

**OTÁVIO COSTA FARIA**

**ÁUDIO ESPACIAL:**

**Aspectos psicoacústicos, desenvolvimento histórico e  
tópicos recentes**

**Trabalho de Conclusão de Curso**

**São Paulo**

**2020**



**OTÁVIO COSTA FARIA**

**ÁUDIO ESPACIAL:**

**Aspectos psicoacústicos, desenvolvimento histórico e  
tópicos recentes**

Trabalho de Conclusão de Curso apresentado ao  
Departamento de Música da Escola de Comunicações  
e Artes da Universidade de São Paulo para obtenção do  
título de \_\_\_\_\_

Orientador(a): Prof(a). Dr(a)\_\_\_\_\_

São Paulo

2020

## FICHA CATALOGRÁFICA

Faria, Otávio Costa

Áudio Espacial: Aspectos psicoacústicos, desenvolvimento histórico e tópicos recentes / Otávio Costa Faria; orientador, Sílvio Ferraz. -- São Paulo, 2020.

59 p.

Trabalho de Conclusão de Curso - Departamento de Música/Escola de Comunicações e Artes / Universidade de São Paulo.

Bibliografia

Versão corrigida

1. Áudio Espacial 2. Binaural 3. *Ambisonics*

I. Ferraz, Sílvio II. Título.

CDD 21.ed. - 780

## RESUMO

FARIA, Otávio Costa. *Áudio Espacial: Aspectos Psicoacústicos, Desenvolvimento Histórico, e Tópicos Recentes*. 2020, 54 p. Trabalho de Conclusão de Curso (Graduação em: Música) – Departamento de Música, Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2020.

**Resumo:** As aplicações de áudio espacial e imersivo tem crescido nas últimas décadas com o avanço de formatos e processamentos digitais de gravação e reprodução, ao lado de novas propostas para se abordar os problemas fundamentais da representação e percepção de espaço pela escuta. No entanto, o conhecimento comum compartilhado no áudio não aborda questões essenciais a estes aspectos de gravação e reprodução, necessitando assim de pesquisas específicas nestas áreas para um entendimento dos fundamentos desta área. Este trabalho tem o objetivo de servir como um primeiro contato a este campo, procurando apresentar estes fundamentos, visando ser uma introdução através da apresentação de aspectos psicoacústicos, históricos, e de implementações recentes quanto às tecnologias usadas na gravação e reprodução de representações espaciais através do áudio. Usando como base textos gerais de língua inglesa que abordem os assuntos comuns no campo do áudio espacial, como os trabalhos de Francis Rumsey, Roginska Agnieszka e Paul Geluso, o presente trabalho procura, através de uma revisão bibliográfica em seus respectivos campos, mapear uma bibliografia geral que sirva como suporte para uma introdução a esta área de estudo, buscando, quando possível, a contraposição com uma documentação original, relevante a cada campo respectivamente abordado, e a literatura de pesquisas recentes que possa complementá-la. Resultados apontam para paradigmas relevantes ao áudio espacial, indicando tendências significativas às diferentes implementações atuais.

**Palavras-chave:** Áudio. Áudio Imersivo. Gravação. Binaural. Ambisonics.

## ABSTRACT

**Abstract:** The applications of immersive and spatial audio have grown in the last decades following the growth of digital formats and digital processing of audio recording and reproduction, as well as new methods of approaching fundamental problems of perception and representation of space in listening. However, despite this growth, the common shared knowledge in audio does not encompass essential topics of this form of recording and reproduction, thus requiring specific research within these areas in order to obtain a understanding of the fundamentals of this field. This work's objective is to serve as a first contact to this field of study, with the intention of presenting these fundamentals through a presentation of aspects such as psychoacoustics, history, and recent implementations of the technologies used in the recording and reproduction of spatial representations in audio. Referencing English language texts by Francis Rumsey, Roginska Agnieszka and Paul Geluso, the present work seeks to, by means of a bibliographical revision of its respective fields of study, map a general bibliography that may serve as a support to an introduction to this field of study, trying to, whenever possible, oppose original documentation relevant to each respective approached field to the recent research literature that may complement it. Results point towards relevant paradigms of spatial audio, indicating significant tendencies of several current implementations.

