e disparidade de níveis de intensidade em sua programação, especialmente após a virada de século, quando muitos de seus sistemas foram digitalizados. A introdução da norma ITU-R BS. 1770 nos Estados Unidos e da norma complementar europeia EBU R-128 estabeleceu parâmetros de níveis adequados baseados em intensidade sonora (*loudness*), passando de um sistema calibrado em picos máximos (*peak-normalization*) para um sistema de normalização de intensidade (*loudness-normalization*). Esses sistemas utilizam algoritmos que consideram a percepção humana de dinâmicas e frequências. Um cálculo é realizado a cada parte da programação que será transmitida (blocos de programas, comerciais, vinhetas) e um ajuste automático é aplicado para regular os diferentes materiais no mesmo nível de intensidade (ROBJOHNS, 2014, p. 168).

A indústria de dispositivos de reprodução de áudio adotou versões similares desses procedimentos. A implementada pela Apple em seus dispositivos (tanto móveis quanto em computadores) é chamada de Sound Check, que ajusta o volume das músicas para um patamar comum, baseado em um cálculo prévio da intensidade sonora de cada faixa. Essa

5 *lossy* (com perda), em oposição à *lossless* (sem perda) é um formato digital de áudio com partes de seu conteúdo original retiradas, visando otimizar o espaço que ele ocupa na memória de um dispositivo.

tecnologia de normalização de intensidade, porém, acaba anulando as vantagens psicoacústicas de aumentar o volume de uma música, como explica Bobby Owsinski:

O Sound Check automaticamente reduz o volume de músicas muito altas e aumenta as mais baixas para que elas soem com a mesma intensidade. Por conta da menor dinâmica da música alta, ela parece soar sem vida e até mais baixa do que a que foi menos comprimida, porque tem menos picos. Isso faz as pessoas repensarem a importância de uma masterização alta, já que pode ser contraproducente tocá-la em um sistema que controla os níveis. O Spotify é outro serviço que utiliza um tipo de Sound Check e outros podem também passar a utilizar algo semelhante (OWSINSKI, 2017, p. 147).

Os problemas causados pelo excesso de compressão, longamente analisados na tese de doutorado do australiano Robert W. Taylor (2018), *Understanding the Use of Hyper-Compression in Music Production*, podem ser resumidos à criação de artefatos sonoros devido à distorção dos sistemas de reprodução, redução de profundidade, palco sonoro e texturas, obscurecendo detalhes e retirando a definição e a separação dos instrumentos.

Em sistemas sem normalização de intensidade, a hipercompressão, apesar de danosa ao material sonoro, consegue chamar a atenção do ouvinte, graças à espantosa diferença de volume em relação a materiais mais suaves executados consecutivamente. Porém, nos reprodutores que utilizam tecnologias de normalização de intensidade, como os dispositivos da Apple e serviços de *streaming* como o Spotify⁶, a vantagem da masterização alta desaparece, restando apenas os problemas, bastante evidenciados na comparação com materiais que possuem maior extensão dinâmica, além de transientes e picos menos alterados.

Para masterizar músicas nos novos sistemas de reprodução com normalização de intensidade são necessárias novas ferramentas de medição que observem as intensidades, e não os picos. Paralelamente às tecnologias de normalização de intensidade (como o Sound Check), implementadas nos dispositivos reprodutores, os estúdios veem novas maneiras de mensurar a intensidade do áudio, surgidas em meados da década de 2010, que propõem alterar a maneira como engenheiros de mixagem e masterização buscam o nível adequado de intensidade em seus trabalhos. A edição inglesa de fevereiro de 2014 da revista especializada *Sound on Sound* traz, em sua matéria de capa (ROBJOHN, 2014), uma discussão a respeito da “guerra de volumes” na qual, segundo especialistas, uma novidade parecia alterar os rumos do problema do excesso de compressão na música: a adoção de padrões

6 A tecnologia de normalização de intensidade utilizada pelo Spotify não tem um nome específico, sendo apenas descrito como Audio Normalization nas configurações do aplicativo.

de medida baseados em nível de intensidade (*loudness level*), e não mais em nível de pico (*peak level*).

Assim, no meio da década de 2010, a “guerra de volumes” é declarada como “em vias de acabar” por Bob Katz e outros profissionais, graças, em parte, aos dispositivos portáteis do século XX - seja pela escolha ativa de uma lista de músicas armazenadas, seja pela opção passiva de uma *playlist* customizada por um serviço de *streaming*. A sobreposição de músicas com muita disparidade de intensidade parece ter provocado a adoção da normalização de intensidade nos reprodutores de áudio portáteis (KATZ, 2015, p. 251).

1.6. A masterização como processo artístico

Como vimos, os dispositivos de reprodução com normalização de intensidade diminuem a importância de atingir sempre “o volume mais alto” nas masterizações, o que abre espaço para outras preocupações estéticas na finalização do áudio, como o equilíbrio de frequências e a profundidade e distribuição do palco sonoro virtual. O controle de dinâmicas, que deixa de ter como objetivo atingir altos níveis de intensidade via redução de sua extensão, agora tem função revigorada na manipulação do contraste e do impacto entre os eventos sonoros de uma mixagem.

Essa manipulação estética tende a passar despercebida por leigos. Na fé de uma reprodução de *alta-fidelidade* (*hi-fi*), o ouvinte em muitas situações esquece que, como Fernando Iazzetta (2012, p. 22) lembra, a alta-fidelidade prometida pelos equipamentos *hi-fi* é uma construção tautológica das qualidades sonoras que aprendemos a ouvir. Assim, uma das principais funções da masterização é compatibilizar esteticamente um áudio para uma escuta já acostumada a determinadas sonoridades e gramáticas - especificidades de um instrumento ou de um gênero musical, como a força de transientes em instrumentos de percussão ou a quantidade de graves presentes em determinado gênero (SHELVOCK, 2012, p. 54).

A masterização com recursos digitais permite um controle do sinal de áudio impossível no domínio analógico. Os mesmos algoritmos que criaram a “guerra de volumes” oferecem uma manipulação até então inédita do fonograma, aumentando a precisão das alterações artísticas que o engenheiro de masterização pode realizar na obra que tem em mãos. Além disso, a masterização no domínio digital não precisa ser realizada em tempo real: o profissional tem na tela de seu computador uma representação do fonograma e pode acessá-lo sem a necessidade de navegar fisicamente por um rolo de fita magnética.

Muitos profissionais também têm trocado seus processadores de efeito, como equalizadores e compressores, por ferramentas digitais. Dessa forma, há uma enorme facilidade em voltar para uma sessão de masterização e fazer pequenas alterações: o que antes era uma enorme lista com as configurações de cada equipamento a serem alteradas manualmente, agora é um arquivo de computador que carrega instantaneamente todos os parâmetros utilizados. Essas novidades do domínio digital ampliam a liberdade de escolhas no processo, abrindo oportunidades para a criatividade e experimentação dos profissionais.

Na sociedade pós-industrial, as etapas da produção musical misturam-se, e a mesma pessoa pode ser compositora, produtora, empresária e distribuidora de sua obra - incluindo aí todas as etapas que anteriormente eram realizadas em estúdio, com diferentes profissionais especializados para cada etapa. Mesmo sem necessariamente dominar todos os aspectos técnicos da masterização, ao finalizar um áudio, realizando ajustes em um áudio estéreo já mixado, e compactando-o para um formato de distribuição como o *mp3*, um músico independente está, de fato, masterizando seu trabalho.

Matt Shelvock (2012) demonstra em sua dissertação a importância da masterização na estética final de um fonograma, colocando o engenheiro de masterização junto aos outros artistas envolvidos em um trabalho de estúdio, como coautores da estética daquele resultado. Mesmo que o trabalho desse profissional deva ser “invisível” (KATZ, 2015, p. 77), dele depende parte da adequação estética de uma produção a determinado gênero.

1.7. Aspectos do áudio na masterização

Os autores dividem as funções e ações de um engenheiro de masterização a partir de diferentes critérios. Jay Hodgson, em sua obra *Understanding Records* (2010), entende três áreas de funções: ajustes no áudio (compressão, equilíbrio de frequências), tarefas administrativas (metadados, *backup* de arquivos) e prática de monitoramento (habituar-se com os alto-falantes usados e a acústica da sala, calibrando os ouvidos). Já Bobby Owsinski, em *The Mastering Engineer's Handbook* (2017), separa as responsabilidades em monitoramento e tarefas administrativas. Bob Katz (2015), por sua vez, resume as funções do engenheiro de masterização de forma similar, porém não separa as ferramentas (compressão, equalização), mas seu uso (edição, espacialização, nivelamento). As tarefas administrativas e o cuidado com o monitoramento também são citadas como responsabilidade desse profissional e são tratadas em capítulos separados em seu livro.

Parece-nos natural a atenção dada pelos autores ao monitoramento, já que, concor-

dando com Bob Katz, a masterização é a “[...] última oportunidade de melhorar o som ou corrigir problemas dentro de um espaço acusticamente projetado - em um microscópio do áudio” (KATZ, 2015, p. 1). Por mais que a fruição musical aconteça em ambientes, situações e equipamentos imperfeitos, há consenso entre os profissionais que essa a etapa de manipulação do fonograma deve ser realizada em condições as mais ideais possíveis, praticamente laboratoriais: níveis de ruído ambiente extremamente baixos e equipamentos que não distorçam a dinâmica e o equilíbrio de frequências do material reproduzido. Um sistema de reprodução “transparente”, de alta-fidelidade em relação ao fonograma, associado ao hábito do engenheiro com a resposta do equipamento, permite “julgamentos subjetivos o mais objetivamente possível” (KATZ, 2015, p. 147). Apesar da simbiose entre o engenheiro e seu equipamento serem fundamentais para a tomada de decisões estéticas na masterização, não cabe aqui analisarmos esse equilíbrio, bastando entender o engenheiro e sua prática de monitoramento da mesma forma que um músico e seu instrumento: quanto melhor forem eles e sua relação, melhor será a performance.

As tarefas administrativas, por sua vez, pertencem a um conjunto de ações que não nos interessa para a reflexão sobre a relação da masterização com as diferentes situações de escuta: elas são ações técnicas que não influenciam a estética final do áudio. Algumas tarefas de edição, como *fades* e eliminação de defeitos pontuais, também se encontram fora do escopo desta pesquisa: as relações analisadas são entre os processamentos gerais do áudio e as situações de escuta e, apesar de influenciarem esteticamente o resultado final, essas edições são localizadas e específicas, não aplicadas ao longo de todo o áudio.

Matt Shelvock considera em seu estudo que “o equilíbrio tonal e a intensidade do áudio estão principalmente nas mãos do engenheiro de masterização” (SHELVOCK, 2012, p. 56). Assim, iremos a seguir analisar brevemente os principais aspectos do áudio que são manipulados pela masterização e as ferramentas que alteram a sonoridade do áudio como um todo, como o controle de frequências (equalização) e o controle de dinâmicas (compressão). Algumas ferramentas de medição de nível de intensidade, por serem fundamentais na mudança de paradigma proposto pelos dispositivos de reprodução que ativamente alteram o nível de intensidade dos fonogramas, também estão contempladas. O objetivo, longe de pretender exaurir o assunto, é estabelecer um léxico dos processamentos utilizados em masterização e como eles podem influenciar um fonograma, estabelecendo algumas noções de masterização que serão úteis para a compreensão de outras partes da dissertação.

1.7.1. Medição de intensidade

Ferramentas de auxílio visual podem ser utilizadas durante a criação de um fonograma, ajudando os profissionais a verificar, por exemplo, a distribuição de energia entre as faixas de frequência de um áudio, a relação de fase entre dois canais ou até mesmo a forma da onda. A medição de nível do sinal, presente em grande parte dos equipamentos de áudio, como mesas de som e processadores, é a principal ferramenta de auxílio visual de quem manipula o áudio.

Como comentamos anteriormente, no domínio analógico, as medições de nível de sinal eram usualmente, até o final dos anos 1990, realizadas em equipamentos que, pela própria materialidade dos ponteiros mecânicos, marcavam um nível médio de intensidade, ignorando os picos. São os *average-reading VU meters*. Ainda para leituras no domínio analógico, surgem os *peak meters*, que conseguem, eletronicamente, ler os picos e mostrar o nível de intensidade do sinal com luzes. Por serem muito caros no domínio analógico, eles só foram amplamente utilizados em suas versões digitais (OWSINSKI, 2017). Com a “guerra de volumes” atingindo seu auge na era dos *compact-discs*, o áudio passou a transitar com frequência perto do limite digital.

No domínio digital existe a necessidade de controlar, com muita precisão, o nível máximo de uma onda, pois passar desse ponto (zero dBFS) causa distorções desagradáveis (*digital clipping*). No domínio analógico, esses picos acabavam sendo suavizados pelos equipamentos, tornando-se, inclusive, desejados. A medição de picos, além da leitura de cada amostra de representação do nível do pico da onda (*Peak Level*, ou PL), pode também ser realizada considerando a interpolação dessas amostras (*True Peak Level*, ou TPL), o que leva eventualmente a um pico de onda, na prática, acima do representável digitalmente, acima de zero dBFS. O resultado normalmente é distorção no conversor de áudio digital para o domínio analógico, o que deve ser evitado - por isso a importância dessa medição interpolada (KATZ, 2015, p. 222).

Os sistemas mais atuais, adaptados à realidade da normalização de intensidade que os dispositivos de reprodução de arquivos de música e os sistemas de *streaming* proporcionam, oferecem leituras de sinal baseadas na percepção de intensidade (*perceived loudness*). São algoritmos baseados no padrão proposto pela norma BS 1770-3 da ITU (*International Telecommunications Union*). Essas normas definem como medir *Program Loudness* ou *Integrated Loudness*⁷ (PL, intensidade do programa, que é um cálculo médio da intensidade,

7 Loudness Units below Full Scale (LUFS) é uma escala de intensidade compensada pela escala

medido em LUFS) e *True Peak Level* (TPL - nível real de pico, medido em dBTP). A norma R-128 da AES (*Audio Engineering Society*) e da EBU (*European Broadcasting Union*) define como medir *Loudness Range* (LRA, extensão dinâmica, em dB), que é a distância entre a percepção mais suave e a mais forte de um áudio⁸.

Além dessas medidas padronizadas, utilizaremos em nossas análises e exercícios a ideia, proposta por Bob Katz em seu livro, de *Peak to Loudness Ratio* (PLR) - algo como Relação Pico para Intensidade. É uma medida da diferença, em dB, entre o *True Peak Level*⁹ (igual ou abaixo de zero dBFS, medido em dBTP) e o *Program Loudness*, medido em LUFS. Apesar das unidades serem diferentes, Bob Katz analisa, apoiado em estudos de terceiros, que há razoável correlação entre uma medição psicoacústica de microdinâmicas e esse método de cálculo simplificado (KATZ, 2015, p. 222).

Essas quatro medidas - *Program Loudness* (PL), *True Peak Level* (TPL), *Loudness Range* (LRA) e *Peak to Loudness Ratio* (PLR) - são muito úteis na masterização para dispositivos com normalização de intensidade. De acordo com Bob Katz (2015), Bobby Owsinski (2017), Hugh Robjohns (2014) e Robert Taylor (2018), elas são o futuro da medição de níveis para a masterização, pois já são o presente da fruição de músicas comercializadas em algumas plataformas *on-line*. Elas serão largamente utilizadas no estudo de caso e nos exercícios de masterização para ilustrar as diferenças e mudanças obtidas nos áudios.

1.7.2. Dinâmica do áudio

A manipulação na dinâmica da intensidade de um áudio é ferramenta essencial nas gravações para conter sua extensão dinâmica dentro da faixa de extensão dinâmica disponível do suporte de fixação. Os compressores eram utilizados inicialmente para ajustar o sinal nas transmissões de rádio e estão na cadeia de processamento de áudio desde então,

K, levando em consideração a percepção humana.

8 Outras duas medidas propostas pela AES/EBU R-128 são relativas a tempos curtos, e não a leituras integrais do fonograma. O Short-Term Loudness (abreviado como S) mede a intensidade em LUFS de janelas de tempo de três segundos, enquanto o Momentary Loudness (abreviado como M) mede em LUFS a intensidade em janelas de 400 milissegundos, algo similar aos VU-meters analógicos, porém com a compensação da escala K para considerar as diferenças na percepção humana, em função da frequência e intensidade do sinal.

9 Por ser uma medida interpolada dos valores das amostras do áudio, os valores de True Peak Level podem eventualmente ultrapassar a escala digital, que vai de infinito negativo até zero dBFS. Por isso, é utilizada a escala dBTP para medir TPL, suportando também valores positivos. Para o cálculo de PLR, desconsideram-se valores de TPL acima de zero dBTP, já que valores positivos apenas distorcem o conversor digital para analógico, sem aumentar a sensação de intensidade para o ouvinte.

utilizados para controle ou como ferramenta criativa (KATZ, 2015, p. 82). Na definição de Robert Taylor, a compressão é “[...] um tipo de controle automático de ganho, que analisa o sinal e age de acordo com parâmetros definidos pelo operador” (TAYLOR, 2018, p. 119).

A compressão pode ser vista como um processamento - realizado por equipamentos analógicos ou algoritmos digitais - que altera proporcionalmente (*ratio*) o ganho de um sinal de áudio, quando ele atinge um determinado valor (*threshold*), começando a agir após um determinado tempo de reação (*attack*) e durante um determinado tempo (*release*). Outros parâmetros podem estar disponíveis em função do modelo e da tecnologia que ele usa para realizar esse processamento. Em compressores ópticos, por exemplo, uma lâmpada é utilizada para excitar um sensor de luz, o que gera curvas de respostas similares às de um filamento incandescente acendendo e apagando. Outros projetos levam em consideração a resposta não-linear de válvulas para, eletricamente, alterar o ganho do sinal de áudio. No domínio digital, existem compressores extremamente rápidos, que se antecedem ao comportamento da onda, além de emuladores de compressores analógicos, cujas resposta e sonoridade fazem parte da gramática sonora das referências musicais gravadas no domínio analógico.

Os compressores, quando utilizados em arranjos complexos de processamento com sinais paralelos e diferentes níveis de ganho, podem manipular de outras formas a dinâmica. Além de diminuir os picos, eles podem funcionar como compressores de dinâmica das partes menos intensas (*upward compression*) ou mesmo aumentando a dinâmica do áudio (*upward* e *downward expansion*). O assunto é longo e foge de nosso escopo, porém uma análise profunda do uso criativo de compressores pode ser vista no capítulo 5 do livro de Bob Katz (2015).

No processo de masterização, os compressores geralmente são utilizados com suavidade e parâmetros lentos, apenas controlando o fluxo de dinâmica de um áudio. Dessa forma, eles controlam o que Bob Katz chama de macrodinâmica do fonograma, alterando essencialmente o *Loudness Range* (KATZ, 2015, p. 74). A manipulação da microdinâmica, por sua vez, vem do uso de parâmetros de tempo mais rápidos e valores de *threshold* mais altos, fazendo o compressor agir apenas nos picos de um sinal de áudio, achatando-os. Isso permite uma redução no *Peak to Loudness Ratio* e, conseqüentemente, um aumento do *Program Loudness*. Um compressor utilizado com um *ratio* muito alto, com uso voltado para o controle de picos, é chamado de *limiter*.

A redução excessiva da microdinâmica está ligada a adjetivações não muito positivas por parte de especialistas e acadêmicos. Enquanto Hugh Robjohns (2014, p. 172) define um áudio muito comprimido como “maçante e débil”, Bobby Owsinski (2008, p. 258) o qualifica como “sem vida”. Bob Ludwig, também engenheiro de masterização, está “[...] convencido de que o excesso de compressão destrói a longevidade de uma obra” (OW-SINSKI, 2008, p. 36). Em sua tese, Robert Taylor denomina hipercompressão o excesso de redução da microdinâmica:

A hipercompressão é acusada de remover dinâmica e fazer a música soar “achatada”, criando confusão musical e reduzindo profundidade e textura, tirando dela sua excitação e poder emotivo, produzindo sons não coesos (devido à mudança contínua de frequência de resposta dos compressores multibanda), amplificando a informação do canal central e reduzindo a largura do estéreo e o impacto dos transientes (TAYLOR, 2018, p. 366).

Do ruidoso cilindro de cera aos arquivos de alta resolução que podem ser comprados na internet, as gerações de suportes e dispositivos de reprodução sonora vieram aumentando sua faixa de extensão dinâmica e permitindo ao áudio ali fixado uma extensão dinâmica cada vez mais próxima das capacidades do ouvido humano - um sistema digital 24 bits possui resolução suficiente para representar variações maiores do que podemos perceber¹⁰.

A realidade da produção musical, porém, não aponta para uma utilização plena dessa capacidade: de acordo com Robert Taylor, a música comercial costuma utilizar apenas 20% da faixa de extensão dinâmica disponível em um *compact-disc* (TAYLOR, 2018, p. 116). Bob Katz (2015) analisa o uso de pouca extensão dinâmica, relatando em seu livro o estudo realizado por Rudolf Ortner (2012) em sua dissertação de Mestrado. Katz usa esse levantamento para medir, de 1979 até 2011, as médias de *Momentary Loudness*, *Loudness Range* e *Program Loudness*, concluindo que, nesse período, os níveis de intensidade das faixas, refletidos nos valores de *Momentary* e *Program Loudness* aumentaram significativamente. No período de 1979-81, o valor de M era 12 LUFS, e o de PL, 17,7 LUFS, passando, no período de 2009-11 para, respectivamente, 5,1 e 8,9 LUFS. O *Loudness Range*, por sua vez, permaneceu praticamente igual, apesar da melhora nos suportes: foi de 6,5 dB em 1979-81 para 6,4 dB em 2009-11, variando para 5,7 dB em períodos intermediários (KATZ, 2015, p. 246). Nesse período, a mídia utilizada mudou do vinil para o *compact-disc*, sem

10 Os limites de audibilidade do ser humano são estimados entre zero dB (o som mais baixo audível) e 120 dB (o limiar da dor), para um tom de referência de 1000 Hz. Um formato de áudio digital de 16 bits possui uma faixa de extensão dinâmica de 96 dB, enquanto um formato de 24 bits possui 120 dB (POHLMANN, 2011, p. 3).

nenhuma mudança aparente no *Loudness Range*, ou seja, a mudança do suporte, com maior faixa de extensão dinâmica, não alterou a extensão dinâmica do áudio fixado, como pareceria lógico na adoção de um suporte com maior capacidade. A escuta mediada por alto-falantes parece ter criado uma cultura de escuta musical com pouca extensão dinâmica, independente da capacidade do suporte onde está fixado o áudio.

1.7.3. Frequências

A predominância do controle de dinâmica como assunto em publicações, *posts* de redes sociais e tutoriais em vídeo na internet é compreensível, já que, por mais que se queira fugir da hipercompressão em alguns gêneros musicais, a percepção do “quanto mais forte melhor” é uma realidade sensorial (Taylor, 2018, p. 21). Além disso, outras alterações como equalização e palco sonoro são afetadas diretamente pelo excesso de compressão, diminuindo a atenção dos ouvidos não treinados em masterização - o que não significa que não tenham impacto na percepção do ouvinte, apenas são engolidos pelo achatamento obsessivo do sinal de áudio.

No campo da masterização a equalização é uma ferramenta constantemente utilizada. Bobby Owsinski (2008, p. 23) afirma usá-la apenas menos do que a compressão. Bob Katz (2015, p. 56) considera ser o equilíbrio tonal a maior razão para seus clientes procurá-lo.

Como já comentamos, a mixagem se preocupa com as relações entre os instrumentos e outros eventos sonoros dentro da obra. A masterização tenta aumentar a coerência desse áudio com a gramática sonora dos ouvintes, compatibilizando algumas características com as referências do mesmo gênero musical. No domínio das frequências, o equilíbrio tonal é conseguido através de ajustes nas faixas de frequência desejadas, sempre com muita delicadeza, pois, em algumas circunstâncias, como o áudio já está mixado, modificar a intensidade de uma frequência buscando amenizar um instrumento leva a alterar as características sonoras dos outros elementos que compartilham a mesma faixa do espectro (SHELVOCK, 2012, p. 24). A experiência e sensibilidade do engenheiro de masterização têm um papel muito importante no sucesso do equilíbrio tonal de uma música. Esse equilíbrio, além de variar entre gêneros musicais, também tende a mudar em função da época em que foi realizada a gravação e a masterização.

Como iremos analisar em seguida no estudo de caso realizado sobre masterizações diferentes do mesmo material, não é apenas o *Program Loudness* que varia conforme a tendência estética do momento.

1.8. Estudo de caso: Tales from Topographic Oceans

O extenso estudo de Rudolf Ortner (2012), utilizado por Bob Katz e Robert Taylor em seus textos, dedica-se a analisar a dinâmica de diferentes músicas e como ela se comportou, em média, ao longo do tempo. Verificar como uma mesma gravação foi masterizada¹¹ em relançamentos ao longo dos anos é uma outra forma de observarmos como diferentes escolhas estéticas influenciam o resultado final do fonograma, seguindo as tendências de masterização de cada período.

A análise de um caso específico de masterizações sucessivas parece-nos interessante para verificar algumas das mudanças vistas no comportamento geral do estudo de Rudolf Ortner, mas também como oportunidade de comparação gráfica do equilíbrio tonal das versões, algo que parece razoável apenas nesse tipo de análise - não há como comparar de forma objetiva o equilíbrio tonal de músicas diferentes. Essa análise também permite observar como o discurso dos encartes de algumas das versões se encaixa com um argumento de vendas baseado na ideia de *hi-fi*, com promessas de performance e “transparência” da mídia.

Para este estudo de caso, escolhemos *The Revealing Science of God (Dance of The Dawn)*, a primeira faixa do álbum *Tales From Topographic Oceans*, do grupo de rock progressivo Yes. Lançado originalmente em vinil no final de 1973, ele foi relançado oficialmente na Europa e nos Estados Unidos em *compact-disc* com quatro versões diferentes: uma edição de 1988, duas novas masterizações em 1994 e 2003, e uma remixagem em 2016. As informações relativas às edições usadas neste estudo foram retiradas de seus materiais impressos.

Além dessas versões em suporte físico, a versão disponível na plataforma de *streaming* Spotify é a mesma de 2003, dado a seu tempo de execução, as faixas-bônus do álbum e a distribuidora creditada no *copyright*¹². Existem, ainda, diversas versões alternativas em outros países, como as do Japão, com várias masterizações próprias.

Há quatro versões supostamente diferentes disponíveis na loja virtual da Apple, a iTunes Store. Uma primeira é parte de um pacote com vários outros discos de estúdio do

11 O termo remasterização, comumente utilizado, induz a ideia de um trabalho de masterização refazendo um anterior, como se partisse dele, o que frequentemente não é o que acontece - fitas com a mixagem, e não as masterizações anteriores, são fontes habituais de tais remasterizações. Por isso, optamos por evitar esse termo.

12 O levantamento das versões disponíveis on-line, tanto do Spotify quanto da Apple, foi realizado em maio de 2019.

grupo, igual a uma edição física de 2013. Outra, vendida como “Masterizada para iTunes”, não indica suas origens. Uma terceira, possui uma marcação “*Deluxe Edition*” na capa. As três são, segundo o *site* do grupo (YES, 2019), originadas da masterização de 2003. O mesmo ocorre em outras plataformas, como Deezer, Rdio e Tidal, além de *sites* de venda de arquivos de alta resolução, como HD Tracks e HiRes Audio. A quarta versão disponível na iTunes Store é a remixagem de 2016.

Como as versões disponíveis nessas plataformas eletrônicas são as mesmas masterizações realizadas para suporte físico, a análise de equilíbrio tonal e dinâmica das versões disponíveis nessas plataformas eletrônicas não acrescenta ao estudo - as eventuais alterações ocasionadas no áudio, devido a compressões de arquivos, estão fora do escopo desta análise.

O trabalho realizado por Steve Wilson em 2016 nos proporciona um material bastante interessante: além do trabalho de restauro e remixagem feito a partir das fitas magnéticas multicanal originais das sessões de gravação de 1973, Wilson recuperou uma fita matriz, já mixada, e um disco de vinil da primeira edição inglesa. Essa fita de mixagem e o disco de vinil da primeira prensagem foram digitalizados e disponibilizados em *blu-ray*, em áudio de alta resolução.

A fita de mixagem é uma cópia da mixagem final, enviada aos estúdios de masterização para a transferência e criação das matrizes de replicação de vinil. Existem inúmeras versões em vinil por conta disso (OWSINSKI, 2008, p. 64), e até hoje grupos de discussão na internet realizam comparações das diferentes versões de seus discos preferidos produzidas em cada país.

Além do acesso às várias versões em *compact-disc*, a disponibilização dessa versão digitalizada do disco de vinil e da fita de mixagem nos permitem observar com rara proximidade as diferenças de sonoridade em cada etapa. Uma comparação direta de *Program Loudness* entre essas digitalizações e as versões em *compact-disc* não parece fazer muito sentido, já que a fita de mixagem nunca foi pensada para ser ouvida pelo público sem uma masterização, e a digitalização do disco de vinil passa por decisões de nível de ganho em amplificações analógicas, que poderiam atrapalhar a comparação. A análise de outros elementos do áudio, como equilíbrio tonal, *Loudness Range* e *Peak to Loudness Ratio*, porém, nos parece muito indicativa das mudanças que passou esse áudio em diferentes suportes e versões ao longo dos anos.

O anexo desta dissertação permite a escuta do mesmo trecho de aproximadamente 45 segundos de cada uma das versões da música. As primeiras seis faixas (faixas 01 a 06 do anexo) apresentam os áudios como foram extraídos de seus suportes originais, enquanto as seis faixas seguintes (faixas 07 a 12 do anexo) contêm os mesmos áudios normalizados por intensidade, para facilitar a comparação do equilíbrio tonal das versões. Duas pastas podem ser encontradas no anexo: uma possui as faixas na resolução de *compact-disc* utilizada para as análises. A outra possui as faixas compactadas para o formato aac, da Apple, fornecidas para facilitar a eventual escuta em dispositivos portáteis¹³.

1.8.1. Quatro versões e duas curiosidades

Utilizando uma ordem cronológica de criação (e não de lançamento), o primeiro áudio que entra em nossa análise é a *mixtape* - fita de mixagem (faixas 01 e 07 do anexo), retirada do disco *blu-ray* da edição especial de 2016. O encarte nos informa que

durante a busca das fitas para a remixagem, encontrou-se uma fita matriz estéreo em alta resolução, apresentada aqui como a “fita de mixagem original em estéreo”. Esta é a primeira vez que essa matriz está sendo disponibilizada na era digital. Ela é fornecida sem nenhuma masterização adicional e é tocável na resolução de até 24 bits/192 KHz¹⁴, dependendo da capacidade do equipamento de reprodução (YES, 2016).

Assim, consideramos que o áudio disponibilizado no disco *blu-ray* é a versão digital mais fiel possível da mixagem original, pré-masterização. Ela é um retrato do resultado da mixagem, e nos oferece um olhar sobre o material disponível para os engenheiros de masterização das versões em *compact-disc* de 1988, 1994 e 2003. O encarte não credita a ninguém a digitalização dessa fita magnética.

O segundo áudio analisado é a digitalização da versão inglesa do disco de vinil (faixas 02 e 08 do anexo), extraída do mesmo disco *blu-ray* da fita de mixagem. Além desse vinil há outro: uma versão promocional para as rádios norte-americanas. Como as músicas foram editadas nessa outra versão - cortes e ênfases diferentes, oferecendo edições mais curtas de parte das músicas para tocar nas rádios - esse áudio não foi considerado em nosso estudo. Porém, o encarte apresenta as duas digitalizações dos discos de vinil em conjunto, “oferecidos como curiosidade de áudio dentro de um contexto, e não como um tipo de ‘experiência audiófila’” (YES, 2016). Além disso, o encarte explicita que “a transferência da

13 A versão compactada também está disponível em <https://bit.ly/2YE7xM9>

14 Bits e KHz são, respectivamente, as unidades de resolução e taxa de amostragem de um áudio digital, a forma de descrever a representação da onda sonora.

versão inglesa é pensada para dar uma ideia de como um disco original, limpo e em muito boa condição, soaria em um equipamento comum” (YES, 2016).

Curiosamente, a digitalização do vinil é chamada de “transferência”, escapando do termo associado com computadores e oposto ao som analógico, mostrando todo um discurso *hi-fi* saudosista que percorre todo o encarte, em oposição ao encarte da edição lançada 28 anos antes, que veremos a seguir. A ficha técnica do encarte também tenta escapar da afirmação de que o áudio do *blu-ray* é “digital”: ao invés de creditar o engenheiro de masterização ou de digitalização dos discos de vinil, o encarte avisa: “transferências de vinil preparadas por John Kimber” (YES, 2016). O áudio do disco de vinil tem resolução inferior ao da fita de mixagem, 24 bits/48 KHz (e não 24/96 como o encarte credita, em um erro de informação), mas ainda assim maior do que a resolução de um *compact-disc* (16 bits/44.1 KHz).

