

UNIVERSIDADE FEDERAL DE MINAS GERAIS

Escola de Música

Programa de Pós-graduação em Música

Fabio Wanderley Janhan Sousa

**MÍDIAS AUDIOVISUAIS ADAPTATIVAS PARA REALIDADE VIRTUAL:
o *espaço extrínseco* de uma modalidade artística emergente**

Belo Horizonte

2019

Fabio Wanderley Janhan Sousa

**MÍDIAS AUDIOVISUAIS ADAPTATIVAS PARA REALIDADE VIRTUAL:
o espaço *extrínseco* de uma modalidade artística emergente**

Versão Final

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

Linha de Pesquisa: Sonologia

Orientador: Prof. Dr. João Pedro Paiva de Oliveira

Co-orientador: Prof. Dr. Jalver Machado Bethônico

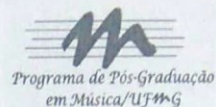
Belo Horizonte

2019

S725m	Sousa, Fábio Wanderley Janhan. Mídias audiovisuais adaptativas para realidade virtual [manuscrito]: o espaço extrínseco de uma modalidade artística emergente. / Fábio Wanderley Janhan Sousa . - 2019. 201 f., enc.; il. Orientador 1: João Pedro Paiva de Oliveira. Orientador 2: Jalver Machado Bethônico. Linha de pesquisa: Sonologia. Tese (doutorado) – Universidade Federal de Minas Gerais, Escola de Música. Inclui bibliografia. 1. Música - Teses. 2. Música e tecnologia. 3. Mídia digital. 4. Audiovisual. 5. Som - Registro e reprodução - Técnicas digitais. 6. Realidade virtual na arte. I. Oliveira, João Pedro. II. Bethônico, Jalver. III. Universidade Federal de Minas Gerais. Escola de Música. IV. Título. CDD: 789.96
-------	---



Universidade Federal de Minas Gerais
Escola de Música
Programa de Pós-Graduação em Música



Tese defendida pelo aluno FÁBIO WANDERLEY JANHAN SOUSA, em 22 de novembro de 2019, e aprovada pela Banca Examinadora constituída pelos Professores:

Prof. Dr. João Pedro Paiva de Oliveira
Universidade Federal de Minas Gerais
(orientador)

Prof. Dr. Jalver Machado Bethônico
Universidade Federal de Minas Gerais
(Escola de Belas Artes)
(coorientador)

Prof. Dr. Daniel Eduardo Quaranta
Universidade Federal de Juiz de Fora

Prof. Dr. Jônatas Manzolli
Universidade Estadual de Campinas

Prof. Dr. José Henrique Padovani Velloso
Universidade Federal de Minas Gerais

Prof. Dr. Eduardo Campolina Vianna Loureiro
Universidade Federal de Minas Gerais



RESUMO

O presente trabalho tem como objetivo principal sistematizar e aplicar, em produções próprias, uma estética da espacialização do áudio em mídias audiovisuais adaptativas para realidade virtual (VR). Para isso apontamos, como uma ferramenta articulatória, o conceito de *espaço extrínseco*, espaço definido pelo compositor de música eletroacústica no que diz respeito a localização de um determinado objeto sonoro, seu tamanho, movimentação e distância. Baseados nas figuras de espaço de Vande Gorne e nas USTs de Hautbois, propomos o que denominamos de USEs, Unidades Semióticas Espaciais, para sistematização do discurso semântico desse espaço. Com esses conceitos definidos, os aplicamos na descrição de projetos que envolvem tanto estéticas mais realistas, como o desenvolvimento de videocliques musicais em 360° gravados ao vivo, quanto estéticas mais abstratas, como é o caso de alguns estudos de música eletroacústica audiovisual tridimensionais para realidade virtual.

PALAVRAS-CHAVE: Espacialização de Áudio. Realidade Virtual. Áudio Imersivo. Mídia Adaptativa. Ambisonics. Sistemas Binaurais. Audiovisual.

ABSTRACT

The present work aims to systematize and apply, in the author's productions, an aesthetics of audio spatialization of adaptative audiovisual media for virtual reality (VR). For that we point out, as an articulatory tool, the concept of *extrinsic space*, space defined by the composer of electroacoustic music regarding to the location of a particular sound object, its size, movement and distance. Based on Vande Gorne's figures of space and Hautbois' USTs, we propose what we call SSUs, Spatial Semiotic Units, to systematize a semantic discourse of this space. With these concepts in place, we apply them to the description of projects that involve both more realistic aesthetics, such as the development of live-recorded 360° music video clips, and more abstract aesthetics, such as some studies of three-dimensional audiovisual electroacoustic music for virtual reality.

KEYWORDS: Audio Spatialization. Virtual Reality. Immersive Audio. Adaptative Media. Ambisonics. Binaural Systems. Audiovisual.

LISTA DE ILUSTRAÇÕES

Figura 1 - Pierre Henry usando o potentiomètre d'espace no concerto nas Salle de L'Ancien, Conservatório de Paris, 1952.	21
Figura 2 - Caricatura do órgão ocular de Louis-Bertrand Castel, por Charles Germain de Saint Aubin.	27
Figura 3 - Dispositivo de Wilfred.	27
Figura 4 - Imagens obtidas pelo dispositivo de Wilfred.	28
Figura 5 - Frame de um vídeo criado com o Lumigraph de Elfriede Fischinger.	28
Figura 6 - “Las Meninas” de Diego Velázquez.	37
Figura 7 - Modos de escuta segundo John Coulter, baseado nos modelos de Schaeffer e Chion.	39
Figura 8 - Fluxograma do áudio para a realidade virtual através do plugin O3A da Blue Ripple.	54
Figura 9 - Óculos 3D de Charles Wheatstone.	55
Figura 10 - Obra de ficção científica de Weinbaum.	55
Figura 11 - Sensorama de Morton Heilig.	56
Figura 12 - Headsight da Philco.	56
Figura 13 - A Espada de Dêmocles de Ivan Sutherland.	57
Figura 14 - O data suit da fabricante Atari.	58
Figura 15 - O data glove da Atari na capa da Scientific American.	58
Figura 16 - VIEW.	59
Figura 17 - Visually Coupled Airborne Systems Simulator, precursor do Super Cockpit.	59

Figura 18 - O projeto Super Cockpit, desenvolvido por Thomas Furness entre os anos de 1986 e 1989.	59
Figura 19 - Virtual Boy da Atari.....	60
Figura 20 - Oculus da Samsung.	61
Figura 21 - Arranjos de câmeras para se obter os 360° de imagem de alta resolução.....	62
Figura 22 - Sansumg gear 360, Nikon Keymission, 360Fly, GoPro Fusion e LG360Cam. Algumas das diversas câmeras 360° com apenas duas lentes disponíveis no mercado.	63
Figura 23 - Fluxo de trabalho em Ambisonics.	74
Figura 24 - Componentes W, X, Y e Z do formato-B Ambisonics de 1ª ordem, representando as componentes de pressão omnidirecional e três componentes de velocidade ortogonais do campo sonoro 3D, respectivamente.	75
Figura 25 - Harmônicos esféricos de sistemas baseados em Ambisonics nas ordens de 0 à 4.	76
Figura 26 - Microfones da Soundfield, SPS 200 e ST450 MKII.....	78
Figura 27 - Microfones AMBEO da Senheiser e NT-FS1 da Rode.....	78
Figura 28 - Plugins desenvolvidos por Dave Malham na Universidade de York – Inglaterra.	79
Figura 29 - Plugins desenvolvidos por Bruce Wiggins na Universidade de Derby – Inglaterra.	80
Figura 30 - Sistema de gravação e reprodução Binaural.	83
Figura 31 - Teoria do cancelamento cruzado.	85
Figura 32 - Decodificador Binaural do ATK for Reaper (com várias opções de HRTFs) e decodificador Binaural do FB360 (com uma única opção).	88

Figura 33 - Aglutinação das terminologias do conceito de espaço em música proposta pelo presente autor.	97
Figura 34 - Relações entre os conceitos de espaço em música proposta pelo presente autor.	99
Figura 35 - Linha do tempo descrevendo a prática composicional de manipulação do espaço na tradição musical ocidental.....	116
Figura 36 - Os locais definidos para os "músicos" 1 (vermelho), 2 (verde) e 3 (amarelo) que são reproduzidos no espaço, de acordo com as seções estabelecidas pela música - vista superior.....	122
Figura 37 - Frame de uma das músicas gravadas com o grupo de Jazz e QR Code para link.....	145
Figura 38 - Frame de um dos vídeos produzidos – projeto Semana Internacional do Café 2017 Sebrae e QR Code para link.	148
Figura 39 - Frame de um dos vídeos produzidos – projeto Corpografias e QR Code para link.....	149
Figura 40 - Partitura da obra Letter Piece de Matthew Shlomowitz.	151
Figura 41 - Frame do vídeo produzido baseado na obra Letter Piece de Matthew Shlomowitz e QR Code para link.....	152
Figura 42 - Frames de cada um dos cinco estudos e QR Code para link.	157
Figura 43 - Frames do estudo From the Inside e QR Code para link.	160
Figura 44 - Interfaces do usuário de alguns dos plugins ATK for Reaper.	188
Figura 45 - Plugin Audio 360 Spatialiser, do pacote de plugins FB360.	191
Figura 46 - Plugin Audio 360 Control, do pacote de plugins FB360.	192
Figura 47 - Plugins Audio 360 Converter, Audio 360 Meter e Audio 360 Stereo Loudness, do pacote de plugins FB360.	193

Figura 48 - Codificadores do pacote de plugins o3acore.	194
Figura 49 - Transformadores do pacote de plugins o3acore.	195
Figura 50 - Visualizadores do pacote de plugins o3acore.	196
Figura 51 - Decodificadores do pacote de plugins o3acore.	197
Figura 52 - Codificadores do pacote de plugins Ambix, desenvolvidos por Kronlachner.	198
Figura 53 - Transformadores do pacote de plugins Ambix, desenvolvidos por Kronlachner.	199
Figura 54 - Decodificadores do pacote de plugins Ambix, desenvolvidos por Kronlachner.	200
Figura 55 - Pacote de plugins multicanais para processamento de sinais em Ambisonics, desenvolvidos por Kronlachner, e oferecidos juntamente com o pacote Ambix.	200
Figura 56 - Plugin para teste da dearvr.	201

Sumário

INTRODUÇÃO	12
1. CONTEXTUALIZAÇÃO DO OBJETO DE ESTUDO	18
1.1 Classificação de manifestações artísticas audiovisuais	19
1.2 Intermedialidade e os dois níveis das relações audiovisuais	30
1.3 Audiovisual adaptativo: imersão através da interatividade	43
1.4 O desenvolvimento das tecnologias de espacialização sonora para produções audiovisuais	49
1.4.1 O áudio espacializado para cinema	49
1.4.2 O áudio espacializado para jogos digitais	51
1.5 O desenvolvimento do VR – uma perspectiva histórica	54
1.5.1 2017 "o ano do VR"	55
1.5.2 Ferramentas para geração e manipulação de Vídeo em VR	62
1.5.3 Mídia audiovisual adaptativa para VR: um híbrido de vídeo-música e arte sonora	66
1.6 Espacialização sonora para realidade virtual	69
1.6.1 Ferramentas contemporâneas para manipulação do áudio tridimensional	70
1.6.2 O Sistema <i>Ambisonics</i>	74
1.6.3 Sistemas <i>Binaurais</i>	81
2. O ESPAÇO EXTRÍNSECO NA MÍDIA AUDIOVISUAL ADAPTATIVA	89
2.1 Os diversos conceitos de “espaço” na música	89
2.1.1 Transgressões dos conceitos de espaço	100
2.1.2 A espacialização sonora em tempo real como prática interpretativa da música eletroacústica	103

2.2 Em busca de uma estética do <i>espaço extrínseco</i>	109
2.2.1 Estratégias de espacialização.	110
2.2.2 A sintaxe do <i>espaço extrínseco</i>	115
2.2.3 A semântica do <i>espaço extrínseco</i>	128
2.2.4 Unidades semióticas espaciais (USEs)	134
3. DESENVOLVIMENTO DE PROJETOS	143
3.1 Grupo 5/4 – gravação realista de um grupo de Jazz.....	143
3.2 Comerciais SEBRAE	145
3.3 Projeto <i>Corpografias</i>	148
3.4 <i>Letter Piece</i> (Matthew Shlomowitz) - As Is.....	150
3.5 <i>cinq études pour une autre réalité</i>	152
3.6 <i>From the Inside</i>	157
CONSIDERAÇÕES FINAIS E PESQUISAS FUTURAS.....	161
BIBLIOGRAFIA	164
ANEXOS	180
Anexo 1 – 15 figuras do espaço	180
Anexo 2 – 19 USTs.....	184
Anexo 3 – Descrição dos plugins utilizados nos projetos	188

INTRODUÇÃO

Em seu trabalho sobre o rádio e o cinema como artes-relé, Schaeffer anuncia que “uma filosofia do cinema e do rádio ainda está por se fazer” (SCHAEFFER, 2010, p. 27) e discorre sobre o fato de que estes seriam os aspectos mais característicos da civilização moderna, “o sinal dos tempos”, pois se observava sons e imagens em toda parte. Seu trabalho aponta para a busca de uma estética daquelas novas formas de se fazer arte e para isso ele observa que:

Convém (...) estudar sucessivamente as características do *objeto* ao qual o rádio e o cinema se dedicam enquanto instrumento, determinar as características desse *instrumento*, analisar a *transformação* que o objeto sofre após a passagem por ele. Será preciso finalmente examinar as condições psicológicas e sociais nas quais se encontra o público enquanto *receptor*. Só então, após essa exploração científica, poderá depreender-se uma técnica aplicada que definirá os *gêneros*, as *regras de uso* e de *execução*, e fornecerá elementos objetivos de apreciação dos *estilos*. (SCHAEFFER, 2010, p. 35)

Com isso ele nos traça o caminho para um dos nossos objetivos: o de elevar esteticamente as técnicas de manipulação e geração de material audiovisual para realidade virtual (VR). Schaeffer (2010, p. 53-54) considera tanto o rádio quanto o cinema, instrumentos, e enumera três fases de desenvolvimento dos mesmos: na primeira fase, o instrumento deforma a arte, no entanto perdoamos qualquer ruído no material produzido, uma vez que o mesmo ainda é admirado como novidade tecnológica; na segunda fase, na tentativa de transmitir a arte, de se igualar ao modelo, nos damos conta das suas deficiências e deformações, o acusando de ainda ter muito em que se aperfeiçoar; na terceira fase, graças ao conhecimento do instrumento, de suas possibilidades e limitações, o mesmo passa a informar a arte, ou seja, transmite, à sua forma, o que já se ouvia ou via anteriormente, mas passa a exprimir também o que não se via ou ouvia antes, dando sua própria contribuição ao desenvolvimento das artes.

Ao nosso ver, esses três estágios de um instrumento tecnológico, propostos por Schaeffer, fazem parte da evolução de qualquer nova mídia, se iniciando com uma manipulação livre e inocente da nova tecnologia, em um nível material, de descobertas, e, uma vez que já a entendemos e dominamos, aos poucos construímos

uma prática em que estabelecemos códigos que devem ser compreendidos, a fim de se aproveitar ao máximo a experiência estética proposta pelo artista / compositor.

Podemos afirmar que, a cada dia, o nosso objeto de estudo, a mídia audiovisual adaptativa, aquela destinada à reprodução através de um *smartphone* acoplado a um óculos de realidades virtual e fones de ouvido, ganha mais colaborações estéticas em direção ao seu terceiro estágio de desenvolvimento, que se caracteriza pelo “conhecimento do instrumento” e onde ele seria capaz de informar as artes, incorporando à elas a sua própria contribuição. Atualmente consideramos que, em seu segundo estágio, nosso objeto ainda apresenta padrões muito familiares ao espectador, com resquícios fortes do cinema e resultados da manipulação de novas tecnologias por amadores. Podemos observar claramente em plataformas de *streaming*¹ de audiovisual no formato digital (Ex.: YouTube 360°, Vimeo) uma série de experiências utilizando câmeras 360°, já bastante acessíveis e de fácil manipulação. Em um canal do YouTube² dedicado para esse tipo de material vemos de passeios em montanhas russas a gravações de final de semana com a família, na maior parte das vezes com o áudio finalizado e transmitido em estéreo, exatamente da mesma forma como as câmeras o entregam, sem qualquer intenção composicional, do ponto de vista estético, ou mesmo explorando a interatividade segundo o movimento de cabeça do ouvinte³, o que seria inerente a tal tecnologia. Tais vídeos, apesar de se utilizarem de uma novidade tecnológica, não exploram suas potencialidades. O que seria esperado de um produto com a finalidade de utilização em *smartphones* associados a dispositivos de interatividade, como os óculos de realidade virtual, por exemplo, nas mãos de amadores, produzem materiais muitas vezes extremamente banais, obscurecendo seu potencial artístico.

De maneira análoga àquilo que Schaeffer propõe às artes-relé, em um terceiro nível, o audiovisual para VR deve “informar a arte”, nesse caso deveríamos trazer informações extras, ampliar dados estéticos, criando relações e dialogando com outros conhecimentos e outras áreas, propondo novas sensações, percepções e expressões, fazendo o ouvinte, crítico, entender um ponto de vista ou criar sua própria interpretação do que vê / ouve / lê. Nesse terceiro estágio, o audiovisual para VR

¹ Transmissão em tempo real pela internet.

² <https://www.youtube.com/channel/UCzuqhhs6NWbgTzMuM09WKDQ>

³ Importante recurso possibilitado pela tecnologia em questão e ainda muito pouco assimilado, como será observado mais a frente nesse trabalho.

deveria utilizar mais do que relações temporais, sincronismos básicos e realistas, que apresentem análises óbvias. É nesse ponto que o processo de espacialização do áudio e a exploração do espaço sonoro, associado ao vídeo em 360°, pode fazer seu papel de criar relações inesperadas, de fazer com que o ouvinte utilize da sua memória e percepção espacial para ir além do evidente, além das relações dedutíveis, como na natureza, onde usamos da nossa audição e conhecimento do ambiente para associar ou dissociar uma imagem a um evento sonoro.

Quando compomos uma obra audiovisual escolhemos materiais e os relacionamos a partir de um pensamento, seja ele qual for, que vai além do que simplesmente aceitar aquele som que veio com aquela imagem. Criamos signos que serão reconhecidos ou não, criamos relações que serão identificadas ou não, mas que não são aleatórias, estão na obra a partir de um desejo do compositor de apresentá-las juntas.

Com isso em mente, podemos dizer que o nosso principal objetivo com esse trabalho seria o de refletir sobre a estética do que definimos como *espaço extrínseco*, no segundo capítulo, a partir do domínio da técnica de manipulação de áudio espacializado. Pretendemos desenvolver sua compreensão estética ao terceiro nível proposto por Schaeffer, ou seja, um nível em que tal manipulação possa representar algo além do observável no cotidiano, que possa trazer aos sentidos novos significados sejam eles quais forem, mas que vão além da irrefletida elaboração de movimentação espacial com o objetivo de trazer a atenção para uma novidade tecnológica, ou representar o real.

Nesse terceiro nível podemos esperar que sejam criadas confirmações, negações, decepções, surpresas, que sejam criados ambientes virtuais através da utilização de imagens e sons, onde o espectador é levado a aprender o funcionamento de cada obra de arte, quais sons se relacionam com quais imagens, e como elas se relacionam entre si através da sua distribuição no *espaço extrínseco*.

Uma vez que a tecnologia aqui estudada permite uma exploração da sensação de espaço observada pelo ouvinte, no presente trabalho, partimos dos diversos conceitos de *espaço* explorados principalmente na música eletroacústica para definir tal parâmetro como uma outra mídia básica⁴, além do som e da imagem, e estabelecer

⁴ Conceito extraído de Elleström (2017).

uma relação entre o que denominamos *espaço extrínseco*, tanto do ponto de vista *sintático* quanto do ponto de vista *semântico*, no processo de composição.

Essa nossa busca por uma separação entre *espaço intrínseco* e *extrínseco*, assim como de uma análise *sintática* ou *semântica*, tem seus fundamentos principalmente em uma mudança da abordagem artística que se desenvolveu no século XX. A exploração do conceito de *espaço* não é exclusivo da área da música. Segundo Alves (2011), enquanto as obras artísticas em geral, desenvolvidas no século XIX, apontam para uma percepção do espaço como dependente de uma interpretação, ou seja, um espaço idealizado e transmitido pelo artista somente através de referências, no século XX, surgem obras que se direcionam para a exploração do processo e, com isso, passam a explorar um espaço real, presente, material. “O espaço deixa de ser entendido como uma envolvente passiva ou neutra e começa, de diversos modos, a constituir-se como parte integrante da obra – substituindo-se assim um conceito de *espaço ideal* por um conceito de *espaço real*” (ALVES, 2011, p. 6).

Nessa exploração de um espaço real, o compositor tem em mãos a possibilidade de um novo meio expressivo, uma nova mídia, cuja linguagem própria ainda está por se fazer. Com o intuito de direcionar o entendimento e o desenvolvimento desse elemento expressivo emergente, pegamos emprestado três conceitos da linguística que, segundo Morris (BLÜHDORN, 1997, p. 152) seriam, na verdade, três subdisciplinas da semiótica: a sintaxe se ocupando da relação dos signos entre si; a semântica, dos signos e seus designados; e a pragmática, dos signos e seus usuários. A pragmática em si, indicaria uma perspectiva específica e uma atitude reconhecível com relação ao sistema de signos, e seria “caracterizada essencialmente pela concepção da dependência contextual do sentido discursivo, da racionalidade dependente do contexto e pela orientação da compreensão” (PARRET, 1984, p. 46). O fato de a pragmática ser “governada por um sistema de regras que são independentes das dimensões sintática e semântica, e simultaneamente correlacionadas com elas” (PARRET, 1984, p. 47), ser o “ambiente em que a sintaxe e semântica se situam” (BLÜHDORN, 1997, p. 158), é o que nos aproxima de uma análise, por esse viés, das composições eletroacústicas, uma vez que o compositor cria essa linguagem, esse conjunto de regras, dentro de sua obra. A pragmática

aplicada aqui seria portanto a da perspectiva do compositor, que analisa e desenvolve suas obras considerando uma certa sintaxe e semântica.

No percurso do nosso trabalho, objetivando uma contextualização do nosso objeto de estudos, realizamos, no capítulo 1, um breve mapeamento conceitual sobre as manifestações artísticas audiovisuais que envolvem elementos visuais e sonoros, e que se aproximam do nosso objeto. Em seguida abordamos conceitos relacionados a intermedialidade, uma vez que as relações entre as duas mídias, som e imagem, já exploradas por diversos autores, no servirão de guia para estruturar as formas como o *espaço extrínseco*, apresentado por essas duas mídias, poderá ser articulado. O estudo sobre o próprio conceito de interatividade pretende definir, dentro de uma variedade de abordagens, um recorte ao qual o presente trabalho está inserido, se mostrando uma característica essencial do que denominamos de mídia audiovisual adaptativa. Um breve histórico sobre o desenvolvimento das tecnologias de espacialização sonora, pretende apontar conexões possíveis: com o cinema, onde se buscaram relações entre som e imagem no domínio do espaço, a serem apresentadas ao ouvinte em ambientes próprios (salas de cinemas) e domésticos, apesar de suas limitações; e com a indústria dos jogos digitais, onde a interatividade com o jogador realiza um papel importante também no processo de relação entre som e imagem.

Nos direcionamos, posteriormente, aos desenvolvimentos relacionados à realidade virtual propriamente dita, passando por um histórico e contextualizando o que chamamos de mídia audiovisual adaptativa inclusive no que diz respeito ao seu aparato técnico. As ferramentas para geração de material visual são descritas brevemente aqui com o intuito de apresentar os recursos que exploramos nas nossas produções. Logo em seguida descrevemos as principais ferramentas para a manipulação de áudio em formato *Ambisonics* e *Sistemas Binaurais* para reprodução através de fones de ouvido, que é como a maior parte das mídias audiovisuais adaptativas para *VR* serão reproduzidas. O intuito nessa etapa não é o de aprofundarmos no conhecimento dos sistemas, uma vez que isso já foi realizado por diversos autores, mas sim dar um panorama geral das tecnologias e ferramentas que utilizaremos nas nossas criações.

O segundo capítulo se divide em duas principais partes. Na primeira, com um dos principais tópicos que deu origem à estruturação desse trabalho, nos direcionamos para o conceito de *espaço* na música, principalmente na música

eletroacústica, onde ele já foi vastamente explorado. A partir de definições extraídas de diversos trabalhos e compositores, entre eles Smalley, Barrett, Henriksen, Chion, Vaggione, Vande Gorne, Dignart, etc., elaboramos um agrupamento desses conceitos em quatro categorias principais de *espaço* que englobem todos os conceitos visitados: *espaço externo*, *espaço interno*, *espaço intrínseco* e *espaço extrínseco*. A definição desses conceitos, provenientes principalmente de trabalhos a respeito da espacialização de obras eletroacústicas e sua difusão como performance musical, tem como propósito principal o de apresentar o *espaço extrínseco* tridimensional das mídias audiovisuais adaptativas como elemento de coesão e elevação da prática da criação para a realidade virtual.

A segunda parte se direciona completamente à discussão estética baseada na divisão entre duas abordagens: uma delas *sintática*, direcionada ao material, e a outro *semântica*, relacionada ao significado. Nessas duas abordagens serão enquadrados alguns conceitos encontrados na literatura e artes pictóricas, assim como discutidas as possibilidades de articulação do *espaço*. As definições das USEs (unidades semióticas espaciais), entendidas também como construções espaciais mínimas percebidas como entidades capazes de gerar significado, estruturadas a partir de uma interseção entre as USTs (unidades semióticas temporais) de Hautbois (2010, 2014) e as figuras de espaço de Vande Gorne (2002), são apresentadas como um direcionamento terminológico e estruturante adotado na análise e desenvolvimento das composições apresentadas no terceiro capítulo.

No terceiro capítulo são descritos alguns dos projetos em que trabalhamos durante o período do doutorado. Neles, descrevemos a exploração das ferramentas para manipulação do audiovisual imersivo, assim como a utilização dos conceitos e terminologias relacionados à *semântica* e às USEs, objetivando principalmente a exploração das possibilidades estéticas da manipulação das mídias audiovisuais adaptativas.

Na conclusão, apresentamos observações relacionadas às limitações da mídia audiovisual adaptativa em seu atual estado e ao processo de sua criação, assim como apontamos algumas possibilidades de trabalhos futuros.

1. CONTEXTUALIZAÇÃO DO OBJETO DE ESTUDOS

Os meios de comunicação contemporâneos se pautam pelo uso simultâneo da imagem e do som. Também as manifestações populares articulam signos audiovisuais para expressar sua cultura e sua Identidade. (...) A visualidade e o som vêm ao encontro de forças poderosas dentro de nós, às quais somos extremamente sensíveis e que, por isso, conseguem fácil aceitação social. Utilizadas conjuntamente em uma síntese audiovisual, imagem e música atingem dois órgãos do sentido. Em suas diferentes formas, a articulação audiovisual de signos possui, em cada uma delas, potencialidades específicas. (...) A realidade AV é tão rica, tão diversa, tão profunda em seus vários modos de presença sensível, em suas operações constitutivas, quanto o mundo total da arte. (...) Como melhor representante desta multiplicidade, o audiovisual é síntese reveladora de nossa sociedade e deste tempo, não só no que esta sociedade humana tem de realidade técnica objetiva, mas também no que diz respeito às suas representações, ao seu fluxo e ao seu imaginário. (...) O AV tornou-se a forma mais adequada à expressão e ao conhecer da pluralidade humana de nossa sociedade; cada uma de suas manifestações tem seu modo diverso – com uma perspectiva própria – de contar a história do homem no século XX. (BETHÔNICO, 2001, p. 19, 20)

O novo milênio, principalmente essa segunda década, se iniciou com uma explosão de material audiovisual distribuído pela internet. Além da possibilidade de se transmitir ao vivo não somente áudio mas também vídeo, observamos o fato de a manipulação de áudio e vídeo se tornar mais simples e acessível a usuários com pouca experiência, prometendo resultados impressionantes. Com isso, diversas portas se abriram, entre elas, a distribuição de música com imagem pelos profissionais da área de entretenimento. Não falamos aqui somente de videocliques, mas também de obras eletroacústicas com material visual, antes acessíveis somente em concertos especializados. O surgimento de novos gêneros musicais / audiovisuais se intensificou e o que denominamos aqui de música eletroacústica audiovisual⁵ encontrou seus meios de atingir o grande público.

Na transformação de processos técnicos em artísticos, a gravação e manipulação do áudio se desenvolveu na composição eletroacústica e a gravação e manipulação de vídeo na música visual (ou *Visual Music*⁶). É nessa trajetória de desenvolvimento, ou evolução, de procedimentos técnicos em recursos estéticos, que

⁵ Conceito definido por Andrew Hill para descrever “obras de música eletroacústica audiovisual que exploram as possibilidades que a combinação de suas duas mídias baseadas em tempo (som e imagem em movimento) permitem.” (HILL, 2010, p. 37). Original: “Electroacoustic audio-visual music works explore the possibilities that the combination of their two time-based media (sound and moving image) allow”.

⁶ “A música visual pode ser definida como imagens visuais baseadas no tempo que estabelecem uma arquitetura temporal de maneira semelhante à música absoluta.” (EVANS, 2005, p. 11). Original: “Visual Music can be defined as time-based visual imagery that establishes a temporal architecture in a way similar to absolute music.”

se encontram as mídias audiovisuais adaptativas para a realidade virtual, mídia emergente desde o ano de 2016. Refletir sobre tal manifestação artística exige um estudo, mesmo que breve, com o intuito de constituir um panorama contextualizador, sobre as manifestações precursoras, e estas, por se apresentarem esparsas na literatura, demandam uma pesquisa bibliográfica preliminar como a que se segue.

1.1 Classificação de manifestações artísticas audiovisuais

Bethônico (1995) conceitua o audiovisual e discute diferentes usos do termo com suas conotações negativas e positivas.

Audiovisual 'diz-se da mensagem constituída da combinação de som e imagem' (FERREIRA, 1986, p. 199), sem importar a existência ou a ausência de próteses tecnológicas, se durável ou efêmera, folclórica ou infográfica, dominante ou dominada. Das diversas grafias encontradas, áudio visual, áudio-visual e audiovisual, preferimos essa última, não só porque é encontrada no 'Aurélio', mas por causa da concepção eisensteiniana de montagem onde dois significados justapostos geram um terceiro ainda maior que a simples soma das partes. (BETHÔNICO, 1995, p. 105)

Essa possibilidade de duas coisas que se somam, se tornarem algo maior do que a soma das partes, é desenvolvida também em estudos relacionados a hibridização e interatividade, como em Kwastek (2003, p. viii), que afirma que “a característica essencial da hibridização das artes é que a forma híbrida não pode ser dividida em seus componentes individuais: o todo é mais do que a soma de suas partes”⁷. Tais estudos serão aprofundados na seção 2.3.

Voltando historicamente no desenvolvimento desses estudos, vemos o embrião da parte sonora do audiovisual no desenvolvimento do fonógrafo de Thomas Edison em 1877, passando pelo desenvolvimento do transmissor do rádio (1896) e do manifesto de música futurista de Russolo (*L'Arte dei Rumori*⁸ - 1913). Obras como *Étude aux chemins de fer* de Pierre Schaeffer e *Poème Électronique* de Edgar Varèse, se apresentam como o que seria chamado por Trevor Whishart em 1996 de arte

⁷ Original: “The essential characteristic of the hybridization of the arts is that the hybrid form cannot be broken down into its individual components: the whole is more than the sum of its parts.”

⁸ Trad.: A arte dos ruídos

sonora, uma forma distinta de música que explora as ideias apresentadas por John Cage de que o som por si mesmo poderia ser capaz de comunicar⁹.

O que esses pioneiros [John Cage, Edgard Varèse, Steve Reich] fizeram, no entanto, foi estabelecer, em suas maneiras próprias muito diferentes, a crença de que o som por si só poderia ser arte: as formas muito específicas nas quais a música organiza o som nem sempre são inteiramente necessárias e, como sugeriu Cage, dada a oportunidade, o som pode falar por si mesmo.¹⁰ (GIBBS, 2007, p. 35)

O termo arte sempre foi historicamente utilizado para designar trabalhos relacionados a pinturas ou até mesmo a esculturas (GIBBS, 2007, p. 43). Em meados do século XX, quando a tecnologia possibilitou que o som fosse gravado e transformado em um objeto (pedaço de fita ou um sulco em um disco de vinil), passível de manipulações - sejam elas tão simples como o cortar e colar de pedaços de fitas magnéticas ou complexas como uma granulação ou reverberação - sendo modelado, esculpido e moldado, de acordo com estéticas artísticas emergentes, o antigo conceito de música se expandiu, se aproximando do que viria a ser denominado de arte sonora¹¹.

Para além da tradicional composição instrumental, e para além do registro fonográfico, essas manipulações sonoras foram aos poucos se incorporando aos processos de composição musical de vanguarda. Diferenças surgem então entre as abordagens de alguns compositores, derivadas principalmente no que diz respeito ao material sonoro explorado e às estratégias de manipulação. Entre tais abordagens, nomeamos aqui três delas, apontadas por Gibbs (2007), ao descrever esse processo de desenvolvimento da música eletroacústica, partindo da instituição da Radio e através do experimentalismo de Schaeffer.

A denominada *Musique Concrète* de Pierre Schaeffer, na França, explorava principalmente sons gravados e manipulados através de processos artesanais com a própria fita magnética. A música eletrônica, decorrente dos estudos de Werner Meyer-

⁹ "(...) Cage efetivamente criou a ideia de que o som em si mesmo poderia comunicar e, talvez o mais importante (para nós pelo menos), que ele poderia ser a base para uma forma de arte distinta." Original: "(...) Cage effectively created the idea that sound by itself could communicate and, perhaps more importantly (for us at least), that it could be the basis for a distinct art form." (GIBBS, 2007, p. 34)

¹⁰ Original: "What these pioneers [John Cage, Edgard Varèse, Steve Reich] did, however, was to establish, in their very different ways, the belief that sound by itself could be art: the very specific ways in which music organises sound are not always wholly necessary and, as Cage suggested, given the opportunity, sound can speak for itself."

¹¹ "Sound Art" ou "Sonic Art".

Eppler e Herbert Eimert, em Colônia, se apropriou da exploração de osciladores e filtros e, posteriormente, de sons provenientes dos mais diversos tipos de síntese, e foi popularizada por Stockhausen (HOLMES, 2008, p. 55). A composição de paisagens sonoras, por sua vez, direcionada pelos estudos de Murray Schafer, Barry Truax e Hildegard Westerkamp, também exploram a manipulação de gravações, no entanto, adotando um viés ecológico e mais naturalista. Aos poucos essas três vertentes se somaram ao que denominamos atualmente de música acusmática, sendo esta a música fixada em suporte (CD, fita magnética, arquivo de áudio) e reproduzida através de alto falantes.

Historicamente, o primeiro concerto público a apresentar obras eletroacústicas espacializadas foi organizado por Pierre Schaeffer, Pierre Henry e Jacques Poullin, em maio de 1952, na *Salle de L'Ancien*, no conservatório de Paris. O repertório executado no mencionado concerto contava com obras tais como *Timbres-durees* de Messiaen (MURRAY, 2016), os *Estudos Concretos* de Boulez e *Konkrete Etude* de Stockhausen.



Figura 1 - Pierre Henry usando o *potentiomètre d'espace* no concerto nas Salle de L'Ancien, Conservatório de Paris, 1952¹².

¹² Fonte: BATES, 2009, p. 127

Dentre as primeiras obras compostas especificamente para sistemas de reprodução com espacialização podemos citar *Pfingstatorium: Spiritus Intelligentiae Sanctus* de Ernest Krenek (1955), para um sistema estereofônico, *Gesang der Jünglinge* de Stockhausen (1955-1956), para cinco grupos de alto-falantes dispostos em torno do público (MENEZES, 1998, p. 22), e *Poeme Électronique* (1958) de Edgard Varèse. Esta última foi, provavelmente não só a composição mais icônica, mas uma experiência multissensorial completa, por explorar um sistema de áudio multicanal associado a efeitos de iluminação e projeções de slides (GIBBS, 2007, p. 30) em um espaço arquitetônico, o *Pavilhão Philips*.

De uma forma semelhante à consideração de que o fonógrafo deu origem a música eletroacústica, o Cinema, talvez possa ser considerado um precursor das artes visuais em movimento que se utilizam de ferramentas tecnológicas, tais como a arte digital, a videoarte, as instalações, as intervenções, as artes performáticas, entre outras. Segundo Gibbs, nos anos 1960, “a arte contemporânea como um todo estava em um estado de fluxo e novas formas emergiam praticamente todo dia”¹³ (GIBBS, 2007, p. 32), e a combinação de áudio e vídeo, aconteceu em diversos níveis, de diversas formas, e recebeu, ao longo de todo esse tempo, diferentes terminologias.

Nessa histórica busca de identificação e definição terminológica de manifestações artísticas audiovisuais, algumas delas estão estreitamente relacionadas com o nosso objeto de estudos. A denominada *Ambient Music* de Brian Eno, que, apesar de originalmente somente sonora, teve diversos vídeos desenvolvidos pelo próprio compositor para acompanhá-la, é conhecida pela sua característica de ser “uma forma musical originalmente concebida para ser parte de um ambiente, sala de espera ou elevador, ao invés de ser ouvida em sua forma pura”¹⁴ (GIBBS, 2007, p. 32). A atividade / profissão / processo denominado de *Sound Design*, “a criação de som com o propósito externo a si mesmo ao invés de uma autônoma obra de arte”¹⁵ (GIBBS, 2007, p. 36), se aproxima do nosso objeto por sua estreita relação com o filme e com o cinema, apesar do termo ser utilizado também pela indústria na criação de identidade sonora de marcas.

¹³ Original: “(...) contemporary art as a whole was in a state of flux and new forms emerged almost daily.”

¹⁴ Original: Ambient music - “A musical form that was originally designed to be part of the sonic environment rather than to be listened to in its own right.”

¹⁵ Original: Sound Design - “The creation of sound for a purpose external to itself rather than as a freestanding piece of art.”

Ainda direcionado à manipulação de sons, outro termo importante, a arte sonora, apresenta algumas discordâncias conceituais entre os autores, principalmente se expandirmos para o contexto atual. Nas palavras de Bernd Schulz (1999, p. 25), o “inexato termo de arte sonora”, é “uma forma híbrida que se desenvolveu nas fronteiras entre música e artes visuais”. Observamos, no entanto, uma descrição de Emerson com uma visão um pouco mais ampla:

Acredito que a música seja um subconjunto da ‘arte sonora’ e que a ‘arte sonora’ seja um subconjunto da ‘Paisagem Sonora’ e que a ‘Paisagem Sonora’ seja realmente o mundo ao nosso redor, virtualmente completo.¹⁶ (EMMERSON, apud GIBBS, 2007, p. 65)

Enquanto Schulz coloca a arte sonora na fronteira entre a música e as artes visuais, Emerson a vê como algo que engloba a música. Para Emerson a música estaria então entre as diversas manifestações artísticas sonoras humanas, dessa forma, esculturas que produzem sons (denominadas também de esculturas sonoras)¹⁷ estariam lado a lado com a música pop, a eletroacústica, a música instrumental e a denominada paisagem sonora do compositor canadense Murray Schafer. Arte sonora seria o nome dado ao que engloba tudo isso. A definição apresentada por Schultz, no entanto, nos é favorita principalmente por ser mais utilizada no meio artístico e ser mais condizente com as produções contemporâneas que se auto denominam artes sonoras.

Ainda nessa busca de compreender o território audiovisual e sua relação conceitual com a arte sonora, recorreremos a Leite (2004) e à sua classificação de modalidades e modelos do audiovisual. Apesar de existirem outras classificações na literatura, escolhemos iniciar por essa autora pelo fato de suas caracterizações estarem estreitamente relacionadas com a música eletroacústica, o que se mostrará importante para a definição e expansão do conceito do nosso objeto.

Leite descreve várias situações em que a imagem e o som podem se relacionar:

(1) no processo de criação de uma partitura, quando uma imagem pode dar origem à

¹⁶ Original: “So I think that music is a subset of sonic art and sonic art is a subset of soundscape and soundscape is really the world around us, virtually complete.”

¹⁷ Como exemplos de “Esculturas Sonoras” podemos citar os trabalhos de Zimoun (www.zimoun.net/), Michele Spanghero (www.michelespanghero.com), Jack Pavlik e suas “esculturas sonoras cinéticas” (<https://www.youtube.com/watch?v=C6hOKQL5SDA>), Tonkin Liu com sua obra “Singing Ringing Tree” (<http://www.tonkinliu.co.uk/project/singing-ringing-tree>), ou até mesmo Martin Molin e sua “Marble Machine” (<https://www.youtube.com/watch?v=IvUU8joBb1Q>). Tais trabalhos se confundem em definição também com o termo “instalação sonora”.

criação de materiais sonoros ou à construção de estruturas, (2) no processo de produção quando o som é concebido para uma imagem concluída ou vice-versa, (3) no processo em que um som real de uma imagem é tomado como material para a composição, (4) no processo de composição onde a imagem e a música interagem entre si e são concebidos simultaneamente ou (5) quando a imagem ilustra uma música previamente composta. Dos processos apresentado pela autora, o que mais nos interessa seria o (4) onde as composições da imagem e da música são concomitantes, por vermos aí uma possibilidade de elevação da produção do audiovisual para realidade virtual ao nível artístico, ao invés de restrito a produção em massa de conteúdo¹⁸.

A autora elabora também uma distinção entre duas modalidades, a *Music Vídeo* da música vídeo, que apesar da proximidade pela tradução, se apresentam como duas coisas completamente diferentes, sendo a *Music Vídeo* melhor traduzida para o que conhecemos como videoclipes, direcionado a música de massa, e a música vídeo, com uma relação mais estreita com a música eletroacústica.

Realizamos essa comparação entre esses termos com o intuito de distinguir o nosso objeto de outros, mais comumente abordados por trabalhos na área da comunicação social. Trabalhos como o de Berry e Shelton (1999) também exploram o que seria categorizado como *Music Video* (videoclips, vídeos de músicas como os exibidos pela emissora de TV MTV), terminologia utilizada novamente por Borém (2016), quando o mesmo descreve ferramentas para sua análise (MaPAs e EdiPAs¹⁹).

Segundo Leite (2004) ainda, a música vídeo se apresenta como um gênero que engloba diversas modalidades do fazer musical contemporâneo que possuem relações diretas com a música eletroacústica: *tape music*, música mista, *live electronic music*²⁰. Esse direcionamento para um objeto relacionado com a música eletroacústica também é desenvolvido por Kasulin (2015) e Lima (2011). Este último

¹⁸ Pressupomos aqui uma diferença principal entre um processo de composição audiovisual com objetivo artístico e um processo direcionado a comunicação em massa, uma vez que em processos direcionados à comunicação em massa, na maioria das vezes, temos uma equipe trabalhando em função de orientações de um cliente e que muitas vezes não têm a visão do todo, sendo o trabalho baseado em padrões e clichês já testados comercialmente.

¹⁹ MaPAs (Mapas de Performance Audiovisual) e EdiPAs (Edição de Performance Audiovisual) são ferramentas para análise de um produto audiovisual. Apesar de estarem sendo utilizadas para análise de videoclipes no trabalho apresentado por Borém, se vislumbra a utilização de tais ferramentas também para obras eletroacústicas visuais, de forma semelhante ao processo sugerido por Tagg (TAGG, 1982) para a análise e estruturação de obras puramente instrumentais por leigos.

²⁰ Também denominado por outros autores simplesmente de *live electronics* (MENEZES, 1996).

propõe que o termo definido por Leite, música vídeo, ao ser fixado em uma mídia fixa seria denominado vídeo música, e se relacionaria historicamente com outras produções videográficas onde o áudio e o vídeo “se conectam de forma indissociável e não-hierárquica”.

Segundo o autor, o denominado vídeo música seria uma “modalidade de produção audiovisual em que as imagens visuais e sonoras se articulam em um processo contínuo e integrado, hibridizando-se em uma obra na qual a experiência perceptiva é trans-sensorial” (LIMA, 2011, p. 8). Ele caracteriza a vídeo música pela forte relação entre os processos, tanto de manipulação sonora da música eletroacústica quanto das imagens que as acompanham. E também, como proposto por Bethônico (1995) e Kwastek (2013), reforça a hibridização entre imagens visuais e sonoras, que fariam com o vídeo-música se apresente como

(...) um híbrido entre imagens visuais e sonoras que não apenas se articulam como somatório de eventos a partir das suas semelhanças e contradições significativas, mas criam através do próprio processo de hibridização um produto único e indissociável cujos elementos constituintes não apenas contribuem para a formação da experiência estética, mas representam um corpo totalmente novo e diferente daquele produzido por suas partes isoladas. (LIMA, 2011, p. 44)

O próprio autor nos apresenta alguns exemplos do que ele denomina de vídeo música, entre eles estão os trabalhos de Mary Ellen Bute (Ex.: *Synchromy n°4*), Walter Ruttmann, Oskar Fischinger, Norman McLaren, Malcolm Le Grice, Michael Snow, Ken Jacobs e Abigail Child. Muitos desses autores são mencionados também em outros trabalhos, no entanto possuem seus trabalhos classificados como *Visual Music* ou *Seeing Sound*, uma outra categoria encontrada na literatura que possui características importantes na definição do nosso objeto de estudos:

A música visual pode ser definida como imagens visuais baseadas no tempo que estabelecem uma arquitetura temporal de forma semelhante à música absoluta. Normalmente é não narrativo e não representacional (embora não seja necessário). A música visual pode ser acompanhada de som, mas também pode ser silenciosa.²¹ (EVANS, 2005, p. 11)

²¹ Original: “Visual music can be defined as time-based visual imagery that establishes a temporal architecture in a way similar to absolute music. It is typically non narrative and non-representational (although it need not be either). Visual music can be accompanied by sound but can also be silent.”

Nos voltando um pouco para a análise desse termo, *Visual Music*, mais difundido nas artes visuais, o gênero em si tem sua origem em produções puramente imagéticas, com características inerentemente abstratas, e provavelmente foi utilizado pela primeira vez por Roger Fry²² em 1912 para descrever a obra de artistas pós-impressionistas franceses, que seguiam exemplos de Kandinsky e Picasso (REKVELT, 2018):

(...) Eles não procuram imitar a forma, mas criar forma; não para imitar a vida, mas para encontrar um equivalente para a vida. Com isso, quero dizer que eles desejam fazer imagens que, pela clareza de sua estrutura lógica e pela sua unidade de textura, devem atrair a nossa imaginação desinteressada e contemplativa com algo da mesma vivacidade que as coisas da vida real apelam para nossas atividades práticas. Na verdade, eles não visam a ilusão, mas a realidade. O extremo lógico de tal método seria, sem dúvida, a tentativa de abandonar toda semelhança com a forma natural e criar uma linguagem de forma puramente abstrata - uma música visual; e as obras posteriores de Picasso mostram isso com clareza suficiente. Eles podem ou não ser bem sucedidos em sua tentativa.²³ (FRY, 1920, p. 157)

No desenvolvimento da denominada *Visual Music*, ao longo da história, diversos artistas e cientistas investiram em experimentos que fizessem algum tipo de associação entre som e imagem, objetivando a criação de um dispositivo que funcionasse em tempo real, direcionado à performance, alguns deles gerando simultaneamente som e imagem, outros gerando somente a imagem. Provavelmente os mais conhecidos foram os que resultaram no *Clavecín Ocularie* de Louis Bertrand Castel (1728), no *Clavilux* de Thomas Wilfred (1919), no *Nourathar* de Mary Hallock-Greenewalt (1924) e no *Lumigraph* de Elfriede Fischinger (1955).

Enquanto no dispositivo de Castel, do início do século XVIII, talvez o mais simples entre eles, existia uma associação direta entre a altura gerada pelo órgão e a cor emitida pelo mesmo, o dispositivo de Thomas Wilfred,

(...) rejeitava uma correspondência mais direta das cores com alturas definidas, emprestando termos como ritmo, performance ou mesmo a introdução do 'tempo' em obras visuais como interface com o sonoro. Deu

²² Pintor e crítico de arte inglês (1866-1934).

²³ Original: "They do not seek to imitate form, but to create form; not to imitate life, but to find an equivalent for life. By that I mean that they wish to make images which by the clearness of their logical structure, and by their closely-knit unity of texture, shall appeal to our disinterested and contemplative imagination with something of the same vividness as the things of actual life appeal to our practical activities. In fact, they aim not at illusion but at reality. The logical extreme of such a method would undoubtedly be the attempt to give up all resemblance to natural form, and to create a purely abstract language of form — a visual music ; and the later works of Picasso show this clearly enough. They may or may not be successful in their attempt."

grande ênfase ao estudo da forma e do movimento utilizando-se de filtros e projetores controláveis por *sliders* para projetar figuras geométricas sobre uma tela, modificando sua forma, seu movimento e a intensidade da cor. (SANCHES, 2010, p. 7)



Figura 2 - Caricatura do órgão ocular de Louis-Bertrand Castel, por Charles Germain de Saint Aubin²⁴.

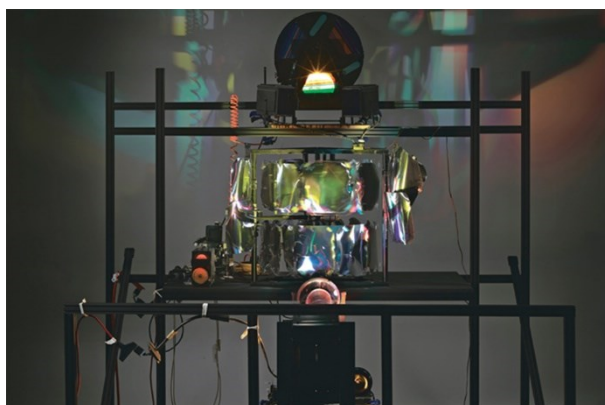


Figura 3 - Dispositivo de Wilfred²⁵.

²⁴ Fonte: https://commons.wikimedia.org/wiki/File:A_caricature_of_Louis-Bertrand_Castel%27s_%22ocular_organ%22.jpg

²⁵ Fonte: <https://www.artforum.com/print/reviews/201706/thomas-wilfred-68697>

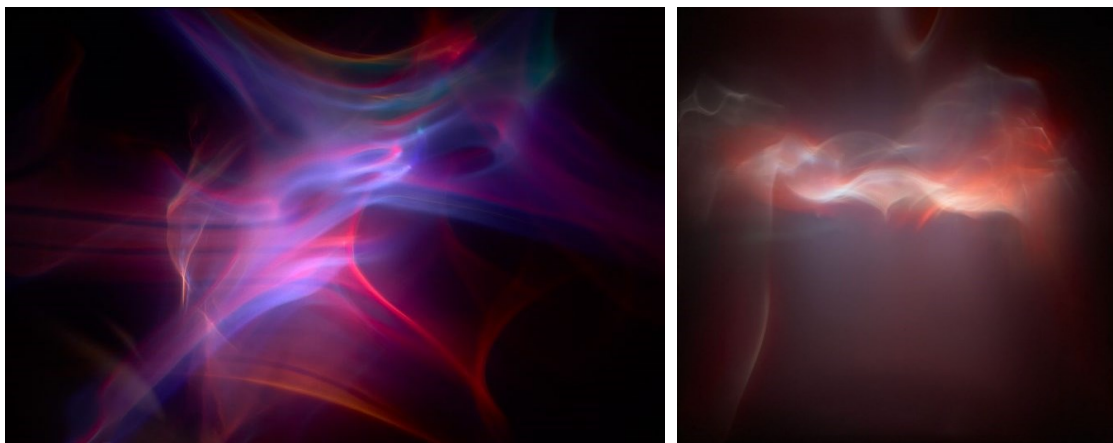


Figura 4 - Imagens obtidas pelo dispositivo de Wilfred²⁶.

O instrumento de Mary Hallock-Greenewalt relacionava variações de intensidade luminosa com variações sonoras, e o instrumento de Fischinger, por sua vez, gerava somente imagens ao ser operado por dois artistas. Uma das maiores questões levantadas por tais desenvolvimentos estão relacionadas ao fato de que “a associação arbitrária entre a cor de um acorde ou altura nos exemplos acima são altamente individuais e possuem pouca base lógica, além de talvez alguma forma de subconsciente ou condicionamento cultural”²⁷ (HILL, 2010, p. 38).

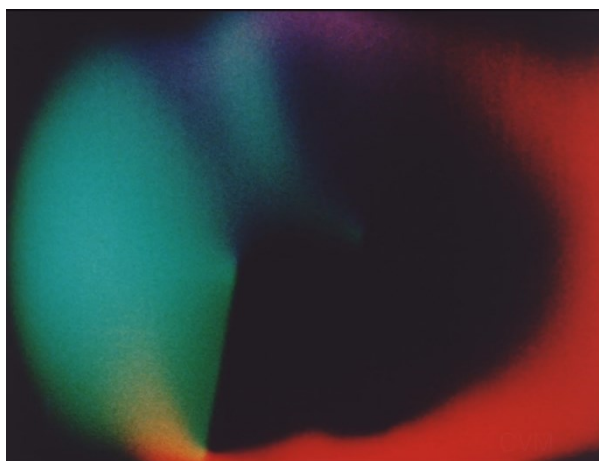


Figura 5 - Frame de um vídeo criado com o Lumigraph de Elfriede Fischinger²⁸.

A *Visual Music*, apesar de ter surgido em torno desses desenvolvimentos, chega ao século XXI mais próxima então do que Andrew Hill denomina de música

²⁶ Fonte: <https://www.theartblog.org/2018/01/thomas-wilfreds-early-influential-light-art-from-1919-rediscovered-in-lumia-at-the-smithsonian-american-art-museum/>

²⁷ Original: “The arbitrary association of colour to chord and colour to pitch in the above examples is highly individualistic and possesses little logical basis, apart perhaps from some form of subconscious or cultural conditioning.”

²⁸ Fonte: <https://vimeo.com/13787280>

eletroacústica audiovisual (HILL, 2010). Segundo o autor, uma característica fundamental desse tipo de manifestação está na associação entre imagem e som, e na sua criação a partir de ferramentas de manipulação digital dos mesmos, possibilitado pelo desenvolvimento tecnológico dos computadores. Dessa forma, ele situa também, tanto a performance de VJs²⁹ em boates, assim como o desenvolvimento de algoritmos para a visualização do espectro sonoro em uma mesma categoria. Em sua tentativa de classificar melhor seu objeto, assim como pretendemos fazer com o nosso, Hill desmembra as manifestações audiovisuais relacionadas com a *Visual Music*³⁰ em quatro categorias:

- A. A abordagem puramente visual da *Visual Music*, exemplificada pelas obras de Thomas Wilfred, Kandinsky e Klee. Obras que objetivam emular a música ou contém estruturas inspiradas pelas musicais, mas que não contém som em si.
- B. Composições visuais sobre uma música pré-existente como as primeiras obras de Orkar Fischinger, as interpretações artísticas de *Fantasia* de Walt Disney, ou videoclipes como os apresentados pela MTV.
- C. A composição do som e imagem informada pelas tradições da música cujos materiais são estruturados ao longo do tempo. Esta forma é definida aqui como música audiovisual por conter elementos tanto sonoros quanto imagéticos. O som e a imagem são considerados componentes igualmente unidos no contexto de uma obra e ambos são estruturados musicalmente.
- D. A síntese de materiais visuais através de algoritmos a partir do som e da representação do som visualmente. Essa categoria inclui softwares de visualização tais como os encontrados no *Windows media player*, osciloscópios e algoritmos computacionais que renderizam visualmente a forma de onda ou o espectrograma de um material sonoro.

²⁹ “*VJing* é a manipulação ou seleção de recursos visuais, da mesma forma que o *DJing* é uma seleção e manipulação de áudio. Um dos elementos-chave na prática do *VJing* é a mistura em tempo real de conteúdo de uma ‘biblioteca de mídias’. Além da seleção de mídia, o *VJing* implica principalmente em processamento em tempo real do material visual. O termo também é usado para descrever o uso performativo de software generativo.” Original: “*VJing* is the manipulation or selection of visuals, the same way *DJing* is a selection and manipulation of audio. One of the key elements in the practice of *VJing* is the realtime mix of content from a “library of media.” In addition to the selection of media, *VJing* mostly implies realtime processing of the visual material. The term is also used to describe the performative use of generative software.” (375 WIKIPEDIANS, 2010, p. 17f)

³⁰ Aqui o autor agrupa o que é denominado de *visual music*, *absolute film*, *cinedance*, *light shows*, *lumia*, *videomusique*, e *áudio-visual music* em gêneros que objetivam a exploração a estruturação de material visual de uma forma musical.