**Key-words:** Audio. Immersive Audio. Recording. Binaural. Ambisonics.

# SUMÁRIO

Listas de Figuras	p. 8
Introdução	p. 9
Capítulo 1: Fundamentos da Acústica Espacial	p. 12
1.1 Codificação espacial e função de transferência	p. 14
1.2 Condicionamentos comuns da escuta espacial	p. 15
1.3 Escuta no plano horizontal	p. 18
1.4 Escuta no plano vertical	p. 21
1.5 Escuta em distância	p. 12
1.6 Aspectos multimodais da escuta espacial	p. 23
Capítulo 2: Histórico de Sistemas de Áudio Espacial	p. 26
2.1 Theatrophone	p. 27
2.2 Alan Blumlein	p. 30
2.3 Pesquisas estereofônicas nos Laboratórios Bell	p. 31
2.4 <i>Fantasia</i> (1940) e o sistema Fantasound	p. 34
2.5 Sistemas quadrifônicos	p. 37
2.6 Desenvolvimento do áudio ambissônico	p. 39
2.7 Propostas de áudio multicanal <i>surround</i>	p. 42
Capítulo 3: Tópicos Recentes em Sistemas no Áudio Espacial	p. 45
3.1 Implementações de áudio binaural	p. 46
3.2 Implementações de áudio ambissônico	p. 49
Conclusão	p. 54
Referências Bibliográficas	p. 55

## LISTA DE FIGURAS

- Fig. 1 – Distribuição da pressão sonora para diferentes ressonâncias naturais de um modelo do pavilhão auricular. As linhas tracejadas indicam planos nodais de pressão sonora p. 15
- Fig. 2 – Localização de eventos sonoros ao longo do plano vertical em função de sinais de banda estreita em diferentes frequências, indicando distância em função da frequência central de um sinal p. 17
- Fig. 3 – Localização entre a distância de uma fonte sonora e a percepção do evento sonoro para diferentes tipos de fala, emitidos à frente do ouvinte p. 18
- Fig. 4 – Diferenças de amplitude interaural, captadas por microfones dispostos em uma esfera, em função do ângulo de incidência lateral e frequência do sinal p. 19
- Fig. 5 – Diferenças de tempo de bandas de frequência, captadas por microfones dispostos em uma esfera, em função do ângulo de incidência lateral e frequência do sinal p. 20
- Fig. 6 – Ilustração da percepção oposta de aspectos interaurais de codificação da escuta para sinais à frente e atrás de um ouvinte quando este rotaciona em relação à fonte sonora, permitindo a resolução de ambiguidades de localização entre os planos frontal e traseiro p. 25
- Fig. 7 – Diagrama explicativo do sistema de reconstrução de frente de onda inicialmente proposto por Fletcher. p. 33

## INTRODUÇÃO

Desde o seu início, a gravação e reprodução de áudio estiveram diretamente ligadas ao conceito de espaço acústico, tanto no processo de arquivamento de uma performance ou de um local captado, obtidos no processo de gravação, como também no processo de reconstrução ou síntese de sinais fixados em uma mídia através de alto-falantes. De fato, a própria escuta humana, inherentemente espacial em sua natureza, atribui a toda percepção auditiva, acústica ou não, qualidades espaciais que a qualificam dentro da nossa percepção de espaço. Aspectos como localização, distância, tamanho, fonte, dentre outros, são qualidades da nossa percepção que são inconscientemente processados pelo sistema auditivo e atribuídos a toda experiência auditiva que temos (BLAUERT, 1983, p. 3-4).

Sendo de tal modo inerente à nossa percepção sonora, estes aspectos espaciais da escuta mostram-se como um valioso campo de estudo para se obter um entendimento pleno de como o sistema auditivo, através de diversos mecanismos e tendências de comportamento, altera e condiciona a nossa percepção.