A versão existente há mais tempo, das analisadas neste estudo, é a primeira versão europeia do álbum em *compact-disc* (YES, 1988), lançada em 1988 (faixas 03 e 09 do anexo). Apesar de existirem desde 1982, em 1988 as vendas de *compact-discs* já haviam ultrapassado as de discos de vinil, mas ainda estavam bem atrás das vendas de fitas cassete (BYRNE, 2012, p. 209). O encarte do disco reflete o esforço em vender o *compact-disc* como uma nova geração do *hi-fi*, ocupando uma considerável área da contracapa do encarte, que reproduzimos aqui pela curiosidade histórica e o discurso *hi-fi*, ainda pretendendo ensinar o usuário a utilizar a novidade tecnológica:

O Sistema de Áudio Digital em *Compact-Discs* oferece a melhor reprodução de som possível - em uma unidade de suporte de som pequena e conveniente. O desempenho notável do *Compact-Disc* é o resultado de uma combinação única de reprodução digital com óptica de laser. Para obter os melhores resultados, você deve aplicar o mesmo cuidado das gravações convencionais [discos de vinil] ao armazenar e manusear o *Compact-Disc*. Nenhuma limpeza adicional será necessária se o *Compact-Disc* for sempre segurado pelas bordas e colocado em sua caixa imediatamente após o uso. Se o *Compact-Disc* ficar sujo com impressões digitais, poeira ou sujeira, ele poderá ser limpo (sempre em linha reta, do centro para a borda) com um pano macio e seco, limpo e sem fiapos. Nenhum solvente ou abrasivo deve ser usado no disco. Se você seguir estas sugestões, o *Compact-Disc* irá proporcionar uma vida inteira de puro prazer de escuta (YES, 1988).

A capa traseira da caixa também tenta vender a “superioridade *hi-fi*” da nova mídia:

A música neste *Compact-Disc* foi originalmente gravada em equipamento analógico. Nós tentamos preservar, tanto quanto possível, o som da gravação original. Devido à sua alta resolução, no entanto, o *Compact-Disc* pode revelar as limitações da sua fita de origem (YES, 1988).

Apesar de informar o usuário sobre os detalhes de uso da mídia e louvar suas qualidades - a ponto de culpar previamente a fita original se o ouvinte encontrasse algum defeito - o encarte se limita a fornecer essas informações sobre sua edição, repetindo o conteúdo do encarte original do vinil, sem creditar ninguém pela versão digital. O conteúdo gráfico do encarte nem é adaptado para o tamanho da nova mídia, sendo apenas uma cópia em miniatura, quase ilegível, do original. Dessa forma, não existem créditos pela digitalização, não se menciona nem mesmo se ela foi feita a partir de uma fita magnética ou de um disco de vinil. A escuta do áudio, porém, sugere ser a fonte uma fita magnética, devido à presença de conteúdo de alta frequência, que é inexistente nos discos de vinil pela incapacidade física do material em representar tais agudos, e à falta dos cliques e chiados irregulares da escuta em vinil - se hoje é possível uma edição minuciosa dessas imperfeições, em 1988, essa tecnologia deixava rastros perceptíveis (KATZ, 2015, p. 112), o que não se observa na escuta do material.

Essa primeira geração de lançamentos em *compact-discs*, de álbuns originalmente lançados em discos de vinil, foi vendida apenas como transferências, sem maiores detalhes, focando apenas no valor *hi-fi* da nova mídia. Passada essa primeira onda de troca de mídia, alguns álbuns foram sendo relançados como “remasterizados” - talvez a primeira leva de discos com essa denominação.

Em 1994, o *compact-disc* já havia ultrapassado as vendas das fitas cassete em mais que o dobro de unidades nos Estados Unidos (BYRNE, 2012, p. 209), e a gravadora Atlantic relança seu catálogo do grupo Yes com o apelo comercial da “remasterização” - expondo esse fato em etiquetas autoadesivas na embalagem, e em grandes créditos na parte de trás da caixa do disco, ocupando duas linhas e todo o comprimento disponível: “Remasterização digital a partir das fitas matriz originais por Joe Gastwirt no estúdio Ocean View Digital” (YES, 1994). Essa edição (faixas 04 e 10 do anexo) não conta com conteúdo específico sobre a mídia, como a versão de 1988. Além dos textos originais do disco de vinil, o encarte possui apenas essas novas informações e a devida adaptação do encarte para o formato do *compact-disc*.

A terceira versão do disco a ser lançada em *compact-disc* (faixas 05 e 11 do anexo) data de 2003 (YES, 2003). Com o conteúdo do encarte ampliado com textos e fotos inéditos, a masterização de Bill Inglot e Dan Hersch é apenas mais um detalhe no lançamento, levando o *Program Loudness* aos padrões do período sem maior alarde. Maior ênfase é

dada ao fato de ter sido utilizada uma fita matriz contendo uma introdução até então não mostrada ao público: por sua longa extensão, a faixa incluída nesse estudo (e também as três outras faixas desse álbum) ocupava um lado inteiro do disco de vinil, com aproximadamente 20 minutos. Por não caber no vinil, essa introdução de mais de dois minutos foi cortada do disco original, ressurgindo apenas nessa edição de 2003. Essa introdução foi retirada e ignorada no estudo dos áudios, mas ela serve de assinatura da versão, comprovando seu uso nas versões disponíveis *on-line*.

Como já comentado, Robert Taylor (2018, p. 98) indica 2004 como o auge da “guerra de volumes”. Assim, essa masterização de 2003 foi realizada perto do momento de maior hipercompressão da história da indústria fonográfica. Mesmo sendo o rock progressivo um gênero musical sem as demandas de sonoridade de gêneros mais populares - e partir de uma produção realizada décadas antes, trazendo seu público para uma sonoridade dos anos 1970 - essa edição não consegue escapar das tendências de seu tempo, como veremos na análise de dinâmica das edições.

A última versão utilizada nesse estudo não pode ser comparada diretamente: não se trata de uma masterização, mas de uma nova mixagem (faixas 06 e 12 do anexo). O músico Steven Wilson é o responsável pela edição de 2016 - a mesma que apresenta as digitalizações dos dois primeiros áudios deste estudo, entre outros “garimpos arqueológicos” como versões demo e instrumentais. A partir das fitas multicanal das sessões de gravação, ele mixou novamente as músicas, tendo muito maior liberdade para restaurar o áudio e dar um tratamento contemporâneo ao material captado. Steven Wilson cria sua versão do álbum que, apesar de respeitar o trabalho original, não pode ser diretamente comparada às outras masterizações. Os valores resultantes do estudo de níveis de intensidade e equilíbrio tonal, porém, ilustram as preocupações em recuperar a dinâmica perdida desde o início da era digital na música comercial.

Para uma análise mais precisa dos áudios, as entradas e saídas da música (*fades in* e *out*), realizadas com alguma diferença em cada masterização, foram cortadas dos áudios utilizados, usando eventos sonoros facilmente localizáveis na forma de onda, como referência de tempo. Essas versões “sem as pontas” foram as utilizadas para os cálculos estatísticos de dinâmica e a criação dos gráficos de distribuição espectral. Esse recorte também permite comparar a velocidade de reprodução das fitas magnéticas no momento da digitalização - no caso do vinil, a velocidade final entre a transferência para a matriz e a

digitalização do disco quase quarenta anos depois. O disco de vinil tem aproximadamente 12 segundos a mais do que a versão de 1994, uma variação de menos de um por cento, visto que a faixa tem pouco mais de vinte minutos de duração total¹⁵. Para viabilizar as comparações de equilíbrio tonal, as versões que possuíam taxa de amostragem e resolução mais altas que um *compact-disc* - caso das versões *mixtape* e disco de vinil - foram reduzidas por *software* para 16 bits / 44,1 KHz.

1.8.2. Equilíbrio tonal

O equilíbrio tonal não pode ser medido tão objetivamente quanto o nível de intensidade de uma faixa, sendo o ouvido a melhor ferramenta para seu julgamento (KATZ, 2015, p. 56). É possível, porém, verificar as diferenças de equilíbrios tonais entre os elementos deste estudo, pois eles partem de um mesmo áudio trabalhado diversas vezes.

Mesmo para leigos, a melhor ferramenta de comparação entre essas versões ainda é o ouvido. Apresentamos aqui um auxílio a essa comparação: um gráfico da análise espectral de uma faixa é sobreposto ao outro, permitindo visualizar os diferentes comportamentos do equilíbrio tonal entre algumas versões. Importante ressaltar que mudanças sutis nessa curva são facilmente audíveis. Como comentado, amostras dos áudios, com normalização de intensidade para facilitar a comparação, estão disponíveis para a escuta.

Todos os gráficos foram realizados usando as ferramentas de análise do *software* especializado RX7 da fabricante iZotope. Para a obtenção dos gráficos de distribuição espectral foram utilizadas janelas do tipo Blackman com tamanho de 8.192 amostras. A adequação para impressão e a sobreposição dos gráficos foi realizada posteriormente em *software* de edição de imagens.

15 Já sem as entradas e saídas dos áudios, a fita de mixagem tem aproximadamente 20 min e 09 s, a versão retirada do disco de vinil tem 20 min e 19 s, as versões de 1988 e 2003 tem 20 min e 08 s e a de 1994 tem 20 min e 07 s. A remixagem tem 20 min, mas ela tem partes levemente reposicionadas durante a peça, então a comparação direta não é adequada.

Fig. 1 - espectrograma comparativo entre as versões mixtape e disco de vinil

Fonte: produção do autor

Os espectros da fita de mixagem (faixa 07 do anexo) e do vinil (faixa 08 do anexo) parecem bastante similares (figura 1). Há um pouco mais de graves na curva de espectro do vinil, o que pode significar menos agudos - o que importa é o equilíbrio, principalmente nesse exemplo, que inclui a digitalização posterior dos dois suportes analógicos. Porém podemos observar o quanto essa curva, com agudos fortemente decrescendo, segue um padrão oposto às diferenças que surgirão nas versões em *compact-disc*.

Fig. 2 - espectrograma comparativo entre as versões disco de vinil e 1988

Fonte: produção do autor

A comparação entre a versão extraída de disco de vinil (faixa 08 do anexo), de 1973 (digitalizada em 2016) e a do primeiro lançamento em *compact-disc*, de 1988 (faixa 09 do anexo), revelam mais semelhanças do que diferenças: o equilíbrio tonal é muito similar (figura 2). Na versão em vinil, pode-se ouvir alguma distorção nas partes de maior intensidade - possivelmente resultado de compressão analógica quando da masterização para a matriz do disco -, enquanto na versão em *compact-disc* ouve-se claramente a falta de chiado nas partes menos intensas, possivelmente causada pela redução de ruído, muito utilizada na época das primeiras digitalizações para *compact-disc* (KATZ, 2015, p. 112). As curvas sobrepõem-se bastante na faixa entre 5 KHz e 10 KHz e a versão do vinil tem maior intensidade de sinal. Após 10 KHz, porém, o sinal da curva do vinil decai em uma curva bastante acentuada. Fora a eliminação excessiva de ruídos, levando parte dos agudos da fita magnética, os responsáveis por essa versão parecem ter, de fato, se atentado às suas próprias observações na embalagem: “tentando preservar, o quanto possível, o som da gravação original” (YES, 1988).

Fig. 3 - espectrograma comparativo entre as versões 1988 e 1994

Fonte: produção do autor

Na escuta comparativa entre os *compact-discs* de 1988 (faixa 09 do anexo) e de 1994 (faixa 10 do anexo), há diferenças sutis entre os médios de cada versão (notável entre 0,5 KHz e 2,5 KHz), mas o mais evidente é a volta do ruído original das fitas magnéticas, e da maior presença de agudos, visível na tendência de separação das duas curvas a partir de aproximadamente 6 KHz (figura 3). O acúmulo de energia após 20 KHz, da versão de

1994, parece ser resultado de um trecho da música onde há um solo de sintetizador, com muita sobra de altas frequências; em uma análise de espectrograma esse trecho é claramente visível. As versões em vinil e a de 1988 têm um filtro bastante atuante nessa faixa de frequência acima de 20 KHz; a de 1994 é a que menos se preocupa com isso.

Fig. 4 - espectrograma comparativo entre as versões 1994 e 2003

Fonte: produção do autor

Os agudos também se fazem presentes na comparação entre as versões de 1994 (faixa 10 do anexo) e 2003 (faixa 11 do anexo): A versão mais recente tem bastante mais presença, como se tivesse sido retirado um obstáculo acústico da anterior. A partir de 1 KHz é possível ver, no gráfico (figura 4), as curvas separando-se até aproximadamente 15 KHz, quando elas se cruzam novamente com a versão de 2003, sendo claramente filtrada nos superagudos - o que não é muito sensível, já que muito pouco se ouve acima de 15 KHz.

O equilíbrio tonal foi sendo modificado, ampliando os agudos a cada versão. Apesar de algum reforço nos graves até 0,25 Hz, em relação à versão de 1994, ela tem menos energia do que a anterior na faixa entre 0,25 Hz e 0,5 Hz, aproximadamente. Isso reflete-se diretamente na sensação de equilíbrio entre graves e agudos.

Fig. 5 - espectrograma comparativo entre as versões 2003 e 2016

Fonte: produção do autor

A versão de 2016 (faixa 12 do anexo), como já comentamos, é uma nova mixagem. Assim, algumas comparações não se aplicam, como o tempo da música, que foi remontada completamente, sendo aproximadamente sete segundos mais curta do que as outras versões em *compact-disc*. A escuta atenta dessa versão deixa evidente um salto qualitativo no áudio em relação à versão de 2003 (faixa 11 do anexo), porém a análise gráfica (figura 5), demonstrando o quanto ela pode iludir, mostra o espectro dos áudios mais próximos do que nas outras comparações. Observando melhor, porém, vê-se que, de aproximadamente 0,11 KHz até 8 KHz - ou seja, a maior parte do espectro - a versão remixada de 2016 possui menor energia, chegando a ter até 5 dB a menos em sua parte central - de leitura dificultada pela grande variação interna aos espectros, que se sobrepõem. Em relação à versão de 2003, ela tem maiores baixas e altas frequências, enquanto as médias todas são menores - um equilíbrio bastante diferente da versão anterior.

1.8.3. Dinâmica

A comparação entre a dinâmica das versões analisadas é mais objetiva do que a do equilíbrio tonal, já que há métodos padronizados para o cálculo dos diferentes níveis de intensidade. Os valores foram obtidos utilizando o mesmo *software* de análise dos gráficos de espectro, o RX7 da fabricante iZotope, com exceção da última linha: *Peak to Loudness Ratio*, obtida a partir da subtração do maior valor (negativo ou zero, ignorando positivos) de *True Peak Level* com o valor de *Loudness Range*, seguindo a proposta de Bob Katz para uma medição simples entre o intensidade da faixa e seu maior pico, oferecendo uma

medida de compressão dos picos (KATZ, 2015, p. 223). As formas de onda e os valores calculados nessa parte do estudo são dos áudios cortados para conterem o mesmo material sonoro, porém sem alteração de intensidade.

Tabela 1 - Estatísticas das diferentes versões da faixa.

versão	True Peak Level L (dBTP)	True Peak Level R (dBTP)	Sample Peak Level L (dBTP)	Sample Peak Level R (dBTP)	Program Loudness (LUFS)	Loudness Range (LU)	Peak to Loudness Ratio
mixtape	-1,01	-1,04	-1,02	-1,02	-15,00	11,50	13,99
vinil	-4,74	-4,58	-4,75	-4,58	-20,50	9,90	15,92
1988	-0,03	-0,02	-0,07	-0,07	-14,80	9,80	14,78
1994	-0,18	-0,17	-0,50	-0,50	-13,30	11,60	13,13
2003	-0,27	+0,27	-0,48	-0,48	-12,80	10,80	12,80
2016	-2,54	-2,44	-2,58	-2,55	-17,60	13,00	15,16

Fonte: produção do autor.

Além da análise numérica, a forma de onda dos áudios permite a visualização de algumas características, como diferenças no corpo entre trechos e achatamento de picos. Importante remarcar que a forma de onda não tem relação direta com o *Program Loudness*, já que transientes podem “desenhar” um perfil maior em um *software* de visualização, sem carregar maior energia total do que um áudio mais comprimido (KATZ, 2015, p. 79).

A seguir veremos as formas de onda de cada áudio e estabeleceremos comparações entre elas, baseadas nos números da Tabela1, acima, e ilustradas pelas diferenças gráficas das formas de onda, ilustradas a seguir nas Figuras 6 a 11.

Fig. 6 - forma de onda da versão mixtape

Fonte: produção do autor

Fig. 7 - forma de onda da versão disco de vinil

Fonte: produção do autor

Fig. 8 - forma de onda da versão de 1988

Fonte: produção do autor

Fig. 9 - forma de onda da versão de 1994

Fonte: produção do autor

Fig. 10 - forma de onda da versão de 2003

Fonte: produção do autor

Fig. 11 - forma de onda da versão de 2016

Fonte: produção do autor

A primeira medida da Tabela 1 é o valor de *True Peak Level*. A norma EBU R-128, cuja primeira versão data de 2010, recomenda não exceder -1 dBTP - assim como vários profissionais da área de masterização, como Bobby Owsinski que, já em 2008, alertava para as possíveis distorções de conversores, compactadores e processamentos após a distribuição do áudio (OWSINSKI, 2008, p. 92). Bob Katz, por outro lado, observa a relação entre *True* e *Sample Peaks*: “em geral, quanto mais um áudio for processado, distorcido ou possuir

altas frequências, maior será o *True Peak* em comparação com o *Sample Peak*” (KATZ, 2015, p. 222).

Todas as versões masterizadas para *compact-disc* utilizadas neste estudo são anteriores à norma EBU R-128. A versão de 1988 quase chega em zero dBFS, mas como não havia uma cultura de excesso de compressão, quase não há diferença entre os valores de *True Peak* e o *Sample Peak*. Pelo contrário: como vimos anteriormente, essa edição preza pela tentativa de alterar o menos possível o áudio original - com exceção do ruído, que talvez não fosse considerado parte do áudio original naquela época, apesar de ser originado pelos equipamentos de estúdio e pela fita magnética.

Nas versões de 1994 e 2003, o valor de *Sample Peak* distancia-se um pouco, indo para aproximadamente conservadores -0,50 dBFS. Porém, nas duas versões, percebe-se o uso de *limiters*, achatando os picos das ondas mais altas e causando distanciamento entre *True Peak* e *Sample Peak*. Enquanto a versão de 1994 mantém o valor máximo de *True Peak* ainda dentro da faixa negativa, a versão de 2003 passa de zero, chegando em + 0,27 dBTP no canal direito, consequentemente ocasionando distorções em processamentos posteriores e na conversão para sinal analógico durante sua reprodução. É possível perceber graficamente nas formas de onda dessas versões momentos com tantos picos achatados que se observa claramente o *limiter* atuando - no trecho próximo ao 17º minuto, isso fica bastante evidente nas imagens.

A remixagem de 2016 diminui bastante o uso de *limiters*, praticamente eliminando novamente a distância entre *True* e *Sample Peaks*. É preciso lembrar, porém, que essa edição é claramente audiófila, possuindo versões em alta resolução e *blu-ray*, de onde foram retiradas as versões da fita de mixagem e do disco de vinil.

O gênero musical da peça estudada não colabora muito para a análise do *Program Loudness*. De fato, um fonograma que contenha várias partes, com *pianissimos* e *fortissimos*, terá uma média inadequada para ser utilizada em regulagem de volume: as partes ficarão, respectivamente, abaixo e acima da meta de ajuste, sendo sua intensidade sonora prejudicada em ambas as passagens - se o material só tem *pianissimos* e *fortissimos*, não existe material com a média calculada. Assim, a média de LUFS entre duas partes de um fonograma com intensidades muito diferentes não representa nenhuma delas, gerando um valor equivocado de *Program Loudness* (KATZ, 2015, p. 220).

É o caso dessa peça. Para a comparação entre as diferentes masterizações da mesma

peça, porém, a análise de *Program Loudness* continua válida: os trechos são os mesmos, apenas com nível de intensidade maior ou menor conforme as versões.

Como já dissemos, a nova mixagem de 2016 possui um diferente equilíbrio dos instrumentos, e as versões da fita de mixagem e do disco de vinil passaram por digitalizações posteriores, atrapalhando a comparação direta. Assim, a comparação é melhor realizada entre as três masterizações em *compact-disc*. A análise dos três valores - cronologicamente - 14,80, - 13,30 e - 12,80 LUFS - ilustra o que Bob Katz (2015) conclui a partir do levantamento de Rudolf Ortner: os valores médios de *Program Loudness* dos anos de 1991-1993 a 2003-2005 passam de - 13,60 para - 9,10 LUFS, ou seja, um salto de 4,50 LUFS em quinze anos, enquanto nosso estudo vê um aumento de 2 LUFS no mesmo lapso de tempo, porém de 1988 a 2003. Um estudo sistemático dos valores de *Program Loudness* em masterizações sucessivas de um mesmo material pode explicar essa diferença, que pode estar relacionada a especificidades desse gênero musical, mas talvez se explique por uma menor adesão das (re)masterizações às tendências sazonais do que as novidades musicais do mesmo período.

Uma constatação interessante da comparação de *Loudness Range* entre as masterizações estudadas é que as versões de vinil e de 1988 possuem quase o mesmo valor de LR, os menores encontrados em nossa análise. Isso significa que, apesar do aumento da compressão de picos das versões de 1994 e 2003, elas conseguiram manter mais *Loudness Range* do que as duas versões anteriores, a de vinil e a de 1988. A versão de 1994, inclusive, tem um LR maior do que a própria versão de fita de mixagem - o que pode ser conseguido com técnicas de compressão e expansão da extensão dinâmica do áudio. Pode ser também que a fita matriz obtida em 1994 seja de melhor qualidade do que a digitalizada em 2016, o que parece improvável dado o acesso ao material original da edição de 2016. Se Steve Wilson conseguiu até as sessões originais de gravação, parece improvável que não tenha tido acesso à melhor fita de mixagem possível.

O *Loudness Range* da versão remixada em 2016 é maior, mas isso em nada surpreende, por ser um trabalho de reconstrução a partir das sessões de gravação, com minúcias audiófilas no restauro do áudio e na finalização do material.

1.8.4. Conclusões do estudo

Uma conclusão deste estudo é encontrada na diferença dos valores de *Program Loudness* das versões da música e o valor médio de PL apresentado por Rudolf Ortner (2012).

Como explanado durante a análise, o *Program Loudness* de uma música com partes muito distintas induz a uma ideia bastante errada de seus níveis de intensidade, por se tratar de uma forma de média ponderada. Isso é um problema para algumas peças de rock progressivo e também em outros gêneros, como músicas de orquestra, que exploram a extensão de dinâmica típica de um conjunto grande de instrumentistas. Esse problema parece ainda não ter sido resolvido pelos sistemas de normalização de intensidade como o Sound Check.

Cabe ao engenheiro de masterização o equilíbrio entre a conservação de um áudio e sua adequação ao tipo de suporte de fixação utilizado, o que nos remete ao problema observado quando do levantamento das versões disponíveis. Datada de 2003, praticamente no auge da “guerra de volumes”, a versão utilizada para os sistemas de *streaming* e *download* é também a mais comprimida e com pior relação *Peak to Loudness Ratio*. Além disso, o *True Peak Level* acima de zero dBTP, como mencionado, causa distorções quando se faz compactação para formatos *lossy* como *aac* e *mp3* - exatamente o que acontece nessas plataformas. Em outras palavras, o consumidor desses serviços é prejudicado por problemas similares aos de 1988, de mudança de mídia: a primeira versão em *compact-disc*, apesar do discurso *hi-fi*, oferecia um álbum com uma redução artificial de ruídos. A versão disponível *on-line*, mesmo não sendo comercializada em mídia física, oferta aos consumidores o resultado da “guerra de volumes” ocorrida com os *compact-discs*.

Mikko Ojanen, analisando diferentes digitalizações e masterizações de uma peça do compositor eletroacústico Erkki Kurenniemi, conclui que as mudanças estéticas realizadas nas masterizações são similares a “diferentes interpretações de uma peça em tempos diferentes” (OJANEN, 2015, p. 14), e que uma masterização “pode ser vista como a interpretação musical de um engenheiro de masterização, refletindo a qualidade da tecnologia usada, assim como os ideais estéticos do momento” (OJANEN, 2015, p. 1). Parece-nos importante a atualização das masterizações, não apenas para resolver os problemas técnicos das adequações a novos tipos de suporte - como parece acontecer com o *streaming* de masterizações muito próximas ou acima de zero dBTP - mas também para que a estética da reprodução reflita os dias em que ela acontece. Particularmente a análise das diferentes versões aqui estudadas permitiu observar como variaram as masterizações em cada relançamento, alterando aspectos do equilíbrio tonal e da dinâmica do fonograma original. Uma sugestão para investigações futuras seria avaliar se essas mudanças nas masterizações respondem a tendências estéticas sazonais.

Por fim, ainda no tema do discurso *hi-fi*, é interessante notar que as mudanças de mídia - tanto a primeira versão em *compact-disc*, de 1988, quanto a disponibilização, em *blu-ray*, da fita de mixagem e do disco de vinil - oferecem em seus textos uma narrativa alinhada com a ideia da alta fidelidade, na qual transferências de um tipo de suporte para outro são realizadas apenas tecnicamente, sem decisões estéticas por parte dos executores. Vimos que isso não é exatamente verdade, por exemplo na edição de 1988, na qual há explícita redução de ruído, seguindo padrões da época (KATZ, 2015, p. 112).

A edição de 1988 fala de alta tecnologia e transferência de alta-fidelidade, sem creditar ninguém na produção da versão digital, e a ficha técnica do *blu-ray*, de 2016, chega a omitir a função dos engenheiros: os créditos apontam “*Vinyl transfers prepared by John Kimber*” (YES, 2016), enquanto a versão remixada é apenas “transferida sem alterações” para *compact-disc*: “*2016 CD master prepared by Neil Wilkes (flat master transfer)*” (YES, 2016). Os dois estão, por definição, masterizando esses áudios, mas são creditados praticamente como *cutters* - os responsáveis pela transferência falsamente transparente para vinil, antes do surgimento da função de engenheiro de masterização.

O mesmo encarte omite os créditos da digitalização das fitas utilizadas na remixagem, assim como a digitalização da fita de mixagem original, não deixando indícios de quem manipulou os rolos de fita magnética. O *hi-fi*, com seu discurso cientificista e tecnológico, acaba por ignorar as pessoas que manipularam os fonogramas originais, sem as quais não existiria essa edição especial em *blu-ray*.

2. Dois caminhos da escuta

Apesar do termo *hi-fi* ter surgido nas décadas de 1930 e de 1940 e de ter sido disseminado apenas após a Segunda Guerra Mundial, a ideia da alta-fidelidade é utilizada como técnica de venda de equipamentos de reprodução sonora desde ao menos os fonógrafos da década de 1910 (IAZZETTA, 2009, p. 113). De fato, os *Tone Tests* de Thomas Edison, eventos de comparação entre a performance de seus fonógrafos e cantores ao vivo, não apenas serviam de demonstração da capacidade de seu equipamento, mas funcionavam como formação de uma cultura de escuta mediada, estabelecendo equivalências entre a performance musical e a reprodução sonora (STERNE, 2003, p. 261).

Alex Cummings sugere uma origem ainda mais antiga do *hi-fi* como argumento de vendas: a *Victor Talking Machine Company*, fabricante de discos e fonógrafos, adquire em 1900 os direitos de uma pintura que mostra um cachorro ouvindo atentamente um fonógrafo de Edison, um gravador e reproduzidor que usava cilindros de cera como suporte. O logo da companhia foi adaptado desse quadro, sendo o cilindro de cera substituído por um fonógrafo de disco. O título da obra, *His Master's Voice* - a voz de seu dono - foi adotado como slogan, reforçando a ideia de um aparelho capaz de iludir até um cão com sua qualidade de reprodução (CUMMINGS, 2013, p. 16). Para Jonathan Sterne, “a própria ideia de que um som reproduzido poderia ser fiel a um som original é um subproduto da cultura e da história da reprodução sonora” (STERNE, 2003, p. 283).

A alta-fidelidade não é simples de ser aferida, como argumenta Michel Chion:

Alguém que ouve uma orquestra num sistema de som em sua sala de estar dificilmente pode compará-la com uma orquestra tocando em sua porta. Deve-se saber, de fato, que a noção de alta-fidelidade é puramente comercial, e não corresponde a nada preciso ou verificável (CHION, 1994, p. 98).

O termo *hi-fi* traz, portanto, essa ideia de fidelidade de reprodução em relação a uma fonte sonora original - os próprios *Tone Tests* de Thomas Edison exigiam algum esforço para funcionar, tanto do cantor ao vivo, aproximando sua performance ao desempenho do aparelho, quanto dos espectadores, tentando adaptar sua escuta ao equipamento (STERNE, 2003, p. 262). Parte de um culto à tecnologia, o *hi-fi* funciona como estratégia de mercado da indústria fonográfica para propagandear seus produtos, muito mais do que um método de classificação objetiva de características de equipamentos. Dos cilindros de cera aos *home theaters* da virada do século passado, cada geração de dispositivos domésticos de reprodução sonora usa o mesmo argumento de vendas, apelando para o discurso tecnoló-

gico da “alta-fidelidade”.

2.1. O discurso da alta-fidelidade: fiel em relação a quê?

Mesmo quando o termo *hi-fi* passou a ser amplamente utilizado, a partir da década de 1950, a escuta presencial da música já era rara, com a maioria das pessoas tendo um acesso muito maior ao rádio e aos discos do que à música ao vivo. A referência dessa alta-fidelidade, portanto, já havia se perdido, e o *hi-fi* se voltava a um discurso de limpeza sonora (IAZZETTA, 2012, p. 22). Assim, a referência da alta-fidelidade deixa de ser a performance ao vivo e passa a ser a própria escuta mediada.

Em seu livro *The Audible Past*, Jonathan Sterne (2003, p. 226) discorre sobre a relação entre um cantor de rádio e seu ouvinte, conectados através desse meio. A performance não é para o ouvinte, é para o próprio meio, que é a natureza da conexão entre as pontas. Com gravações, a relação é a mesma: quem consome música mediada não ouve um conjunto de músicos, mas sim a performance de um fonograma - o que inclui esses músicos, mas também microfones e sistemas de gravação, fixação e reprodução, com escolhas estéticas e sonoridades dos equipamentos utilizados.

Do tipo de microfone escolhido ao suporte de fixação utilizado, o percurso do áudio na produção de um fonograma afeta suas características sonoras. Da mesma forma, os sistemas de reprodução são responsáveis por moldar a sonoridade de um áudio fixado em suporte em função de suas peculiaridades - a qualidade da agulha ou do motor de rotação de uma vitrola, a estabilidade elétrica de um sistema de amplificação. O discurso comercial do *hi-fi*, vendendo sua ideia de equipamentos “fiéis”, acaba descrevendo situações de escuta laboratoriais - uma escuta típica de um estúdio de masterização, onde a transparência do sistema de amplificação é crucial para a execução correta do trabalho (KATZ, 2015, p. 274), e não uma escuta de fruição.

Os sistemas de monitoramento de um estúdio de masterização são muito similares às salas audiófilas de escuta, trocando leitores comerciais de mídias por computadores e processadores utilizados na manipulação do áudio. Os profissionais que escrevem sobre isso, porém, não descrevem seus equipamentos como *hi-fi* - esse termo praticamente não aparece em seus textos. Eles centram suas recomendações em equipamentos de resposta plana e salas de escuta que não causem problemas acústicos, chamando esses equipamentos de “alta resolução” (KATZ, 2015, p. 273). Enquanto Bobby Owsinski (2008, p. 16) utiliza sete páginas para reforçar a necessidade de utilizar um ótimo sistema de monitoramento, Bob

Katz (2015) escreve dois capítulos inteiros a respeito de caixas de som e salas acústicas. Apesar de não classificarem diretamente seus equipamentos como *hi-fi*, certamente partes de seus discursos replicam características desse universo, mesmo que utilizem um viés mais técnico e menos assumidamente fetichista. Eles julgam a “transparência e fidelidade” de seus sistemas baseados em suas próprias experiências, e não em panfletos de venda ou revistas especializadas – nas quais alguns desses profissionais escrevem para o público amador. Como resume Jonathan Sterne, “fidelidade sonora tem muito mais a ver com fé na função social e organização de máquinas do que com a relação entre o som e sua fonte” (STERNE, 2003, p. 114).