Assim como em Hill, nosso maior interesse está na categoria C, onde os artistas / compositores têm, através de aparatos tecnológicos digitais, a possibilidade de criar ambas as partes visual e sonora de suas obras audiovisuais concomitantemente, através de um processo criativo que transite entre os dois suportes, ao invés de serem concebidos de forma independentes. Dessa forma o artista tem potencial para um total controle do material, das ferramentas de manipulação e do resultado final de seu trabalho. Considerando isso podemos então enumerar as características principais do nosso objeto de estudos:

- A. Origem na música eletroacústica e afastamento do produto final de possíveis semelhanças com o produzido pela indústria do cinema ou de videocliques;
- B. Material audiovisual final gerado a partir da manipulação, pelo compositor, de materiais primários gravados ou sintetizados, através de recursos tecnológicos, na maioria das vezes no computador;
- C. Produto final fixado em uma mídia, possibilitando o maior controle do compositor do material apresentado ao ouvinte.

Com essas características podemos observar que, tanto a música vídeo quanto a vídeo música e a *Visual Music* são nomenclaturas válidas. Apesar disso, adotaremos o termo música eletroacústica audiovisual³¹ para definir e expandir o nosso objeto de estudos, pelo simples fato de este ser um termo mais descritivo e claro. Tanto o termo música vídeo quanto o vídeo música podem ser confundidos pelos videocliques se traduzidos para o inglês, o que pretendemos evitar, e a *Visual Music* possui, historicamente, uma grande ênfase no visual, o que também não desejamos.

1.2 Intermedialidade³² e os dois níveis das relações audiovisuais

Dentre as diversas justificativas apresentadas por Bethônico (2001) do porquê de se estudar o audiovisual, direcionamos nosso trabalho para o aprofundamento na técnica e estética de sua composição em busca de uma sofisticação tanto da tecnologia quanto do desenvolvimento da sua linguagem (tanto leitura quanto

³¹ Essa *música eletroacústica audiovisual* será reproduzida através de uma *mídia audiovisual adaptativa para realidade virtual*, objeto principal do nosso trabalho, englobando aplicações inclusive além da presente definição.

³² O termo intermedialidade vem aqui substituir o antigo termo intertextualidade, uma vez que o primeiro engloba o segundo de acordo com a literatura mais recente.

escritura). Com isso visamos beneficiar tanto a própria comunidade dos que produzem material audiovisual quanto dos que consomem esse material. Para alcançarmos essa sofisticação, precisamos direcionar parte dos nossos estudos para as possíveis relações entre som e imagem, buscando identificar o que poderia fazer com que um objeto audiovisual³³ (híbrido composto de um objeto sonoro acompanhado de um gesto visual) seja percebido como mais do que a soma entre suas partes.

Quando falamos sobre essas possíveis relações entre duas mídias (como definido por Elleström³⁴ e Wolf³⁵), ou até mesmo no processo analítico de obras que envolvam ambas, observamos duas possibilidades de abordagem: uma delas partindo de um ponto vista macro, e outra de um ponto de vista micro. Denominamos aqui de um ponto de vista macro, aquelas relações que envolvem duas obras distintas entre si, sejam elas desenvolvidas por autores diferentes ou simplesmente separadas no tempo e no espaço mas que dialogam entre si de alguma forma. Tal abordagem nos aproxima tanto dos principais de estudos mais abrangentes sobre intermedialidade (RIPPL, 2015. CLÜVER, 2006; ELLESTRÖM, 2017; WOLF, 1999), quanto do conceito de dialogismo desenvolvido por Baktin (FIORIN, 2006).

Já do ponto de vista micro, nos focamos nos elementos internos da obra e como eles se relacionam entre si, ou seja, uma abordagem da obra ou processo criativo a partir de suas características internas de estruturação. Tal abordagem pode ser exemplificada pelos trabalhos de Chion (1983, 1994, 2009) e Cook (1998), entre outros trabalhos sobre intermedialidade e semiologia, mas principalmente pelos métodos de análise sistemática abordados por Grela (1986), Bethônico (2001) e pelo modelo da análise tripartite de Nattiez (2014).

Essa nossa separação entre relações micro e macro, pode ser comparada à abordagem descrita no trabalho de Lafetá (2011, p. 40)³⁶, onde ele discursa sobre uma categoria de referenciação de uma mídia em outra, extraída dos trabalhos de

³³ O que denominamos, aqui, de um objeto audiovisual, se aproxima do que Lima denomina de “bloco audiovisual” (LIMA, 2011, p. 65).

³⁴ “Mídia significa *meio, intervalo, entre-espço*, e assim por diante. A definição padrão encontrada nos dicionários enfatiza que uma mídia é um canal para a mediação de informações e entretenimento (o *utile dulci* de Horácio), de forma a ser completamente possível incluir as formas artísticas entre outras mídias.” (ELLESTRÖM, 2017, p. 54)

³⁵ “[M]ídia” pode ser definida [...] como um meio de comunicação convencionalmente distinto, especificado não só por canais (ou um canal) de comunicação particular(es) mas também pelo uso de um ou mais sistemas semióticos que servem para transmitir mensagens culturais.” (WOLF, 1999; apud CLÜVER, 2006, p. 34)

³⁶ O trabalho de Lafetá (2011) explora as fronteiras entre as diferentes artes passando por um panorama sobre a intermedialidade e se propõe a estudar o hibridismo de Peter Greenway.

Cluver e Rajewsky. Segundo ele, uma referência intermediática acontece quando outra mídia está referenciada de forma alusiva (acontece entre mídias diferentes) enquanto que uma referência intramidiática faz referência a um texto ou mídia anterior, encontrado no mesmo tipo de mídia, seja através da repetição de uma técnica, citação, etc.

Com o intuito de exemplificar esse conceito de referenciação no nosso contexto, tomamos como exemplo um simples som de chuva no telhado. Se apresentado ao ouvinte como uma fonte sonora pontual, reproduzida por um sistema monofônico, pode se remeter a uma chuva em um local afastado. A mesma referenciação ao fenômeno meteorológico pode ser feita quando apresentamos o mesmo material sonoro através de um sistema multicanal que o envolva, como na instalação sonora *Promenade* de Dominique Gonzalez-Foester (Inhotim Collection, 2007), onde tal som é reproduzido através de 18 alto-falantes instalados no teto de uma sala extremamente reverberante, no entanto, sua percepção será completamente diferente, transportando o espectador ao local onde a chuva está acontecendo. Esse objeto sonoro pode ser manipulado pelo compositor de forma que leve o ouvinte do primeiro lugar, como observador de uma chuva distante, ao segundo, o imergindo no ambiente sugerido. Podemos dizer que, uma vez reconhecido o objeto sonoro, a associação ao evento meteorológico reconhecido (chuva) é uma referenciação intermediática, na qual o espectador associa o som ouvido ao fenômeno reconhecido, enquanto que o deslocamento realizado através da manipulação da apresentação desse objeto sonoro em um único alto-falante ou vários, se dá por uma referenciação intramidiática, realizada por uma comparação entre as duas situações supostas.

Apesar de um tipo de análise, seja ela ao nível macro ou micro, realizando referenciações intermediática ou intramidiática, respectivamente, não excluir completamente a outra, nosso direcionamento se dará predominantemente pela segunda via, a de Grela e Nattiez, ou seja, através de referenciações intramidiáticas, objetivando entender as relações internas dos elementos de uma criação plurimidiática³⁷. Ainda assim, processos descritos e considerados inerentes ao nível da análise macro, intermediática ou dialógica (FIORIN, 2006), são extremamente válidos para análises desenvolvidas para os elementos internos da obra. Com relação a tais processos, podemos associar facilmente, por exemplo, a transposição, a

³⁷ Criação composta de diversas mídias básicas como estabelecido por Elleström (2017).

justaposição, a combinação e a fusão, propostas por Arvidson (2007) em um nível intermediário, com os processos de justaposição, superposição e inclusão, apresentados por Grela (1986) em um contexto intramidiático.

Nos voltando para essas relações intramidiáticas, precisamos definir inicialmente o que é mídia. Nos estudos sobre intermedialidade sistematizados por Elleström (2017) podemos extrair, para um melhor entendimento e futura aplicação em nosso objeto, o que o autor explora como classificação de mídias, modalidades das mídias e aspectos qualificadores das mídias³⁸.

Nos iniciando com a classificação das mídias, observamos que o autor as divide entre as mídias básicas, as mídias técnicas e as mídias qualificadas. As mídias básicas³⁹, segundo Elleström (2017), seriam o som, a imagem e a palavra. Tal conceito é abordado também por Lúcia Santaella (2001), em seu livro “Matrizes da linguagem e pensamento”, onde ela descreve o que, para ela, seriam as três modalidades de linguagem reconhecidas pelo ser humano: a matriz visual, a sonora e a verbal.

A descrição de cada uma dessas matrizes da linguagem apresentadas por Santaella pode ser resumida da seguinte forma: a matriz sonora pode se consistir de voz, música e ruídos; por sua vez, a matriz visual é identificada através de imagens sintéticas, de registro, não-representativas, figurativas ou mistas, animadas, estáticas, editadas ou não; com relação a matriz verbal, apesar do trabalho da autora apresentá-la através de textos, no nosso objeto de estudos ela irá aparecer de formas mais ou menos sutil, seja através da apresentação do título da obra, através do conteúdo verbal presente nos próprios sons escolhidos pelo compositor, através de imagens que façam alusões a algum material textual, em qualquer língua que seja, ou através de textos que acompanhem a própria obra e direcionem a apreciação do ouvinte.

³⁸ Esses conceitos de aspectos qualificadores não serão aprofundados na presente tese. Resumidamente, Elleström (2017) descreve os aspectos qualificadores das mídias como sendo contextuais ou operacionais, tendo estreita relação tanto com a materialidade da mídia em si quanto com a sua referencialidade. Tal conceito só se mostra importante na presente pesquisa no nível em que as mídias das quais tratamos aqui são observadas em um contexto cultural específico e segue uma tendência de desenvolvimento técnico também específico da nossa época e cultura.

³⁹ Propomos nesse trabalho que o espaço extrínseco, como será descrito adiante, tanto visual quanto sonoro, em uma mídia audiovisual adaptativa para realidade virtual, possa ser analisado como uma mídia básica, que se expressa através de uma mídia técnica, no nosso caso, a mídia exibida em um ambiente de realidade virtual. Como exemplo desse tratamento do espaço como mídia, tomamos a técnica de criação de perspectiva em uma pintura: a profundidade, um dos parâmetros da percepção de espaço, observada em qualquer pintura que explore tal recurso ótico está estritamente ligada puramente a nossa percepção e conhecimento de mundo, sendo identificada como uma mídia básica, enquanto que, materialmente, do ponto de vista da mídia técnica, ela é inexistente.

Elleström (2017, p. 60), define como mídias técnicas⁴⁰ os objetos ou realizadores de uma mídia, ou seja, um livro, um CD, um arquivo.... Elleström afirma que "todas as mídias precisam de mídias técnicas para serem realizadas"⁴¹. É necessário aqui separarmos a materialidade da percepção das mídias, apesar de aparentemente impossível uma vez que só percebemos a mídia através de sua materialidade, com o intuito de definir quais qualidades ou características pertencem efetivamente aos aspectos materiais de uma mídia e quais deles fazem parte da nossa percepção (ELLESTRÖM, 2017, p. 57), ou seja, o que seria a mídia básica e o que seria a mídia técnica.

As manifestações audiovisuais que pretendemos estudar podem ser identificadas como o que Elleström denomina de mídias qualificadas, ou como uma obra híbrida, segundo Lafetá (2011) e Santaella, por integrar todas as três matrizes da linguagem proposta pela última, ou as três mídias básicas. Santaella propõe que

(...) o primeiro grande poder definidor da hipermídia está na hibridização das matrizes de linguagem e pensamento, nos processos sógnicos, códigos e mídias que ela aciona, e, conseqüentemente, na mistura de sentidos receptores, na sensorialidade global, (...) na medida mesma em que o receptor ou leitor imersivo interage com ela, cooperando na sua realização. (SANTAELLA, 2001, p. 391-392)

O trabalho de Julio Lafetá (2011) se aproxima dessa definição de mídias qualificadas de Elleström (2017), quando, em seu texto, a definição de intermedialidade aparece da seguinte forma: "como conceito, intermedialidade implica todos os tipos de inter-relação e interação entre mídias; uma metáfora frequentemente aplicada a esses processos fala de cruzar as fronteiras que separam as mídias" (LAFETÁ, 2011 p. 24)⁴².

⁴⁰ Encontramos autores que denominam tal tipo de mídia como suporte, mas preferimos a terminologia de Elleström para mantermos uma certa coerência com o restante da terminologia adotada.

⁴¹ Devido a sua natureza (mídia digital), e pela forma como ela se apresenta ao espectador (através de streaming), nosso objeto se aproxima também do conceito de hipermídia adotado e desenvolvido por Santaella (2001) e descrito por Winck (1997): "(...) a hipermídia consiste de informação digital, isto é, dados eletromagnéticos transcódificados numericamente num espaço a n dimensões. (...) Esse suporte físico, de caráter eletrônico, inaugura um espaço de representação ontologicamente infinito" (WINCK, 1997, p.3; apud SANTAELLA, 2001, p. 391).

⁴² Na histórica busca do homem pelo rompimento de fronteiras entre as mídias, temos como referência importante a ideia de obra de arte total (Gesamtkunstwerk) de Richard Wagner, onde música, teatro, canto, dança e artes plásticas coexistem. Essa mesma ideia reaparece em outras manifestações artísticas que envolvem diversas mídias, resumidas em som e imagem em movimento.

Além das definições de mídias básicas, qualificadas e técnicas, um dos conceitos mais importantes que buscamos no trabalho de Elleström (2017, p. 58) diz respeito às modalidades de mídia, definidas pelo autor como sendo “os fundamentos essenciais de todas as mídias, sem os quais a midialidade não pode ser compreendida e que, juntos, constroem um complexo midiático que integra a materialidade, a percepção e a cognição”. Segundo o autor essas modalidades, divididas entre modalidade material, modalidade sensorial, modalidade espaço-temporal e modalidade semiótica, permeiam uma escala entre os três níveis: material ou tangível, perceptual e cognitivo, ou conceitual. Todas as expressões artísticas e mídias, sejam elas básicas, qualificadas ou técnicas, são necessariamente realizadas nessas quatro modalidades, dentro das quais ele separa aquelas que seriam as pré-semióticas (as três primeiras) e a semiótica⁴³.

Tabela 1: Modalidades de mídias de Elleström (2017).

Modalidades de mídias	
Pré-semióticas	Material (material, perceptual, conceitual)
	Sensorial (material, perceptual, conceitual)
	Espaço-temporal (material, perceptual, conceitual)
Semióticas	Semiótica (material, perceptual, conceitual)

A modalidade material, segundo ele, seria aquilo que restaria de uma mídia se não existissem mais as pessoas para interpretá-las, e podemos citar aqui os corpos humanos, as superfícies planas ou tridimensionais, as ondas sonoras, o laser ou a luz como exemplos dessa modalidade.

Já a modalidade sensorial diz respeito às formas como uma mídia atinge o espectador, ou seja, através dos cinco sentidos, e, como as outras modalidades, é subdividida em três níveis, o nível dos dados sensoriais (o que a mídia atinge fisicamente no ouvinte), o nível dos receptores (como o corpo reconhece aqueles estímulos) e o nível da sensação (como o receptor interpreta aquele estímulo).

A modalidade semiótica se elabora no nível do significado, fruto da compreensão da mídia por um sujeito e sua visão de mundo. Vale a pena observar

⁴³ Esses modos se aproximam dos dois níveis que queremos estabelecer aqui para uma análise do material audiovisual, o sintático e o semântico, que será desenvolvido no primeiro capítulo.

aqui também uma questão referente ao fato de que, como todo som possui uma mesma interface material, constituída de movimentos provenientes de ondas sonoras que excitam nossos receptores auditivos, a forma de se separar um texto lido de sons não verbais está em como eles são analisados do ponto de vista da modalidade semiótica. Enquanto a compreensão de um texto falado se relaciona com a decodificação da linguagem em que foi elaborado, a compreensão de sons não verbais (ex.: música) se dará na interpretação de suas semelhanças e contiguidade, ou proximidade (ELLESTRÖM, 2017, p. 83).

Já a modalidade espaço-temporal consiste em “estruturar a percepção dos dados sensoriais da interface material em experiências e concepções de espaço e tempo” (ELLESTRÖM, 2017, p. 61). Segundo o autor, todas as mídias possuem aspectos nos dois modos (espaço e tempo), seja em um nível material ou em um nível semiótico. Dentro da modalidade espaço-temporal, Elleström (2017, p. 68) aponta três níveis da espacialidade na mídia, aquele nível em que o espaço se apresenta como característica inerente da mídia (material), aquele nível em que o espaço se apresenta como aspecto fundamental da cognição humana, ou seja, como o expectador recebe fisiologicamente os estímulos audiovisuais, e o nível em que o espaço é um aspecto interpretativo do que a mídia representa (este ele denomina de espaço virtual).

Para uma análise do espaço, tomemos uma pintura como exemplo: em um nível material podemos indicar as dimensões tridimensionais da pintura, 30cm x 60cm x 3cm, por exemplo, no nível perceptual podemos observar a perspectiva da imagem representada na tela, levando o receptor a reconhecer primeiro plano e planos de fundo, em um nível conceitual o receptor pode ter tido a sensação de estar imerso em uma outra realidade onde as imagens representadas na pintura se movimentam talvez pelas características do traçado, etc. Um bom exemplo de análise do espaço visual em uma pintura, e da transgressão dos níveis de espaço mencionados acima é desenvolvido por Michel Foucault ao descrever a tela “As meninas” do pintor espanhol Diego Velázquez.



Figura 6 - “Las Meninas” de Diego Velázquez⁴⁴.

No momento em que colocam o espectador no campo de seu olhar, os olhos do pintor captam-no, constroem-no a entrar no quadro, designam-lhe um lugar ao mesmo tempo privilegiado e obrigatório, apropriam-se de sua luminosa e visível espécie e a projetam sobre a superfície inacessível da tela virada. Ele vê sua invisibilidade tornada visível ao pintor e transposta em uma imagem definitivamente invisível a ele próprio. (...) Olhamo-nos olhados pelo pintor e tornados visíveis aos seus olhos pela mesma luz que no-lo faz ver. (FOUCAULT, 1999, p. 5-6)

O que queremos chamar a atenção com essa descrição é como essa modalidade espaço-temporal pode aproximar a percepção do receptor da modalidade semiótica, que é onde o receptor resgata seu conhecimento de mundo e suas experiências anteriores. Nessa pintura o papel do expectador se confunde com o papel de um elemento do quadro, ou seja, enquanto esperamos nos deparar com uma imagem imutável, representação de algo do mundo do pintor, nos vemos na verdade

⁴⁴ Fonte: www.museodelprado.es/coleccion/obra-de-arte/las-meninas/9fdc7800-9ade-48b0-ab8b-edee94ea877f

como um elemento da própria pintura, representada naquele quadro do qual vemos somente as costas. O objeto da pintura, ao qual está a atenção do pintor, seriam o rei e a rainha, posicionados exatamente no local em que se encontra o espectador, cujos reflexos podem ser observados em um provável espelho ao fundo da sala. O espaço representado na pintura se mistura com o espaço em que o espectador se encontra, no entanto isso ocorre somente em um nível semântico, proveniente do nosso conhecimento de mundo após a observação da obra em questão.

Essa compreensão da obra como resultado de uma análise da mesma pelo sujeito quando comparada ao seu conhecimento de mundo é mencionada por Elleström (2017, p. 69), que afirma que o mundo em si não possui significado algum, e que tal significado é produto de um sujeito que carrega consigo suas características sociais e cria suas ideias a respeito do que está sendo observado. Todo o significado em si, nesse nível semiótico, é produto da mente do sujeito e de sua interpretação individual ao atribuir sentido às coisas, ações, ocorrências e artefatos que observa.

Fazemos aqui um paralelo entre dois dos três níveis em que as modalidades de mídia de Elleström permeiam, o material e o conceitual⁴⁵, e as possibilidades de relacionamento entre duas mídias, apontadas por John Coulter (2007, p. 4). Ele recorre aos modos como a nossa percepção pode atuar, argumentando que os sons possuem um potencial duplo, o abstrato e o concreto (ou referencial) e que essas qualidades do som formam a base dos modos de escuta propostos por Schaeffer e Chion⁴⁶. Ele sugere também que a atenção do espectador pode ser direcionada para um desses modos (atenção seletiva) ou para ambos os modos (atenção dividida). De uma forma semelhante o autor categoriza a percepção visual, pautada no abstrato ou no referencial, e com isso estrutura quatro formas de pareamento das mídias audiovisuais: abstrato-abstrato, abstrato-referencial, referencial-abstrato e referencial-referencial.

⁴⁵ O nível perceptual estaria relacionado à fisiologia do expectador e ao sistema de reprodução ao qual ele é submetido, e não se encaixa na abordagem de Coulter.

⁴⁶ Dentro do contrato audiovisual descrito por Chion “a relação audiovisual não é natural mas sim um tipo de pacto simbólico no qual o espectador considera que os elementos sonoros e imagéticos participam de uma e da mesma entidade ou mundo” (CHION, 1994, p. 222). Nesse contexto o autor descreve três modos de escuta da música acusmática, uma delas, a escuta casual, consiste no reconhecimento das fontes sonoras, a segunda, a escuta semântica, envolve a decodificação dentro de uma abordagem da sensibilidade e da linguagem, e a terceira, a escuta reduzida, sintática, morfológica, se destina a escuta no som por si mesmo e sua estrutura, independentemente de causa ou significado.

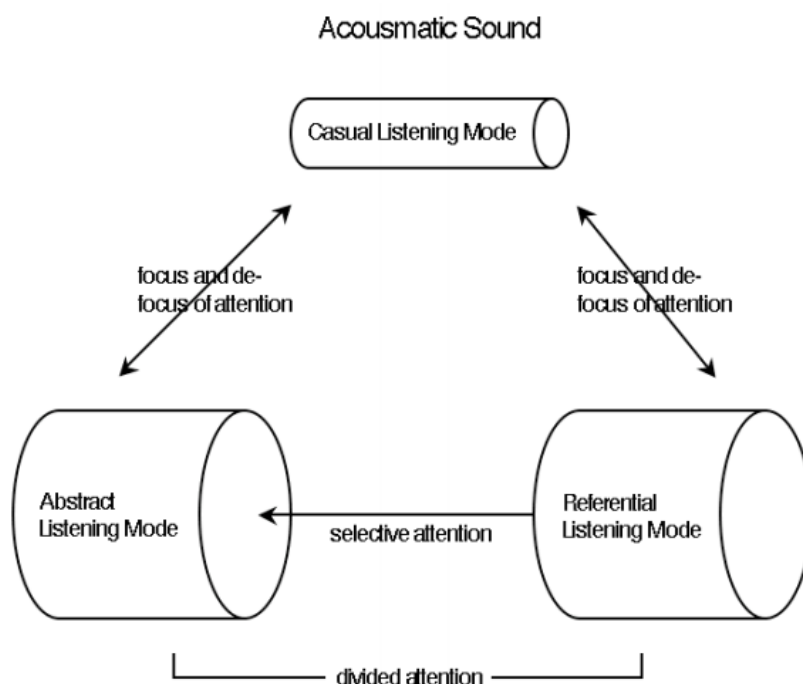


Figura 7 - Modos de escuta segundo John Coulter, baseado nos modelos de Schaeffer e Chion⁴⁷.

Aprofundando-se nas relações entre as duas modalidades de mídia, sonora e visual, o autor estabelece e descreve o que seriam relações concomitantes (que vão juntas, acompanhantes) e isomórficas (correspondentes em forma, na natureza e produtos de suas operações). Segundo ele, pares de mídia exibindo relacionamentos concomitantes são integrados através de um processo de realce e mascaramento; enquanto pares de mídia exibindo relações isomórficas contêm recursos que atuam como catalisadores no processo de integração. Essas qualidades incluem características formais compartilhadas, como duração, intensidade, proximidade e assim por diante (abstrações), bem como significados compartilhados (referências) (COULTER, 2010, p. 27).

Todos esses autores perceberam a dificuldade, e a necessidade, de se separar a análise do que chamaríamos de materialidade da mídia, de seu significado, mas é François Bayle (1986), que nos dá uma pista decisiva da existência dessa separação e de como ela se aproxima da semiótica, ou seja, direcionada para a percepção simbólica de um objeto sonoro e a imagem que ele carrega consigo. Em seu trabalho ele define o que seria o *I-Son*, um objeto que descreve como os sons na música acusmática podem se relacionar com um ícone, um arquétipo ou uma metáfora.

⁴⁷ Fonte: COULTER, 2007.

Quer o objeto seja encontrado ou fabricado, a imagem é ajustada à sua função de estimular a percepção, de se transformar em algo que seja reconhecível na realidade de acordo com três aspectos característicos: *icônico, indexador, simbólico*.⁴⁸ (BAYLE, 2007, p. 244)

Segundo Bayle (1986, p. 249), nenhum objeto sonoro é destituído de algum significado, mesmo que este seja criado ou retirado da realidade e destituído de sua origem e contexto imediato, tendendo a se tornar um objeto de linguagem. Ele eventualmente chegará a exercer uma função no imaginário do receptor nos indicando uma referência semiótica. Para ele a imagem, referência semiótica, é uma apresentação que pode ser uma coisa em si (corpórea, material) ou uma evocação (mental). Nos dias de hoje, nos quais podemos produzir tais imagens através da tecnologia devemos distinguir claramente uma configuração externa desse objeto-imagem da sua configuração mental pois todo som-imagem (*I-son*) se duplica sendo ao mesmo tempo um objeto e um fenômeno.

Esse conceito do som como um objeto de linguagem, introduzido por Bayle, é dividido pelo próprio autor em três níveis: um nível de imagem (*im-sound*), onde o som possui um referencial cru ou icônico; um nível de diagrama (*di-sound*), onde um índice diagramático é criado a partir da manipulação de filtragem e dilatação do objeto sonoro; e um nível metafórico (*me-sound*), onde tratamos do figurativo e auto-referencial, onde a metáfora lida diretamente com o conceito de signo.

De uma forma semelhante Bayle descreve o espaço, reafirmando que tanto ele quanto o tempo em si, podem ser redimensionados nesses três níveis, nos direcionando para a distinção e estudo dos dois níveis mais extremos, o físico e mensurável, e o fenomenológico. Em ambos os níveis, abstrato ou referencial, material ou semiótico, relações das mais diversas podem ser exploradas e, de uma forma geral, observamos nos trabalhos que envolvem um produto plurimiédiático uma coexistência entre esses dois níveis. Acharmos importante então, para uma clarificação conceitual, a separação desses dois níveis quando se tratando de música eletroacústica audiovisual principalmente pelo fato de que cada compositor tende a criar seu próprio repertório de relações, nem sempre acessível ao receptor.

No que diz respeito às relações entre o som e a imagem, Bethônico (2001, p. 124) nos apresenta três possibilidades: a comparação, a montagem e a tradução. A

⁴⁸ Original: "Whether the object is found or fabricated, the image is adjusted to its function of stimulating perception, of turning to it a face that is real according to the three characteristic aspects of all meaningful tracing: *iconic, indexive, symbolic*."

primeira delas, a comparação, coloca as características de um código perante as do outro, e dessa forma se reconhecem suas semelhanças e diferenças; a segunda forma de relação seria uma montagem, na qual ambos os signos atuam em conjunto de forma integrada na percepção do espectador; a terceira, a tradução, se baseia em associações diversas construídas em um nível conceitual, sejam elas criadas pelo compositor (na escritura) ou pelo expectador (na leitura). (BETHÔNICO, 2001, p. 125).

Em um processo de comparação, as analogias são identificadas pela imaginação, reconhecimento de similaridades e diferenças, baseadas no conhecimento do mundo real. As comparações no entanto podem atingir o limite da racionalidade, sendo uma relação pessoal, subjetiva, “abalando o pensamento institucionalizado” (BETHÔNICO, 2001, p. 132-133). Elas operam cognitivamente a partir da identificação dos termos envolvidos, mapeamento da fonte em relação ao fenômeno em observação, avaliação das inferências sugeridas pelo mapeamento e pelo aprendido, e podem ser delimitadas pela similaridade, estrutura e propósito. Uma relação através de montagem ou tradução, por outro lado, irá considerar relações entre signos, seus objetos e significados, com o objetivo de estudar mudanças e translações que o conjunto audiovisual sofre, focando a análise na gramática desses eventos, na escritura do audiovisual, quando existente (BETHÔNICO, 2001, p. 140).

De certa forma, além dessas relações, o compositor tem nas mãos uma outra ferramenta que seria a de criar suas próprias relações e “ensinar” o ouvinte a percebê-las e reconhecê-las. Segundo Bethônico, essa “articulação gestáltica”, que pode se apresentar em diversos níveis de força, depende de uma “pregnância” para funcionar como linguagem e sugerir um sentido ao ouvinte:

A articulação gestáltica da composição é responsável pela diferenciação das forças e pela predominância de umas sobre outras: a segregação entre imagem e som é tanto mais forte quanto menor for a sincronia (lei da proximidade) e a possibilidade do significado de uma das linhas sobressair-se ao da outra depende de sua pregnância. Não só as relações do sentido de um fragmento com o repertório do leitor fortalecem sua pregnância. Também a articulação da obra pode valorizar um sentido: por continuidade (quando já surge previamente no contexto da obra); por proximidade (quando está temporalmente perto de várias referências deste sentido); por semelhança (quando há identificação por analogia com alguma referência na obra ou na cultura) e por segregação (quando por contraste faz-se sobressair o fragmento responsável pelo sentido). (Bethônico, 2001, p. 226)

Um outro autor importante que pretende sistematizar a análise das relações entre os signos sonoros e visuais - no nosso ponto de vista um pouco simplista mas

extremamente útil para análises menos aprofundadas - é Nicholas Cook (1998), que estabelece três formas de relação entre o som e a imagem: (1) conformidade - quando os meios operam em concordância; (2) controvérsia - quando entram em contradição; e (3) complementação - quando operam entre as duas primeiras.

Pretendemos aplicar essas mesmas propostas de relação aqui apresentadas entre o som e a imagem, do ponto de vista material e conceitual, ao que denominamos de espaço extrínseco e, para isso, adotaremos uma divisão que iremos desenvolver mais à frente, e que denominamos de *abordagens sintática e semântica do espaço* em obras tridimensionais, guiados pela ideia de que o espaço de um objeto audiovisual também tem o potencial de assumir essas duas diferentes funções, a de um ícone (sintaxe do espaço), a de um símbolo (semântica do espaço), e onde o índice seria a própria percepção do espectador associada ao seu conhecimento de mundo e experiências anteriores. Passaremos a nos referir então a essas duas abordagens do espaço do material audiovisual de *sintaxe* (abordagem sintática), o ponto de vista material, e *semântica*, o ponto de vista dos significados. A abordagem sintática estará sempre relacionada ao nível da natureza do material, do imediatamente perceptível, enquanto a semântica irá se encontrar no nível da decodificação do material pelo conhecimento e histórico do espectador, inserido em uma prática pragmática ou apresentada pelo compositor / artista.

1.3 Audiovisual adaptativo: imersão através da interatividade

Na música, uma parte do que denominamos música eletroacústica se direciona para a composição de obras que se estruturam a partir da interação de um instrumentista (o som emitido por ele e seus gestos captados pelos mais diversos tipos de sensores) com um software específico. Costumamos nos referir a esse subgênero como música eletroacústica com processamento em tempo real, *live electroacoustic music*, ou simplesmente *live electronics* (MENEZES, 1996, p. 247). Como características principais está o desenvolvimento de um *patch*, software programado comumente no MAX/MSP ou Pure Data para ser executado durante a performance. Com isso a obra apresenta algumas características de imprevisibilidade, provenientes da resposta do software às entradas fornecidas pelo músico. Essa imprevisibilidade já foi historicamente considerada por alguns compositores como um ponto forte de uma performance eletroacústica (MANOURY, 1988), mas também pode ser considerada o ponto fraco dessa vertente, uma vez que o compositor deixa de ter controle absoluto sobre o resultado de sua obra.

Dentre as mais diversas possibilidades de atuação do computador na música, Drummond (2009, p. 132) aponta aquelas exploradas pelos primeiros praticantes que ainda se mostram presentes nos dias de hoje: o “conceito de controle compartilhado, instrumentos inteligentes, ambientes de conversação colaborativa e a indefinição das distinções entre a construção de instrumentos, performance, improvisação e composição”⁴⁹. Para o autor o termo “interativo” é aplicado nas mais diversas áreas que envolvem as novas mídias, de sistemas de mapeamento reativos mais simples com relações diretas entre entrada e som de saída a sistemas baseados em inteligência artificial que são capazes de aprender e se comportar de maneira autônoma, orgânica e intuitiva. O autor aponta também a exploração de sistemas interativos no desenvolvimento de instrumentos musicais digitais que realizam um mapeamento entre gesto e resultado sonoro, assim como para composições interativas, ambientes colaborativos e modelos de conversação.

Segundo Rowe (1993, p. 1), um sistema musical interativo seria aquele cujo comportamento se altera em resposta a uma entrada, ou estímulo, musical. Ao

⁴⁹ Original: “the concept of shared control, intelligent instruments, collaborative conversational environments, and the blurring of the distinctions between instrument building, performance, improvisation and composition”.

contrário do definido pelo autor, nosso objeto de estudos não é um sistema musical interativo em sua conceituação, mas seria talvez um sistema de reprodução audiovisual interativo. No nosso caso a entrada não é musical mas sim os dados de orientação da cabeça do espectador, que altera o que é apresentado em seu campo de visão, adaptando também a reprodução sonora, segundo sua orientação. Drummond (2009, p. 1), por sua vez, afirma que um sistema interativo teria o potencial para variação e uma certa imprevisibilidade em sua resposta, de acordo com o contexto. Nosso objeto também não se enquadraria completamente em tal definição uma vez que não existe imprevisibilidade em sua resposta, mas sim uma adaptação às limitações naturais da percepção humana, permitindo que o ouvinte oriente sua escuta e campo de visão em um ambiente tridimensional. Apesar das diversas outras possibilidades de interação possíveis através do giroscópio de um smartphone em um óculos de VR, objetivamos uma criação audiovisual em si que é completamente finalizada pelo compositor / artista, e na qual o que se altera a partir das informações de posição do espectador é a porção do ambiente tridimensional apresentada ao ouvinte, interagindo com seu movimento de cabeça, intensificando o que chamamos de imersão⁵⁰.

No entanto, um conceito mais próximo, que poderia ser aplicável ao nosso objeto, a mídia audiovisual adaptativa para realidade virtual, seria o de arte de mídia interativa⁵¹, desenvolvido por Katja Kwastek. Segundo ela, a arte de mídia interativa combina elementos tanto das artes visuais, quanto performáticas e das artes baseadas no tempo⁵² (KWASTEK, 2013, p. XVIII). Esse tipo de arte híbrida consiste de obras em que o espectador é coautor das mesmas, a contemplação e a ação não se excluem mas, ao contrário, devem ser consideradas mutuamente aspectos de uma experiência estética. Nesse tipo de obra, para fazer um paralelo às artes performáticas, o espectador se torna o performer através de uma mediação tecnológica, ou seja, de processos que respondem às suas ações mediadas tecnicamente. O conceito de interatividade adotado inicialmente por Kwastek diz respeito a relação de duas coisas, pessoas ou sistemas, independentes em si, que se relacionam um com outro. Esse conceito está estreitamente relacionado a uma via de

⁵⁰ Conceito relacionado principalmente à indústria de jogos, explorada na seção 2.4.2.

⁵¹ Original: *interactive media art*

⁵² Original: *time-based arts*

mão dupla, ou seja, de uma reciprocidade, na qual, o aparato técnico responde ao espectador e vice-versa.

Ela elabora também uma diferenciação entre os termos interação e interatividade, no qual o primeiro descreve um processo que realmente acontece e o segundo um processo em potencial. As características principais da interação podem ser identificadas nas trocas e presença em tempo real, no controle e resposta do sistema, e nos processos de seleção e interpretação (KWASTEK, 2013, p. 7). Inicialmente, obras com algum nível de interação eram comumente denominadas de responsivas ou reativas, e a diferenciação entre arte de mídia interativa e arte interativa surge da necessidade de diferenciar obras baseadas em processamento computacional ou não, respectivamente.

A autora aponta o questionamento dos artistas do século XX com relação a arte direcionada ao objeto em contraponto a arte orientada ao processo ou evento, valorizando os resultados estéticos do acaso, muitas vezes incluindo também algum fator de automatização na obra. Kwastek elabora também que:

Embora um dos objetivos da arte moderna e contemporânea tenha sido desafiar sua restrição tradicional à representação, a representação de situações espaciais e a representação de processos temporais permanecerão sempre possíveis campos de interesse das artes. Por meio da tecnologia interativa, as representações espaciais e temporais podem ser organizadas de novo - por meio de ilusões de espaço imersivas e formas lineares de narrativa⁵³ (KWASTEK, 2013, p.35)

A inclusão do espectador como coautor ainda veio depois do vislumbre da utilização do próprio espaço no processo composicional, como podemos observar em obras como as pinturas hipersensitivas completamente brancas de Robert Rauschenberg. Em tais obras as sombras projetadas na tela pelo público em si, dependente da iluminação específica da hora do dia, eram o objeto de contemplação. De forma semelhante a obra 4'33" de John Cage obriga o ouvinte a perceber auditivamente o espaço que o rodeia.

Se aproximando um pouco mais do nosso objeto de estudos, Kwastek enumera dois desenvolvimentos, relativamente recentes e relativos às representações de ilusão

⁵³ Original: "Although one of the goals of modern and contemporary art has been to challenge its traditional restriction to representation, the representation of spatial situations and the representation of temporal processes will always remain possible fields of interest of the arts. By means of interactive technology, spatial and temporal representations can be newly organized—through immersive space illusions and alinear forms of narrative."

tradicionais, que permitem que o receptor explore uma obra ativamente: o primeiro sendo a possibilidade de acentuação de efeitos imersivos e o segundo sendo a processualidade dos espaços simulados. Deles, o primeiro se aproxima mais do nosso objeto uma vez que todo o processo composicional já foi finalizado pelo compositor, restando ao expectador explorar o resultado com uma certa imersão.

Tendo isso tudo em vista, e observando que existem diferentes níveis de interação, retomamos a definição de Santaella que, a partir da sua descrição das três matrizes de linguagem e pensamento, se apropria e explora o conceito de hipermídia como uma manifestação artística computacional que explora tais matrizes em sua plenitude e que prevê a interatividade, onde “quanto maior a interatividade, mais profunda será a experiência de imersão do leitor, imersão que se expressa na sua concentração, atenção e compreensão da informação” (SANTAELLA, 2001, p. 394).

Muitas manifestações audiovisuais não se enquadrariam no conceito da autora de hipermídia, por não apresentarem qualquer interatividade com o ouvinte. De certa forma, nosso objeto de estudos, a música eletroacústica audiovisual para realidade virtual, poderia se encaixar em tal conceito uma vez que, além de objetivar a hibridização de linguagens, ainda inclui a interatividade, decorrente da adaptação tanto do áudio quanto do vídeo, conforme a movimentação da cabeça do ouvinte.

Steuer (1992) por sua vez, estabelece que a interatividade é a “extensão em que os usuários podem participar modificando a forma e o conteúdo do ambiente mediado em tempo real”⁵⁴ (STEUER, 1992, p. 84), e estabelece três fatores contribuintes importantes nessa interatividade: a velocidade (taxa em que as informações inseridas pelo usuário são assimiladas pelo ambiente), a amplitude (número de possibilidades de ação do usuário) e o mapeamento (habilidade do sistema de modificar o ambiente de forma natural). A ideia de mediação em tempo real é para nós um dos princípios mais importantes para a análise do nosso objeto, assim como o conceito de interruptibilidade (PRIMO, 2019, p. 5), que consiste na possibilidade de que o processo seja interrompido sem precisarmos esperar que algo aconteça, comumente, em outros meios, uma linha de comando ou um botão de desliga.

⁵⁴ Original: “Interactivity is the extent to which users can participate in modifying the form and content of a mediated environment in real time.”

No contexto da produção da denominada *Visual Music* podemos mencionar como uma proposta que envolve interatividade, a de Terry Trickett⁵⁵, que cria imagens a partir do sinal captado de seu clarinete e de outros mecanismos. Segundo ele, “ao combinar o desenvolvimento de softwares e o clarinete, eu crio [ele cria] performances que demonstram uma estreita afinidade entre a música e a imagem visual”⁵⁶, uma modalidade artística que ele denomina de *Visual Music*, apesar de diferir bastante de outras manifestações, já vistas, que utilizam a mesma terminologia.

Outra proposta que envolve a produção de material visual em tempo real, estabelecendo uma certa interação, aqui sem muita participação do espectador, pode ser observada nos aplicativos de visualização de áudio, que, comumente, recebem um áudio qualquer (arquivo de áudio escolhido pelo usuário) e transforma parâmetros de um vídeo a partir das informações de intensidade do áudio. Podemos citar aqui o conhecido visualizador de áudio do *Windows Media Player* e outros pacotes de visualizadores tais como o *G-Force*, o *Morphyre*, o *Pulsing Colors*, entre outros⁵⁷.

Experiências muito semelhantes mas que trazem algum interesse a mais pela escolha do algoritmo relativo a música, ou até mesmo apresentando algumas composições feitas especialmente para um determinado algoritmo de visualização podem ser observadas nos trabalhos de Joshue Ott⁵⁸ (vídeo) com Garth Stevenson (música), ou Giovanni Muzio⁵⁹. Esses trabalhos, apesar de ainda interagirem, muitas das vezes, somente com a amplitude do sinal de áudio, já apresentam uma certa elaboração. Eles costumam ser desenvolvidos através do software *Processing*, dedicado ao processamento de imagens em tempo real. Atualmente podemos observar também esse mesmo tipo de algoritmo em softwares dedicados a VJs, como o *Resolume Arena*, extremamente popular no meio da música eletrônica dançante. Outros trabalhos de interação do vídeo com o áudio podem ser desenvolvidos a partir do software *VVVV*, mais direcionado para a arte generativa como pode ser observado no trabalho de Lasal⁶⁰, Udo Kaiser⁶¹ ou Tasvisuals⁶².

⁵⁵ <http://www.trickettimages.com/>

⁵⁶ Original: “It shows how, by combining computational design and clarinet playing, I create performances that demonstrate a close affinity between music and visual imagery.”

⁵⁷ <https://support.microsoft.com/pt-br/help/17878/visualizations-for-windows-media-player>

⁵⁸ <https://www.youtube.com/channel/UC8BoYhg4n7qhHmk6c6HclcQ>

⁵⁹ https://www.youtube.com/channel/UCB0_Xos5U8R_rqvusG49tBg

⁶⁰ <https://www.youtube.com/channel/UCVIXMwQYBwfibYvdJvijk7A>

⁶¹ <https://www.youtube.com/channel/UC7VbfRT0ycGamtB6HjtVbQw/videos>

⁶² <https://www.tasvisuals.com/>

Não podemos deixar de mencionar aqui ferramentas que geram tanto o áudio quanto o vídeo a partir da ação do usuário, através de aplicativos, que podem ser considerados potenciais instrumentos audiovisuais para performances ao vivo. Nessa categoria podemos mencionar entre os mais significativos o *Mazetools*⁶³ e o *Patatap*⁶⁴.

É consciente nossa escolha, entre as infinitas possibilidades do que vemos em sistemas musicais interativos, entre outros, da manipulação de somente uma dessas diversas dimensões interativas, a rotação de campos sonoros em *Ambisonics*, dentre vários outros explorados em outros contextos e níveis de interatividade. Esse recorte se deve principalmente ao fato de estarmos nos direcionando ao estudo de uma aplicação específica, onde o formato final objetivado é um vídeo a ser transmitido via *streaming* que, por enquanto, nas plataformas atuais, suporta somente a possibilidade de manipulação desse parâmetro. Nada impede, no entanto, devido à possibilidade de transcrição⁶⁵ de uma criação audiovisual tridimensional com material sonoro realizado em *Ambisonics*, que o compositor desenvolva versões de sua obra para outras aplicações, sejam elas instalações sonoras, exposições em salas de concerto ou cinema.

No contexto do nosso objeto, mídias para realidade virtual, o termo interatividade aparece presente no trabalho de Begault (2000), destinado a uma análise sistemática do áudio para o VR. Em sua pesquisa direcionada para a criação de ambientes tridimensionais para treinamento de pilotos, afirma que o som de uma mídia audiovisual tridimensional é uma importante via de retorno que pode ser redundante a um evento visual ou pode fornecer retorno referente às ações ou situações que estão fora do campo de visão do ouvinte⁶⁶. O autor menciona também, várias vezes, a palavra interatividade na sua descrição de um sistema de simulação de realidade virtual (ou sistema multimídia, como é denominado algumas vezes), incluindo uma máquina de realidade (*reality engine*) que teria a função de receber sinal dos sensores e adaptar as ilusões visuais, auditivas e táteis, para o receptor.

Segundo Begault (2000, p. 12), o primeiro sistema de realidade virtual, projetado com algum nível de interatividade, teria sido o *Convolvotron* (WENZEL,

⁶³ www.mazetools.com

⁶⁴ <https://patatap.com>

⁶⁵ O termo *transcrição* foi adotado também pelo presente autor em sua dissertação de mestrado (SOUSA, 2010b) para descrever o processo de transformação de um áudio formatado especificamente para um tipo de sistema para um outro tipo de sistema de reprodução. Parte da literatura denomina esse processo também de *Upmix* ou *Downmix*.

⁶⁶ Chion chamaria tais sons de *off-screen*.

WIGHTMAN e FOSTER, 1988), com capacidade para processamento do áudio em formato *Binaural*, que utilizava muito do processamento computacional disponível, exigindo que o áudio fosse codificado, degradando-o e o fazendo menos convincente. Além da importante função de imergir o receptor, vastamente explorada pela indústria dos jogos digitais, e que pode ser observada também no material audiovisual para VR, esse áudio tridimensional adaptativo que estudamos possui outras aplicações.

Como já brevemente mencionado, o processamento para reprodução de uma mídia audiovisual adaptativa para realidade virtual, através do qual iremos objetivar a interatividade entre o ouvinte e o material audiovisual, se baseia não na geração de material audiovisual em si, mas sim na manipulação de matrizes de áudio codificadas em *Ambisonics*, a partir da movimentação do ouvinte ao observar um ambiente virtual em três dimensões, assim como o que é apresentado em seu campo de visão.

1.4 O desenvolvimento das tecnologias de espacialização sonora para produções audiovisuais

Um áudio realmente de alta qualidade irá fazer as pessoas dizerem que os jogos possuem boas imagens, no entanto boas imagens não farão com que o áudio soe melhor, elas o farão soar pior. Logo, no mundo virtual que nós estamos construindo, uma das maiores preocupações deve ser um ambiente sonoro rico.⁶⁷ (LAUREL, 1993 apud BEGAULT, 2000, p. 10)

Pretendemos aqui descrever um pouco sobre o desenvolvimento do áudio multicanal que foi utilizado por artistas audiovisuais, do cinema aos jogos digitais, dando mais ênfase aos formatos que são de maior interesse e importância para o áudio utilizado na realidade virtual.

1.4.1 O áudio espacializado para cinema

Apesar de os primeiros experimentos com áudio multicanal (ex.: *Fantasound* de 1940) tentarem incluir um canal surround aos dois canais do padrão estereofônico já consolidado, a implementação de um sistema para aplicações comerciais teve seu início provavelmente diante do paradoxo do sistema estéreo apresentar um "buraco"

⁶⁷ Original: "Really high-quality audio will actually make people tell you that games have better pictures, but really good pictures will not make audio sound better, they make sound worse. So in the (virtual) world we're building, one of the features is a rich auditory environment."

entre as caixas (L e R) proveniente da distância entre as caixas, cada uma situada de um lado da tela, e o próprio tamanho da tela do cinema, fazendo com que esse “buraco” seja ainda mais perceptível. Devido a peculiaridade de, no cinema, o áudio ser apresentado para um grande público, se mostra comum o fato de um espectador perceber o som vindo mais de um lado do que do outro, devido a sua proximidade a um determinado alto-falante.

No entanto, apesar desse fenômeno existir (muito percebido também em sistemas de sonorização ao vivo de grandes escalas), nosso cérebro possui certa capacidade de ignorar tais problemas. Begault (2000, p. 12), discorre sobre esse fenômeno e o denomina de *visual capture*⁶⁸, que seria a habilidade do ser humano de associar um som a uma imagem mesmo que eles não apresentem um local do espaço exatamente coerentes - ele fala de associarmos o som vindo de um alto-falantes à boca do personagem ou a um carro que passa mesmo que eles não estejam posicionados exatamente no lugar onde a imagem é apresentada.

Historicamente, enquanto alguns desenvolvedores se voltaram para o sistema quadrifônico que, nos anos de 1970, encontrou dificuldades de se mostrar relevante devido a diversos fatores, entre eles sua falta de homogeneidade, outros se direcionaram para a tela do cinema, com o intuito de minimizar problemas relacionados a uma insuficiência da ilusão de imagens fantasma⁶⁹, por exemplo, criadas no centro dos dois alto-falantes.

A inclusão de um terceiro canal entre os canais L e R para o qual é dedicado a maior parte do diálogo na estética cinematográfica, objetiva conseguir fixar a voz dos personagens da narrativa ao centro, independentemente da posição do ouvinte, e foi implementada nos sistema 5.1 e posteriormente no 7.1 (HOLMAN, 2008, p. 57). Com o desenvolvimento das tecnologias de áudio multicanal, foram adicionados canais traseiros, inicialmente um único, eventualmente multiplicado em diversos alto-falantes e, posteriormente, dois canais surround, possibilitando a criação de imagens fantasma e movimento também atrás do ouvinte. Um sexto canal direcionado aos efeitos de

⁶⁸ Chion em 2009 denomina esse mesmo fenômeno de *spatial magnetization*, “um fenômeno psicofisiológico universal que funciona de forma mais eficaz quando os sons são sincronizados com as imagens.” (CHION, 2009, p. 248)

⁶⁹ Chamamos de *imagens fantasma* o recurso sonoro que nos permite posicionar ou localizar fontes sonoras provenientes de algum ponto entre dois alto-falantes. Tal fenômeno é resultante da interação dos alto-falantes envolvidos e se mostra menos confiável do que quando posicionamos uma fonte sonora exatamente no local de um alto-falantes. Apesar disso é um recurso extensamente explorado e base de muitos sistemas multicanais como o VBAP, por exemplo.

baixas frequências também foi implementado, transformando o cinema em uma experiência não somente audiovisual mas também sensorial.

Apesar de muitos acreditarem que o formato 5.1 do cinema se tornou um padrão por se apresentar superior, sistemas multicanais que pretendem oferecer uma experiência completamente imersiva ao espectador devem ser homogêneos e coerentes (MALHAM, 1998, p. 174). Segundo Malham, tal sistema não deveria apresentar buracos entre os alto-falantes, o que é possível somente ao posicioná-los a um mesmo ângulo entre si, igual ou inferior a 60°, e utilizando alto-falantes *full range*, ou seja, que respondam a toda a faixa dinâmica audível.

O cinema nunca se propôs a fornecer um sistema de áudio homogêneo e coerente, mas sim a impressionar o ouvinte que, tendo sua atenção fixada na tela através dos diálogos, eventualmente surpreendido por efeitos especiais extremamente graves ou se movimentando ao seu redor, alcança seu objetivo com um sistema de três canais frontais, dois traseiros e um *subwoofer*, o consolidado 5.1

Sistemas efetivamente homogêneos como o *Ambisonics* ou *Wave Field Synthesis*, nunca se tornaram formatos finais comerciais pelo fato de não terem recursos técnicos e financeiros para seus desenvolvimentos para o cinema, mas foram extremamente utilizados e explorados nos nichos de música eletroacústica e instalações sonoras. Com o desenvolvimento dos jogos digitais e da produção de material audiovisual para a realidade virtual, esses formatos finalmente emergem no mercado.

1.4.2 O áudio espacializado para jogos digitais

A indústria de jogos foi a que mais se desenvolveu com relação ao áudio tridimensional e impulsionou o aparecimento de tecnologias direcionadas à realidade virtual. O principal motivo para esse desenvolvimento diz respeito à necessidade de imersão do jogador na trama do jogo, o que é obtido através da implementação de um sistema de reprodução sonora imersivo e interativo.

Segundo Huiberts (2010), são diversos os conceitos relacionados ao fenômeno de imersão, muitas vezes este sendo definido como uma “experiência de perder a noção de corporeidade no presente enquanto concentrado em um ambiente

mediado⁷⁰” (DOVEY e KENNEDY, 2006, p.146), ou como onde “a pessoa é envolvida no sentimento de isolamento do mundo real⁷¹” (PATRICK et al., 2000, p. 2). Sem dúvida as áreas que mais estudam a imersão é a de desenvolvimento de jogos e a de mídias digitais, sendo ela considerada por muitos o Santo Graal dos jogos digitais (ROLLINGS e ADAMS, 2003), e o estado em que o indivíduo está completamente concentrado e absorvido pelo jogo. Denominado por outros autores também de incorporação (CALLEJA, 2007, p. 88) ou fluxo (CSIKSZENTMIHALYI, 1997), esse estado é alcançado quando o jogador tem a experiência de estar participando do evento, desconectado do mundo real, e é um dos objetivos de um bom desenvolvedor de jogos, uma vez que seu produto será, inclusive, avaliado pelos próprios jogadores através desse parâmetro.

Os autores descrevem duas principais formas de interatividade entre o som do jogo e o jogador. A primeira delas, mais explorada, diz respeito a resposta sonora que o jogador tem a partir de suas ações. Um exemplo simples seria, em um jogo de tiro, o jogador ouvir o som do tiro ao disparar sua arma. Outros exemplos podem ser observados também quando a música de fundo muda quando o jogador se aproxima de algum perigo ou de algum inimigo.

Essa primeira estratégia de interatividade objetiva a criação de uma imersão maior a partir do áudio e tem sido desenvolvida a partir da criação de diferentes faixas de áudio sintetizado em tempo real ou pré-gravadas e que respondem de forma adaptativa. Com o intuito de seguir uma ação na tela em constante mudança, uma música e o som que podem se adaptar à experiência do jogador tem sido integrada ao mecanismo do jogo e se estabelecendo rapidamente como um novo padrão (MARKS, 2009, p. 6). Essa interatividade do som, em resposta aos gestos e percursos definidos pelo jogador, tem a principal finalidade de aumentar a imersão do jogador.

A segunda forma de *interação* diz respeito ao espaço sonoro em que o jogador está imerso e em como ele responde quando o mesmo atua ao longo do jogo. Fontes sonoras localizadas a esquerda por exemplo, passam a ser vistas e ouvidas a frente do jogador quando ele desloca seu olhar para ela. Sons provenientes de uma fonte sonora observada a frente ao longo de um percurso por exemplo passam para trás quando o jogador passa por elas e segue em seu percurso.

⁷⁰ Original: “(...) the experience of losing a sense of embodiment in the present whilst concentrating on a mediated environment (...).”

⁷¹ Original: “(...) one in which a person is enveloped in a feeling of isolation from the real world (...).”

Essa segunda forma de interação é a que mais interessa a esse trabalho. A *Blue Ripple Sound* foi a primeira empresa reconhecida pelo mercado a desenvolver plugins comerciais para manipulação do áudio tridimensional. Lançou seu primeiro plugin em 2013, direcionado para a indústria dos jogos, uma vez que os mesmos ofereciam ótimo poder de processamento em tempo real e renderização em 3D antes mesmo dos engenheiros de som começarem a trabalhar com sistemas *Ambisonics*⁷². Antes do desenvolvimento de tais plugins aplicados à realidade virtual, tais formatos, até pouco tempo, eram explorados quase que unicamente em ambientes acadêmicos, experimentais e para fins de pesquisa, como nos já mencionados trabalhos da compositora britânica Natasha Barrett, do compositor irlandês Enda Bates, entre diversos outros.

Em um esquema do funcionamento de um dos seus plugins, podemos observar claramente como tal adaptação se desenvolve. No estúdio de criação o engenheiro de som cria uma mixagem multicanal, seja ela em 5.1, 7.1, sistema octofônico, com ou sem informações de altura, e o codifica em *Ambisonics* de primeira, segunda ou terceira ordem, tendo como arquivo final um áudio multicanal codificado e adaptativo. Esse áudio codificado em *Ambisonics* é inserido em um mecanismo, ou *game engine* (ex.: *Unity*, *Unreal engine*, *Blender*), que controla o jogo e recebe também outros parâmetros, tais como os dados de câmeras, giroscópios ou acelerômetros provenientes da interfaces com o jogador. Uma parte dessa interface é dedicada à identificação da posição e dos movimentos de cabeça e corpo do jogador e irá manipular o campo sonoro tridimensional criado em *Ambisonics* (através dos processos de *tilt*, *rotation* e *roll*⁷³), antes de realizar sua decodificação para fones de ouvido, ou seja, sua transcrição para um sistema *Binaural*.

⁷² Informações obtidas via e-mail com membros da empresa.

⁷³ Esses processamentos de inclinação, rotação e rolamento, respectivamente, são de fácil realização em sistemas baseados em *Ambisonics* e um dos principais motivos para sua adoção nas tecnologias de áudio tridimensional.