2.2. A ideia de alta-conveniência: portabilidade e praticidade

Não parece haver uma definição universal de *hi-fi* utilizada pela indústria fonográfica. A cada geração de equipamentos, o mercado oferece a última atualização em tecnologia de reprodução, prometendo maior e maior qualidade, e mais e mais “fidelidade” - naturalmente sempre pelo alto custo da tecnologia de ponta que está sendo ofertada no momento. As propagandas de equipamentos de som são povoadas desde sempre com esse discurso comercial.

Os rádios da década de 1920, inicialmente, eram vendidos em *kits* de montagem, voltados a um público bem específico, os *hobbistas* em eletrônica (IAZZETTA, 2009, p. 114). Em outras indústrias, acontece o mesmo: igualmente vendidos em *kits* de montagem, os computadores pessoais eram brincadeira *nerd* no final da década de 1970. O discurso tecnológico também faz parte da ideia de *hi-fi*, tentando diferenciar seu equipamento do produto comum, acessível e popular.

Além do discurso de fidelidade das propagandas dos primeiros reprodutores sonoros, alguns autores identificam outras características. Enquanto Rafael Alexandre da Silva enxerga no rádio “[...] um passo adiante na aproximação de uma democratização musical” (SILVA, 2015, p. 254), David Byrne elogia a portabilidade dos cilindros de cera: “Os cilindros de Edison não soavam realmente tão bem quanto músicos ao vivo, mas você podia carregá-los e tocá-los onde quisesse” (BYRNE, 2015, p. 124).

Alguns dispositivos de reprodução sonora, porém, apesar de utilizarem tecnologias inovadoras, não pretendem redefinir o *hi-fi* ou mesmo serem categorizados como tal. Vários equipamentos comercializados ao longo de toda história da escuta mediada parecem tentar vender uma outra qualidade: a conveniência. Jonathan Wyner, entrevistado por Bob

Katz, aponta o que ele enxerga como grande responsável pelo sucesso de alguns formatos de suporte de áudio:

[...] o sucesso geral não é determinado pela fidelidade [...] vamos pensar um pouco na história dos formatos para consumidores. Nos últimos quarenta anos, quando algum formato teve sucesso no mercado somente pela melhora na fidelidade, sem ter oferecido também uma vantagem óbvia em termos de conveniência ou economia para o consumidor? (KATZ, 2015, p. 167).

De fato, em 1955, a Sony lançou seus primeiros “rádios de bolso”, utilizando a então nova tecnologia dos transístores para diminuir custos e tamanhos dos aparelhos (CUMMINGS, 2013, p. 84). A própria fita cassete, desenvolvida na década de 1960 pela Philips, é uma mídia inteiramente voltada para a praticidade: ela não é melhor do que o vinil em nenhum aspecto técnico, mas sim nas qualidades de uso, isto é, tamanho, resistência e preço - além de serem graváveis, uma capacidade que foge ao vinil (CUMMINGS, 2013, p. 87). Entre vitrolas portáteis e *boomboxes*, todo um nicho de equipamentos com foco na praticidade segue paralelo aos lançamentos mais vistosos do mundo do *hi-fi* - chamaremos esse nicho de *alta-conveniência*, englobando equipamentos geralmente portáteis.

As chamadas “caixas de som *bluetooth*”, como denominadas por seus principais fabricantes, como Sony e JBL, por sua vez, são dispositivos portáteis de reprodução de áudio para pequenas audiências. Elas são herdeiras diretas dos *boomboxes*, os aparelhos de rádio e fita cassete associados com a cena urbana norte-americana do final da década de 1970. Esses podem ser vistos como dispositivos responsáveis por uma ideia de fruição musical coletiva em espaços públicos, uma escuta urbana, móvel - não uma escuta da cidade, mas uma escuta dentro de seu contexto.

Existe toda uma gama de equipamentos práticos que, ao invés de se apoiar em um discurso de “alta-fidelidade” como argumento de venda, oferece características de uso como diferencial. Na primavera japonesa de 1980, um novo dispositivo, desenvolvido nessa linha de alta-conveniência, começa a tornar-se tão popular que provoca e reflete mudanças culturais, inaugurando uma categoria própria no que Shuhei Hosokawa define como *musica mobilis*, “[...] a música cuja fonte, voluntária ou involuntariamente, é movida de um ponto a outro, coordenada pelo transporte físico do possuidor dessa fonte” (HOSOKAWA, 2012, p. 105). Entre ruídos da cidade, músicos de rua e música mediada por alto-falantes, surge o transeunte, que ouve música em seu fone de ouvido. O Walkman leva a escuta privada para as ruas da cidade. Essa nova situação de escuta musical se dá graças ao encontro de diversas características de conveniência em um só dispositivo e por isso Hosokawa o con-

sidera o “objeto definitivo” para a *escuta privada*, ou seja, uma escuta mediada para apenas um indivíduo.

A privatização da escuta, porém, já acontecia em dispositivos não portáteis, utilizando equipamentos *hi-fi* com fones de ouvido - que antes do Walkman eram, como veremos a seguir, grandes, pesados e inadequados para a fruição musical móvel. Já existiam também dispositivos portáteis, como os citados *boomboxes* e rádios de pilha. A união da privatização e da portabilidade, porém, surge com o desenvolvimento do Walkman.

Muito já foi escrito sobre esse dispositivo, mas não foram encontrados estudos sistemáticos sobre os fones de ouvido, transdutores individuais, a parte realmente responsável pela privatização da escuta. Vários textos, porém, abordam o tópico de maneira indireta, tais como o surgimento da fonografia (STERNE, 2003) e do *mp3* (STERNE, 2012), o consumo musical (TAYLOR, 2001), o Walkman (HOSOKAWA, 2012; ACOSTA, 2000; CHAMBERS, 2004; DU GAY et al., 1997; BULL, 2005), e os meios locativos (TUTER; VARNELIS, 2006; BEHRENDT, 2015; VERSTRAETE, 2017).

A seguir, levantaremos brevemente a história e a cultura de escuta em fones de ouvido, encontrando suas origens e funções paralelas à escuta privada, e levantando diversas camadas de valor que podem ser associadas a eles ao longo do tempo, desde seu uso como aparelho militar até como acessório de moda.

2.3. A cultura de escuta por meio de fones de ouvido¹⁶

A origem dos fones de ouvido remete ao que Jonathan Sterne denomina *função timpânica*, que surge no final do século XVIII e começo do XIX, com a identificação e aproveitamento dos mecanismos de transmissão das ondas sonoras entre diferentes meios físicos nas operações auditivas do corpo humano. De acordo com o autor, a partir desses mecanismos, foram derivadas várias técnicas de pesquisa, bem como experimentos científicos, que promoveram a construção de novos equipamentos extensores da audição:

A história do isolamento e a reprodução da função timpânica nos leva de volta à construção do som e da audição como objetos de conhecimento e experimentação no final do século XVIII e no século seguinte. A função timpânica emergiu do encontro da acústica moderna com a otologia, a fisiologia e a pedagogia dos surdos (STERNE, 2003, p. 22).

Leon Scott é o inventor do fonautógrafo, o primeiro dispositivo com transdução so-

16 Este subcapítulo foi adaptado do artigo escrito pelo pesquisador juntamente com Julián Jaramillo Arango e Esteban M. V. Astorga, e apresentado no congresso Sonologia 2019 (BARBOSA; ARANGO; ASTORGA, 2019).

nora. Analisando seus trabalhos, Sterne vê que o mecanismo do ouvido pode ser usado com vários propósitos. Sterne, então, realiza uma arqueologia de dispositivos auditivos, incluindo na história da amplificação auditiva o estetoscópio e a comunicação vias redes telegráfica e telefônica, desenvolvidas no século XIX. Segundo Sterne (2003), esses dispositivos contribuem para a criação dos primeiros equipamentos de reprodução sonora, comercializados pela Edison na última década do mesmo século.

Seguindo o raciocínio de Sterne, a função de uma tecnologia se configura antes de sua materialização em dispositivos e produtos. A função da tecnologia seria “[...] um conjunto de princípios operacionais e filosóficos comuns e, mais importante, encarnações e intensificações de tendências que já existiam em outras partes da cultura” (STERNE, 2003, p. 34).

O autor sugere que o interesse pelo canal auditivo como instrumento de pesquisa desencadeou a proliferação da criação de equipamentos de áudio. Os fones de ouvido não são exceção, sendo um resultado material da descoberta da função timpânica.

Vale ressaltar que, ao encontrar essa conexão entre o estetoscópio e os equipamentos de áudio de Edison, Sterne inverte o papel causal que a tecnologia tem para a cultura. Além disso, pode-se argumentar que Sterne segue o caminho inverso de outros pensadores do áudio, contrabalançando a explicação hegemônica do pós-guerra das origens da cultura de áudio, notavelmente os teóricos culturais da Escola de Frankfurt, como Walter Benjamin ou Theodor Adorno, ou até mesmo compositores das chamadas escolas de música eletroacústica, como Pierre Schaeffer ou Karlheinz Stockhausen. Para esses pensadores e artistas, o pressuposto é que a tecnologia de reprodução sonora desencadeou uma série de práticas culturais e artísticas. Em contraste, de acordo com Sterne,

“[...] as tecnologias de reprodução sonora timpânica são melhor compreendidas como o resultado da proliferação de um conjunto específico de práticas e entendimentos práticos sobre som e ouvido, não como sua causa” (STERNE, 2003, p. 35).

2.3.1. Uma *commodity* tecnológica

Além das relações entre a função timpânica e a criação dos primeiros dispositivos de reprodução sonora, também identificamos uma necessidade social de comunicação como chave para o desenvolvimento de tecnologias de equipamentos sonoros do século XIX, como os fones de ouvido. Essa necessidade, alimentada com as descobertas do sistema auditivo, levou inventores a criarem produtos tecnológicos de áudio para uma modernidade

crescente.

O *Electrophone System* é uma das primeiras invenções da tecnologia de áudio a operar no Reino Unido no final do século XIX. Este sistema de transmissão de som usou as linhas da *National Telephone Company*, difundindo a seus assinantes um sinal de áudio estéreo ao vivo, com apresentações de música em salas de concerto, teatros e igrejas, bem como notícias do mercado financeiro (ESTREICH, 2013).

Fig. 12 - Soldados britânicos hospitalizados, utilizando o Electrofone System.

Fonte: Central News Photo Service (CNPS, 1917)

A figura mostra um modelo de fone de ouvido usado pelos assinantes do *Electrophone System*. Sua forma revela uma proximidade com o estetoscópio, ecoando a ideia de uma extensão auditiva externa ao corpo. O *Electrophone System* também foi uma das primeiras tentativas de comercialização de áudio, mas o empreendimento desapareceu quando as transmissões comerciais de rádio se popularizaram.

Os fones de ouvido foram originalmente desenvolvidos para fins de comunicação militar e civil, onde os *patchbays* telefônicos eram operados manualmente. A necessidade de um dispositivo que não ocupasse as mãos levou a alguns dos primeiros projetos de fones de ouvido. O *Bi-telephone*, patenteado em 1891 por Ernest Mercadier, antecipou os *earbuds*, pequenos fones de ouvido de inserção auricular. Ele também deixa as mãos livres, mas ainda se assemelha bastante a um estetoscópio médico.

Fig. 13 - Patente do Bi-telephone de Ernest Mercadier (esq.) e do dispositivo de Nathaniel Baldwin (dir.).

Fonte: MERCADIER (1891) e BALDWIN (1913)

A patente de 1915 de Nathaniel Baldwin mostra uma ligação mais próxima com um fone de ouvido moderno. Seu dispositivo em formato de tiara alcançou grande sucesso comercial, devido à qualidade de sua amplificação sonora. A história de seu desenvolvimento, depois de várias rejeições, passa por Baldwin enviar seu protótipo para a marinha norte-americana. Após comentários e sugestões de seu novo cliente, ele melhorou o dispositivo, alcançando um *design* mais confortável e recebendo um pedido de 100 unidades da Marinha dos EUA (HOWETH, 1963, p. 149). Baldwin começou construindo suas primeiras encomendas em sua própria casa, ampliando sua produção em fábricas no estado norte-americano de Utah. Com a popularidade do rádio amador como *hobby*, em 1922, ele chegou a produzir 150 fones de ouvido por dia. Construiu mais duas fábricas nas cidades de Chicago e Holladay, e assinou contratos de vendas com empresas do Japão e do Canadá (SINGER, 1979, p. 51).

2.3.2. Equipamentos de escuta

Nos anos 1930, com a expansão das indústrias fonográfica e de rádio, observamos o *design* de novos dispositivos em módulos separados. O módulo final de áudio no equipamento, transduzindo energia elétrica em som, passa a ser comercializado separadamente,

no formato de caixas de som ou de fones de ouvido. No campo do *Sound Design*, o conceito de *Auditory Display* (que traduzimos aqui como “Equipamentos de Escuta”) engloba os fones de ouvido como um caso particular de dispositivo:

Auditory Display abrange todos os aspectos de um sistema de interação homem-máquina, incluindo a configuração, alto-falantes ou fones de ouvido, modos de interação com o sistema de exibição e qualquer solução técnica para coleta, processamento e computação, necessários para obter som em resposta a dados. (HUNT; HERMANN, 2011, p. 1).

O conceito de *Auditory Display*, relevante no campo de pesquisa de *design*, nos ajuda a perceber novas funções do fone de ouvido, pois ele é utilizado como equipamento de transdução para som de uma imensa quantidade de produtos de consumo utilizados no dia-a-dia. Essa ampliação do uso leva a mais desenvolvimentos tecnológicos. Assim, em 1937, a empresa alemã Beyerdynamic lança o primeiro fone de ouvido dinâmico (BEYER-DYNAMIC, 2014), o DT-48. Monofônico, o equipamento introduz a tecnologia de transdução dinâmica, que até hoje é utilizada na maioria dos fones de ouvido comerciais, devido ao seu custo reduzido e baixa distorção.

2.3.3. O (pseudo)cientificismo mercantil

Os interesses econômicos da indústria fonográfica e do rádio na década de 1940 também participaram da transformação social dos fones de ouvido. Com a proeminência dos modelos unidirecionais de comunicação, os setores de publicidade e *marketing* passaram a elaborar um novo universo simbólico associado ao áudio, materializado na noção de alta-fidelidade, que proporcionou uma nova camada de valor para os fones de ouvido. A ideia do áudio *hi-fi* reúne valores e ideologias circuladas nas campanhas publicitárias da época, que atingiram os consumidores em seu ponto mais fraco. A proliferação de aparelhos domésticos, destinados à classe média norte-americana, tornou os fones de ouvido um dos padrões dessa escuta tecnológica. De acordo com Fernando Iazzetta,

O termo *hi-fi* (high-fidelity) foi introduzido pela indústria fonográfica como uma das estratégias de mercado para convencer seus consumidores de que a música reproduzida por aparelhos como fonógrafos, gramofones e, posteriormente, toca-discos era “fiel”, ou seja, era uma representação adequada da música realizada ao vivo (IAZZETTA, 2009, p. 114).

Analisando as campanhas publicitárias de dispositivos de áudio na sociedade americana do pós-guerra, o pesquisador Timothy Taylor destaca a noção de cientificismo mercantil como “uma crença, negativa ou positiva, nas qualidades infáveis da ciência e tecnologia”

(Taylor, 2001, p. 76). Sendo importantes dispositivos dentro do sistema de áudio doméstico *hi-fi*, os fones de ouvido foram além dos símbolos de diferenciação social, adquirindo um valor relacionado a um lazer futurista, proporcionado pelos avanços da ciência. Além disso, Taylor aponta questões de gênero na materialização deste cientificismo mercantil no áudio *hi-fi*, identificando a consolidação de uma figura masculina como principal consumidora dessa tecnologia.

O cientificismo mercantil no cenário doméstico para homens é provavelmente melhor representado pela moda *hi-fi* do final da década de 1940 e da década de 1950. Muitos dos primeiros entusiastas do *hi-fi* foram homens que aprenderam sobre tecnologia de áudio no exército, durante a Segunda Guerra Mundial. Eles mantiveram esse interesse tornando-se hobbistas de *hi-fi* quando voltaram para casa (TAYLOR, 2001, p. 79).

De acordo com Taylor, dentro dos valores do período, um fone de ouvido *hi-fi* “[...] permite que seu usuário escape da produção de mídia de massa, destinada ao público feminino” (TAYLOR, 2001, p. 80). Dessa forma, o homem tinha acesso a um dispositivo qualificado para ouvir produtos musicais de “alta cultura”. Os eletrodomésticos para cozinhar e limpar, por sua vez, eram geralmente associados a uma “versão feminina” do *hi-fi*:

[...] O público masculino utilizou-se dessas tecnologias para confrontar o domínio feminino dentro do ambiente familiar, apoiando-se na obscuridade que velava o funcionamento daqueles aparelhos e no volume sonoro que podiam produzir, afrontando, desse modo, a ordem e a tranquilidade imposta pela dona de casa (IAZZETTA, 2009, p. 121).

O que Taylor e Iazzetta sugerem é que, durante as décadas de 1940 e 1950, a noção consumista de *hi-fi* contribuiu para a definição de identidades de gênero. Particularmente, ao transformar a casa do usuário em um ambiente de lazer tecnológico, o equipamento de áudio *hi-fi*, como os fones de ouvido, ajudou a exaltar a masculinidade do usuário.

Durante a década de 1950, transístores gradualmente substituíram as válvulas dentro dos equipamentos de áudio. Componentes transistorizados construídos pela AT&T permitiram a construção de amplificadores de áudio mais leves, menores e mais eficientes, que foram rapidamente implementados em aparelhos auditivos para mitigar a perda de audição. A partir disso, de acordo com Santiago Acosta, “[...] em 1952, os senhores Morita e Ibuka desejavam adquirir os direitos de produção [desses aparelhos], e foram concedidas licenças a uma pequena empresa japonesa, recém-nomeada Sony, por US\$ 25.000” (ACOSTA, 2000, p. 20).

Graças ao transistor, durante as décadas de 1950, 1960 e 1970, corporações como a Sony lançaram no mercado rádios menores, além de dispositivos com capacidades de gravação e navegação no suporte, como rebobinar e avançar. As variedades funcional e visual dos equipamentos de áudio também sugerem uma transformação nos hábitos do usuário, refletindo a inserção do consumo de áudio no estilo de vida. Como consequência, os fones de ouvido aderiram a esse processo e a marca Koss lançou modelos como o SP/3 e o ESP/6.

A ideia do *hi-fi* era frequentemente evocada para vender esses dispositivos. O SP/3, criado em 1958, foi o primeiro fone de ouvido estéreo a ser vendido (KOSS, 2015). Já o ESP/6, lançado em 1968, e também estava alinhado a uma prática de escuta audiófila. Esse modelo foi o primeiro fone de ouvido eletrostático lançado no mercado, isto é, cuja indução magnética ocorre no próprio diafragma, que está composto por finas camadas de superfícies condutoras que vibram quando excitadas pelo sinal de áudio elétrico.

Em relação à audiofilia, os fabricantes de dispositivos de áudio passaram anos otimizando as tecnologias e oferecendo a seus consumidores produtos com maior qualidade. Com o aumento exponencial da qualidade (e dos preços aceitos pelo consumidor), surge um nicho de mercado para o consumo dessas tecnologias de ponta. Um exemplo é o primeiro modelo com cancelamento de ruídos ambientes, fabricado pela Bose e lançado em 1995.

Vale ainda mencionar que, em 1966, dois anos antes do lançamento do ESP/6, a Koss lançou os famosos Beatlephones, cuja principal inovação não era técnica, mas social, comercializando o equipamento para novos públicos. Os Beatlephones tinham o corpo de cor pintado de azul e fotos dos membros da banda, proporcionando, através do patrocínio de artistas, uma nova camada de valor aos fones de ouvido (STUFF, 2013).

2.3.4. Uma estratégia urbana

Com o lançamento do Walkman em 1979 e sua rápida popularização a partir do ano seguinte, o consumo de dispositivos de áudio torna-se mais complexo. Explorando a portabilidade, a Sony muda o paradigma de reprodutores de áudio como aparelhos domésticos *hi-fi*. O modelo de fone de ouvido supra-aural do Walkman era mais leve e menor, uma vez que foi projetado para ser usado em ambientes urbanos.

O Walkman delineou uma transformação cultural que não ficou indiferente aos estudos acadêmicos. Assim, uma série de teorias foram elaboradas, apontando diferentes aspectos que são relevantes para o nosso pensamento sobre as funções e camadas de

valor dos fones de ouvido. Essas teorias acumulam-se no campo de estudos que aborda as culturas de escuta móvel. O texto a seguir pretende contribuir com esse campo, discutindo algumas das principais ideias e autores.

2.3.4.1. Os efeitos do Walkman

O texto seminal de Shuhei Hosokawa, *The Walkman Effect* (HOSOKAWA, 2012), escrito em 1984, foi o primeiro a relacionar o então novo dispositivo à vida urbana, discutindo o Walkman como um equipamento do transeunte pela cidade moderna. Hosokawa discute conceitos como portabilidade, singularidade, autonomia e construção/desconstrução de significado como fatores chave no desejo de isolamento social. O autor esboça toda uma teoria musical sobre portabilidade, a *musica mobilis*, levando em consideração muitas relações entre a fruição musical e as estruturas urbanas.

Do ponto de vista tecnológico, Hosokawa afirma que o Walkman foi uma regressão, pois não tem operação de gravação nem possui terminal transdutor próprio, um alto-falante acoplado. A idéia de reduzir as funções foi concebida pelo presidente da Sony, Akio Morita, ao caminhar pelas ruas de Nova York. Além disso, a campanha publicitária do Walkman apontou para um estilo de vida urbano em que a simplificação tecnológica se torna um aspecto positivo. Ao coletar argumentos conceituais e técnicos, Hosokawa sugere que o Walkman é uma estratégia, uma tática para a vida urbana contemporânea.

O britânico Ian Chambers (2004) estendeu alguns dos pensamentos abordados no texto original de Hosokawa, notavelmente discutindo a construção de uma identidade diaspórica da vida em trânsito. Para Hosokawa, a construção da identidade levantada pelo uso do Walkman materializa-se na noção de singularidade, enquanto Chambers chama a atenção para a desestabilização do local causada pelo Walkman:

Sua qualidade incomum reside em confundir, propositalmente, os limites prévios, a partir da sua provocante aparência de “fora de lugar”. Agora, a confusão de “lugar”, de vozes, histórias e experiências que falam de “fora de lugar” faz parte do conjunto de sentidos mais amplo da semântica contemporânea e da crise política. Uma hierarquia espacial anterior precisou, aos poucos, confrontar um excesso de linguagens que emergiram das histórias e linguagens do feminismo, dos direitos sexuais, da etnicidade, da raça e do meio ambiente, que transbordam e minam sua autoridade. O Walkman é, portanto, um ato político? É certamente um ato que inconscientemente se entrelaça com muitas outras micro-atividades, conferindo um sentido diferente à polis. Ao produzir um sentido diferente de espaço e tempo, ele participa na reescrita das condições de representação, sendo que “representação” indica claramente tanto as dimensões semióticas do cotidiano quanto a potencial participação em uma comunidade política (CHAMBERS, 2004, p. 100).

Nos campos da sociologia e dos estudos culturais, o Walkman forneceu uma motivação para a elaboração de trabalhos acadêmicos detalhados e didáticos, abordando tópicos e métodos de pesquisa em Ciências Sociais. O texto *Doing Cultural Studies, The Story of the Sony Walkman* (DU GAY et al., 1997) reúne cinco ensaios e cinco autores (Paul Du Gay, Stuart Hall, Linda Janes, Hugh Mackay e Keith Negus), que analisam o fenômeno cultural do Walkman a partir de perspectivas como representação, identidade, produção, consumo e regulação:

Escolhemos o Walkman porque é um típico artefato cultural e um meio da cultura moderna. Através do estudo de sua “história” ou “biografia” pode-se aprender muito sobre as maneiras pelas quais a cultura funciona em sociedades modernas recentes como a nossa (DU GAY et al., 1997, p. 2).

Embora o principal tema do texto sejam os estudos culturais, e não o Walkman, a escrupulosa pesquisa desenvolvida pelos autores permite uma análise mais detalhada das várias nuances culturais do Walkman e dos fones de ouvido. O estudo engloba a análise de material publicitário e jornalístico, entrevistas com executivos e *designers* da Sony, bem como uma bibliografia que confronta muitos dos teóricos dos estudos culturais do século XX. Dessa forma, o texto responde de forma sistemática ao complexo problema do significado cultural do Walkman.

O capítulo 5 é particularmente relevante para a discussão, pois discute o consumo e a evolução histórica das teorias que o explicam. O consumo é caracterizado nas sociedades contemporâneas como uma prática produtiva que envolve apropriação e resistência. O capítulo 6 trata dos mecanismos da sociedade para regular o uso do Walkman. Segundo os autores, o dispositivo provoca uma ruptura nas classificações estabelecidas, que separa as esferas pública e privada. Os autores discutem o debate sobre modos apropriados e inapropriados de comportamento público, surgido logo após o lançamento do Walkman na Grã-Bretanha.

O livro de Michael Bull, *Sounding out the city - Personal Stereos and the Management of Everyday Life* (2000), segue a mesma direção que o estudo realizado por Du Gay et al. (1997), sendo posicionado no campo dos estudos culturais. A originalidade de Bull reside na metodologia de pesquisa, uma vez que o autor apela para métodos etnográficos para chegar às suas conclusões. Ao fazê-lo, o autor aponta novos argumentos sobre o uso do Walkman, a partir de entrevistas e grupos focais, entre outros métodos empíricos de pesquisa de Ciências Sociais.

Um comentarista mais recente de Hosokawa, Pieter Verstraete (2017), reflete sobre as qualidades do Walkman para construir e desconstruir valores do dispositivo:

Este passeio de áudio (ou teatro) ativa a sensação ambígua de estar no comando do ambiente e flerta com o voyeurismo, oferecido pelo sigilo do equipamento móvel de música, que ajuda a pessoa a se sentir segura enquanto olha descaradamente para os outros, que se tornam parte de seu teatro secreto (VERSTRAETE, 2017, p. 13).

Cerca de vinte anos após o Walkman, algumas tecnologias amadureceram e convergiram para renovar a *musica mobilis* na cultura urbana, adotando um suporte de fixação de áudio invisível, intangível, virtualizado. Como Jonathan Sterne explica em seu livro *Mp3: The Meaning of a Format* (2012), a cultura da troca de músicas no final do século XX renasceu com novas formas de armazenar, transportar e copiar música.

2.3.4.2. Os fones de ouvido na era do computador

Como veremos em detalhes ainda neste capítulo, com a proliferação de computadores pessoais com funções multimídia e leitores de *compact-disc*, as músicas passam a ser extraídas de seu suporte físico e virtualizadas em arquivos digitais de computador. Com o desenvolvimento de algoritmos que reduzem o tamanho dos arquivos digitais de áudio - entre eles, o formato *mp3* -, as músicas puderam ser compactadas em arquivos menores, sendo facilmente armazenadas e, graças à popularização da internet nos países desenvolvidos, também distribuídas entre aqueles interessados em acumular arquivos *mp3* em seus computadores.

Logo surgiram dispositivos portáteis capazes de reproduzir o novo formato de áudio. Eles eram extremamente limitados no começo, quase como protótipos funcionais. Um dos primeiros dispositivos capazes de reproduzir arquivos *mp3* foi o Rio Multimedia, lançado em 1998 (STERNE, 2012, p. 203). Tinha 32 megabytes de memória, o suficiente para disponibilizar cerca de meia hora de música no *bitrate* utilizado da época. Apenas três anos depois, a Apple começou a vender seus reprodutores de *mp3*. O primeiro iPod, de 2001, tinha cinco gigabytes de espaço - ou “mil músicas no seu bolso”, como a campanha de *marketing* propagandeava.

Em consonância com as tentativas da indústria fonográfica de conter a pirataria, a Apple lança a plataforma de vendas digitais da iTunes Store em 2003, estabelecendo um novo mercado (CUMMINGS, 2013, p. 220). Isso permitiu a comercialização dos fonogramas digitais para um público mais amplo do que os *geeks* da internet: a Apple vende um

milhão de arquivos na primeira semana, 25 milhões em oito meses (HILL, 2013) e, em 2008, se torna a maior vendedora de música dos Estados Unidos (APPLE, 2008).

No mercado de computadores pessoais, o *design* de produtos da Apple sempre foi um diferencial, sendo parte de sua estratégia de *marketing*. Seu “tocador de mp3” tinha que ser único e incomum.

A Sony popularizou os fones de ouvido super-aurais - pequenos e discretos para a época - vendendo-os com o Walkman. Os aparelhos, embora portáteis, dificilmente eram ocultáveis, e o Walkman era pensado para ser usado fixado à roupa - alguns modelos contavam com cliques e alças para tal. Os iPods, lançados mais de vinte anos depois, usavam componentes muito mais miniaturizados. Embora as imagens publicitárias mostrassem silhuetas segurando o dispositivo na mão, até mesmo o *slogan* das mil músicas sugeria o bolso como o local mais adequado para mantê-lo.

Devido ao seu tamanho e descrição, o apelo de exclusividade do produto teve que ser refletido na parte mais visível do conjunto: desde o seu primeiro modelo, o iPod foi vendido com um par de inovadores fones de ouvido brancos. A campanha publicitária do iPod de 2003 ilustra como a Apple pretendia usar os fones de ouvido para se destacar da concorrência.

Fig. 14 - Campanha publicitária do iPod da Apple.

Fonte: Elvert Barnes (2006).

Com uma campanha publicitária aproximadamente cem vezes mais cara do que a de seus concorrentes (ISAACSON, 2011, p. 392), a Apple não apenas se posicionou confortavelmente no mercado de reprodutores portáteis como também colaborou para renovar e multiplicar o uso de fones de ouvido nas ruas. Os produtos da marca Apple, diferenciados socialmente pelo uso de fones de ouvido brancos, foram copiados pelos concorrentes pouco tempo depois, comprovando a eficiência da estratégia.

Os fones de ouvido, que lentamente saíram de aplicações militares para desempenhar um papel decisivo na privatização da música, agora também serviam como objeto de ostentação na paisagem urbana.

2.3.4.3. A transformação em item de ostentação

Logo após os primeiros anos de venda de iPods e outros “tocadores de mp3”, o rapper norte-americano Dr. Dre, o produtor musical Jimmy Iovine e o engenheiro Noel Lee se uniram para criar fones de ouvido que combinavam qualidade de som com moda e força de marca. Apelando aos amantes do gênero hip-hop, a resposta de áudio dos fones de ouvido foi projetada para melhorar os padrões de ritmo e as frequências dos graves. A marca Beats by Dr. Dre lançou seu primeiro par de fones em 2008, associando seus produtos com celebridades da música e do esporte, em campanhas agressivas com inserção de produtos em eventos, dando aos fones de ouvido o status de acessório de moda e item de luxo. Eles são circum-aurais (que cobrem toda a orelha), maiores do que os fones de ouvido da Apple e, portanto, muito chamativos e sem nenhuma discrição.

Fig. 15 - Jogador de basquete LeBron James com seus fones de ouvido do Dr. Dre.

Segundo os criadores, o objetivo não era competir com marcas famosas de fones de ouvido como Sennheiser ou Bose, mas disputar o dinheiro dos jovens que comprem um par de tênis. O sucesso do empreendimento refletiu nos números da indústria, com um aumento de 32% na receita das vendas do equipamento. Os fones de ouvido agora passam a ter outra função social, como reflete Greenburg:

Assim como um tênis Nike era tanto uma declaração de moda quanto um equipamento necessário à prática de esportes, a Beats - com sua linha de produtos disponíveis em uma miríade de cores - tornou-se rapidamente um acessório e um dispositivo de áudio. “Mudamos a maneira como os fones de ouvido são parte do estilo de vida”, diz Lee. “Estabelecemos que fones de ouvido são legais de serem usados ao redor do pescoço, mesmo que eles não estejam conectados”. Brian Dunn, executivo-chefe da Best Buy, acrescenta: “Não se trata apenas do som; é também uma questão de moda” (GREENBURG, 2013).

Fica claro, no trecho acima, que os fones de ouvido adquirem um novo valor, para além da fruição musical ou do isolamento do ruído urbano. Na década de 1960, eles já estavam associados a bandas famosas como os Beatles, mas desde o lançamento do iPod, ganham uma nova camada de significado, com um apelo claro de status e moda, perdendo por vezes até mesmo sua função de reprodução de som.

2.3.4.4. Áudio localizado

Alguns dispositivos digitais de áudio fornecem operações que ampliam a função de reprodução de arquivos de áudio. Tanto no setor de jogos quanto no de aplicativos móveis, a experimentação criativa no campo do áudio interativo cresceu nas últimas décadas. Na construção dessas novas narrativas, a integração de um microfone ao fone de ouvido produziu uma nova e poderosa interface homem-computador. Esse novo módulo de mão-dupla deu origem a novas formas de explorar realidades virtuais e mistas, nas quais nem a visão nem as mãos estão ocupadas. Entre essas múltiplas e novas narrativas, discutiremos rapidamente os fones de ouvido no contexto do áudio locativo.

A mídia localizada tira proveito da computação sensível ao contexto (TOWNSEND, 2006), trabalhando a partir do pressuposto de que o acesso onipresente à internet está mudando nossa relação com o espaço, ao sobrepor um segundo mundo virtual ao físico (TUTERS; VARNELIS, 2006). Sob essas condições, os fones de ouvido tornam-se uma ferramenta indispensável para o transeunte acessar as novas tecnologias de cidades conectadas, “inteligentes”.

A especificidade da mídia sonora localizada, ou do áudio locativo, tem sido discutida

por Frauke Behrendt (2015) a partir de quatro categorias diferentes: som colocado, mobilidade sonorizada, plataformas sonoras e instrumentos musicais. A autora destaca diferentes usos para informações de GPS (sistema de localização geo-referencial) em aplicativos de áudio, no que chama de sons localizados:

A popularidade dessa categoria é demonstrada pelo fato de que, ao longo da última década, surgiram vários temas ou subgêneros da categoria “som localizado”, incluindo sons mais narrativos, como históricos, turísticos, educativos, ficcionais, jogos, e outros menos narrativos, como música e som experimental (BEHNRENDT, 2015, p. 9).

Na experiência de áudio locativo, os fones de ouvido desempenham um papel diferente do de um reproduutor normal de música. O fone de ouvido, acrescido de um microfone, permite que o usuário envie e receba informações sem ocupar nem a visão ou as mãos. A relevância do GPS e das informações *on-line* sobre a cidade está crescendo, e as ferramentas para acessar essas informações estão ficando cada vez mais sofisticadas. Espera-se que, em breve, esse seja um campo de exploração ainda maior para os *designers* de fones de ouvido de áudio localizado.

2.3.5. Conclusões sobre os fones de ouvido

Como observamos neste subcapítulo, os fones de ouvido podem ser vistos como um ponto de convergência onde se encontra um conjunto de valores tecnológicos e culturais, onde coexistem questões sobre música, mercado, alteridade, cidadania, gênero e moda. Além desses cruzamentos, também lembramos da questão do suposto empoderamento causado pela fruição musical com mediação tecnológica, desvelando as contingências culturais e sociais que interagem na definição de sua função. Pudemos também, como assinala Sterne, pensar como “[...] uma prática social de quase dois séculos - o isolamento em um mundo de sons, iniciado por médicos no início do século XIX - pôde ser articulada de novas maneiras” (STERNE, 2003, p. 337).

Além do uso científico abordado por Sterne ou do sucesso comercial dos fones de ouvido de Baldwin para uso militar, os fones de ouvido juntaram-se à evolução da civilização humana recente, tornando-se um símbolo de modernidade. Nas primeiras décadas do século XX, eles eram um dispositivo de comunicação rudimentar, usado nos sistemas de rádio e telefonia. Com a expansão das indústrias fonográfica e de rádio, foram prontamente explorados como um dispositivo de fruição musical - é quando eles assumem sua função principal ao longo do século XX.

Embora o consumo de música seja um valor da classe média do pós-guerra, os fones de ouvido foram transformados em uma espécie de mercadoria de escuta tecnológica fundamentada em uma confiança inefável na inovação científica e materializada no conceito de *hi-fi*. Além disso, de acordo com Taylor e Iazzetta, durante os anos 1950 e 1960, os fones de ouvido eram um definidor de gêneros, uma vez que estavam associados a um universo masculino.

A miniaturização do transistor trouxe novos modos de uso para os fones de ouvido, principalmente com a popularização do Walkman nos anos 1980. Eles transformaram-se em uma estratégia urbana para os transeuntes em circulação pela cidade moderna. Assim, valores como mobilidade e portabilidade ganharam peso como valores de consumo. A função dos fones de ouvido tornou-se uma questão de estilo de vida.

A partir da década de 1980, os fones de ouvido são usados deliberadamente como isoladores de contexto, não sem provocar uma ruptura nos comportamentos tradicionais entre espaços públicos e privados. A combinação iPod e iTunes explorou as tecnologias de internet de áudio em arquivos, e mudou o cenário do mercado de consumo, oferecendo fones de ouvido como um extra.

Enquanto a privatização da escuta musical era simbolizada por fones de ouvido da marca da maçã, Dr. Dre lançava seus fones de ouvido com apelo à moda, não demorando para que eles se tornassem itens de luxo. Enquanto isso, fones de ouvido equipados com microfones também passaram a ser adotados em narrativas interativas, como jogos e aplicativos para *smartphones*. Particularmente no campo emergente do áudio locativo, os fones de ouvido desempenham um papel decisivo como interface de computador.

Como alguns desses valores dos fones de ouvido têm uma validade temporária, eles se sobrepõem na contemporaneidade. Como resultado, podemos concluir que os fones de ouvido têm muitas funções possíveis, relacionadas aos hábitos de consumo de áudio. A partir dessa perspectiva, pode-se interpretar a história dos fones de ouvido como também uma história do consumismo, *marketing* e publicidade de áudio. Enquanto os usuários do Walkman trocaram seus sistemas *hi-fi* caseiros por um sistema mais conveniente para o uso urbano, os usuários de fones de ouvido da marca Beats by Dr. Dre absorveram completamente essa função urbana, adicionando novas conotações a eles. A evolução do uso dos fones de ouvido mostra como o consumismo tornou-se uma operação ativa: uma ação em que a identidade é definida e construída.

Por fim, perguntamos-nos: por que o Walkman recebeu tanta atenção acadêmica, enquanto é difícil encontrar estudos dedicados aos fones de ouvido? É interessante notar que algumas das análises dedicadas ao Walkman e ao iPod poderiam ser atribuídas não aos dispositivos reprodutores, mas aos fones de ouvido, como a qualidade do isolamento, a ruptura do público e do privado ou mesmo a associação com um estilo de vida. Desse ponto de vista, o Walkman poderia ser interpretado como um fetiche acadêmico, que acabou por mascarar a relevância dos fones de ouvido como um símbolo da cultura moderna.

2.4. Da escuta privada à personalizada

O Walkman e seu compacto fone de ouvido alteram aspectos práticos da escuta musical móvel, levando a escuta privada para os espaços públicos. Há melhoras técnicas trazidas pelo Walkman, como fones de ouvido menores e mais leves, e a miniaturização do leitor de fita cassete, para que ele pudesse ser acoplado ao corpo. Segundo Shuhei Hosokawa, essas melhoras são secundárias em comparação com a mudança social que esse aparelho proporciona (HOSOKAWA, 2012, p. 107).

Uma plateia de concerto ou de cinema em silêncio, fruindo um espaço sonoro coletivo, está, segundo Jonathan Sterne, atomizando sua experiência coletiva em escutas individuais (STERNE, 2003, p. 161). Essa cultura de escuta individual reflete-se na privatização da escuta trazida pelo fone de ouvido, para, mais tarde, ser levada às ruas com o Walkman. Rafael Alexandre da Silva resume esse caminho da escuta do coletivo ao privado:

Aquele passo dado pela criação do gramofone em direção à democratização do acesso à música, e também à virtualização do intérprete, agora avançava em direção a um novo rumo: a uma experiência individual de vivência musical. O que antes fora, desde sempre, desde os homens das cavernas até os príncipes das cortes europeias, algo que se fazia exclusivamente em conjunto, que é ouvir música, pela primeira vez na história do homem foi feito de maneira individual (SILVA, 2015, p. 257).

O usuário sempre teve alguma capacidade de decisão na escolha do que seria reproduzido em dispositivo. Ao selecionar um cilindro de cera ou gravar uma fita cassete com suas canções preferidas, ele exerce seu poder sobre o conteúdo sonoro. Naturalmente, cada mídia permite diferentes liberdades nessa escolha: carregar centenas de discos de vinil ou navegar no conteúdo de uma fita cassete até o ponto desejado podem ser um problema. Assim, diferentes mídias precisam de distintos manuseios, gerando contingências na escuta, como veremos a seguir.

2.4.1. Manuseio e práticas de escuta

Os discos de acetato de 78 rotações por minuto, do começo do século XX, exigiam um complexo manuseio antes de serem tocados. O disco era escolhido a partir das informações de texto e imagens da capa e da lombada, delicadamente retirado de sua proteção e posicionado no prato do toca-discos. Com ele girando, a agulha era manualmente posicionada e, finalmente, do áudio registrado era feito som. Como esses discos tinham capacidade para comportar menos de quatro minutos de áudio, ou apenas uma ou duas músicas por lado, para ouvir outras composições, era necessário trocar o disco e repetir todo o procedimento.

O tempo de preparação para a escuta de uma música gravada não foi muito alterado com a popularização, na década de 1950, dos *long play* de vinil em 33 rotações por minuto. Eles continham várias músicas em cada lado de suas 12 polegadas, com aproximadamente de 20 a 24 minutos por lado, mas, além de todo aquele manuseio dos *singles* em acetato, era necessário posicionar a agulha na posição correta do disco para escolher a canção desejada, obrigando o usuário a observar o espaçamento maior entre os sulcos, uma indicação visual da mudança das faixas no álbum.

As *jukebox*, comuns em locais públicos nos Estados Unidos da década de 1950 e 1960, automatizaram a mecânica desse processo, tocando os *singles* escolhidos a partir de botões (CASTRO, 2014, p. 162). Quando acionado, o mecanismo posicionava o disco selecionado no eixo giratório, a agulha era colocada na posição exata e a música, reproduzida. Apesar do suporte ser o mesmo, seu reproduutor distinto gerava outras maneiras de interagir com ele.

As fitas cassete, criadas nos anos 1960 mas popularizadas na década seguinte (BYRNE, 2012, p. 109), pela própria natureza do suporte, possuíam outras características de manuseio: são compostas por um invólucro plástico que protege uma delicada fita magnética, na qual o áudio é registrado. Fisicamente mais resistentes e menores do que os discos, as fitas podem ser manuseadas com menos cuidado. A escolha da música, porém, tende a ser ainda mais lenta: por se tratar de uma fita contínua com as músicas em sequência, é necessário utilizar botões para acelerar ou retroceder a fita, navegando até o ponto de reprodução desejado.

Já os *compact-discs*, que surgem nos anos 1980, retomam, em versão digital, a delicadeza física do disco de vinil. Os procedimentos para seu manuseio são um tanto diferentes,

mas a necessidade de manipular a mídia apenas pelas bordas, evitando a superfície de leitura, revela a sua fragilidade.

Sua dimensão, um disco de 12 centímetros de diâmetro, permite que o usuário o posicione no reprodutor com apenas uma mão. O diodo que emite o laser de leitura do disco pode ficar visível durante a manipulação do *compact-disc*. Em operação, porém, o laser é ocultado pelo corpo do reprodutor, substituindo o antigo posicionamento manual da agulha. A seleção das músicas é feita de maneira quase instantânea, com a utilização de botões de controle que fazem a reprodução saltar rapidamente de faixa em faixa.

Como ilustramos brevemente, a materialidade de cada tipo de suporte de áudio determina limites de manuseio e práticas de escuta. A virtualização do fonograma vai separar o áudio de um suporte físico palpável. Logo, surgem outras possibilidades de práticas de escuta, ainda dependentes das limitações dos reprodutores, mas não da fisicalidade dos suportes de fixação de áudio.

2.4.2. Virtualização do fonograma

A segunda metade da década de 1990 testemunha a popularização, nos Estados Unidos, do uso da internet a partir dos navegadores Netscape, lançado ainda em 1994, e Internet Explorer, introduzido pela Microsoft como parte dos sistemas Windows em 1995. A conexão à rede de computadores era extremamente lenta, como lembra Jonathan Sterne (2012, p. 202): “em 1994, a maior parte dos usuários de internet ainda usava conexões discadas, e variações de alguns quilobytes faziam grande diferença no tempo de enviar ou receber um arquivo.”

Nesse contexto, o formato de arquivo digital de áudio MPEG-1 Audio Layer III (popularmente conhecido como *mp3*), criado poucos anos antes, torna-se peça essencial para a troca de arquivos digitais de áudio pela internet. Uma música de três minutos arquivada em um *compact-disc* ocupa 31,75 megabytes. Essa mesma música, compactada em *mp3* com 128 kbps de *bitrate*¹⁷, ocupa aproximadamente 2,81 megabytes¹⁸. E, após a conversão, que inclui alguma perda de informações do áudio devido à compactação para *mp3*, o arquivo digital de áudio pode ser copiado livremente, sem qualquer degradação, como qualquer

17 O *bitrate* de um arquivo *mp3* é a medida de fluxo de dados, em quilobites por segundo (kbps), e está diretamente relacionado ao tamanho final do arquivo, implicando também na qualidade de armazenamento do áudio. Segundo Jonathan Sterne (2012, p. 14), 128 kbps era *bitrate* utilizado como padrão até recentemente.

18 Cálculos de tamanho de arquivo feitos a partir de informações do site <http://www.audiomountain.com/tech/audio-file-size.html>

outro arquivo de computador - diferentemente das cópias em fita cassete, onde cada geração de cópia significa uma perda de qualidade.

O que antes era, portanto, um grande arquivo extraído de um *compact-disc*, agora pode ser enviado e recebido pela internet em alguns minutos. Assim como com as fitas cassete, a conveniência de adquirir e acumular músicas ilegal e gratuitamente passa a ser mais importante do que a reprodução de alta-fidelidade:

Do mesmo modo que a indústria fonográfica durante o século XX transformou cada ouvinte num colecionador ansioso por possuir mais gravações do que era capaz de ouvir de fato, atualmente há uma verdadeira caçada musical na internet. Quanto mais músicas o ouvinte possui, melhor, independentemente de sua qualidade musical ou sonora (IAZZETTA, 2009, p. 128).

Ainda preso dentro de um reproduzidor volumoso (o computador pessoal), mas livre das limitações de um suporte físico, como um disco de vinil ou um *compact-disc*, o fonograma agora está gravado em *bits* eletrônicos invisíveis, sem fisicalidade aparente. A esse arquivo digital de áudio, sem suporte palpável, denominamos *fonograma virtualizado*. Nesse formato invisível, ele pode ser colecionado, copiado, manipulado e consumido de formas anteriormente impossíveis.

2.4.3. *Singles, álbuns e playlists*

A configuração física dos primeiros suportes de fixação de áudio comercializados, como o cilindro de cera e o disco de goma-laca, limitava o tempo de reprodução a uma ou duas músicas. Segundo Mark Katz (2010), as composições de música popular das primeiras décadas do século XX, exemplarmente as de *jazz* nos Estados Unidos, foram fortemente influenciadas pela necessidade de caberem dentro dos menos de quatro minutos de um disco de 78 rotações. Nas apresentações ao vivo, essa limitação desaparecia, pois os artistas podiam permanecer tocando por muito mais tempo. Mas são os discos, com suas músicas de duração limitada, que sobreviveram como testemunhas dessas performances sonoras:

Com algumas exceções, a duração das gravações e a brevidade dos solos permaneceram constantes pelos trinta e um anos entre as primeiras gravações de *jazz* e a introdução dos discos *long play*, em 1948. [...] A maneira como o começo do *jazz* é entendido [...] muitas vezes passa por sua preservação fonográfica, que esteve por mais de três décadas sujeita a uma severa e arbitrária limitação de tempo (KATZ, 2010, p. 86).

Mesmo com a disseminação dos *long play* na década de 1950, a duração de aproximadamente três minutos para as músicas populares não foi alterada. Mark Katz considera que é comercial a razão dessa estabilidade, pois o lucro principal da indústria vinha, ainda,

da venda de *singles* de 78 rotações por minuto, para uso em máquinas de *jukebox* (KATZ, 2010, p. 40). Mas o fato é que o *long play* criou um novo padrão de consumo de músicas, em pacotes de canções, com a venda de conjuntos de músicas: o álbum. Nele, a organização linear das músicas sugere uma ordem de escuta.

Esse formato de vendas, inicialmente dedicado a coletâneas e a nichos de menor vendagem do que a da música popular, torna-se predominante na década de 1960, com a utilização do formato por artistas como The Beach Boys e The Beatles na construção de um discurso que permeasse as canções. A partir deles, a indústria passa a utilizar o álbum como seu principal produto, mais por motivos pragmáticos do que artísticos: o *long play*, na década de 1960, era vendido ao consumidor final por até cinco vezes o preço de um *single* (HEYLIN, 2012, p. 8).

Apesar das dificuldades no manuseio das mídias, como descrevemos anteriormente, e do domínio na venda de álbuns desde a década de 1960, a escuta dos fonogramas individuais nunca deixou de ser realizada: o espaçamento físico entre as faixas do disco de vinil, os botões de acelerar e retroceder nos reprodutores de fita cassete e os de salto de faixa nos tocadores de *compact-disc* são prova disso.

Dos suportes físicos para áudio de grande sucesso comercial, o que melhor se dispõe a ser copiado é a fita cassete. É impossível realizar uma gravação caseira em disco de vinil, e os gravadores de áudio em *compact-disc* não se tornaram populares - talvez pela proximidade cronológica com a geração seguinte de tecnologias, que iria virtualizar tanto o fonograma quanto as trocas de músicas.

A expressão *mixtape* vem do uso da fita cassete como suporte para cópias de músicas, definida por Alex Cummings como “uma coleção própria de músicas gravadas em fita cassete e compartilhadas em um pequeno círculo de amigos¹⁹” (CUMMINGS, 2013, p. 155). As *mixtapes* são resultado do desejo de consumidores de música em separar e reagrupar canções, ativamente selecionando e reordenando áudios, apesar da degradação resultante da cópia analógica em equipamentos caseiros.

Em 1999, o universitário Shawn Fanning reuniu algumas ferramentas já existentes criando um *software* para troca direta de arquivos na internet. O Napster verificava quais arquivos no formato *mp3* estavam disponíveis entre os usuários e estabelecia uma conexão direta entre eles, que descarregavam suas canções preferidas em seus computadores

19 Esse termo também se refere a cópias em fitas cassete de obras realizadas e distribuídas por artistas independentes.

(CASTRO, 2014, p. 137). Na virada do século XX, o formato *mp3* já era muito usado e se estabelecia como padrão de arquivo de áudio, graças à pirataria on-line (STERNE, 2012, p. 207). Essa troca de arquivos *mp3* renovou o hábito da escuta de músicas selecionadas, como faziam os que escolhiam a faixa do disco de vinil ou aceleravam a reprodução da fita cassete até a canção desejada.

Por não haver a necessidade de piratear pela internet um álbum completo, e nem existir um suporte físico que exige sua ordenação linear, os arquivos *mp3* podem ser acumulados sem uma organização específica. A *playlist* surge como uma sugestão de seleção e ordem na escuta desses arquivos. No contexto do fonograma virtualizado, ela é um arquivo de computador que contém a ordem de arquivos digitais de áudio a serem reproduzidos pelo dispositivo. Ela não possui fisicalidade, nem mesmo contém os arquivos, surgindo como uma versão virtualizada da *mixtape*. Uma outra modalidade de ordenação das músicas independe da ação ordenadora do usuário: o *shuffle* reproduz em sequência aleatória os arquivos de áudio presentes no dispositivo, proporcionando uma nova forma de encadear os fonogramas virtualizados (KATZ, 2010, pg. 189).

2.4.4. iPod

Em pouco tempo, a reprodução de arquivos digitais de áudio deixou de ser possível somente dentro do computador pessoal, sendo trasladada a outros aparelhos. Os primeiros modelos de reprodutores portáteis de arquivos digitais de áudio, quase protótipos funcionais, eram extremamente limitados. Um dos primeiros capazes de reproduzir arquivos no formato *mp3* foi o Rio, da Diamond Multimedia, lançado em 1998 (STERNE, 2012, p. 203), que possuía 32 megabytes de memória interna, o equivalente a aproximadamente 30 minutos de música no *bitrate* habitual da época de seu lançamento, 128 kbps. Mas isso foi suficiente para inaugurar as vendas de “tocadores de *mp3*”. O áudio digital estava virtualizado e podia ser copiado dentro de um dispositivo que não precisava conter fisicamente uma mídia, como um *compact-disc* ou uma fita cassete, tornando-se extremamente portátil.

A indústria fonográfica não absorveu imediatamente a ideia de que seu produto circulasse na internet, legalmente ou não. Sua resposta ao movimento de cópias ilegais - após tentar em vão conter a pirataria e a distribuição legal de fonogramas virtualizados - é criar um novo mercado (CUMMINGS, 2013). Em abril de 2003, a empresa norte-americana de tecnologia Apple, que já vendia seu reprodutor portátil iPod, lança sua plataforma de ven-

das de música, a iTunes Store (CHEN, 2010), simbolizando um novo momento e modelo de negócios da indústria fonográfica, com vendas *on-line*.

O fonograma virtualizado, inicialmente uma brincadeira ilegal entre estudantes em dormitórios de faculdades norte-americanas (STERNE, 2012, p. 205), era agora domado pelas grandes gravadoras e a indústria de tecnologia. Esse controle, porém, permitiu a comercialização do fonograma virtualizado como produto e, conseqüentemente, sua propagação por uma audiência mais ampla. Oferecendo álbuns completos e *singles*, a Apple vendeu um milhão de arquivos na primeira semana após o lançamento do iTunes Store, e 25 milhões apenas oito meses depois (HILL, 2013). Em 2008, a Apple tornou-se a maior vendedora de músicas dos Estados Unidos (APPLE, 2008).

Michael Bull, em seu artigo *No Dead Air! The iPod and the Culture of Mobile Listening* (2005), entende que a capacidade de armazenamento de um reprodutor portátil de arquivos digitais de áudio é fator disruptivo na liberdade do usuário na escolha das músicas que irão acompanhá-lo:

Com o surgimento da tecnologia *mp3*, é dado ao usuário em trânsito um acesso inigualável à sua coleção de músicas. Gerações anteriores de dispositivos móveis, apesar de prover portabilidade, limitavam o consumidor a poucas escolhas, devido ao seu formato, enquanto equipamentos como o iPod da Apple permite aos usuários armazenar até dez mil músicas. Esses arquivos podem ser organizados através de *playlists* em qualquer ordem desejada (BULL, 2005, p. 343).

Para ele, o iPod permite que o usuário exerça controle sobre sua escuta durante o transporte pela cidade, através da criação de uma “bolha privada de escuta” (BULL, 2012, p. 207).

O iPod²⁰ - ou qualquer outro dispositivo reprodutor de arquivos digitais de áudio - é essencialmente um “Walkman de arquivos digitais de áudio”, com capacidade de armazenamento inimaginável em décadas anteriores. Michael Bull considera que, em relação ao Walkman, o acesso a inúmeras músicas altera a maneira como o usuário interage na cidade durante seu trajeto. Ele está focado em situações específicas de escuta, típicas do dispositivo portátil: uma escuta de fones de ouvido, em lugares públicos, onde pode ou não haver o desejo de isolamento total de músicas que o usuário disponibilizou previamente para ouvir, encadeando-as manualmente ou através de *playlists* (BULL, 2005, p. 344).

20 Optamos por utilizar ao longo do texto o nome iPod para representar os dispositivos portáteis de reprodução de arquivos de áudio como descritos por Michael Bull, simples “tocadores de mp3” com capacidade de armazenamento até então inédita.

2.4.5. Smartphones

O surgimento dos primeiros reprodutores portáteis de *mp3* é uma faceta da miniaturização de componentes e dispositivos eletrônicos do final do século XX e começo dos anos 2000. Aparelhos relativamente simples, como agendas eletrônicas, começam a perder espaço para equipamentos mais sofisticados como *handhelds* (ou *palmtops*, computadores de dimensões reduzidas), com telas coloridas e sensíveis ao toque.

No mesmo período, os aparelhos celulares passam a fazer parte da vida cotidiana. Vários dispositivos acabam convergindo para o que hoje se entende por *smartphone*, aparelho que reúne funções de telefone celular, computador pessoal com conexão à internet e capacidade multimídia, com *softwares* para realizar funções especializadas, e *hardwares* como câmera, microfone e reprodução de áudio - seja por alto-falante, conector de fone de ouvido ou transmissor de curta distância (*bluetooth*).

Os *smartphones* são dispositivos capazes de, entre outras funções, descarregar dados de uma conexão, armazená-los, processá-los e reproduzir áudio para o usuário, utilizando uma interface bastante customizável de interação, a tela sensível ao toque.

Quando Michael Bull escreveu seu artigo *The Audio-Visual iPod*, em 2008, os *smartphones* ainda eram pouco desenvolvidos, e o iPhone, também da Apple, tinha sido lançado no ano anterior. No tocante à reprodução de arquivos digitais de áudio, eles ainda não apresentavam funções muito diferentes dos iPods descritos por Michael Bull.

Desde então, porém, o *smartphone*²¹ tornou-se um dispositivo com capacidades inéditas: trata-se essencialmente de um computador portátil, capaz de executar tarefas através de *softwares* denominados, por razões de *marketing*, de aplicativos, ou *apps*, baseados em seu sistema operacional, com grande capacidade de armazenamento e possibilidade de conexão contínua com a internet. A seguir, analisaremos brevemente essas características do *smartphone* e suas possíveis relações com mudanças nas situações de escuta privada no contexto urbano.

2.4.5.1. Armazenamento

Graças ao barateamento de chips de memória, os dispositivos portáteis de reprodução aumentaram exponencialmente sua capacidade de armazenamento nas primeiras décadas do século XX. Como citado anteriormente, o primeiro modelo de iPod foi lançado

21 Enquanto o termo iPod foi adotado por Michael Bull e utilizado igualmente nessa pesquisa, optamos por utilizar o nome genérico do dispositivo seguinte, *smartphone*, sem apontá-lo a uma marca determinada.

pela Apple em 2001, com memória de cinco gigabytes²² (SANTOS, 2014, p. 86), espaço suficiente para aproximadamente 40 discos em formato *mp3* de qualidade máxima - ou “1.000 músicas no seu bolso” (em qualidade não definida), como propagandeava a campanha publicitária e o website da empresa.

Dezesseis anos depois, em 2017, a Apple lançou o modelo de *smartphone* iPhone X, com até 256 gigabytes de memória, podendo guardar mais de dois mil álbuns de música, com a qualidade máxima do formato *mp3*.

Na venda de fonogramas virtualizados em plataformas como iTunes Store ou HDTracks, no lugar de adquirir uma mídia física, o consumidor compra o direito de descarregar o fonograma em seus dispositivos de reprodução. Essa modalidade de venda torna possível a comercialização de álbuns em escala menor do que a exigida para mídias físicas, já que os custos de replicação são nulos e os de transporte, estoque e lojas físicas muito reduzidos (CUMMINGS, 2017, p. 234).

Essa liberdade facilita a oferta, abrindo espaço para nichos de mercado, como o de alta resolução. Nele, a conveniência de formatos como o *mp3* ou o *aac* dá lugar a formatos que requerem mais espaço e utilizam resoluções e taxas de amostragem maiores que a do *compact-disc*²³, descrevendo a onda de áudio com maior precisão.

Os fonogramas digitais *hi-def* (do inglês alta definição) servem a um mercado de audiófilos, tocando em seus sistemas *hi-fi* arquivos que prometem ser melhores do que os disponibilizados em *compact-discs*.

No aspecto do armazenamento, o *smartphone* ampliou sua capacidade em relação ao iPod descrito por Michael Bull em mais de 50 vezes, potencialmente ampliando ainda mais a capacidade do usuário de carregar sua coleção de músicas, com a mesma qualidade ou em versão de alta resolução.

2.4.5.2. Conexão

Para Michael Bull, o principal fator disruptivo do iPod em relação aos dispositivos portáteis anteriores é a grande capacidade de armazenamento, libertando o usuário de ouvir sempre a mesma fita cassete e ampliando seu controle sobre seu trajeto na cidade

22 O equivalente à memória de aproximadamente sete compact-discs.

23 O padrão utilizado no compact-disc possui dois canais (estéreo) de 16 bits de resolução e taxa de amostragem de 44,1 KHz - podendo assim descrever 44.100 vezes por segundo a amplitude da onda de áudio, com 65.536 pontos de resolução. Mas alguns arquivos de alta resolução chegam a 24 bits por 192 KHz, o que utiliza 16.777.216 pontos para descrever a amplitude, 192 mil vezes a cada segundo.

(BULL, 2005, p. 344). O iPod que Bull analisa, no entanto, é um dispositivo sem conexão direta com a internet, que reproduz o que o usuário tiver colocado nele²⁴. Por maior que seja a memória de um dispositivo desses, seu conteúdo é limitado pelo universo finito e definido das músicas que o usuário já conhece e possui.

Com a difusão dos *smartphones*, os dispositivos portáteis de reprodução integram um sistema que tem cada vez mais suas funções atreladas à conexão com a internet. O transeunte apreciador de música trocou seu Walkman por um *smartphone* que cumpre, entre outras, a tarefa de tocar música. Os *smartphones* estão cada vez mais presentes no cotidiano das pessoas: segundo Don Reisinger (2017), a marca de um bilhão de usuários de *smartphones* com o sistema Android, da Google²⁵, foi atingida em 2014. Em 2017, por sua vez, a Apple contava com 700 milhões de usuários de iPhones vendidos.

Nesse contexto, aparecem os serviços de *streaming*, como Spotify, Deezer, Tidal e Apple Music. Neles, os fonogramas não são adquiridos, mas seu acesso e usufruto é permitido durante o período de assinatura do serviço - o usuário escuta os fonogramas virtualizados por fluxo de mídia, ou seja, sem armazená-los definitivamente em seus dispositivos.

O álbum *Thriller*, do músico Michael Jackson, é o mais vendido de toda a história da indústria fonográfica, com mais de 66 milhões de cópias (GUINNESS, 2017). Esse número, porém, é pequeno perto do número de inscritos em serviços de *streaming* de música, como o Spotify, com 157 milhões de usuários ativos em 2018 (IFPI, 2018, p. 48).

O *streaming* não se propõe apenas a reproduzir as músicas escolhidas pelo usuário, mas também a apresentar peças inéditas, ampliando o repertório do ouvinte. Dependendo do serviço contratado, o usuário pode escolher o que vai ouvir ou fica dependente da seleção feita pelo fornecedor do serviço. Nesse caso, *playlists* criadas por inteligência artificial, baseadas nos gostos do usuário (e também outras montadas por interesses comerciais), permitem que ele descubra músicas que ele não teria acesso se dependesse apenas de seu repertório de músicas adquiridas previamente - uma liberdade muito maior do que a que Michael Bull enxergou no iPod.

Quando o usuário utiliza esse tipo de *playlist*, porém, ele abre mão do controle sobre

24 Os modelos mais recentes de iPod, porém, são muito parecidos com os smartphones da Apple. Quando citamos o iPod no texto, estamos nos referindo a modelos similares ao descrito por Michael Bull.

25 A empresa norte-americana Google Inc. desenvolve o sistema operacional Android para uso em smartphones de diversos fabricantes, enquanto a Apple implementa seu sistema IOS apenas no próprio smartphone que comercializa, o iPhone.

quais músicas vai ouvir, para se deixar guiar por uma *playlist* que não criou. Mesmo que esse serviço seja personalizado, ele está optando por uma maior conveniência, e parcialmente abrindo mão do controle de sua escuta.

A melhora nas taxas de transferência de dados pela internet, tanto em conexão fixa (via cabo ou fibra óptica) quanto por telefonia celular, tem ampliado a capacidade dos serviços de *streaming* de áudio. Essa melhora, porém, ainda não se refletiu na transferência, em tempo real, de fonogramas virtualizados em resolução maior do que a de um *compact-disc*. Curiosamente alguns serviços, como o Deezer, anunciam serviços especiais de *streaming* com qualidade de *compact-disc* como sendo de “alta-fidelidade”. Nas explicações fornecidas, porém, nota-se que esses serviços prometem uma alta-fidelidade ao som do *compact-disc*, em uma tentativa de discurso *hi-fi* dentro de um serviço de alta-conveniência:

Quando uma música é compactada para *mp3* ou para o *streaming*, um algoritmo retira detalhes do som para gerar um arquivo menor. O *streaming* padrão e arquivos *mp3* utilizam 320 kbps e são suficientes para seus ouvidos entenderem a música, mas o serviço hi-fi transfere em 16 bits e 44,1 KHz em qualidade FLAC, para você ouvir todos os detalhes que existem nas fontes analógicas originais. É como usar óculos, mas para seus ouvidos (DEEZER, 2017).