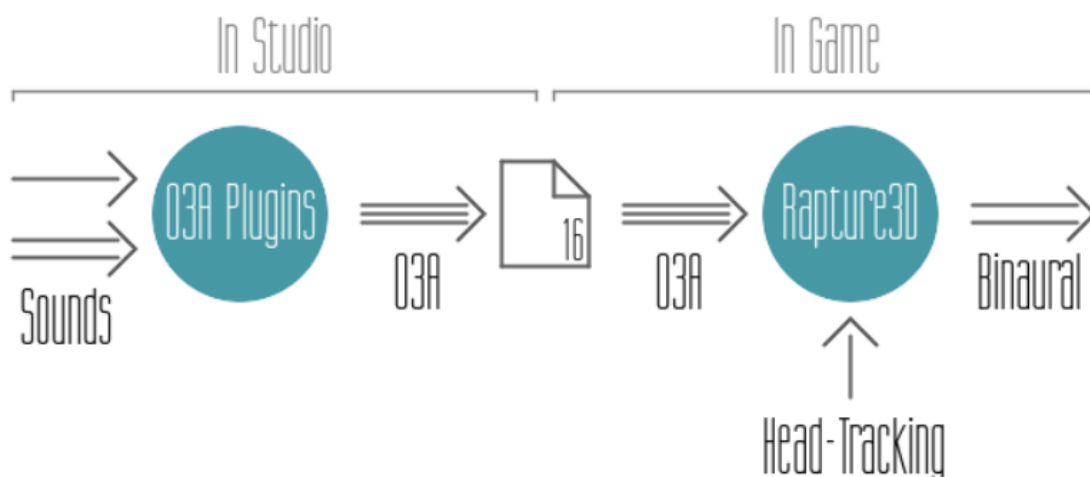


Figura 8 - Fluxograma do áudio para a realidade virtual através do plugin O3A da Blue Ripple⁷⁴.

1.5 O desenvolvimento do VR – uma perspectiva histórica

Apesar de estarmos sempre falando de realidade virtual, alguns autores discordam com relação a essa nomenclatura. Begault (2000, p. 7) sugere o termo *Sistema de Ambiente Virtual*⁷⁵, uma vez que a realidade em si é muito difícil de simular, sendo mais fácil simular algo que não seja a realidade. Já Char Davies (2002)⁷⁶ prefere chamá-la de *Espaço Virtual Imersivo*⁷⁷ e propõe a terminologia também para o ambiente onde suas obras são apresentadas: uma *Arena Espaço-temporal Audiovisual*⁷⁸.

Para além das discussões a respeito da terminologia, é observável que as novas tecnologias envolvendo a produção audiovisual para VR, tem sido o foco de diversas mídias desde 2016, quando do início da disponibilização de dispositivos populares para a visualização de material audiovisual em 3D, assim como do desenvolvimento do mesmo. A possibilidade de interatividade, antes limitada devido ao poder de processamento dos computadores, é uma das maiores responsáveis pela sua visibilidade, assim como a expansão do uso de *smartphones* e o acesso a mídias com reprodução de áudio tridimensional através de fones de ouvido.

⁷⁴ Fonte: <http://www.blueripplesound.com/>

⁷⁵ Original: Virtual Environment System

⁷⁶ Apesar de sua formação inicialmente direcionada para a pintura, passou a se dedicar a obras audiovisuais, sendo reconhecida internacionalmente por obras tais como "Osmoses" e "Ephemere" (<http://www.immersence.com/index.html>).

⁷⁷ Original: Immersive Virtual Space

⁷⁸ Original: Spatial Temporal Audiovisual Arena

1.5.1 2017 "o ano do VR"

Um dos primeiros dispositivos desenvolvido com o intuito de fornecer uma imagem com mais profundidade e volume - o que seria uma das características objetivadas pelas tecnologias da realidade virtual - foi o óculos 3D desenvolvido pelo inglês Charles Wheatstone em 1838 e que consistia de um óculos estereoscópico com lentes levemente anguladas que realizavam tal função. Em 1935 Stanley G. Weinbaum, escritor estadunidense de ficção científica, escreveu *Pygmalion's Spectacles*, uma obra de ficção que descrevia um óculos com o qual o portador podia reproduzir uma série de imagens, como um filme, que inclui as sensações de cheiro, tato e som, com o qual ele podia interagir.



Figura 9 - Óculos 3D de Charles Wheatstone⁷⁹.

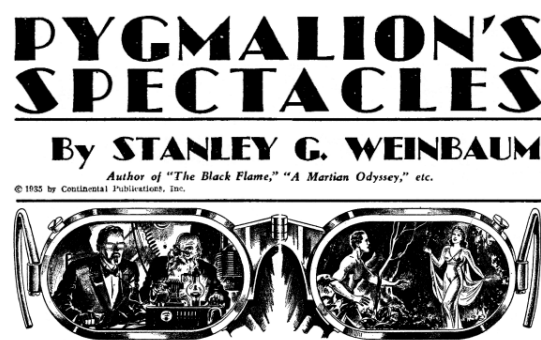


Figura 10 - Obra de ficção científica de Weinbaum⁸⁰.

Nos anos 1970's foi desenvolvido o *Sensorama*, por Morton Heilig, uma máquina mecânica que reproduzia imagens m 3D e sons que, juntamente com o movimento de uma cadeira e a simulação de odores e ventos, foi imaginado como um cinema do futuro. Não foi manufaturado em larga escala mas teve grande influência nos desenvolvimentos futuros.

⁷⁹ Fonte: KLEINA, 2017

⁸⁰ Fonte: BARONE, 2019



Figura 11 - Sensorama de Morton Heilig⁸¹.

Em 1961, a fabricante Philco desenvolveu o *Headsight*, um projetor de estímulo visual e reconhecimento de objetos. Um dos objetivos do dispositivo era possibilitar a movimentação de uma câmera remota, e pretendia ser utilizado em treinamentos militares e de voo.



Figura 12 - Headsight da Philco⁸².

⁸¹ Fonte: BARONE, 2019

⁸² Fonte: D'ANGELO, 2016

Em 1965, o técnico em informática estadunidense Ivan Sutherland desenvolveu o *Head-Mounted-display*, que evoluiu para o também conhecido como espada de Dâmocles, o primeiro *headset* funcional de VR em 1968.

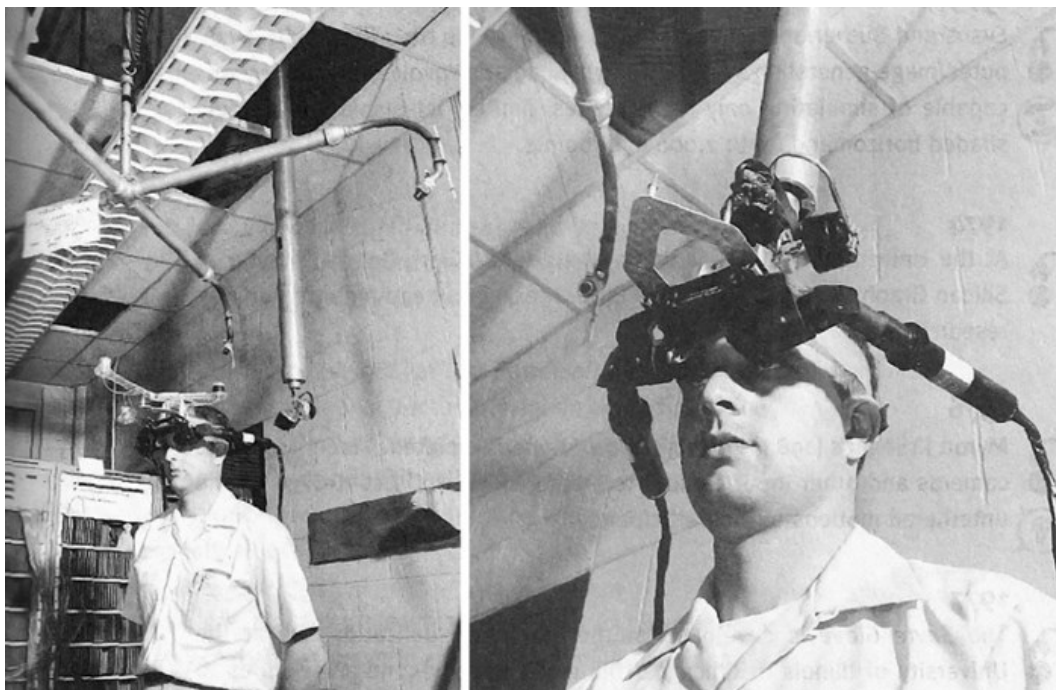


Figura 13 - A Espada de Dêmocles de Ivan Sutherland⁸³.

Em 1985, a desenvolvedora de jogos digitais *Atari VPL Research* oficializa o termo *VR* e lança o *data glove*, o *eye phone*, e o *data suit*, dispositivos para realização de interface com ambientes virtuais.

⁸³ Fonte: BARONE, 2019



Figura 14 - O data suit da fabricante Atari⁸⁴.



Figura 15 - O data glove da Atari na capa da Scientific American⁸⁵.

Em meados dos anos 1980's, a Nasa Ames desenvolve o *Virtual interface environment workstation*, mais conhecido como *VIEW*, um sistema de visualização estereoscópico que, associado a luvas com sensores de movimento (*data gloves*) possibilitavam ao usuário interagir com o ambiente virtual criado pelo computador. Entre 1986 e 1989, foi desenvolvido por Thomas Furness, um *Super Cockpit* para as Forças Aéreas Americana, um sistema capaz de projetar informação como mapas 3D, radar e dados aviônicos, num espaço virtual 3D que o piloto poderia ver e ouvir em tempo real. O sistema de detecção de movimentos, os controles de voz e outros sensores, permitiam ao piloto controlar o avião com a fala, gestos e movimentos oculares.

⁸⁴ Fonte: BARONE, 2019

⁸⁵ Fonte: LOWOOD, 2018

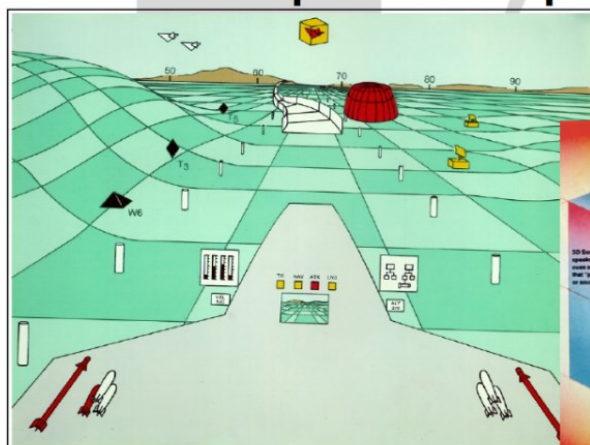


Figura 16 - VIEW⁸⁶.



Figura 17 - Visually Coupled Airborne Systems Simulator, precursor do Super Cockpit⁸⁷.

The Super Cockpit (1980's)



■ **Furness - USAF**



Figura 18 - O projeto Super Cockpit, desenvolvido por Thomas Furness entre os anos de 1986 e 1989⁸⁸.

Nos anos 1990's o grupo *Virtuality* começou a desenvolver uma série de dispositivos e arcades com a tecnologia, inclusive para a Sega, com um *headset* para o Megadrive, no entanto foi somente em 1995 que a Nintendo desenvolveu um dispositivo que saiu realmente da fase de protótipo, o *Virtual Boy*. Esse dispositivo falhou como jogo digital por causar dores de cabeça ao usuário e não ser nada

⁸⁶ Fonte: ROSSON, 2014

⁸⁷ Fonte: PEREIRA, 2019

⁸⁸ Fonte: BYE, 2019

confortável de se usar. Posteriormente a Atari lançou um dispositivo similar denominado *Jaguar VR* mas que não chegou ao grande mercado.



Figura 19 - Virtual Boy da Atari⁸⁹.

Filmes como *O passageiro do futuro* de 1992 e a série de TV *VR Troopers* dos anos 1990's também ajudaram a alimentar as ideias a respeito dos desenvolvimentos tecnológicos para a realidade virtual e a criar o imaginário que temos nos dias de hoje.

Foi somente em Agosto de 2012 que Palmer Luckey resolveu criar um projeto de financiamento coletivo do *Oculus* que, depois de seu período de arrecadação, começou a ser enviado aos compradores em Março 2013 incluindo um código para o desenvolvedor. Em Setembro de 2014 esse kit de desenvolvimento se tornou público e se teve início ao *boom* de games e aplicativos desenvolvidos para o dispositivo, principalmente com a compra do projeto pelo *Facebook*. A versão para o consumidor no início de 2016 ainda era de \$600 (seiscentos dólares), relativamente caro, e a necessidade de um computador com alto poder de processamento ainda faziam do *Oculus* pouco acessível.

⁸⁹ Fonte: KLEINA, 2017



Figura 20 - Oculus da Samsung⁹⁰.

Mas obviamente não foi só o *Oculus* que se desenvolveu nesse período, diversos outros fabricantes, principalmente do setor de jogos digitais entraram nessa corrida pelo mercado do VR. O *Freefly VR* foi lançado em 2013 com o grande diferencial sendo a utilização de um *smartphone* no lugar de um console de jogos, não foi muito bem sucedido mas abriu as portas para o *Cardboard* de 2014 e o *Daydream* de 2016, ambos da *Google*, e o *Gear VR* da *Samsung* de 2015. A *Sony*, consagrada fabricante de jogos digitais com seu console *Playstation*, lançou em outubro de 2016 o *Playstation VR*. A *HTC*, lançou o *VIVE* para ser usado inclusive em sua famosa plataforma de jogos e aplicativos online, o *STEAM*.

Dessa forma é considerado por muitos entusiastas que somente em 2017 a tecnologia disponível, assim como os diversos lançamentos observados, chegou a realizar de maneira satisfatória os anseios dos desenvolvedores, em sua maioria de games (principal aplicação das tecnologias de realidade virtual). Da mesma forma observamos no desenvolvimento de dispositivos que gravam em 360°, começando pela *Nokia* com a câmera *OZO* (já descontinuada). Outras fabricantes a seguiram nesse segmento, incluindo a *GoPro*, a *Samsung* e a *Microsoft*.

⁹⁰ Fonte: www.oculus.com

1.5.2 Ferramentas para geração e manipulação de Vídeo em VR

As ferramentas de geração e manipulação de vídeos tridimensional para VR também estão em constante desenvolvimento e as dividimos entre as que exploram imagens pré-gravadas e aquelas que geram imagens sintetizadas.

Para a captura de imagens em 360° a principal ferramenta na atualidade é uma câmera dedicada. Depois de breve período de propostas que incluíam arranjos de câmeras que variavam de quatro a dezesseis câmeras independentes, câmeras com duas lentes grande angulares se tornaram as principais ferramentas da atualidade.



Figura 21 - Arranjos de câmeras para se obter os 360° de imagem de alta resolução⁹¹.

⁹¹ Fonte: <https://www.thingiverse.com/thing:1017856>, <https://mashable.com/2014/04/29/jaunt-360-degree-virtual-reality-oculus-rift/>



Figura 22 - Samsung gear 360⁹², Nikon Keymission⁹³, 360Fly⁹⁴, GoPro Fusion⁹⁵ e LG360Cam⁹⁶. Algumas das diversas câmeras 360° com apenas duas lentes disponíveis no mercado.

Depois de capturado, o material precisa passar por um processamento de emenda e colagem entre os sinais das duas lentes, normalmente realizado por software desenvolvido pelo próprio fabricante, gerando um vídeo equiretangular⁹⁷ que poderá ser manipulado em softwares de edição de vídeo, tais como o *Adobe Premiere* e o *Sony Vegas*⁹⁸. Só muito recentemente tais softwares desenvolveram ferramentas para manipulação desse material, normalmente em formato 4K, uma resolução bem alta, e, em decorrência disso, ainda possuem poucos plugins e possibilidades de manipulação.

⁹² Fonte: <https://www.samsung.com/br/wearables/gear-360-2017/>

⁹³ Fonte: <https://www.nikonusa.com/en/nikon-products/product/action-camera/keymission-360.html>

⁹⁴ Fonte: <https://www.360fly.com/shop/cameras.html>

⁹⁵ Fonte: <https://gopro.com/en/br/shop/cameras/fusion/CHDHZ-103-master.html>

⁹⁶ Fonte: <https://www.lg.com/br/acessorios-para-celular/lg-LGR105>

⁹⁷ Formato retangular com uma proporção de 2:1.

⁹⁸ Apesar desses dois softwares não serem gratuitos, existem outros, apesar de poucos, que são gratuitos e aceitam a manipulação de vídeo em 360°. Entre eles podemos citar o Shotcut, disponível em <https://shotcut.org/>, que, no entanto, não aceita áudio de 4 canais no formato Ambisonics.

Falamos de poucas manipulações aqui, considerando que os tradicionais cortes, emendas, superposições, enquadramentos, *stretching* e *reverse*, já são bem assimiladas da mesma forma que em vídeos tradicionais. Dentre as outras possibilidades, normalmente em formato de plugin proprietário, podemos apontar como principais, das presentes no *Sony Vegas*, por exemplo, a de rotação do campo visual tridimensional nos três eixos e a de criação de um “pequeno planeta”, que afasta o campo visual simulando um ambiente planetóide como se o víssemos de cima. Já o *Adobe Premiere* oferece um número maior de opções, onde as de criação de anomalias digitais (*glitches*), controle de brilho, desfoque, desvios cromáticos, controle de nitidez, redução de ruídos e, principalmente, a transformação de plano em esfera e esfera em plano, se destacam.

Com relação aos softwares que geram material visual através de síntese ou modelagem 3D, observamos que, apesar de serem muito mais difíceis de manipular, seu resultado se aproxima mais do que objetivamos como música eletroacústica audiovisual (HILL, 2010), principalmente pelas suas características abstratas. Entre os gratuitos podemos citar o *Blender*⁹⁹, o *Unity*¹⁰⁰ e o *Mandelbulb 3D / Mandelbulber*¹⁰¹. Já entre os pagos, os mais conhecidos são o *Adobe After Effects*, o *Cinema 4D*¹⁰², o *Xenodreams*¹⁰³ (gratuito para teste), o *Maya*¹⁰⁴ e o *Houdini*¹⁰⁵, da *SideFx*.

O software *Adobe After Effects*, transita um pouco entre esses dois mundos, no qual é possível gerar algum material 3D mas também manipular vídeos bidimensionais, adaptando-os para a visualização num ambiente de realidade virtual. No presente trabalho, no entanto, foram explorados principalmente materiais visuais proveniente do *Mandelbulb 3D*, que já prevê uma saída simulada de uma câmera 360°, manipulados posteriormente no *Sony Vegas* e no *Adobe Premiere*. Observada a dificuldade no domínio da manipulação de softwares de modelagem 3D, deixamos essa produção para oportunidade futuras.

Quando nos referimos à produção de material visual 3D na verdade estamos nos direcionando para um material que se apresente ao redor do ouvinte, 360° com

⁹⁹ <https://www.blender.org/>

¹⁰⁰ <https://unity.com/pt>

¹⁰¹ <http://www.mandelbulb.com/2014/mandelbulb-3d-mb3d-fractal-rendering-software/>

¹⁰² <http://www.cinema4d.com.br/>

¹⁰³ <http://xenodream.com/downloads.htm>

¹⁰⁴ <https://www.autodesk.com.br/products/maya/overview>

¹⁰⁵ <https://www.sidefx.com/products/houdini/>

altura. A percepção de profundidade, realizada na natureza pela nossa visão estereoscópica, baseada na combinação da luz recebida pelos dois olhos, no audiovisual para realidade virtual, é realizada pela própria plataforma de streaming, através de uma divisão do material para as duas lentes do óculos.

O que manipulamos no software na realidade é uma imagem equiretangular que se fecha em um globo 360° em torno do espectador quando transposta para a plataforma ou player. Tal manipulação exige um certo treinamento para aprendermos a identificar, nesse plano bidimensional, o que será apresentado acima ou abaixo, à frente ou atrás, no campo visual, quando transposto. Observamos também uma difícil transposição desse vídeo em formato equiretangular para outras mídias tradicionais, que não seja um recorte do todo tridimensional.

Todo esse trabalho de manipulação, hoje, é possível de ser realizado em computadores domésticos, independente se o trabalho será apresentado em seu formato final em salas de concerto ou, no nosso caso, através das plataformas de *streaming*. Por outro lado, a criação de performances ao vivo, com material pré-selecionado, onde a improvisação exige ainda mais de seu domínio sobre as ferramentas e materiais, como é o caso dos *VJs* e do grupo *As Is*¹⁰⁶, também passa por uma etapa de estudo e introspecção no material selecionado em ambientes domésticos.

Nessa improvisação, o processo de espacialização do áudio se mostra um pouco mais difícil, mas não impossível, e ainda pode se utilizar de algumas ferramentas específicas, como é o caso do *IsPace*¹⁰⁷. Sua limitação se encontra no próprio controle das diversas fontes sonoras e no detalhamento das movimentações das mesmas que, quando em uma obra pré-concebida em estúdio pode ser muito mais refinada, do ponto de vista da elaboração e da resolução espacial, com possibilidade de se definir um método de espacialização para cada seção de uma obra, por exemplo. Caso a obra esteja direcionada para a criação em tempo real de material visual, com algoritmos de interação (*live electronics*), vemos como principal problemas o processo de renderização, de geração das imagens visuais, que, nesse

¹⁰⁶ Grupo de música experimental nascido do grupo de pesquisa Intersignos, na Escola de Belas Artes da UFMG, que realiza performances audiovisuais com projeções e espacialização do áudio ao vivo. Alguns de seus trabalhos foram registrado nos formatos estéreo e Binaural e podem ser encontrados no YouTube (www.youtube.com/channel/UCnUC_kp2MYesIFiavO8dTIQ).

¹⁰⁷ Software desenvolvido em conjunto com o grupo de pesquisa Intersignos, e utilizado em diversas performances do grupo *As Is*.

caso podem perder em qualidade e resolução, ou demandar equipamentos bem mais robustos.

1.5.3 Mídia audiovisual adaptativa para VR: um híbrido de vídeo-música e arte sonora

Precisamos ter em mente que, enquanto a nossa visão é limitada inicialmente a 180°, podendo ser ampliada com movimentos de cabeça, a nossa audição tem a característica de estar sempre recebendo sinais dos 360° ao nosso redor, de cima e de baixo. As manifestações artísticas audiovisuais tem, até hoje, explorado essas características fisiológicas sem que o espectador, maravilhado com as novidades tecnológicas, questione a coerência do que lhe é apresentado. Seguindo o proposto por Schaeffer, antes de poder efetivamente contribuir com uma nova estética artística, o instrumento almeja transmitir a arte da forma mais fiel possível, desafiando suas limitações. Nesse processo identificamos então o que outras mídias deixam a desejar, ou seja, as limitações que elas apresentam que a impedem de alcançar a fidelidade da representação.

Imagine uma tela de projeção de cinema, com um sistema de reprodução de áudio em 5.1. Ouvimos sons de helicópteros vindos de trás em direção a tela. A primeira contradição que podemos apontar está no fato de não termos caixas, nesse sistema de reprodução de áudio, sobre a cabeça do ouvinte, o que poderia fazer tal movimento mais convincente. Outra incongruência com relação a imagem pode ser observada no fato de que, se virarmos nossa visão para trás, não veremos tal helicóptero¹⁰⁸. Obviamente a adoção de sistemas de projeção visual e de reprodução de áudio obedecem a uma série de demandas mercadológicas, de disponibilidade tecnológica, entre outras coisas. No entanto, queremos apontar aqui que talvez seria muito mais coerente se tivéssemos, juntamente com uma tela frontal e bidimensional, um sistema de reprodução no qual os alto-falantes estivessem posicionados ao redor da mesma ao invés de ao redor do público, possibilitando a reprodução de objetos sonoros sempre frontais mas com a possibilidade de se apresentarem mais acima ou mais abaixo, de acordo com a imagem.

¹⁰⁸ Na literatura o que denominamos aqui de incongruências, entre elas o nomeado som *off screen* (CHION, 1994), entre outros, podem se apresentar como recurso estético. Não os desprezamos, no entanto, ao apontar no percurso proposto por Schaeffer o objetivo de imitar a realidade.

Uma novidade na reprodução videográfica, que se popularizou pouco antes do VR, foi o denominado vídeo 3D, que, com a ajuda de um óculos, nada mais é do que uma ampliação da noção de perspectiva, e que não chega realmente a envolver o espectador nas três dimensões (largura, altura e profundidade) por não apresentar qualquer material atrás, em cima ou embaixo do espectador. Os desenvolvimentos na área do áudio por sua vez já possibilitam há muito tempo que, através de um sistema multicanal, o espectador seja imerso em um campo sonoro tridimensional.

Essas limitações das mídias, ou dos sistemas, utilizadas para reprodução de material audiovisual, se apresentam como obstáculos técnicos colocados à configuração de um ambiente imersivo de projeção audiovisual que não permite correlacionar diretamente eventos visuais com eventos sonoros. Incoerências como as apontadas acima são responsáveis por gerar espectadores com pouco ou nenhum senso crítico, por outro lado, pode-se dizer que elas estimulam a criatividade. Ouvir um objeto sonoro sendo reproduzido acima do ponto de escuta, sem que nele tenhamos um objeto visual pode ser interessante até certo ponto, onde a tecnologia deixa de ser novidade, e tal fenômeno passa a ser completamente incoerente.

Já mesmo na produção que observamos para o que chamamos de mídia audiovisual adaptativa, vemos inúmeras produções para realidade virtual com o áudio ainda em formato estereofônico (ou até monofônico) parecem apresentar, nesse domínio do *espaço extrínseco*, somente relações de controvérsia entre o áudio e o vídeo. Entendemos a novidade que a produção audiovisual para realidade virtual ainda representa, no entanto, ferramentas para o desenvolvimento de áudio tridimensional estão cada vez mais populares e em sua grande maioria gratuitas. Uma elaboração de áudio tridimensional para vídeos para VR ainda não é prioridade da maioria e apontamos como principais motivos: as exigências da indústria pela alta velocidade em que se deve disponibilizar o material audiovisual, sobrando pouco tempo para um tratamento cuidadoso do áudio; a ênfase histórica e ainda hegemônica do visual sobre o sonoro; e a falta de conhecimento estético para explorar os recursos de mixagem de áudio tridimensional, assim como essas outras relações de conformidade e complementação.

Seguindo o que Schaeffer propõe com relação aos três níveis de desenvolvimento de uma nova mídia, depois de passado o momento em que ela é admirada como novidade, sua busca em direção a representação da realidade para

depois poder se tornar um instrumento artístico, exige que ela seja o mais congruente possível.

Na busca pelo desenvolvimento de material audiovisual ao terceiro nível proposto por Schaeffer, onde a técnica informa a arte, esses paradoxos se tornam obstáculos pois, em seu segundo nível, já observamos tais imperfeições, incapacidades de corresponder à realidade. Uma vez essas imperfeições superadas e a realidade sendo possível no novo aparato tecnológico, as portas realmente se abrem para a criação artística. Através de mídias audiovisuais tridimensionais adaptativas para a realidade virtual, uma projeção visual tridimensional, associada a uma reprodução de áudio também tridimensional, tende a minimizar as contradições, como as descritas acima, também observadas em outras mídias, e deixa o compositor livre para explorar o material em um nível semântico, inclusive criando contradições propositalmente, dentro do contexto de sua obra.

Nesse novo aparato, nessa mídia emergente, a exploração do espaço tridimensional apresentado pelo audiovisual, assim como sua interatividade (por mais trivial que seja), respondendo aos movimentos do ouvinte e possibilitando uma maior imersão, possui papel crucial no desenvolvimento e na criação de uma retórica da utilização do espaço que incite significados para o espectador. Na produção de audiovisual para VR vemos então a possibilidade do fim das incoerências técnicas, onde dispomos de um sistema completo de reprodução tridimensional imersivo e responsivo, em que uma composição eletroacústica audiovisual que explore tais possibilidades se aproxima de uma instalação audiovisual inserida em uma outra realidade.

Apesar da proposta inicial de explorarmos a mídia audiovisual adaptativa como produto final, alguns podem argumentar com relação a necessidade de aprimoramentos, observáveis em qualquer sistema com o qual viermos a trabalhar. Considerando isso, podemos pensar na possibilidade de que a produção de música eletroacústica audiovisual explore as mídias audiovisuais adaptativas com o intuito de divulgação, com essa consciência de futuros aprimoramentos, mas se garantindo através da manutenção de formatos em altas resoluções, a serem exploradas em outros contextos ou outros sistemas.

1.6 Espacialização sonora para realidade virtual

Diversas técnicas e ferramentas para se realizar a espacialização sonora ao vivo ou em material pré-gravado foram desenvolvidas ao longo dos últimos 50 anos. Algumas dessas técnicas se direcionaram especificamente para o cinema, tal como o padrão 5.1, que acabou se tornando o padrão para surround para aplicações domésticas. Outras se voltaram para aplicações mais específicas, como o *Ambisonics* e o *VBAP*, principais técnicas adotadas para a difusão de música eletroacústica e instalações sonoras com grandes arranjos de alto-falantes.

Questões relativas à psicoacústica da espacialização sonora, tais como o funcionamento da percepção humana do espaço, seus mecanismos e variáveis que a afetam, já foram vastamente exploradas tanto pelo presente autor em outros trabalhos (SOUSA 2010a, 2010b, 2011, 2013), quanto por outros autores (WALLACH 1939, 1940, BLAUERT, 1969, 1974, BROWN e WANG, 2005, PLACK, 2005, KENDALL, 1995; MALHAM, 1998; BEGAULT, 2000). O mesmo podemos dizer com relação ao funcionamento técnico de cada um dos diversos sistemas de reprodução de áudio espacializado assim como as idiossincrasias que cada sistema de reprodução multicanal apresenta (BARRETT, 2002; BAALMAN, 2010, BATES, 2009; ROBUSTÉ, 2014).

De uma forma generalizada, do ponto de vista artístico, alguns autores e compositores já observaram que técnicas baseadas na manipulação de amplitude entre os sinais enviados para os alto-falantes (ex.: *VBAP – Vector Based Amplitude Panning*) oferecem uma localização mais precisa de fontes sonoras, enquanto perdem no quesito envolvimento. Devido a isso tais sistemas são preferivelmente explorados quando há a necessidade de reprodução de fontes sonoras pontuais e bem definidas.

Por outro lado técnicas baseadas em manipulação de fase, como o *Ambisonics*, oferecem uma localização mais pobre¹⁰⁹ mas uma imersão muito maior do ouvinte no espaço sonoro, sendo preferivelmente explorados quando da necessidade de reprodução de fontes sonoras em movimento ou para a criação de campos sonoros mais homogêneos. A utilização de sistemas como esse auxilia também na atenuação de artefatos que ocasionalmente surgem na manipulação de fontes sonoras entre dois

¹⁰⁹ Tal problema costuma ser superado quando existe a exploração de fontes sonoras em movimento ou da movimentação do campo sonoro como um todo.

ou mais alto-falantes, como os característicos “buracos” quando o ângulo entre os mesmos é maior do que 60°.

A representação de distância através de simulações de reverberação, absorção do ar e efeito *Doppler*¹¹⁰, que podem favorecer trajetórias de fontes sonoras, também são facilmente implementadas nos mais diversos sistemas, tanto nos estereofônicos, quanto nos baseados em *VPAB* ou *Ambisonics* (BATES, 2009, p. 211).

A percepção do ouvinte fora do *sweet spot*¹¹¹ do sistema de reprodução, também deve ser levado em conta ao se decidir qual técnica adotar. Enquanto sistemas baseados em panoramas de amplitude (estéreo, *VBAP*, 5.1) possuem esse *sweet spot* mais limitado, sistemas baseados em *Ambisonics* de ordens mais elevadas oferecem uma performance com o *sweet spot* ampliado, atendendo de forma mais coesamente posições dos ouvintes, dentro ou até mesmo fora do arranjo de alto-falantes (MALHAM, 1992, p. 1).

Na produção de áudio para realidade virtual, duas técnicas de espacialização têm sido vastamente utilizadas em conjunto. O já mencionado *Ambisonics*, desenvolvido na década de 1970 por Michael Gerzon (1973, 1975) e os sistemas *Binaurais*, direcionados para a reprodução de campos sonoros tridimensionais através de fones de ouvido, que começaram a ser concebidos no final dos anos 1880 (BELL, 1880) e aperfeiçoados no início dos anos 1930 por Alan Blumlein (1933), na EMI.

Antes de os aprofundarmos na parte dos desenvolvimentos na área do áudio, passamos brevemente pelos desenvolvimentos e ferramentas que exploramos na produção da parte visual.

1.6.1 Ferramentas contemporâneas para manipulação do áudio tridimensional

Temos que ter em mente que qualquer tecnologia, depois de certo tempo, se torna ultrapassada. Além disso, é fato que uma composição com áudio espacializado criado usando certas tecnologias da época corre o risco de perder sua integridade quando transferida para outra (BAALMAN, 2010, p. 209). A proposição aqui de criação

¹¹⁰ O efeito Doppler consiste em uma alteração de altura (pitch) proveniente do próprio movimento da fonte sonora em relação ao ouvinte (RUMSEY, 2001 p. 217; CHOWNING, 1971)

¹¹¹ O que denominamos de *sweet spot* é a melhor posição para se ouvir um determinado sistema de reprodução baseado em alto-falantes. Em um sistema estereofônico, por exemplo, o *sweet spot* está localizado em um dos vértices de um triângulo equilátero, onde os outros dois são os alto-falantes. Em um sistema octofônico circular horizontal o *sweet spot* estaria bem no centro do círculo.

de material audiovisual para VR não está isenta disso. Vale observar que, atualmente, apesar de termos sistemas extremamente avançados de reprodução de áudio tridimensional, o *Ambisonics* de primeira ordem ainda é o único aceito pelo YouTube 360°. Isso nos força a direcionar nossa produção para tal sistema na expectativa de que em breve venham a ser aceitos novos formatos, inclusive o denominado *HOA* (*High Order Ambisonics*¹¹²).

Outras plataformas e estratégias alternativas de transmissão de vídeos no formato 360° foram pesquisadas, no entanto, o YouTube 360° se mostrou a plataforma mais amigável e difundida. O *Valiant360*¹¹³, por exemplo, se mostra como um ótima ferramenta *github* para a implantação de vídeos em 360° em diversos navegadores de internet mas apresenta o empecilho de ser controlado somente pelo mouse - fator que pode diminuir consideravelmente a imersão do espectador por não explorar dispositivos próprios para VR. O *Omnitone*¹¹⁴ por sua vez, apesar de já oferecer suporte para *HOA* (*Ambisonics* de segunda e terceira ordem), também necessita de alguma programação para ser implementado em um navegador padrão, se limitando à página, na internet, do desenvolvedor de conteúdo para sua difusão. O *Vimeo360* também se mostra como uma opção, principalmente por distribuir conteúdo visual com altíssimas resoluções¹¹⁵, exceto pelo fato de que ele ainda suporta somente áudio estereofônico, não possuindo suporte para *Ambisonics* até a data do presente trabalho¹¹⁶.

O maior atrativo, e principal motivo para direcionarmos nosso trabalho para o YouTube 360°, permanece na sua atual difusão e facilidade de operação¹¹⁷. Para sua implementação necessitamos apenas que o desenvolvedor do material audiovisual siga alguns procedimentos específicos em softwares de edição de áudio e vídeo já dominados, para cumprir os requisitos exigidos pelo sistema, e fornecer tanto o vídeo em 360° quanto o áudio (baseado em *Ambisonics* de primeira ordem) controlados pelo giroscópio presente na maioria dos *smartphones*.

¹¹² Sistemas baseados em *Ambisonics* de ordens superiores à primeira.

¹¹³ <https://flimshaw.github.io/Valiant360/>

¹¹⁴ <https://googlechrome.github.io/omnitone/#home>

¹¹⁵ Vale observa que a distribuição por essa plataforma está isenta de propagandas, às custas de uma assinatura paga mensal.

¹¹⁶ Consulta feita no site do desenvolvedor na data de 31-07-2019.

¹¹⁷ Além dessas duas principais características, a possibilidade de inclusão de palavras chaves (tags) para otimização nos mecanismos de busca da internet, também se mostra um grande benefício quando tratamos da divulgação de um trabalho artístico.

Quando manipulamos o áudio tridimensional temos que atentar para questões da psicoacústica com o intuito de nos aproximarmos o melhor possível do que percebemos na natureza. Mesmo que nossa intenção não seja recriar algo real mas sim preservar um espaço composicional virtual, o que nos faz perceber o espaço na natureza deve ser profundamente entendido. Por exemplo: para simular distância de uma fonte sonora recorreremos à manipulação da própria amplitude do sinal de áudio, à simulação tanto de primeiras reflexões quanto da razão entre som direto e som reverberado, à atenuação de altas frequências, ao uso de efeito *Doppler*, etc. (BATES, 2009, p. 62). Begault (2000, p.40) menciona inclusive o fato do efeito *Doppler* ser um dos principais mecanismo de percepção de fontes sonoras em movimento, o que tem que ser levado em conta no processo de produção do material sonoro.

Begault (2000, p. 8), em sua época, afirma e descreve a necessidade de diversos sensores, sejam eles para o rastreamento do movimento da cabeça do ouvinte, dos membros ou até mesmo para análise de voz, no processo de manipulação do campo sonoro tridimensional em tempo real. A grande vantagem no uso do atual *smartphone* é sua capacidade de integrar sensores, microfones, uma tela de alta resolução e saídas de áudio de alta definição em sua construção, conjugando uma grande facilidade de uso à uma alta difusão nos dias atuais. Essas características fazem dele, atualmente, a melhor ferramenta para a disseminação da tecnologia e de mídias audiovisuais adaptativas desenvolvidas para VR. As possibilidades de interação colocam esse formato em um nível onde tecnologia e arte se fundem, possibilitando um novo mundo de criações audiovisuais antes indisponível. Nesse mundo o domínio de ferramentas de manipulação de áudio e de vídeo se tornam os instrumentos dos compositores contemporâneos.

Begault (2000) também afirma que o futuro do desenvolvimento de material multimídia pra realidade virtual está no desenvolvimento dos softwares, o que já observamos nas atualizações mais recentes (2018) de softwares tais como o *Adobe Premiere* e *After Effects*, entre diversos outros, assim como o desenvolvimento de plugins para manipulação do áudio 3D (*FB360*, *o3acore*, *Ambix*, *ATK for Reaper*, etc...). Os principais instrumentos do artista desenvolvedor de material audiovisual para VR passam a ser portanto os softwares de manipulação digital tanto de áudio quanto de vídeo.

Historicamente, a primeira ferramenta de espacialização de áudio foi o *potentiomètre d'espace* desenvolvido por Pierre Schaeffer e Pierre Henry em Paris, 1951, que controlava um quinto canal de áudio móvel criando um contraste em relação a outros quatro fixos (PALOMBINI, 1999). Posteriormente uma ferramenta utilizada por Stockhausen, baseada em uma mesa rotativa com um alto-falante e quatro microfones¹¹⁸ foi utilizada na espacialização da obra *Kontakte* (1959-1960). Atualmente, para gravar material de áudio já tridimensional, temos a nossa disposição microfones tetraédricos que nos darão quatro sinais de áudio em *Ambisonics* de primeira ordem, que já pode ser associado diretamente a uma imagem tridimensional para VR. Outra opção seria a de se espacializar um som monofônico ao redor do ouvinte através de um codificador *Ambisonics*, o que pode ser obtido com diversos *plugins VSTs* digitais, cada vez mais populares a partir dos anos 2010.

Na criação dos projetos que serão descritos no capítulo 3, realizamos a mixagem da parte sonora dentro de um sistema baseado em *Ambisonics*, que será, no momento da escuta, através do canal do YouTube, com o auxílio de óculos para realidade virtual e fones de ouvido, transcrita para um sistema *Binaural*. Os vídeos foram desenvolvidos com materiais gerados em softwares de síntese de imagens (4.4 e 4.5), gravados com câmeras 360° (4.1, 4.2 e 4.6), ou extraídos de um ambiente virtual gerado dentro da plataforma de criação de jogos digitais, *Unity* (4.3), e finalizados em softwares específicos de montagem. Com o intuito de melhorar o entendimento do processo em si, a seguir descrevemos o funcionamento tanto dos dois sistemas de espacialização utilizados na criação de mídias audiovisual imersiva para realidade virtual, quanto das ferramentas que utilizamos. Descrevemos brevemente também as ferramentas exploradas para a geração dos vídeos.

¹¹⁸ “Foi já em 1959, paralelamente ao texto de Stockhausen sobre a questão da música no espaço, que surge então, na rádio NWDR de Colônia, no Estúdio de Música Eletrônica dirigido por Eimert, a famosa *mesa rotativa* ou *alto-falante rotativo* (*Rotiertisch* ou *Rotationlautprecher*, respectivamente). Acionada pela mão, a mesa rotativa, redonda e apoiada sobre uma única perna central de encaixe, girava rodeada por quatro microfones dispostos simetricamente ao seu redor, conectados cada qual a um canal de um gravador magnético quadrifônico. Em seu centro, dispunha-se um alto-falante emitindo algum sinal sonoro. Consequentemente, após o acionamento do motor rotativo da mesa, os microfones captavam sequencialmente e regularmente o som emitido, conforme velocidade de giro da mesa. Se posteriormente dispostos em torno do ouvinte, os quatro alto-falantes correspondentes às quatro pistas do gravador reproduziam o som emitido em rotação na quadrafonia.” (MENEZES, 1998, p.25)

1.6.2 O Sistema Ambisonics

O sistema *Ambisonics* propõe uma abordagem hierárquica com relação à localização espacial sonora, sua transmissão, reprodução e armazenamento, que seja aplicável a sistemas monofônicos, estéreo, surround horizontal ou reprodução perifônica¹¹⁹ completa incluindo informações no eixo vertical, referentes à altura (GERZON, 1973, 1975). Essa abordagem hierárquica consiste no fato de que, se temos somente o primeiro canal, a componente W, temos um sinal monofônico; se temos dois canais, W e X, conseguimos uma reprodução espacial estereofônica semelhante a que se obtém com o padrão estéreo M-S por exemplo; se temos três canais, W, X e Y, conseguimos uma reprodução surround horizontal, como dois padrões estéreo M-S conjugados; já com os quatro canais, W, X, Y e Z, a reprodução perifônica total pode ser obtida.

O fluxo de trabalho geral com sinais de áudio em sistemas baseados em *Ambisonics* se processa através do endereçamento tanto de material monofônico, quanto de material gravado por microfones em formato-A¹²⁰, já codificado em formato-B¹²¹, para os mais diversos transformadores¹²². A saída desses transformadores é direcionada para decodificadores que irão desmembrar o campo sonoro tridimensional criado de acordo com a configuração de alto-falantes definida como saída.

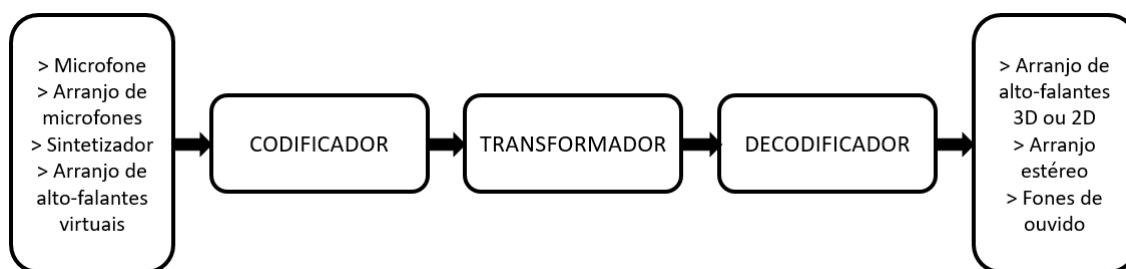


Figura 23 - Fluxo de trabalho em Ambisonics¹²³.

¹¹⁹ Termo utilizado para projeções tridimensionais que incluem altura.

¹²⁰ O formato-A do Ambisonics, se refere aos quatro sinais provenientes diretamente das quatro capsulas de um microfone tetraédrico, nomeados de LFU (*left front up*), RFD (*right front down*), LBD (*left back down*) e RBU (*right back up*).

¹²¹ O formato-B diz respeito já às componentes omnidirecional (W) e bidirecionais nos eixos X, Y e Z.

¹²² Chamamos de transformadores as ferramentas que manipulam o campo sonoro tridimensional como um todo, já com todas as fontes monofônicas codificadas e somadas a eventuais campos sonoros provenientes de microfones.

¹²³ Fonte: adaptado de <http://www.ambisonictoolkit.net/documentation/workflow/>

Na etapa que chamamos de codificação, as componentes direcionais de um som monofônico são codificadas vetorialmente num conjunto de harmônicos esféricos¹²⁴ (figura 24), no caso do sistema de 1ª ordem, quatro, que são conhecidos como formato-B (MALHAM, 1998). O sinal no formato-B pode ser manipulado como uma única entidade através de rotações, inclinações ou até mesmo convoluções com outros sinais no mesmo formato, manipulação utilizada para gerar reverberações tridimensionais, por exemplo.

Caso se pretenda captar um ambiente, ou campo sonoro completo, de forma tridimensional, pode-se fazê-lo através da utilização de um microfone de quatro cápsulas em formato tetraédrico, que já irá fornecer quatro canais de áudio no formato conhecido como formato-A que, transformado no denominado formato-B através de cálculos matemáticos realizados por ferramentas específicas, também pode sofrer as referidas transformações.

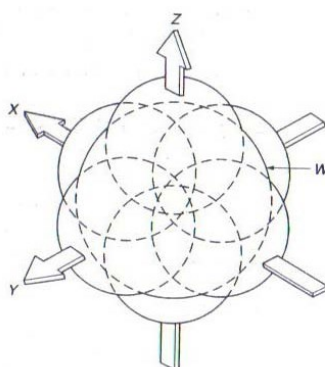


Figura 24 - Componentes W, X, Y e Z do formato-B Ambisonics de 1ª ordem, representando as componentes de pressão omnidirecional e três componentes de velocidade ortogonais do campo sonoro 3D, respectivamente¹²⁵.

Na etapa de decodificação, o sinal gerado para cada alto-falante vai depender das configurações estabelecidas no decodificador, entre elas a quantidade de alto-falantes e a posição de cada um deles, ou se para um arranjo estereofônico simples, ou até mesmo para fones de ouvido. Essa possibilidade de reprodução em diversos

¹²⁴ Segundo Nettingsmeier (2010, p. 5), podemos comparar os harmônicos esféricos aos harmônicos de uma corda, ou harmônicos naturais, por exemplo. A fundamental seria uma esfera perfeita, o primeiro conjunto de harmônicos esféricos ocorre ao longo dos eixos X, Y e Z, como se prendêssemos um anel no equador de um balão, gerando algo próximo do padrão polar bidirecional de um microfone. O próximo conjunto de harmônicos esféricos seriam obtidos como se prendêssemos dois anéis ao redor do balão, nos meridianos 0° e 90°, formando um tipo de trevo, um quadrupolo tesserar, cujo lobos são mais direcionais do que os de primeira ordem. E assim sucessivamente...

¹²⁵ Fonte: RUMSEY, 2001, p. 113

sistemas diferentes se apresenta como uma das maiores vantagens dos sistemas baseados em *Ambisonics*, fazendo com que este seja um formato “à prova de futuro”¹²⁶ (NETTINGSMEIER, 2010).

O que denominamos de *HOA*, *High Order Ambisonics* ou *Ambisonics* de ordens superiores, se refere à sistemas de codificação *Ambisonics* que utilizam mais de quatro canais, sendo a segunda ordem composta por 9 canais, a terceira ordem por 16, a quarta ordem por 25 canais, e assim por diante. Tal codificação possibilita maior resolução espacial, ou seja, reprodução de sons de forma mais pontuais, às custas do aumento de canais utilizados para sua transmissão e armazenamento.

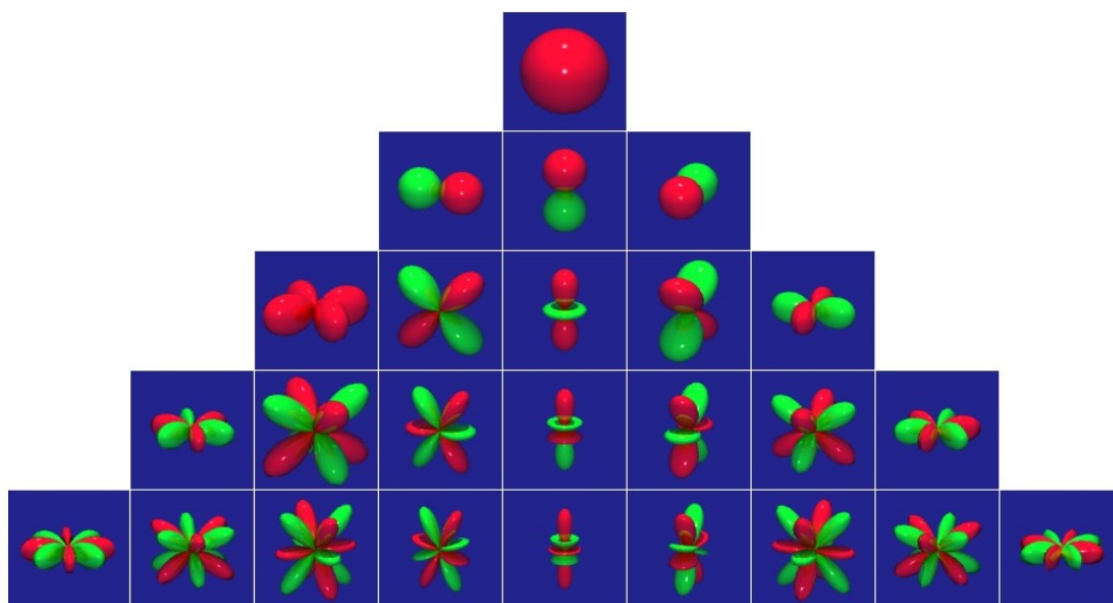


Figura 25 - Harmônicos esféricos de sistemas baseados em Ambisonics nas ordens de 0 à 4¹²⁷.

O motivo principal pelo qual o sistema *Ambisonics* não foi largamente utilizado pela indústria na época de seu desenvolvimento é desconhecido e não oficializado. No entanto, através de conversas em congressos na área de engenharia de áudio com profissionais mais experientes, pressupõe-se que, devido ao seu surgimento quase que simultâneo com o Dolby surround 5.1, o trabalho de Michael Gerzon (1973, 1975) tenha sido abafado pelo grande investimento realizado pela indústria em ferramentas direcionadas aos desenvolvedores e usuários do sistema Dolby 5.1.

Atualmente vemos nitidamente um crescimento no interesse com relação a essa tecnologia por três principais motivos: primeiro pelo fato de ela ser a única que

¹²⁶ Original: *future-proof*

¹²⁷ Fonte: NETTINGSMEIER, 2010

permite que o áudio tridimensional seja adaptável a qualquer configuração de alto falantes, permitindo inclusive uma certa transcrição entre formatos de reprodução, e permitindo que a codificação das informações espaciais seja feita de uma forma que abstrai as condições concretas dos diferentes espaços em que o áudio possa vir a ser projetado, adaptando-o de maneira dinâmica a diversas condições específicas; segundo, pelo fato de o *Ambisonics* ser interativo exigindo pouco poder de processamento, ou seja, é possível movimentar os três eixos do campo sonoro tridimensional criado virtualmente através de um giroscópio, por exemplo, que responda ao movimento da cabeça do ouvinte - aplicação evidente em produções para VR; terceiro, o crescente desenvolvimento de ferramentas digitais gratuitas e de fácil acesso para sua manipulação.

Como dito anteriormente, na geração de material sonoro em formato *Ambisonics*, nos são apresentadas duas principais opções, a de se gravar o ambiente real tridimensional através de um microfone específico, o microfone tetraédrico, ou a de se espacializar tridimensionalmente um objeto sem qualquer informação referente ao espaço. Pelo fato de não termos tido acesso a um microfone tetraédrico, todo o nosso material sonoro foi desenvolvido por essa segunda via, através tanto da espacialização de fontes sonoras monofônicas com o uso de plugins, quanto da manipulação de reverberadores também no formato-B *Ambisonics*, ou presentes dentro dos próprios plugins. De qualquer forma segue uma breve descrição dos microfones para captação em *Ambisonics* assim como dos quatro pacotes de *plugins VST*, disponíveis atualmente e de fácil acesso, que testamos na nossa etapa de criação.

A *Soundfield*¹²⁸, primeira fabricante de microfones tetraédricos nos oferece algumas opções de microfones, sendo elas, um microfone cuja saída em formato-A pode ser manipulada por software, e outros que incluem decodificadores com saídas em 5.1, estéreo ou mono diretamente do processador. Seguindo o caminho de microfones controlados por software que oferecem inicialmente os quatro canais do *Ambisonics* de primeira ordem em formato-A, estão a Sennheiser com o microfone denominado AMBEO, e mais recentemente a Rode com o NT-FS1.

¹²⁸ <http://www.soundfield.com>



Figura 26 - Microfones da Soundfield, SPS 200 e ST450 MKII¹²⁹.



Figura 27 - Microfones AMBEO da Senheiser¹³⁰ e NT-FS1 da Rode¹³¹.

¹²⁹ Fonte: <https://www.soundfield.com/#!/products>

¹³⁰ Fonte: <https://pt-br.sennheiser.com/microphone-3d-audio-ambeco-vr-mic>

¹³¹ Fonte: <https://www.rote.com/ntsf1#overview>

Para a codificação de sinais de áudio monofônicos para um sistema tridimensional em *Ambisonics* de primeira ordem recorremos aos mais diversos *plugins VSTs* digitais. Há pouco tempo atrás, poucos eram os *plugins* que trabalhavam com tal formato, no entanto com a expansão do áudio para *VR*, diversos desenvolvedores investiram em ferramentas desse tipo, com interfaces mais amigáveis e intuitivas. Vale observar que selecionamos apenas alguns dos mais recentes *plugins* desenvolvidos, considerando que desde há um bom tempo diversos pesquisadores tem se dedicado ao desenvolvimento de tais ferramentas (Dave Malham¹³², Bruce Wiggins¹³³, Daniel Courville¹³⁴). No anexo 3, fazemos uma breve descrição dos que selecionamos, sendo aconselhável uma pesquisa no site oficial ou manual para maiores detalhes do funcionamento e operação dos mesmos. Nosso principal objetivo é o de oferecer um panorama geral das possibilidades técnicas e estéticas e das ferramentas atualmente disponíveis, principalmente das gratuitas.

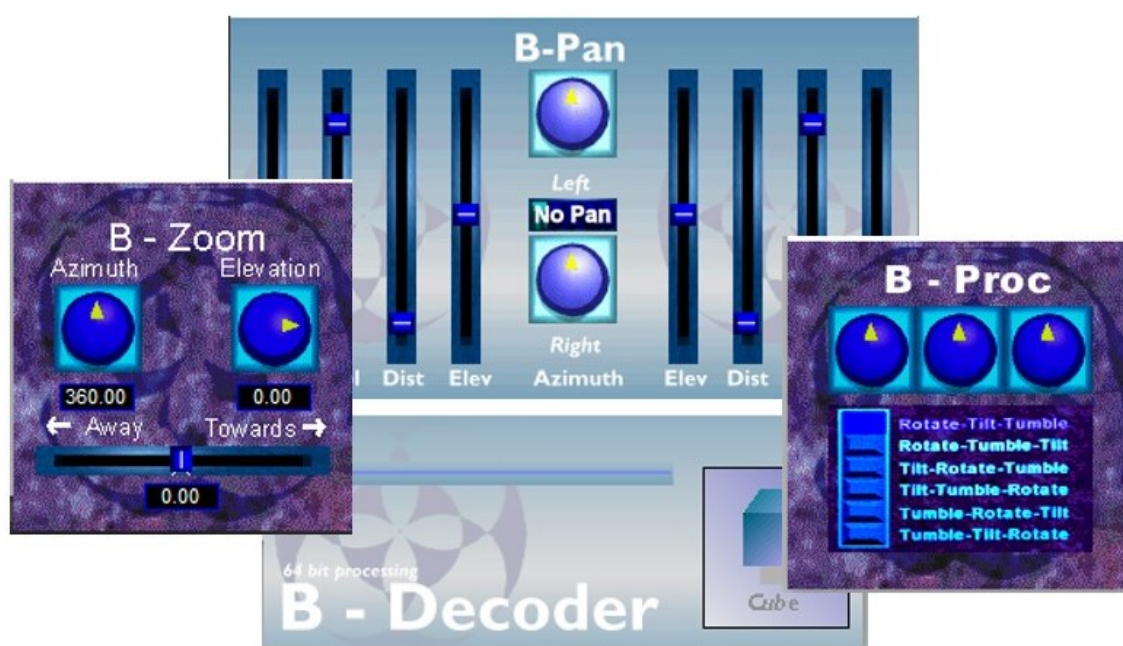


Figura 28 - Plugins desenvolvidos por Dave Malham na Universidade de York – Inglaterra.