2.4.5.3. Processamento

Em 1993, mesmo ano do lançamento do formato *mp3*, a Intel, principal fabricante de processadores para computadores pessoais, comercializou a primeira geração de seu processador Pentium, com capacidade de 60 MHz, que funcionava em um computador de mesa (DEFFREE, 2018).

Em 2001, o primeiro dispositivo portátil de reprodução da Apple, o iPod, era vendido com quase o dobro da capacidade de processamento²⁶. O *smartphone* iPhone XS MAX da Apple, lançado em 2018, possui um processador de aproximadamente 2,49 GHz²⁷, cerca de 41 vezes mais rápido do que o Pentium de 1993²⁸.

O poder de processamento dos *smartphones* atingiu, na década de 2010, capacidade para executar tarefas muito exigentes, como rodar aplicativos de jogos, edição fotográfica ou mesmo criação musical, com síntese e *sampleamento* de áudio. A própria decodifica-

26 A primeira geração do iPod utiliza dois processadores ARM 7TDMI, com 90 MHz cada (EVERYMAC, 2018a).

27 A Apple não divulga a velocidade de seus processadores atuais, mas alguns websites fazem cálculos estimativos baseados em performance (EVERYMAC, 2018b).

28 A comparação é apenas ilustrativa, dado que outras tecnologias presentes nos aparelhos atuais, como processadores de vídeo, na verdade ampliam essa diferença.

ção de formatos compactados (como o *mp3*) exige que os dispositivos façam cálculos matemáticos, algo corriqueiro para esses processadores. Assim, nos *smartphones*, é possível fazer alterações em tempo real nos áudios armazenados na memória interna ou recebidos por *streaming*.

As resoluções e taxas de amostragem do áudio digital utilizadas pelos consumidores comuns permanecem similares desde a sua padronização para o *compact-disc*. Já a capacidade de manipulação de áudio em tempo real aumentou devido ao avanço da tecnologia dos *smartphones*, que são verdadeiros computadores de bolso.

2.4.6. A escuta personalizada

Já era possível escolher músicas em um Discman (tocador portátil de *compact-disc*) ou trocar de fitas em um Walkman, mas os dispositivos portáteis de fonograma virtualizado do início do século XX, como iPods e *smartphones*, trazem uma outra perspectiva à ideia de personalização. Com um controle muito mais completo nas mãos, e a capacidade de escolher entre centenas de músicas armazenadas na memória (ou milhares delas nos serviços de *streaming*), esses aparelhos inauguram o que chamaremos de *escuta personalizada*.

A escuta privada, que saiu às ruas com o Walkman na década de 1980, existia anteriormente: o fone de ouvido era usado em sistemas de som domésticos dos anos 1960 e rádios amadores dos anos 1950. Mesmo antes da invenção do fone de ouvido, ao auscultar um paciente com o estetoscópio - classificado por Jonathan Sterne (2003) como um dos primeiros dispositivos de escuta mediada - o médico estava realizando uma escuta privada.

Para Shuhei Hosokawa, a importância do Walkman é que esse dispositivo seria a síntese da escuta móvel. Ele argumentava que, mesmo se novas tecnologias melhorassem o aparelho, elas seriam apenas secundárias:

Não se pode imaginar qualquer avanço que constitua mais do que exemplos de “progresso” secundário, como a invenção de aparatos mais leves, de maior fidelidade e menores, de fones de ouvido sem fio, de dispositivos à prova d’água [...]. O Walkman é, sem dúvida, o objeto definitivo para a escuta privada [...]
(HOSOKAWA, 2012, p. 106)

Hosokawa dá ênfase à portabilidade do dispositivo, que ele marca como diferencial essencial em relação aos equipamentos de reprodução anteriores. Aproximadamente duas décadas depois, o foco de Michael Bull não é mais a portabilidade, mas o poder dado ao usuário graças à liberdade de escolha musical do iPod e suas consequências no gerenciamento de seu trajeto pela cidade:

A criação de um mundo sonoro personalizado através do uso do iPod cria uma forma de solidão acompanhada para seus usuários, na qual eles se sentem capacitados, no controle, e autossuficientes enquanto atravessam a cidade. [...] A natureza móvel e contingente da jornada é vivenciada exatamente como seu oposto, precisamente pela criação e gerenciamento de seu próprio mundo sonoro. (BULL, 2005, p. 353)

Como comentado, o surgimento do *streaming* amplia questões sobre liberdade de escolha: se por um lado ele fornece acesso a músicas desconhecidas pelo usuário, a descoberta desse novo repertório exige o abandono do controle, valorizado por Michael Bull, ao deixar o algoritmo do serviço selecionar o que será reproduzido. Esse abrir mão do controle é ainda maior quando da escuta em modo *shuffle*, que estabelece uma ordem aleatória nas músicas disponíveis na memória de um dispositivo. A respeito o modo *shuffle*, Gisela Castro considera que,

[...] surpreendendo o ouvinte com encadeamentos e combinações pouco previsíveis ou usuais [...] a entrega do controle da sequência de faixas musicais que se vai escutar atende a uma crescente demanda do consumidor de música por conforto e conveniência (CASTRO, 2014, p. 164).

Assim como a escuta em modo *shuffle*, a escuta de *streaming* com descoberta de repertório implica em menor poder decisório, logo em menor controle. Ao permitir que o algoritmo do serviço de *streaming* proponha uma seleção de músicas, ficando apenas com a incumbência de pular as faixas indesejadas, o usuário tem à disposição uma forma de escuta com uma nova camada de conveniência: no modo *shuffle*, a ordem é escolhida pelo dispositivo, a partir das músicas que o usuário colocou na memória do dispositivo, e na escuta de *streaming*, a própria lista de músicas deixa de ser escolhida pelo usuário, abrindo possibilidades de novas descobertas musicais. Se o Walkman é a síntese de uma escuta móvel, o iPod e o *smartphone* apresentam-se como dispositivos para uma escuta personalizada.

A escuta personalizada, além de móvel e privada, passa por um amplo poder de escolha do usuário, como a seleção das músicas, através de uma *playlist* ou de um serviço de *streaming* - o que inclui a opção de escutar músicas que o usuário nem mesmo conhece, selecionadas para seu perfil. Essa escuta personalizada, porém, não leva em conta o contexto em que ela ocorre: um *smartphone* não altera a reprodução sonora em função da localização do usuário, de sua movimentação ou do ruído do ambiente onde ele se encontra, apesar de o dispositivo possuir sensores para todos esses fatores.

2.4.7. Situações de escuta

O Walkman inaugurou a escuta privada em espaços públicos, e o iPod proporciona uma escuta personalizada, baseada nas escolhas de repertório feitas pelo usuário. Esses novos hábitos de escuta móvel privada, descritos por Michael Bull (2000, 2005, 2012), são ampliados pelas possibilidades do *smartphone*. A incorporação do dispositivo portátil de reprodução de áudio por um aparelho onipresente na vida cotidiana leva a novas situações de escuta: o *smartphone* acompanha constantemente o usuário, que utiliza sua função de reproduzidor de música em diversos contextos.

Para o exercício de masterização que apresentaremos no terceiro capítulo, fizemos uma seleção de três situações de escuta privada na cidade, descritas a seguir. A partir dos três níveis de escuta identificados por Barry Truax (TRUAX, 2001, p. 163) - *listening-in-search*, *listening-in-readiness*, *background listening* -, serão descritas três situações de escuta privada em ambiente urbano, ou seja, voltadas apenas para a escuta musical, de fone de ouvido, na cidade. Elas não pretendem exaurir todas as possibilidades de escuta e nem são excludentes entre si, mas se distinguem pela disparidade de intenções na fruição musical, o que se refletirá em diferentes tratamentos de masterização no terceiro capítulo.

2.4.7.1. A escuta que mascara

Um usuário de dispositivo portátil de reprodução está no meio da multidão, ouvindo sons só a ele perceptíveis. Seu isolamento acústico cria uma barreira entre ele e a multidão, estabelecendo relações de presença e ausência, realidade e fantasia. Essa escuta, que substitui os sons externos por músicas, é explorada por Michael Bull (2012) em seu texto *The audio-visual iPod*. O dispositivo portátil de reprodução sonora é utilizado como um catalisador de uma experiência estética do usuário, que altera sua percepção com as escolhas musicais que faz, manipulando suas sensações e interpretações do espaço em que se encontra. “O mundo é arrastado para a narrativa individual do usuário, ao invés das ruas o atraírem para seu domínio” (BULL, 2012, p. 206).

Uma escuta privada como trilha sonora individual, segundo a pesquisa de Michael Bull, funciona melhor sem interferências. Ele entrevista usuários que preferem estar isolados dos sons externos, mascarando a rua com a música de seus fones de ouvido:

O mundo e a experiência do usuário ganham significado por meio de seu mundo sonoro, envolvente e privado. Os usuários do iPod, invariavelmente, preferem ouvir sua música em alto volume, gerando uma esmagadora sensação de presença, e ao mesmo tempo bloqueando qualquer som ambiente que possa manchar o intenso e fortalecedor prazer dessa experiência (BULL, 2012, p. 204).

Essa situação de escuta, que denominaremos aqui de *escuta que mascara*, é essencialmente um esforço de isolamento sonoro da rua, com as músicas escolhidas pelo usuário substituindo os sons da cidade.

2.4.7.2. A escuta de fruição

Na vida acelerada das grandes cidades, há pouco tempo para a apreciação musical. Como ilustra Jonathan Sterne (2012, p. 239), “ [...] o mais próximo que muitas pessoas chegam de ouvir música sem fazer mais nada é em trânsito, seja utilizando um reproduutor portátil, ou no carro”. Ao contrário do uso do iPod observado por Michael Bull, no qual o usuário mascara os sons urbanos com as músicas, existem situações de escuta no ambiente urbano nas quais ela acontece como atividade principal do usuário. É o caso de quando se ouve música à espera de um ônibus ou durante o trajeto do metrô. A esse tipo de situação, chamaremos de *escuta de fruição*.

Essa escuta tem como principal objetivo a apreciação musical: ela se situa em algum lugar entre as opostas *escuta analítica* e *escuta distraída*, como consideradas por Barry Truax (2001, p. 163). O som não é nem necessariamente destrinchado em busca de informações, nem está à mercê da atenção variável do ouvinte, ocupado com uma outra atividade. A *escuta de fruição*, como aqui propomos, lida com a escuta como fonte de prazer, independente da quantidade de detalhes que o ouvinte conseguirá distinguir - o que varia em função das interferências externas, que podem ocorrer sem portanto interromper essa fruição.

Na escuta que mascara, o isolamento é quase tão importante quanto o material ouvido; na de fruição não há uma fantasia de realidade dependente do isolamento, apenas um momento de prazer proporcionado pela escuta - uma descoberta musical, uma memória afetiva disparada pela escolha do repertório, ou simplesmente uma companhia familiar no inóspito ambiente urbano.

Ela pode ser comparada à leitura de livros ou jornais nos espaços urbanos. Como lembra Jonathan Sterne (2012, p. 236), “ [...] a leitura em trânsito é considerada uma maneira de administrar a alienação e a falta de ocupação [durante os trajetos pela cidade].” A escuta de fruição é a atividade principal, concentrada, até ser interrompida, exatamente como uma leitura. Textos muito complexos podem não ser adequados à leitura no meio do metrô. Uma tentativa de escuta estrutural, adorniana, talvez também não seja viável, ao contrário de uma escuta atenta, que pode ser interrompida ao chegar em seu destino.

2.4.7.3. A escuta distraída

Uma *escuta distraída* ocorre como atividade secundária, complementar às de fruição e a que mascara. O usuário de um dispositivo portátil de reprodução sonora, em seu trajeto pela cidade, pode variar entre essas três escutas, ao atravessar uma rua ou quando está chegando a seu destino e precisa desembarcar do ônibus, voltando sua atenção para o trajeto. Em algum momento, querendo escapar da realidade onde está, ativa uma escuta que mascara e se isola acusticamente do ambiente, fantasiando histórias para os outros transeuntes. Jonathan Sterne (2012, p. 225) trata essa escuta distraída como uma atividade secundária e simultânea a outras no contexto urbano:

Os que vivemos em ambientes industrializados desenvolvemos, a partir da onipresença da música em nossos cotidianos, um modo de escuta independente de características específicas da música. Neste modo [de escuta], ouvimos “junto a” (ou simultaneamente com) outras atividades.

Ela não exige um isolamento do ambiente, ao contrário: os sons da escuta privada podem se misturar com os da cidade, distraindo o usuário sem aliená-lo do contexto urbano, o que pode ser importante para pedestres e ciclistas, por exemplo. Por mais que isso desagrade a audiófilos como Bob Katz, essa é a escuta mais comum atualmente. “É um pensamento desalentador, mas hoje a maior parte da escuta musical deixou de ser uma experiência solene, de primeiro plano” (KATZ, 2015, p. 7).

2.5. Os desafios da masterização hoje

O sucesso comercial do Walkman, lançado em 1979, pode ser visto como a aceitação, por parte dos consumidores, de uma alternativa ao discurso da alta-fidelidade. Embora já existissem vários reprodutores portáteis como vitrolas e rádios de pilha, o Walkman simboliza, pelo seu uso voltado à mobilidade urbana, o começo da busca por um dispositivo de alta-conveniência.

Os dados oficiais da *Recording Industry Association of America* (RIAA - Associação da Indústria Fonográfica da América), cobrindo o período de 1973 a 2018, apontam como o Walkman influenciou a venda de fitas cassete nos Estados Unidos (RIAA, 2019). Apesar de serem comercializadas desde 1964 para reprodução em aparelhos caseiros e automóveis (CUMMINGS, 2013, p. 87), apenas em 1983 - três anos depois do lançamento do Walkman nos EUA - a venda de álbuns em fitas cassete superou a quantidade de álbuns comercializados em vinil - 236,5 milhões de unidades contra 209,6. A partir daí, por quase uma década, a fita cassete, utilizada no Walkman e também em sistemas caseiros e auto-

móveis, dominou o mercado como mídia mais vendida.

Esses números apontam que, na década de 1980, a alta-conveniência predominou sobre a alta-fidelidade - por mais que coexistisse um discurso *hi-fi* na venda dos dispositivos reprodutores de fita cassete. Alex Cummings resume o jogo de mudanças das mídias em uso: “Os anos oitenta testemunharam o auge da popularidade das fitas cassete, enquanto os álbuns em disco de vinil perdiam espaço e a indústria fonográfica promovia o novo *compact-disc*” (CUMMINGS, 2013, p. 168).

A popularização do *compact-disc*, com vendas fortemente baseadas em um discurso *hi-fi*, acontece apenas na década seguinte. Em 1992, os álbuns vendidos em *compact-disc* tomam a dianteira no mercado, vendendo 407,5 milhões de unidades contra 366,4 milhões de álbuns em fitas cassete.

O ano de 2000 registrou a maior venda de álbuns em *compact-disc* da história: 942,5 milhões, o que também representou o recorde anual de unidades vendidas em qualquer mídia física. As vendas de álbuns em formato de arquivo digital nunca foram tão volumosas, chegando a 118 milhões de unidades, em seu auge, em 2013. Em compensação, em 2012, foram vendidos 1,4 bilhão de músicas individuais (*singles*) em formato de arquivo digital de áudio (RIAA, 2019).

Esses números tratam exclusivamente do mercado norte-americano, mas ilustram o caminho da alta-conveniência: o sucesso do Walkman e seus similares na reprodução portátil de fitas cassete, o ressurgimento de uma mídia física *hi-fi* e a expansão de vendas de um formato que estava em baixa desde os anos 1960, o *single*, agora em versão digital e virtualizada. O *single* em arquivo digital pode ser visto como um formato alinhado com a ideia de alta-conveniência: o usuário adquire apenas uma música, sem precisar pagar por um álbum com várias faixas indesejadas, nem ter que se deslocar até uma loja física.

Uma outra maneira de consumir música é através de *streaming*. Os já citados 157 milhões de usuários ativos, apenas do serviço Spotify em 2018 (IFPI, 2018, p. 48), mostram a força do modelo de negócios, alugando o acesso às músicas, sem vendê-las em arquivo digital ou mídia física. Esse serviço apresenta novos significados para a alta-conveniência: o usuário não precisa de espaço físico ou de memória digital para guardar suas músicas, que são acessíveis a partir de diversos dispositivos, bastando para isso que disponha de uma conexão com a internet. Ele também não precisa escolher em que músicas investir seu dinheiro, já que, pagando pelo serviço, tem acesso às músicas que conhece e a outras

que o serviço oferece. Por outro lado, o usuário deixa de possuir a mídia física ou o arquivo digital, sendo obrigado a pagar pelo serviço enquanto quiser manter esse acesso.

Como comentamos anteriormente, a ação do usuário em selecionar músicas específicas para sua escuta já acontecia na escolha das faixas dos discos de vinil, passando pelas *mixtapes* de fita cassete e as *playlists* digitais. Nos discos de vinil e nas fitas cassetes, além do tempo de troca de mídia e navegação até a faixa desejada, não havia um problema sério de diferença de intensidade entre diferentes gravações em suporte analógico - uma *mixtape* de fontes analógicas não deveria apresentar disparidades de intensidade entre as faixas.

Isso, no entanto, não acontece no domínio digital. Bob Katz (2015) afirma que, no vinil, a “diferença entre o mais forte e o mais suave álbum de música *pop* não era maior do que 6 dB, o que é perfeitamente razoável” (KATZ, 2015, p. 243); já no domínio digital, a diferença passa a ser de até 20 dB. Como vimos no primeiro capítulo, a “guerra de volumes” ampliou a diferença de intensidades desde o surgimento do *compact-disc*, e, dependendo das músicas escolhidas pelo usuário - e da distância cronológica entre as masterizações dessas faixas - há uma disparidade sensível de *Program Loudness* entre as músicas. Ferramentas como o Sound Check tentam resolver essa disparidade entre as músicas com a normalização de intensidade, que acaba expondo os efeitos do excesso de compressão de extensão dinâmica das músicas. A normalização de intensidade elimina o que muitos artistas envolvidos com a música comercial viam como obrigação, ter um *Program Loudness* competitivo, para se voltar a outras dimensões estéticas de suas obras:

Com a normalização de intensidade, os artistas podem optar por fazer gravações com qualquer quantidade de compressão e recuperam uma maior capacidade de expressão. A normalização restaura a liberdade de escolha para produzir gravações não distorcidas e com microdinâmica (KATZ, 2015, p. 251).

A mudança de compra de álbuns para *singles* nas vendas digitais também trouxe preocupações novas quanto ao equilíbrio tonal das músicas. Antes, cada álbum era pensado como um conjunto, em que o ouvido tem tempo de se acostumar e entrar em sintonia com as diferentes propostas de cada música (KATZ, 2015, p. 47). Em tempos de *singles* e *playlists*, com músicas de diversas origens, no entanto, o (des)equilíbrio tonal das músicas serve como elemento de contraste entre elas, nem sempre de forma positiva.

Na fase de masterização, o equilíbrio tonal tem que ser pesado em função do gênero musical e da intenção da produção (KATZ, 2015, p. 56), que tenta se diferenciar dentro de uma sequência desconhecida de músicas. Os autores que são referência no tema, porém,

não fazem essa distinção estética de equilíbrio tonal entre masterização para suportes fixos, como *compact-disc*, ou para arquivos digitais ou *streaming*, claramente ignorando a diferença na escuta de álbuns completos ou *singles* sequenciados. Tanto Bob Katz quanto Bobby Owsinski limitam-se a falar tecnicamente da masterização para saídas digitais, sem tecer considerações a respeito da diferença entre ouvir um álbum completo ou *singles* variados. Homero Lotito, engenheiro de masterização brasileiro, chega a afirmar, em entrevista de 2009, que, para ele, “o conceito básico da masterização continua sendo o de criar algo que possa ser ouvido em sequência. Não considero a manipulação de uma música isoladamente como sendo uma masterização” (MARINO, 2009, p. 81). Para ele, a masterização remete ao equilíbrio e coesão entre as faixas de um álbum, e não na otimização do áudio para um suporte específico. Talvez o maior desafio atual da masterização seja exatamente aprender a olhar as situações de fruição do material criado, e pensar em como os processos de masterização podem ser adaptados às novas situações de escuta.

2.5.1. Hi-fi no *smartphone*?

A maior parte das situações de reprodução sonora ocorre fora de ambientes acusticamente tratados e sem equipamentos *hi-fi*. Apesar de possuírem limitações, elas podem transmitir o que realmente importa em uma música. Para David Byrne, a mensagem cultural transmitida pelas canções que ele ouvia no rádio era o que importava, e a qualidade da reprodução sonora não era essencial:

Eu ouvi pela primeira vez músicas de *rock*, *pop* e *soul* em um rádio transistorizado de baixa qualidade, e elas mudaram completamente minha vida. A qualidade do som era atroz, mas aquele som metálico estava transmitindo um monte de informações. Apesar de ter sido uma transmissão de áudio, a mensagem social e cultural embutida na música me eletrificou tanto quanto o som. Esses componentes extramusicaís carregados junto com a música não exigiam um sinal de alta resolução - bom o suficiente era bom o suficiente. Não estou dizendo que o som metálico deva ser considerado satisfatório ou desejável, ou que nunca devemos nos esforçar mais do que “bom o suficiente”, mas é incrível quanto pode ser transmitido com qualidade de áudio de baixa fidelidade (BYRNE, 2012, p. 126).

Para pensar sobre a escuta, é importante pensar *por que* alguém utiliza um dispositivo de reprodução, e não apenas na qualidade com a qual esse áudio está sendo reproduzido. Fernando Iazzetta discorre sobre a transição da importância da qualidade de áudio para a qualidade de escuta:

Músicas codificadas em *mp3* são escutadas em automóveis misturadas ao ruído urbano, em tocadores digitais enquanto se faz exercícios físicos ou pelos alto-falantes precários encontrados na maioria dos computadores pessoais, sem que

os ouvintes se incomodem com sua qualidade sonora. O que está em jogo não é mais uma qualidade de áudio, já que as tecnologias atuais parecem oferecer muito mais do que um ouvinte médio necessita, mas uma qualidade de escuta que está totalmente sujeita ao ambiente e aos contextos em que ocorre (IAZZETTA, 2009, p. 128).

Uma década depois, com o aumento da velocidade de conexão da internet e da capacidade de armazenamento dos dispositivos, é possível transferir e armazenar arquivos digitais de áudio de tamanho maior. Isso permite utilizar *bitrates* maiores e oferecer ao usuário uma qualidade de reprodução maior do que parecia ser suficiente em 2009. Mark Katz (2010) tenta enxergar, em seu livro *Capturing Sound*, os possíveis futuros para o que chama de *era da pós-fidelidade*, na qual são utilizados equipamentos aqui denominados de dispositivos de *alta-conveniência*:

A pós-fidelidade veio para ficar? Em uma visão de longo prazo, vejo três possibilidades. Uma é que a pós-fidelidade marca um ponto de virada, no qual o som gravado diverge de maneiras, cada vez mais pronunciadas, de seu modelo anterior, a música ao vivo. Outra possibilidade é que a era da pós-fidelidade seja temporária, e que novos avanços tragam uma nova era *hi-fi*. Uma terceira possibilidade é que os conceitos de fidelidade e conveniência percam seus significados distintivos e oposicionistas. Se a música nunca sair do nosso lado, o conceito de conveniência como um valor especial ficará sem sentido. Ao mesmo tempo, um som de alta qualidade (qualquer que seja sua definição) não exigirá nenhum esforço particular para ser obtido ou acessado, e tampouco será um conceito particularmente significativo (KATZ, 2010, p. 219).

Ele considera que se estava abrindo mão da fidelidade na reprodução, usando formatos de arquivo com perda de dados, como o *mp3*, ou fones de ouvido de pouca qualidade para ouvir música durante um traslado, em nome da conveniência. Porém, ao mesmo tempo, ele prevê a possibilidade de que a pós-fidelidade seja apenas um momento anterior aos avanços tecnológicos necessários para unir novamente a qualidade e a conveniência, anulando essa dicotomia.

O “bom o suficiente” de David Byrne tinha qualidade sonora ruim para seus próprios padrões, consciente de uma reprodução deficiente. Com as mudanças dos dispositivos de reprodução desde a juventude de David Byrne, nos anos 1970, talvez essa reprodução “boa o suficiente” já tenha chegado aos aparelhos portáteis, oferecendo “muito mais do que um ouvinte médio necessita”, como entende Fernando Iazzetta, e anulando a oposição citada por Mark Katz entre fidelidade e conveniência. Um elemento que contribui para o encontro (ou perda de significado) desses opostos é a melhora na tecnologia de compactação dos arquivos: o usuário comum, hoje, consegue realmente distinguir entre diferentes qualidades de reprodução?

A compactação para *mp3* (e outros formatos digitais de áudio com eliminação de dados) pode gerar artefatos sonoros perceptíveis - *bitrates*²⁹ mais altos diminuem esses efeitos, e até mesmo Bobby Owsinski afirma que um áudio compactado em *mp3*, com a maior qualidade possível, pode ser indistinguível de seu *compact-disc* de origem (OW-SINSKI, 2008, p. 94). Jonathan Sterne argumenta que as alterações sonoras realizadas pela compactação podem acabar incorporadas na cultura de escuta e serem consideradas parte das músicas, assim como aconteceu com a distorção harmônica e compressão de extensão dinâmica causada por válvulas (STERNE, 2012, p. 234).

Além dos problemas gerados pela compactação do áudio para sua transmissão eletrônica, alguns estudos questionam a capacidade do ouvinte comum em perceber diferentes taxas de compressão da extensão dinâmica. A compressão modifica características do áudio, e uma escuta atenta revela essas alterações. Mas isso alteraria de alguma forma a qualidade de escuta para o usuário comum?

Robert Taylor (2018, p. 132) resume em sua tese as conclusões de alguns dos estudos realizados nessa área. De seu levantamento, vale destacar as contribuições de Hjortkjær e Walther-Hansen (2014), que afirmam que um ouvinte comum não distingue entre grandes quantidades de compressão, e as de Taylor e Martens (2014), que concluem que as taxas aceitáveis de compressão variam em função do gênero musical. O próprio Taylor, porém, não considera as pesquisas totalmente conclusivas, não sendo ainda possível quantificar os limites de compressão de extensão dinâmica para que um ouvinte leigo seja capaz de perceber alterações na qualidade de reprodução. Estudos mais aprofundados - e voltados especificamente para situações de escuta portáteis - podem revelar surpresas ao senso comum sobre a real pertinência da compressão na qualidade da escuta em situações reais.

O discurso das comunidades audiófilas e dos engenheiros de masterização são bastante similares, criticando a compressão e tomando como óbvia a necessidade de gravação, masterização e reprodução completamente “transparentes”, reproduzindo o discurso da alta-fidelidade a uma fonte sonora. Um exemplo desse discurso pode ser encontrado em revistas especializadas, inclusive brasileiras. A revista *Audio Video Magazine*, especializada em testes de equipamentos audiófilos, entrevistou Homero Lotito, engenheiro de masterização com estúdio em São Paulo:

29 O bitrate de um arquivo *mp3* é a medida de fluxo de dados, em quilobites por segundo (kbps), e está diretamente relacionado ao tamanho final do arquivo, implicando também na qualidade do áudio armazenado.

Há uma tendência de aceitação maior de um som que não seja totalmente chapado. Ainda está muito distante daquilo que eu acho que deveria ser, mas há uma certa melhora. Certos clientes de música instrumental já valorizam este aspecto da dinâmica e a situação é maravilhosa. Algumas gravações saem com uma dinâmica que até alguns ouvintes audiófilos poderão se surpreender. A postura destes músicos é solicitar que a gravação retrate corretamente a forma como a música foi tocada por eles (MARINO, 2009, p. 83).

Vê-se claramente o uso de um discurso *hi-fi*, no qual o objetivo da reprodução é retratar fielmente o momento da captura sonora. Essa lógica, porém, não leva em conta a situação da escuta real. Ao priorizar um nicho específico de consumidores, distancia-se do ouvinte comum e ignora que a escuta móvel está cada vez mais presente na vida dos consumidores médios - como demonstram as vendas de arquivos digitais de áudio e assinaturas de serviços de *streaming*.

Bernie Grundman, um dos engenheiros entrevistados por Bobby Owsinski em seu livro, define a masterização como “[...] uma maneira de maximizar a música para fazê-la mais efetiva para o ouvinte” (OWSINSKI, 2008, p. 3). Pareceria natural, seguindo essa definição, que a masterização deveria conservar as dinâmicas da mixagem para uso em arquivamento e escutas audiófilas em salas especiais. Mas mesmo o audiófilo - e constante advogado das grandes extensões dinâmicas - Bob Katz atenta que “excesso de extensão dinâmica é tão ruim quanto sua falta” (KATZ, 2015, p. 267). Assim, cabe ao engenheiro de masterização o ajuste da extensão dinâmica para a que considerar adequada, não sendo necessária a exploração de toda a faixa disponível apenas porque a tecnologia permite. Voltando à definição de Grundman, portanto, o resultado final deve proporcionar a escuta mais efetiva possível para o ouvinte, o que inclui pensar na faixa de extensão dinâmica disponível na escuta - e não no suporte.

Para a escuta casual em situações com ruído ambiente, masterizações com muita extensão dinâmica podem causar alguns problemas. Ao argumentar sobre a dificuldade de uma escuta estrutural em ambiente urbano, Otávio Santos acaba descrevendo o mascaramento das passagens mais suaves como um empecilho para essa e outras escutas:

Consideremos o parâmetro musical da dinâmica: no caso de um ouvinte selecionar uma sinfonia cujo início seja em *pianissimo*, este não será capaz de seguir a música estando dentro de um ônibus em horário de rush. Ao aumentar o volume (intensidade) [da música], indiretamente [o ouvinte] transforma o *pianissimo* em *mezzoforte*. Poucos segundos após seu início, a sinfonia, então, apresenta uma cadência em *fortissimo*. O susto fará o ouvinte automaticamente reduzir o volume, e o *fortissimo* será transformado novamente em *mezzoforte*. Neste contexto, a estrutura e o sentido da obra foram completamente alterados, dificultando em graus extremos a prática de uma escuta estrutural (SANTOS, 2014, p. 26).

Na situação descrita acima, existe apenas uma pequena faixa de dinâmica disponível, devido ao ruído do trânsito em que o usuário se encontra, o que gera mascaramento dos sons mais suaves. Mas é possível criar variações de dinâmica a partir de masterizações que levem em conta essa situação de escuta, mantendo as partes mais suaves da música acima do ruído ambiente - simulando o que o usuário tenta realizar ao manipular o ganho com o botão de volume. A variação será menor do que a original, mas um material preparado para ser ouvido nessas situações extremas pode ser bastante mais adequado do que forçar o usuário a variar manualmente o ganho de seu dispositivo. Ouvir um material com uma extensão dinâmica menor do que a original parece mais vantajoso do que ouvi-lo apenas parcialmente. O exemplo é de variação de dinâmica entre movimentos de uma sinfonia, mas poderia também ser aplicado aos instrumentos mais suaves de um arranjo musical, que são mascarados pelo ambiente enquanto os mais fortes ainda são audíveis, como será exemplificado no próximo capítulo. Nesse caso, compressões de dinâmica específicas para determinadas faixas de frequência podem ser necessárias, entre outros ajustes de equilíbrio tonal.

Arquivos digitais de áudio com muita extensão dinâmica possuem um *Program Loudness* baixo, devido à necessidade de “espaço” para a extensão dinâmica operar - é necessário deixar *bits* disponíveis para descrever os picos mais fortes do material. Isso não é problema em situações de escuta com pouco ruído ambiente, bastando o ajuste do ganho desejado do amplificador utilizado na reprodução - a resolução do *compact-disc*, suficiente para descrever sinais com 96 dB de diferença, é teoricamente mais do que suficiente para cobrir a “distância” de 80 dB entre o *pianissimo* e o *fortissimo* de uma orquestra (POHLMANN, 2011, p. 4). A utilização de um áudio com muita extensão dinâmica em ambientes ruidosos, porém, gera dois problemas. Por um lado, se a amplificação necessária é muito maior do que a projetada para o dispositivo, seu ganho máximo pode não ser suficiente para atingir a intensidade adequada. Por outro, se a extensão dinâmica for muito grande, e o ganho estiver ajustado para as passagens suaves, as passagens mais fortes do áudio podem ficar excessivamente altas, levando à manipulação do botão de volume descrita por Otávio Santos.

As situações de escuta privada possibilitadas pelos dispositivos portáteis parecem opostas às condições laboratoriais sugeridas pelo discurso *hi-fi*. No lugar de equipamentos que prezam pela fidelidade de reprodução, são utilizados dispositivos que prezam

pela conveniência. No lugar de ambientes silenciosos que proporcionam concentração e atenção para a fruição musical, a agitação da cidade, com seus sons e distrações, coloca a escuta como atividade secundária frente aos estímulos e tarefas da cidade e do trabalho.

Levando em consideração as situações de escuta que eles proporcionam, não parece adequado pensar em uma reprodução de alta-fidelidade para os dispositivos portáteis proporcionam. Desde as campanhas publicitárias do Walkman os fabricantes dos dispositivos portáteis propõem essa escuta de alta-conveniência, com a música acompanhando outras atividades. Por outro lado, os engenheiros de masterização, como já verificamos, parecem ainda tender a finalizar os arquivos digitais de áudio buscando otimizar a experiência do ouvinte para condições idealizadas de escuta.

Como analisaremos a seguir, muitos elementos da escuta mediada têm origem em culturas de escuta e suportes de fixação de áudio que não correspondem à realidade da escuta privada e personalizada proporcionada pela presença dos *smartphones* na vida cotidiana. Entre esses elementos que não refletem as situações de escuta atuais está a própria masterização.

2.5.2. Problematizações sobre os fonogramas virtualizados

Tocar é o verbo utilizado pela indústria fonográfica em seus primórdios para tentar transferir a ideia de ação no uso de um instrumento musical para o manuseio de um fonógrafo (IAZZETTA, 2009, p. 33). O termo *faixa*, associado à música, vem da mudança de densidade dos sulcos em um disco de vinil, indicando visualmente a posição de transição entre as peças. *Capa* é outra terminologia proveniente da era dos discos, quando a proteção física do vinil servia também de superfície de exposição de imagens e informações. Além desse léxico, que foi apropriado das tecnologias que os antecederam, os fonogramas virtualizados e a cultura em torno de seu uso carregam consigo outras idiosincrasias de momentos anteriores da indústria fonográfica.

O limite de duração de um áudio depende da capacidade do suporte em que ele está fixado. Para dispositivos como o iPod, o limite é o tamanho que o áudio ocupa na memória³⁰. No entanto, o tempo de uma canção popular mantém-se similar há quase um século: cerca de três minutos (BYRNE, 2012, p. 13). Apesar da necessidade da indústria do *jukebox*

30 Em um cálculo aproximado, um arquivo de mp3 em 128 kbps ocupa 56.3 MB para armazenar uma hora de áudio. Nessa taxa, um iPod com 32Gb de capacidade poderia armazenar um só fonograma com mais de 24 dias de duração. Vale observar que essa capacidade também depende do tamanho máximo permitido pelo sistema de arquivos do dispositivo.

ter quase desaparecido há tempos (como comentamos anteriormente, ela utilizava discos de 78 rotações), uma cultura de música popular com essa duração parece ter se estabelecido e perdurado. Como sugere David Byrne, “talvez [...] todos tenhamos internalizado esse aspecto arbitrário da música gravada, e agora achamos as exceções estranhas e inusitadas” (BYRNE, 2012, p. 92).

O formato *álbum* também faz parte das heranças do vinil, não apenas pelo próprio formato de conjunto de canções, como por sua duração similar. Os *compact-discs*, que poderiam armazenar até 70 minutos, costumavam ter sua duração subutilizada na música comercial. Da mesma maneira, não há nada que impeça obras musicais de serem distribuídas pela internet com uma duração muito diferente, o que parece não acontecer com frequência. A própria comercialização de arquivos digitais de áudio, que vem sendo substituída pela licença de acesso às músicas em um servidor via *streaming*, parece um reflexo de um momento anterior, quando os principais dispositivos de reprodução dos usuários não estavam praticamente o tempo todo conectados à internet.

Esses elementos são parte intrínseca da cultura de fruição de música mediada, desenvolvida ao longo do século XX. Mas assim como os hábitos de escuta e as tecnologias foram sendo modificados - o rádio invadindo a sala de estar no passado ou o incremento da venda de *singles* em relação à venda de álbuns atualmente -, partes dessa cultura arrastam-se como sobras atávicas de tecnologias anteriores.

Ao longo da história da música mediada, os processos de masterização vêm sendo atualizados aos novos formatos que surgem. O então inédito patamar de nível de ruído e a ampla capacidade de reprodução de frequências trazidas pelo *compact-disc*, por exemplo, foram sendo absorvidos conforme esse suporte era popularizado. Algumas mudanças nas práticas de masterização para uma realidade de fonogramas virtualizados são encontradas nos textos dos autores estudados.

Bobby Owsinski, por exemplo, adicionou, na quarta edição de seu livro (2017), um capítulo que trata exclusivamente de masterização para iTunes. Ele descreve a proposta da Apple como um conjunto de normas para que os arquivos digitais de áudio tenham alta qualidade. Em síntese, ele explica que a Apple exige arquivos de alta resolução para oferecer o selo de qualidade *Mastered for iTunes* nas músicas disponíveis em sua loja virtual. Os procedimentos solicitados garantem uma melhor performance do codificador utilizado pela Apple, o formato *aac*, e permite que a Apple, armazenando em seus servidores a

versão enviada em alta resolução, possa atualizar os arquivos compactados que ela vende quando atualizar seu algoritmo de codificação, controlando assim a qualidade final dos arquivos (APPLE, 2012, p. 3).

Em sua documentação sobre *Mastered for iTunes*, a Apple ilustra como as músicas vendidas e distribuídas por ela serão utilizadas em diversas situações:

Por ser um formato altamente portátil, os arquivos vendidos no iTunes podem ser ouvidos em um amplo espectro de situações. Assim, um ouvinte pode usar pequenos earbuds dentro de um vagão de metrô, enquanto outro pode ouvir atentamente uma cantata de Bach em seu sistema de som *[hi-fi]*. Da mesma forma, um estudante universitário pode ouvir Miles Davis usando fones de ouvido *[hi-fi]* na biblioteca do campus (APPLE, 2012, p. 4).

Também é possível perceber nesse texto a oposição entre os equipamentos tidos como de alta-fidelidade e os de alta-conveniência, explicitando três situações de uso. Mas, ao aconselhar sobre o uso de compressão na masterização, Bob Katz demonstra o quanto a ideia de uma masterização única, que responda a situações díspares de escuta, pode gerar problemas:

Sem compressão, é extremamente difícil para o engenheiro de masterização atender ao mesmo tempo às necessidades de ouvintes críticos, sentados em salas acústicas, e de ouvintes casuais, praticando corrida ou dirigindo em um carro barulhento. É nosso dever satisfazer o produtor musical e as necessidades dos ouvintes, por isso devemos continuar a usar a quantidade de compressão necessária para fazer uma gravação soar bem em casa. Mas tente evitar usar mais compressão do que a necessária para a escuta em sistemas domésticos; [...] Se comprometimentos forem necessários para a escuta casual ou no carro, tente usar técnicas de sons transparentes para aumentar as passagens suaves, como a compressão paralela. (KATZ, 2015, p. 100)

Sua recomendação em voltar a masterização para a escuta doméstica explicita qual situação de escuta ele prioriza em seu trabalho. A própria escuta em fones de ouvido não parece ser essencial: ele acredita que o desafio da masterização é “produzir material que seja bem reproduzido em diferentes tipos de suporte e diferentes ambientes de escuta” (KATZ, 2015, p. 12). O próprio termo utilizado na frase em inglês, *listening environments* (ambientes de escuta), parece pressupor uma escuta mediada por alto-falantes, ou seja, dentro de um ambiente - ao contrário do que aconteceria em *listening situations* (situações de escuta), que também incluiriam as escutas mediadas por fones, nas quais o som é reproduzido no ouvido do usuário, e não no ambiente.

Bob Katz atenta-se às novas tecnologias dos dispositivos, como o Sound Check, mas não em seus usos, não se preocupando com as mudanças nas situações de escuta, ainda

que reconheça que elas existam (KATZ, 2015, p. 7). Owsinski tampouco tangencia o assunto, considerando “apenas útil” saber como a masterização soa em um fone de ouvido (OWSINSKI, 2017). Os autores ignoram, portanto, adequações de masterização que poderiam ser feitas exclusivamente para fones de ouvido, como se esse dispositivo de escuta fosse exótico ou raro. Ironicamente, esses mesmos dois autores comentam longamente sobre masterização para escuta em múltiplos alto-falantes (*surround*), considerada, por eles próprios, como um nicho de mercado³¹.

De fato, se as situações de escuta foram multiplicadas desde que a Sony miniaturizou o reproduutor de fita cassete, lançando o Walkman como novo dispositivo de apreciação musical (HOSOKAWA, 2012, p. 104), a masterização parece ainda voltada para um suporte (mesmo que seja um arquivo digital de áudio), e não para situações de escuta. Um exemplo disso é o depoimento concedido por Doug Sax a Bobby Owsinski, que estava realizando seis masterizações diferentes para um mesmo trabalho. Elas eram dedicadas, porém, a tipos diferentes de mídia e não a situações de escuta distintas (OWSINSKI, 2017). Além de raras exceções para nichos específicos, como o audiófilo³², cada masterização ainda é destinada a ser um produto diferente, isto é, vendido separadamente, e não a ser uma adaptação do áudio para diferentes situações de escuta.

Os engenheiros de masterização já criam versões diferentes dos fonogramas para usos e suportes distintos. Versões de músicas para a transmissão em rádio (*radio mix*), por exemplo, sempre levaram em consideração os processamentos pelos quais o áudio passaria durante sua transmissão - Bob Katz analisa longamente esses processamentos, fazendo recomendações específicas, como evitar a hipercompressão (KATZ, 2015, p. 356). A fixação dos fonogramas em suportes como discos de vinil, fitas cassete ou *compact-discs*, por sua vez, também demanda diferentes masterizações.

Observa-se, no entanto, que a comercialização de arquivos digitais de áudio frequentemente utiliza a mesma versão do *compact-disc* (como demonstrado no Capítulo 1, com o caso do álbum *Tales from Topographic Oceans*). Isso ocorre inclusive na iniciativa

31 Utilizando um mecanismo de pesquisa na versão digital de seu livro, Bob Katz utiliza mais de 200 vezes o termo *surround* e apenas 20 o termo *headphone* (e uma o termo *earbud*). Na edição de 2008 do livro de Bobby Owsinski são encontradas 167 vezes a palavra *surround* e apenas duas vezes a palavra *headphone*.

32 Algumas edições especiais de álbuns para fãs contêm versões das mesmas músicas em *surround* e em estéreo. É o caso das versões disponíveis na edição especial de *Tales from Topographic Oceans* de 2016, algumas delas analisadas no Capítulo 1, que não apresenta diferentes masterizações, mas sim distintas mixagens e gravações alternativas.

Mastered for iTunes, que curiosamente solicita masterização específica para fonogramas virtualizados, mas oficialmente verifica apenas a resolução do arquivo - e não sua masterização - para fornecer o selo de qualidade (APPLE, 2012). Também se verifica esse problema em *sites* como HDTracks e serviços de *streaming*, como o Spotify.

Áudios masterizados para *compact-disc*, principalmente no auge da “guerra de volumes”, podem possuir níveis de compressão de extensão dinâmica inadequados para os compactadores *lossy*. Os algoritmos dos compactadores têm dificuldade de distinguir os sinais que poderiam ser eliminados do material sonoro, descartando, assim, partes que deveriam ser mantidas, gerando uma perda de qualidade descontrolada. Surgem, mesmo nas configurações mais conservadoras dos algoritmos, artefatos sonoros não-musicais, típicos das configurações de baixa qualidade (TAYLOR, 2018, p. 122).

Masterizações específicas para esses formatos são recomendadas pelos autores estudados. As gravadoras, porém, nem sempre se preocupam com isso. O próprio álbum *Tales from Topographic Oceans* é um exemplo. Segundo o site do grupo Yes, as vendas pela internet utilizam a versão de 2003 - inclusive as de alta resolução (YES, 2019) -, que possui a maior compressão de extensão dinâmica e *True Peak* acima de zero dBTP, o que pode gerar os problemas de compactação e distorções, como foi descrito no Capítulo 1 e será demonstrado no Capítulo 3.

Como vimos, as práticas de masterização parecem ainda estar se adaptando a uma realidade de fruição musical sem mídia física. Iniciativas como a *Mastered for iTunes* tentam resolver problemas ocasionados pela compactação de áudio com dinâmicas excessivamente comprimidas, mas eles continuam acontecendo. As práticas de masterização ainda precisam ser revistas para essas plataformas: “*Downloads* e *streaming* significam um aumento do uso de formatos compactados de áudio como o *mp3*, o que gera comprometimentos, o que nos obriga a realizar ajustes em nossas práticas [de masterização]” (KATZ, 2015, p. 5).

Hugh Robjohns, em sua reportagem sobre o fim da “guerra de volumes”, vislumbra um futuro próximo, em que se disponibilizarão versões para diferentes situação de escuta, com normalização de intensidade fornecida pelos dispositivos de arquivos digitais de áudio:

Eu suspeito que um modelo similar [ao das masterizações para rádio e vinil] acontecerá na indústria da masterização daqui a alguns anos, à medida que a normalização de intensidade for se tornando disseminada, talvez com versões masterizadas separadas, otimizadas para dispositivos com normalização de intensidade [*loudness-normalization*] e para dispositivos com normalização por picos máximos [*peak-normalization*]. Também posso prever alguns artistas e

selos revisitando seu material de catálogo [álbuns antigos], e remasterizando-o para soar melhor em sistemas com normalização de intensidade (ROBJOHNS, 2014, p. 175).

Atualmente, o usuário que adquire um álbum ou um *single* em uma plataforma de vendas *on-line* tem acesso a apenas uma masterização. Se antes as mídias físicas impunham algumas situações de escuta, agora o mesmo fonograma virtualizado circula por diferentes dispositivos de reprodução. O usuário possui uma mesma versão para utilizar em seu sistema *hi-fi* doméstico e em seu *smartphone* com fones de ouvido durante trajetos urbanos.

A reprodução sonora em dispositivos portáteis não apenas melhorou de qualidade, como se tornou uma das mais comuns formas de fruição musical (KATZ, 2015, p. 7). Além do fornecimento de masterizações separadas para dispositivos diferentes, como sugere Hugh Robjohns, parece-nos importante pensar em masterizações diferentes para diferentes situações de escuta. Fornecer múltiplos arquivos digitais de áudio é uma solução simples para evitar os comprometimentos sonoros que se fazem necessários ao se otimizar um áudio para várias situações díspares. Uma outra parte da solução, proporcionada pela capacidade tecnológica dos dispositivos portáteis atuais, é processar o áudio em tempo real, adequando-o às situações do ouvinte.

Exploraremos na sequência algumas dessas possibilidades de variações nas masterizações, antes da realização dos exercícios que colocam em prática a manipulação do áudio para uma *masterização contextual*.

2.5.3. Por uma masterização contextual

A segunda edição do livro de Bob Katz (2007) foi escrita antes do surgimento do *streaming* e de tecnologias de normalização de intensidade³³. Nele, o autor retrata o problema da masterização para *compact-disc*, ao tentar contemplar todas as situações de escuta:

Há, é claro, ocasiões específicas em que é necessária muita compressão: música ambiente, festas, bares e *jukebox*, aparelhos em automóveis, corredores usando fones de ouvido, alto-falantes nas lojas de discos [...], e assim por diante. Em cada um desses casos, deveria ser possível produzir um *compact-disc* com compressão de extensão dinâmica sob medida para esse fim, ou instalar um compressor na *jukebox* ou no equipamento de reprodução. Certamente isso é muito menos prejudicial do que comprometer a qualidade da música para todos os ouvintes. O que desejamos é um substituto de baixa fidelidade para a fita cassete [o *compact-disc* com compressão]. Ironicamente, o *compact-disc* transformou-se em seu pior inimigo, pois ele não consegue ter diferentes qualidades para diferentes necessidades (KATZ, 2007, p. 188).

33 A terceira edição, de 2015, foi atualizada com as tecnologias mais recentes. O capítulo de onde essa citação foi retirada foi totalmente reescrito.

O fonograma virtualizado, de uso ainda mais flexível do que o *compact-disc*, parece ter ampliado o problema. A fisicalidade dos suportes determinam limitações em seu manuseio e induzem a algumas práticas de escuta - não é muito simples correr no parque carregando um toca discos em funcionamento, nem muito prático levar centenas de fitas cassete para decidir o que ouvir dentro de um vagão em movimento. Já um iPod pré-carregado com as músicas do usuário é facilmente adaptável à maioria das práticas de escuta mediada existentes, bastando alternar seu terminal de saída para a necessidade do momento - fones de ouvido ou alto-falantes amplificados. Um *smartphone* com serviço de *streaming* permite acesso a um repertório ainda mais amplo, bastando uma conexão com a internet.

Essa capacidade de adaptação a diversas circunstâncias permite que esses dispositivos portáteis de reprodução de áudio possam ser utilizados tanto em situações de escuta atenta, audiófilas, quanto em situações de escuta móvel privada. A utilização de um mesmo arquivo digital de áudio para todas as situações parece antiquado, uma sobra atávica da comercialização de mídias físicas, quando um fonograma tinha a prática de escuta limitada por sua materialidade. Da mesma forma que os engenheiros de masterização criavam versões específicas para a transmissão radiofônica, para o disco de vinil e para a fita cassete, o fornecimento de masterizações adequadas para situações de escuta díspares parece uma solução que evita os comprometimentos sonoros de uma masterização “universal”, que funciona (mal) em todas as circunstâncias.

IPods têm ampla capacidade de armazenamento, e *smartphones* podem receber essas diferentes versões por *streaming*, sem nem ao menos precisar mantê-las na memória, por isso fornecer múltiplas masterizações convenientes a cada situação de escuta não parece ser um problema.

Por outro lado, mesmo com essa ampla capacidade de armazenamento, naturalmente, é inviável contemplar individualmente cada situação de escuta com uma masterização específica. Nos *smartphones* existe a possibilidade de processamento do áudio em tempo real, permitindo que algumas situações de escuta sejam ajustadas automaticamente ou configuradas pelo usuário. Dessa forma, não é preciso realizar versões sob medida, como sugeriu Bob Katz - o *smartphone* poderia possuir algoritmos de processamentos para reproduções com menos compressão, hipercompressão, ou variações intermediárias. Outra alternativa seria alternar entre as versões desejadas de masterização no *streaming*.

Uma masterização voltada para situações específicas de escuta, levando em conta

tanto o ambiente quanto o desejo de escuta do usuário, é denominada nesta pesquisa de *masterização contextual*. Naturalmente, a masterização contextual pode ser pensada para todas as situações de reprodução sonora. O recorte desta pesquisa, porém, limita-se à análise da escuta móvel privada, não tratando, por exemplo, da interação acústica entre reprodutores domésticos e o ambiente em que se encontram, ou da escuta coletiva, como festas dançantes e música ambiente para lojas.

Já existem algumas tecnologias para tentar otimizar a escuta em dispositivos portáteis. Onipresente nos dispositivos reprodutores, o controlador de áudio mais comum é o botão de volume, que regula o ganho de amplificação do áudio. Botões atenuadores de ganho e reforçadores de graves existem nos reprodutores automotivos há bastante tempo. Em entrevista para Nick Southall (2013), Bob Katz via no processamento em tempo real a solução para reproduzir a música com menor extensão dinâmica e, assim, adequar o áudio à situação de escuta:

Nick Southall: Será que algo disso [a qualidade do arquivo de áudio] faz muita diferença se as pessoas estiverem ouvindo em configurações não ótimas, em sua mesa no escritório, no carro, com fones de ouvido em um trem, etc?

Bob Katz: Claro que, quanto pior o ambiente de reprodução, piores os problemas de som. Passagens suaves em locais barulhentos não são um problema resolvido pelo iTunes Radio ou pelo Sound Check. Alguns dos outros serviços de *streaming* usam *limiters* e compressores, e eu não apoio a redução da qualidade de áudio para o menor denominador comum durante a transmissão. A solução tem que acontecer no reprodutor: no carro, na pista de corrida, na academia, etc. E isso vai acontecer. Em um lugar barulhento, a compressão no reprodutor seria ligada para que você não precise controlar o volume durante passagens suaves. A Apple tem a capacidade de, por exemplo, usar o microfone embutido no iPhone como um sensor de ruído ambiente (SOUTHALL, 2013).

Como podemos perceber, o problema dele não é com a compressão em si; é com a oferta de uma masterização não otimizada para a melhor experiência de escuta, sempre partindo do princípio de que o consumidor deve ter acesso a apenas uma masterização. Poucos anos depois, na terceira edição de seu livro, Bob Katz atualiza sua fala, com propostas de futuras implementações para a adequação do áudio à situação de escuta, vislumbrando soluções no dispositivo para situações de escuta que demandam compressão:

Estamos no começo da normalização de intensidade. Em algum momento no futuro, esperamos que o dispositivo de destino (o rádio, o computador, o tocador de música portátil) seja inteligente o suficiente para aumentar passagens suaves em um local barulhento (como uma academia ou seu carro), para que os produtores possam criar material com uma ampla faixa de dinâmica, que possa ser alterada por tecnologias inteligentes em locais que exijam uma faixa de dinâmica menor. O Bose AutoPilot já é uma tecnologia muito eficiente de

controle de ganho, que não apresenta artefatos evidentes de compressão e pode aumentar o ganho em um carro quando ele está barulhento, mas evita que o som fique alto demais quando ele para em um semáforo, por exemplo. Espero que esse tipo de tecnologia migre para os reprodutores portáteis, pois isso marcará o fim dos produtores exigindo que as masterizações sejam feitas pensando em motoristas e corredores (KATZ, 2015, p. 222).

Um problema comum da escuta em dispositivos portáteis é a qualidade de reprodução dos fones de ouvido utilizados e sua diversidade. Eles possuem grande variação de resposta de frequências, principalmente nas mais graves e, dependendo do tipo de encaixe na orelha do usuário, possuem níveis diferentes de isolamento do ruído ambiente. Assim, uma masterização contextual pode levar em conta o fone de ouvido utilizado, ou permitir ao usuário sua configuração para ajustes em tempo real, com compensação de frequências para escutas mais atentas, ou considerações de isolamento para taxas de compressão de extensão dinâmica.

Como Bob Katz lembra na entrevista a Nick Southall (2013), as tecnologias de normalização de intensidade, já presentes nos reprodutores, não resolvem o problema de redução de faixa dinâmica ambiente, devido ao ruído urbano. Soluções que utilizam variações entre masterizações com maior ou menor compressão parecem simples, e podem ser implementadas como um botão de volume, alternando entre as duas versões.

A implementação de tecnologias de masterização contextual permitiria aos profissionais voltarem-se a decisões artísticas, sabendo que, em condições não ideais de escuta, seus arquivos de áudio serão adequados pelo *smartphone*, utilizado agora não mais somente como um dispositivo de reprodução, mas como um dispositivo de otimização da escuta.

3. Exercícios de masterização

Os serviços de *streaming* e as vendas de arquivos digitais de áudio são, no final da década de 2010, grande parte das trocas comerciais da indústria fonográfica³⁴ (RIAA, 2019). Os formatos podem ter mudado, mas a indústria continua fornecendo masterizações baseadas principalmente no consumo de mídia física, como verificamos nos dois primeiros capítulos. A cultura da escuta móvel privada, porém, privilegia a qualidade da escuta sobre a qualidade de reprodução (IAZZETTA, 2009, p. 128), opondo-se à ideia do discurso *hi-fi*, isto é, de melhora tecnológica a cada geração de equipamentos. Uma situação de escuta móvel privada apresenta contingências diferentes de uma escuta *hi-fi*: fones de ouvido pequenos, arquivos digitais de áudio *lossy* e ambientes ruidosos estão longe de configurar a situação idealizada pregada pelo discurso de alta-fidelidade.

Para ilustrar algumas das discussões dos capítulos anteriores e aliar o conhecimento prático em masterização com a análise das situações de escuta, realizamos neste capítulo alguns exercícios de masterização³⁵, descritos nas páginas a seguir. Eles buscam demonstrar os problemas expostos anteriormente na pesquisa e apresentar propostas do que definimos como masterização contextual. Seleccionamos fonogramas virtualizados, procurando diversidades de gênero musical, época da gravação e, principalmente, de aspectos de dinâmica das músicas, isto é, nas relações entre *Program Loudness* (PL, intensidade do programa, medida em LUFS) e *Loudness Range*³⁶ (LRA - faixa de extensão de intensidade, medida em dB). Após a descrição das faixas selecionadas, explicamos o procedimento geral dos exercícios e os resultados obtidos em cada um deles.

A masterização inclui vários aspectos não contemplados nesta pesquisa. A transformação de uma coleção de músicas em um álbum - primeira definição de masterização para Bobby Owsinski (2008, p. 3) -, por exemplo, não é abordada nesta análise. Sua segunda definição, porém, voltada ao ajuste fino de nível e equilíbrio tonal no preparo das músicas para a distribuição (OWSINSKI, 2017), carrega consigo a preocupação com ajustes

34 Podemos incluir nessa afirmação alguns outros modelos de fruição musical não abordados diretamente na pesquisa, como vídeos no youtube e rádios on-line.

35 Como os fonogramas virtualizados utilizados já foram masterizados antes de serem disponibilizados, tecnicamente o que estamos realizando aqui são exercícios de remasterização. O termo escolhido, masterização contextual, inclui nele a ideia de que, no futuro, a masterização poderia ser finalizada no dispositivo, com os serviços de streaming e venda de fonogramas fornecendo arquivos prontos para serem processados em tempo real.

36 Essas e todas as outras medidas utilizadas nos exercícios foram descritas no primeiro capítulo desta dissertação.

que podem ser processados nos dispositivos (ou fornecidos por streaming em configurações personalizadas), o que faz parte do escopo desta pesquisa.

Realizamos nove exercícios, focados nas alterações que um dispositivo portátil pode realizar para personalizar uma escuta móvel privada. Os primeiros três exercícios são dedicados à manipulação da dinâmica: o primeiro ilustra a normalização de intensidade realizada por ferramentas como o Sound Check; o seguinte, a alteração da dinâmica com compressores de diferentes características; e o terceiro é um exemplo de mascaramento com ruído urbano e de uso de compressão para diminuir a extensão de dinâmica, mitigando o problema.

Os três exercícios seguintes exemplificam: processamentos de espacialização de áudio para fones de ouvido; equilíbrio tonal com equalização para otimização de performance do equipamento utilizado; e comparações de qualidade entre diferentes *bitrates* do formato digital de áudio *lossy* mais conhecido, o mp3.

Três masterizações contextuais são realizadas ao final do capítulo, visando combinar esses processamentos com as situações de escuta levantadas no Capítulo 2 (escuta que mascara, escuta de fruição e escuta distraída), demonstrando suas possibilidades.

Como no estudo de caso do Capítulo 1, os resultados estão disponíveis no DVD anexo³⁷ desta dissertação, que contém duas versões de todas as faixas, em formatos diferentes: os arquivos em resolução e taxa de amostragem altas, como utilizados nos exercícios, e uma versão no formato aac da Apple (com extensão de arquivo *.m4p*), compactado em sua qualidade máxima, a partir dos arquivos *hi-res* 24 bits / 96 KHz.

O formato compactado foi escolhido por converter diretamente áudios de formatos *hi-res* para formatos de consumo, a fim de facilitar sua execução em equipamentos comuns de reprodução de áudio, inclusive portáteis, e porque é considerado praticamente transparente, sem alterar significativamente o áudio na compactação (CAMILLERI, 2009, p. 51). Apenas os exercícios de compactação de arquivos, com diferentes qualidades, apresentam conversões no formato *mp3*, e assim foram mantidos nas duas versões.

Os resultados dos exercícios podem ser escutados em alto-falantes, mas os efeitos são mais perceptíveis em fones de ouvido. Alguns exercícios, como o de espacialização, e alguns detalhes nos exercícios de masterização contextual funcionam apenas na escuta com fones de ouvido, devido ao uso de recursos psicoacústicos para a manipulação do palco sonoro virtual.

37 A versão compactada deles por ser obtida em <https://bit.ly/2YE7xM9>

3.1. Faixas utilizadas nos exercícios

Apresentamos abaixo a seleção de faixas utilizadas nos exercícios e o que motivou sua escolha. As práticas de masterização recomendam processar sempre os áudios com a melhor resolução disponível (OWSINSKI, 2008, p. 10). Por isso, todos os áudios utilizados vieram de arquivos de alta resolução, vendidos em *sites* especializados como HDTracks, ou extraídos de suportes físicos de alta resolução, como *audio blu-ray*. Na Tabela 2, quando o ano original de lançamento não é o mesmo que o indicado na coluna da versão utilizada, o material pode ter passado por uma masterização diferente da versão original. No caso das músicas que apresentavam o mesmo ano nas duas colunas, a versão *hi-res* foi comparada com a versão em *compact-disc*, podendo-se afirmar que se trata da mesma masterização, com a resolução reduzida para a mídia (KATZ, 2015).

Tabela 2 - Relação das faixas utilizadas nos exercícios.

Faixa	Intérprete	Álbum	Ano original de lançamento	Versão utilizada (ano)	Resolução original (bits/KHz)
Sunrise	Norah Jones	Feels Like Home	2004	Blue Note 724358480016 (2012)	24/192
Take Five	The Dave Brubeck Quartet	Time Out	1959	Columbia CS 8192 (2013)	24/176
Coriolan Overture, Op. 62	Luzerner Sinfonie-orchester & James Gaffigan	Beethoven: 4 Ouvertüren / Overtures	2018	Sony Classical G0100038 99360A (2018)	24/96
Thank You	Led Zeppelin	Led Zeppelin II	1969	Atlantic HD603497 920068 (2014)	24/96
Beat it	Michael Jackson	Thriller	1982	Epic G0100029 27138N (2015)	24/176
Get Lucky	Daft Punk	Random Access Memories	2013	Columbia 0886443 927087 (2013)	24/88
Heart is a Drum	Beck	Morning Phase	2014	Capitol HDTracks (2014)	24/96
Shape of You	Ed Sheeran	Divide	2017	Atlantic HDTracks (2017)	24/44
Garota de Ipanema	Rosa Passos & Ron Carter	Entre Amigos	2003	Chesky SACD291 (2005)	24/96
Smells Like Teen Spirit	Nirvana	Nevermind	1991	Geffen B0015887-02 (2011)	24/96

Fonte: produção do autor.

As escolhas representam diversas combinações de características de áudio, especialmente observando suas dinâmicas: queríamos testar os processamentos desenvolvidos nos exercícios de masterização contextual em materiais bastante diferentes. A escolha também passou por critérios subjetivos e pessoais, informados abaixo. A análise das faixas, realizada através do *software* RX7 da iZotope, forneceu-nos as seguintes informações:

Tabela 3 - Estatísticas das faixas utilizadas nos exercícios.

Faixa	True Peak Level L (dBTP)	True Peak Level R (dBTP)	Sample Peak Level L (dBTP)	Sample Peak Level R (dBTP)	Program Loudness (LUFS)	Loudness Range (LU)	Peak to Loudness Ratio	possíveis <i>clippings</i> (L + R)
Sunrise	-0,56	-1,38	-0,57	-1,40	-14,7	3,7	14,14	0 + 0
Take Five	-1,32	-5,19	-1,34	-5,21	-21,2	7,2	19,88	0 + 0
Coriolan Overture, Op. 62	-0,40	-0,41	-0,43	-0,43	-14,4	20,5	14,00	0 + 0
Thank You	-0,10	+0,14	-0,10	-0,01	-13,7	16,8	13,70	3 + 0
Beat it	-1,54	-1,76	-1,73	-1,76	-15,7	6,4	14,16	0 + 0
Get Lucky	+0,50	+0,46	-0,10	-0,10	-10,9	3,8	10,90	391 + 548
Heart is a Drum	+0,11	+0,08	0,00	0,00	-8,6	3,5	8,60	33 + 24
Shape of You	+1,50	+1,37	-0,09	-0,09	-6,9	3,7	6,90	5286 + 7227
Garota de Ipanema	-2,74	-1,80	-2,75	-1,81	-19,7	5,7	16,96	0 + 0
Smells Like Teen Spirit	-0,05	-0,05	-0,08	-0,09	-7,1	6,0	7,05	0 + 0

Fonte: produção do autor.

A primeira faixa, *Sunrise* (faixa 13 do anexo), é da cantora *jazzy* Norah Jones. Apesar do *Program Loudness* (PL) baixo para os padrões da música *pop*, -14,7 LUFS, ela apresenta um *Loudness Range* (LRA) bastante baixo, 3,7 LU, valor próximo do das músicas *pop* da seleção.