¹³² https://www.york.ac.uk/inst/mustech/3d_audio/vst/welcome.html

¹³³ https://www.brucewiggins.co.uk/?page_id=78

¹³⁴ http://www.radio.uqam.ca/ambisonic/b2x_v2.html

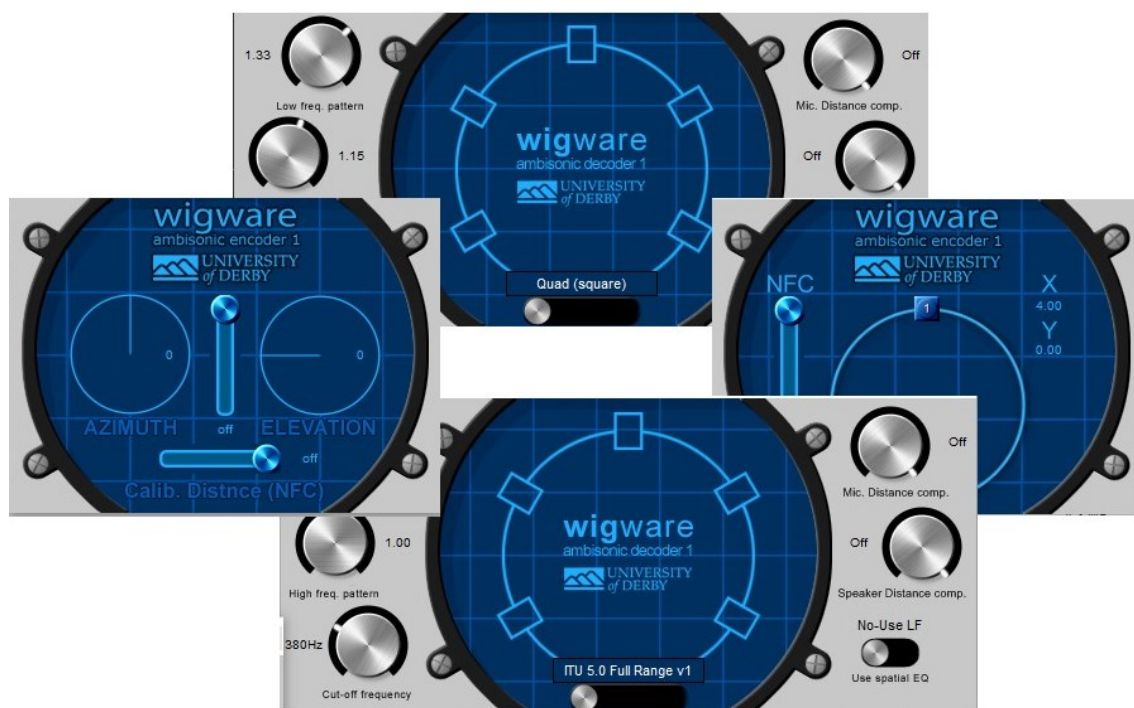


Figura 29 - Plugins desenvolvidos por Bruce Wiggins na Universidade de Derby – Inglaterra.

Alguns dos plugins descritos já possuem um simulador de espaço, incluindo reverberações, primeiras reflexões e ressonâncias, como é o caso do *FB360*. É importante observar aqui que tais recursos costumam ser necessários principalmente em softwares denominados *motores de jogos (game engines)*, como a *Unity*, e que realizam a espacialização de áudio orientada a objeto, ou seja, o usuário define a posição do objeto sonoro (comumente associado a um objeto visual ou diegético) e a posição do ouvinte. O sistema realiza a espacialização sonora gerando, inclusive, conteúdo sonoro referente ao ambiente físico em que o personagem se encontra, e codifica tudo em *Ambisonics*, formato que será manipulado de acordo com o movimento do ouvinte (muitas vezes coincidente com o movimento do personagem do jogo) e, em uma etapa final, decodificado para o sistema de reprodução de áudio disponível, seja ele multicanal baseado em alto-falantes ou sistemas *Binaurais*.

Essa reverberação, simulação do espaço físico em que a fonte sonora se encontra, pode ser realizada também através da convolução de um sinal em formato-B, constituído de uma mixagem já espacializada de diversos sinais monofônicos, com uma *IR* em formato-B para obter reverberações extremamente convincentes. Alguns dos pacotes descritos acima já oferecem tais possibilidades, no entanto, um par de um plugin qualquer (ex. *SIR*, *Waves IR*, etc.) que realize convolução em estéreo (dois canais por vez) já é capaz de realizar tal operação, bastando ao compositor carregar

os arquivos das *IRs*, que podem ser obtido em diversos sites dedicados¹³⁵, de dois em dois, para completar as quatro convoluções (nos canais W, X, Y e Z).

1.6.3 Sistemas Binaurais

Nesse início de milênio, um outro advento que vem influenciar o processo de produção e distribuição do audiovisual foi o desenvolvimento e democratização dos aparelhos celulares. Quando estes se tornaram *smartphones*, capazes de reprodução não somente do material sonoro (já sendo distribuído pela internet), quanto de material de vídeo, favorecidos pela exibição em suas telas que chegam a 7 polegadas de tamanho, essa democratização de material audiovisual se intensificou ainda mais.

Tais aparelhos, atualmente disseminados de uma forma sem precedentes, associados ao acesso à internet em velocidades cada vez mais altas, permitem que o ouvinte receba conteúdo audiovisual a qualquer momento e em qualquer lugar. Em tais dispositivos a reprodução do áudio se dá muito comumente por meio de fones de ouvido que já vem com o próprio aparelho e já se tornaram objeto de fetiche para muitos usuários. A possibilidade de alcançar o público através desse meio trouxe à tona um sistema de reprodução de áudio por muito tempo esquecido e que possibilita a reprodução de áudio tridimensional 360° em torno do ouvinte e inclusive com altura, o sistema *Binaural*.

Recentemente também, tem se dado muita atenção ao denominado áudio 8D (ou *8D Tunes*)¹³⁶, no qual, com fones de ouvido pode-se ouvir uma música espacializada tridimensionalmente, prometendo uma experiência imersiva, ou até mesmo “sentir sons com o cérebro”¹³⁷. Apesar do sensacionalismo das reportagens, das contradições e especulações encontradas na internet de como se adaptar ou realizar uma gravação para tal aplicação, podemos afirmar que o processo em si não passa de uma espacialização para sistemas *Binaurais* baseada em *HRTFs*¹³⁸. Muitos canais especializados tem surgido e publicado versões espacialização de músicas conhecidas, e os próprios artistas tem entrado nessa onda de publicar suas música

¹³⁵ <http://cnmat.berkeley.edu/projects/pairs>

<https://www.waves.com/downloads/ir-convolution-reverb-library>

¹³⁶ <https://www.youtube.com/channel/UCrRpYEytlHGyDgNWO6VbHIQ>

¹³⁷ https://www.correiobraziliense.com.br/app/noticia/diversao-e-arte/2018/10/23/interna_diversao_arte,714557/ja-ouviu-musica-em-8d-tecnica-permite-sentir-sons-com-o-cerebro.shtml

¹³⁸ *Head Related Transfer Function* (função de transferência relativa a cabeça).

em 8D. No processo, ao se aplicar um codificador *Ambisonics* ao áudio estéreo tradicional saído do estúdio, adicionando algum movimento (muitas vezes bem aleatório) e uma reverberação no formato-B antes da codificação para *Binaural*¹³⁹ já se obtém o reverenciado áudio 8D. Uma outra forma de se obter um resultado semelhante seria utilizar diretamente um plugin que realize um panorama *Binaural* como, por exemplo, o *Orbit*, oferecido gratuitamente pela Sennheiser¹⁴⁰, com o prejuízo de não se ter a profundidade oferecida pela simulação de reverberação.

Com sua origem real na década de 1920, nos laboratórios Bell, Harvey Fletcher já tentava gravar e reproduzir um ambiente tridimensional utilizando somente dois canais (MALHAM, 2010, p. 1). Cerca de 11 anos depois, no mesmo local, Alan Blumlein retoma seus experimentos. Ambos realizaram gravações exploravam o uso de *Dummy Heads*¹⁴¹ ou arranjos de microfones específicos, baseados por exemplo, no posicionamento um par de microfones estéreo na mesma distância média observada entre os ouvidos de um ser humano, com a inclusão de uma barreira acústica para simular a cabeça e o torso, e da utilização de uma orelha sintética para a simulação de suas reflexões, obtendo uma captação que simulasse os mecanismos de escuta que nos fazem perceber a tridimensionalidade sonora na natureza. Inicialmente, tal técnica seria capaz de reproduzir todos os parâmetros responsáveis pela percepção espacial da escuta humana (*ITD*¹⁴², *ILD*¹⁴³, distância, reflexões, etc.) (RUMSEY, 2001, p. 13). Tais gravações eram, depois, reproduzidas através de fones de ouvido.

¹³⁹ Observar que no resultado final desse processo, onde se obtém um áudio de dois canais, não teremos interatividade decorrente do movimento de cabeça do ouvinte.

¹⁴⁰ <https://en-us.sennheiser.com/ambeo-orbit>

¹⁴¹ Manequins em formato de torso e cabeça com microfones inserido em seus ouvidos desenvolvidos para criação de material binaural, ou para a gravação de *HRTFs*. Ex.: Neumann KU100 e 3diosound (<https://en-de.neumann.com/ku-100>, <https://3diosound.com/>)

¹⁴² *Interaural Time Difference* (diferença de tempo entre os ouvidos).

¹⁴³ *Interaural Level Difference* (diferença de intensidade entre os ouvidos).

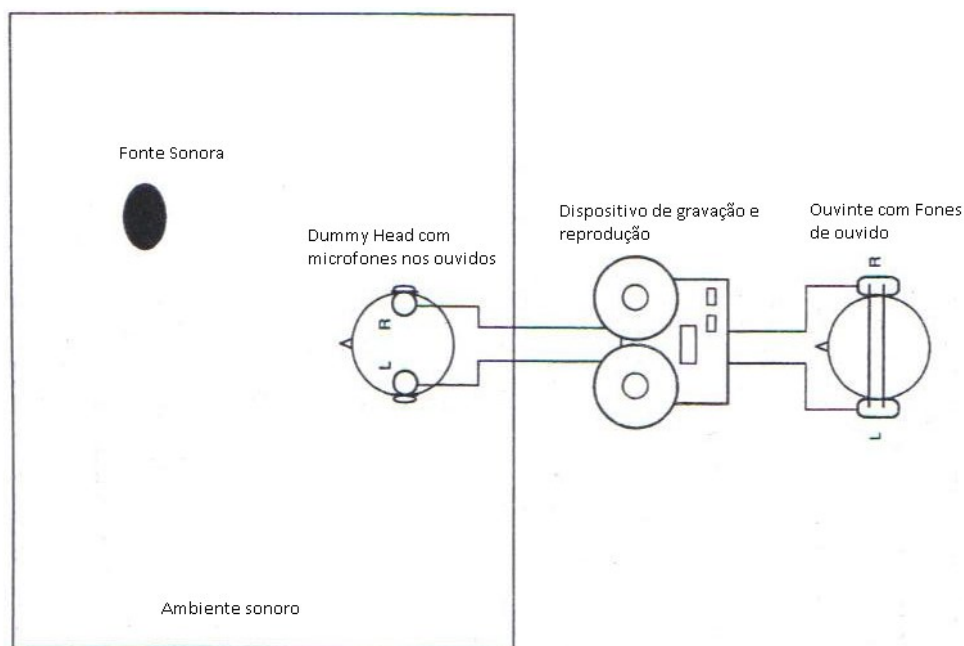


Figura 30 - Sistema de gravação e reprodução Binaural¹⁴⁴.

Um dos principais motivos para que o áudio *Binaural*, apesar de fascinar a maior parte dos pesquisadores, não ter sido implementado e adotado comercialmente foi o fato de o mesmo, quando desenvolvido para fones de ouvido, não ser imediatamente compatível com alto-falantes (RUMSEY, 2001, p. 64). Isso acontece devido à sua própria natureza, de gravar isoladamente o que chegaria a cada orelha, diferindo em muito de uma reprodução por alto-falantes onde ambas as orelhas recebem material proveniente de ambos os transdutores.

Somado a isso, observamos que, na época, não existiam mídias próprias para sua gravação, tendo sido exploradas somente em demonstrações técnicas (MALHAM, 2010, p. 1). O fato de que a ilusão tridimensional oferecida por tais sistemas desaparece quase completamente quando o ouvinte move a cabeça, desestruturando as relações entre os sinais recebidos pelos dois ouvidos e a lógica do cérebro humano, fazendo com que o som seja percebido internamente à cabeça do ouvinte, também não favorecia a adoção do sistema. Tal adoção exigiria então que o mesmo fosse interativo, adaptando-se a posição da cabeça do ouvinte, “enganando” o cérebro ao simular as alterações provenientes da movimentação e mantendo assim uma percepção sonora tridimensional convincente com uma relativa percepção das fontes sonoras externas à cabeça.

¹⁴⁴ Fonte: adaptado de RUMSEY, 2001, p. 66

Existem, no entanto, diversas vantagens e desvantagens de se utilizar sistemas *Binaurais* para a reprodução de áudio tridimensional através de fones de ouvido em detrimento de sistemas baseados em alto-falantes. Begault (2000) nos chama a atenção para o fato de que, quando um sistema é baseado em alto-falantes, uma vez que o som será reproduzido em um ambiente desconhecido, a sala e os alto-falantes em si irão impor naturalmente transformações não lineares no áudio reproduzido antes que o mesmo chegue ao ouvinte. Tais transformações podem ser compensadas no design ou no controle de um sistema de áudio tridimensional baseado em fones de ouvido (BEGAULT, 2000, p. 152), ou seja, o material desenvolvido para reprodução em tais sistemas não estaria tão susceptível à influência do *espaço externo* e não apresentaria variações na escuta dependente da posição do ouvinte na sala (BEGAULT, 2000, p. 174).

Em seu trabalho sobre áudio para realidade virtual, Begault (2000, p. 174) vislumbra uma série de aplicações para os sistemas de realidade virtual 3D baseados na reprodução de áudio *Binaural*. Ele começa com a descrição de aplicações aeronáuticas, destino principal do seu trabalho, para posteriormente abordar outras aplicações tais como a da indústria do entretenimento.

Segundo ele, a premissa básica de um bom sistema de reprodução *Binaural* considera reproduzir da melhor forma possível, os parâmetros de escuta espacial que chegariam aos ouvidos do ouvinte em situações reais. Isso pode ser obtido ao se gravar uma resposta a um impulso (*HRTF*) do ouvinte para cada ângulo horizontal e vertical de uma fonte sonora, criar um banco de dados ao qual outras fontes sonoras serão convoluídas ao serem posicionadas em tais coordenadas antes de reproduzidas por fones de ouvido.

Parâmetros para a realização de simulação convincente de distância também envolvem a manipulação da intensidade do sinal da fonte sonora assim como a relação entre o som direto e o som refletido. As denominadas *BRIRs*¹⁴⁵, que se constituem basicamente de uma *HRTF* que já inclui informações sobre material refletido em uma determinada sala para uma determinada posição do ouvinte e fonte sonora, também podem ser utilizadas para se alcançar tal espacialização.

A maior desvantagem apontada pelo autor se mantém então no fato de que um material desenvolvido para sistemas *Binaurais* e reprodução através de fones de

¹⁴⁵ *Binaural Room Impulse Response* (resposta de impulso binaural de uma sala).

ouvido não é imediatamente compatível com sistemas baseados em alto-falantes. Begault (2000, p. 175) menciona, no entanto, uma possibilidade de se adaptar o áudio *Binaural* à reprodução através de alto-falantes por meio do que ele denomina de cancelamento cruzado.

Sistemas que realizem essa compensação pelo cancelamento cruzado, conhecidos também como sistemas transaurais ou *ambiophonics*, pretendem compensar o fato de que o que chega em cada ouvido nessa situação será sempre uma combinação do que sai dos dois alto-falantes. Por questões matemáticas deve-se saber exatamente a posição do ouvinte em relação aos alto-falantes, possibilitando o cálculo da compensação realizada pelo sistema. Tais sistemas são reconhecidos principalmente pelo seu *sweet spot* extremamente pequeno em decorrência de seus princípios de funcionamento e, como Begault, não nos aprofundaremos neles, que serão objetos de pesquisa de outros autores (MILLER, 2002).

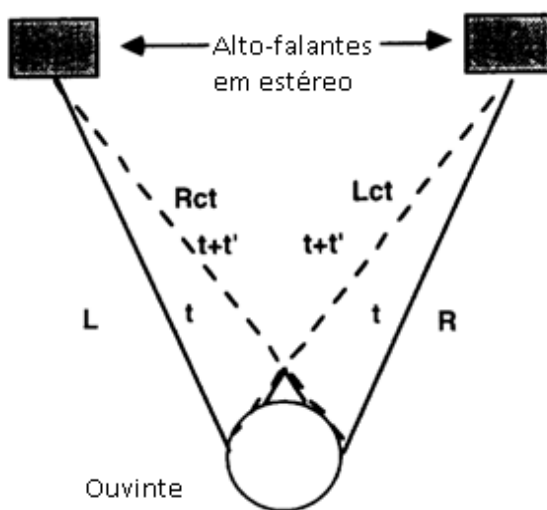


Figura 31 - Teoria do cancelamento cruzado¹⁴⁶.

Ao aprofundarmos nos estudos dos sistemas *Binaurais*, observamos que uma das principais discussões entre os pesquisadores diz respeito a utilização pelo ouvinte de sua própria *HRTF*, gravada em laboratório, ou da utilização de uma *HRTF* generalizada. Enquanto algumas pesquisas apontam a maior precisão quando se utiliza a *HRTF* do próprio usuário, por exemplo (MOLLER, 1992, 1996), outras confirmam o fato de que o usuário pode se adaptar a uma *HRTF* que não a dele após

¹⁴⁶ Fonte: adaptado de BEGAULT, 2000, p. 175

algum tempo de treinamento (BEGAULT, 1991). Miller (2002) e Moller (1992, 1996) consideram também, o processo de gravar as *HRTFs* para cada ouvinte, muito pouco práticos principalmente se o objetivo for a da reprodução audiovisual para o grande público, e diversas pesquisas que pretendem calculá-las a partir de fotos já se desenvolveram (ASSELOT, 2008). A movimentação das fontes sonoras e simulação de outros parâmetros, tais como a amplitude da reverberação e do som direto, assim como a simulação do efeito *Doppler* surgem como parâmetros capazes de suprimir a necessidade de *HRTF* individualizadas dos sistemas *Binaurais* para fones de ouvido (CHOWNING, 1971).

Outra questão sempre apontada é a de que, mesmo se objetivarmos a reprodução para um único indivíduo através de fones de ouvidos, os sistemas *Binaurais* esbarram na questão da interatividade, inexistente caso não seja implementado algum algoritmo para realizar os ajustes necessários, em tempo real, para manter a ilusão da tridimensionalidade.

O desenvolvimento da realidade virtual observada nos últimos 10 anos e direcionada para os smartphones, vem trazer esses sistemas de volta ao foco dos pesquisadores uma vez que os mesmos podem oferecer um sinal referente a orientação do ouvinte. Isso possibilita que as alterações para manter a ilusão de tridimensionalidade através dos fones de ouvido sejam realizadas.

Dentre as considerações de Begault (2000) para o design de sistemas de áudio tridimensionais, está mencionada a necessidade de se decidir se o processo será realizado com um número limitado de medições de posições *HRTFs* ou se será feita a interpolação entre diversas medições através de um software. Adicionalmente, ele observa que qualquer sistema que prevê movimentação de fontes sonoras ou da cabeça do ouvinte (para a implementação da interatividade) necessita de uma resolução, nesse caso do ponto de vista da quantidade de ponto no espaço dos quais se tem gravados *HRTFs*, mais refinada.

Considerando que em uma composição eletroacústica audiovisual para realidade virtual a maior parte das fontes sonoras se movimentam, assim como o próprio ouvinte, para confirmar a percepção da origem da fonte sonora e garantir algum nível de imersão, tal processamento se torna mais do que necessário (BEGAULT, 2000, p. 132). Begault (2000) foca seu trabalho na interpolação de *HRTFs*, o que é muito pesado do ponto de vista de processamento computacional.

Sistemas *Ambisonics* não são tratados pelo autor mas, por possibilitarem a manipulação de todo o campo sonoro tridimensional como um todo com pouco recurso de processamento, se apresentam como uma solução pra diversos dos empecilhos que ele encontra e descreve no processo de implementação de um sistema de realidade virtual 3D baseado em fones de ouvido.

Em aplicações para *VR* teremos um *smartphone* acoplado a um óculos de realidade virtual preso à cabeça do ouvinte. A movimentação do mesmo será transformada em coordenadas que controlam a rotação do campo sonoro tridimensional criado em *Ambisonics*, antes de ter o áudio decodificado para *Binaural*. Nesse processo se considera os alto-falantes, fontes sonoras secundárias dos sistemas baseados em *Ambisonics*, como fontes sonoras primárias do processamento para sistemas *Binaurais*. Com isso podemos considerar que todo o campo sonoro em si está em constante movimento, se ajustando continuamente de acordo com a movimentação do ouvinte, com o intuito de fornecer uma percepção externalizada estática, e favorecendo inclusive a localização das fontes sonoras independentemente das *HRTFs* utilizadas pelo sistema. Vale reforçar aqui que, como descrito acima, a produção de uma mídia audiovisual adaptativa para realidade virtual não será compatível com a reprodução através de alto-falantes.

Com relação à percepção da externalização das fontes sonoras e do campo sonoro como um todo em um Sistema Binaural baseado em fones de ouvido, ou seja, representá-lo de uma forma convincente o suficiente fora da cabeça do ouvinte, dependemos de quatro principais variáveis. A principal delas, solucionada nessa aplicação para *VR* é a movimentação das fontes sonoras primárias ou secundárias (sinais dos alto-falantes provenientes da codificação de um sinal de um sistema baseado em *Ambisonics*) de acordo com a movimentação da cabeça do ouvinte. A segunda variável seria o tempo que o ouvinte precisa de contato com o sistema em si para se treinar na percepção espacial com uma outra *HRTF*, de preferência com fones *in-ear* para que a sua própria seja excluída da equação. Uma reverberação convincente, já possível com *IRs* gravados em *Ambisonics*, seria uma terceira variável importante de ser considerada no processo de composição. E a quarta seria a e de uma imersão suficiente forte para “convencer” o espectador de que tais sons não se originam dos alto-falantes do fone de ouvido mas sim de fontes sonoras externas. Aplicações para *VR* possuem grande potencial para solucionar tais variáveis,

principalmente as da movimentação do campo sonoro e o da reverberação por convolução ao adotar um sistema intermediário baseado em *Ambisonics*. O uso de um mesmo conjunto de *HRTFs* pelo player depende do próprio desenvolvedor e a questão da imersão varia de acordo com a pré-disposição do ouvinte.

Na decodificação do sinal *Ambisonics* para *Binaural*, temos a disposição diversas ferramentas embutidas nos próprios pacotes de *plugins* já mencionados para manipulação de material em formato *Ambisonics*. A principal diferença entre tais decodificadores desse tipo, irá se apresentar na quantidade de opções, sendo vezes mais elaborado como é o caso do *ATK for Reaper* (com diversas *HRTFs* a escolha do usuário), outras menos, como no *FB360*.

Observamos que uma ação importante do desenvolvedor de conteúdo (compositor / engenheiro de som), esteja na escolha de qual codificador utilizar para poder se adaptar a ele e realizar mixagem espacializadas cada vez mais consistentes, como é feito com qualquer outro sistema de monitoração de áudio profissional. No que diz respeito a qual desses decodificadores o YouTube 360 adota, fica uma questão aberta para futuras pesquisas, tendo certeza de que qualquer que seja, contanto que seja sempre o mesmo, os espectadores irão se adaptar a ele na medida do tempo de exposição.

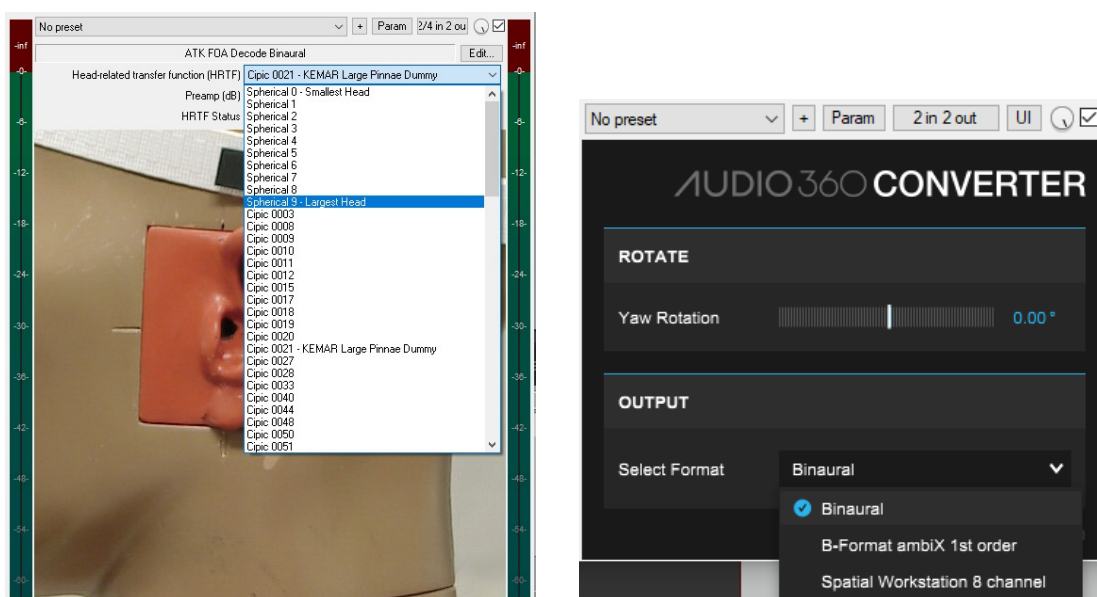


Figura 32 - Decodificador Binaural do *ATK for Reaper* (com várias opções de HRTFs) e decodificador Binaural do *FB360* (com uma única opção).

2. O ESPAÇO EXTRÍNSECO NA MÍDIA AUDIOVISUAL ADAPTATIVA

O espaço como um parâmetro musical é, contudo, muito mais abrangente e mais difícil de se definir. Ele incorpora as dimensões de objetos sonoros individuais, as relações entre objetos sonoros e até as relações entre objetos sonoros e o espaço acústico em que são percebidos. (...) Ele [o espaço como parâmetro musical] engloba todos os aspectos de um som e, ainda assim, os aspectos espaciais de um som muitas vezes não são conscientemente percebidos como separados do próprio som. No nosso desenvolvimento evolutivo, o 'onde' é tratado imediatamente pelo nosso subconsciente, enquanto o 'o quê' e o 'porquê' se tornaram focos principais da nossa atenção consciente.¹⁴⁷ (BATES, 2009, p. 1)

2.1 Os diversos conceitos de “espaço” na música

Quando pesquisamos o termo *espaço* na literatura relacionada com a música, encontramos uma infinidade de conceitos¹⁴⁸ e é principalmente por isso que vemos a necessidade de definir o que chamamos de *espaço* em música, antes de pormenorizar desenvolvimentos estéticos de espacialização.

Dentre os diversos autores que dissertam a respeito do conceito de *espaço*, Chion (1998, p.31) aponta a dicotomia entre *espaço interno* e *espaço externo*, ao se

¹⁴⁷ Original: "Space as a musical parameter is, however, much broader and more difficult to define. It incorporates the dimensions of individual sound objects, the relationships between sound objects and even the relationship between the sound objects and the acoustic space in which they are heard. (...) It encompasses every aspect of a sound and yet the spatial aspects of a sound are often not consciously perceived as separate from the sound itself. In our evolutionary development, the where is dealt with immediately by our subconscious while the what and why became the primary focus of conscious attention."

¹⁴⁸ Dentre eles podemos citar: espaço externo, espaço de escuta, espaço percebido, espaço físico, espaço interno, espaço composto, espaço componível, espaço sobreposto, espaço da obra, coerência espacial, polifonia do espaço, espaço extrínseco, textura espacial, espaço-morfologia, espaço vazio, espaço percorrido, espaço ocupado, espaço preenchido, posicionamento espacial, espaço intrínseco, espaço espectral, espaço de projeção, espaço referencial, espaço movente, espaço imaginativo, espaço madrigalesco, espaço identificável, espaço móvel, espaço vinculado à fonte, espaço de representação, espaço real, espaço perspectivo, espaço estereofônico, espaço panorâmico, espaço prospectivo, espaço mimético, espaço acústico, espaço artificial, espaço geométrico, espaço ilusório, espaço irreal, forma espaço, espaço geometria, espaço alusão, espaço ilusão, espaço fonte, espaço ambiofônico, espaço paisagem, espaço metáfora, espaço artifício, espaço virtual, espaço tridimensional, temática do espaço, movimento no espaço, acústica do espaço, espaço sonoro, espaço sonoro real, articulação do espaço, espaço articulável, espaço natural, transposição do espaço, espaço vazio, espaço profundo, espaços simultâneos, espaço octofônico, espaço fechado, espaço aberto, espaço resultante, espaço vertical, espaço desordenado, espaço textural, espaço percorrido, espaço ocupado, espaço preenchido, espaço abstrato, espaço imaginário, espaço metafórico, espaço arquitetônico, espaço de projeção estereofônico, espaço eletroacústico, impressão de espaço, espaço de composição, espaço indiferente, espaço vivido, espaço ouvido, projeção espacial, espaço frontal, espaço traseiro, escritura do espaço, cenografia do espaço, espaço do material, espaço global, consonância espacial, espaço individual, espaços múltiplos, espaço amplo, espaço pessoal, espaço perceptível, espaço conhecido, envolvimento do espaço, presença espacial, cena espacial, espaço estreito, espaço distal.

referir especificamente à música concreta e Smalley (1997, p. 122)¹⁴⁹, por sua vez aponta a influência do que denominamos de *espaço externo* na percepção da obra musical, principalmente quando uma obra, trabalhada pelo compositor dentro de seu próprio estúdio, é transferida para um espaço com um sistema de difusão de grande escala. Tomamos então como princípio a sistematização de tais termos elaborada pelo compositor Horacio Vaggione (2000, p. 4), que define o *espaço interno* como o espaço representado pela própria obra acusmática, constituindo-se de sons que possuem características espaciais individuais e que se relacionam entre si e ao longo da obra. E, em contrapartida, o conceito de *espaço externo* que diz respeito ao local onde a obra está sendo executada, imprimindo em uma obra acusmática suas características próprias e a modificando antes de chegar ao ouvinte.

O que assumimos, juntamente a Vaggione (2000) como *espaço externo* tem então essa estreita relação com o que o público ouve durante a difusão sonora tanto de obras acusmáticas durante sua execução em sala de concerto, quanto de obras puramente instrumentais. O espaço físico onde a peça está sendo apreciada assume um papel de filtro da obra, apontando uma certa fragilidade da composição, que se transformará em função das características próprias do local onde é projetada (e.g. se o espaço for mais reverberante do que o estúdio utilizado pelo compositor durante o processo de criação, este evidenciará menos os detalhes da peça). Ao contrário da música instrumental, tocada ao vivo em concerto, onde o executante tem a possibilidade de transformar a execução da obra em função da acústica da sala, na música acusmática pré-gravada tal possibilidade não tem um paralelo direto, ficando a obra de certa forma refém do espaço em que vai ser tocada.

No caso da tradição da difusão em concerto de obras acusmáticas em *Acousmoniums*, ou orquestras de alto-falantes, a partir de múltiplos alto-falantes, como proposta pelo GRM¹⁵⁰, entre outros, tenta-se, através da manipulação do material da composição, fazer um diálogo entre o *espaço interno* e o *espaço externo* da obra. Isso é feito a partir da exploração da própria acústica da sala onde a obra é reproduzida, através do controle delicado do som destinado a cada alto-falante, com o intuito de dar a conhecer melhor as características do *espaço interno* da obra, explorando

¹⁴⁹ Denominados *espaço composto* e *espaço de escuta*, respectivamente, pelo autor.

¹⁵⁰ *Groupe de Recherches Musicales*. Trad.: Grupo de pesquisas musicais (França).

artisticamente as características do espaço físico disponível, aproximando-se assim das possibilidades adaptativas da interpretação instrumental.

Refletindo sobre uma possível fidelidade à obra, podemos contrapor essa influência do *espaço externo* na recepção da obra pelo ouvinte, em alguns casos considerada negativa e prejudicial ao entendimento da obra acusmática do ponto de vista do previsto pelo compositor, ao considerarmos a performance de obras que chamamos de *site-specific*, que tem no local de exibição ou instalação o ponto forte de sua performance. Tais obras tem como princípio de elaboração, a exploração de características particulares de determinado espaço com parte da obra em si. Obras como *Adsonore*, desenvolvida por Natasha Barrett (2005), ou as mencionadas por Klein (2009) são somente uma amostra das diversas possibilidades vislumbradas pelos artistas que optam por desenvolver seus trabalhos destinados a um espaço específico.

Robusté (2014) desenvolve sua tese de doutorado, sobre a audição espacial e a percepção sonora na composição musical, a partir dessa consideração de que o *espaço externo* transforma a percepção sonora da música reproduzida. A partir disso ele desenvolve suas composições explorando tais características do espaço físico da performance ao usar somente materiais invariáveis no tempo, tais como ondas sinusoidais e ruídos. O objetivo principal do seu trabalho seria o de mostrar a possibilidade de se estabelecer "relações causa-efeito entre o movimento do som e a percepção de variações de características do som, variações do ritmo e textura", assim como o de "criar campos sonoros difusos, independentemente dos níveis de reverberação da sala, e de criar espaços sonoros com uma certa profundidade, sem utilizar nenhum tipo de *delay* ou reverberação artificiais".

Dentro do que denominamos aqui de *espaço interno* observamos o que Dignart denomina de coerência ou incoerência espacial, "definida a partir de relações espaciais entre as fontes sonoras num contexto composicional" (DIGNART, 2015, p. 92). Dentro do *espaço interno*, Mary (2013) aponta também o que se pretende pelo compositor como uma polifonia do espaço (evolução espacial simultânea e independente de vários elementos constituintes da obra) ao se realizar uma orquestração eletroacústica (fusão e fissão espectral, ou seja, soma e dissociação de componentes formativas do som que podem criar timbres mais ou menos complexos).

Ainda dentro do conceito de *espaço interno* de uma obra musical, o compositor tem a possibilidade de utilizar o deslocamento do som no espaço para criar objetos composicionais reconhecíveis para o ouvinte, e que têm o potencial de adquirir um significado semântico, muitas vezes associados à elementos derivados da própria tradição analítica e composicional do ocidente: tensões espaciais (apresentação ambígua de fontes sonoras) ou relaxamentos espaciais (articulação espacial coerente, desobstruída e estável); conflito espacial (também definido na literatura como dissonância espacial - SMALLEY, 2008, p. 123) ou consonância espacial (DIGNART, 2015, p. 86); entre outros.

Tais termos e conceitos, *tensão e relaxamento, conflito e consonância*, carregados de um significado imposto pela tradição da música ocidental de concerto, têm um potencial referencial, útil para a comunicação de ideias relacionadas ao espaço e sua manipulação, mas que são utilizados para descrever parâmetros muito subjetivos e diversos aos para os quais eles foram originalmente criados. Apesar disso, a liberdade e as novas possibilidades que a tecnologia permitiu na produção musical pós meados do século XX, oferecem ao compositor o potencial de expansão desses conceitos para novas relações, que veremos estar relacionadas tanto ao nível do timbre quanto ao modo de difusão ou reprodução sonora, por exemplo

Esse *espaço interno* como definido aqui ainda pode ser subdividido entre o *espaço intrínseco* e o *espaço extrínseco*. O que alguns autores (DIGNART, 2015; HENRIKSEN, 2002; CHION, 2009) denominam de *espaço intrínseco*, o que se relaciona à questões espectromorfológicas¹⁵¹ do objeto sonoro ou fonte sonora, sendo sua altura (*pitch*), amplitude, formato e densidade (timbre) constituintes desse espaço percebido somente em um nível metafórico, conceitual. Já o *espaço extrínseco* diz respeito à noções ligadas ao espaço como percebido fisicamente, à localização de um determinado objeto sonoro ou fonte sonora: seu movimento, sua direção e distância fazem parte desse conceito de espaço.

Dentro do conceito de *espaço intrínseco*, podemos observar no processo de análises melódicas e harmônicas, direcionadas à música baseada em notação tradicional, o uso do termo intervalo para definir se um espaço entre duas notas, do ponto de vista da altura, é maior ou menor do que entre outras duas notas. Uma outra

¹⁵¹ Espectromorfologia: conceito desenvolvido por Smalley (1997, p. 107) para definir a evolução timbrística e de amplitude de um objeto sonoro ao longo do tempo, características intrínsecas de um objeto sonoro.

utilização do termo *espaço* nesse domínio pode ser observada na relação temporal entre duas notas, ou até mesmo à duração de uma nota em si.

Como exemplo do *espaço intrínseco* direcionado à percepção de altura, tomamos como exemplo Lewin (2007), que descreve essa utilização desse *espaço* de uma forma sistemática e se utiliza de recursos matemáticos para estabelecer o que ele denomina de Sistema Intervalar Generalizado e desenvolver suas transformações. O termo *espaço* com essa mesma abordagem também é utilizado por Xenakis:

A noção de linha, na verdade, implica em movimento, então Xenakis observa que a forma particular de um glissando depende da velocidade de um ponto sonoro hipotético que se move continuamente dentro de um **espaço** imaginário temporal de altura.¹⁵² (ILIESCU, 2005, p. 2, grifo nosso)

No *Guia para os objetos sonoros*¹⁵³ de Michel Chion (1983), também vemos um bom exemplo da utilização do termo *espaço* dentro dessa definição que estabelecemos de *espaço intrínseco*. Referindo-se à obra de Schaeffer e ao que ele denomina de *campos perceptuais da altura*, da intensidade e do tempo ou duração, o autor estabelece as três dimensões em que os objetos sonoros podem ser manipulados e exemplifica uma delas:

Vamos ouvir um objeto tão simples quanto um glissando de violino. Qual é o critério dominante em cada instante desse som? Altura. O que varia? Altura novamente. Em que **espaço** isso varia? No domínio da altura. (CHION, 1983, p. 119, grifo nosso)

O que Schaeffer chama de tipos e classes da massa de um objeto sonoro, em seu *Tratado dos objetos musicais*¹⁵⁴ (aqui extraído do *Guia para os objetos sonoros* de Chion) também se enquadra nessa mesma classificação de *espaço intrínseco*. Os descritos “tipos de massa” (tônica, complexa, variável, outras) e as “classes de massa” (tom puro,ônico, grupoônico, canelado, grupo nodal, nó, ruído branco) são as classificações em que esse conceito de *espaço intrínseco* aparece de forma mais evidente, por estarem estreitamente relacionadas ao conteúdo espectral, à morfologia, do objeto sonoro.

¹⁵² Original: “The notion of line actually implies that of motion, so Xenakis remarks that the particular shape of a glissando depends on the speed of a hypothetical sound point that continuously moves within an imaginary pitch-time space.”

¹⁵³ Original: *Guide to Sound Objects*

¹⁵⁴ Original: *Traité des objets Musicaux*

Uma expansão desse mesmo conceito de Chion, referente ao domínio das frequências, agora aplicado à música acusmática, pode ser observada no conceito de Henriksen (2002, p. 43) e Smalley (1997) de *espaço espectral*, que consiste no conjunto de alturas utilizadas na obra, na geração de noções de esvaziamento ou preenchimento, entre outros. Segundo suas próprias palavras: “o espaço espectral cobre a distância entre o som audível mais grave e o mais agudo” (SMALLEY, 1997, p. 121). Smalley também define quatro qualificadores do espaço espectral, que dizem respeito a distribuição de energia no domínio da frequência. São eles:

- 1) Esvaziamento/Preenchimento – Se o espaço é extensivamente coberto e cheio, ou se as espectromorfologias ocupam áreas menores, criando lacunas dando a impressão de esvaziamento ou talvez de isolamento espectral.
- 2) Difusão/Concentração – Se o som é espalhado ou disperso através do espaço espectral ou se ele é concentrado ou condensado em regiões.
- 3) Estratos/Interstícios – A estratificação do espaço espectral em filetes estreitos ou largos separados por espaços intermediários.
- 4) Sobreposição/Cruzamento – Como os estratos ou espectromorfologias invadem o espaço espectral um do outro, se se movem em torno ou através da região um do outro. Isso está diretamente relacionado com os processos de movimento e crescimento. (SMALLEY, 1997, p. 121)

Portanto, como demonstram esses exemplos, temos muitos conceitos morfológicos e sintáticos, dentro daqueles citados na nota da pag. 96, que carregam uma significação espacial metafórica, sem efetivamente descreverem uma articulação no espaço físico tridimensional. Nosso trabalho, no entanto, não quer discutir esse espaço intrínseco, mas sim o extrínseco, no qual encontramos muita contradição na literatura, principalmente ao compararmos a definição adotada por Barrett e Smalley, e a de Henriksen. Por razões que ficarão mais claras adiante, adotamos aqui o definido por Henriksen:

O *espaço extrínseco* tem a ver com o som **no** espaço. Refere-se ao espaço que envolve o som como resultado de seu comportamento espacial e sua interação com o ambiente em que ele existe. O *espaço extrínseco* é percebido em termos de movimento, distância e direção dos sons. A localização no espaço físico e / ou virtual é baseada no *espaço extrínseco*.¹⁵⁵ (HENRIKSEN, 2002, p. 18, grifo nosso)

¹⁵⁵ Original: “Extrinsic space has to do with the sound in space. It refers to the space surrounding the sound as a result of its spatial behaviour and interaction with the environment in which it exists. Extrinsic space is perceived in terms of movement, distance and direction of sounds. Location in physical and/or virtual space is based on extrinsic space.”

Com sua origem nas antífonas religiosas, onde essa designação se refere ao posicionamento, dentro de um ambiente acústico, de determinada fonte sonora ou grupos de fontes sonoras durante a execução de uma obra, esse tipo de espaço, foi especialmente desenvolvido a partir do início do século XX, tornando-se um elemento importante, ou mesmo determinante, no pensamento composicional desse novo tempo. É importante observar que a elaboração do *espaço extrínseco* só se dá no *espaço externo*, quando uma obra é executada em um ambiente qualquer, ou em um sistema de reprodução qualquer.

Partimos então da espacialização de áudio como tratada pela música acusmática, principalmente pelo fato de ela ter sido uma das primeiras a desenvolver e explorar tanto técnica quanto esteticamente as possibilidades de sistemas multicanais de difusão sonora. Com o desenvolvimento desse gênero musical e a implementação da reprodução através de alto-falantes, a exploração do *espaço extrínseco* de uma obra foi-se tornando cada vez mais diversificada, tornando-se, muitas vezes, essencial para a construção e o entendimento da mesma. É precisamente nesse tipo de música, onde a inexistência de fontes sonoras instrumentais fisicamente conectadas ao som projetado permite uma libertação do ouvinte para novas formas de escuta, que a expansão dos conceitos e métodos de difusão sonora se realizou com mais eficácia.

Segundo Henriksen (2002, p. 33), é o *espaço extrínseco* que nos permite perceber informações relativas a direção, altura, distância e movimento de um objeto sonoro, assumindo que o mesmo é ouvido como parte de um campo sonoro construído e que só pode ser observado em relação a um ponto de escuta. Refere-se então a um som *no* espaço, espaço físico em que nos posicionamos para ouvir algo e, apesar de estar intimamente relacionado com informações espectromorfológicas, resulta da interação entre o objeto sonoro com o ambiente em que ele existe, e com o sistema de reprodução, que nos dá pistas psicoacústicas de sua localização. O autor estabelece ainda algumas categorias do som dentro do *espaço extrínseco* tais como sons direcionais e não direcionais, e os sons em movimento (de profundidade, lateral, de elevação, dispersão e convergência) (HENRIKSSSEN, 2002, p. 40-41).

Um outro trabalho que apresenta um conceito que podemos incluir no que definimos aqui como *espaço extrínseco*, apesar de direcionado aos estudos de

instrumentos acústicos, é o de Koechlin (2006, p. 139)¹⁵⁶, ao se referir ao volume (referente não à intensidade mas à percepção de dimensões) de uma fonte sonora. Sabemos pelos estudos de acústica que fontes sonoras mais graves se espalham mais uniformemente e de forma omnidirecional, e que quanto mais agudo for o som produzido por uma fonte sonora, mais direcional tende a ser sua projeção. Tais características fazem parte do que percebemos com relação às dimensões de uma fonte sonora e são válidas tanto para fontes sonoras acústicas quanto para sons reproduzidos por alto-falantes e que, no contexto de composições acusmáticas, se torna extremamente importante na escolha do posicionamento de um objeto sonoro no *espaço extrínseco*. Tal parâmetro também aparece com o nome de calibre na obra de Schaeffer (CHION, 1983, p. 118) e em diversos livros que tratam de análise subjetiva de parâmetros espaciais de áudio gravado, sendo expandido nesses para o volume, ou tamanho percebido, tanto de fontes sonoras individuais, quanto de grupos ou salas de performance (RUMSEY, 2002; BECH e ZACHAROV, 2006). Essa dimensão de tamanho do objeto sonoro, também é definida aqui como estando no domínio do *espaço extrínseco*.

Ainda se referindo ao conceito de *espaço extrínseco*, Merlier (2006, p. 132; *apud* DIGNART, 2015, p. 145) classifica a variedade de formas em que ele pode se apresentar, no que ele denomina de parâmetros característicos da percepção espacial, divididos entre estáticos e dinâmicos. Tal classificação se atém ao nível sintático do que definimos como *espaço extrínseco*. Merlier divide tais parâmetros entre os estáticos como sendo os relacionados à matéria (densidade, textura), ao formato (largura, tamanho, área), a posição, a distância, se interior / exterior, ao envolvimento, ao ambiente / lugar; e os dinâmicos seriam os que possuem estreita relação com o movimento das fontes sonoras apresentadas (mobilidade, trajetória, direcionalidade, crescimento), variação de forma, de direção, de lugar, de densidade, aproximação / afastamento.

¹⁵⁶ Referência a republicação de parte do texto original do autor na antologia de Paul Mathews de 2006. O original em si é de um tratado de quatro volumes datados de 1954-9 (KOECHLIN, 1954-9).

Tabela 2: Parâmetros característicos da percepção espacial (adaptação de MERLIER, 2006, p. 132).

Parâmetros característicos da percepção espacial	
Estáticos	Dinâmicos
Matéria (densidade, textura)	Movimento (mobilidade, trajetória,
Formato (largura, tamanho, área)	direcionalidade, crescimento)
Posição	Variação de forma
Distância	Variação de direção
Interior/ exterior	Variação de lugar
Envolvimento	Variação de densidade
Ambiente/ lugar	Aproximação/ afastamento

A tabela de Merlier apresentada acima, assim como outras duas encontradas em seu trabalho, abrangem uma conceituação de espaço desenvolvida por diversos autores mas que não deixa claro cada uma das categorias como as que propomos aqui. A partir dessa proposta, apresentamos então uma forma de aglutinar os diversos outros conceitos de espaço em música e os enquadrámos em quatro categorias inter-relacionadas: o *espaço interno* e o *espaço externo*, o *espaço intrínseco* e o *espaço extrínseco*.

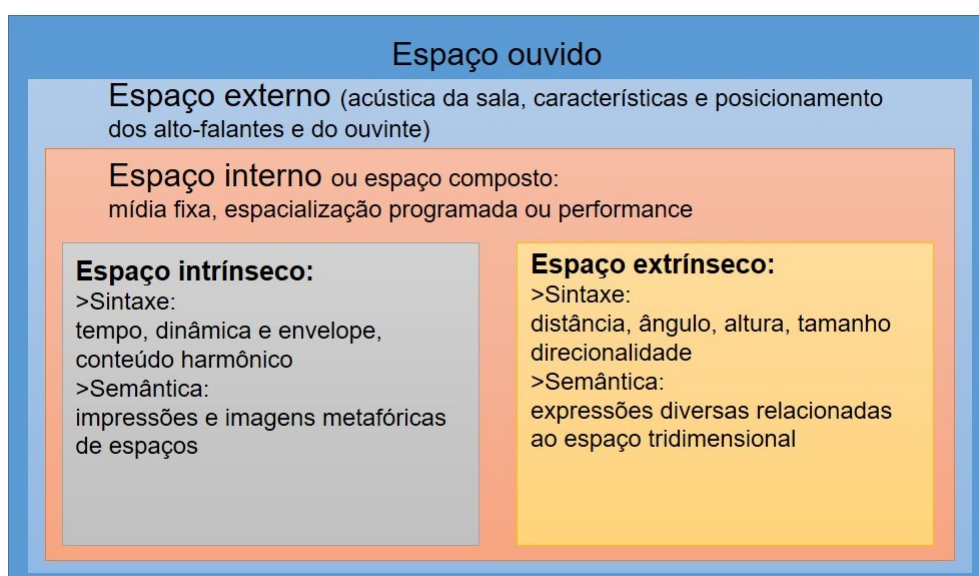


Figura 33 - Aglutinação das terminologias do conceito de espaço em música proposta pelo presente autor.

Apesar de segregados nessa representação, os diferentes conceitos de espaço se comunicam, interferem e sofrem interferências entre si. O *espaço interno* de uma obra, como proposta pelo compositor, é influenciado pelo *espaço externo* quando da reprodução em concerto, ou executado em um outro sistema de reprodução que não o utilizado no processo composicional. De uma forma reversa o compositor pode pensar em manipular o *espaço interno* de sua obra objetivando sua reprodução em um determinado *espaço externo*, aproveitando ou se adaptando a suas características, como já visto em obras *site-specific*.

Dentro do *espaço interno* de uma obra, se encontram ambos os *espaços intrínseco* e *extrínseco* sobre os quais o compositor tem controle em seu estúdio. A necessidade de se separar as abordagens sintática e semântica, surge da proposta de unir conceitos estabelecidos tanto por compositores ingleses (Smalley e Barrett) quanto franceses (Bayle, Schaeffer e Chion). Dentro do conceito de *espaço intrínseco* estão: do ponto de vista sintático, o espaço dos domínios das alturas, dinâmicas, temporais, e conteúdo harmônico (espectral); e, do ponto de vista semântico, as impressões ou imagens sonoras a quais os objetos sonoros nos remetem, fruto do nosso conhecimento de mundo e inerentes às características morfológicas que identificamos no próprio objeto. Já no conceito de *espaço extrínseco* estão: do ponto de vista sintático, a própria movimentação nos três eixos físicos do espaço (distância, ângulo e altura), assim como o próprio tamanho percebido da fonte sonora e qualquer direcionamento de movimento que ela venha a apresentar; e, do ponto de vista semântico, diversas expressões referentes ao espaço tridimensional, entre eles, a possibilidade da criação de uma ilusão de um espaço real, a geração de espaços envolventes ou pontilistas, assim como uma estruturação de gestos espaciais como os proposto pelas 15 *figuras do espaço* de Vande Gorne¹⁵⁷, dentre outros.

Dessa forma o gráfico abaixo ilustra, de forma breve, as possibilidades de relação entre os diferentes conceitos de espaço: do *espaço interno* com o *externo*, e vice-versa, assim como, dentro da obra, os *espaços intrínseco* e *extrínseco* se relacionando mutuamente.

¹⁵⁷ Os 15 gestos descritos por Vande Gorne são: misturar, desmanchar, ênfase, tremular, oscilar, balançar, onda, rotação, espiral, rebote, inserção/ ruptura, aparecimento/ desaparecimento, explosão, acumulação, invasão e serão abordados na seção 3.2.3.



Figura 34 - Relações entre os conceitos de espaço em música proposta pelo presente autor.

Nessa síntese, podemos observar também que uma ou outra abordagem observada na literatura foge à nossa conceituação, ou melhor, combina dois ou mais desses conceitos principais. Uma combinação interessante, por exemplo, entre dois desses conceitos de espaços (o *espaço intrínseco* e o *espaço extrínseco*), pode ser observado no trabalho de Normandeau (2009), reconhecido compositor de música eletroacústica canadense:

Isso é o que chamo de espacialização do timbre: todo o espectro de um som é re combinado apenas virtualmente no espaço da sala de concertos. Cada ponto representa apenas uma parte do conjunto. Não é uma concepção de espaço que é adicionada no final do processo de composição - uma abordagem frequentemente vista, especialmente hoje com software multipista - mas uma espacialização verdadeiramente composta. É um parâmetro musical exclusivo da música acusmática.¹⁵⁸ (NORMANDEAU, 2009, p. 3)

Essa estratégia de filtrar o som da obra e distribuir suas diferentes regiões em diferentes pares ou grupo de alto-falantes, no entanto, já havia sido descrita e

¹⁵⁸ Original: "This is what I call timbre spatialisation: the entire spectrum of a sound is recombined only virtually in the space of the concert hall. Each point represents only a part of the ensemble. It is not a conception of space that is added at the end of the composition process – an approach frequently seen, especially today with multitrack software – but a truly composed spatialisation. It is a musical parameter that is exclusive to acousmatic music."

implementada por Christian Clozier (1998) no Gmebaphone de Bourges, na França, em um dos primeiros *Acousmoniuns* implementados:

O princípio fundamental do Gmebaphone é a divisão e, em seguida, a adição eletrônica de sons por um conjunto de filtros específicos do extremo grave ao extremo agudo (como na óptica, onde um prisma divide a luz em cores fundamentais), em registros de timbres que são transmitidos e projetados acusticamente por alto-falantes especializados, que não têm mais função individual e idêntica, mas sim a de reprodução dos elementos básicos do sintetizador acústico que é o Gmebaphone, sintetizador que também suporta adaptações às características acústicas dos diferentes locais de concerto.¹⁵⁹ (CLOZIER, 1998, p. 56)

2.1.1 Transgressões dos conceitos de espaço

Partindo agora para as possibilidades de transgressão dos conceitos de *espaço* já estabelecidos, podemos observar, durante o processo de espacialização do material sonoro, uma série de possibilidades limítrofes, ou seja, transformações de um tipo de *espaço* em outro, eventualmente interessante do ponto de vista estético, outras vezes indesejado.

Retomando o conceito de *espaço externo* e refletindo sobre como as transformações que o mesmo impõe à obra podem afetar o nosso objeto de estudos, o áudio tridimensional para realidade virtual, observamos que, pela sua natureza reduzida aos fones de ouvido, conseguimos, até certo ponto, transformar o *espaço externo* em *espaço interno*. Nesse processo excluimos diversos problemas relatados por compositores ao ouvirem suas obras em espaços maiores, quando difundidas para grandes públicos. Em contrapartida, observamos esse aumento de responsabilidade sobre o compositor na criação de espaços virtuais convincentes e, eventualmente, na necessidade de criar uma versão para a realidade virtual e outra versão para concerto ou execução através de alto-falantes. Ao nos limitarmos a reprodução através dos fones de ouvido e sistemas *Binaurais*, quase isentamos a obra da interferência do *espaço externo*. Obviamente isso não ocorre por completo pois, uma vez que cada fone de ouvido, de marca e modelo distintos, possui uma certa

¹⁵⁹ Original: “Le principe fondamental du Gmebaphone porte sur la division puis une addition électronique des sons par un ensemble de filtres spécifiques de l'extrême grave à l'extrême aigu (ansi qu'en optique un prisme divise la lumière en couleurs fondamentales), en registres de timbres qui à la diffusion sont projetés acoustiquement par des haut-parleurs spécialisés, lesquels n'ayant plus de fonction individuelle et identique, constituant pour la reproduction les éléments de base du synthétiseur acoustique qu'est le Gmebaphone, synthétiseur qui par ailleurs prend en charge l'adaptation aux contraintes acoustiques des différents lieux de concert.”

coloração, o espectador também observará diferenças na percepção de uma obra com diferentes fones. No entanto, o compositor pode considerar, para fins práticos, que tais colorações são mínimas e, processos como o descrito acima (espacialização do timbre), por exemplo, passam a fazer parte do *espaço interno* da obra em si, agora, independente de um sistema de filtragem, hardware ou alto-falantes físicos específicos. Tais manipulações sonoras agora, passam a ser responsabilidade do compositor, que deverá realizar as devidas filtrações em software, assim como a espacialização dos sinais resultantes nos infinitos pontos ao redor do ouvinte, que receberá o resultado final através somente de um par de fones de ouvido.

Com relação à manipulação do que denominamos de *espaço externo*, podemos agregar a resposta acústica de uma determinada sala ao *espaço intrínseco* de um objeto sonoro. Quando uma determinada fonte sonora se propaga em um espaço acústico, ela cria não apenas reflexões primárias e secundárias como, também, o reforço e atenuação em determinadas regiões do espectro ao longo do tempo. A resposta da acústica de uma sala a um impulso ou ruído, ou mais comumente uma varredura contínua de frequências¹⁶⁰, pode ser gravada e transformada no que denominamos de *IRs*. Tais *IRs*, ou respostas a impulso, podem ser aplicados a outro áudio qualquer e fazer soar como se aquele objeto sonoro tivesse sido gravado em tal espaço acústico, o que é o princípio básico da maioria dos processadores de *reverb* baseados em convolução (DERUTY, 2010; ADRIAENSEN, 2006). Esse processo faz com que essa resposta da sala passe então a fazer parte do timbre propriamente dito, criando um novo objeto sonoro, com parte do seu *espaço intrínseco* transformado. Tais procedimentos são comumente aplicados a objetos sonoros, pelos compositores de música acusmática, com dois objetivos distintos: o de representar um *espaço externo* reconhecido ou desejável, em um nível semântico, ou o de alterar o timbre e envelope dos objetos sonoros dentro da composição.