A faixa *Take Five* (faixa 14 do anexo) é a gravação mais antiga da seleção, gravada por

The Dave Brubeck Quartet em 1959. Sua sonoridade original é mantida nessa versão, sem muita manipulação digital. Encontramos nela valores esperados para gravações históricas de jazz, com PL baixo e LRA intermediário. Apesar de comprimida - presumidamente com compressão de válvula e fita magnética - ela ainda mantém picos bastante elevados em relação ao corpo principal, como se pode ver na Figura 16.

A faixa seguinte, ao contrário de *Take Five*, é uma gravação recente de um material tipicamente audiófilo e com grande extensão dinâmica. Trata-se da interpretação de Luzerner Sinfonieorchester e James Gaffigan para a *Op. 62 de Beethoven - Coriolan Overture* (faixa 15 do anexo), que apresenta uma das mais amplas dinâmicas encontradas na pesquisa, com impressionantes 20,5 LU de LRA. Ela utiliza bem a faixa de extensão dinâmica disponível, oferecendo um ótimo material para sistemas *hi-fi* e, como veremos no exercício de mascaramento do ambiente, uma experiência desastrosa para uma escuta móvel privada.

A escolha de *Thank You* (faixa 16 do anexo), da banda de rock Led Zeppelin, surgiu na busca de alguma música do gênero que possuísse dinâmicas similares às da música erudita. Essa gravação possui grande LRA - alguns trechos do material chegam a se aproximar do nível de ruído das fitas magnéticas em que foram gravados.

Bob Katz cita como exemplo da “guerra de volumes” as diferentes masterizações do álbum *Thriller*, do cantor Michael Jackson, afirmando que a primeira versão em *compact-disc* da faixa *Beat it* (faixa 17 do anexo) possuía um PL de apenas -18,28 LUFS, sendo levada a -7,35 LUFS na versão de 2009, disponível em uma coletânea oficial (KATZ, 2015, p. 242). A versão escolhida para os exercícios deste capítulo é a de uma masterização conservadora de 2015, voltando a valores de -15,7 LUFS de PL e 6,4 LU de LRA. De nossa seleção, ela apresenta o valor de *Program Loudness* mais próximo da referência da tecnologia Sound Check, que é, de aproximadamente, -16,5 LUFS (KATZ, 2015, p. 221) - ou seja, com a normalização de intensidade ativada em um dispositivo da Apple, essa é a música da seleção que menos sofre alteração de intensidade. Seu *Loudness Range*, por outro lado, é elevado para uma música *pop* dançante, talvez reflexo da época em que foi gravada, em contraste com outras faixas da lista, o que veremos a seguir.

A dupla Daft Punk é conhecida por sua música *pop* dançante de alta qualidade técnica, especialmente o álbum *Random Access Memories*, de 2013, do qual foi extraída a faixa *Get Lucky* (faixa 18 do anexo). Apesar de seu som bastante definido e limpo, a masterização beirando os limites do áudio digital ultrapassa o valor de zero dBTP, gerando várias possíveis distorções no conversor digital-analógico, como comentado em capítulos

anteriores. Apesar desses *clippings*, a faixa mantém valores baixos de PL para músicas *pop*, se compararmos com as seguintes da lista, e um LRA próximo a essas mesmas músicas, indicando áudios bastante comprimidos - apesar de esse valor baixo ser bem menos evidente na escuta da faixa *Sunrise*, que possui valor similar de LRA.

As duas faixas seguintes, *Heart is a Drum* (faixa 19 do anexo) e *Shape of You* (faixa 20 do anexo), respectivamente dos artistas Beck e Ed Sheeran, foram escolhidas por terem valores similares em algumas de suas estatísticas: possuem LR e LRA bem próximos, típicos de músicas *pop* de meados de 2010. A segunda faixa, porém, parece utilizar-se propositalmente do recurso de distorção causado por *clipping*, ocorrendo milhares de vezes durante a música. Em *Heart is a Drum*, apesar de ser comprimida e tangenciar zero dBFS, essa distorção ocorre raramente, de forma aparentemente acidental.

A dupla que encerra a lista foi escolhida por seu contraste: a primeira delas, *Garota de Ipanema* (faixa 21 do anexo), em gravação de Rosa Passos e Ron Carter, é uma versão explicitamente audiófila, utilizando pouca compressão e mantendo os picos dos ataques dos instrumentos. A outra, *Smells Like Teen Spirit* (faixa 22 do anexo), é uma versão masterizada em 2011 (e ainda mais comprimida) do sucesso de 1991 da banda Nirvana, um dos melhores exemplos de compressão utilizada para criar impacto sonoro encontrados na pesquisa. A mixagem e a simplicidade de arranjo de *Garota de Ipanema* também é interessante nos exercícios de espacialização, quando utilizamos técnicas de manipulação do palco sonoro virtual para algumas situações de escuta.

Como podemos conferir nas estatísticas apresentadas na Tabela 3, as músicas escolhidas têm diferenças grandes entre si e são suficientemente díspares para ilustrar os problemas levantados na pesquisa. Os trechos das faixas selecionadas estão normalizados por intensidade e podem ser encontrados no anexo de arquivos digitais de áudio, com a numeração de 13 a 22.

3.2. Preparação e *softwares*

Após a seleção das músicas, elas foram extraídas de seus álbuns de origem e tiveram levantadas suas estatísticas. A taxa de amostragem de alguns dos áudios foi então convertida para 96 KHz, já que elas não eram todas iguais, visando facilitar a realização dos exercícios - todos os arquivos tinham resolução de 24 bits, mantida até a etapa final. A perda de qualidade é insignificante nas faixas que tiveram sua taxa diminuída para 96 KHz, e, para as aumentadas, isso apenas evita algumas degradações devido aos processamentos

realizados posteriormente (KATZ, 2015, p. 319). As faixas foram, então, normalizadas por intensidade (no nível utilizado pelo Sound Check, -16,5 LUFS) para serem processadas com *headroom* suficiente para evitar *clipping*, e facilitar a comparação na escuta - com as exceções indicadas, como no exercício de compactação de arquivo. Os procedimentos de recepção e finalização dos arquivos foram realizados no *software* RX7 da fabricante iZotope, e os processamentos intermediários foram realizados dentro da plataforma de áudio Studio One, da PreSonus.

Além do RX7 e do Studio One, os processamentos foram realizados com os seguintes *softwares*: MCompressor, da MeldaProduction; Pro-C2, da Fabfilter; Morphit, da ToneBoosters; CanOpener, da GoodHertz; e L3LL Ultramaximizer, da Waves. As configurações de parâmetros que funcionaram melhor foram selecionadas e ajustadas para ilustrar os assuntos discutidos nesta pesquisa. Os áudios foram tratados com os diferentes processamentos necessários para cada exercício no estúdio Audioclicks, em São Paulo. Os exercícios empíricos de mascaramento do ambiente, como percepção de espacialização e equilíbrio tonal dos áudios, que exigiam testes comparativos em trajetos pela cidade, testando a masterização contextual, foram realizados também na cidade de São Paulo, entre abril e julho de 2019.

Além dos monitores de áudio Genelec 8130a do estúdio de masterização onde foram realizados os processamentos, foram também utilizados dois fones de ouvido diferentes para referência e testes fora do estúdio: um *earbud* da Apple³⁸ e um supra-auricular fechado da Beyerdynamic, modelo DTX501p.

Visando facilitar a análise e comparação das faixas, após a realização dos processamentos de cada exercício, as músicas foram cortadas em trechos de 40 segundos, selecionados para destacar os detalhes mais significativos de cada exercício.

3.3. Normalização de intensidade

O primeiro exercício realizado é uma simulação de normalização de intensidade. As dez músicas, na ordem apresentada anteriormente, tiveram seu *Program Loudness* medido e ajustado para -16,5 LUFS - valor aproximado de PL para o qual a Apple altera as faixas com a tecnologia Sound Check. Apenas após esse ajuste, os trechos selecionados foram

³⁸ Mais precisamente o modelo EarPod, fornecido a partir do iPhone 5, lançado em 2012. A partir de 2017 o modelo teve seu plug modificado, mas as características sonoras permanecem inalteradas. Esse earbud também é distribuído junto à compra dos iPods, desde 2012. Antes disso o modelo utilizado, tanto no iPhone quanto no iPod, era também intra-auricular, porém com outra curva de resposta.

recortados e montados em sequência. A primeira sequência (faixa 23 do anexo) apresenta os trechos das músicas em sua versão original. A seguinte (faixa 24 do anexo) simula a normalização de intensidade feita pelo Sound Check.

Fig. 16 - Forma de onda da sequência dos trechos selecionados.

Fonte: produção do autor.

Fig. 17 - Forma de onda dos mesmos trechos, com aplicação de normalização de intensidade.

Fonte: produção do autor.

Como podemos perceber na comparação gráfica, há grande redução de intensidade nos trechos das músicas que continham maior valor de *Program Loudness*. Duas delas,

Take Five (a segunda) e Garota de Ipanema (a penúltima), possuíam PL abaixo de -16,5 LUFS e, por isso, não tiveram suas intensidades alteradas, assim como ocorre no uso da tecnologia Sound Check. Isso significa que, na escuta de uma playlist com o Sound Check ativado, apenas elas são reproduzidas com a intensidade diferente das outras músicas.

O quinto trecho, *Beat it*, foi o que sofreu menor redução de PL, o que pode significar que o engenheiro de masterização dessa versão de 2015 estava preocupado em utilizar o máximo de extensão dinâmica possível, sem se preocupar com a comparação com outros álbuns em sistemas sem normalização de intensidade. É preciso checar as estatísticas do álbum inteiro para reforçar tal afirmação, mas dificilmente essa música está entre as mais suaves de *Thriller*.

O antepenúltimo trecho, *Shape of You*, ilustra a distorção no conversor digital-analógico, comentada no Capítulo 1. Sem normalização de intensidade, o *True Peak* do sinal passa de zero dBTP, o que ocorre milhares de vezes nessa música, como já indicado na Tabela 3. O mesmo não ocorre na versão normalizada por intensidade, na qual se nota uma mudança de timbre dos instrumentos, ocasionada com essa mudança.

Importante ressaltar que a normalização de intensidade foi realizada aqui por faixa, e não por álbum, que é uma opção ao utilizar o Sound Check em um dispositivo da Apple. Essa escolha é relevante porque, na última década, no formato de arquivos digitais, os *singles* foram muito mais vendidos do que os álbuns (RIAA, 2019). Além disso, o serviço de *streaming* da Apple possibilita optar entre a normalização de intensidade por faixa ou por álbum, enquanto seu principal concorrente, o Spotify, só permite a normalização por faixa.

3.4. Alteração de extensão dinâmica

A compressão de extensão dinâmica é uma das ferramentas mais utilizadas em toda a cadeia de gravação. Do amplificador de guitarra elétrica ao compressor da masterização, ocorrem muitas manipulações na dinâmica de um sinal. Esse exercício é uma demonstração de como um compressor pode alterar a sonoridade de um áudio - não necessariamente para melhor.

Não faz parte do escopo desta pesquisa um estudo aprofundado dos tipos, efeitos e utilizações da compressão aplicada à masterização do áudio. Mas alguma explanação se faz necessária, pois ela interfere diretamente no resultado sonoro da compressão de extensão dinâmica utilizada neste e em outros exercícios a seguir.

Um compressor de dinâmica age como um redutor proporcional (*ratio*) do ganho do

signal de áudio, a partir de um determinado valor (*threshold*). Ele aguarda um tempo antes de entrar em funcionamento após ser acionado (*attack*) e, após o valor do sinal do áudio sair da linha do *threshold*, ele continua agindo por um determinado tempo (*release*). Os tempos de *attack* e *release* dos compressores são fundamentais para seu funcionamento desejado, sendo parâmetros críticos no resultado sonoro da compressão. Em mixagem, esses valores podem variar de zero (ação instantânea, possível apenas no domínio digital) a muitos segundos. Resumidamente, o compressor reduz apenas os picos das ondas quando seu *threshold* está alto. Ao contrário, quando ele está baixo, o compressor altera a maior parte do áudio. O nome técnico que Bob Katz dá a esse tipo de compressão, que é a mais comum, é *downward compression* (KATZ, 2015, p. 76). Quando a taxa de compressão (*ratio*) é muito alta, o compressor é denominado limiter. Outras formas de alteração da extensão dinâmica são, por exemplo, o aumento da intensidade do material quando ele está abaixo do valor de *threshold* (*upward compression*), e seu aumento quando ele está acima do *threshold* (*upward expansion*). Esse último caso não é tecnicamente uma compressão, mas sim uma expansão de extensão dinâmica - um dos motivos para tratarmos o tema como alteração de extensão dinâmica, e não apenas compressão. Na Figura 18, apresenta-se uma possível representação visual desses processamentos.

Fig. 18 - Representações visuais de quatro tipos de compressão.

Fonte: produção do autor.

Neste exercício, o trecho do *jazz Take Five* (faixa 25 do anexo), de The Dave Brubeck

Quartet, foi submetido a três diferentes compressões³⁹. O primeiro exercício de compressão de expansão dinâmica (faixa 26 do anexo) é facilmente perceptível na comparação com a referência original, mas ainda funciona bem. Com seus 4,7 LU de RLA (calculados a partir da música inteira, e não apenas do trecho), o processamento comprime bastante o áudio original, deixando-o bem achatado - o que não é muito característico de gravações de *jazz*, principalmente as antigas. O uso dessa compressão, bastante transparente, ilustra como algo considerado como correto para os padrões atuais de música *pop* pode soar inadequado em outros gêneros. A música fica sem a dinâmica original, mas nada comparável a *Shape of You* (faixa 20 do anexo), de Ed Sheeran.

A compressão seguinte (faixa 27 do anexo) demonstra os efeitos de *pumping* na primeira metade do trecho (OWSINSKI, 2008, p. 261) e de hipercompressão na segunda metade para, utilizando a expressão de Bobby Owsinski (2008, p. 258), deixar o áudio “sem vida”, ilustrando um uso normalmente não desejado da compressão.

Alguns *softwares* de alteração de extensão dinâmica eliminam *ratio* e *threshold* e executam o processamento a partir de curvas de entrada e saída desenhadas graficamente, o que permite emular a resposta não-linear de compressores analógicos, ou criar uma mescla de compressão e expansão, que se alterna ao longo da faixa de extensão dinâmica. O último áudio deste exercício (faixa 28 do anexo) mostra a aplicação desse processamento não-convencional de dinâmica. Como ele foi criado durante a pesquisa, chamaremos apenas de *compressão em curva*, o que não o descreve com precisão, mas é suficiente para sua identificação. A compressão em curva foi desenvolvida a partir de testes realizados em várias músicas (inclusive nas dez selecionadas neste capítulo), sendo aplicada em quatro exercícios: neste mesmo (de alteração de extensão dinâmica), no seguinte (mascaramento do ambiente) e nos de masterização contextual para a escuta que mascara e para a escuta distraída.

O *software* utilizado para a implementação dessa curva foi o MCompressor da empresa MeldaProduction. Como podemos observar na Figura 19, na interface gráfica do *software*, quando o sinal está próximo do silêncio ($-\infty$ dBFS) não há alteração. Quando o nível do sinal do áudio está suave, o processamento faz com que o nível do sinal aumente fortemente. Em níveis próximos de zero dBFS, o sinal volta a não ser alterado. Partimos das sugestões de parâmetros de compressão feitas por Bob Katz (2015, p. 105) para definir

39 Com o objetivo de utilizar os mesmos parâmetros em todos os exercícios, essa faixa foi inteira normalizada para -16,5 LUFS.

nossos tempos de *attack* e *release*. A curva foi pensada como possibilidade a partir de três necessidades iniciais: evitar o aumento excessivo do ruído de fundo da gravação, elevar os níveis médios do material, e manter sem muita alteração as partes mais fortes da onda. Testamos diversas combinações de técnicas de compressão, e chegamos nessa configuração em curva. O ganho de entrada do processamento é regulado para atingir os patamares aproximados de “suave” e “forte” de um material com intensidade média de -16,5 LUFS. O ganho de entrada e o ganho produzido pelo processamento são compensados com uma redução de sinal na saída, produzindo resultados bastante consistentes entre as faixas, como analisaremos nos exercícios de masterização contextual.

Fig. 19 - Representação visual da compressão em curva.

Fonte: produção do autor, a partir do *software* MCompressor.

3.5. Mascaramento do ambiente

O exercício de mascaramento é uma demonstração de como a diminuição da extensão dinâmica pode ser utilizada a serviço da qualidade da escuta em situações de escuta móvel privada, mesmo piorando a qualidade de reprodução do áudio. O trecho escolhido para exemplificá-lo é *Thank You*, da banda Led Zeppelin (faixa 29 do anexo), cuja forma de onda, já normalizada para -16,5 LUFS, pode ser vista na Figura 20.

Fig. 20 - Forma de onda do trecho original.

Fonte: produção do autor.

Foi utilizado o processamento de *compressão em curva* para elevar os níveis de sinal das passagens suaves, como descrito no exercício anterior. A presença de forte ruído de fundo, resultado de processos analógicos durante a produção da música, chama a atenção, mesmo na versão original, e é acentuada com a aplicação da *compressão em curva*, como pode ser percebido no exercício (faixa 30 do anexo) e na Figura 21.

Fig. 21 - Forma de onda do trecho com a compressão em curva aplicada.

Fonte: produção do autor.

A comparação dessas duas versões mostra como, na segunda versão, as partes mais suaves tiveram sua intensidade aumentada, enquanto as mais fortes pouco mudaram. A escuta em situações de ambiente ruidoso explicita o quanto essa compressão em curva pode facilitar a escuta das partes suaves, destacando-as do ruído ambiente. Para simular essa situação, um mesmo ruído urbano foi sobreposto às duas versões já apresentadas, resultando na versão original em ambiente ruidoso (faixa 31 do anexo) e na versão comprimida no mesmo ambiente ruidoso (faixa 32 do anexo). O efeito é facilmente perceptível aos 14 segundos dos trechos apresentados, quando a voz do cantor quase desaparece no ruído, e o som do teclado fica submerso no ambiente urbano - o que não acontece com a mesma intensidade na versão comprimida.

Importante destacar que as faixas completas (não os trechos) estão com o mesmo *Program Loudness*, o que significa que, reproduzidas em um dispositivo com normalização de intensidade, as duas versões apresentariam a diferença de intensidade aqui exposta. As variações do ruído ambiente geram mascaramentos temporários e, naturalmente, diferentes fones de ouvido permitem diferentes isolamentos acústicos com a cidade, mas o objetivo aqui é apenas ilustrar o efeito de mascaramento. Como analisaremos no exercício de masterização contextual, o problema do mascaramento é complexo demais para ser resolvido com um processamento prévio à situação.

3.6. Equilíbrio tonal

Cada um dos módulos de um sistema de reprodução de áudio (um amplificador e um alto-falante, por exemplo), comercializados separadamente, não consegue levar em consideração as curvas de resposta do outro. Assim, sem saber qual alto-falante será utilizado, o projetista de um amplificador de áudio não pode compensar as limitações do transdutor. De forma similar, o engenheiro de masterização não consegue prever em quais sistemas o áudio em que está trabalhando será reproduzido. Sua solução, desde o surgimento da profissão, é assumir um sistema de reprodução ideal, com uma resposta plana de referência (KATZ, 2015, p. 279). Mas, conhecendo qual sistema de reprodução está sendo utilizado, é possível compensar algumas das limitações na resposta dele - é isso o que permite fazer, por exemplo, o equalizador de um sistema de som doméstico.

Alguns *softwares* permitem compensar a curva de resposta de determinados equipamentos já analisados pelo desenvolvedor - como equalizadores pré-configurados para determinados alto-falantes. O presente exercício de equilíbrio tonal serve para demonstrar a

diferença de resposta que pode ser obtida com uma equalização adequada, compensando as deficiências de um módulo de transdução - neste caso, foi considerado o EarPod, fone de ouvido auricular fornecido desde 2012 pela Apple com seus *smartphones*.

O primeiro trecho disponível para a escuta é a referência utilizada, *Heart is a Drum* do cantor Beck, sem alterações além da normalização de intensidade (faixa 33 do anexo). A ele, aplicamos o processamento do *software* Morphit da empresa ToneBoosters (Figura 22), ajustado para compensar a resposta de frequências do EarPod (faixa 34 do anexo).

O *software* não altera os limites físicos de resposta do fone de ouvido, mas faz o melhor possível para aproximar sua curva de resposta à curva de um sistema de resposta plana. Mesmo se o *software* não consegue cumprir completamente sua promessa, porque isso depende do modelo de fone de ouvido utilizado, certamente ele proporciona uma mudança na qualidade de reprodução. A mudança mais perceptível no equilíbrio tonal acontece nos sons mais graves da versão processada (faixa 34 do anexo), um reflexo direto desses fones auriculares, pequenos e leves, que têm dificuldade em reproduzir frequências mais baixas.

Fig. 22 - Curva de compensação de resposta de frequências para o Earpod da Apple.

Fonte: produção do autor, a partir do *software* Morphit.

A escuta deste exercício em um modelo de fone de ouvido diferente do indicado não deve alterar muito o resultado. O que ele se propõe a mostrar é a diferença entre a escuta com monitores calibrados (mesmo que rudimentarmente) e descalibrados, ilustrando o quanto os transdutores que utilizamos no dia-a-dia variam a qualidade de reprodução sonora, alterando significativamente os sons por eles reproduzidos.

3.7. Espacialização

A localização espacial de uma fonte sonora é realizada através da interpretação da diferença entre os sons que chegam aos dois ouvidos (SOUSA, 2010). As reflexões no ambiente, a distância e o ângulo dos ouvidos em relação à fonte e as alterações na movimentação da cabeça permitem ao cérebro estabelecer qual a origem de determinado som no espaço.

Dois alto-falantes emitindo um mesmo sinal confundem essa percepção, criando uma “fonte sonora fantasma”, localizada entre eles (KATZ, 2015, p. 291). Variando a intensidade do sinal entre os alto-falantes, é possível alterar a posição percebida dessa fonte sonora fantasma, um dos recursos mais comuns de mixagem para distribuir os canais dos instrumentos no palco sonoro virtual.

A mixagem, normalmente realizada com alto-falantes, desenha o palco sonoro virtual, considerando as interações dos dois canais de um áudio estéreo com o ambiente. Os fones de ouvido, porém, reproduzem os canais separadamente para cada ouvido, alterando a localização do palco sonoro virtual. Posicionado normalmente no espaço entre os alto-falantes, portanto fora da cabeça do ouvinte, o palco sonoro é agora introjetado para o espaço entre seus ouvidos. Consequentemente, a sensação de distância da cabeça é anulada, como iremos ilustrar.

Uma maneira de contornar o problema da localização do palco sonoro virtual na escuta com fones de ouvido é utilizar tecnologias que simulam a interação dos dois canais de áudio com um ambiente e, logo, simular os sons dos dois alto-falantes chegando aos dois ouvidos, com as sutis diferenças que decorrem do posicionamento da cabeça em relação aos alto-falantes.

Uma das diferentes técnicas que produzem esse resultado é chamada de *headphone crossfeed*, já implementada com maior ou menor sucesso em diversos equipamentos, digitais e analógicos. Quando utilizado a contento, o *headphone crossfeed* posiciona o palco sonoro na frente do ouvinte, localizando-o para um posicionamento similar a de um palco sonoro virtual criado por alto-falantes em um ambiente acusticamente adequado (GOODHERTZ, 2018).

Uma versão digital de *headphone crossfeed*, chamada função de transferência relativa à cabeça (ou HRTF, na sigla em inglês), é utilizada, por exemplo, em espacialização de eventos sonoros em jogos eletrônicos. Os filtros utilizados para o *headphone crossfeed*

são relativamente simples de serem implementados, pois simulam fontes sonoras fixas (os alto-falantes), em oposição a elementos móveis de jogos eletrônicos com visão em primeira pessoa, nos quais complexas interpolações se fazem necessárias (SOUSA, 2010, p. 2).

A empresa norte-americana GoodHertz desenvolveu um algoritmo de *headphone crossfeed*, visando simular um ambiente acústico em fones de ouvido, para melhorar a experiência de espacialização de quem realiza mixagens de áudio com fone de ouvido. Ele foi implementado para uso em *softwares* de mixagem multipista, com o nome de CanOpener Studio, e também possui uma versão como aplicativo de reprodução de áudio para dispositivos portáteis no sistema iOS, com o nome CanOpener - for Headphones.

O exercício aqui apresentado permite comparar a profundidade de palco sonoro, com e sem a utilização desse processamento, e exige, portanto, a utilização de fone de ouvido para ser apreciado adequadamente. Os trechos de *Garota de Ipanema* em sua forma original (faixa 35 do anexo) e da versão com o *headphone crossfeed* aplicado (faixa 36 do anexo) precisam ser escutados com atenção, mas uma vez que a espacialização é compreendida, a diferença fica clara. Uma maneira simples de identificá-la é ouvindo as nuances da localização do contrabaixo e do prato. Na versão original, eles se opõem espacialmente, cada um dominando a escuta em um ouvido no início do trecho, e podem ser localizados nas laterais da cabeça. Já na versão processada com *headphone crossfeed*, eles fazem parte de um mesmo palco sonoro virtual, localizado à frente do ouvinte, revelando inclusive a profundidade desse palco. Algumas músicas produzidas com algumas tecnologias mais novas - e talvez voltadas para um público que consome música principalmente com fone de ouvido - utilizam esse processamento em alguns elementos das peças, o que ocorre no trecho de *Heart is a Drum* (faixa 19 do anexo).

A terceira escuta proposta neste exercício (faixa 37 do anexo) é uma versão do mesmo trecho, com o mesmo conteúdo nos dois ouvidos, ilustrando a diferença de percepção na localização dos instrumentos em um material monofônico. Esses processamentos são utilizados nos exercícios de masterização contextual.

3.8. Compactação de arquivo

Como descrito no Capítulo 2, o formato de áudio utilizado no *compact-disc* descreve a onda sonora digitalizada com resolução de 16 bits e taxa de amostragem de 44,1 KHz. Esse formato ocupa aproximadamente 10 megabytes de memória por minuto de áudio. Parte da mudança de hábitos de consumo de música aconteceu a partir da virtualização do

fonograma e da troca de arquivos de áudio pela internet. Mas, para essa troca de arquivos ser possível, foram utilizados alguns formatos de compactação de arquivo de áudio, sendo o *mp3* o mais conhecido deles⁴⁰. Esses formatos *lossy* (compactação com eliminação de conteúdo) dominam a troca e a comercialização de arquivos digitais de áudio, em uma valorização da alta-conveniência sobre a alta-fidelidade: um arquivo compactado ocupa menos espaço em memória, podendo ser transferido mais rapidamente e acumulado em quantidades maiores do que arquivos de áudio não compactados⁴¹.

O exercício apresentado aqui procura mostrar as diferentes quantidades de perda de informações de áudio, em função do *bitrate* escolhido na compactação para *mp3*. O trecho escolhido é de *Beat it*, de Michael Jackson, apresentado aqui sem normalização de intensidade, para permitir que o codificador execute a compactação da forma como ela é realizada na extração de áudio de um *compact-disc*⁴² (faixa 38 do anexo).

A escuta do primeiro resultado do exercício de compactação, utilizando um *bitrate* de 64 kbps, deixa evidente os artefatos sonoros produzidos, principalmente as sobras sonoras nas frequências médio-agudas, e a ausência das muito agudas, eliminadas pela compactação. Também podemos perceber facilmente a alteração de alguns timbres, como o da caixa da bateria (faixa 39 do anexo).

Outros dois diferentes níveis de compactação são oferecidos aqui para a escuta, 128 kbps (faixa 40 do anexo) e 320 kbps (faixa 41 do anexo). Os trechos apresentam menos eliminação de conteúdo, o que dificulta a percepção dos artefatos - principalmente no último, em que eles só ocorrem com alguma clareza em alguns tipos de eventos sonoros. Durante muito tempo, 128 kbps foi considerado o *bitrate* padrão para a compactação em *mp3* (STERNE, 2012, p. 14). Já 320 kbps é a melhor qualidade possível para esse formato de compactação⁴³.

40 Jonathan Sterne faz uma análise profunda do formato *mp3* e das causas e consequências da virtualização do fonograma em seu livro *mp3: the meaning of a format* (STERNE, 2012).

41 Existem diversos formatos de compactação *lossless*, ou seja, sem eliminação de conteúdo. Apesar desse tipo de arquivo ocupar menos memória do que um formato sem compactação, formatos *lossy* são ainda menores, e por isso mais utilizados.

42 Apesar da versão utilizada aqui ser originalmente de alta resolução, para a compactação em *mp3* ela foi reduzida ao padrão do *compact-disc*, 16 bits e 44,1 KHz, não havendo diferença significativa ou sensível com a versão fornecida em mídia física.

43 Chris Camilleri (2009) faz uma análise profunda das diferentes compactações e perdas dos formatos *mp3* e *aac*.

3.9. Masterizações contextuais

Os exercícios apresentados até aqui ilustraram processamentos que podem ser realizados para aumentar a qualidade de escuta ou a de reprodução em dispositivos portáteis de escuta privada. Eles podem ser implementados em tempo real pelo dispositivo, ou pela escolha de diferentes arquivos digitais de áudio em um serviço de *streaming*, utilizando a versão com os processamentos adequados ao usuário.

Uma masterização contextual, baseada na situação de escuta do usuário, pode incluir alguns desses processamentos em configurações específicas de situações. Os três exercícios finais são testes propositivos de como algumas das alterações vistas anteriormente podem ser combinadas, permitindo alterações da qualidade de escuta e de reprodução de um dispositivo portátil.

Os exercícios não têm a pretensão de esgotar o assunto, ou de resolver o problema proposto com uma “máquina de masterização” pronta para implementação em *smartphones*. Eles pretendem indicar relações entre processamentos de masterização e sua influência na escuta móvel privada, enfatizando especificamente as situações de escuta analisadas no Capítulo 2.

De forma similar aos exercícios anteriores, as faixas foram previamente normalizadas para uma intensidade similar à do Sound Check, -16,5 LUFS. Porém, ao contrário do que a tecnologia da Apple faz, visando manter o mesmo *Program Loudness* para todas as músicas, as gravações *Take Five* e *Garota de Ipanema*, que possuíam valores abaixo da meta do Sound Check, tiveram também seu PL alterado para -16,5 LUFS. A Apple evita realizar esse aumento porque isso pode causar *clipping* de algumas faixas com PL baixo e *Sample Peak* alto - exatamente o que acontece com a *Take Five*. Para evitar tal *clipping*, foi utilizado um *limiter*, ajustado apenas para segurar esses picos excessivos, um procedimento conservador em masterização, quando utilizado com parcimônia (KATZ, 2015, p. 90).

Para a comparação dos processamentos realizados nos três exercícios a seguir, recomendamos a escuta da montagem original (normalizada por intensidade) dos trechos das dez músicas (faixa 42 do anexo), seguida da comparação com os resultados obtidos nos exercícios (faixas 43, 44 e 45 do anexo).

3.9.1. A escuta que mascara

O primeiro exercício de masterização contextual busca mascarar os sons ambientes com as músicas escutadas em fone de ouvido. O efeito foi conseguido com a redução do

Loudness Range das músicas. Foi utilizada uma mesma configuração de processamentos para todas elas, o que poderia ser um processamento pré-programado, aplicado em tempo real nos *smartphones*. Além disso, uma manipulação do palco sonoro virtual buscou aumentar as sensações de distanciamento do ambiente e de isolamento, oferecidas pela escuta com fone de ouvido (BULL, 2012).

A partir da análise das dez faixas selecionadas para os exercícios, procuramos buscar um processamento de alteração da extensão dinâmica que, na contramão do que a “boa masterização” recomenda, reduzisse o *Loudness Range*, aproximando as passagens suaves das mais fortes, e evitando assim que, tendo sido regulado o volume para passagens médias, as suaves não fossem mascaradas pelo ruído ambiente.

Um resultado colateral positivo dessa compressão é a eliminação de picos excessivamente grandes para a extensão dinâmica disponível no ambiente urbano, como comentado anteriormente. Grandes extensões dinâmicas são desejadas para escutas atentas em sistemas de reprodução *hi-fi*. Em uma escuta privada em ambiente urbano, porém, elevações súbitas de intensidade podem tornar-se excessivas e agressivas aos ouvidos, pois ultrapassam o alto nível de intensidade frequentemente utilizado para obter o isolamento do ambiente (BULL, 2005). Esse processamento opõe-se às ideias de alta-fidelidade e preservação do registro na masterização, propondo alterações estéticas que visam a melhora da qualidade de escuta.

Ao mesmo tempo que desejávamos comprimir as músicas que apresentavam uma extensão dinâmica inadequada aos ruídos da cidade, foi necessário manter o *Loudness Range* de músicas que já estavam bastante achatadas. A solução encontrada, descrita no exercício de alteração de extensão dinâmica, foi utilizar a *compressão em curva*, como denominamos o processamento de compressão variável.