Natasha Barrett (2002, p. 320), ao dissertar sobre a criação de uma *ilusão de espaço*, a primeira das abordagens de espaço descritas pela autora, menciona e dá um exemplo da transformação de uma reverberação que simula um espaço pequeno e que pode ser interpretada como parte do objeto sonoro em si, parte do que denominamos de *espaço intrínseco*. Como dito anteriormente, tal prática é comum entre os compositores de música eletroacústica com o intuito de prolongar o objeto

¹⁶⁰ *Continuous Frequency Sweep*.

sonoro para além da sua duração original. Um processo contrário também pode ser proposto se considerarmos que a percepção do *espaço extrínseco* (ângulo, altura e distância), em sistemas *Binaurais*, é obtido através da incorporação de *HRTFs* (*IRs* criados a partir de gravações com *Dummy Heads* de fontes sonoras posicionadas em um determinado ponto no espaço de escuta) aos objetos sonoros que se queira espacializar. Tais *HRTFs* são representadas também como filtros dinâmicos, que atuam no domínio da frequência, e que são direcionados para cada uma das duas orelhas do ouvinte. Dessa forma, uma manipulação no domínio do *espaço intrínseco* de um objeto sonoro, estaria nos servindo para adicionar informações referentes ao *espaço extrínseco* do objeto sonoro.

O exemplo provavelmente mais comum de uma transformação no *espaço extrínseco*, seria o de transformar o *espaço extrínseco* em si, em *espaço intrínseco*. Quando determinado objeto sonoro realiza trajetórias cada vez mais rápidas, chega um determinado ponto em que não distinguimos mais sua localização se alterando, mas sim uma mudança, ou modulação, no timbre do que ouvíamos anteriormente. Tal transformação é muito conhecida pelos guitarristas de rock com o nome de efeito *rotary*, em que o som da guitarra viaja de um alto-falantes para outro do sistema estéreo em diferentes velocidades. Diversos compositores exploraram tal transformação em suas composições, dentre eles podemos citar Stockhausen que, em sua obra *Kontakte*, utiliza uma série de movimentos espaciais. Segundo ele, tais movimentos seriam capazes de articular durações, parâmetro naturalmente pertencente ao que denominamos de *espaço intrínseco*¹⁶¹ (BATES, 2009, p. 138). Felder, ao descrever a obra *Sirius*, também de Stockhausen, afirma que “o som se move tão rápido em rotações e rampas e todo tipo de movimento espacial que parecia estar parado, mas vibrava”¹⁶² (FELDER, 1977, apud BATES, 2009, p. 143). Como um exemplo de trabalho mais recente nesse domínio podemos citar a pesquisa de Ryan McGee (2019), onde esse fenômeno é levado ao extremo, gerando diferentes timbres e sendo ainda associado a geração de imagens, processo que o autor denomina de síntese por modulação espacial¹⁶³.

¹⁶¹ Original: “Changes in spatial location can articulate durations in exactly the same way.”

¹⁶² Original: “The sound moves so fast in rotations and slopes and all sorts of spatial movements that it seems to stand still, but it vibrates”

¹⁶³ Original: “Spatial Modulation Synthesis”

Essa transformação em específico, da manipulação do *espaço extrínseco* até que ele seja percebido como intrínseco, é mencionada por Barrett (2002, p. 320) e desenvolvida por Robusté (2014). Baseado na proposta apontada por Stockhausen, no que ele denominava de “novo começo”¹⁶⁴, Robusté aponta a “necessidade de encontrar uma nova abordagem estética que permita incorporar a espacialização sonora para modificar a percepção de intensidade, altura e variação da cor tonal, e enfatizar essas variações ao incorporá-las como parte do discurso musical”¹⁶⁵ (ROBUSTÉ, 2014, p. 13).

Ao contrário de Smalley, que afirma que “a percepção espacial está inextricavelmente ligada ao conteúdo espectromorfológico, e a maioria dos ouvintes não pode facilmente apreciar o espaço como uma experiência em si mesma”¹⁶⁶ (SMALLEY, 1997, p. 122), Robusté desenvolve sua tese defendendo que seu

objetivo é usar a espacialização sonora como parte da estrutura formal do próprio som, em vez de usá-lo como parte da estrutura formal da música. Isto implica no fato de que a espacialização sonora e o movimento sonoro alteram a percepção dos atributos sonoros, por um lado, e criam os eventos musicais próprios, por outro.¹⁶⁷ (ROBUSTÉ, 2014, p. 13)

O autor elabora obras baseadas em sons simples (sinusoidais e ruídos) com o intuito de testar tal abordagem e conclui que as variações ao se utilizar tais sons são muito sutis, deduzindo que as mesmas seriam mascaradas quando do uso de sons mais complexos.

2.1.2 A espacialização sonora em tempo real como prática interpretativa da música eletroacústica

Formulei rapidamente a opinião de que a espacialização era um aspecto crítico e necessário da música eletrônica, já que o movimento espacial dos

¹⁶⁴ Original: “Neuanfang” ou “New Beginning”

¹⁶⁵ Original: “(...) to the necessity to find new aesthetic approaches that permit to incorporate sound spatialization to modify the perception of loudness, pitch and tone color variations and to emphasize these variations to incorporate them as part of a musical discourse.”

¹⁶⁶ Original: “Spatial perception is inextricably bound up with spectromorphological content, and most listeners cannot easily appreciate space as an experience in itself.”

¹⁶⁷ Original: My aim is to use sound spatialization as part of the formal structure of sound itself, instead of using it as part of the formal structure of music. This implies that sound spatialization and sound movement alter the perception of the sound attributes on the one hand, and create the musical events themselves on the other

sons proporcionava fisicalidade e dinamismo, o que aumentava muito a experiência auditiva.¹⁶⁸ (BATES, 2009, p. 18)

A manipulação em tempo real do *espaço extrínseco* como elemento estético pela música eletroacústica é único e inerente a esse gênero mas enfrenta a dificuldade da disseminação. No mercado musical, apesar da redução de vendas de mídias físicas e aumento da busca pela música pela internet, esse gênero ainda se beneficia de concertos ao vivo que exploram a espacialização em tempo real, transformando a difusão de música eletroacústica em uma performance (BATES, 2009, p. 23). A música acusmática, com sua característica própria de se difundir através de alto-falantes, transformou a prática da difusão sonora em uma prática interpretativa, que se utiliza dos diversos sistemas de alto-falantes, sejam eles estereofônicos ou multicanais, para transformar a escuta de uma determinada obra em uma experiência mais rica.

Nesse processo de difusão sonora, definida por alguns autores como uma projeção e espalhamento do som em um ambiente acústico para um grupo de ouvintes, em oposição à escuta dentro de um espaço confinado e individual, estão embutidos conhecimentos tanto da obra quanto do sistema de reprodução disponível, através do qual o performer manipula tanto o *espaço extrínseco* da obra em si, quanto adapta a mesma ao *espaço externo* da obra, o espaço de escuta (AUSTIN e SMALLEY, 2000).

A criação ao vivo e em frente ao público de gestos musicais no espaço, ou da criação de novos gestos, leva em consideração tanto a estrutura interna da obra quanto a resposta que se pretende obter do espaço de escuta, do concerto. A execução da difusão sonora no espaço traz interesse aos concertos de música acusmática por ser uma “interpretação” de uma obra conhecida do performer. A imprevisibilidade da abordagem do espaço escolhida pelo performer faz com que cada apresentação seja única.

Muitas obras acusmáticas já são concebidas e espacializadas para sistemas multicanais, outras, no entanto são finalizadas em um formato estereofônico que, ao ser reproduzido através de tais sistemas, é multiplicado para tirar proveito da

¹⁶⁸ Original: “I very quickly formed the opinion that spatialization was a critical and necessary aspect of electronic music, as the spatial movement of sounds provided a physicality and dynamism that greatly enhanced the listening experience.

variedade de alto-falantes disponíveis, possibilitando um maior envolvimento do público e a prática de uma interpretação da obra difundida.

Quando a obra já é concebida para um sistema multicanal, ainda assim ela pode ser multiplicada, se o sistema disponível permitir. Como exemplo disso podemos imaginar uma obra composta originalmente para um sistema quadrifônico sendo reproduzida por um arranjo de 32 alto-falantes, ou uma obra desenvolvida inicialmente para um sistema com alto-falantes posicionados somente no eixo horizontal, sendo expandida para um sistema com altura (anéis de caixas acústicas montados em cima e embaixo do público).

Nesse contexto da performance espacial da música acusmática, Smalley (AUSTIN e SMALLEY, 2000) descreve quatro razões para que tal multiplicação de canais ocorra e aponta algumas possibilidades interpretativas desse tipo de abordagem de um sistema multicanal:

- 1) Para que se tenha uma boa imagem estéreo da obra reproduzida (como pretendido pelo compositor), os alto-falantes no formato estereofônico, devem estar distanciados entre si, e em relação ao ouvinte, de uma determinada maneira que, em grandes espaços, se apresenta de uma forma diferente para cada ouvinte, para alguns mais próximo, para outros mais afastado. Dessa forma, quanto mais afastado o ouvinte se encontra dos alto-falantes menor é a imagem estéreo percebida. A utilização de diversos pares de alto-falantes permite que essa deficiência seja suprimida ou compensada.
- 2) Na composição da obra existem diversas perspectivas espaciais pretendidas pelo compositor. Quando essa obra é reproduzida por múltiplos sistemas, os mesmos podem ser utilizados para exagerar essas perspectivas reproduzidas, adaptando o *espaço extrínseco* representado na música ao *espaço externo* em que ela é reproduzida.
- 3) Um outro objetivo de se espacializar uma obra já fixada em mídia estereofônica pode ser a de criar uma dramatização do espaço. Uma vez que os diversos alto-falantes podem criar uma variedade de perspectivas, tão grande quanto o próprio sistema multicanal disponível, a textura espacial e sua tipologia podem ser manipuladas pelo performer de diversas maneiras diferentes.

- 4) A faixa dinâmica espacial prevista pelo compositor no seu estúdio de trabalho, sem sombra de dúvida é menor do que a disponível nas salas de apresentação. Com isso a necessidade de se expandir essa faixa dinâmica espacial da obra se torna evidente para a adaptação da mesma à realidade da performance.

Inicialmente a obra musical acusmática possui um formato fixado pelo compositor em uma mídia física, analógica ou digital, na qual já está contido e realizado tanto o seu planejamento artístico do ponto de vista do *espaço intrínseco* (alturas, timbres, ritmos gestos e texturas) quanto do ponto de vista do *espaço extrínseco*. O ouvinte que simplesmente ouve uma reprodução dessa mídia já estará se relacionando com a obra na perspectiva original do compositor. No entanto, a espacialização sonora como performance, além de ampliar aspectos considerados importantes pelo intérprete, enriquecendo a percepção da obra em si, vem também, historicamente, resolver um grande problema no que diz respeito a aceitação da música acusmática e dessa prática em situação de concerto, problema este apontado por Amorim quando ele afirma que:

(...) a música eletroacústica, na sua forma mais primitiva, elimina o intérprete e atribui uma nova função para o compositor, que passa a ser responsável não só pela construção do material sonoro enquanto estrutura formal, mas também como um objeto físico que contém latente a performance da obra. (AMORIM, 2013 p.54)

O fato de se ter uma obra fixada, passível de ser apreciada em qualquer momento, torna o concerto de música acusmática um pouco vazio e sem sentido. Por que ir ao teatro ouvir uma música que posso ouvir no conforto da minha sala de estar ou do meu estúdio particular? Somente pra ouvir uma obra em um espaço melhor e com equipamentos melhores? Diversos autores e compositores consideram a presença física e a atuação de um intérprete nesse tipo de concerto, um catalisador para a escuta, que consegue chamar a atenção do ouvinte para sua atitude física, além, é claro, de ser o responsável por adaptar a obra ao espaço de escuta, manipulando a velocidade dos movimentos e as mudanças de planos, utilizando ferramentas tais como o posicionamento das caixas, o controle de volume e a equalização (VANDE GORNE, 2002; AMORIM, 2013, p.55).

A espacialização em tempo real se apresenta então como uma das formas de fazer com que um concerto de obras acústicas se torne uma experiência válida ao adicionar uma certa incerteza e uma certa maleabilidade à escuta. Podemos acrescentar aqui que experiências com a iluminação do espaço de concerto e / ou a inclusão de vídeos possuem um objetivo semelhante, assim como a composição de obras mistas que somam o som de um instrumento acústico com uma parte fixada¹⁶⁹.

No desenvolvimento da prática de música eletroacústica, observamos uma dicotomia entre a música concreta de Pierre Schaeffer e a música eletrônica de Stockhausen, inicialmente no que diz respeito a escolha e desenvolvimento de materiais sonoros para a composição. Essa dicotomia pode ser observada também na abordagem da manipulação do *espaço extrínseco* nesse tipo de música por esses dois personagens particularmente. Stockhausen optava por explorar um sistema complexo de alto falantes como um sistema homogêneo, permitindo-o posicionar e movimentar fontes sonoras em qualquer ponto em torno do público, incorporando o *espaço extrínseco* como um novo parâmetro musical a ser serializado por exemplo. Já Schaeffer utilizava diferentes alto-falantes com o intuito de explorar suas características tonais, assim como sua posição no espaço diante dos ouvintes como ferramenta para intensificar uma característica já presente no objeto sonoro em si. Nas palavras de Bates (2009, p. 154): “a abordagem de Schaeffer enfatizava a escuta reduzida, que é o foco consciente das propriedades intrínsecas do som”¹⁷⁰.

Tomando como referência a primeira abordagem, observamos o desenvolvimento de diversos sistemas homogêneos baseados em *Ambisonics* (GERZON, 1973, 1975), *Vector Based Amplitude Panning* - VBAP - (PULKKI, 1997), *Distance Based Amplitude Panning* – DBAP – (LOSSIUS, 2009) e *Wave Field Synthesis* - WFS - (BOONE e VERHEIJEN, 1993) que podem ser implementados em locais que possuem uma infraestrutura direcionada para tal, como o *AlloSphere* (*University of Santa Barbara*¹⁷¹), o *BEAST* (*University of Birmingham*¹⁷²), o *Rymer*

¹⁶⁹ “Nas palavras do compositor argentino Mário Davidovsky (1934), a introdução do eletrônico de mãos dadas com o intérprete ao vivo iria diminuir o impacto de caixas de som gritando para você. A mistura de sons eletrônicos com instrumentos, que são mais familiares, torna os eletrônicos mais aceitáveis por parte do ouvinte”. (AMORIM, 2013, p.56)

¹⁷⁰ Original: “Schaeffer’s approach emphasized reduced listening, that is the conscious focus on the intrinsic properties of the sound.”

¹⁷¹ <http://www.allosphere.ucsb.edu/index.html>

¹⁷² <http://www.beast.bham.ac.uk/>

Auditorium (York University¹⁷³) ou, como divulgado comercialmente, com sistemas da *IOSONO*¹⁷⁴.

A abordagem de uma composição multicanal para um arranjo simétrico e homogêneo de alto-falantes possibilita efeitos espaciais extremamente elaborados mas traz à tona questões relativas a recepção da própria obra pelos ouvintes. Uma vez que os ouvintes irão se encontrar em diferentes posições em relação ao *sweet spot* do sistema, eles estarão submetidos a sons resultantes diferentes (BATES, 2009, p. 22), o que pode não ser o ideal para uma apreciação uniformizada da reprodução da obra.

Já a segunda abordagem, a da exploração de "orquestras de alto-falantes", ou *Acousmonium*, como ficou conhecido internacionalmente o sistema concebido por François Bayle em 1974 (MENEZES, 1996, p. 27), é composto por diversos pares de alto-falantes, com diferentes características tonais, e focada em adaptar uma obra estereofônica a um espaço específico preservando a imagem estéreo para todos os ouvintes (BATES, 2009, p. 21). Considerado o legado da música concreta, foi explorado desde as primeiras performances de Pierre Schaeffer com o *potentiomètre d'espace*, e possui uma forte comunidade internacional de compositores adeptos. Entre tais sistemas podemos citar o *Audior* de Eraldo Bocca¹⁷⁵, originário da Itália (Milão) mas com características móveis, e o teatro do grupo *Musiques & Recherches*¹⁷⁶, Bélgica – Bruxelas.

Segundo Bates:

Essa estética [a de uma orquestra de alto-falantes] foca no desenvolvimento temporal, espectral e espacial dos sons ao invés da relação entre esses parâmetros, e o processo de difusão é então usado para exagerar o conteúdo espectral, espacial e de dinâmica do material musical já existente em uma determinada obra.¹⁷⁷ (BATES, 2009, p. 150)

¹⁷³ <https://www.york.ac.uk/music/about/facilities/performance/#tab-2>

¹⁷⁴ <http://www.iosono-sound.com/home/>

¹⁷⁵ <https://www.audior.eu/acusmonium/>

¹⁷⁶ <http://www.musiques-recherches.be/fr/acousmonium/dispositions-techniques>

¹⁷⁷ Original: "This aesthetic focuses on the temporal, spectral and spatial development of sounds rather than the relationship between these parameters, and the diffusion process is therefore used to exaggerate the dynamic, spectral and spatial content of the musical material already present in the work."

Apesar de o processo de difusão sonora claramente adicionar algo a performance de obras acusmáticas, o performer está sempre limitado a enfatizar algum material pré-existente na obra que está sendo difundida (BATES, 2009, p. 154).

No nosso objeto de estudos, a mídia audiovisual adaptativa para realidade virtual, vemos a possibilidade de disseminação de obras eletroacústicas já espacializadas para qualquer arranjo de alto-falantes. Sua apresentação ao espectador, passa por um processo de remediação, conceito extraído de Rajewsky (2005, p. 60), ao qual denominamos um tipo particular de relação intermédias, que, através do processo de remodelação, tanto a antiga quanto a nova mídia evoluem em uma luta para o reconhecimento cultural. Aqui uma reprodução direcionada inicialmente para sistemas multicanais de alto-falantes, é remediada para um par de fones de ouvido, possibilitando então que uma produção musical que necessitaria de um aparato tecnológico caro e complexo tenha seu conteúdo distribuído com o mínimo de perdas por um dispositivo popular, barato e de fácil acesso. Temos que observar no entanto que perdemos, nesse processo, tanto o elemento da performance¹⁷⁸, quanto aquele fator variável do *espaço externo* que, através de fones de ouvido, afeta minimamente o conteúdo da mídia em si.

2.2 Em busca de uma estética do espaço extrínseco

(...) o movimento de fontes sonoras deve ser entendido em um contexto em que o árbitro principal é o grau em que o movimento, localização e imagem auditiva podem ser entendidos pelo ouvinte, com sucesso, como sendo integrado ao significado da experiência musical. Em outras palavras, uma vez que podemos movimentar uma fonte sonora ao redor do ouvinte e sabemos como fazê-lo, precisamos – em uma infeliz mistura de metáforas – nunca perder de vista o porquê de o estarmos fazendo.¹⁷⁹ (HARRISON, J. e WILSON, S. 2010, p. 183).

¹⁷⁸ A possibilidade de se transmitir via internet, áudio realizado e espacializado em um sistema baseado em *Ambisonics* em tempo real, assim como o vídeo equiretangular em formato 4K, ainda não existe com as ferramentas de que dispomos atualmente. Como observado nas produções desenvolvidas, o Youtube precisa de algum tempo (cerca de uma hora) para receber um áudio multicanal no formato em *Ambisonics* e transformá-lo em *Binaural* interativo para os fones de ouvido.

¹⁷⁹ Original: “Once again, this [moving sound sources] must be understood in a context in which the main arbiter is the degree to which movement, location and aural image can be successfully understood by the listener as being integrated within the meaning of the musical experience. In other words, whilst we know we can move sound around and we know how to do it, we must – in an unfortunate mixing of metaphors – never lose sight of why we are doing it”.

Como descrito na seção 2.1.2, a performance da música acusmática em si, há muito foi observada como deixando a desejar no que diz respeito à expressividade visual (BATES, 2009; AMORIM, 2013). Enquanto na música instrumental, o ouvinte é capaz de reconhecer os sons percebidos como reações aos gestos do performer sobre seu instrumento, o que torna a performance mais viva, nos concertos de música acusmática, essa relação de causa e efeito se perdeu completamente. Com isso em mente, Bates afirma que “a espacialização dinâmica dos sons eletrônicos, particularmente seu princípio de estruturação gestual / ambiental, como sugerido por Denis Smalley, pode potencialmente substituir o componente visual faltante e prover uma fisicalidade necessária e potencialmente emocionante”¹⁸⁰ (BATES, 2009, p. 216).

Pierre Schaeffer, em 1952, em seu trabalho que pretendia resumir os princípios básicos e métodos da música concreta, introduz pela primeira vez o conceito de um movimento sonoro ao longo de uma trajetória, e a criação de um relaxamento através do contraste entre localizações espaciais estáticas e fontes sonoras móveis controladas por um performer (BATES, 2009, p. 128). Begault (2000, p. 1) por sua vez, propõe que um sistema sonoro tridimensional possui limitações somente quanto à imaginação do compositor e às ferramentas que ele tem disponíveis. Inspirado por Begault e dando continuidade à proposta inicial de Schaeffer, pretendemos nessa parte do presente trabalho focar em questões estéticas da espacialização sonora, no que se refere especificamente à dimensão musical do *espaço extrínseco*, onde a observamos ainda pouco sistematizada.

2.2.1 Estratégias de espacialização.

Diversas estratégias para a espacialização sonora têm sido apresentadas tanto da área da música instrumental (BRANT, 1967, MERIC, 2011) quanto da música eletroacústica, área que faz uma certa divisa com a engenharia de som (BAALMAN, 2010; BARRETT, 2002; BEGAULT, 2000). Apresentamos a seguir a descrição de algumas delas que selecionamos no trabalho de quatro autores. Não pretendemos validar ou invalidar qualquer uma dessas estratégias, mas sim, a partir da observação de uma certa mistura entre abordagens em um nível sintático (relativo ao nível

¹⁸⁰ Original: “The dynamic spatialization of electronic sounds, particularly the gestural/environmental structuring principle suggested by Denis Smalley, can potentially replace or substitute for this missing visual component and provide a necessary and potentially thrilling physicality”.

material) com um nível semântico (relativo ao significado), justificar a necessidade da sua separação, para um melhor entendimento e desenvolvimento, como será explorado a seguir.

Marije Baalman (2010, p. 209-210), em seu trabalho sobre técnicas de composição espacial e tecnologias de espacialização sonora, define seis possibilidades de espacialização sonora: (1) Gestual-coreográfico – quando a criação de trajetórias introduzem uma coreografia sonora à peça, sendo que a mesma apresente algum significado; (2) Paramétrico-serial - utilização da localização da fonte sonora como mais um parâmetro musical a ser serializado; (3) Criação de campos sonoros - difusão ou distribuição da energia sonora, com objetivo de criação de amplos e envolventes campos sonoros; (4) Ambiência - simulação acústica realizada através da aplicação de reverberação e ecos realistas; (5) Acústica-expandida - expansão do espaço acústico, afinando o mesmo a partir da performance musical ao excitar ressonâncias específicas já existentes (ex.: *Répons* de Pierre Boulez); (6) Topológico-referencial - alusão a outros espaços ao se utilizar sons remissivos de espaços e ambientes específicos¹⁸¹.

Dessas seis abordagens descritas pela autora, deixamos de lado a abordagem 5, uma vez que a mesma explora o *espaço externo* a obra¹⁸², o qual já vimos ser extremamente limitado na nossa mídia audiovisual adaptativa para realidade virtual, nos direcionando então para a manipulação do *espaço interno*. Entre essas que envolvem o *espaço interno* da obra, as estratégias 1, 2 e 3 se direcionam para o *espaço extrínseco*, enquanto a 4 e 6 ao *espaço intrínseco*. Ainda poderíamos classificar tais estratégias entre aquelas que exploram a manipulação do *espaço extrínseco* em sua materialidade (1, 2 e 3) e aquelas que preveem uma abordagem mais direcionada ao semântico ou referencial (1, 4 e 6), ou seja, ao que o *espaço extrínseco* apresentado pode expressar, simular ou aludir. A estratégia 1, enumerada pela autora faz parte dessas duas abordagens, uma vez que, em sua própria definição, ela propõe que o movimento por si mesmo apresente um significado.

Outro autor que tenta responder a essa nossa pergunta é Begault (2000, p. 1), que se limita a quatro principais abordagens da espacialização sonora: (1) Replicação - na qual se objetiva a reprodução ao mais fiel possível de um evento real, incluindo

¹⁸¹ Esta última abordagem é descrita também em BARRETT, 2002.

¹⁸² Uma vez que a produção de audiovisual para realidade virtual exclui quase completamente a influência do *espaço externo*, tal estratégia não faz sentido ser desenvolvida aqui.

movimentação de fontes sonoras e simulação do ambiente na qual ela se situa; (2) Criação – que consiste na possibilidade de obtenção de espaços sonoros desconhecidos, não necessariamente realistas, assim como movimentação de fontes sonoras de formas inesperadas; (3) Transmutação – transição entre espaços sonoros devagar ou rapidamente, realistas ou não; (4) Representação – caso em que se propõe a representação de um percurso pelo qual um indivíduo passaria, com transições entre eles e como se o ouvinte estivesse ouvindo através dos ouvidos do performer por exemplo.

Podemos observar que todas essas proposições de Begault se direcionam para o âmbito puramente semântico, uma vez que a identificação de cada um desses processos propostos depende sempre de uma interpretação do material ouvido, e não de uma análise meramente estrutural do material por si mesmo, do ponto de vista sintático, como em um tratamento serialista por exemplo.

Natasha Barrett (2002, p. 135-136), por sua vez, desenvolve em seu trabalho quatro principais formas de abordar o espaço na música eletroacústica: (1) Criação de ilusão de um espaço ou da localização espacial de uma fonte sonora - nessa categoria a autora inclui a manipulação do tamanho das fontes sonoras e sua movimentação, assim como o observável na própria natureza do material sonoro; (2) Alusão a um espaço ou localização de uma fonte sonora - segundo a autora a alusão espacial é aberta a reinterpretações, mais do que a ilusão espacial; (3) Simulação de um campo sonoro tridimensional; (4) Possibilidades espaciais contingentes ao desenvolvimento temporal.

No decorrer da descrição das abordagens (1) e (2), a autora menciona a possibilidade de transformação de uma ilusão em uma alusão e afirma que a mesma possui informações que podem não corresponder às leis do som num espaço real, podendo ser insuficientes ou conflitantes. Segundo ela “a totalidade de conexões espaciais inclui cheiro, toque, visão e som, mesmo que o caminho primário para o espaço seja somente aural”¹⁸³ (BARRETT, 2002, p. 135, 136). Nessa afirmação observamos claramente um direcionamento para uma abordagem semântica do

¹⁸³ Original: "(...) the totality of the spatial connection includes smell, touch, sight and sound, even though the primary path to that space may have been solely aural."

espaço extrínseco, relacionado com uma série de associações provenientes da percepção sonora, quase que pressupondo uma sinestesia¹⁸⁴.

No trabalho de Barrett, podemos observar que suas quatro abordagens para o espaço se misturam entre os níveis sintático e semântico. A primeira delas, a ilusão de um espaço, se dá puramente em um nível sintático uma vez obedecidas e recriadas as “leis” reais de espacialização, que incluem propriedades de reverberação, efeito *Doppler* quando a fonte se encontra em movimento, efeitos próprios da transmissão sonora tais como perda de energia em frequências agudas proporcional à distância, entre outras. A segunda abordagem, a de uma alusão a um espaço, transita entre o sintático e o semântico: uma vez obedecidas as leis da física, aproxima-se de uma ilusão (nível sintático) mas, ao apresentar simultaneamente elementos que a afastam das leis da física e são abertas à interpretação, se aproxima do nível semântico. Podemos exemplificar esse trânsito quando manipulamos uma reverberação para aumentar o tamanho da sala em que a fonte sonora está. Sabemos ser impossível do ponto de vista real, logo, associamos tal processo como uma manipulação do espaço extrínseco intencionada pelo compositor. Percebemos tal manipulação inicialmente como uma estrutura sintática, mas buscamos um significado dentro do contexto em que ela ocorre.

Sua terceira abordagem, a de uma simulação de um ambiente tridimensional depende completamente de uma boa representação das leis da física e se coloca completamente inserida em um nível sintático. Já a quarta abordagem, a das possibilidades dependentes de uma evolução temporal, podem ocorrer tanto no nível sintático, quanto no semântico. Sua proposta de exploração da memória do ouvinte pelo compositor por exemplo, relacionada com essa quarta abordagem, propõe trazer à tona espaços apresentados anteriormente e assimilados no nível semântico, enquanto que as transformações de *espaço extrínseco* em *intrínseco* (ver 3.1.1) por exemplo ocorrem em um nível sintático.

No trabalho de Smalley (1997), o autor aponta a sintaxe do material sonoro, como sendo aquela denominada “real”, e a semântica do mesmo, “analógica”, se refere ao nível das analogias, das metáforas que a manipulação do *espaço extrínseco*

¹⁸⁴ Relação espontânea individual entre percepções de diferentes sentidos (ex.: um som remeter a uma cor, ou um cheiro remeter a uma imagem).

pode provocar. Essa combinação entre estes dois aspectos da percepção espacial são aplicados no processo composicional de forma simultânea e nunca excludentes.

Essa mistura de abordagens do espaço nos níveis semântico e sintático, observada nos trabalhos descritos aqui, é uma das principais razões que nos levam a descrever as duas possibilidades de abordagem do *espaço extrínseco*, assim como apontar autores que o exploram de maneira semelhante, e possíveis relações desse tipo de separação analítica em outras áreas para guiar a nossa própria. A separação dessas abordagens nesse momento se dá somente com o intuito de facilitar o entendimento dos processos apresentados pelos autores que guiaram a pesquisa.

Temos que considerar, nesse contexto, que toda a linguagem do material audiovisual para realidade virtual, que se articule em termos paramétrico-morfológicos, pode ser discutida através de metáforas relacionadas ao espaço, no entanto sem realmente expressar um espaço tridimensional, no domínio do *espaço extrínseco*. Uma abordagem do *espaço extrínseco* em si depende de uma elaboração material, sintática, concebida especificamente (semantizada) como informação espacial.

Na busca pela função semântica do *espaço extrínseco*, diferenciando do que seria sintático, estrutural, pode ser interessante adotarmos a dicotomia entre gesto e textura, que a música eletroacústica e a música instrumental contemporânea, em uma fuga da abordagem das tradições tonais e seriais, passou a explorar, e aplicá-la no nosso objeto de estudos. Sem nos aprofundarmos muito nessa questão, uma vez que observamos que cada compositor e teórico elabora seu próprio discurso com relação ao que seria o gesto (CASTELLANI, 2010), assumimos o conceito estabelecido por Smalley:

Um gesto é, portanto, uma trajetória de movimento de energia que excita o corpo sonoro, criando vida espectromorfológica. Do ponto de vista tanto do agente quanto do ouvinte passivo, o processo do gesto musical é tátil e visual, bem como aural.¹⁸⁵ (SMALLEY, 1997, p. 111)

Segundo Bates (2009, p. 155), uma composição pode explorar tanto estruturas baseadas em gestos, o que implicaria em um certo nível de movimentação para a

¹⁸⁵ Original: "A gesture is therefore an energy-motion trajectory, which excites the sounding body, creating spectromorphological life. From the viewpoint of both agent and watching listener, the musical gesture-process is tactile and visual as well as aural."

frente, resultado de ímpetus externos, quanto estruturas baseadas em texturas, que focam na atividade interna de um objeto sonoro que aparenta atuar sem qualquer estímulo externo aparente. Esses mesmos princípios de gesto e textura são sugeridos por Denis Smalley como princípios formais de uma composição.

Expandindo esse conceito, inicialmente elaborado dentro do que denominamos de *espaço intrínseco*, para o *espaço extrínseco*, podemos realizar gestos (movimentos bem definidos) através do sistema de reprodução, em diferentes velocidades, aumentando ou diminuindo a percepção de tamanho da fonte sonora, etc. A exploração desses gestos espaciais ou a criação de texturas (movimentação mais indefinida e sem direção) no *espaço extrínseco* é que nos guiarão para a elaboração do que denominaremos de Unidades Semióticas Espaciais (USEs).

2.2.2 A sintaxe do espaço extrínseco

A sintaxe, ou gramática, diz respeito à relação dos signos entre si, o estudo da manipulação dos mesmos de uma forma mais estrutural, formal, no nível da matéria. No nosso contexto, observamos a relação dos objetos sonoros com os objetos visuais, explorada pela maior parte dos trabalhos (CHION, 1994; COOK, 1998) no domínio do *espaço intrínseco*. Aqui apresentamos sua expansão, já proposta esparsamente por diversos autores, para a relação dos gestos espaciais entre si, no domínio do *espaço extrínseco*.

Observamos uma sintaxe do *espaço extrínseco* principalmente quando os autores e compositores se referem a movimentação de uma fonte sonora, a um ambiente sonoro mais envolvente ou uma fonte sonora pontual (VANDE GORNE, 2002, 2017; BARRETT, 2002, 2005). Tal proposta não surgiu imediatamente em todos os níveis que conhecemos hoje (ângulo, altura, distância, tamanho da fonte sonora, etc.) o que podemos exemplificar com o fato de não se poder manipular a dispersão sonora de um instrumento acústico, por exemplo, sendo estática sua percepção de tamanho (ou dependente da região que o mesmo ocupa no espectro de frequências).

Segundo Begault (2000, p. 192), a manipulação do *espaço extrínseco* no contexto musical foi historicamente tão inevitável quanto o desenvolvimento da manipulação da altura, do timbre e do ritmo. Ao longo do seu variado desenvolvimento histórico, três técnicas de organização do *espaço extrínseco* na música ocidental se

sucederam: a alternância entre localizações, a sugestão de ambientes e fontes sonoras distantes, e o movimento de fontes sonoras em si.

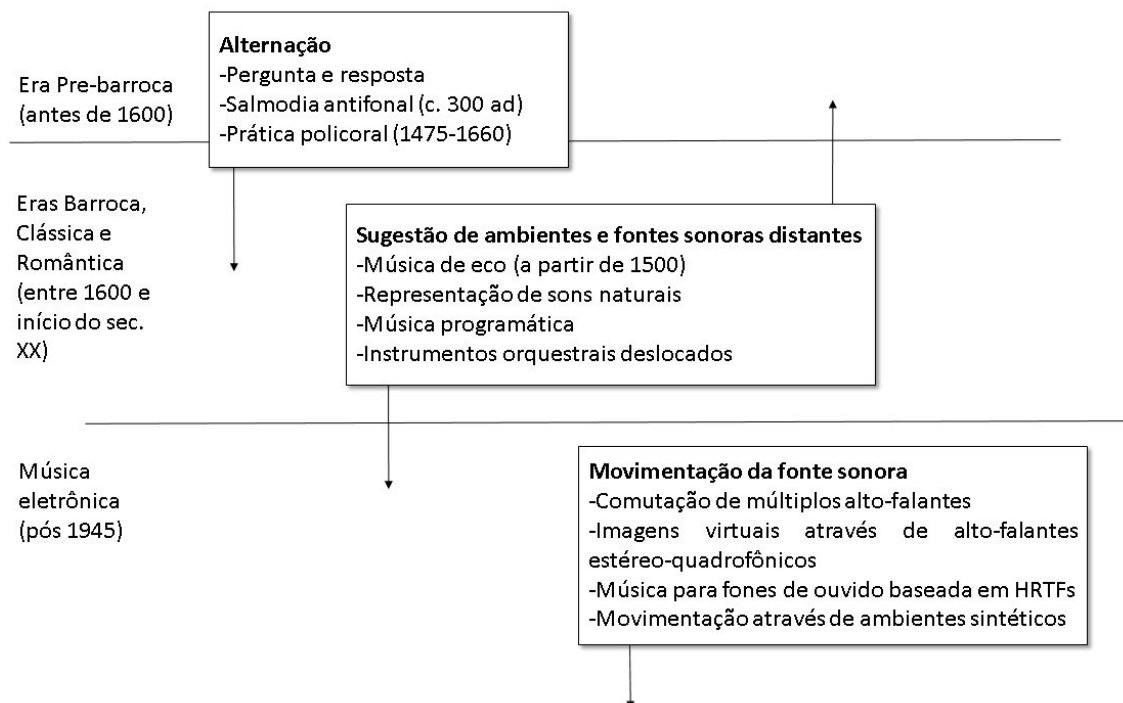


Figura 35 - Linha do tempo descrevendo a prática composicional de manipulação do espaço na tradição musical ocidental¹⁸⁶.

Atualmente na música eletroacústica, com o uso de alto-falantes, temos todas essas estratégias ao nosso dispor. Podemos explorar uma alternância até mesmo em um sistema estereofônico; podemos sugerir ambientes e fontes sonoras distantes através da manipulação da intensidade do sinal do objeto sonoro, da relação entre o sinal direto e o reverberado (utilizando inclusive reverberações com extrema fidelidade ao real), ou da utilização de efeito Doppler; podemos movimentar fontes sonoras não somente no plano horizontal mas também acima ou abaixo do ouvinte dependendo do sistema de reprodução multicanal que tivermos a nossa disposição.

No que diz respeito a percepção dessa manipulação do *espaço extrínseco*, podemos partir da psicologia da *Gestalt*, a exemplo de Bregman (1990), para desenvolver estratégias de espacialização. Enumeramos aqui quatro princípios da percepção sonora humana que devem ser considerados nesse processo, que se aplicam tanto a música instrumental quanto eletroacústica ou acusmática: (1) o

¹⁸⁶ Fonte: adaptado de BEGAULT, 2000, p. 192

princípio da proximidade: elementos posicionados mais próximos espacial ou temporalmente, foram provavelmente gerados pelo mesmo evento; (2) o princípio da similaridade: sons com timbre ou frequências similares provavelmente pertencem a um mesmo evento; (3) o princípio da boa continuidade e conclusão: sons gerados pelo mesmo evento tendem a ser contínuos e seguir um ao outro; e (4) o princípio do destino comum: sons com trajetória de frequências, dinâmica ou rítmica semelhantes, provavelmente se originaram do mesmo evento.

Tomando inicialmente um exemplo da música instrumental, uma vez que, no desenvolvimento estético musical, as relações começaram a se tornar gradativamente mais complexas, a separação espacial das fontes sonoras, com o intuito de manter o entendimento da obra pode ser considerada uma ferramenta de extrema importância. Henry Brant (1967), talvez tenha sido um dos primeiros compositores a tentar uma sistematização da utilização do *espaço extrínseco* em obras instrumentais, afirmando, entre outras coisas, que a percepção de diferentes camadas de material musical podem ser intensificadas através da separação dos músicos. Pode ter sido inconsciente mas, de forma prática ele já estava explorando e manipulando o *espaço extrínseco* baseado nos quatro princípios de Bregman, enumerados acima, em suas composições.

Essa exploração do espaço físico da apresentação musical com o intuito de separação de fontes sonoras que possuem espectromorfologias muito semelhantes, permite ao compositor obter polifonias e texturas antes difíceis de se conseguir. Tal distribuição das fontes sonoras acústicas, instrumentos, no espaço físico, muitas vezes definida pelo compositor em uma bula, ou definida por uma tradição (como na orquestra), faz parte de uma manipulação de como o *espaço extrínseco* da obra irá se apresentar ao espectador.

Natasha Barrett (2002, p. 318), já na música eletroacústica, também observa a importância da espacialização no que diz respeito à possibilidade de aumentar o número de fontes sonoras identificáveis simultaneamente, ou seja, como Brant, do uso da separação espacial com o intuito de evitar mascaramentos espectrais e criar uma textura temporal que permita contrapontos mais ricos. Da mesma forma que podemos separar duas fontes sonoras, podemos também juntá-las explorando o mesmo princípio, como observado por Smalley (1997, p. 122), ao afirmar que o compositor de música acusmática tem em suas mãos a capacidade de unir duas

fontes sonoras distintas, ou duas espectromorfologias, colocando-as em um mesmo ponto no espaço (*espaço extrínseco*).

Bates (2009, p. 119) menciona também a obra de Charles Ives, *The unanswered question*, como exemplo, onde a manipulação do *espaço extrínseco* é utilizado para clarificar e definir as diversas camadas musicais sobrepostas. Segundo ele, em algumas obras a separação espacial é utilizada para criar a sensação de distância e perspectiva, em outras ainda, para facilitar a performance de trechos dissonantes ou com material musical pouco similar, com o intuito de beneficiar tanto músicos quanto ouvintes.

Um outro exemplo importante, com um uso extremamente interessante do *espaço extrínseco* para separação fontes sonoras pode ser observado na obra *Gruppen* (1955-57), de Stockhausen, onde a separação espacial das três orquestras tem o propósito de clarificar as relações entre as três densas camadas de texturas (BATES, 2009, p. 133). Segundo Harley (1994, p.153) a espacialização em *Gruppen* tem como objetivo principal articular melhor a música uma vez que o pontilhismo criado pelo sistema serial adotado pelo compositor tende a gerar uma certa estaticidade. O próprio Stockhausen (1959/1961 p. 70) descreve ser “possível articular estruturas pontilistas longas fazendo-as perambular no espaço, ao movimenta-las de um lugar a outro”¹⁸⁷. Segundo ele esse processo facilitaria a escuta de duas ou mais camadas que se constituam de um mesmo padrão sonoro.

Harley (1994, p. 156) menciona ainda que Stockhausen não reconhece a distância como parâmetro espacial passível de espacialização, somente a localização. Isso porque, pelo fato de a distância estar estreitamente relacionada com a intensidade (sons mais afastados soam menos intensos e vice-versa), ela não seria serializada com efetividade, se confundindo com o parâmetro intensidade ou dinâmica¹⁸⁸. Dentre as abordagens de movimento exploradas por Stockhausen na obra *Gruppen*, Harley aponta um processo de rotação no naipe dos metais, por exemplo, no grupo 119 da partitura, onde acordes que passam lentamente de uma orquestra para outra criam uma percepção de movimento.

¹⁸⁷ Original: “(...) it is possible to articulate longer pontilistic structures by having them wander in space, by moving them from one place to another (...)”.

¹⁸⁸ Essa percepção desse parâmetro o coloca na verdade no limite entre o que definimos como *espaço extrínseco* e *intrínseco*, sendo facilmente objeto de transformações como as descritas em 3.1.1. A manipulação da amplitude de um sinal pode exprimir intensidade, dinâmica, dentro de um contexto do *espaço intrínseco*, ou distância, dentro do contexto do *espaço extrínseco*.

Apesar de suas aparentes vantagens, a espacialização das diversas fontes sonoras com o intuito de clarificar cada material ou camada musical pode trazer consigo novos desafios do ponto de vista da performance, principalmente quando a levamos aos extremos como foi feito na obra *Gruppen*. Entre elas:

- A coordenação rítmica é difícil de se alcançar quando os músicos se encontram espacialmente separados por grandes distâncias¹⁸⁹, superiores a 10 metros, por exemplo.

- A proximidade ou afastamento espacial pode ser considerada equivalente à de alturas, permitindo uma maior complexidade de texturas. Se considerarmos um material sonoro com conteúdo espectral em uma faixa muito estreita de frequências, ele pode se misturar, como se fosse produzido por um mesmo instrumento (princípio da proximidade de Bregman), se reproduzido por instrumentos próximos. No entanto tal material poderia ser explorado sem problemas se distribuído entre instrumentos relativamente distantes entre si.

- Cada situação e posição de escuta é diferente e não existe uma única ótima posição de escuta, de forma que cada ouvinte tem uma percepção diferente do todo da obra (BRANT, 1967, p. 224).

Essas mesmas questões observadas por esses compositores quanto à espacialização de fontes sonoras instrumentais podem ser percebidas também quando tratamos de obras eletroacústicas ou acusmáticas. No entanto, para essas últimas, a espacialização em sistemas de alto-falantes multicanais maiores, com mais canais, se mostra como uma solução prática.

Brant (1967) descreve também a exploração de padrões geométricos para mapear trajetórias espaciais e essa estratégia foi adotada por diversos outros compositores que utilizaram de abstrações geométricas na tentativa de organizar sistematicamente as relações espaciais ao longo de suas obras (BATES, 2009, p. 125). Vande Gorne (2002) descreve essa mesma estratégia ao definir o que ela chama de *espaço geométrico* como sendo aquele em que

(...) o movimento é parte da forma quando se torna figura, repetição, transição de disparo, de ruptura, etc. Aqui, a geometria do espaço não é um suporte, é

¹⁸⁹ Uma solução que encontramos para essa questão é a transmissão ao vivo da regência, de um dos regentes para os outros, estratégia que pode ser observada no vídeo de “Répons” de Pierre Boulez (<https://www.youtube.com/watch?v=OQE5TYnD58k>).

um objeto musical real e abstrato que conduz a escuta e estrutura a percepção ao longo do tempo.¹⁹⁰ (VANDE GORNE, 2002, p. 5)

Essa abordagem de gerar formas geométricas é um bom exemplo de como a exploração do *espaço extrínseco* pode se encontrar no limiar do que definimos aqui de sintaxe e semântica. Apesar de a definição dos pontos no espaço em que os objetos sonoros irão aparecer ou percorrer ser estruturada de uma forma sintática, sua interpretação pode gerar um significado que está no nível semântico, criando uma coreografia, ou cenografia, sonoro-espacial.

Apesar de a primeira obra eletroacústica a explorar de uma forma sistematizada um espaço com mais de dois canais de áudio ter sido *Gesang der Jünglinge* de Stockhausen (1955/1956) - difundida originalmente entre cinco grupos de alto-falantes - a primeira obra a realmente serializar a projeção sonora no espaço foi em *Kontakte*, onde seis movimentos espaciais distintos foram utilizados: rotação, loop, alternância, distribuição estática com fontes duplicadas, distribuição estática com fontes diferentes, e uma fonte única estática. Tais movimentos são tratados de forma serial do ponto de vista composicional, assim como o ritmo e as alturas (BATES, 2009, p. 138).

Essa abordagem serialista do espaço foi muito questionada por outros autores e compositores uma vez que existe uma distância importante a ser observada entre o que foi planejado pelo compositor e o que o público realmente constata audivelmente. As próprias limitações, e resolução espacial, da escuta humana, levam a crer que algumas nuances seriam impossíveis de serem percebidas pelo público. A partir dessa observação, Stockhausen passou a incorporar em suas composições, ao invés de posições absolutas das fontes sonoras, outros parâmetros espaciais, tais como a velocidade de um movimento, a sua direção e o seu movimento angular. Parâmetros relacionados a distância contudo passaram a ser usados de uma forma muito mais intuitiva e dramática (BATES, 2009, p. 145, 146).

A localização angular precisa não se destina a ser percebida com precisão por cada membro da audiência; em vez disso, uma mudança na localização espacial é usada para indicar uma determinada duração temporal. Assim, mesmo que cada ouvinte perceba um movimento espacial ligeiramente

¹⁹⁰ Original : “Le mouvement fait partie de la forme lorsqu'il devient figure, répétition, transition, rupture, déclenchement, etc. Ici, l'espace géométrie n'est donc pas un support, c'est un objet musical réel et abstrait qui conduit l'écoute et structure la perception par son évolution dans le temps”.

diferente, todos perceberão o mesmo movimento espaço-temporal.¹⁹¹
(BATES, 2009, p. 213)

Xenakis também rejeitou a abordagem serialista do espaço, sua abordagem para tal parâmetro musical, influenciada pela sua formação em arquitetura e vasto conhecimento na área de exatas, “usava de técnicas matemáticas formalizadas para desenhar massas sonoras geométricas em grande escala”¹⁹² (BATES, 2009, p. 141). Ele também sugeriu que formas geométricas poderiam ser criadas através da projeção sonora em trajetórias definidas pelos alto-falantes, propondo estratégias para a produção de um círculo, de uma superfície, entre outros. Em *Terretektorh* ele propõe que diversos padrões geométricos e movimentos poderiam ser implementados considerando cinco principais formas de distribuição espacial: pontos distribuídos estocasticamente, planos sonoros com movimento interno, sons estáticos, linhas individuais densamente tecidas e glissandos contínuos.

Como um exemplo dessa abordagem mais próxima da serial, do que Vande Gorne denomina de *espaço geométrico*, revisitamos a obra *The Seven Sins*¹⁹³ do presente autor. Nela, a utilização intencional de séries de três, quatro e cinco pontos no espaço pretendem trazer significados provenientes da retórica da cultura ocidental. A obra, com seu formato algorítmico que inclui um fator aleatório a cada performance, é dividida em sete seções que se desenvolvem com a exploração dos mesmos 14 *samples* gravados a partir de um grupo de gamelões¹⁹⁴. Essas amostras são divididas em dois grupos com características semelhantes, mas com diferenças intrínsecas, e escolhidas de tal maneira que sua organização pode ser considerada serial no sentido de que uma amostra nunca toca novamente antes que todas as sete amostras tenham sido tocadas uma vez. De maneira semelhante, são definidos os conjuntos de três intervalos temporais entre as amostras. Os sete pontos no espaço foram escolhidos de tal forma que se pode desenhar três triângulos diferentes, um quadrado, um círculo

¹⁹¹ Original: “The precise angular location is not intended to be accurately perceived by each member of the audience, instead, a change in spatial location is used to indicate a certain temporal duration. So even though each listener will perceive a slightly different spatial movement, they will all perceive the same spatio-temporal movement”.

¹⁹² Original: “(...) use of formalized mathematical techniques to design large-scale geometrical sound masses (...)”.

¹⁹³ Composição algorítmica estreada em 2011 na universidade de York.

¹⁹⁴ Original: Gamelan Group. Grupo de instrumentos de percussão originários principalmente da Indonésia, composto principalmente de metalofones, xilofones, flautas de bambu, instrumento de corda friccionada (rebab) e vocais.

e uma estrela de sete pontas, e este conjunto de posições é definido pelas formas com a qual cada seção está relacionada.

Têm-se então três "músicos virtuais" (ou *sample players*) que tocam sozinhos em uma seção, e na sequência se somam uns aos outros, a única exceção é a primeira seção quando um certo caos é intencionalmente reproduzido. Cada "músico" tem seu local no *espaço extrínseco* onde suas notas irão soar e que são escolhidas a partir do algoritmo de acordo com uma forma geométrica relacionada a cada seção.

A primeira seção não tem uma forma particular e se encaixa no conceito de caos com o qual tudo começou (descrito na criação do mundo da Bíblia Cristã e em outros escritos antigos). A segunda e a terceira seções são baseadas em três conjuntos diferentes de posições que criam formas triangulares (forma geométrica ternária com expressivo significado divino em diversas culturas) no espaço de escuta. A quarta e quinta seções são baseadas em três conjuntos de quatro posições que criam formas quadradas (forma representativa das artes e ciências desenvolvidas pela humanidade) e a sexta e sétima, envolvendo todos os sete pontos escolhidos no espaço, criam formas circulares (nos remetendo ao infinito) e, de certa forma, pelo menos perceptualmente, retornando ao ambiente caótico da primeira seção, que introduz a obra.

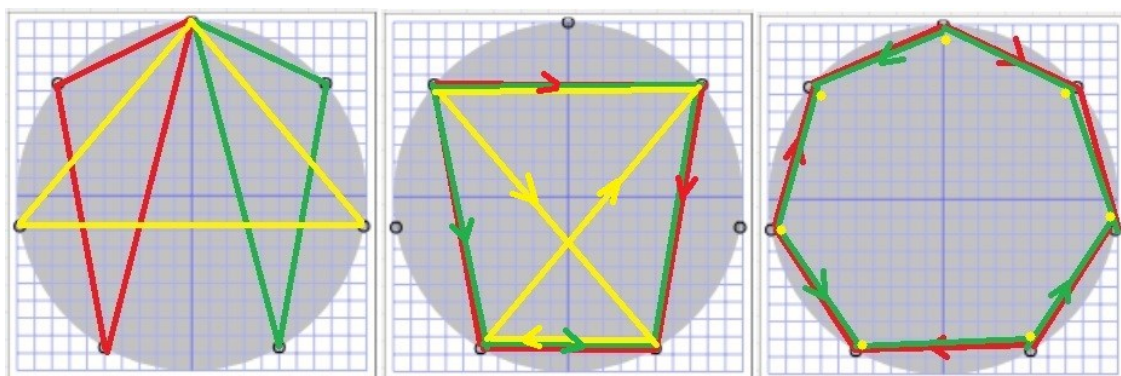


Figura 36 - Os locais definidos para os "músicos" 1 (vermelho), 2 (verde) e 3 (amarelo) que são reproduzidos no espaço, de acordo com as seções estabelecidas pela música - vista superior.

Esse *espaço geométrico*, que considera o mesmo como uma estrutura, podendo se apresentar como linhas, planos, superfícies e volumes¹⁹⁵, é uma das quatro *espécies de espaço*, ou *categorias do espaço*, descritas por Vande Gorne

¹⁹⁵ A autora sugere inclusive que no máximo quatro dessas estruturas sejam exploradas simultaneamente.

(2002), dentre as quais temos outras três: o *espaço ambiofônico*, o *espaço fonte* e a *ilusão de espaço*.

O *espaço ambiofônico*, com o principal objetivo de gerar um certo envolvimento do público, é explorado, por exemplo, na esfera construída por Karlheinz Stockhausen na exposição mundial de Osaka em 1970, que incluía uma parede interior coberta de alto-falantes em 360° tanto em cima quanto em baixo, ou no Pavilhão Philips construído por Le Corbusier para a Feira Mundial de Bruxelas em 1958. Essa categoria de espaço se caracteriza principalmente por não permitir a identificação clara da origem das fontes sonoras (VANDE GORNE, 2002), podendo ser, no entanto, seccionado, passível de gerar expressividade através da divisão ou da (re)composição de um espaço. Tal categoria também será abordada nas obras descritas no capítulo 3, em especial aplicado a algumas texturas dos *cinq études pour une autre réalité*, e nos experimentos com os comerciais do Sebrae, onde essa categoria de espaço será explorada para não distrair a atenção do ouvinte de uma mensagem narrada.

O que a autora descreve como *espaço fonte*, abordagem oposta à anterior, onde localizamos com precisão as fontes sonoras observadas, encontra-se como exemplo mais expressivo nos já mencionados *Acusmoniuns*, sendo Pierre Henry um dos primeiros a explorar tal abordagem em suas obras. Neles se difundem obras originalmente estereofônicas em tantos alto-falantes quanto os que estiverem disponíveis, criando trajetórias, texturas, preenchimento e esvaziamentos de acordo com o observado no material do espaço intrínseco da obra, seus gestos e texturas compostos a partir da manipulação de timbre, altura e durações. Uma outra possibilidade dessa mesma abordagem é descrita pela autora na geração de movimento espacial, no mascaramento, na projeção, na ênfase (foco e espalhamento), e na criação de movimento em texturas granulares através de pequenas nuances.

Uma outra categoria de espaço seria o que a autora descreve como *ilusão de espaço* que, segundo ela, engloba tanto a geração da percepção de profundidade, quanto a percepção do som como imagem, como representação. Essa abordagem se aproxima mais de uma abordagem semântica do *espaço intrínseco*, podendo ser aplicada também quando pensamos em uma certa aproximação da realidade, principalmente no que diz respeito à utilização de uma reverberação convincente em

um material gravado, em uma associação profunda do conhecimento da reverberação de um espaço visualizado e sua correspondente sonora. Exemplificamos tal abordagem na experiência imersiva apresentada no trabalho 4.1, onde um videoclipe gravado ao vivo se beneficia de uma reverberação por convolução em formato-B *Ambisonics* para se mostrar mais realista.

Inserindo a já referida dicotomia observada no tratamento do espaço, apresentada pelas músicas de Schaeffer e de Stockhausen, em um contexto, agora, analítico sintático, levantamos uma pergunta: o movimento espacial, ou sua distribuição, surge do objeto sonoro em si ou é imposto a ele? A organização e distribuição do material sonoro de obras eletroacústicas no *espaço extrínseco*, que seguiu basicamente essas duas vertentes, a de Schaeffer - abordagem estética que pretende utilizar o espaço como uma performance para sublinhar ou exagerar conteúdos pré-existent no próprio espaço intrínseco dos objetos sonoros - e a de Stockhausen - organização da distribuição do material musical no espaço físico, seguindo uma proposta direcionada à estética do serialismo e colocando o espaço extrínseco imposto aos objetos sonoros no mesmo nível de outros parâmetros musicais - se mostra como os primórdios de uma sistematização da sintaxe do *espaço extrínseco*.

Também com o intuito de sistematizar aqui uma terminologia de possibilidades dentro da sintaxe do *espaço extrínseco*, tomamos de início Henriksen (2002, p. 33) que, em seu trabalho, o divide em duas categorias: estacionário e em movimento. Sendo o estacionário subdividido em direcional ou não-direcional, e o em movimento aquele que se aproxima a partir de uma certa distância, o que possui movimento lateral, o que possui movimento frente-trás, o que apresenta elevação, é dispersante ou convergente.

Ainda nessa estruturação sintática do *espaço extrínseco*, Smalley (1997, p. 123, 124) define cinco trajetórias possíveis de um objeto sonoro: aproximação, afastamento, cruzamento, rotação e perambulação¹⁹⁶. Essas abordagens sintáticas do *espaço extrínseco* apontam em direção ao trabalho de Annette Vande Gorne (2002, p. 9-11), que, ao discorrer sobre as técnicas de espacialização de uma obra primariamente estereofônica em uma “orquestra de alto-falantes”, ou *Acousmonium*, passa por diversas estratégias que nos auxiliam como um manual, um compêndio de

¹⁹⁶ Original: approach, departure, crossing, rotation, wandering.

15 gestos de espacialização, ou *figuras do espaço*¹⁹⁷, para a performance em tempo real. Em sua descrição de cada *figura do espaço* encontramos a conceituação de um gesto no espaço, sua forma de realização na mesa de mixagem (relacionada à sintaxe, à morfologia), e sua função musical (relacionada a semântica). Apesar de sua abordagem direcionada para a performance ao vivo podemos facilmente extrapolar essa abordagem de gestos espaciais para a composição fixa em sistemas multicanais já incluindo a espacialização.

Apesar de extremamente importantes para o desenvolvimento do nosso trabalho, observamos nas 15 figuras do espaço de Vande Gorne, algumas redundâncias e possibilidades de aprimoramento. Pelo fato de serem pensados para a performance ao vivo, através da manipulação de *faders* em uma mesa de som, temos que considerar que todos eles são manipulações de amplitude de um sinal de áudio direcionado para um ou mais alto-falantes, fazendo com que o que os diferencie esteja limitado à velocidade e à direção objetivada pelo performer (proveniente da escolha de pares ou grupos de alto-falantes no qual ele realiza o movimento).

Quando a autora desenvolve os gestos “misturar” e “desmanchar” por exemplo, como gestos inversos, propomos que eles sejam agrupados em uma mesma unidade, proveniente de uma mesma sintaxe, baseado em um mesmo processo de manipulação, no entanto propondo semânticas opostas. Os gestos de “tremulação” e “oscilação” também nos aparentam gestos muito semelhantes nos quais somente o parâmetro velocidade os diferencia. De forma semelhante vemos os gestos “balanço” e “onda”. No gesto “espiral” observamos claramente a associação de um gesto de “rotação” com um gesto de trajetória linear, que é considerado por ela como um novo gesto ao invés da junção de dois gestos mais simples.

Observada uma certa sintaxe do *espaço extrínseco*, podendo esta ser aplicada tanto a objetos sonoros tridimensionais quanto objetos visuais tridimensionais, quando pretendemos uma análise sintática da relação entre os *espaços extrínsecos* apresentados pelo áudio e pelo vídeo na nossa mídia emergente, retomamos as três relações básica estabelecida por Nicholas Cook (1998): (1) conformidade - quando os meios operam em concordância; (2) controvérsia - quando entram em contradição; e (3) complementação - quando operam entre as duas primeiras. Tais relações não são completamente estanques podendo se apresentar mais próxima de uma ou de outra,

¹⁹⁷ Ver tradução adaptada no anexo 1.

eventualmente se iniciando em uma delas e se desenvolvendo para outra, mas que nos serve como um princípio de estudo nessa área e nos apontam diversas questões interessantes.

Ao expandir a aplicação dessas três relações que, nos estudos do mencionado autor, se referem, inicialmente, ao *espaço intrínseco* do material audiovisual, para o *espaço extrínseco*, apresentado por esses dois domínios (áudio e vídeo), passamos de uma aplicação da tecnologia simplesmente pelo que ela nos possibilita, para realmente desenvolver gestos espaciais tanto no áudio quanto no vídeo que dialogam e podem gerar significado.

Antes de avançar nesse outro nível de exploração do *espaço extrínseco*, o da abordagem semântica, levantamos algumas considerações técnicas e criativas, principalmente se tratando de material distribuído massivamente, através de alguns exemplos. Em videocliques de música, em que é apresentado um áudio em 5.1, podemos observar que muito frequentemente o ponto de escuta simplesmente não coincide com o ponto de vista, ou seja, enquanto ouvimos uma mixagem do áudio com os instrumentos distribuídos no *espaço extrínseco* reproduzido pelo sistema 5.1 de uma determinada forma (invariavelmente fixa), as imagens que vemos estão sempre mudando de posição, hora vemos o cantor em primeiro plano, hora todos os músicos de costas com o público ao fundo, hora com o ponto de vista do público com a banda toda de frente. Um dos principais motivos para que isso ocorra, do ponto de vista prático, se deve à presença de uma necessidade de dinamicidade do ponto de vista (que podemos observar também em filmes de ação) que, se aplicada ao ponto de escuta, transformaria o áudio de uma forma indesejada.

Essas relações de controvérsia são comuns, e podem ser observadas também ao assistirmos um filme no cinema em um tela a nossa frente, que não apresenta material visual ao nosso lado, acima ou atrás, e ao ouvir sons reproduzidos através de um sistema 5.1 que preenche toda a nossa volta. Nesses casos estamos submetidos a idiossincrasias que, apesar de incoerentes, controversas, aprendemos a apreciar.

Outra questão semelhante pode ser observada nos próprios concertos de música eletroacústica audiovisual. Muito comumente se é exibido uma parte visual em uma tela frontal enquanto o áudio, mais elaborado espacialmente, é reproduzido em um sistema multicanal com 8, 16 ou até 48 alto-falantes distribuídos em torno dos

ouvintes de forma a ocupar um *espaço* que não corresponde ao que se observa na parte visual. Tais expressões artísticas não perdem qualquer mérito com a observação dessas incongruências mas nos levam à busca de um sistema homogêneo em que possamos nos libertar de tais limitações ao invés de incorporá-las à obra.

Bates (2009, p. 159) também observa um fenômeno de incoerência semelhante, no entanto, no que diz respeito à performance de obras mistas que combinam instrumentos acústicos com áudio pré-gravado já espacializado. Em tais performances um dos principais desafios observados diz respeito a supressão da desconexão proveniente do posicionamento estático dos músicos que entra em conflito com a parte acusmática espacializada, realizando movimentos em torno do público. Essa desconexão entre o foco visual frontal direcionado para o performer e a parte eletrônica costuma ser parcialmente solucionada com a amplificação suave da parte instrumental, fazendo com que ele se misture melhor com a eletrônica, ou com a colocação de um dos alto-falantes mais próximo ao performer (SMALLEY, 1996, p. 546). Apesar da autora descrever o sucesso de diversos compositores em tal empreitada¹⁹⁸, quando obras que envolvem conteúdo visual previamente desenvolvidos são apresentadas, a dificuldade reaparece, ainda com maior intensidade.

Observamos essas dificuldades de superar as incongruências até mesmo na própria vídeo-música, se mantendo com uma série de limitações, entre elas, no que diz respeito ao suporte através da qual é veiculada. Apesar de existirem abordagens imersivas mais voltadas para instalações sonoras, que tentam superar tais limitações ao incluir um alto nível de interatividade, contando inclusive com processos de *video-mapping* e sensores dos mais diversos, grande parte da produção ainda se vê refém do que Caesar (2013) denomina de tecnografia do suporte, uma dependência do que se produz ao meio em que será veiculado. Segundo Caesar essa tecnografia do suporte seria “aquilo que se identifica num produto, seja ele estético, seja ele de qualquer outra ordem, aonde se percebe uma presença muito explícita, as vezes implícita também, da máquina que o produziu.” (CAESAR, 2013).

As demandas técnicas, de poder de processamento e hardware dedicado, associados a um grande número de projetores e espaço específico para reprodução

¹⁹⁸ A exemplo da obra *Kontakte* de Stockhausen, que conta como um dos elementos que favorece a unificação da parte eletrônica com a instrumental, um alto-falante no palco, aproximando os dois mundos e os expandindo para as laterais, com dois alto-falantes, e para o fundo com somente um.

(domos e / ou telas muito grandes), por exemplo, fazem com que muitos compositores prefiram explorar somente recursos que se apresentem como mais factíveis. A proposta de um sistema em que tais limitações não existam permitiria que o compositor explorasse uma estruturação, uma sintaxe do *espaço extrínseco*, sem tais preocupações.

2.2.3 A semântica do espaço extrínseco

O domínio da semântica se apresenta na relação dos signos com seus designados, ou seja, do gesto do objeto sonoro no espaço com o que ele emana na percepção do espectador através de uma interpretação. Tal domínio é explorado por diversos autores mais profundamente no que diz respeito ao *espaço intrínseco* e pretendemos aqui especular sobre uma semântica do *espaço extrínseco*, proveniente de uma retórica (reunião de regras relativas a eloquência), de um discurso da exploração dos gestos espaciais, que poderá ser melhor observado também na descrição dos projetos desenvolvidos.

Já no tratado dos objetos sonoros de Schaeffer (1966), o autor identifica quatro tipos de escuta (ouvir, escutar, entender e compreender), onde a compreensão acontece quando tratamos o objeto como um signo, transformando a escuta em uma escuta semântica, se referindo a valores e a construção de sentido. Podemos expandir essa mesma proposta à exploração do *espaço extrínseco* como parte da obra musical, carregado de um significado, explorado de uma forma a acrescentar valor e possuir seu próprio sentido musical.

De uma forma geral, é extremamente difícil discursar sobre uma análise semântica separadamente da sintaxe, uma vez que toda e qualquer opção estética realizada em um nível sintático, relacionada a forma e estrutura de uma obra, passou inevitavelmente por uma observação semântica, por mais subjetiva que seja, do compositor. Da mesma forma observamos quando discursamos sobre o *espaço extrínseco* separadamente do *espaço intrínseco*. Como visto anteriormente eles irão coexistir e interagir entre si dentro da concepção da obra, e terão um impacto de forma conjunta na percepção do ouvinte do gesto musical, espacializado ou não, em ambos os domínios.

Em se tratando do material audiovisual, em abordagens mais simples, ou mais realistas, o compositor cria relações entre o material visual e o sonoro, e busca explorar o *espaço extrínseco*, inicialmente, através de relações de conformidade, de forma que um acompanhe sempre o outro. Vemos possibilidades de manipulação do *espaço extrínseco* além dessa relação básica, com possibilidades de criação de surpresas e decepções. Podemos exemplificar isso quando, por exemplo, apresentamos um evento sonoro, anteriormente apresentado em concordância com um material visual, percorrendo ambos uma determinada trajetória. Quando em uma reapresentação, o elemento sonoro traz ilusoriamente uma imagem ausente, junto com seu movimento, causando uma surpresa quando o ouvinte percebe que um material visual em si, como esperado, não está mais lá. O contrário também pode acontecer. Podemos apresentar um material visual em determinado ponto do *espaço extrínseco* no qual seu correspondente sonoro, previamente apresentado e já associado, não está presente ou foi substituído, causando estranheza. Essas mencionadas surpresas ou estranhezas são possíveis somente se o ouvinte percebeu aquela manipulação do *espaço extrínseco* com um certo valor semântico, relacionado estreitamente ao material do *espaço intrínseco*.

O campo sonoro-visual possibilitado pela tecnologia da realidade virtual expande as possibilidades de surpreender o ouvinte nesse sentido uma vez que ele não consegue ver todo o ambiente, sendo necessário movimentar sua cabeça para o ponto de atenção, se surpreendendo ou se decepcionando ao constatar que a imagem que devia estar acompanhando um evento sonoro não está lá ou foi modificada de alguma forma, e vice-versa.

Pretendemos aqui, então, diferenciar o que outros autores discorrem com relação ao *espaço extrínseco* de uma obra no que diz respeito a sua função estrutural, sintática, da sua função semântica. Começaremos com a já mencionada abordagem da terminologia do *espaço extrínseco* observada nos trabalhos de Natasha Barrett (2002) e Denis Smalley (1997), da qual discordamos e denominaríamos, na verdade, de uma abordagem semântica do *espaço intrínseco*, da espectromorfologia do objeto sonoro.

Barrett (2002, p. 314) define o conceito de *espaço extrínseco* ao relacioná-lo com o espaço ao qual ele alude, ou seja, através da nossa percepção do mesmo e do nosso conhecimento de mundo, identificamos, no material ouvido, relações com

espaços reais conhecidos. Segundo ela, “o extrínseco é a capacidade sonora de implicar, referir-se ou associar-se a algo que não esteja empiricamente presente no espectro. Nossa memória brinca com o extrínseco, trazendo as associações que temos com o som” ¹⁹⁹. Esse conceito de *espaço extrínseco* relacionado ao nível semântico, é adotado também por Smalley (1997, p. 110) e se aproxima também do que Nattiez (1990) denomina de aspecto musical intrínseco e extrínseco explorado em sua obra direcionada a uma análise semiótica da música instrumental.

De forma semelhante ocorre com o seu desenvolvimento no que Barrett denomina de consonância espacial e conflito espacial²⁰⁰, que se dá baseada na expectativa do ouvinte do que seria observado no espaço aludido, na reprodução convincente ou não de espaços reais, etc. ou seja, em um nível semântico, assim como na forma sistematizada com que Smalley define e discute as formas espaciais baseadas no conceito de “ligação-fonte” ²⁰¹. “Smalley usa o termo ligação-fonte para descrever essa tendência natural de relacionar sons a supostas fontes e causas, ou relacionar sons uns aos outros devido a uma origem compartilhada”²⁰² (BATES, 2009, p. 154)²⁰³.

Na já mencionada transformação (em 2.1.1) que Barrett descreve, de um objeto sonoro ressonante em um espaço, ou seja, um objeto sonoro no qual foi aplicado uma reverberação qualquer, passando a apresentar informações de um espaço reconhecido (BARRETT, 2002, p. 320), observamos que tal reconhecimento pressupõe um processo de atribuição de sentido e se dá em um nível semântico, ou seja, sem o nosso conhecimento do mundo, tal objeto seria analisado simplesmente como um objeto estendido, ou prolongado, em uma análise sintática.

Smalley (1997) apoia essa sua abordagem no fato de que, para ele, a imagem ou reconhecimento de um espaço vem à tona mesmo sem estar sendo apresentado por um sistema multicanal, e é essa “espacialização” que se apresenta somente em um nível semântico é que encontramos o que é denominado por Smalley de “ligação-

¹⁹⁹ Original: “The extrinsic is the sound capacity to imply, to refer, or to associate with something other than that empirically present in the spectrum. Our memory plays with the extrinsic, bringing forth associations we have with the sound.”

²⁰⁰ Consonância espacial: quando um som real se apresenta na ilusão de forma como esperado. / Conflito espacial: quando a informação espacial inerente ao objeto está em conflito com sua disposição espacial. (BARRETT, 2002, p. 319)

²⁰¹ Original: “source-bonding”

²⁰² Original: “Smalley uses the term source-bonding to describe this natural tendency to relate sounds to supposed sources and causes, or to relate sounds to each other due to a shared origin.”

²⁰³ Esse conceito pode ser encontrado também em SMALLEY, 1997, p. 110.

fonte”. Esse imaginário proveniente da identificação de fontes sonoras e os espaços a que aludem não é herança exclusiva do receptor, mas sim uma das principais ferramentas do compositor que “cria os jogos de combinatórias e ideias musicais que podem ser significativas, alegóricas, metafóricas e até mesmo figurativas”²⁰⁴ (JUSTEL, 2011, p. 112)

Como uma justificativa para a nossa classificação e renomeação da abordagem desses compositores, recorreremos a Alves (2011) que nos descreve a passagem de um período em que a pintura deixa de ser representacional e passa a explorar uma linguagem própria se tornando abstrata. Tal processo aconteceu com artistas como Kandinsky, Klee e Mondrian, por exemplo.

Desse modo, nas suas pinturas, torna-se perceptível um gradual afastamento de uma representação mimética da realidade, através de um processo de redução e abstração. A pintura deixa de remeter para algo que lhe é exterior e passa a constituir-se como pura presença, como imanência. (ALVES, 2011, p. 32).

A ideia de *espaço extrínseco* apresentada por Barrett e Smalley possuem um forte lastro nessa necessidade de representação, não nos servindo para descrever o proposto pela maioria dos artistas que exploram o espaço real como parte da linguagem musical eletroacústica contemporânea. No entanto tal abordagem não pode ser completamente abandonada pois ainda tem uma função de grande impacto na percepção da noção de espaço das obras.

Retomando o nosso conceito assumido de *espaço extrínseco*, observamos que são realmente poucos os autores que realmente dissertam sobre sua semântica, e para orientar nossa abordagem vamos partir também das artes visuais, onde o *espaço extrínseco* é relativamente bem descrito.

Podemos retomar, como um exemplo adaptável à música, a noção de perspectiva. Essa ferramenta das artes visuais se encontra no limiar do que definimos aqui como nível sintático e semântico do *espaço extrínseco*. Ao mesmo tempo que ela *não* nos oferece em si, uma materialidade do nível de profundidade de uma arte pictórica, ela é reconhecida por qualquer um como contendo tal informação espacial da obra. Um outro recurso das artes visuais que está nesse limiar entre o nível

²⁰⁴ Original : “Ce-la n’implique pas, pour autant, que l’imaginaire soit du patrimoine exclusif du récepteur. Le compositeur, lui aussi travail avec l’imagination, puisqu’il met en jeu des combinatoires et des idées musicales porteuses de sens, allégoriques, métaphoriques et même figuratives.”

semântico e sintático do *espaço extrínseco*, descrito por Arnheim (2005, p. 218) diz respeito à relação entre figura e fundo de uma obra visual. Devido ao nosso conhecimento de mundo podemos concluir, ao observar uma imagens e o tamanho de seus elementos, quais estão na frente e quais estão ao fundo.

Há, então, regras definidas de acordo com as quais os fatores perceptivos determinam a localização em profundidade dos planos frontalmente orientados no espaço pictórico. Os artistas aplicam estas regras intuitiva ou conscientemente para tornar as relações de profundidade visíveis. Ao olhar para fotografias ou pinturas representativas, o observador é auxiliado até certo ponto pelo que conhece sobre o espaço físico com base em sua própria experiência. Ele sabe que uma figura humana grande significa estar mais próxima do que uma casa pequena. O artista, contudo, não pode confiar muito no mero conhecimento. Se ele quiser que uma figura se sobressaia num fundo, deve usar o efeito visual direto dos fatores perceptivos. (ARNHEIM, 2005, p. 226)

Podemos obter tais efeitos, tanto a perspectiva quanto essa percepção de figura e fundo no domínio sonoro, recorrendo à manipulação de intensidade das fontes sonoras quando comparadas entre si ou ao que chamamos de reverberação no áudio que, mesmo monofônico, pode trazer tais informações espaciais, de um espaço pequeno ou grande, com uma reverberação que reconhecemos facilmente provenientes de tais espaços. Vande Gorne (2017) é a autora que mais se aproxima de uma abordagem da *perspectiva* do ponto de vista semântico do *espaço extrínseco* ao dissertar sobre sua criação a partir da apresentação de dois planos sonoros diferentes que serão superpostos para a apreciação do ouvinte. Sua abordagem apresenta paralelos na descrição de um fenômeno semelhante na manipulação sonora em obra acusmáticas, possíveis tanto no *espaço intrínseco*, principalmente através da manipulação de intensidade, quanto no *espaço extrínseco*, no que ela denomina de manipulação do espaço em um nível ornamental. Ao movimentarmos uma fonte sonora, seja ele qual for, por mais estático que seu material interno tenha se apresentado até então, quando se inicia algum movimento, tal objeto sonoro se torna figura principal, foco de atenção do ouvinte.

Em outras áreas o espaço também é explorado de forma artístico-semântica e, entre eles, passamos pelo trabalho de Bachelard (1993) que, apesar de direcionado a conceitos observados na literatura, nos traz observações que podem ser válidas na construção de uma semântica do *espaço extrínseco* no audiovisual. O trabalho de Bachelard é focado no desenvolvimento de uma fenomenologia da imagem poética e,

segundo ele, essa imagem de um espaço existe antes mesmo do pensamento, do raciocínio sobre o percebido, se apresentando quase que como uma *Gestalt*. Nada é capaz de preparar essa imagem poética, “nem a cultura, no modo literário nem a percepção no modo psicológico”, ela “não está submetida a um impulso. Não é o eco do passado. É antes o inverso: pela explosão de uma imagem, o passado longínquo ressoa em ecos e não se vê mais em que profundidade esses ecos vão repercutir e cessar” (BACHELARD, 1993, p. 183).

Esse conceito de Bachelard de uma imagem poética, ou seja, uma imagem imaginada pelo leitor e provocada por um conjunto de palavras, versos ou parágrafos, nos é útil pois, como já visto, diversos compositores de música eletroacústica (Barry Truax, Natasha Barrett, Denis Smalley) se referem ao espaço em um nível semântico a partir de imagens provocadas pelos objetos sonoros em si. Essa imagem provocada nem sempre está associada ao que denominamos de *espaço extrínseco* mas sim às características intrínsecas reconhecidas pelo ouvinte em cada objeto sonoro utilizado na composição.

Truax [TRUAX, 1996] argumenta que é difícil evitar o significado representacional dos sons ambientais, afirmando que ‘o som ambiental adquire seu significado tanto em termos de suas próprias propriedades quanto em termos de sua relação com o contexto’. Apesar dessa dificuldade, os registros ambientais também têm sido utilizados de maneira mais simbólica, na qual os diferentes sons e espaços gravados são usados para contar uma espécie de narrativa.²⁰⁵ (BATES, 2009, p. 157)

Lucena (2007), em seu trabalho sobre a fenomenologia apresentada no livro de Bachelard, afirma que “tudo é valor humano; o espaço não pode ser unicamente exterior pois é vivido, imaginado, recordado interiormente.” (LUCENA, 2007, p. 9). Essa afirmação da autora nos leva de volta ao que já afirmamos anteriormente, uma semântica do *espaço extrínseco* é extremamente difícil de ser definida uma vez que é indissociável dos componentes *intrínsecos*.

Ao comparar signos da miniatura e da imensidão de espaços imaginados, Bachelard aponta para uma certa relatividade que pode ser observada também na

²⁰⁵ Original: “Truax has argued that the avoiding the representational meaning of environmental sounds is difficult, stating that ‘environmental sound acquires its meaning both in terms of its own properties and in terms of its relation to context’. Despite this difficulty, environmental recordings have also been used in a more symbolic fashion in which the different recorded sounds and spaces are used to tell a sort of narrative.”

música, quando espaços pouco maiores, quando apresentados logo após *espaços* menores, podem ser interpretados como maiores do que realmente são. Isso se dá pois nossa percepção não é exata mas explora sempre comparações, tendo também uma memória limitada com relação a estímulos apresentados a curto prazo. A manipulação do tamanho da fonte sonora, ou de grupos de fontes sonoras, se dá de forma semelhante ao descrito por Bachelard no nosso descrito *espaço extrínseco*, no qual o sistema de reprodução é a ferramenta através da qual o compositor o manipula. Podemos tomar como exemplo o som de um molho de chaves que, quando reproduzido por um sistema multicanal de grandes proporções, pode ter seu tamanho alterado dependendo da proposição do compositor.

2.2.4 Unidades semióticas espaciais (USEs)

Com o intuito de elaborar um paralelo e estabelecer a origem dessa proposta de relacionamento entre os *espaços extrínsecos* apresentados pelo áudio e pelo vídeo, ainda embrionária, retomamos um pouco o nível do *espaço intrínseco*, visitando a pesquisa desenvolvida no laboratório de música e informática de Marseille (MIM), sobre as denominadas “unidades semióticas temporais” (USTs)²⁰⁶. Tais unidades, definidas por musicistas e pesquisadores, nos aponta um caminho para uma estruturação sintática já associada à uma função semântica. Em seu trabalho eles definiram 19 USTs como sendo categorias de segmentos musicais que adquirem significado através de sua organização temporal (HAUTBOIS, 2010, 2014).

Segundo Hautbois (2014), é argumentado que a existência de figuras temporais guiam a percepção auditiva. De uma forma semelhante, na literatura da percepção visual, existem figuras geométricas básicas, consideradas unidades primitivas, que permitem a decomposição do espaço, seja através da correspondência a modelos conhecidos, através da decomposição em contornos primitivos, ou descrições estruturais de formas básicas. O desenvolvimento da USTs parte do princípio definido por Schaeffer de objeto sonoro, como sendo um evento sonoro percebido como um todo independente de sua origem ou significado, uma *Gestalt*. A classificação das USTs de acordo com Hautbois se baseia no espectro e evolução temporal, e

²⁰⁶ Ver tradução adaptada no anexo 2

estabelece o objeto sonoro como uma unidade morfológica que se destaca de um fluxo sonoro sob as leis da *Gestalt*.

As USTs tentam dar uma consistência à definição de Schaeffer com um número pequeno de figuras temporais vislumbrando o menor segmento correspondente a uma concepção semantizada, relacionando uma interpretação conceitual a determinada característica estrutural de um gesto. Os pesquisadores do MIM propõem um tipo de objeto semiótico, as USTs, baseados em trabalhos de Sciarrino, Wishart, Smalley, Baley, entre outros.

Segundo os pesquisadores, "as USTs são formas sonoras que transmitem significado através da evolução dinâmica do seu padrão ao longo do tempo"²⁰⁷ (HAUTOBOIS, 2014, p. 100) que, ao contrário dos objetos sonoros de Schaeffer, não são isoladas de seu contexto, mas sim baseadas em uma atitude de escuta que não foca na altura, nem na harmonia ou timbre, mas no efeito geral de todos esses parâmetros ao longo do tempo. As USTs tem sido consideradas uma eficiente ferramenta de análise tanto morfológica quanto no nível semântico, têm sido também exploradas no processo de educação musical e se tornado uma ferramenta composicional (HAUTOBOIS, 2014, p. 103). Quatro características morfológicas foram usadas na classificação: duração, se pode ser delimitado no tempo ou não, o número de fases (etapas) e a matéria sonora; duas características cinéticas: o tipo de aceleração e a progressão temporal; e características semânticas: direção do processo, movimento e energia sonora.

Descritos e explorados os conceitos já estabelecidos tanto no que diz respeito à sintaxe e à semântica do *espaço extrínseco*, encontramos uma interseção entre essas definições de USTs, gestos observáveis no âmbito do *espaço intrínseco*, com as 15 *figuras do espaço* desenvolvidas em um nível sintático do por Vande Gorne, sistematizando 8 construções espaciais que denominamos de *Unidades Semióticas Espaciais* (USEs). Tais gestos se baseiam não somente na prática de espacialização em tempo real mas também na percepção de estruturas mínimas a serem exploradas na etapa de espacialização de uma obra eletroacústica, experimentadas de forma prática tanto em concerto de música acusmática, quanto nas produções para sistemas baseados em *Ambisonics* para realidade virtual.

²⁰⁷ Original: "(...) TSUs are sound forms that convey meaning through the dynamic evolution of their pattern over time".

As USEs pretendem ser unidades estruturantes de gestos no *espaço extrínseco* consideradas significativas e que, quando combinadas, seja através de superposições, justaposições, ou outras relações, resultam em uma narrativa, num discurso próprio da obra, emanado com a participação ativa do *espaço intrínseco*, e explorado pelo compositor através de uma retórica dos seus usos.

As USTs mencionadas anteriormente se dividem em duas grandes categorias, cada uma subdividida em outras três: invariantes (1) por repetição, (2) por estagnação e (3) por efeito caótico; e variantes (4) com desenvolvimento uniforme, (5) com desenvolvimento frustrado e (6) com equilíbrio interrompido. Tomando essa estrutura com o propósito de estruturar também a nossa categorização de possibilidades de manipulação do *espaço extrínseco*, as USEs, nos aproximamos de Henriksen (2002), ao estabelecer aquelas manipulações estacionárias (estático, expansão / compressão, ênfase e granulação) e aquelas em movimento (balanço, direção, atravessar / saltar e rotação).

Observamos que um limiar entre o que define uma ou outra USE pode estar relacionado ao contexto. Se tomarmos como exemplo as USEs “balanço”, “direção” e “atravessar / saltar”, o que irá as diferenciar entre si pode estar na velocidade do evento, na sua repetição ou não. Um gesto frente-trás por exemplo, se único, será observado como “direção”, se rápido demais, pode ser observado como “atravessar / saltar”, se repetido em velocidade moderada, pode ser observado como um “balanço”. Segue-se uma descrição mais detalhada da nossa proposta de USEs, fazendo também um paralelo com os dois principais trabalhos que as orientaram, seguida de uma tabela resumida. Não desconsiderando a complexidade de suas combinações, tentamos ordená-las, em sua unidade, da mais simples à mais complexa:

>USE: Estático

Sintaxe: Fonte sonora fixa num ponto do espaço ou onipresente. A fonte sonora não apresenta qualquer movimento no *espaço extrínseco*.

Semântica: Imobilidade, obsessão, quietude, expectativa.

Observações: Pode ser uma ferramenta para serialização, quando pontual e bem definido, como descrito na obra *The seven sins* do presente autor.

Paralelos: Associamos essa USE à UST invariante por repetição “obsessivo”, assim como às USTs invariantes por estagnação “estacionário”, “flutuando” e “em suspensão” nas quais só existem movimento do material intrínseco do objeto sonoro.

Na obra *Gruppen* de Stockhausen aparece com o nome de “isolamento” (HARLEY, 1994), e na *Kontakte*, com as seguintes variações: “fonte única estática”, “distribuição estática com fontes diferentes” e “distribuição estática com fontes duplicadas” (BATES, 2009, p. 138).

>USE: Expansão / Compressão

Sintaxe: Aumento ou diminuição do número de alto-falantes na reprodução de uma fonte sonora, podendo ser obtido através de outras técnicas dependendo do sistema.

Semântica: Manipulação (estiramento ou diminuição) do tamanho da fonte sonora, grupo ou espaço de reprodução.

Observações: Sucessivas compressões e expansões, dependendo da velocidade, podem ser percebidas como um movimento de direção ou atravessar por exemplo.

Paralelos: Uma “expansão” muito rápida consistiria no que Vande Gorne denomina de “explosão”, assim como o que ela denomina de “misturar”, “desmanchar”, “acumulação” e “invasão”, podem ser observadas como variações dessa manipulação de “expansão” e “compressão”. Associamos essa USE também às quatro USTs variantes com desenvolvimento frustrado, pelas suas características relacionadas ao aumento ou decaimento energético, que associamos ao aumento ou diminuição do tamanho da fonte sonora, grupo ou espaço de reprodução. Na obra *Gruppen*, Stockhausen chamaria essa USE de “fusão”.

>USE: Ênfase

Sintaxe: Aumento da intensidade sonora em determinada região do campo sonoro.

Semântica: Destacamento de estruturas específicas, esforço para iniciar um movimento.

Observações: Depende de um campo sonoro completo já estabelecido

Paralelos: Relação direta com o que Vande Gorne chama de ênfase, sendo esse gesto muito característico para romper com a reprodução uniforme de um campo sonoro, chamando a atenção do ouvinte para determinado detalhe.

>USE: Balanço

Sintaxe: Normalmente realizado entre lados opostos de um sistema de reprodução (L-R ou F-B).

Semântica: Destacamento de estruturas percebidas, esforço para iniciar um movimento.

Observações: Essa USE pode ser explorada numa composição acusmática quando o material em si possui movimentação no domínio do *espaço intrínseco* muito suave ou quando se pretende dar um pouco de movimento a texturas muito estáticas, criando mais interesse na sua percepção. Se caracteriza também pela sua previsibilidade de movimentação seja rítmica ou relacionada ao próprio material sonoro.

Paralelos: Engloba as figuras do espaço que Vande Gorne denomina de “balanço” e “onda”, entendemos que ambos são diferentes níveis de uma mesma operação, a associamos também à UST invariante por repetição “ondulando”.

>USE: Direção

Sintaxe: Trajeto retilíneo contínuo, de um alto-falantes para outro, ou entre conjuntos de alto-falantes.

Semântica: Trajetória, deslocamento de atenção.

Observações: Se diferencia da USE “atravessar / saltar” principalmente no que diz respeito a sua velocidade. Se o movimento direcional for muito rápido, o mesmo pode ser observado também como um salto. Pode acontecer em qualquer eixo do campo sonoro.

Paralelos: Pode ser associada à UST variante com desenvolvimento uniforme “avançando” pela sua característica direcional, ao movimento, nas obras de Stockhausen, de “alternância”, ou nas trajetórias descrita por Smalley de “aproximação” e “afastamento”.

>USE: Atravessar / Saltar

Sintaxe: Gesto brusco de transferência de energia de uma região para outra do sistema, em alguns casos possível somente em software, de preferência com o auxílio de um simulador de efeito Doppler.

Semântica: Mudança de sessão, ruptura, surpresa.

Observações: É constituída de uma mudança repentina, relativamente rápida, dos alto-falantes responsáveis pela reprodução de determinado objeto sonoro. Com o auxílio da simulação do efeito Doppler pode gerar movimentos extremamente convincentes inclusive de estar atravessando a posição do ouvinte.

Paralelos: Engloba o que Vande Gorne denomina de “rebote”, “inserção / ruptura”, “aparecimento / desaparecimento”. Pode ser associada também à trajetória descrita por Smalley como “cruzamento”.

>USE: Rotação

Sintaxe: Trajeto circular, constante, acelerando ou freando, subindo ou descendo, comprimindo ou expandindo, ao redor do espectador.

Semântica: Quando somente em um plano implica em um confinamento, uma preparação, quando associada à USE “direção” assume a função de um direcionamento.

Observações: Esta USE, quando conjugada com a USE “direção”, seja ela realizada em qualquer um dos três eixos espaciais e em qualquer sentido, para cima, para baixo, para dentro ou para fora do sistema de reprodução, possui um paralelo direto com o que Vande Gorne denomina de “espiral”. Na performance de espacialização em tempo real com auxílio de uma mesa de mixagem pode ser obtida também através de sucessivas “ênfases”.

Paralelos: Podemos associá-la à UST invariante por repetição “rodando”, assim como às USTs variantes com desenvolvimento uniforme “trajetória infinita” e “peso”, quando muito insistente e com pouca variação. Tal USE pode ser encontrada na descrição da espacialização das obras *Gruppen* e *Kontakte* (nesta última também com o nome de “loop”), de Stockhausen, assim como na descrição das trajetórias possíveis de um objeto sonoro de Smalley.

>USE: Granulação

Sintaxe: Alternância, em diferentes velocidades, da posição da fonte sonora.

Semântica: Alternância, indecisão, desfazimento do espaço, tensão ou preparação para um gesto.

Observações: A granulação espacial tem sido objeto de estudos de alguns autores (WILSON, 2008) e ainda pouco explorada principalmente devido a demanda de processamento. O processo de dividir a fonte sonora em pequenos trechos e espacializá-los simultaneamente exige uma ferramenta especializada ou um bom trabalho artesanal. Seu embrião está, no entanto, em movimentos curtos e rápidos da posição vetorial da fonte sonora que, por se tornar muito rápido, deixa de ser percebido como uma estrutura de trajetória e se torna o que denominamos de granulação.

Paralelos: engloba as denominadas “tremulação” e “oscilação”, descritas por Vande Gorne e a associamos às USTs invariantes por efeito caótico (“divergente” e “caótico”), podendo ser também associada ao que Smalley denomina de “perambulação”.

Tabela 3: 8 USEs propostas pelo presente autor

USE	Sintaxe	Semântica
Estático	Fonte sonora fixa num ponto do espaço ou onipresente	Imobilidade, obsessão, quietude, expectativa
Expansão / Compressão	Aumento ou diminuição do número de alto-falantes na reprodução de uma fonte sonora, podendo ser obtido através de outras técnicas dependendo do sistema.	Manipulação (estiramento ou diminuição) do tamanho da fonte sonora, grupo ou espaço de reprodução
Ênfase	Aumento da intensidade sonora de determinada região do campo sonoro	Destacamento de estruturas percebidas, esforço para iniciar um movimento
Balanço	Normalmente realizado entre lados opostos de um sistema de reprodução (L-R ou F-B)	Movimento direcionado, diálogo, pergunta e resposta
Direção	Trajetó retilíneo contínuo, de um alto-falantes para outro, ou entre conjuntos de alto-falantes.	Trajetória, deslocamento de atenção
Atravessar / saltar	Gesto brusco de transferência de energia de uma região para outra do sistema, em alguns casos possível somente em software, de preferência com o auxílio de um simulador de efeito Doppler	Mudança de sessão, ruptura, surpresa
Rotação	Trajetó circular, constante, acelerando ou frenando, subindo ou descendo, comprimindo ou expandindo	Quando somente em um plano: confinamento, preparação. Quando associada à “direção”: direcionamento
Granulação	Alternância, em diferentes velocidades, da posição da fonte sonora	Alternância, indecisão, desfazimento do espaço, tensão ou preparação para um gesto

As USTs variantes com o equilíbrio interrompido (“caindo” e “propulsão”), descritas por Hautbois, se convertidas em USEs, se constituiriam de duas ou mais USEs, podendo ser associadas a um conjunto de dois ou mais gestos no *espaço extrínseco*. Da mesma forma podemos deduzir diversas outras combinações dessas USEs que propomos para o desenvolvimento de gestos espaciais mais complexos. Um objeto sonoro inicialmente estático, gerando uma certa expectativa, pode começar a se movimentar lentamente, incitando um certo direcionamento e saltar, ou se dispersar no espaço através de uma granulação. Como outro exemplo podemos tomar uma granulação focalizada em uma região do sistema de reprodução se comportando em conjunto e expandindo seu movimento, se direcionando para outra região,

chegando inclusive a realizar uma rotação constante ou que se redireciona para cima ou para baixo, para dentro ou para fora, em uma espiral, do sistema de reprodução.

Uma vez que, na prática de espacialização ao vivo, o controle que temos é somente o de amplitude do sinal direcionado para cada alto-falantes, muitas vezes esses gestos se confundem, ficando estreito o limiar que os diferencia. Contínuas “expansões” e “compressões”, por exemplo, dependendo da velocidade, podem gerar um efeito de “ênfase” ou “granulação”, quando realizados muito rapidamente. Em softwares, essa “granulação” pode ser levada a extremos bem interessantes chegando a transformar o que seriam gestos bem definidos, em texturas esparsas no campo sonoro.

A criação de gestos e texturas no *espaço extrínseco* que estejam em concordância ou discordância com seus respectivos gestos e texturas sonoras, observados no material *intrínseco*, assim como da parte visual de uma música eletroacústica audiovisual, por exemplo, agora podem ser relacionados. Podemos pensar em um gesto sonoro rápido e expansivo no domínio do *espaço intrínseco* por exemplo, acompanhado de um gesto visual com as mesmas características, assim como um gesto de “expansão” no domínio do *espaço extrínseco*. Tais gestos estariam em concordância e fariam certa lógica para o ouvinte. Já se apresentarmos esses gestos sonoro e espacial concordantes com um gesto visual estático, sem movimentação alguma, tal combinação seria encarada com uma certa dissonância ou contraste, recurso expressivo igualmente importante na criação do discurso da obra. Um outro gesto sonoro pode ser apresentado com um *espaço extrínseco* discordante, por exemplo, e sendo complementado ou continuado por um gesto visual, como um rastro.

Temos portanto, no nosso objeto de estudos, quatro dimensões espaciais que podem ser consideradas mídias básicas dessa modalidade audiovisual emergente: o gesto ou textura sonoro e visual em ambos os *espaços intrínseco* e *extrínseco*. Se adotarmos a simplificação de interação entre mídias abordadas por Cook (1998), essas quatro mídias, das quais duas emergentes (*espaço extrínseco* sonoro e visual), podem estar em concordância, discordância ou se complementando. Dessa forma, a espacialização sonora no áudio para VR, além do papel de realizar uma imersão, como nos jogos digitais, fazendo com que o ouvinte se torne imerso em uma realidade fictícia, possibilita diversas outras combinações obtidas a partir do que propomos aqui.

Entre as diversas produções observadas na internet, que realmente exploram o áudio adaptativo, várias delas favorecem uma recriação realista do mundo, onde o ambiente é gravado e reproduzido da forma mais fiel possível inclusive com o que é visto pelo ouvinte. No entanto, prevemos uma exploração do *espaço extrínseco* que vá além. O espaço observado no vídeo, quando associado ao do áudio, pode ser usado para se desenvolver uma narrativa, explorando gestos retóricos e criando relações mais complexas, dependendo do objetivo do compositor. A possibilidade de se criar essas outras relações, é que traz o potencial de elevação da produção de obras audiovisuais para VR para o nível artístico apontado por Schaeffer.

3. DESENVOLVIMENTO DE PROJETOS

Como uma das propostas principais do presente trabalho, além de nos direcionarmos para conceitos e procedimento provenientes da prática da música eletroacústica audiovisual destinada a reprodução através de VR, alguns experimentos foram desenvolvidos. Tais estudos possuem como principal objetivo o de desenvolver habilidades específicas da técnica e estética em questão e que, além de gerar produtos que pudessem ser reconhecidos pelo mercado, propiciassem reflexões estéticas advindas da prática.

3.1 Grupo 5/4 – gravação realista de um grupo de Jazz

Nessa primeira tentativa de manipulação de *mídia audiovisual adaptativa* para realidade virtual, foi objetivada a geração de um videoclipe, gravado ao vivo no auditório da Escola de Música a UFMG. O grupo se constitui de quatro instrumentistas que se dispuseram uniformemente em torno de uma câmera 360 (ponto de vista do receptor). Alguns microfones próximos²⁰⁸ foram posicionados e esse áudio foi mixado em *Ambisonics* de primeira ordem, único formato de áudio interativo que o YouTube reconhece por enquanto. Um reverberador baseado em convolução particionada, com *IRs*²⁰⁹ gravados em *Ambisonics*, também foi utilizado na mixagem, assim como o plugin codificador *Planewave* do *ATK for Reaper* para o posicionamento das fontes no campo sonoro tridimensional.

Do ponto de vista estético é importante relatar que uma produção desse tipo, com foco em uma reprodução realista se limita a um tipo de replicação (BEGAULT, 2000) de um ambiente tridimensional. Para isso é importante que as posições dos músicos se mantenham, independentemente da movimentação do ouvinte, o que é alcançado com a interação do mesmo com o campo sonoro criado em *Ambisonics* e onde uma reverberação convincente desempenha função essencial. A adoção de reverberação por convolução a partir de quatro *IRs* também no formato-B *Ambisonics*, gravadas originalmente com um microfones tetraédrico, nos possibilita isso de forma satisfatória. Nesse tipo de trabalho ainda não existe qualquer elaboração estética do

²⁰⁸ *Spot mics*

²⁰⁹ Extraído de bibliotecas livres na internet (Ex.: <http://www.openairlib.net/auralizationdb>)

ponto de vista composicional na manipulação e associação entre som e imagem uma vez que o propósito é que o produto final tenha fortes vínculos com o que seria uma experiência real.

O áudio em *Ambisonics* foi todo gerado e finalizado dentro do *Reaper*²¹⁰. Como saída obteve-se um áudio em .wav de quatro canais que foram importados pro *Adobe Premiere*. O processo de criar um projeto no *Adobe Premiere* para montagem do material para realidade virtual não é muito diferente do para outras mídias, exceto pelo fato de que deve-se observar alguns detalhes para garantir que esse material seja identificado como uma mídia interativa para realidade virtual que inclui também o áudio adaptativo.

Após importar o áudio multicanal e o vídeo 360, a criação de uma sequência deve ser monoscópica²¹¹, em 4K e *Ambisonics* (um dos *presets* observados no menu de criação de sequências). Após colocar as mídias na sequência criada e sincronizá-las, a monitoração do áudio pode ser realizada com um *plugin* específico de decodificação do material *Ambisonics* para fones de ouvido nativo do software, denominado *Binauralizer*. Opcionalmente pode-se aplicar um *plugin* transformador do campo sonoro em *Ambisonics* para se regular o eixo do áudio com o eixo do vídeo, identificando e posicionando a frente como referência, por exemplo.

Após a aplicação de textos (letreiros com o nome da música e nome dos integrantes, informações sobre o projeto, entre outros materiais textuais), deve-se, no processo de exportação, observar se o áudio de saída está configurado para AAC, 48KHz, 4.0 ch, 512 de *bitrate* - padrão aceito pelo player do *YouTube*. Deve-se marcar também, na mesma janela as indicações de que o áudio está em formato *Ambisonics* (adaptativo) e o vídeo em *VR*. Para garantir que essas informações estarão no arquivo que será enviado para a plataforma deve-se utilizar um software de inserção de metadados, tal como o *Spatial Media Metadata Injector*²¹², uma ferramenta sugerida pelo próprio website de orientações do *YouTube*²¹³. Vale observar que o *preview* do

²¹⁰ DAW (*digital audio workstation* – área de trabalho de áudio digital) da fabricante Cockos (<https://www.reaper.fm/>)

²¹¹ Ao contrário da visualização estereoscópica, onde temos um material visual direcionado para cada um dos dois olhos (processo realizado pelo *player VR*) para obtermos a percepção da profundidade, a sequência monoscópica é o padrão para a manipulação dentro dos softwares de edição. Nesse formato as arestas do retângulo visualizado serão posteriormente unidas para a formação da esfera em 360° do *VR*.

²¹² <https://github.com/google/spatial-media/tree/master/spatialmedia>

²¹³ <https://support.google.com/youtube/answer/6395969?co=GENIE.Platform%3DDesktop&hl=en>

editor de vídeos do próprio *YouTube* não suporta vídeos em 360° *VR*, logo a visualização deve ser conferida em uma página fora da página de desenvolvedor, ou seja, o usuário deve fazer o *logout* do sistema e conferir o resultado como um usuário externo através de um link público ou privado. O processamento do áudio em *Ambisonics* para *Binaural* também não é imediato, precisando cerca de uma hora após a importação da mídia para que seja observada a interatividade a partir dos dados do giroscópio do *smartphone*.

É importante observar que, no produto audiovisual finalizado, vídeo equiretangular para *VR* com quatro canais de áudio em *Ambisonics* de primeira ordem, é a própria plataforma de streaming (no caso, o YouTube) que faz a conversão e manipulação em tempo real a partir de informações recebidas pelo giroscópio do *smartphone*. Dessa forma observamos que o mesmo produto pode não funcionar em outros players²¹⁴.



Figura 37 - Frame de uma das músicas gravadas com o grupo de Jazz e QR Code para link²¹⁵.

3.2 Comerciais **SEBRAE**

Foi-nos solicitado realizar a espacialização em *Ambisonics* de um material de áudio previamente desenvolvido por uma produtora, para três comerciais do SEBRAE, com o intuito de divulgação da produção de café da região Sul de Minas Gerais. O

²¹⁴ Essa incompatibilidade com outros players foi observada na reprodução do link do vídeo do YouTube através do WhatsApp, por exemplo.

²¹⁵

<https://www.youtube.com/watch?v=ybQ5wRgNeRs&list=PLOSzS2LT3tVaRxmWRbWo3ypE6euw1-eXJ>

projeto se constituía principalmente de material musical, *foley*²¹⁶, uma locução e algumas vozes, e o processo de espacialização foi realizado dentro do *Reaper* com as ferramentas do *ATK for Reaper*.

Nesse processo vale a pena ressaltar que a locução particularmente deveria estar sempre presente, e uma das formas encontradas para se obter tal efeito foi colocá-la posicionada em cima do ouvinte, de forma que mesmo que ele direcionasse sua visão para um dos lados (movimento mais comum na exploração do ambiente virtual), não haveria tanta alteração na percepção timbrística da locução. Apesar de ter sido uma ideia extremamente coerente com a proposta, o contratante não se satisfez e me pediu que a posicionasse na frente do ouvinte.

Tal realização faz com que, ao mover a cabeça, a locução se movimentasse para o lado oposto, com o intuito de manter sua posição no campo sonoro criado. A percepção desse fenômeno traz a atenção para a própria locução, o que nos pareceu inicialmente indesejado, mas mesmo assim, foi a solução mais aceita pelo contratante, o que especulamos se dever ao favorecimento da manutenção do timbre da locução, que provavelmente faz parte da marca comercial do solicitante.

Uma outra solução possível, observada posteriormente à finalização do trabalho, seria colocar o locutor totalmente, e somente, no canal omnidirecional do campo sonoro tridimensional em *Ambisonics* de primeira ordem no formato-B (W). Isso pode ser feito manualmente através do endereçamento dos sinais dentro da própria DAW, ou através de um dos transformadores do campo sonoro gerado pelo *Ambisonics* oferecido pelo *ATK for Reaper*, denominado *Transform DirectO* que oferece essa opção, basicamente anulando as componentes X, Y e Z do sinal que se pretende omnidirecional. Tal recurso foi explorado em trabalhos posteriores principalmente no desenvolvimento de texturas.

Esse tipo de produção, diferentemente da anterior, não possui uma trilha sonora que necessite estar diretamente associada a uma imagem. Optamos por uma distribuição dos instrumentos da trilha sonora em torno do ouvinte, no ambiente tridimensional, o mais homogêneo possível, pretendendo preencher todo o espaço disponível e favorecendo o processo de imersão. Como a trilha já havia sido composta

²¹⁶ O que denominamos de *foley* consiste em material sonoro proveniente de fontes sonoras do ambiente, máquinas, ruídos, passos, etc.

não existia muita possibilidade de ampliar o conteúdo instrumental, com isso trabalhamos somente com o material musical que já estava composto.

O que se propõe então é a criação de um *espaço ambiofônico* (VANDE GORNE, 2002) que envolva o ouvinte sem que, no entanto, tire a atenção da locução, onde se encontra o principal conteúdo a ser passado para o espectador. A exploração de movimentos lentos e uniformes também cumpre esse objetivo de não desviar a atenção do ouvinte da locução. USEs tais como a direção, atravessar / saltar, granulação ou ênfase se mostrariam completamente inapropriadas em tal processo.

Apesar das possibilidades apresentadas de movimentação da trilha musical, o que foi realmente realizado com essa porção do material sonoro foi uma espacialização estática, distribuindo todo o material musical no campo sonoro tridimensional de forma uniforme. Tal abordagem se dá pelo fato de não existir nenhuma necessidade real de realização de uma *Visual Capture* (BEGAULT, 2000) ou *Spatial Magnetization* (CHION, 2009), uma vez que não temos nenhum elemento da trilha musical sendo representada na imagem. Não existe, portanto, qualquer relação de conformidade, controvérsia ou complementação do material sonoro musical com o visual.

Existe, no entanto, uma preocupação dessa natureza com relação ao material sonoro presente no *foley* e nas falas de personagens visíveis no vídeo, que necessitam sim de uma certa conformidade entre som e imagem, com o intuito de traduzir uma certa replicação (BEGAULT, 2000) do ambiente visualmente apresentado. Estes foram então associados espacialmente aos elementos visuais aos quais se referem, buscando uma certa coerência sintática entre o material visual e sonoro desses elementos no *espaço extrínseco*. Esse tratamento foi obtido através da realização de automações nos parâmetros do plugin de espacialização escolhido para condizer com o percebido na parte visual do trabalho.



SCAN ME

Figura 38 - Frame de um dos vídeos produzidos – projeto Semana Internacional do Café 2017 Sebrae e QR Code para link²¹⁷.

3.3 Projeto Corpografias

O projeto *Corpografias*²¹⁸, coordenado pela Profa. Dra. Graziela Andrade do Departamento de Fotografia, Teatro e Cinema da UFMG e pelo Prof. Dr. Jalver Bethônico, se desenvolveu no período de maio a agosto de 2018, com minha participação na composição sonora multicanal. A instalação ficou disponível no espaço pelo período de duas semanas, na galeria da Escola de Belas Artes.

A exposição em si consistia de um vídeo interativo acessado pelo *Occulus* com um smartphone acoplado, através do qual o expectador se via em um ambiente virtual desenvolvido dentro da *Unity*. Sua movimentação dentro do salão (de dimensões aproximadas de 6m por 10m) era enviada para o programa, que realizava movimentação similar dentro do ambiente virtual criado. Nele, o expectador tinha a oportunidade de se direcionar para pontos escolhidos no espaço que o iriam levar a um outro espaço (de um total de oito), onde havia uma performance de uma das bailarinas criadoras. Suas criações foram gravadas no ambiente real em 2D e manipuladas visualmente de diferentes formas e com diferentes estéticas, explorando tanto o vídeo quanto material textual, poemas e as próprias vozes das bailarinas, em

²¹⁷ <https://www.youtube.com/watch?v=7wRK-g4tKFM&list=PLOSzS2LT3tVYK0DDiuDyb1FEOnguouCP->

²¹⁸ Ficha Técnica: Graziela Andrade (Concepção); Ester França, Jalver Bethônico, Marcelo Padovani (Direção); Gilberto Goulart (Vídeo); Marcelo Padovani, Ítalo Travenzoli (Programação, Desenvolvimento e Cenários Digitais); Ana Paula Lima, Anna Vitória, Carla Aquino, Lara Gama, Lara Santos, Luana Vitra, Luísa Machala, Rayane Calixto (Bailarinas Criadoras); Jalver Bethônico (Gravação e Edição de Vozes); Fábio Janhan (Música e Edição em Ambisonics); Aida Bethônico, Anna Vitória, Ester França, Lara Gama, Lara Santos, Luana Vitra, Luísa Machala, Rayane Calixto (Vozes); Fábio Janhan, Jalver Bethônico, Rafael Trindade, Rafael Sodrê (Captação Sonora Espacializada).

adaptações independentes para a tridimensionalidade do espaço virtual criado para cada uma delas.

O áudio multicanal, objetivado para ser uma paisagem sonora com leves distorções e variações, foi criado a partir de material gravado dos próprios ambientes em que foram realizadas as performances, e foi mixado em *Ambisonics* de primeira ordem e difundido através de oito caixas fixadas nas paredes do salão. Uma vez que a movimentação natural do expectador estaria alterando sua percepção sonora do ambiente, a mixagem era decodificada de forma fixa para o arranjo de alto-falantes em formato de um octaedro.

O *espaço ambiofônico* (VANDE GORNE, 2002) criado, se mistura com a ideia de *espaço fonte* (VANDE GORNE, 2002), onde determinados gestos espaciais destoam da textura criada. Não existe qualquer relação entre os *espaços extrínsecos* do material visual com o sonoro. As relações se mantêm puramente no domínio do *espaço intrínseco* pela sua origem comum, as diversas localidades onde foram realizadas as performances. Sem se prender a qualquer obrigatoriedade de relação entre o sonoro e o visual, as mais diversas USEs são exploradas livremente no domínio musical numa abordagem *gestual-coreográfica* (BALLMAN, 2010) própria.

O fato de ter sido mixado em *Ambisonics* foi proposital pois um dos objetivos do projeto era o de se produzir também alguns vídeos (a serem disponibilizados no YouTube), extraídos da experiência interativa em tempo real da proposta do projeto. Tais vídeos foram gravados diretamente da *Unity* e, nesse caso, com a perda da interação do espectador com o ambiente virtual criado no domínio visual, no entanto, com o áudio em *Ambisonics* adaptativo, a imersão se mantém no nível sonoro.

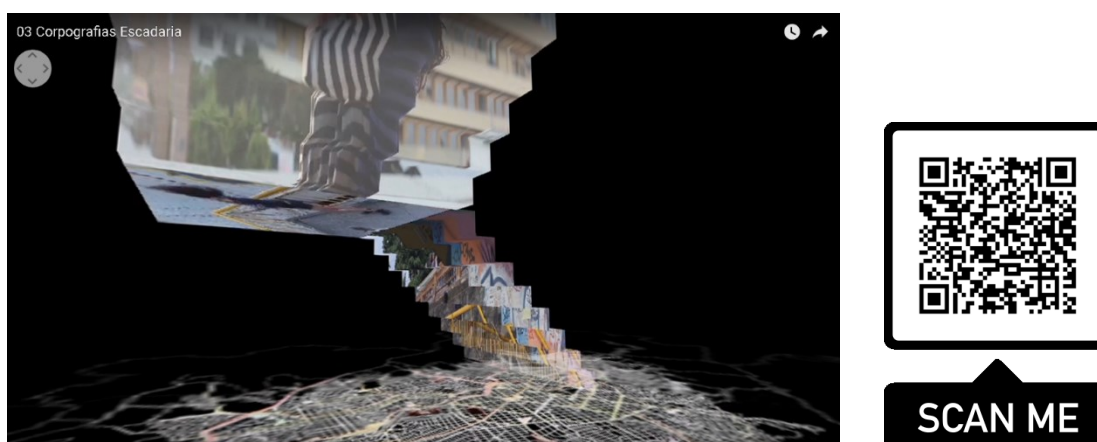


Figura 39 - Frame de um dos vídeos produzidos – projeto Corpografias e QR Code para link²¹⁹.

3.4 Letter Piece (Matthew Shlomowitz) - As Is

Esse trabalho foi um recorte pessoal da obra *Letter Piece* de Matthew Shlomowitz, que se constitui basicamente de um conjunto de cinco gestos, nomeados pelas letras do alfabeto de A a E, e organizados de forma que um performer e um músico interpretem combinações coincidentes, ou não, dessas letras. Os próprios intérpretes criam os gestos, e a obra em si se mostra um exercício de sincronicidade e assincronicidade, um exercício de relacionamento entre os gestos de dois mundos paralelos. O grupo As Is, de performance experimental, criou e realizou uma performance completa, com os cinco gestos sonoros e visuais em 2D, dos quais foram extraídos os gestos sonoros presentes nesse vídeo, que é apresentado com gestos visuais criados em 3D, através do software *Mandelbulb*, pelo presente autor.

Os objetos sonoros gravados com o grupo foram montados no *Reaper* e exportados em *Ambisonics* de primeira ordem para sincronismo com a parte visual no *Adobe Premiere*. Os objetos visuais foram produto de uma primeira experimentação com o software *Mandelbulb*, no qual imagens fractais são criadas a partir de fórmulas matemáticas e exploradas espacialmente através de um navegador. Determinados quadros podem ser selecionados para a criação de uma animação que interpola as diferentes imagens geradas e selecionadas.