A manipulação do palco sonoro virtual também se mostrou uma ferramenta interessante para aumentar a sensação de isolamento na escuta de fones de ouvido. Mas, ao contrário das primeiras ideias para o exercício, um palco bastante aberto, criando a ilusão de uma redoma sonora ao redor da cabeça do usuário, funcionou melhor em nossas experimentações do que a redução dos canais estéreo para uma reprodução monofônica, o que parece posicionar o áudio dentro da cabeça do usuário. No exercício realizado, quanto mais externo à cabeça estava o palco sonoro virtual, mais as músicas pareciam barrar os sons ambientes.

Naturalmente, não se pode prever as variações de intensidades do ruído urbano. Algumas soluções, porém, podem ser pensadas para o processamento em tempo real, como por exemplo, adaptando a sugestão imaginada por Bob Katz: o microfone acoplado ao *smartphone* mede o ruído do ambiente e compensa automaticamente a quantidade de compressão utilizada. Sugerindo uma variação de ganho, Katz acaba propondo uma forma de compressão que varia em função do ruído ambiente (SOUTHALL, 2013).

O resultado do exercício (faixa 43 do anexo) pode ser comparado com a montagem original (faixa 42 do anexo) e experimentado em situações reais de ambiente ruidoso.

3.9.2. A escuta de fruição

Este exercício é um esforço para analisar a escuta de fruição em dispositivos portáteis com fone de ouvido. Ela parece-nos ser a escuta mais atenta dentre as possíveis em ambiente urbano, e pode ser observada na cidade por muitos dos usuários de dispositivos portáteis com fones de ouvido supra-auriculares - o que é aceito socialmente desde sua popularização como item de ostentação, como vimos no Capítulo 2. Algumas marcas têm lançado modelos preocupados com o conforto para o uso nas ruas, alinhando portabilidade com qualidade de reprodução sonora.

Buscando uma reprodução de qualidade em equipamentos de alta-conveniência, uma masterização contextual pode otimizar o arquivo de áudio para a escuta com fone de ouvido. A partir de processamentos em tempo real, e evitando distorcer a reprodução como no exercício anterior, pode-se adequar o áudio ao fone utilizado, para fornecer a reprodução o mais fiel possível do fonograma virtualizado, a fim de unir (ou anular) as ideias de alta-fidelidade e alta-conveniência (KATZ, 2010, p. 219).

Uma estratégia para otimizar a performance dos fones de ouvido, como vimos no exercício de equilíbrio tonal, é utilizar um processamento em tempo real para adequar o áudio à curva de resposta do fone de ouvido. Além desse processamento, podemos utilizar a manipulação do palco sonoro virtual para apresentar ao usuário uma escuta mais próxima da experiência com alto-falantes, usando a tecnologia de *headphone crossfeed*, como realizamos no exercício de espacialização. Para ilustrar essa masterização contextual para escuta de fruição, optamos por apresentar apenas o processamento de *headphone crossfeed*. Não aplicamos a adequação de curva de resposta do fone de ouvido para não distorcer o equilíbrio tonal do leitor que não possuir exatamente o mesmo modelo de fone de ouvido utilizado no exercício.

Quanto melhor for a resposta do fone de ouvido utilizado, melhor poderá ser percebido o efeito obtido. Uma das maneiras de ouvi-lo é reparar no posicionamento dos eventos sonoros que acontecem no centro do palco sonoro: em condições de escuta de fone de ouvido sem o *headphone crossfeed*, a fonte virtual de alguns elementos, como a voz do artista - frequentemente posicionados ao centro do palco sonoro - parece vir do centro da cabeça do usuário. Com o *headphone crossfeed*, a fonte sonora virtual parece estar localizada à frente da cabeça do usuário. O resultado do exercício (faixa 44 do anexo) pode ser comparado com a montagem original (faixa 42 do anexo).

3.9.3. A escuta distraída

O exercício de escuta distraída pretende criar uma ambientação sonora para ser ouvida em níveis bastante baixos, para uso nos momentos em que a escuta do ambiente externo é importante, como pedestres atravessando ruas ou ciclistas no trânsito.

Como pode ser percebido na escuta do exercício (faixa 45 do anexo), o nível de intensidade da faixa foi reduzido para permitir uma comparação de intensidade com os outros exercícios, sem a necessidade de alteração do ganho no dispositivo. No sentido inverso da masterização contextual para escuta que mascara - na qual um processamento buscou ampliar a sensação de espacialidade -, aqui optamos por converter o áudio, originalmente em estéreo, para um sinal monofônico, pois em nossos testes, a falta de espacialização se mostrou eficiente em mesclar o áudio reproduzido com os sons ambientes. Em situações muito ruidosas, o ambiente externo mascara a escuta distraída, o que nos pareceu aceitável: essa proposta de masterização objetiva gerar o pano de fundo para os eventos sonoros pontuais da cidade, ao contrário do interesse dos entrevistados do Michael Bull (2005, 2012).

Para que as passagens mais suaves das músicas não desaparecessem por completo, uma compressão foi utilizada para diminuir o *Program Loudness* das faixas. Nesse caso, optamos por não utilizar a compressão em curva, escolhendo em seu lugar a primeira compressão experimentada nos exercícios de alteração de extensão dinâmica. O resultado pareceu menos agressivo em intensidades baixas, e o problema do aumento excessivo do ruído de fundo do áudio aqui não é relevante, pois seu ganho total foi diminuído. Sugerimos escutar o resultado desse exercício durante a execução de uma outra tarefa, e compará-la com os trechos inalterados (faixa 42 do anexo), porém com o ganho atenuado. Esse processamento também pode ser utilizado como função de atenuação temporária, como

funcionava o botão de atenuação do primeiro Walkman, que, quando pressionado, permitia ao usuário interagir com as pessoas sem precisar retirar o fone de ouvido (DEMARAS, 2018, p. 28).

3.10. Conclusões dos exercícios

Dos exercícios realizados, podemos tirar algumas conclusões interessantes a respeito das diferenças entre arquivos de áudio com altos valores de *Loudness Range* em relação a materiais mais comprimidos. Os primeiros são extremamente mais flexíveis a processamentos posteriores, enquanto os achatados geram problemas para sua manipulação, já que a distância entre seu nível médio de sinal e o topo dos picos é muito pequena, dificultando um processamento padronizado, e exigindo uma masterização dedicada - o que inviabiliza processamentos em tempo real no dispositivo. Por outro lado, arquivos com *Program Loudness* muito baixos acabam ignorados pela normalização de intensidade, sendo mantidos com a intensidade menor do que os normalizados. Esses casos, porém, são considerados raros por Bob Katz (2015, p. 256), que se preocupa com a possibilidade da Apple realizar no futuro esse ajuste de PL para maior quando necessário - exatamente o que precisamos fazer em duas das faixas escolhidas.

A medição de *Program Loudness* apresenta um porém para alguns tipos de música, como apontado por Bob Katz (2015, p. 220). Arquivos de áudio com partes muito distintas, como foi o caso da faixa *Thank You*, acabam gerando valores de PL inadequados para representar tanto as partes suaves quanto as fortes. A escuta das faixas normalizadas, apresentadas no primeiro exercício de masterização (faixa 24 do anexo), exhibe essa situação, com as partes mais fortes de *Thank You* bem mais fortes do que os outros trechos; e suas partes suaves um pouco abaixo da média das outras. O trecho de *Op. 62 de Beethoven - Coriolan Overture* também passa por esse problema, mas como a expectativa da extensão de dinâmica da música erudita é maior, isso é menos perceptível. A agressividade causada pelo equilíbrio tonal de *Thank You* também colabora com essa diferença de sensação entre os dois trechos.

A normalização de intensidade já é uma realidade nos serviços de *streaming*, que é a forma de consumo de música com maior crescimento no final da década de 2010 (RIAA, 2019). A masterização parece ainda presa a uma realidade de álbuns vendidos em mídias físicas, mas essa situação parece estar sendo modificada pela própria demanda de lançamentos de *singles*, diminuindo a importância dos álbuns e forçando os engenheiros de

masterização e os artistas a olharem mais para como essa música está sendo ouvida: cada vez menos em sistemas *hi-fi* domésticos, e cada vez mais em fones de ouvido pela cidade, usando *smartphones* como dispositivos de reprodução de áudio.

Os ajustes do equilíbrio tonal e da espacialização não são explorados como poderiam para aproveitar ao máximo as características da escuta por meio de fones de ouvido, resultado da comercialização de uma só masterização por música. Edições especiais, de álbuns ou singles masterizados para a escuta em fones de ouvido, poderiam aproveitar, por exemplo, a maior capacidade de espacialização. Na reprodução em alto-falantes podem surgir vários problemas na reprodução dos dois canais de uma música estéreo no ambiente, criando cancelamentos de fase, entre outros. No isolamento dos fones de ouvido, porém, o controle do palco sonoro virtual é muito maior.

Aliado à correção de equilíbrio tonal dos diferentes modelos de fones de ouvido, abre-se uma nova camada de qualidade de reprodução nos dispositivos de escuta privada, com fonogramas adaptados a esses equipamentos. Para masterizações contextuais, as possibilidades são ainda maiores, pois, a partir de processamento (ou do streaming de fonogramas já processados) voltados a situações específicas de escuta, é possível otimizar a qualidade de escuta do usuário para cada necessidade de reprodução sonora dele - um cenário bem diferente da reprodução de um mesmo fonograma para distintos usos.

Considerações finais

Ao longo da pesquisa, encontramos dois discursos antagônicos: o da produção e reprodução sonora associadas a um discurso tecnológico e de alta-fidelidade; e o do uso de equipamentos voltados à portabilidade e à praticidade, como Walkman, iPods e *smartphones* - o que sintetizamos na ideia de alta-conveniência.

A masterização do áudio voltava seus esforços para otimizar a qualidade de reprodução em determinado suporte. A “guerra de volumes” no domínio digital, porém, trouxe deterioração dessa qualidade, priorizando atingir níveis de intensidade cada vez maiores para os usuários dos *compact-discs* (KATZ, 2015).

A virtualização do fonograma ampliou as situações de escuta. O arquivo digital de áudio fez-se presente no cotidiano, com sua reprodução acontecendo em diversas situações que, mesmo que tenham sido inauguradas anos antes pelo Walkman, encontraram, a partir do iPod e do *smartphone*, um público maior. A virtualização também permitiu a ampliação da escolha do repertório via *playlists* e a comercialização de *singles*, novos desafios para a masterização, que agora precisa levar em consideração a fruição de fonogramas em situações móveis e privadas de escuta. O cenário altera-se ainda mais com a chegada dos mecanismos de normalização de intensidade, impondo a mesma percepção de volume a todas as músicas de uma *playlist*.

A curta extensão dinâmica das músicas comercializadas parece ser consequência da escuta mediada por alto-falantes, talvez uma herança da pequena faixa de dinâmica disponível nos primeiros suportes de áudio, como discos de cera e goma-laca, e também da capacidade de transmissão dos rádios (CUMMINGS, 2013). Essa extensão, menor do que a faixa disponível, é oposta ao discurso do hi-fi, que prega uma “fidelidade” à fonte sonora. Por outro lado, como vimos, grandes extensões dinâmicas são incompatíveis com a escuta cotidiana dos usuários comuns, frequentemente realizada em dispositivos móveis com fones de ouvido.

Frequentemente relegado como apenas uma parte de um dispositivo portátil de reprodução sonora, o fone de ouvido foi estudado no Capítulo 2, quando analisamos sua importância como equipamento de áudio responsável pela privatização da escuta. Com os *smartphones* estabelecendo-se como dispositivos de reprodução sonora, as situações de escuta em que o fone de ouvido aparece se tornam ainda mais frequentes.

A masterização pode fornecer os elementos de otimização da experiência de escuta

do usuário nas diferentes situações em que ele consome música: não é mais necessário que o engenheiro de masterização priorize uma situação de escuta em detrimento de outra. Múltiplos fonogramas podem ser distribuídos para responder a distintas situações de escuta, como propõe Hugh Robjohns (2014), ou o fonograma pode ser processado de diversas formas, em tempo real, pelo dispositivo de reprodução, como sugere Bob Katz (2015).

A partir de uma eventual onipresença dos dispositivos e uma facilidade tecnológica em obter qualidade de reprodução, Mark Katz (2010) vislumbra um futuro em que fidelidade e conveniência perdem seus significados, e as qualidades de reprodução e de escuta coexistem em um mesmo dispositivo. Nossa pesquisa procurou encontrar nas ferramentas da masterização disponíveis atualmente alguns elementos para contribuir com esse futuro.

Referências

ACOSTA MAYA, S. The Sony walkman. *Revista Universidad EAFIT*, v. 36, n. 117, p. 19-27, jun. 2012.

APPLE. iTunes Store Top Music Retailer in the US. **Apple**, 2008. Disponível em: <<https://www.apple.com/newsroom/2008/04/03iTunes-Store-Top-Music-Retailer-in-the-US/>>. Acesso em: mai. 2018.

_____. Mastered for iTunes: music as the artist and sound engineer intended. **Apple**, 2012. Disponível em: <https://images.apple.com/br/itunes/mastered-for-itunes/docs/mastered_for_itunes.pdf>. Acesso em: jun. 2019.

BALDWIN, N. **Head-band for Telephone-receivers**. Patente depositada em 1913. Documento disponível em: <<https://patentimages.storage.googleapis.com/7f/cf/ab/698071252b6629/US1127161.pdf>>. Acesso em: fev 2019.

BARBOSA, P. A. Brief overview of the evolution of musical technology: promises and risks for the diversity of cultural expressions. In: RICHIERI HANANIA, L.; NORODOM, A.-T. (Org.). **Diversity of Cultural Expressions in The Digital Era**. Buenos Aires: Teseo Press, 2016, p. 39-58.

_____, P. A.; ARANGO, J. J.; ASTORGA, E. M. V. Headphone listening cultures: contingencies and crossings. In: SONOLOGIA 2019, 2019, São Paulo. **Proceedings...** In press.

BARNES, E. **20.Chelsea.NYC.PM.25March2006**. 2006. FotografiasoblicençaCC [CCBY-SA 2.0]. Original disponível em <<https://www.flickr.com/photos/perspective/14765417491/in/photostream/>>. Acesso em: fev. 2019.

BECK. Heart is a Drum. In: BECK. **Morning Phase**. Los Angeles: Capitol Records, 2014. Áudio de alta resolução, HDTracks.

BEHRENDT, F. Locative media as sonic interaction design: walking through placed sounds. **Journal of Mobile Media**, vol. 9, n. 2, 2015. Disponível em <<http://wi.mobilities.ca/frauke-behrendt-locative-media-as-sonic-interaction-design-walking-through-placed-sounds/>>. Acesso em: jan. 2019.

BEYERDYNAMIC. Once and Today. **Beyerdynamic**, 2014. Disponível em: <<https://global.beyerdynamic.com/company/once-and-today>>. Acesso em: jun. 2019.

BULL, M. **Sounding Out the City: Personal Stereos and the Management of Everyday Life**. Oxford: Berg, 2000.

_____. No Dead Air! The iPod and the Culture of Mobile Listening. **Leisure Studies**, vol. 24, n. 4, p. 343-355, 2005.

_____. The Audio-Visual iPod. 2008 [2012]. In: STERNE, J. (Ed.). **The Sound Studies Reader**. New York: Routledge, 2012.

BYRNE, D. **How Music Works**. San Francisco: McSweeney's, 2012.

CAMILLERI, C. **Embracing mp3**: mastering audio in the age of mobile music. 2009. Dissertação (Mestrado em Tecnologia Musical) - Department of Music and Performing Arts Professions, The Steinhardt School, New York, 2009.

CASTRO, G. **Música Serve Pra Pensar**: comunicação em rede, consumo e escuta musical. São Paulo, 2014.

CHAMBERS, I. The Aural Walk. In: COX, C.; WARNER, D. (Ed.). **Audio Culture**: readings in modern music. New York: Continuum, 2004.

CHEN, B. April 28, 2003: Apple opens iTunes Store. **Wired**, 2010. Disponível em: <<https://www.wired.com/2010/04/0428itunes-music-store-opens/>>. Acesso em: abr. 2018.

CHION, M. **Audio-Vision**: sound on screen. New York: Columbia University Press, 1994.

CNPS - Central News Photo Service. **1917_London_Electrophone_hospital_listeners.jpg**. 1917. Fotografia. Original disponível em <https://commons.wikimedia.org/wiki/File:1917_London_Electrophone_hospital_listeners.jpg>. Acesso em: fev. 2019.

CUMMINGS, A. S. **Democracy of Sound**: music piracy and the remaking of american copyright in the twentieth century. New York: Oxford University Press, 2013.

DAFT PUNK; WILLIAMS, P. Get Lucky. In: DAFT PUNK; WILLIAMS, P. **Random Access Memories**. New York: Columbia Records, 2013. Áudio de alta resolução, Columbia 0886443927087.

DAVE BRUBECK QUARTET, THE. Take Five. In: DAVE BRUBECK QUARTET, THE. **Time Out**. New York: Columbia Records, 1959 [2013]. Áudio de alta resolução, Columbia CS 8192.

DEEZER. Deezer HiFi. **Deezer**, 2017. Disponível em: <<https://www.deezer.com/en/offers/hifi>>. Acesso em: mai. 2018.

DEFFREE, S. 1st Intel Pentium processor is shipped, March 22, 1993. **EDN Network**, 2018. Disponível em: <<https://www.edn.com/electronics-blogs/edn-moments/4369686/1st-Intel-Pentium-processor-shipped--March-22-1993>>. Acesso em: abr. 2018.

DEMARAS, P. Old School. **Tone Audio Magazine**, jun. 2018. P. 28-30. Disponível em: <http://www.tonepublications.com/MAGPDF/TA_089.pdf>. Acesso em: jun. 2019.

DU GAY, P. et al. **Doing Cultural Studies**: the story of the Sony Walkman. London: SAGE Publications, 1997.

ESTREICH, B. The Electrophone System. **Bob's Old Phones**, 2013. Disponível em: <<https://www.telephonecollecting.org/Bobs%20phones/Pages/Essays/Electrophone.htm>>. Acesso em: jan. 2019.

EVERYMAC. Apple iPod (Original/Scroll Wheel) 5 GB, 10 GB Specs. **Everymac**, 2018a.

Disponível em: <<https://everymac.com/systems/apple/ipod/specs/ipod.html>> Acesso em: abr. 2018.

_____. Apple iPhone Xs Max (Global/A2101) 64, 256, 512 GB Specs. **Everymac**, 2018b. Disponível em: <<https://everymac.com/systems/apple/iphone/specs/apple-iphone-xs-max-global-uk-eu-asia-a2101-specs.html>>. Acesso em: jun. 2019.

GOODHERTZ. CanOpener Studio. **GoodHertz**, 2018. Disponível em: <<https://goodhertz.co/canopener-studio>>. Acesso em: jun. 2018.

GREENBURG, Z. Dr. Dre's \$3 Billion Monster: The Secret History Of Beats. **Forbes**, 2013. Disponível em: <<https://www.forbes.com/sites/zackomalleygreenburg/2018/03/08/dr-dre-3-billion-monster-the-secret-history-of-beats-3-kings-book-excerpt/#478d996e258d>>. Acesso em: jan. 2019.

GUINNESS. Best-selling album. **Guinness**, 2017. Disponível em <<http://www.guinnessworldrecords.com/world-records/70133-best-selling-album>>. Acesso em: mai. 2018.

HEYLIN, C. **Sgt. Pepper's Lonely Hearts Club Band**: um ano na vida dos Beatles e seus amigos. São Paulo: Conrad, 2012.

HILL, B. The iTunes influence. **Engadget**, 2013. Disponível em: <<https://www.engadget.com/2013/04/29/the-itunes-influence-part-one/>, <<https://www.engadget.com/2013/04/30/the-itunes-influence-part-two-setting-the-music-free/>> e <<https://www.engadget.com/2013/05/01/itunes-influence-art-in-the-age-of-digital/>>. Acesso em: mar. 2018.

HJORTKJÆR, J.; WALTHER-HANSEN, M. Perceptual effects of dynamic range compression in popular music recordings. **Journal of the Audio Engineering Society**, vol. 60, n. 1/2, p. 37-41, 2014.

HODGSON, J. **Understanding Records**: a field guide to recording practice. New York: Continuum International Publishing Group Ltd, 2010.

HOSOKAWA, S. The Walkman effect. 1984 [2012]. In: STERNE, J. (Ed.). **The Sound Studies Reader**. New York: Routledge, 2012.

HOWETH, L. S. **History of Communications**: electronics in the United States Navy. Washington, D.C.: U.S. Government Printing Office, 1963.

HUNT, A.; HERMANN, T. Interactive Sonification. In: HERMANN, T., HUNT, A., NEUHOF, H. (Eds.). **The Sonification Handbook**. Berlin: Logos-Verlag, 2011.

IAZZETTA, F. **Música e Mediação Tecnológica**. São Paulo: Perspectivas, 2009.

_____. Da escuta mediada à escuta criativa. **Contemporânea Comunicação e Cultura**, Salvador, vol. 10 n. 01, p. 10-34, 2012.

IFPI. Global Music Report 2018: state of the industry. **International Federation of the Phonographic Industry**, 2018. Disponível em <<http://www.ifpi.org/dl/?url=http://www>>.

ifpi.org/downloads/GMR2018.pdf>. Acesso em: mar. 2018.

ISAACSON, W. **Steve Jobs**. New York: Simon & Schuster, 2011.

JACKSON, M. Beat It. In: JACKSON, M. **Thriller**. New York: Epic Records, 1982 [2015]. Áudio de alta resolução, Epic G010002927138N.

JOHNSON, C. **LeBron James warming up before USA vs Spain**. 2012. Fotografia sob licença CC [CC BY-SA 2.0]. Original disponível em <https://commons.wikimedia.org/wiki/File:LeBron_James_warming_up_before_USA_vs_Spain.jpg>. Acesso em: fev. 2019.

JONES, N. Sunrise. In: JONES, N. **Feels Like Home**. Los Angeles: Blue Note, 2004 [2012]. Áudio de alta resolução, Blue Note 724358480016.

KATZ, M. **Capturing Sound**: how technology has changed music (revised edition). Berkeley: University of California Press, 2010.

KATZ, B. **Mastering Audio**: the art and the science - second edition. Burlington: Focal Press, 2007.

_____. **Mastering Audio**: the art and the science - third edition. Burlington: Focal Press, 2015.

KOSS. Koss history. **Koss**, 2015. Disponível em <<https://www.koss.com/history>>. Acesso em: jan. 2019.

LED ZEPPELIN. Thank You. In: LED ZEPPELIN. **Led Zeppelin II**. New York: Atlantic Records, 1969 [2014]. Áudio de alta resolução, Atlantic HD603497920068.

LUZERNER SINFONIEORCHESTER; GAFFIGAN, J. Coriolan Overture, Op. 62. In: LUZERNER SINFONIEORCHESTER; GAFFIGAN, J. **Beethoven: 4 Ouvertüren / Overtures**. New York: Sony Classical, 2018. Áudio de alta resolução, Sony Classical G010003899360A.

MARINO, R. Homero Lotito, o mago da masterização. **Audio Video Magazine**, São Paulo, ed. 152, dez. 2009. Disponível em: <<https://www.referencestudio.com.br/wp-content/PDFsHL/audiovideomagazine-152.pdf>>. Acesso em: jun. 2019.

MERCADIER, E. **Bi-Telephone**. Patente depositada em 1891. Documento disponível em: <<https://patentimages.storage.googleapis.com/dd/da/7d/876557b92866f8/US454138.pdf>>. Acesso em: fev. 2019.

NIRVANA. Smells Like Teen Spirit. In: NIRVANA. **Nevermind**. Santa Monica: Geffen Records, 1991 [2011]. Áudio de alta resolução, Geffen B0015887-02.

OJANEN, M. Mastering Kurenniemi's Rules (2012): the Role of the Audio Engineer in the Mastering Process. **Journal on the Art of Record Production**, vol 10, 2015. Disponível em <<https://www.arpjournal.com/asarpwp/mastering-kurenniemis-rules-2012-the-role-of-the-audio-engineer-in-the-mastering-process/>>. Acesso em: abr. 2018.

ORTNER, R. **Je lauter desto bumm! - The evolution of loud:** Untersuchung zur Evolution der Lautheit und Klangcharakteristik in 60 Jahren westlicher Populärmusik. 2012. Tese (Mestrado em Música) - Department für Kunst- und Bildwissenschaften Zentrum für Zeitgenössische Musik, Donau-Universität Krems, Krems an der Donau, 2012.

OWSINSKI, B. **The Mastering Engineer's Handbook** (2th edition). Boston: Thomson, 2008.

_____. **The Mastering Engineer's Handbook** (4th edition). Boston: Cengage Learning, 2017. Ebook.

PASSOS, R.; CARTER, R. Garota de Ipanema. In: PASSOS, R.; CARTER, R. **Entre Amigos**. New York: Chesky Records, 2003 [2005]. Áudio de alta resolução, Chesky SACD291.

POHLMANN, K. C. **Principles of Digital Audio**. New York: McGraw-Hill, 2011.

REISINGER, D. Here's How Many iPhones Are Currently Being Used Worldwide. **Fortune**, 2017. Disponível em: <<http://fortune.com/2017/03/06/apple-iphone-use-worldwide/>>. Acesso em: mai. 2018.

RIAA. U.S. sales database. **Recording Industry Association of America**, 2019. Disponível em <<https://www.riaa.com/u-s-sales-database/>>. Acesso em: jun. 2019.

ROBJOHNS, H. The end of the Loudness War?. **Sound On Sound**, Cambridge, fev. 2014.

SANTOS, O. L. S. **As transformações da escuta a partir da utilização das mídias portáteis**. 2014. Dissertação (Mestrado em Sonologia) - Escola de Comunicação e Artes da Universidade de São Paulo, São Paulo, 2014.

SHEERAN, E. Shape of You. In: SHEERAN, E. **Divide**. New York: Atlantic Records, 2017. Áudio de alta resolução, HDTracks.

SHELVOCK, M. **Audio Mastering as Musical Practice**. 2012. Dissertação (Mestrado em Artes) - The School of Graduate and Postdoctoral Studies The University of Western Ontario, London (ON), 2012.

SILVA, R. A. Do Gramofone Ao Live Streaming. **Revista Tulha**, v. 1, n. 1, p. 251-263, jan.-jun. 2015.

SINGER, M. Nathaniel Baldwin, Utah Inventor and Patron of the Fundamentalist Movement. **Utah Historical Quarterly**, vol. 47, n. 1, p. 42-53, 1979.

SOUSA, G. H. M. **Auralização de fontes sonoras móveis usando HRTFs**. 2010. Dissertação (Mestrado em Ciências) - Instituto de Matemática e Estatística da Universidade de São Paulo, São Paulo, 2010.

SOUTHALL, N. Dynamic Range Compression: are the loudness wars over? **The Quietus**, nov. 2013. Disponível em: <<http://thequietus.com/articles/13821-loudness-wars-apple-itunes-bob-katz>>. Acesso em: abr. 2019.

STERNE, J. **The Audible Past:** cultural origins of sound reproduction. Durham; London:

Duke University Press, 2003.

_____. **Mp3**: the meaning of a format. Durham; London: Duke University Press, 2012.

STUFF. Headphones: the complete history. **Stuff**, dez. 2013. Disponível em: <<https://www.stuff.tv/features/headphones-complete-history>>. Acesso em: jun. 2018.

TAYLOR, R.; MARTENS, W. L. Hyper-compression in music production: listener preferences on dynamic range reduction. In: AUDIO ENGINEERING SOCIETY, 2014, Berlin. **Proceedings of the 136th Convention of the Audio Engineering Society**. Berlin, 2014.

TAYLOR, R. **Understanding the Use of Hyper-Compression in Music Production**: a systems based approach to examining innovative change in the field of music production. 2018. Tese (Doutorado em Comunicação e Artes) - School of Design, Communication and IT, Faculty of Science and Information and Technology University of Newcastle, Newcastle, 2018.

TAYLOR, T. **Strange Sounds**. New York: Routledge, 2001.

TOWNSEND, A. Locative-Media Artists in the Contested-Aware City. **Leonardo**, vol. 39, n. 4, p. 345-347, 2006.

TRUAX, B. **Acoustic Communication**. Westport: Ablex Publishing, 2001.

TUTERS, M.; VARNELIS, K. Beyond Locative Media: giving shape to the internet of things. **Leonardo**, vol. 39, n. 4, p. 357-363, 2006.

VERSTRAETE, P. The Secret Theatre Revisited: eavesdropping on locative media performances. **Journal of Sonic Studies**, vol. 5, 2017. Disponível em <<https://www.researchcatalogue.net/view/411830/411831>>. Acesso em: fev. 2019.

YES. The Revealing Science of God (Dance of The Dawn). In: YES. **Tales From Topographic Oceans**. New York: Atlantic Records, 1973 [1988]. Compact-disc, Atlantic 781 325-2.

_____. The Revealing Science of God (Dance of The Dawn). In: YES. **Tales From Topographic Oceans**. New York: Atlantic Records, 1973 [1994]. Compact-disc, Atlantic 82683-2.

_____. The Revealing Science of God (Dance of The Dawn). In: YES. **Tales From Topographic Oceans**. Burbank: Elektra/Rhino Records, 1973 [2003]. Compact-disc, Elektra/Rhino 8122-73791-2.

_____. The Revealing Science of God (Dance of The Dawn). In: YES. **Tales From Topographic Oceans**. Buckinghamshire: Panegyric Records, 1973 [2016]. Áudio de alta resolução, Panegyric GYRBD80001.

_____. Tales From Topographic Oceans. **Yes**, 2019. Disponível em: <<http://yesworld.com/discography/tales-topographic-oceans/>>. Acesso em: mai. 2019.

Anexo I - faixas do DVD

O DVD anexado à dissertação possui duas pastas, as duas com os arquivos de áudio utilizados no final do capítulo 1 (estudo de caso: Tales From Topographic Oceans) e ao longo do capítulo 3. Como já descrito na dissertação, uma das pastas contém os arquivos de áudio em sua resolução máxima, enquanto a outra contém os arquivos compactados no formato aac. Abaixo relacionamos as faixas para facilitar o manuseio e escuta do material. As informações são válidas para as duas pastas, assim como para a versão que pode ser obtida via internet⁴⁴.

Tabela 4 - Lista das faixas presentes no DVD anexo à dissertação.

faixa	nome do arquivo e conteúdo
1	trecho versão mixtape (original)
2	trecho versão disco de vinil (original)
3	trecho versão 1988 (original)
4	trecho versão 1994 (original)
5	trecho versão 2003 (original)
6	trecho versão 2016 (original)
7	trecho versão mixtape (normalizada)
8	trecho versão disco de vinil (normalizada)
9	trecho versão 1988 (normalizada)
10	trecho versão 1994 (normalizada)
11	trecho versão 2003 (normalizada)
12	trecho versão 2016 (normalizada)
13	trecho Sunrise (normalizada)
14	trecho Take Five (normalizada)
15	trecho Coriolan op 62 (normalizada)
16	trecho Thank You (normalizada)
17	trecho Beat It (normalizada)
18	trecho Get Lucky (normalizada)
19	trecho Heart Is A Drum (normalizada)
20	trecho Shape Of You (normalizada)
21	trecho The Girl From Ipanema (normalizada)
22	trecho Smells Like Teen Spirit (normalizada)
23	exercício de normalização de intensidade - sequência trechos originais
24	exercício de normalização de intensidade - sequência trechos normalizados

44 Disponível em <https://bit.ly/2YE7xM9>

faixa	nome do arquivo e conteúdo
25	exercício de alteração de dinâmica - original
26	exercício de alteração de dinâmica - compressão normal
27	exercício de alteração de dinâmica - compressão <i>pumping e lifeless</i>
28	exercício de alteração de dinâmica - <i>compressão em curva</i>
29	exercício de mascaramento do ambiente - original
30	exercício de mascaramento do ambiente - comprimida
31	exercício de mascaramento do ambiente - original com ambiente
32	exercício de mascaramento do ambiente - comprimida com ambiente
33	exercício de equilíbrio tonal - original
34	exercício de equilíbrio tonal - ajustada
35	exercício de espacialização - original
36	exercício de espacialização - <i>headphone crossfeed</i>
37	exercício de espacialização - monofônica
38	exercício de compactação - original
39	exercício de compactação - mp3 64 kbps
40	exercício de compactação - mp3 128 kbps
41	exercício de compactação - mp3 320 kbps
42	masterização contextual - sequência original normalizada
43	masterização contextual - escuta que mascara
44	masterização contextual - escuta de fruição
45	masterização contextual - escuta distraída

Fonte: produção do autor.