O posicionamento no campo sonoro tridimensional específico de cada um dos instrumentos do grupo musical, de forma correspondente ao gesto empregado, passa despercebido pelo fato dos objetos sonoros explorados na obra terem uma característica muito entrecortada. A apreciação do tratamento objetivado no *espaço extrínseco* fica, infelizmente, prejudicada, uma vez que sua percepção se mostra pouco evidente. A seleção dos gestos visuais não possui uma narrativa definida mas simplesmente cumprem a função estrutural exigida pela obra em questão, executando a sequência de letras proposta pelo primeiro sistema, enquanto que os músicos e a realização do *espaço extrínseco* sonoro (este implementado com o pacote de plugins *ATK for Reaper*) realizam simultaneamente o segundo sistema proposto na partitura.

A proposta da presente obra apesar de não deixar evidente a manipulação do *espaço extrínseco* sonoro, se mostra um bom exemplo de uma abordagem *paramétrico-serial* (BALLMAN, 2010), ou de um *espaço geométrico* (VANDE GORNE, 2002), onde a própria localização das fontes sonoras se apresentam com um material

musical a ser serializado. Caímos, no entanto, naquela questão já observada por Bates (2009) de que, apesar de explorada por Schoenberg e rejeitada por Xenakis e Schaeffer, o público nem sempre percebe as relações criadas nesse domínio através dessa abordagem em específico.

[1] *Crotchet = 72 (metronomic)*

Performer: A + A + A + B + B + C +
 Musician: A + A + A + B + B + C +

P M: *x2* A + A + A + B + B + C + *x8* C C C C

[2]

P M: A + A A + + B + B C + D + D D +
 A + A A + + B + B C + D + D D +

P M: A + A A + + B + B C + D + D D + + A + A B
 D + D D + + A A + B + C + C D + + D A + B

P M: C + C C + + D + D A + B + B B + + C + C
 B + B C + + C D + A + A + A B + + B C +

P M: *x4* D D D D *Molto rit* D D D D

[3] *A tempo (subito)*

P M: A A + + A B + + B A + + B B + +
 A A + + A B + + B A + + B B + +

P M: A A + + A B + + B A + + B B + +
 + + A A + + A B + + B A + + B B + +

[4] *x4**

P M: A A A E E D + C C C B B A + E D C B A E
 E E E A B + C C C D D E + A B C D E

*1st time: as per usual
 2nd time: performer only speaks letters / musician as per usual
 3rd time: both only speak letters
 4th time: performer as per usual / musician only speaks letters

[5]

P M: A + A + + A + + + A + + + + A + + + + A +
 A A + + A + + + A + + + + A + + + + A +

P M: + + + + + A + + + + + + + + + A + + + + +
 + + + + + A + + + + + + + + + A + + + + +

Figura 40 - Partitura da obra Letter Piece de Matthew Shlomowitz.



Figura 41 - Frame do vídeo produzido baseado na obra Letter Piece de Matthew Shlomowitz e QR Code para link²²⁰.

3.5 cinq études pour une autre réalité²²¹

Esses cinco estudos têm como inspiração os cinco primeiros estudos de Schaefer²²² no título²²³, como primeira abordagem dentro de uma manifestação artística emergente. Esses cinco estudos, de cerca de 2 minutos cada, exploram material visual sintetizado no software *Mandelbulb*, sendo os estudos “*changer d’arches*”²²⁴ e “*sur basse résolution*”²²⁵ tendo sido montados utilizando o *Adobe Premiere*, e os estudos “*d’origine organique*”²²⁶, “*dans le désert*”²²⁷ e “*d’or nous portons la poussière*”²²⁸ montados no *Sony Vegas*.

O estudo “*sur basse résolution*” foi especializado utilizando o pacote de *plugins Ambix*. Nesse processo foi interessante observar que a decodificação *Binaural*, pelo fato de passar por uma decodificação para um arranjo de alto falantes antes da convolução com as *HRTFs*, apresenta mais opções e sonoridades completamente diferentes umas das outras. Decodificar o campo sonoro 3D criado em *Ambisonics* para um arranjo de alto-falantes octofônico horizontal ou em formato de um cubo, por exemplo, se mostram apresentando resultados binaurais completamente diferentes. Essa observação já havia sido realizada pelo presente autor em trabalho anterior

²²⁰ <https://youtu.be/19I2OVJg8XI>

²²¹ Tradução: “cinco estudos para uma outra realidade”.

²²² “*5 Etudes de Bruits*”, trad.: “cinco estudos de ruído”.

²²³ Exemplo de exploração e da importância da matriz verbal (SANTAELLA, 2001), como um terceiro elemento na *música eletroacústica audiovisual*.

²²⁴ Tradução: “mudando em arcos”.

²²⁵ Tradução: “em baixa resolução”.

²²⁶ Tradução: “de origem orgânica”.

²²⁷ Tradução: “no deserto”.

²²⁸ Tradução: “de ouro carregamos o pó”.

(SOUZA, 2011, p. 155) e retorna na utilização dos decodificadores, levantando questões passíveis de pesquisas futuras relativas ao processo de decodificação das plataformas de streaming, principalmente se passam por uma decodificação para fontes sonoras secundárias (alto falantes virtuais) antes do formato binaural, e qual seria ela.

De uma forma geral existe uma exploração de um *espaço ambiofônico* (VANDE GORNE, 2002), através de uma textura direcionada completamente para o canal W do campo sonoro em *Ambisonics* de primeira ordem, a transformando em um elemento envolvente e imutável. Um *espaço fonte* (VANDE GORNE, 2002) é explorado no gestual e na movimentação de outros objetos sonoros. Em alguns trechos se objetivou a criação de uma certa granulação do material sonoro no *espaço extrínseco* (1'05", 1'35" 1'52", 2'00", 2'13"), em concordância com o material visual.

Devido ao fato de não termos uma ferramenta específica, no pacote de plugins utilizado, para a realização do processo de granulação em específico, exploramos duas estratégias alternativas: a primeira delas, realizada através de automação dos parâmetros *azimuth* e *elevation* do plugin utilizado; a segunda através de uma ferramenta de modulação de parâmetros nativa do Reaper aplicado aos mesmos parâmetros. A primeira estratégia se mostrou um pouco complexa e demorada, mas necessária se pretendemos obter controle suficiente para uma manipulação detalhada. A segunda estratégia, da aplicação de uma modulação dos parâmetros *azimuth* e *elevation* do plugin, possui um resultado semelhante através de um processo relativamente rápido, no entanto com pouco controle de detalhes, por ser completamente autônomo. Com isso, essa segunda estratégia foi explorada principalmente em concordância com alguns objetos audiovisuais que se apresentam de forma muito rápida, como *flashes*, sem necessidade de um grande controle. Por chamar muito a atenção do ouvinte para movimentos bruscos, tal manipulação foi utilizada com parcimônia.

Em 1'35" realizamos uma granulação em um objeto sonoro sendo que seu objeto visual correspondente apresenta características estáticas, tal estranhamento, proveniente de uma relação de discordância, destaca um objeto que tem uma única aparição no estudo. Nos instantes 1'52", 2'00" e 2'13" podemos observar um exemplo de como um objeto sonoro, já reconhecido correspondente a um objeto visual, continua o trajeto do mesmo sem que ele esteja visível, ilustrando uma relação de

complementação entre o visual e o sonoro, nesse domínio do *espaço extrínseco*, exemplo também do princípio de continuidade apresentado por Bregman (BREGMAN, 1990).

Do ponto de vista técnico, vale observar que no momento de renderizar o áudio final, os plugins utilizados já fornecem o áudio na ordem WYZX, característica do formato ambix, sendo diretamente compatível com o adotado pelo YouTube. Ao utilizar outros plugins não necessariamente isso irá acontecer, necessitando de algum tipo de conversão, também oferecida pelo pacote de plugins utilizado (*converter*).

No estudo “*d’origine organique*” espacializamos o áudio utilizando o pacote de plugins *ATK for Reaper*. Nessa peça foi explorado também a utilização de um reverberador com respostas a impulsos retirados de bibliotecas em *Ambisonics*, no formato-B, implementado através de uma mandada auxiliar com dois plugins *SIR*²²⁹ estéreo funcionando simultaneamente. Vale observar que o mais importante em tal processo é mandar duas entradas (W e X) da cadeia de efeitos para um dos plugins e outras duas (Y e Z) para o outro, tendo o processamento dos quatro canais em sua totalidade. Tal procedimento de explorar um reverberador baseado em respostas a impulso somente é possível quando se trabalha com *Ambisonics* de primeira ordem, uma vez que não foram disponibilizadas gravações de *IRs* com microfones de ordens superiores.

A obra apresenta duas texturas principais, uma mais grave e outra com características granulares. A primeira delas foi direcionada diretamente para o canal W do campo sonoro em *Ambisonics*, se apresentando onipresente, criando um *espaço ambiofônico*, enquanto a segunda foi espacializada aleatoriamente, de forma concordante com suas características intrínsecas, através da automação dos parâmetros *azimuth* e *elevação*. Como o pacote de plugins adotado não oferece a possibilidade de exploração da USE atravessar, por não apresentar nenhum tipo de simulador de efeito Doppler, muitos dos gestos visuais que atravessariam o espectador tem seu gesto sonoro correspondente espacializado de forma frontal subindo ou descendo no campo sonoro 3D (ex. 2’42”, 2’47”, 3’03”, 3’30”). Gestos sonoros concordantes com um objeto visual específico tais como em 3’00”, 3’11” e 3’39” foram espacializados para condizer com seu correspondente visual, assumindo uma abordagem de replicação (BEGAULT, 2000).

²²⁹ <https://www.siraudiotools.com/sir1.php>

No momento da renderização do projeto foi necessário a inserção de um conversor de formato, dessa vez *FuMa-ambix*, para que a saída tivesse o formato final aceito pelo YouTube. Tal processo foi conferido auditivamente se utilizando de um plugin de decodificação de *Ambisonics* para *Binaural* do próprio pacote ambix (excluído no momento da renderização).

No estudo “*dans le désert*” optamos por utilizar o pacote de plugins da *o3acore*. Nesse estudo, que explora diversos objetos audiovisuais realizando o movimento de atravessar o expectador, a espacialização sonora do plugin escolhido deixou a desejar por não possibilitar tal movimento, observado até então somente no *FB360*. Apesar disso, outras relações foram exploradas.

Em uma textura granular que percorre a obra foi realizada uma espacialização coerente com o visual através de automação, e à algumas texturas visuais onipresentes (5’13”, 5’22” e 6’18”) foram aplicados algum movimento, em discordância com o visual, com o intuito de chamar a atenção do ouvinte para as mesmas, uma vez que os gestos sonoros que permeiam a peça podem se mostrar monótonos. Gestos sonoros espaciais foram explorados para gerar rupturas entre texturas visuais com características diferentes (5’28”, 6’09”) ou mudanças de direção observada em seu conteúdo intrínseco (5’52”, 5’56”).

A saída dos codificadores são no formato *ambix*, e a ausência de um plugin que faça a codificação diretamente para *Binaural* nesse pacote de plugins nos exigiu que aplicássemos um conversor de formato (*ambix>FuMa*) para posteriormente utilizar um decodificador para *Binaural* de outro pacote de plugins como o do *ATK for Reaper* por exemplo. Nesse processo não se observou nenhuma perda ou alteração no áudio significativamente perceptível, mesmo porque, nessa etapa de composição é difícil elaborar experimentos sistematizados o suficiente para avaliar possíveis influências das conversões de formato ou de diálogo entre os diferentes plugins.

O estudo “*changer d’arches*” teve seu áudio especializado através dos *plugins FB360* que, apesar de oferecerem codificação para *Ambisonics* de até segunda ordem, sua renderização final foi realizada somente em primeira devido ao fato de a plataforma de streaming adotada (YouTube) só possuir suporte para este. Nesse estudo o movimento de atravessar o expectador, apresentando uma concordância entre o material visual e sonoro, possibilitado pelo plugin escolhido foi amplamente explorado ao longo da peça como pode ser observado em 6’53”, 6’57”, 7’22” e 7’34”

por exemplo. Em diversos momentos a partir de 7'31" exploramos o que Chion (1994) denominaria de som diegético *off screen*. Apesar de estarmos trabalhando com os 360° tanto visual quanto sonoro, sons que estão relacionados com a temática da obra mas não são visualizados receberiam essa denominação, e podem ser observados no som de martelos batendo, peças de metal caindo, entre outras, que não se mostram visíveis.

Nessa obra não colocamos as texturas diretamente no canal W do campo sonoro *Ambisonics*, como fizemos em algumas outras para dar a sensação de envolvimento ao ouvinte. Ao realizarmos uma automação de seu movimento, muito sutil na maioria das vezes, explorando principalmente o parâmetro distância, abrimos a possibilidade de chamar a atenção do ouvinte, através do movimento, para algo que supostamente seria uma constante. Efeitos de desmanchar, como observado em 7'11" foram realizados através de automação do parâmetro distância também, outra possibilidade vislumbrada seria a de se realizar uma granulação com mais parâmetros.

O estudo "*d'or nous portons la poussière*" também teve seu áudio especializado usando as ferramentas do *FB360*. Duas texturas principais, que separam a obra em duas seções, e que se pretendiam sem uma direção determinada foram direcionadas exclusivamente para o canal W do decodificador. Os gestos principais de atravessar o expectador foram realizados explorando a simulação do efeito Doppler do plugin. A obra como um todo possui características predominantemente texturais, e em decorrência disso, os poucos gestos são extremamente significativos e chamam a atenção do ouvinte de forma mais intensa.

São apresentados basicamente duas USEs, a de atravessar e a granular, sendo a primeira mais predominante e a segunda explorada nos *flashes* observados em 10'24" e 10'28" por exemplo. Nem sempre os gestos de atravessar fazem a trajetória frente-trás como seria mais evidente, mas eventualmente a trajetória de esquerda-direita, como em 11'00". Algum movimento mais aleatório, que explorasse o efeito Doppler do plugin, foi explorado, mas em controvérsia com o material visual, como em 9'49" e 9'52. Outra controvérsia pode ser observada em 9'59" onde o espaço extrínseco do material sonoro se apresenta de forma estática, apesar da movimentação visual. Como visto anteriormente, a saída desse pacote de plugins já é no formato ambix, não colocando maiores dificuldades na renderização.



Figura 42 - Frames de cada um dos cinco estudos e QR Code para link²³⁰.

3.6 From the Inside

Esse breve estudo, de cerca de dois minutos, explora as possibilidades da utilização de uma câmera 360 da Samsung como geradora de material visual tridimensional. Com certeza, a maior dificuldade encontrada no processo foi utilizar a câmera de forma a gerar imagens realmente interessantes do ponto de vista artístico. Na prática, percebemos que sua característica mais individual, a de gravar em 360°, se apresenta também como o maior desafio a ser superado, uma vez ser quase impossível realizar uma gravação aparentemente descompromissada, sem que tenhamos pessoas olhando diretamente para a câmera ou o próprio operador constantemente visível. Para uma gravação realista ela se mostra a ferramenta perfeita, como observada nas gravações dos dois primeiros projetos descritos. Mas, para a composição de *música eletroacústica audiovisual*, encontrar situações em que o material gravado pudesse se aproximar do abstrato (parâmetro adotado como poética composicional), sem evocar diretamente a algo conhecido na natureza, que trouxesse consigo um referencial semântico já definido, foi extremamente difícil.

A estratégia encontrada para solucionar tal problema foi a de se explorar o interior de objetos encontrados dentro de casa, sejam eles armários, painéis, sacolas plásticas, vidros, caixas, que pudessem esconder o seu entorno da própria câmera. Tais objetos foram escolhidos também considerando que não podem ser completamente foscos, uma vez que, se nenhuma luz entra no objeto escolhido, não

²³⁰ <https://youtu.be/rz4ENoCMdKQ>

temos gravação alguma – princípio básico da filmagem ou fotografia de que precisamos de luz incidindo no objeto para ser refletida e formar imagens.

A espacialização foi realizada usando o pacote de plugins *FB360*, com o qual já observamos a dificuldade de se explorar reverberação baseada em convolução, por se tratar de um plugin que oferece o formato-B *Ambisonics* de segunda ordem. O material sonoro a se trabalhar também foi uma questão que demandou muita reflexão uma vez que os sons provenientes da própria filmagem, ou eram muito baixos ou nulos, ou simplesmente não possuíam informação suficiente para serem trabalhados de forma criativa (seja pela baixa qualidade do microfone da própria câmera ou pela sua relação sinal-ruído). A adoção de sons sintetizados, apesar de extremamente distantes, em seu material intrínseco, do visual, se mostrou a opção mais interessante. Uma vez que, no *espaço intrínseco*, na natureza do material sonoro, não seria encontrada qualquer coerência com o visual, devido principalmente às suas naturezas extremamente distantes, buscamos então atingir uma coerência no *espaço extrínseco*.

São exploradas basicamente 12 texturas visuais diferentes e, associada a elas, respectivas texturas sonoras. Em determinados momentos, no entanto, suas aparições são rápidas o suficientes para poderem ser observadas como gestos. Além das texturas, 3 gestos sonoros são associados diretamente a aparição de três dos elementos visuais. O maior interesse surge na forma como tais texturas audiovisuais apresentam seu *espaço extrínseco* nas duas dimensões, visual e sonora.

O estudo apresenta um percurso aparentemente caótico por dentro dos objetos escolhidos para serem filmados, onde eles são eventualmente processados e superpostos em diferentes combinações. Em uma narrativa gestual-coreográfica (BAALMAN, 2010) mais relacionada com o sonoro, são apresentados diversos momentos para ilustrar as relações de conformidade ou controvérsia dos materiais sonoros e visuais tanto no domínio do *espaço intrínseco* quanto *extrínseco*. Segue-se a descrição de momentos específicos onde tais relações se mostram mais evidentes.

Em 0'49'' uma textura sonora associada a uma imagem desfocada em movimento de rotação apresenta uma USE de salto no domínio sonoro, chamando muita atenção para um elemento visualmente pouco evidente. A textura sonora associada à visualização de dentro de um recipiente metálico perfurado, apresenta, em três momentos distintos, 0'34'', 1'05'' e 1'34'', movimentação no *espaço extrínseco*

em controvérsia com o material visual, trazendo a atenção para um elemento, inicialmente em segundo plano.

Em 0'15" a textura sonora associada a visualização obtida do interior de uma construção de plástico escura, em uma relação de contraste, se mostra com um movimento de rotação no *espaço extrínseco* bem lento enquanto o visual faz rotação rápidas; em 0'18", novamente em uma relação de contraste, a rotação do evento sonoro se mostra mais lento e em sentido contrário ao visual; em 0'24", explorando relações mais complexas que se modificam ao longo de apenas 10 segundos, a textura sonora percorre três USEs com sentidos diferentes, inicialmente realizando uma rotação mais lenta que o visual e em sentido contrário, em um segundo momento, mais rápida que o visual e no mesmo sentido, inclusive com modificações no *espaço intrínseco* que se aproximam mais da movimentação visual do que a própria movimentação no domínio do *espaço extrínseco*, e em um terceiro momento, novamente em sentido contrário e mais lento. Em 0'39" essa mesma textura retorna, dessa vez apresentando um movimento de rotação completamente coerente entre o visual e o sonoro, rápido e na mesma direção. Em 1'19" essa textura volta se apresentando coerente com o material visual tanto em seu movimento de rotação, que se inicia lento e vai acelerando, até ser interrompido pela textura que se segue, quanto em seu material no *espaço intrínseco* que apresenta modificações que guiam tal percepção. Em 1'27" ressurge um contraste, enquanto o visual se apresenta com uma movimentação lenta, o *espaço extrínseco* se encontra completamente estático e o *intrínseco* com uma agitação extremamente alta. Em 1'39" uma nova conformidade se apresenta tanto no *espaço extrínseco* sonoro e visual quanto no domínio do *espaço intrínseco*.

Em 0'46", 1'05", 1'11" e 1'32" temos uma textura sonora associada a uma imagem visual obtida de dentro de uma estrutura escura circular e estática. Em todas as aparições realizamos uma movimentação no *espaço extrínseco* do material sonoro explorando a USE granular, em contraste com a parte visual, completamente estática. Apesar de dissonante, tal realização, de certa forma, reforça a associação, inicialmente imprevisível.

Em 1'14" uma textura sonora associada a uma textura visual obtida com a gravação de dentro de um vidro canelado, realiza um movimento de rotação tanto em sentido contrário ao do visual, quanto em uma velocidade diferente, ilustrando, nesse

caso uma controvérsia. Essa mesma textura reaparece em 1'21'', 1'30'' e 1'46'', e é apresentada, no domínio do *espaço extrínseco*, como estática, apesar de um certo movimento lento no correspondente visual. A granulação do vidro canelado da imagem se mostra em concordância com uma certa granulação do material sonoro escolhido, no domínio do *espaço intrínseco*.

A textura sonora associada ao interior de um objeto feito de fios marrons trançados, apresentada em 0'16'', 0'21'', 0'26'', 0'32'' e 1'37'' explora, no domínio do *espaço extrínseco*, a USE em forma de ondas, em concordância com a movimentação do seu componente visual. No domínio do *espaço intrínseco* podemos observar um material sonoro também em forma de ondas, enquanto que no visual se aproximaria mais de um material granular.

O material sonoro apresentado em 0'50'', 0'54'', 0'58'' e 1'45'', associado visualmente com o interior de aparelhos domésticos, possui uma característica granular, no domínio do *espaço intrínseco*, em discordância com o seu *espaço extrínseco*, completamente estático. Por sua vez, tal estaticidade, se mostra em concordância em ambos os níveis do material visual.

Os gestos sonoros, associados ao abrir das portas em 0'01'', 0'10'', 1'53'' e 1'56'', ao aparecimento da cachorrinha em 0'26'' e 0'29'', e à sacola marrom com o logo vermelho em 0'43'', 0'50'', 0'56'', 1'01'', 1'08'' e 1'14'', foram espacializados de forma a condizer com suas correspondentes visuais, buscando uma certa replicação espacial.

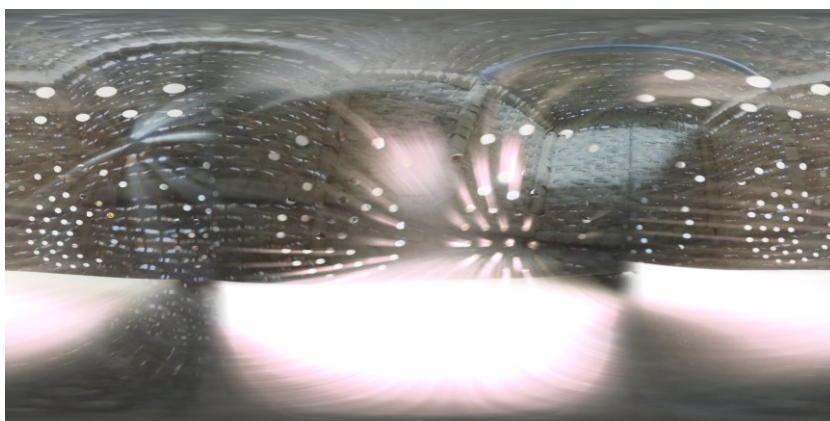


Figura 43 - Frames do estudo From the Inside e QR Code para link²³¹.

²³¹ <https://www.youtube.com/watch?v=gV8W6KzFuzM>

CONSIDERAÇÕES FINAIS E PESQUISAS FUTURAS

No presente trabalho foi realizado um apanhado conceitual e sistematizado em torno do que denominamos de *espaço* na música. A divisão dos diversos conceitos observados em apenas quatro categorias (*espaço externo*, *espaço interno*, *espaço intrínseco* e *espaço extrínseco*), por abarcar quase que a totalidade dos discursos encontrados, se mostrou esclarecedora e essencial para o desenvolvimento tanto do presente trabalho, como de outros que pretendem explorar o *espaço* como elemento da composição audiovisual.

Nos direcionando para um desenvolvimento estético do denominado *espaço extrínseco*, segregamos então as abordagens encontradas em duas: uma delas *sintática*, direcionada à estruturação formal; e a outra *semântica*, direcionada ao significado, ao referencial. É evidente a existência de uma *sintaxe* e uma *semântica* do *espaço intrínseco*, já abarcada por diversos outros trabalhos. Com relação ao *espaço extrínseco*, a *sintaxe* também se mostra bem evidente, no entanto uma *semântica* ainda está sendo construída. Difícil de ser tratada separadamente da *semântica* do próprio *espaço intrínseco*, a *semântica* do *espaço extrínseco* apresenta uma possível via através do que denominamos de USEs (Unidades Semióticas Espaciais), unidades mínimas perceptíveis de gestos e texturas no *espaço extrínseco* que podem ser combinadas para elaboração de gestos espaciais mais complexos.

Com esse arsenal de conceitos definidos, e apresentadas as diversas ferramentas e tecnologias atuais para a manipulação de áudio e vídeo tridimensional, nos voltamos então para o trabalho prático de manipulação de mídias audiovisuais adaptativas e de composição de música eletroacústica audiovisual para a realidade virtual. Focamos em aplicações destinadas à reprodução via *smartphones* acoplados à óculos de VR e fones de ouvido através de plataformas de streaming. Nelas, exploramos a aplicação dos conceitos vistos tanto no processo de criação quanto no processo de descrição e análise. Como um dos resultados principais da pesquisa, apresentamos seis trabalhos com diferentes naturezas, diferentes propósitos e diferentes abordagens, como exemplos dessa mídia emergente, a mídia audiovisual adaptativa.

A exploração do *espaço extrínseco* como uma nova mídia básica, de acordo com o conceito de Elleström (2017), passível de exploração controlada através de

composições específicas para realidade virtual, abre portas para um novo nível de tratamento estético desse parâmetro, não mais refém do espaço acústico ou do sistema de reprodução. O compositor agora tem em mãos uma forma de explorar poeticamente o que antes era experimental e limitado aos laboratórios.

Apesar da necessidade da realização de uma pesquisa quantitativa com a observação também de outros participantes, a espacialização perifônica (com altura) na produção de material audiovisual imersivo foi considerada, empiricamente, ao longo do desenvolvimento dos projetos, de extrema importância no processo de imersão. O desenvolvimento de material audiovisual para realidade virtual sem dúvida se beneficia da exploração de materiais sonoros posicionados acima e abaixo do ouvinte para que seja homogêneo, convincente e contribua para o processo de imersão.

Se compararmos a reprodução de audiovisual tridimensional para VR com a difusão de obras eletroacústicas através de grandes arranjos de alto-falantes, obviamente observamos uma certa diferença com relação à experiência estética. No entanto, a possibilidade de difusão das obras com áudio e vídeo tridimensional através de plataformas de streaming altamente difundidas, como o YouTube, nos aponta um futuro próximo onde mais canais de áudio serão suportados pelas plataformas, possibilitando o adoção de HOA como padrão de áudio, diminuindo cada vez mais a distância entre uma performance de música eletroacústica da experiência virtual. Com essa diferença minimizada, poderemos considerar a finalização uma obra em formato *Ambisonics* que possibilita diversas “transcrições”, podendo transitar entre diversos sistemas de reprodução, sem perdas significativas. Vislumbramos então a possibilidade de adaptação do material visual no formato equirretangular, juntamente com o áudio finalizado em HOA, originalmente objetivado nesse trabalho somente para a reprodução em realidade virtual, sendo adaptados para sistemas de grande escala que explorem telas 360° e arranjos de alto-falantes de diferentes dimensões, e vice-versa.

A nossa percepção do *espaço extrínseco* está, sem dúvida, interligada ao material intrínseco sonoro e a própria disciplina de acústica nos aponta relações entre a direcionalidade dos sons com sua ocupação no domínio das frequências. Com isso, no contexto da reprodução audiovisual para realidade virtual, vale observar também que, questões relativas à percepção da região grave do espectro sonoro e sua

importância para o processo de imersão e percepção do *espaço extrínseco*, devem ser considerados em trabalhos futuros, pois ainda precisam de aprimoramentos. O fato de os fone de ouvido não responderem a frequências muito graves e a nossa própria sensibilidade a sons graves se dar principalmente com o corpo ao invés dos ouvidos, levou a indústria de jogos, por exemplo, a explorar tal fenômeno através de artifícios tais como controles e cadeiras que vibram²³².

As possibilidades de se adicionar essa sensação de grave, assim como o cheiro e sensações táteis a uma obra de arte desenvolvida para realidade virtual, ainda é uma perspectiva futura, no entanto, não muito distante. Já observamos que estão em desenvolvimento as impressoras de cheiro²³³ e máquinas de reprodução de texturas²³⁴. Com a reprodução de audiovisual tridimensional associada a tais desenvolvimentos, uma modalidade artística ainda mais complexa poderia emergir.

²³² <http://www.cioffreda.com/projects/vybe-gaming-pad>

²³³ <https://www.youtube.com/watch?v=jnWIKPq5IHg&feature=youtu.be>

²³⁴ https://www.youtube.com/watch?v=lvtfD_rJ2hE

BIBLIOGRAFIA

375 WIKIPEDIANS. *VJing*. Greyscale Press, 2010.

ADRIAENSEN, Fons. *Acoustical impulse response measurement with ALIKI*. In.: 4th LINUX AUDIO CONFERENCE. **Anais**. Karlsruhe, Germany, 2006. Disponível em: <http://lac.zkm.de/2006/papers/lac2006_fons_adriaensen_01.pdf> Acesso em 9 Set 2019.

ALVES, Ana Margarida Duarte Brito. *O espaço na criação artística do século XX: heterogeneidade, tridimensionalidade, performatividade*. 2011. Dissertação (Doutorado em história da arte contemporânea). Faculdade de ciências sociais e humanas, Universidade de Nova Lisboa.

AMORIM, Felipe. *A Interpretação Fixada*. In.: **Revista Modus**, v. 8, nº 13, 2013. p. 51-58.

ARNHEIM, Rudolf. *Arte e percepção visual: uma psicologia da visão criadora*. Trad.: Ivonne Terezinha de Faria. São Paulo: Pioneira Thomsom Learning, 2005.

ARVIDSON, ASKANDER, BRUHN, FÜHRER. *Introduction*. In.: **Changing Borders: Contemporary Positions in Intermediality**. Intermedial Studies Press, 2007. p. 13-16. Disponível em: <http://www.ht.lu.se/media/utbildning/dokument/kurser/IMSB23/20141/changing_borders_dragen_dragen_.pdf> Acesso em 9 Set 2019.

ASSELOT, M.; Dellepiane, M.; Pietroni, N.; Scopigno, R.; Tsingos, N. *Reconstructing head models from photographs for individualized 3D-audio processing*. In.: **Pacific Graphics** vol 27 nº 7. 2008.

AUSTIN, L. and SMALLEY, D. *Sound diffusion in composition and performance: an interview with Denis Smalley*. In.: **Organised Sound**, v. 24, nº 2. United Kingdom: Cambridge Press, 2000. p. 10-21.

BAALMAN, M. *Spatial Composition Techniques and Sound Spatialization Technologies*. In.: **Organised Sound**, v. 15, nº 3. United Kingdom: Cambridge Press, 2010. p. 209-218.

BACHELARD, Gaston. *A poética do espaço*. São Paulo: Martins Fontes, 1993.

BARONE, Juliane Hernandez. *História da Realidade Virtual*. Disponível em <<https://www.bichosgeeks.com/historia-da-realidade-virtual-parte-1/>> Acesso em 9 Set 2019.

BARRETT, N. 2002. *Spatio-Musical Composition Strategies*. In.: **Organised Sound**, v. 7, nº 3. United Kingdom: Cambridge Press, 2002. p. 313-323.

BARRETT, N. 2005. *Adsonore*. In.: **Organised Sound**, v. 10, nº 2. United Kingdom: Cambridge Press, 2005. p. 111-119.

BATES, E. The Composition and Performance of Spatial Music. 2009. Tese (Doutorado em Música), Department of Music and Department of Electronic and Electrical Engineering, University of Dublin, Ireland.

BAYLE, François. *I-Son*. 1986. Disponível em <http://www.orema.dmu.ac.uk/analytical_toolbox/i-son> Acesso em 9 Set 2019.

BAYLE, François. *Space, and more*. In.: **Organised Sound**, v. 12, nº 3. United Kingdom: Cambridge Press, 2007. p. 241-249.

BECH, Soren and ZACHAROV, Nick. *Perceptual Audio Evaluation – Theory, Method and Application*. John Wiley & Sons: England, 2006.

BEGAULT, D. R. *Challenges to the successful implementation of 3-D sound*. In.: **Journal of the Audio Engineering Society**, v. 39, nº 11, 1991. p. 864-870.

BEGAULT, D. R. *3-D Sound for Virtual Reality and Multimedia*. Ames Research Center / NASA, Moffett Field, California, 2000.

BELL, A. G. *Experiments Relating to Binaural Audition*. In.: **American Journal of Otology**, 1880.

BERRY, V. T.; SHELTON, V. *Watching Music Interpretations of Visual Music Performance*. In.: **Journal of Communication Inquiry**, v. 23, nº 2. SAGE Social Science Collections, 1999. p. 132-151.

BETHÔNICO, Jalver. *Comunicação visual e comunicação sonora: uma relação semiótica*. 1995. Dissertação (Mestrado em Ciência da Informação). Minas Gerais: Universidade Federal de Minas Gerais – Programa de Pós-Graduação em Ciência da Informação.

BETHÔNICO, Jalver. *A Articulação dos Signos Audiovisuais*. 2001. Tese (Doutorado em Comunicação e Semiótica). São Paulo: Pontifícia Universidade Católica – Programa de Pós-Graduação em Comunicação e Semiótica.

BLAUERT, J. *Sound localization in the median plane*. In.: **Acoustica**, v. 22, 1969. p. 205-213.

BLAUERT, J. *Räumliches Hören*. Stuttgart: Hirzel Verlag. 1974. Repr. Trans. Spatial Hearing: the Psychophysics of Human Sound Localization. London: MIT Press. 1983, 1997.

BLUMLEIN, A. D. *Improvements in and relating to sound-transmission, sound-recording and sound-reproducing systems*. UK Patent GB 394325. 1933.

BORÉM, Fausto. *MaPA e EdiPA: Duas Ferramentas Analíticas para as Relações Texto-Som-Imagem em Vídeos de Música*. In.: **Musica Theorica**, Salvador: TeMA, 2016. p. 1-37.

BOONE, Marinus M.; VERHEIJEN, Edwin N. G. *Multi-Channel Sound Reproduction Based on Wave-Field Synthesis*. In.: 95th Convention of the AES. **Anais**. New York. 1993.

BRANT, Henry. *Space as an Essential aspect of Music Composition*. Ed. Schwartz, Elliot. In.: **Contemporary Composers on Contemporary Music**. New York: Holt, Rinehart and Winston. 1967.

BREGMAN, Albert S. *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge: MIT Press. 1990.

BROWN, Guy J. and WANG, DeLiang. *Binaural Sound Localization*. In.: **Computational Auditory Scene Analysis – Principles, Algorithms and Applications**. New Jersey: IEEE Press Wiley Interscience. 2005.

BLÜHDORN, Hardarik. *A relação entre pragmática, semântica e gramática*. In.: **Revista de Estudos da Linguagem**, v. 6, nº 2, Belo Horizonte - Minas Gerais. 1997. p. 150-188.

BYE, Kent. *50 years of VR with Tom Furness: The Super Cockpit, Virtual Retinal Display, HIT Lab, & Virtual World Society*. In.: **Voices of VR Podcast**. Disponível em: <<http://voicesofvr.com/245-50-years-of-vr-with-tom-furness-the-super-cockpit-virtual-retinal-display-hit-lab-virtual-world-society/>> Acesso em 9 Set 2019.

CAESAR, R. *Podcast n.9 Rodolfo Caesar*. In.: **Podcast ArteSonora, ArtRio - Feira Internacional de Arte Contemporânea do Rio de Janeiro**. 2013. Disponível em <<https://soundcloud.com/artesonora/podcast-09-rodolfo-caesar>> Acesso em 9 Set 2019.

CALLEJA, G. *Revising Immersion: A Conceptual Model for the Analysis of Digital Game Involvement*. Situated Play, Proceedings of the Third International Conference of the Digital Games Research Association. **Anais**. The University of Tokyo, Japan. 2007. p. 83-90.

CASTELLANI, Felipe Merker. *Uma abordagem sobre a noção de gesto musical nas poéticas de Luciano Berio e Brian Ferneyhough*. 2010. Dissertação (Mestrado em Música). Universidade Estadual de Campinas, São Paulo.

CHION, Michel. *Guide to Sound Objects*. Paris: Institut National de L`audiovisual & Éditions Buchet/Chastel, 1983.

CHION, Michel. *Audio-Vision: Sound On Screen*. Ed. Claudia Gorbman. Columbia University Press, New York. 1994.

CHION, M. *Les deux espaces de la musique concrète*. In.: **L'espace du son I**, ed. Fancis Dhomont. 1998. p. 31-33.

CHION, Michel. *Film, a sound art*. Trad. Claudia Gorbman. Columbia University Press, New York. 2009.

CHOWNING, John. *The simulation of moving sound sources*. In.: **Journal of the Audio Engineering Society**, v. 19, nº 1. 1971. p. 2-6.

CLOZIER, Christian. *Un instrument de diffusion: le Gmebaphone*. In.: **L'espace du son I**, ed. Fancis Dhomont. 1998. p. 56-57.

CLÜVER, Claus. *Inter textos / Interartes / Inter mídia*. In.: **Aletria**. 14, jul-dez, 2006, p.9-39. Disponível em: <<http://www.periodicos.letras.ufmg.br/index.php/aletria/article/view/1357/1454>> Acesso em 9 Set 2019.

COOK, Nicholas. *Analyzing musical multimedia*. Oxford: Oxford University Press. 1998.

COULTER, John. *The Language of Electroacoustic Music with Moving Images*. In.: **Electroacoustic Music Studies Network**. De Montfort, Leicester. 2007.

COULTER, John. *Electroacoustic Music with Moving Images: the art of media pairing*. In.: **Organised Sound**, v. 15, nº 1. United Kingdom: Cambridge Press, 2010. p. 26-34.

CSIKSZENTMIHALYI, M. *Finding Flow. The Psychology of Engagement with Everyday Life*. New York: Harper Collins Publishers. 1997.

DAVIES, Char. *Char Davies*. Palestra publicada na universidade UCLA Design Media Arts, em 11 de fev de 2002. Disponível em: <<https://www.youtube.com/watch?v=Ztljq2BY1qE>> Acesso em 9 Set 2019.

D'ANGELO, Thiago. *Tecnologias HMD para realidade Aumentada*. 16 jan. 2016. Disponível em <<http://www.decom.ufop.br/imobilis/tecnologias-hmd-para-realidade-aumentada-parte-1/>> Acesso em 9 Set 2019.

DERUTY, Emmanuel. *Creative convolution: new sounds from impulse responses*. In.: **Sound on Sound**. Set 2010. Disponível em: <<https://www.soundonsound.com/techniques/creative-convolution-new-sounds-impulse-responses>> Acesso em 9 Set 2019.

DIGNART, M. C. *Espaço, Gesto e Textura Musical na Música Eletroacústica: uma Abordagem Analítica e Composicional*. 2015. Tese (Doutorado em Música), Departamento de Comunicação e Arte, Universidade de Aveiro: Aveiro, Portugal.

DOVEY, J. And Kennedy, H.W. *Game Cultures: Computer games as new Media*. Berkshire: Open University Press. 2006.

DRUMMOND, Jon. *Understanding Interactive Systems*. In.: **Organised Sound**, v. 14, nº 2. United Kingdom: Cambridge Press. 2009. p. 124-133.

ELLESTRÖM, Lars. *Midialidade: ensaios sobre comunicação, semiótica e intermidialidade*. Porto Alegre: EdiPUCRS, 2017. Disponível em: <<http://ebooks.pucrs.br/edipucrs/livro/midialidade/>> Acesso em 9 Set 2019.

EVANS, Brian. *Foundations of a Visual Music*. In.: **Computer Music Journal**, v. 29, nº 4. Massachusetts Institute of Technology, 2005. p. 11-24.

FELDER, D. *An interview with Karlheinz Stockhausen*. In.: **Perspectives of New Music**, v. 16, nº 1, 1977. p. 85-101.

FERREIRA, Aurélio Buarque de Holanda. *Novo Dicionário da Língua Portuguesa*. 2. ed. re. aum. Rio de Janeiro: Nova Fronteira. 1986.

FIORIN, José Luiz. *Introdução ao pensamento de Bakhtin*. São Paulo: Ática. 2006.

FOUCAULT, Michel. *As palavras e as coisas: uma arqueologia das ciências humanas*. Trad.: Salma Tannus Muchail. 8ª edição, São Paulo: Martins Fontes, 1999.

FRY, Roger Eliot. *Vision and Design*. London: Chatto & Windus, 1920. Disponível em <<https://archive.org/details/visiondesign00fryr>> Acesso em 9 Set 2019.

GERZON, M. A. *Periphony: With-Height Sound Reproduction*. In.: **Journal of the Audio Engineering Society**, v. 21, nº 1. 1973. p. 2-10.

GERZON, M. A. *The design of precisely coincident microphone array for stereo and surround sound*. In.: 50th AES convention. **Anais**. London (UK), Mar 1975.

GIBBS, Tony. *The Fundamentals of Sonic Art and Sound Design*. Ava, Switzerland: 2007.

GRELA, Dante. *Análisis musical: una propuesta metodologica*. In: **Serie 5**. Rosario: Universidad de Rosario, 1986. p. 1-14.

HARLEY, M. A. *Space and Spatialization in Contemporary Music: History and Analysis Ideas and Implementations*. 1994. Tese (Doutorado em Musicologia). Montreal, Canada: School of Music, McGill University.

HARRISON, J. and WILSON, S. *Editorial*. In.: **Organised Sound**, v. 15, nº 3, United Kingdom, Cambridge Press, 2010. p. 183-184.

HAUTBOIS, Xavier. *Les Unités Sémiotiques Temporelles: de la sémiotique musicale vers une sémiotique générale du temps dans les arts*. In.: **Musimédiane** nº 5, 2010. Disponível em <<http://www.musimediane.com/numero5/02-SEMIOGENE/>> Acesso em 9 Set 2019.

HAUTBOIS, Xavier et al. *An experimental validation of Temporal Semiotic Units and Parameterized Time Motifs*. In.: **Musicae Scientiae**. v. 18, nº1. 2014. p 98-123.

HENRIKSEN, F.E. *Space in electroacoustic music: composition, performance and perception of musical space*. 2002. Tese (Doutorado em Música). City University London.

HILL, Andrew. *What is Electroacoustic Audio-Visual Music? Nomenclature and Cognition*. In.: Proceedings of Sound, Sight, Space and Play. **Anais**. De Montford University Leicester, United Kingdom, 2010.

HOLMAN, Tomlinson. *Surround Sound Up and Running*. 2nd edition. Oxford: Focal Press, 2008.

HOLMES, Thom. *Electronic and experimental music: technology, music, and culture*. 3rd ed. New York, NY: Routledge. 2008.

HUIBERTS, Sander. *Captivating Sound: The Role of Audio for Immersion in Computer Games*. 2010. Tese (Doutorado em Artes). Utrecht School of the Arts: Utrecht, The Netherlands.

ILIESCU, MIHU. *Glissandi and Traces: a Study of the Relationship between Musical and Extra-musical Fields*. In.: Definitive Proceedings of the International Symposium Iannis Xenakis. **Anais**. Athens, May 2005.

JUSTEL, E. *Vers une Syntaxe de l'espace*. In.: Dhomont, F. (ed.). **L'espace du son III**, Ohain, Ed. Musiques et Recherches, 2011. p. 111-131.

KASULIN, Aitana. *Video-música entre la conceptualización y la práctica*. 2015. Tese (Mestrado em Administração Cultural). Buenos Aires: Universidad de Buenos Aires, Facultad de Filosofía y Letras.

KENDALL, G. S. *A 3-D Sound Primer: Directional Hearing and Stereo Reproduction*. In.: **Computer Music Journal**, v. 19, nº 4, 1995. p. 23-46.

KLEIN, Georg. *Site-Sounds: On strategies of sound art in public space*. In.: **Organised Sound**, v. 14, nº 1. United Kingdom: Cambridge Press. 2009. p. 101-108.

KLEINA, Nilton. *Mais velha do que você pensa: a história da Realidade Virtual*. 7 Nov 2017. Disponível em < <https://www.tecmundo.com.br/mercado/123579-a-historia-da-realidade-virtual.htm> > Acesso em 9 Set 2019.

KOECHLIN, Charles. *Traité de l'orchestration*. 4 volumes. Paris: M Eschig, 1954–9.

KOECHLIN, Charles. *The Balance of Sonorities: Volume and Intensity*. In.: Mathews, Paul. **Orchestration**. New York: Routledge, 2006. p.139-146.

KWASTEK, Katja. *Aesthetics of Interaction in Digital Art*. Cambridge MIT Press. 2013.

LAFETÁ, Júlio César Alessi de Carvalho, *A intermedialidade no cinema de Peter Greenaway: uma análise intermediária do filme Prospero's books*. 2011. Dissertação (Mestrado em Artes Visuais). Escola de Belas Artes, Universidade Federal de Minas Gerais.

LEITE, Vania Dantas. *Relação som/imagem: um estudo da relação som/imagem na produção musical eletroacústica de compositores brasileiros atuantes no Rio de Janeiro: do gesto instrumental tradicional às interfaces interativas em tempo real*.

2004. Tese (Doutorado em Música). Universidade Federal do Estado do Rio de Janeiro.

LEWIN, D. *Generalized Musical Intervals and Transformations*. New York: Oxford, 2007.

LIMA, Marcelo Carneiro de. *Vídeo-Música*. 2011. Tese (Doutorado em Música). Universidade Federal do Estado do Rio de Janeiro.

LOSSIUS, T.; BALTAZAR, P.; and HOGUE, T. d. I. *DBAP - Distance-based amplitude panning*. In.: International Computer Music Conference (ICMC). **Anais**. Montreal, 2009.

LOWOOD, Henry E. *Virtual Reality*. 2018 In.: **Encyclopedia Britannica**. Disponível em < <https://www.britannica.com/technology/virtual-reality#ref884340> > Acesso em 9 Set 2019.

LUCENA, Karina de Castilhos. *Uma fenomenologia da imaginação através do espaço*. In.: **Revista eletrônica de crítica e teoria das literaturas**. PPG-LET-UFRGS, Porto Alegre, v. 3, nº1. 2007.

MALHAM, D. G. *Experience with Large Area 3-D Ambisonic Sound Systems*. In.: Proceedings of the Institute of Acoustics, vol. 14, nº 5. **Anais**. St. Albans: Institute of Acoustics. 1992. p. 209-216.

MALHAM, D. G. *Tutorial Article: Approaches to Spatialisation*. In.: **Organised Sound**, v. 3, nº 2, United Kingdom: Cambridge Press, 1998. p. 167-77.

MALHAM, D. G. *Surround Sound, Yesterday, Today and Tomorrow*. In.: **Rycote Microphone Data Book**, 2010. Disponível em <<http://microphone-data.com/media/filestore/articles/Surround%20sound-10.pdf>> Acesso em 9 Set 2019.

MANOURY, Philippe. *Le geste, la nature et le lieu: Un Démon das les circuits*. In.: **Traverses**. Paris: Centre Georges Pompidou. 1988. Também publicado em: O gesto, a natureza e o lugar. In.: Menezes, Flo (Trad.). *Música Eletroacústica: História e estéticas*. São Paulo: Fundação Editora da UNESP. 1996.

MARKS, Aaron. *The Complete Guide to Game Audio For Composers, Musicians, Sound Designers, and Game Developers*. 2nd ed. Elsevier Focal Press: Burlington, USA. 2009.

MARY, M. *De la Idea a la Obra: Los Caminos de la Creatividad. Orquestación Electroacústica y Polifonía del Espacio*. In.: **Sonic Ideas**, v. 5, n. 10, México. 2013. p. 55-61.

MCGEE, Ryan. *Spatial Modulation Synthesis*. Disponível em: <<http://www.spatialmodulation.com/>> Acesso em 9 Set 2019.

MENEZES, Flo. *Música Eletroacústica: História e estéticas*. São Paulo: Fundação Editora da UNESP. 1996.

MENEZES, Flo. *Atualidade estética da música eletroacústica*. São Paulo: Fundação Editora da UNESP. 1998.

MERIC, Renaud. *"Music is not a language..." Listening to Xenakis's electroacoustic Music*. In.: Proceedings of the Xenakis International Symposium. **Anais**. London: Southbank Centre. 2011.

MERLIER, B. *Vocabulaire de l'espace en musiques électroacoustiques*. Delatour France, 2006.

MILLER, Robert E. Robin. *Compatible PanAmbiophonic 4.1 and PerAmbiophonic 6.1 Surround Sound for Advanced Television – Beyond ITU 5.1*. In.: SMPTE 144th Technical Conference and Exhibition. **Anais**. Pasadena, CA, 2002.

MISCH, I., Hentschel, F., & Kohl, J. *On the Serial Shaping of Stockhausen's Gruppen für drei Orchester*. In.: **Perspectives of New Music**, v. 36, nº 1. 1998. p. 143-187.

MOLLER, H. *Fundamentals of Binaural technology*. In.: **Applied Acoustics**, v. 36. 1992. p. 171-218.

MOLLER, H. et al. *Binaural Technique: do we need Individual Recordings?* In.: **Journal of the Audio Engineering Society**, v. 44, nº 6. 1996. p. 451-469.

MURRAY, Christopher Brent. *Olivier Messiaen's Timbres-durées*. In.: Dingle, C. e Fallon, R. (ed.). **Messiaen Perspectives 1: Sources and Influences**. Routledge: New York, NY. 2016.

NATTIEZ, Jean-Jaques. *Music and Discourse: Towards a Semiology of Music*. Princeton: Princeton University Press. 1990.

NATTIEZ, Jean-Jaques. *Modelos linguísticos e análise das estruturas musicais*. REIS, Sandra L. de F. (trad.), SAMPAIO, Luiz P. (rev.) In.: **Per Musi**, v. 9. 2014. p. 5-46.

NETTINGSMEIER, Jörn. *Higher order Ambisonics – a future-proof 3D audio technique*. In.: Proceedings of 26 Tonmeistertagung – VDT international convention. **Anais**. Leipzig, Germany. 2010.

NORMANDEAU, R. *Timbre Spatialisation: The Medium is the Space*. In.: **Organised Sound**, v. 14, nº 3. United Kingdom: Cambridge Press. 2009. p. 277-285.

PALOMBINI, Carlos. *Musique Concrète Revisited*. In.: **Electronic Musicological Review**, vol. 4, UFPr Arts Department. 1999. Disponível em <http://www.rem.ufpr.br/_REM/REMr4/vol4/arti-palombini.htm> Acesso em 9 Set 2019.

PARRET, Herman. *Pragmática*. In.: **Cadernos de Estudos Linguísticos**. São Paulo, Unicamp. nº 7. 1984. p. 39-51.

PATRICK E., Cosgrove D., Slavkovic A., Rode J.A. Verratti T., Chiselko G. *Using a Large Projection Screen as an Alternative to Head-Mounted Displays for Virtual Environments*. In: Proceedings of the SIGCHI conference on Human factors in computing systems. **Anais**. ACM press, New York. 2000. p. 478-485.

PEREIRA, Ivo; Nogueira, Nuno. *Realidade Virtual*. Disponível em: <<http://web.ist.utl.pt/ist170613/>> Acesso em 9 Set 2019.

PLACK, Christopher J. *Spatial hearing*. In.: **The Sense of Hearing**. England: Lawrence Erlbaum Associates. 2005.

PRIMO, A. F. T; Cassol, M. B. F. *Explorando o conceito de interatividade definições e taxonomias*. Disponível em <<http://www.ufjf.br/grupar/referencias/>> Acesso em 9 Set. 2019.

PULKKI, V. *Virtual sound source positioning using vector base amplitude panning*. In.: **Journal of the Audio Engineering Society**, v. 45, nº 6. 1997. p. 456-466.

RAJEWSKY, Irina O. *Intermediality, intertextuality and Remediation: A Literary Perspective on Intermediality*. In.: **Intermedialités**. Canadá: Collection Érudit, nº 6. 2005. p. 43-64.

REKVELT, Joost. *Light matters: the origin of the term visual music*. Disponível em <<http://www.joostrekveld.net/?p=1105>> Acesso em 9 Set 2019.

RIPPL, Gabriele. *Introduction: Why Intermediality?* In: RIPPL, Gabriele (Ed.). **Handbook of Intermediality: Literature-Image-Sound-Music**. Berlin: De Gruyter Mouton. 2015. p. 1-31.

ROUSTÉ, Joan Riera. *Audição Espacial e Percepção do Som na Composição Musical*. 2014. Tese (Doutorado em Música). Departamento de Comunicação e Arte, Universidade de Aveiro, Portugal.

ROLLINGS, A. & Adams, E. *Andrew Rollings and Ernest Adams on Game Design*. Indianapolis: New Riders Publishing, USA. 2003.

ROSSON, Lois. *The Virtual Interface Environment Workstation (VIEW)*. 2014. Disponível em <https://www.nasa.gov/ames/spinoff/new_continent_of_ideas/> Acesso em 9 Set 2019.

ROWE, Robert. *Interactive Music Systems*. Cambridge: MIT, 1993.

RUMSEY, Francis. *Spatial Audio*. Oxford: Focal Press. 2001.

RUMSEY, F. *Spatial quality evaluation for reproduced sound: terminology, meaning and scene-based paradigm*. In.: **Journal of Audio Engineering Society**, v. 50, nº 2. 2002. p. 651-666.

SANCHES, Alexandre Martinello. *Estruturação Audiovisual Através de Qualidades Morfológicas*. 2010. Dissertação (Mestrado em Música). Campinas, SP.

SANTAELLA, L. *Matrizes da linguagem e pensamento: Sonora, Visual, Verbal. Aplicações na hipermídia*. São Paulo: Iluminuras / Fapesp, 2001.

SCHAEFFER, Pierre. *Ensaio sobre o rádio e o cinema: estética e técnica das artes-relé 1941-1942*. Trad.: Carlos Palombini. Belo Horizonte: Editora UFMG, 2010.

SCHAEFFER, Pierre. *Traité des objets musicaux: essai interdisciplines*. Trad. : Martinazzo, Ivo. Brasília: Editora Universidade de Brasília (1993). 1966.

SCHULZ, B., Ed. *Robin Minard: Silent Music*. Heidelberg: Kehrer Verlag, 1999.

SMALLEY, D. *The Listening Imagination: Listening in the Electroacoustic Era*. In.: **Contemporary Music Review**, v. 13, nº 2, Harwood Academic Publishers, 1996. p. 77-107.

SMALLEY, D. *Spectromorphology: Explaining Sound-Shapes*. In.: **Organised Sound**, v. 2, nº 2, United Kingdom: Cambridge Press, 1997. p. 107-126.

SMALLEY, D. *Space-form and the Acousmatic Image*. In.: **Organised Sound**, v. 12, nº 1. United Kingdom: Cambridge University Press. 2007. p. 37-58.

SMALLEY, Denis. *Spatial Experience in Electro-acoustic Music*. In: Dhomont, F. **L'espace du son II**. Ohain: Musiques et Recherches. 2008. p.123-126.

SOUSA, F. W. J. *Mixagem em surround 5.1 e transcrição para audição Binaural de músicas eletroacústicas espacializadas: o caso de "4 sketches em movimento"*. In.: Convention paper presented at the 8th AES brazilian convention. **Anais**. São Paulo (Brazil). 2010a.

SOUSA, F. W. J. *Mixagem de áudio multicanal no formato surround 5.1 e sua transcrição para sistemas Binaurais*. 2010b. Dissertação (Mestrado em Música). Belo Horizonte: Universidade Federal de Minas Gerais.

SOUSA, F. W. J. *The development of a 'Virtual Studio' for monitoring Ambisonic based multichannel loudspeaker arrays through headphones*. 2011. Dissertação (Mestrado em Música e Tecnologia). York - Inglaterra: Universidade de York.

SOUSA, F. W. J. *Espacialização de áudio*. In.: Convention paper presented at the 11st AES brazilian convention. **Anais**. São Paulo (Brazil), 2013.

STEUER, Jonathan. *Defining virtual reality: dimensions determining telepresence*. In.: **Journal of Communication**, v. 42, nº 4. 1992. p. 72-93.

STOCKHAUSEN, Karlshinz. *Musik in Raum*. In.: **Die Reihe**. nº 5. 1959. Also published in Stockhausen, *Texte zur elektronische und instrumentale Musik*. Bd. I. 1952-1962. Aufsätze zur Theorie des Komponierens. Cologne: Verlag. M. DuMont Schanberg, 1963. [English transl. by Ruth Koenig: *Music in Space*. Die Reihe nº5 (1961): 67-82].

TAGG, P. *Analysing popular music: theory, method and practice*. In.: **Popular Music**, v. 2. 1982. p. 37-65.

TRUAX, Barry. *Soundscape, Acoustic Communication & Environmental Sound Composition*. In.: **Contemporary Music Review**, v. 15, nº 1. 1996. p. 49-65.

VAGGIONE, Horacio. *Perspectives de l'électroacoustique*. In. : **Chimères**, nº 40, Bourges: Mnemosyne, 2000. p. 1-11.

VANDE GORNE, A. *L'interprétation spatiale*. In.: **Revue DEMéter- Revue Électronique**, Université de Lille-3, 2002. Disponível em: <<http://demeter.revue.univ-lille3.fr/interpretation/vandegorne.pdf>> Acesso em: 9 Set 2019.

VANDE GORNE, A. *Traité D'écriture su support*. Édition Musiques & Recherches: Revue LIEN VIII. 2017. Versão eletrônica disponível em: <<https://sites.inagrm.com/avdg/index.xhtml>> Acesso em: 26 Set 2019.

WALLACH, H. *On sound localization*. In.: **Journal of Acoustic Society of America**, v. 10. 1939. p. 270-274.

WALLACH, H. *The role of head movements and vestibular and visual cues in sound localization*. In.: **Journal of Experimental Psychology**, v. 27. 1940. p. 339-368.

WENZEL, E. M., Wightman, F. L., and Foster, S. H. *A virtual display system for conveying three-dimensional acoustic information*. In.: 32nd Annual Meeting of the Human Factors and Ergonomics Society. **Anais**. Santa Monica: Human Factors and Ergonomics Society. 1988.

WILSON, Scott. *Spatial Swarm Granulation*. In.: Proceedings of the International Computer Music Conference. **Anais**. 2008.

WOLF, Werner. *The Musicalization of Fiction: A Study in the Theory and History of Intermediality*. Amsterdam e Atlanta, GA: Rodopi. 1999.

ANEXOS

Anexo 1 – 15 figuras do espaço

>Tradução e adaptação, do francês, das 15 *figuras do espaço* propostas pela compositora Vande Gorne (VANDE GORNE, 2002):

Gesto: Misturar

Execução: Lenta transição entre dois pares ou grupos de alto-falantes. O gesto deve ser cuidadoso para não criar “buracos” acústicos.

Função: Fortalecimento da mixagem existente no suporte, mudanças de planos, profundidades ou larguras.

Observações: Possibilidade da criação de trajetórias através de mixagens sucessivas.

Gesto: Desmanchar

Execução: Gesto inverso ao misturar, uma vez que, a partir de uma massa dada, apresentada através de um par ou grupo de alto-falantes, se realiza através da redução da amplitude ou supressão de alto-falantes. O gesto pode ser suave, imperceptível, ou abrupto, usando o *mute*, por exemplo

Função: Representação de um desmanche do suporte, alteração do plano de profundidade ou calibre.

Observações: Possibilidade de imposição de uma trajetória ao som através de desmanches sucessivos.

Gesto: Ênfase

Execução: Colocar em evidência uma localização precisa (solistas por exemplo) ou um grupo formando um espaço particular, um volume, um novo calibre, simplesmente aumentando ligeiramente a amplitude dos alto-falantes selecionados, sem alterar os outros.

Função: Pode ser pontual na obra ou ser uma estratégia geral para a projeção, caso em que se abre um pouco os potenciômetros principais (criando uma base) e, em seguida, se destaca alguns pontos da estrutura, as seções da obra, o efeito desejado sobre a percepção do ouvinte.

Observações: A ênfase é um jogo flexível, leve e descontraído na mesa de mixagem.

Gesto: Tremulação

Execução: Rápidas operações de ênfase e não-ênfase consecutivas adicionando aleatoriedade. Pode ser realizada sobre a amplitude do sinal ou no espectro (através de filtragem). Para assegurar a continuidade, manter canais opostos abertos e fixos.

Função: Equivalente espacial de tremolo, realçar um momento de micro montagem ou pontilista (granular).

Observações: ---

Gesto: Oscilação

Execução: Alternância rápida e regular entre dois alto-falantes ou dois grupos de altifalantes, na amplitude ou no espectro. Mais facilmente obtida através de um mixer digital automatizado ou programação de um LFO.

Função: Equivalente espacial da trilo, ou trinado, agitação que prepara uma explosão, uma explosão, ou o contrário, fazendo um refluxo vivo, uma espera.

Observações: ---

Gesto: Balanço

Execução: Alternância lenta e gestual entre duas colunas ou dois grupos de alto-falantes, alternância dinâmica ou espectral.

Função: Evidenciar uma escritura de diálogo, delimitação de limites espaciais, canção de ninar.

Observações: ---

Gesto: Onda

Execução: Trajetos de ida e vinda que atravessa misturas ou desmanches sucessivos, pode ser realizado no eixo frente-trás do palco, no eixo lateral, na rotação ou em partes dela.

Função: Movimento de massa unidirecionalmente previsível.

Observações: A vantagem desse movimento é a de se aderir a arquétipos agógicos conhecidos.

Gesto: Rotação

Execução: Trajeto circular entre quatro alto-falantes através da cena ou, mais frequentemente, em torno do público, realizado através de um gesto ligeiro de acentuação de cada aresta de forma sucessiva.

Função: Evidenciar a rotação interna de um som (deve seguir um ritmo de deslocação), criação de movimentos, por exemplo, com sentido de confinamento.

Observações: ---

Gesto: Espiral

Execução: Trajetória circular sobre a qual é aplicada uma aceleração ou desaceleração, o ponto culminante seria em outro lugar.

Função: Preparação, anúncio, caminho objetivo orientado ou conclusivo.

Observações: ---

Gesto: Rebote

Execução: Salto rápido a partir de um ponto a outro no espaço, a partir de um grupo para o outro, a partir de um solista a um grupo (e vice-versa), com um gesto de alternância rápida na console ou um jogo com o mute. É melhor realizada quando ambos os polos estão distantes.

Função: (Re) lançamento no gatilho. Alcançar uma outra área do espaço sem transição.

Observações: ---

Gesto: Inserção / ruptura

Execução: Em uma área já estabelecida, em um período suficientemente longo, realização de uma transição abrupta ou sobreposição (através do mute) de uma área de caráter e / ou tamanho diferente.

Função: Figura da retórica, colocar em evidência a escritura de inserção sobre um suporte, acentuação. A ruptura pode servir como um desvio abrupto e contrastante em relação a um outro estado.

Observações: Exemplo: uma grande massa difusa inserir um solo (calibre estreito) direcional.

Gesto: Aparecimento / desaparecimento

Execução: Erupção / fechamento não preparado de um estado espacial diferente, sobreposta à anterior ou seguinte. O uso do mute é a melhor maneira de se realizar.

Função: Surpresa, "mágica", acordar da escuta.

Observações: ---

Gesto: Explosão

Execução: Transição abrupta de um espaço limitado ou estreito em um espaço largo e ambiental, não difuso.

Função: Destacar uma massa eruptiva, de uma morfologia característica e enérgica.

Observações: ---

Gesto: Acumulação

Execução: Adição sucessiva de planos e / ou calibres uns sobre os outros para alcançar um tutti espacial.

Função: Destacar um material de som corpuscular (grãos) ou ampliação progressiva de uma trama.

Observações: ---

Gesto: Invasão

Execução: Acumulação rápida, trajeto acumulativo direcionado ao público.

Função: Como "a chegada do trem a estação de Montparnasse", dos irmãos Lumière, efeito de agressão subjetiva.

Observações: ---

Anexo 2 – 19 USTs

>Descrições morfológicas e semânticas de cada uma das 19 USTs definidas por Hautbois, adaptadas e traduzidas (HAUTBOIS, 2014):

A- USTs Invariantes por repetição:

01- UST: Ondulando

Descrição Morfológica: Única fase, não delimitada por tempo, composta da repetição lenta de um motivo de som crescente e decrescente. A forma do perfil pode envolver diferentes critérios morfológicos (massa, dinâmica, grão, etc.).

Descrição Semântica: Cada ciclo transmite a sensação de ser empurrado para a frente e depois levado de volta até o fim. Ficamos com a impressão de que estamos estagnando através desta unidade apesar de sentirmos movimento dentro de cada ciclo.

02- UST: Rodando

Descrição Morfológica: Única fase, não delimitada por tempo, na qual um parâmetro (pitch, timbre) é impulsionado por uma repetição cíclica rápida, juntamente com um empuxo em cada ciclo.

Descrição Semântica: Temos a sensação de que um objeto está girando em torno de si ou no espaço.

03- UST: Obsessivo

Descrição Morfológica: Única fase, não delimitada por tempo, com uma rápida e possivelmente repetição variada de um elemento pulsado.

Descrição Semântica: Sentimo-nos constrangidos por um processo mecânico sobre o qual não conseguimos agir.

B- USTs Invariantes por estagnação:

04- UST: Estacionário

Descrição Morfológica: Unidade não delimitada no tempo com uma lenta evolução temporal. Mostra uma regularidade temporal global, mas elementos aleatórios podem aparecer em outros níveis. Diversas configurações morfológicas podem ser encontradas: pouca alteração na estrutura temporal, detalhes na permanência global ou ciclos lentos com pouca variedade.

Descrição Semântica: A unidade parece não chegar a lugar nenhum: não há progresso, mesmo que haja sempre algo acontecendo.

05- UST: Flutuando

Descrição Morfológica: Unidade não delimitada no tempo com uma lenta evolução temporal. Sucessão irregular de elementos temporais desconexos e relativamente breves.

Descrição Semântica: Apesar dos eventos sonoros que aparecem aleatoriamente, não há sentimento de expectativa ou suspense. A maneira como os objetos são distribuídos no tempo é percebida como um fluxo linear.

06- UST: Em suspensão

Descrição Morfológica: Unidade não delimitada no tempo formada por uma sequência repetida, sem variação significativa, com uma lenta progressão temporal. A matéria sonora e eventos temporais não mudam muito.

Descrição Semântica: O equilíbrio das forças operacionais transmite uma sensação de quietude relacionado a um sentimento de expectativa hesitante. Sabemos que algo vai acontecer, mas não o quê ou quando.

C- USTs Invariantes por efeito caótico:

07- UST: Divergente

Descrição Morfológica: Unidade não delimitada no tempo caracterizada por uma sucessão de breves sequencias temporais sem relação umas com as outras.

Descrição Semântica: Falta de causalidade entre os eventos temporais sucessivos. O grande número de direções, sem aparente conexão, resulta em indecisão.

08- UST: Caótico

Descrição Morfológica: Unidade não delimitada no tempo caracterizada pela densa superposição de sequencias temporais sem relação umas com as outras.

Descrição Semântica: Falta de causalidade entre os eventos temporais superpostos. Isso causa um burburinho indescritível, uma alta saturação, uma sensação de incomunicabilidade e tensão.

D- USTs variantes com desenvolvimento uniforme:

09- UST: Avançando

Descrição Morfológica: Unidade com única fase, não delimitada por tempo, com uma unidade globalmente uniforme e processo ininterrupto em que uma breve sequência repetida renova regularmente a atividade condutora.

Descrição Semântica: Parece regularmente nos empurrar para a frente. Isso nos dá a impressão de que estamos nos movendo decisivamente em uma direção precisa.

10- UST: Trajetória infinita

Descrição Morfológica: Unidade com única fase, não delimitada por tempo, com uma evolução linear e geralmente lenta de um parâmetro sonoro.

Descrição Semântica: O processo deve ser orientado em uma direção (para cima ou para baixo, por exemplo) e, no entanto, parece nunca acabar. O fenômeno sonoro deve ser longo o suficiente para ser percebido como um processo e não um evento efêmero.

11- UST: Peso

Descrição Morfológica: Unidade não delimitada no tempo caracterizada por uma breve sequência repetida lenta e irregularmente.

Descrição Semântica: Parece ter problemas para progredir, apesar de uma energia motriz.

E- USTs variantes com desenvolvimento frustrado:

12- UST: Desaparecendo

Descrição Morfológica: Unidade delimitada por tempo com uma única fase. Extinção progressiva do tipo ressonância, sem qualquer fonte de energia complementar.

Descrição Semântica: É como a trajetória de um veleiro com suas velas para baixo, que segue em frente, desacelerando, por causa do momento. Processo previsível até a sua conclusão.

13- UST: Travando

Descrição Morfológica: Unidade delimitada por tempo com duas fases: um processo já iniciado e que de repente se opõe a outro processo com uma energia decrescente.

Descrição Semântica: Dá a impressão de uma desaceleração forçada. Todo movimento é abruptamente contido até parar.

14- UST: Alongando

Descrição Morfológica: Unidade de única fase delimitada pelo tempo. Crescimento linear de uma característica morfológica, com aumento de energia, em uma estrutura temporal congelada.

Descrição Semântica: Parece ir adiante ao máximo de um processo (ou esforço). Sensação de um estiramento de som envolvendo duas forças opostas.

15- UST: Querendo começar

Descrição Morfológica: Unidade não delimitada por tempo composta por duas fases repetidas. A primeira é uma forma articulada, relativamente curta. A segunda é o oposto da primeira (em massa, dinâmica, homogeneidade). A reiteração entre as duas fases não é estrita: existe pelo menos um parâmetro que varia.

Descrição Semântica: É como algo tentando começar. A reiteração sugere um repetido esforço para alcançar um objetivo.

F- USTs variantes com o equilíbrio interrompido:

16- UST: Caindo

Descrição Morfológica: Unidade de duas fases delimitada pelo tempo. A primeira fase é globalmente uniforme, enquanto a segunda é uma mudança acelerada na altura (para cima ou para baixo). O limite dessa transição é nítida, não contínua.

Descrição Semântica: Um estado de equilíbrio instável. Suspensão seguida de pivotamento (a fase de suspensão é realizada após o fato).

17- UST: Propulsão

Descrição Morfológica: Unidade delimitada pelo tempo com três fases. A primeira fase é um fulcro bastante sustentado: um som prolongado e homogêneo ou uma iteração lenta, globalmente uniforme. A fase seguinte é uma breve aceleração de intensidade, altura ou qualquer outra característica morfológica. A terceira fase é uma ressonância típica ou silêncio.

Descrição Semântica: Sentimos a aplicação de uma força a um estado estável, resultando em um movimento acelerado. Projeção a partir de um ponto de partida.

18- UST: Comprimindo - Alongando

Descrição Morfológica: Unidade delimitada por tempo com duas fases contrastantes. Durante a fase de compressão, a matéria sonora é descontínua e errática. A fase de alongamento é um segmento globalmente uniforme.

Descrição Semântica: Primeiro, há uma sensação de compressão (como se nós pressionássemos algo a um obstáculo), então a barreira é subitamente superada, suprimindo toda a resistência e liberando energia. É uma mudança repentina de energia localizada para energia dispersa.

19- UST: Suspendendo - Questionando

Descrição Morfológica: Unidade delimitada pelo tempo com duas fases contrastantes. A primeira fase é uma sequência temporal breve e geralmente variada e repetida. A segunda fase contrasta morfológicamente com a primeira: um som sustentado e crescente ou um silêncio.

Descrição Semântica: Um movimento de interrupção em uma posição estática.

Anexo 3 – Descrição dos plugins utilizados nos projetos

ATK for Reaper

Provavelmente o pacote de ferramentas, para trabalhar com sinais de áudio em formato *Ambisonics*, mais simples e fácil de utilizar em qualquer *DAW*²³⁵ seja o *ATK for Reaper*. O fato de ser compatível com a maior parte das plataformas que suportam *VSTs* e possuir uma instalação bem simples se associa com sua oferta de um conjunto de codificadores, transformadores e decodificadores de *Ambisonics* de primeira ordem com uma interface gráfica extremamente intuitiva e que ajuda muito na visualização do processo realizado. Exploramos o seu funcionamento nos quatro primeiros trabalhos expostos no capítulo 3, a mixagem de um videoclipe para *VR* de um grupo de Jazz, a mixagem de uma trilha sonora para uma série de três vídeos produzidos pelo Sebrae, na produção das paisagens sonoras para o projeto *Corpografias* e na obra *Letter Piece*.

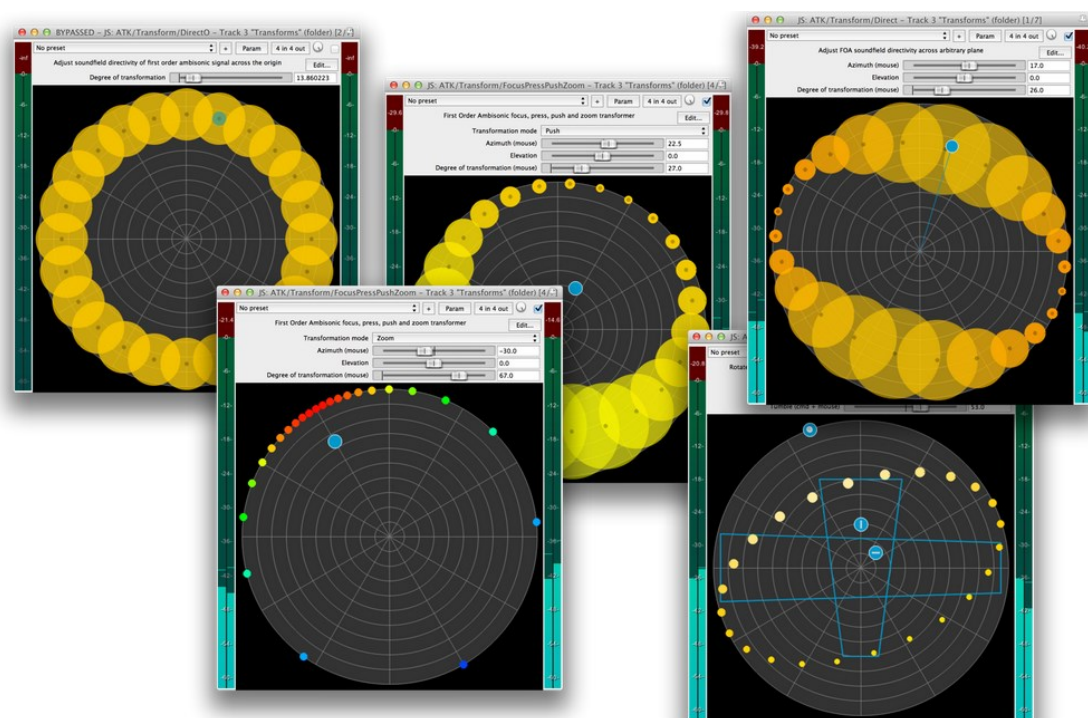


Figura 44 - Interfaces do usuário de alguns dos plugins *ATK for Reaper*²³⁶.

Entre os codificadores ele oferece: o *Omni*, que codifica um sinal mono como um campo sonoro omnidirecional; o *Planewave*, ferramenta utilizada nas produções apontadas acima, que fornece codificação clássica de uma fonte mono como uma

²³⁵ *Digital Audio Workstation*, ou Estação de Trabalho de Áudio Digital. Software de gravação e manipulação de áudio digital multipista (Ex.: Protools, Reaper, Logic, Cubase, Nuendo, Vegas, etc.).

²³⁶ Fonte: <http://www.ambisonictoolkit.net/documentation/reaper/>

onda plana, onde a direção de chegada (azimute e elevação) pode ser definida; o *Spread*, que codifica um sinal monofônico girando suavemente o sinal pelo campo sonoro por faixa de frequências; o *Diffuse*, que aleatoriza a fase do sinal monofônico de entrada para criar um campo difuso; o *Stereo*, que codifica os canais estéreo esquerdo e direito como duas ondas planas vindos da esquerda e da direita - o espalhamento angular entre as duas ondas é parametrizado; o *SuperStereo*, que é um método clássico de expandir o campo sonoro apresentado por sinais estereofônicos para *Ambisonics*; o *UHJ*, que oferece acesso a várias gravações publicadas para audição perifônica em 2D; o *Planewave transcoders*, que oferece a codificação para vários formatos de distribuição padrões: Quad, 5.0 e 7.0 padrões ITU, Pantophonic para 2D e Periphonic para matrizes 3D; e o *AtoB*, que fornece a codificação do formato-A, para os quatro sinais separados e ordenados em W, X, Y e Z, relacionados, criando o campo sonoro tridimensional.

Entre os transformadores do campo sonoro já codificado, o *ATK for Reaper* oferece: o *Direct*, onde é possível ajustar a diretividade do campo sonoro em um plano especificado pelo usuário; o *DirectO*, onde é possível ajustar a diretividade de um campo sonoro de primeira ordem em sua origem, funcionando como um filtro passa-baixa espacial em que, com um grau crescente de transformação, o campo sonoro torna-se menos direcional, eventualmente tornando-se omni; o *FocusPressPushZoom*, consiste em uma interface unificada para quatro diferentes transformações espaciais: *Focus* e *Zoom* são transformações relacionadas à dominância e podem ser descritas como enfatizadores de elementos na direção do interesse, o *Press* e *Push*, ao invés de enfatizar elementos na direção de um alvo, faz com que todos sejam aproximados ou afastados da direção do interesse (com o aumento desse tipo de transformação espacial todas as quatro transformações colapsam o campo sonoro em uma planagem que chega da direção de interesse especificada por azimute e elevação); o *Dominance*, aumenta o ganho de elementos na direção de interesse enquanto diminui o ganho de elementos na direção oposta; o *Mirror*, espelha o campo sonoro através de um plano arbitrário; o *MirrorO*, espelha o campo sonoro através da origem; e o *RotateTiltTumble* que realiza rotações no campo sonoro em *Ambisonics* de primeira ordem nos três eixos.

Além desses transformadores o *ATK for Reaper* oferece também dois efeitos de simulação de campo próximo sendo eles o de proximidade que é capaz de simular o

efeito de aproximação nos sinais codificados, e o de campo próximo, capaz de reduzir ou anular o efeito de proximidade dos sinais codificados.

Os decodificadores oferecidos pelo *ATK for Reaper* são o *Mono*, que retorna um único canal e pode ser usado para se ouvir um determinado ponto no campo sonoro definido por um ângulo e elevação especificados; o *Stereo*, que retorna um par de microfones virtuais; o *UHJ*, que também é um decodificador estereofônico, no entanto mais indicado para o trabalho de produção; o *Quad*, decodificador quadrafônico otimizado com ângulo de alto-falante variável; o *5.0*, que utiliza os decodificadores otimizados de Bruce Wiggins para um sistema com cinco alto-falantes; o *Pantophonic2D*, um decodificador de polígono 2D regular; o *BtoA*, que decodifica novamente o campo sonoro em formato-B para o formato-A; e o decodificador *Binaural*, direcionado para a escuta em fones de ouvido.

FB360

O *Facebook 360 Spatial Workstation* é um pacote de plugins para manipular áudio espacial para vídeo 360 e VR cinematográfico que inclui plugins para as *DAWs* mais populares, em formato *VST*, assim como um reproduutor de vídeo 360 sincronizável com a linha temporal do software de áudio, e utilitários diversos para ajudar a projetar e publicar áudio espacial em vários formatos. O áudio produzido pode ser direcionado diretamente para o *Feed* de notícias do Facebook em dispositivos Android e iOS, no Chrome para computador e no *Samsung Gear VR* por meio de fones de ouvido.

O principal plugin desse pacote é o *Audio 360 Spatialiser*, onde pode-se realizar praticamente todas as operação de movimentação da fonte sonora incluindo ângulo, distância, elevação, efeito *Doppler*, simulação das reflexões de uma sala, entre outros parâmetros. Um interessante recurso desse plugin é o de se poder abrir um vídeo dentro dele mesmo e se delimitar uma área (normalmente um rosto de uma pessoa) com o qual será realizado um rastreamento do canal de áudio, ou seja, uma vez essa área demarcada, o áudio irá se movimentar a partir da movimentação da mesma²³⁷. A entrada desse plugin pode ser um canal mono, estéreo, *Ambisonics* formato-B, 4.0, 5.0, 6.0 ou 7.0. De acordo com a entrada irão aparecer uma, duas ou até sete fontes

²³⁷ A descrição e observação dessa ferramenta em funcionamento pode ser conferida em diversos vídeos tutoriais disponíveis no YouTube (<https://www.youtube.com/watch?v=rdGjedVUvpo>)

sonoras para serem especializadas. Lembrando que no formato-B *Ambisonics*, sinal já transformado do formato-A proveniente de um microfone tetraédrico, essas opções estarão reduzidas a manipulação do campo sonoro nos eixos X, Y e Z, e no envolvimento (dominância do canal W, omnidirecional).

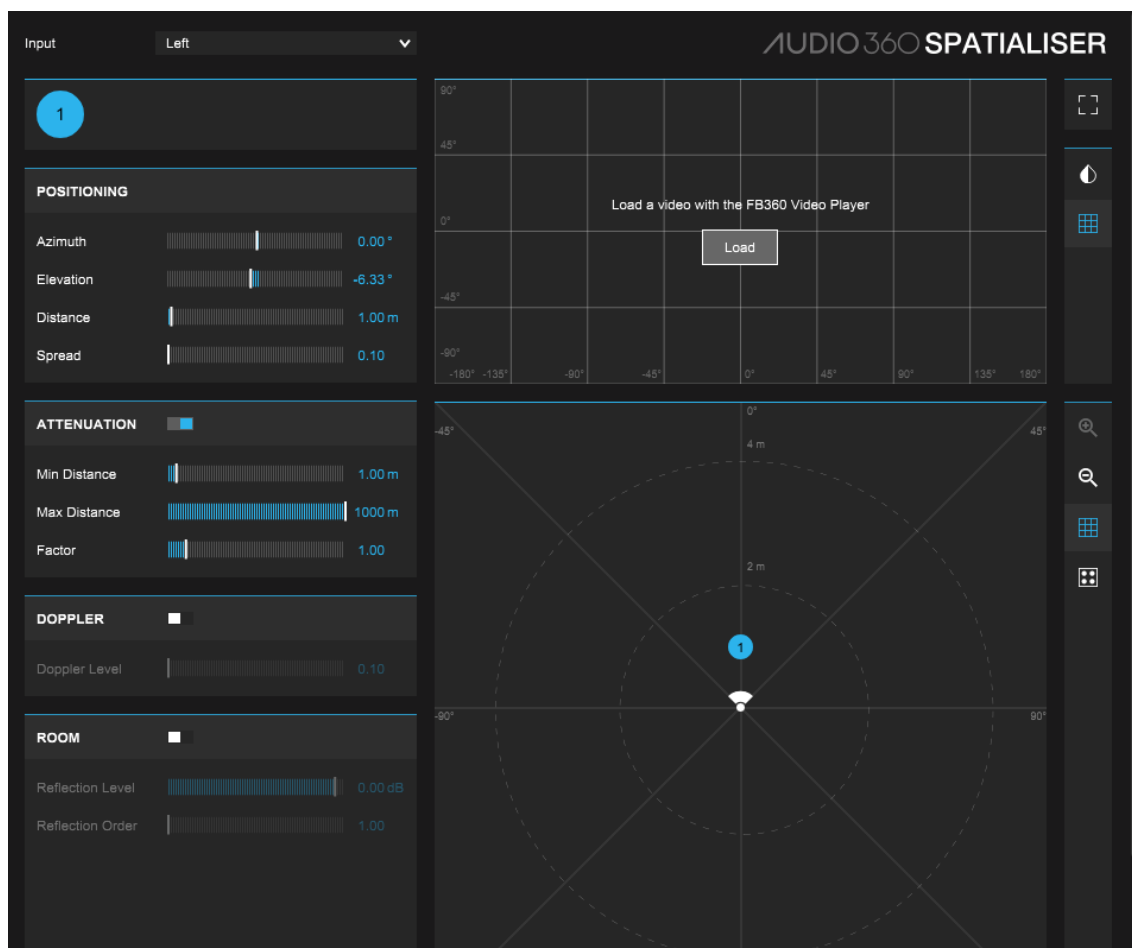


Figura 45 - Plugin Audio 360 Spatialiser, do pacote de plugins FB360.

O *Audio 360 Control*, é o plugin responsável pela manipulação de alguns parâmetros de sala, parâmetros a serem controlados em tempo real por um aplicativo externo através de metadados (movimentação da cabeça do espectador extraída do reprodutor do vídeo em VR) e uma codificação para *Binaural* com objetivo de monitoração.

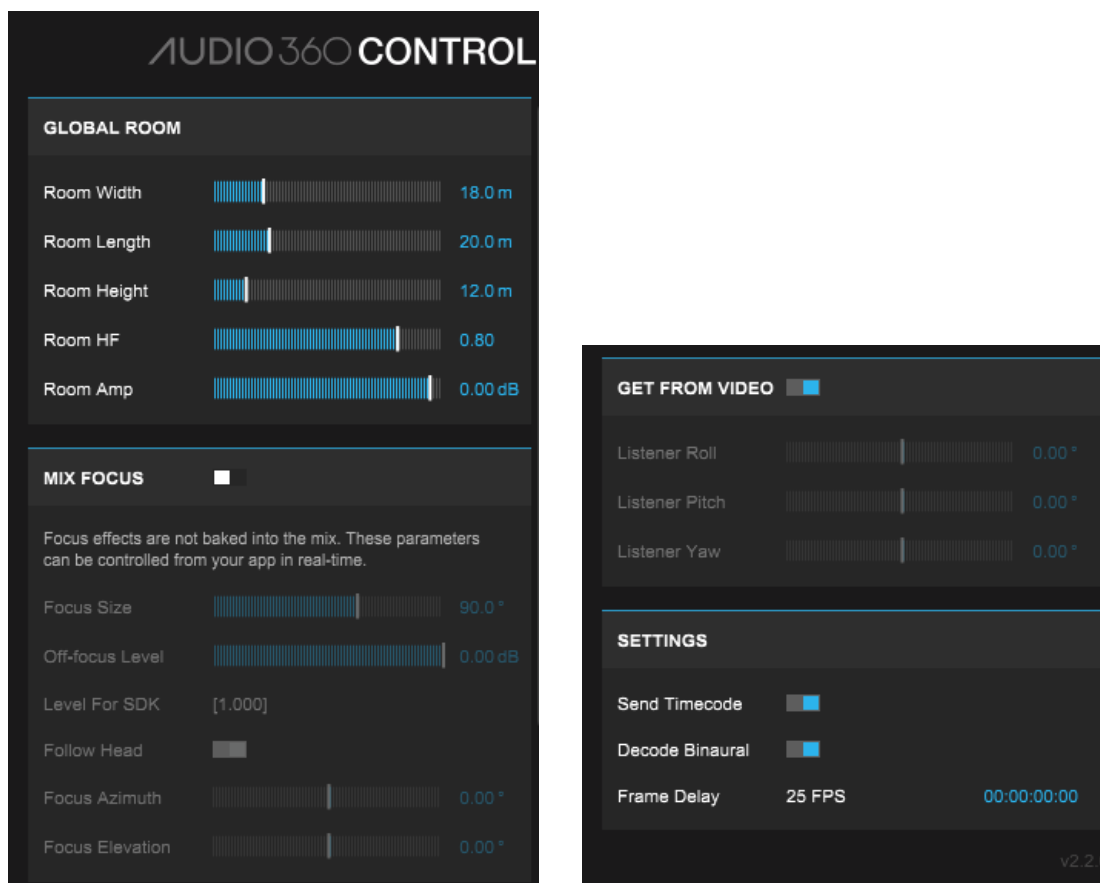


Figura 46 - Plugin Audio 360 Control, do pacote de plugins FB360.

O *Audio 360 Converter* possui um controle do eixo Y, permitindo que a mixagem já finalizada do campo sonoro tridimensional (ou o áudio extraído de uma gravação de um microfone tetraédrico) seja eventualmente realinhado ao vídeo através de uma rotação em tal eixo. Ele permite também que uma entrada em determinado formato seja convertida em *Binaural*, formato-B *Ambix* de primeira ordem ou de segunda ordem (*spatial workstation 8 channel*). Os plugins *Audio 360 Meter* e *Audio 360 Stereo Loudness* são utilizados principalmente no processo de monitoração de intensidade sonora (já em *LUFs*²³⁸, padrão mais recente adotado pela indústria) no processo de masterização.

²³⁸ Abreviação de *Loudness Units Relative to Full Scale*, unidade de medição adotada pela indústria a partir de 2013 como solução para padronização de áudio nas mídias de massa.

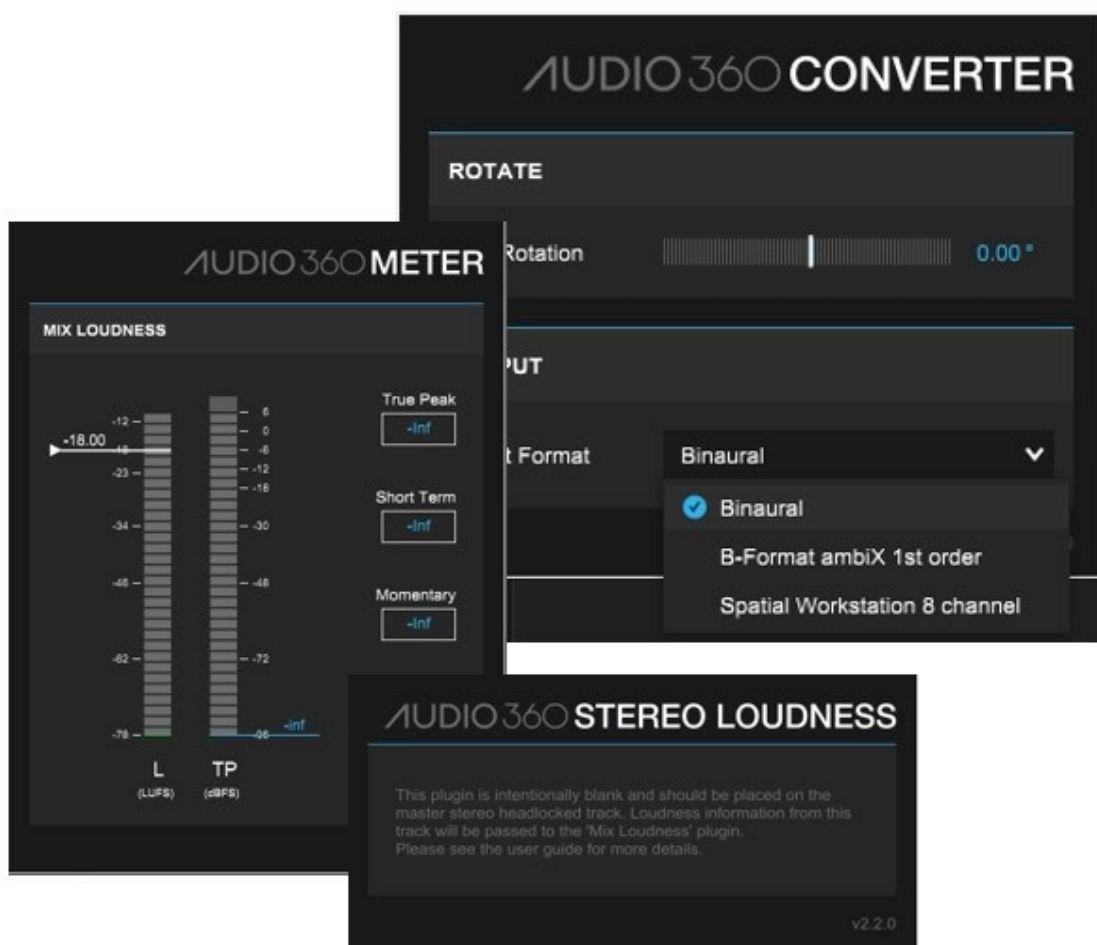


Figura 47 - Plugins Audio 360 Converter, Audio 360 Meter e Audio 360 Stereo Loudness, do pacote de plugins FB360.

o3acore

O conjunto de plugins *o3acore* oferece ferramentas gratuitas para a produção de material com saída em *Ambisonics* de ordens superiores (*HOA*). Os módulos codificadores, denominados *O3A Panner* se apresentam em seis formatos diferentes (*Standard*, *Classic*, *Hemisphere*, *Large*, *Two Channel* e *Eight Channel*) sendo a diferença entre eles somente no número de canais nos quais eles realizam o processamento de codificação do ângulo e elevação simultaneamente e na forma de visualização. A saída de todos eles é em *HOA* de terceira ordem (16 canais):

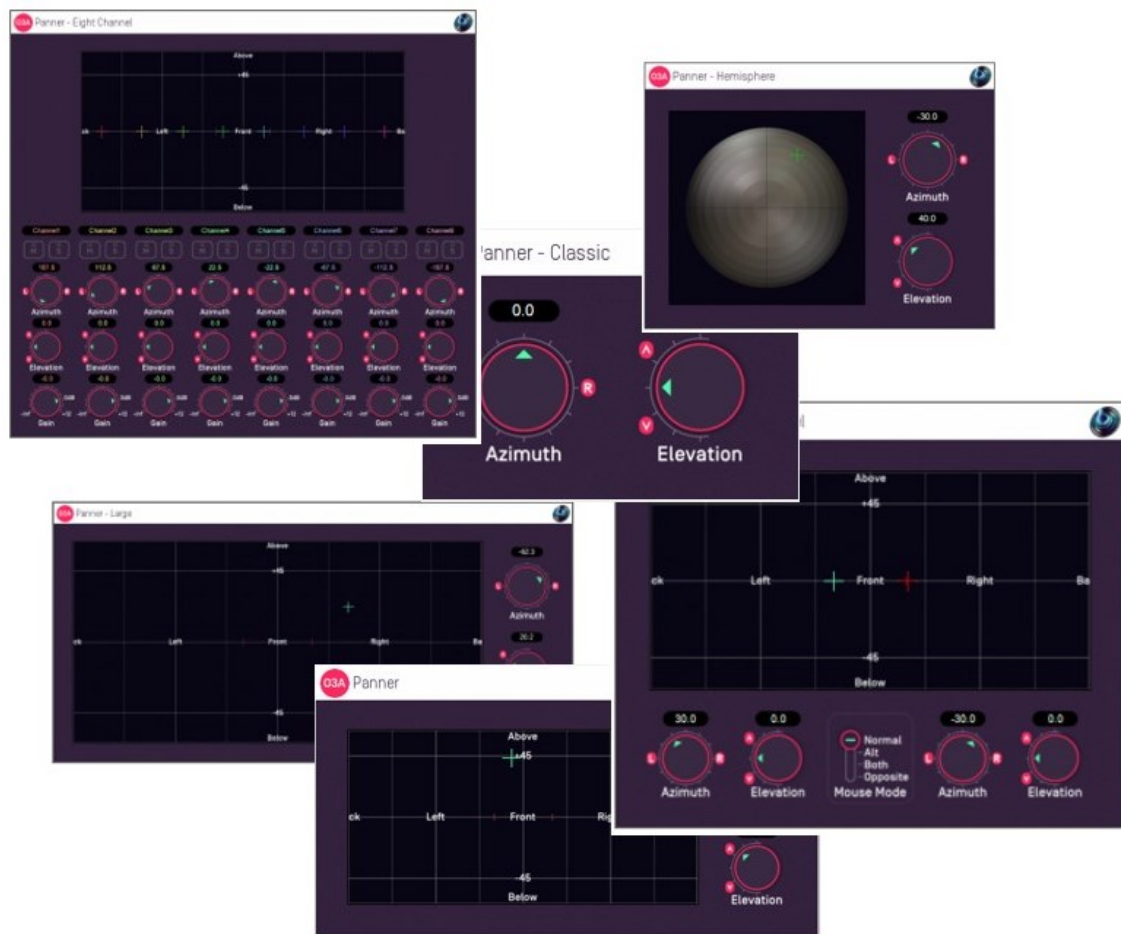


Figura 48 - Codificadores do pacote de plugins o3acore.

Os transformadores vão de um simples ganho à *delays*: o *Beamer* extrai som monofônico de uma direção específica do campo sonoro 3D; o *Virtual Microphone* simula um arranjo de microfones estéreo convencional posicionado dentro do campo sonoro 3D criado; o *Gain* oferece um simples ganho em todo o campo sonoro; o *Look* rotaciona todo o campo sonoro para uma direção específica, como se fosse o próprio movimento da cabeça do ouvinte; o *Rotation* permite a rotação de todo o campo sonoro em qualquer um dos três eixos; o *Spatial Delay* oferece um efeito de atraso em todo o campo sonoro e inclui os parâmetros de realimentação, reflexões, rotação e filtragem.

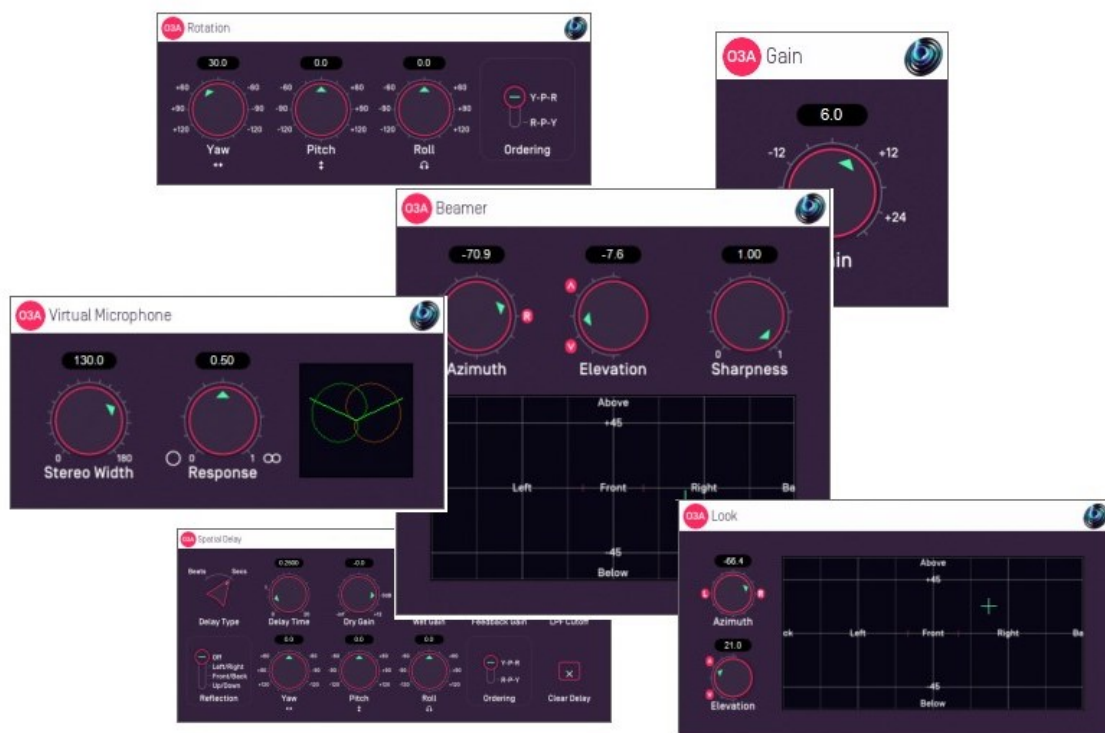


Figura 49 - Transformadores do pacote de plugins o3acore.

Os visualizadores apresentam de medidores simples de energia à modos inéditos de visualização do campo sonoro dependente da frequência. O *Meter* apresenta uma simples análise da energia sonora da mixagem em tempo real; o *Meter-Signal* apresenta um diagnóstico dos componentes do sinal como um todo distribuído nos harmônicos esféricos; o *Visualiser* apresenta uma representação visual em tempo real da mixagem no campo sonoro; o *Visualiser-Hemisphere* apresenta essa mesma representação visual em dois hemisférios ao invés de um plano; o *Colorizer* apresenta essa representação visual dependente da frequência; o *Flare* apresenta essa visualização usando uma análise ativa dependente da frequência e posição no campo sonoro.



Figura 50 - Visualizadores do pacote de plugins o3acore.

Os decodificadores desse pacote de plugins consistem em um *Decoder-Mono*, um *Decoder-Stereo*, um *Decoder-5.1 Basic*, um *Decoder-FuMa*, que converte uma mixagem em *HOA* em padrão *o3a* para o padrão *FuMa*, compatível com o clássico formato-B, e um *Injector-FuMa* que faz o oposto do plugin anterior. Vale observar aqui que esse pacote não oferece decodificadores para grandes arranjos de alto-falantes ou mesmo para sistemas *Binaurais*, o que faz com que os dois últimos plugins sejam extremamente importantes para que haja alguma compatibilidade com outros plugins, de outros pacotes, nessa etapa de decodificação.

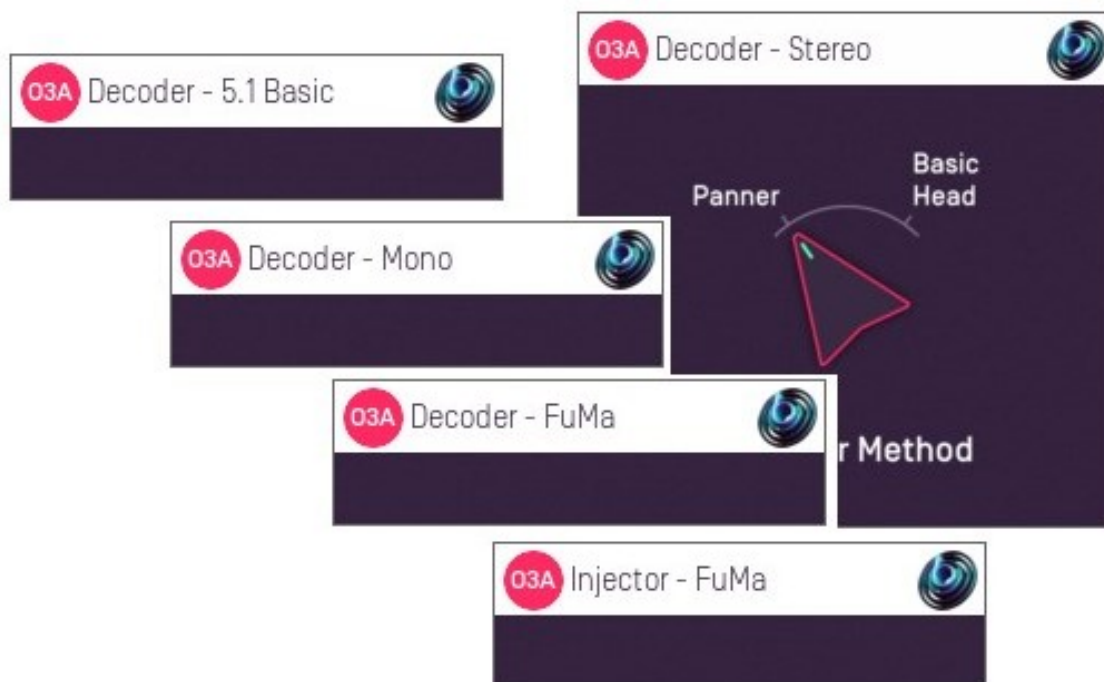


Figura 51 - Decodificadores do pacote de plugins o3acore.

Ambix

Provavelmente este seja o pacote de plugins com mais opções, inclusive referente à ordem do *Ambisonics* com que podemos trabalhar. Ele oferece o mesmo pacote de plugins tanto de primeira ordem, quanto de terceira, quinta e sétima ordens. Resultado da pesquisa de doutorado de Matthias Kronlachner²³⁹, se mostra extremamente versátil, sendo suportado tanto por qualquer sistema operacional (windows, Linux ou Mac OSX), 32 ou 64 bits, quanto por qualquer *DAW* que suporte *plugins VSTs*.

O *ambix-encoder* vem em cinco versões, com suporte de uma a oito entradas incluindo um parâmetro de movimentação automático, com controle de ângulo, elevação e velocidade do movimento. Oferece também um controle de separação entre as fontes sonoras e seu tamanho, visualizado através de um globo hemisférico.

²³⁹ <http://www.matthiaskronlachner.com/?p=2015>

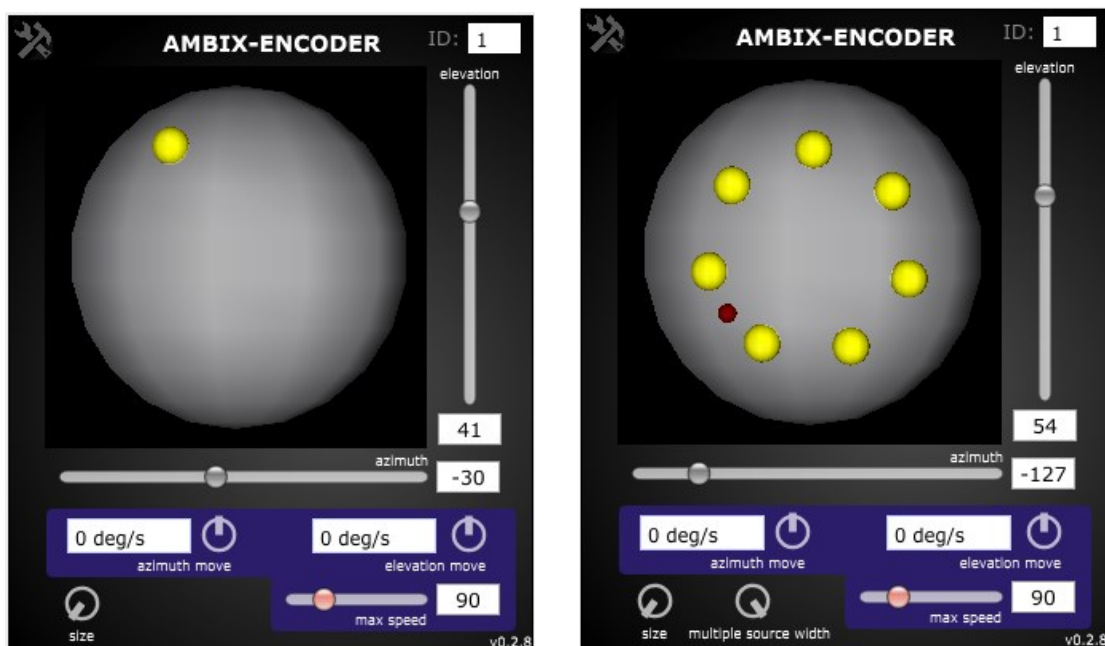


Figura 52 - Codificadores do pacote de plugins Ambix, desenvolvidos por Kronlachner.

Com relação aos transformadores, o autor oferece em seu pacote de plugins, três transformadores: o *warping*, que permite distorcer a imagem espacial para ou de certas direções na esfera, tal distorção da imagem espacial em direção ao equador pode compensar uma imagem espacial elevada durante a reprodução em um sistema irregular; o *directional loudness modifications*, que permite modificações de intensidade direcionais, possibilitando que determinada região da projeção do campo sonoro seja atenuada ou amplificada, importante principalmente quando se explora gravações surround para se enfatizar um instrumento por exemplo ou atenuar uma reverberação indesejada proveniente da parte de trás do microfone (tal manipulação é realizada através da definição de regiões circulares ou retangulares na superfície da esfera e aplicação de um fator de ganho); o *source widening*, que permite a ampliação do tamanho perceptível da fonte sonora através de um parâmetro de controle de largura / difusão, tal ampliação se dá através de um algoritmo que aplica uma rotação dependente da frequência, produzindo uma direção de chegada com dispersão de frequência, podendo ser usado também para sintetizar reflexões precoces difusas; o *rotator*, que permite uma rotação em todo o campo sonoro tridimensional em qualquer um dos três eixos; o *mirror*, que permite que componentes simétricos dos canais *Ambisonics* sejam ajustados; o *vmic*, que possui o mesmo princípio que o *directional loudness modifications*, mas não terá como saída um sinal em *Ambisonics* mas somente uma parte do campo sonoro selecionado, possui uma abordagem

semelhante ao de um microfone virtual, captando somente uma região do campo sonoro; o *maxre*, que realiza uma compensação com o intuito de minimizar os lóbulos laterais e otimizar a concentração de energia no centro do campo sonoro²⁴⁰; e o *converter*, que possibilita conversões entre os diversos formatos de *Ambisonics*.



Figura 53 - Transformadores do pacote de plugins Ambix, desenvolvidos por Kronlachner.

Os decodificadores para sistemas *Binaurais* e para arranjos de alto-falantes fornecidos por Kronlachner são baseados em uma biblioteca fornecida por ele mesmo, sendo necessário fazer o *download* dessa biblioteca e armazená-la no local correto para que os *plugins* a reconheçam.

²⁴⁰ Para mais informações sobre o processo ver: <http://www.aes.org/tmpFiles/elib/20190817/16554.pdf>

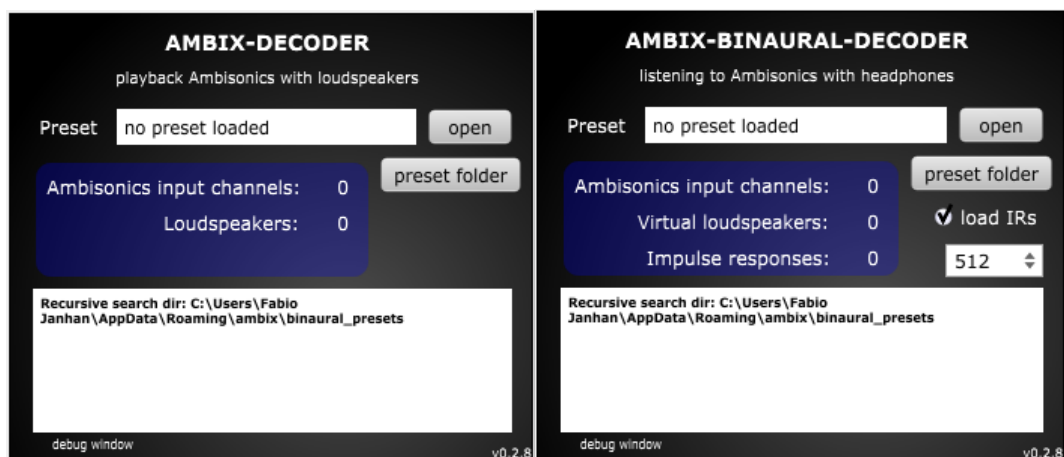


Figura 54 - Decodificadores do pacote de plugins Ambix, desenvolvidos por Kronlachner.

O autor oferece também alguns *plugins* multicanais, tais como *delay*, filtro, medidor de intensidade e processador de convolução, para o processamento e controle do material gerado em *Ambisonics* de primeira ordem (4 canais) ou ordens superiores, chegando a 64 canais, em situações em que diversas instâncias do mesmo plugin são necessárias simultaneamente.

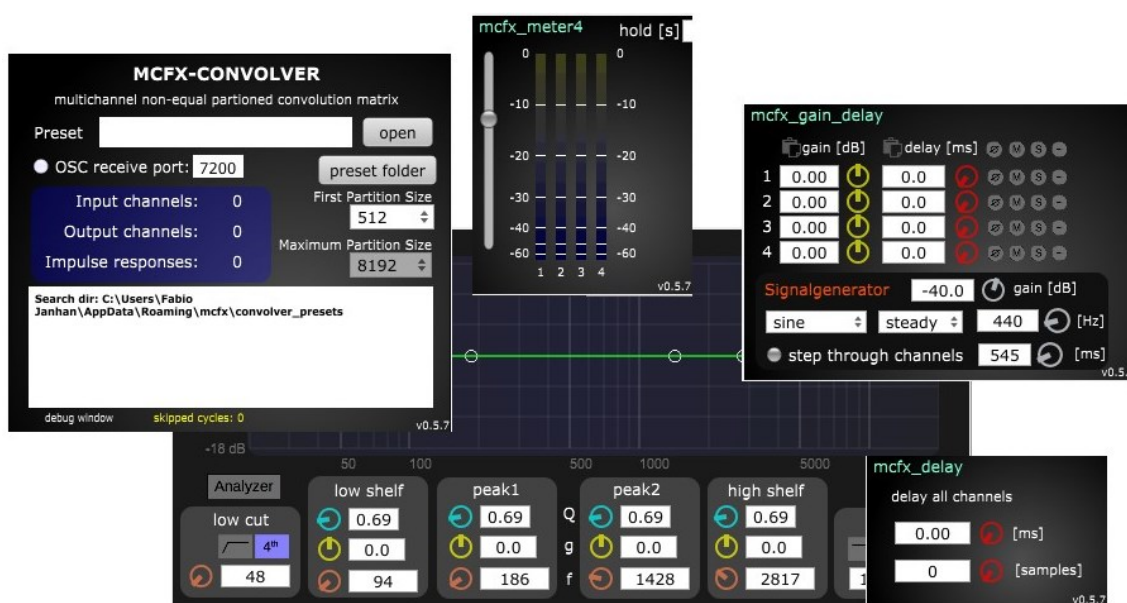


Figura 55 - Pacote de plugins multicanais para processamento de sinais em Ambisonics, desenvolvidos por Kronlachner, e oferecidos juntamente com o pacote Ambix.

Buscamos direcionar nossa abordagem para plugins de acesso gratuito e multiplataforma com o intuito de oferecer caminhos realmente acessíveis. No entanto, vale apontar brevemente alguns plugins que não se enquadram nesse recorte mas são importantes do ponto de vista da variedade e robustez, como por exemplo, o

Gaudio Works²⁴¹ para uso somente sistemas operacionais MacOS. Outros, como o dearvr²⁴² oferecem um pacote pago e somente um plugin gratuito para teste, com limitações em suas funcionalidades, como seu transformador de campo sonoro e conversor de formatos *Ambisonics*. Tais *plugins* não são abordados com profundidade aqui uma vez que não serão testados e explorados no desenvolvimento dos trabalhos (cap. 3), mas merecem um espaço por serem considerados desenvolvimentos de fabricantes renomados no mercado.



Figura 56 - Plugin para teste da dearvr.

²⁴¹ <https://en.gaudiolab.com/sol-vr360-sdk/>

²⁴² <https://www.dearvr.com/>