No entanto, em comparação com outros campos de estudo relacionados à escuta e à percepção auditiva, a ênfase de pesquisas e interesses ao redor de investigações quanto ao funcionamento e comportamento do sistema auditivo em relação ao espaço parece não ter sido a mesma que dada a outros aspectos da escuta. Por ser em parte um modo de percepção que não é facilmente analisado através de aspectos fisiológicos apenas, a percepção espacial da escuta foi, durante um período considerável de tempo, entendida somente de forma limitada. Proposições de modelos para o entendimento de aspectos de segunda ordem, psicoacústicos por natureza, são recentes se comparados com as primeiras propostas de modelos para a escuta espacial baseados exclusivamente em aspectos fisiológicos e acústicos. A dificuldade da análise estatística das diversas variáveis em jogo quando se trabalha com aspectos psicoacústicos em um grande número de indivíduos durante um experimento é uma possível razão pela qual as proposições de modelos para a escuta se mantiveram exclusivamente limitados a parâmetros físicos mais facilmente mensuráveis em situações experimentais (BLAUERT, 1983, p. 10-12).

Estas dificuldades, entretanto, não impediram que tecnologias fossem desenvolvidas com a experiência espacial da escuta como objetivo ao longo das décadas. Tendo tido maior ou menor relevância ao longo do século 20, propostas para sistemas espaciais que, implementassem possibilidades de reprodução de qualidades acústicas espaciais, mesmo que estas qualidades não fossem originais aos espaços acústicos de gravação ou fidedignas a parâmetros de um espaço real, ocorreram ao longo da existência do áudio. Mesmo assim, a proximidade que normalmente se tem com os conceitos relacionados a estas tecnologias e com os requerimentos básicos de conhecimento sobre o funcionamento e modos de se abordar estas ferramentas é comumente menor do que com outros aspectos do áudio.

Esta aparente disparidade entre os campos de conhecimento comum relacionados a aspectos do áudio e os conceitos exclusivos ao áudio espacial se justifica quando vista através da particularidade deste campo ao longo do desenvolvimento do áudio no o século 20: em grande parte, durante a sua existência, o áudio espacial foi, por um lado, algo exclusivo de experimentos laboratoriais que procuravam entender o comportamento da escuta inserida em espaços acústicos, e, por outro, a partir dos anos 50, uma área que se inseriu no campo do entretenimento através de novas propostas do cinema, que procurava se distinguir da experiência da televisão (REEVES, 1999, p. 97). Durante quase a totalidade da existência do áudio, estes dois campos foram em grande parte as aplicações exclusivas de sistemas multicanal de áudio, tendo o mercado fonográfico se mantido em reproduções monofônicas e estereofônicas. Mesmo quando este procurou a introdução de tecnologias multicanal, como na década de 70 através de sistemas quadrifônicos, o alcance destes sistemas foi tímido e breve, não tendo se mantido presente ou alterado a indústria fonográfica de forma considerável como inicialmente se esperava (TORICK, 1998, p. 31).

No entanto, aplicações recentes, impulsionadas pela facilidade de processamentos digitais, tal como por implementações virtuais de diversas formas que permitem a modelagem e derivação de espaços acústicos, mostram o presente como um momento oportuno de desenvolvimento e interesse em sistemas de áudio imersivo. Em parte, livre de aspectos limitantes de sistemas passados, como custo e instalação de equipamentos, estas novas propostas parecem indicar um renovado interesse no desenvolvimento e no acesso a experiências imersivas através da reprodução eletroacústica.

Os processos tradicionais de produção usados para a representação de qualidades espaciais em sistemas de áudio comumente são capazes de reproduzir uma impressão de aspectos de espaços, sem a capacidade de recriar a ilusão de localização bem definida e externalização esperada da escuta em situações não mediadas por meios tecnológicos. Tecnologias específicas, combinadas com o respectivo conhecimento do sistema auditivo, parecem ser necessárias para que se possa recriar uma percepção que seja ilusória, e não sugestiva, da sensação de espaço, tal qual é visto como o objetivo para estas implementações (BLAUERT, 1983, p. 4-5).

Um estudo prévio se apresenta necessário para que se tenha um entendimento pleno dos fundamentos aplicados a estes tecnologias, tais como: quais as implementações básicas com as quais se pode trabalhar, como elas diferem entre si, quais os aspectos psicoacústicos que estão em jogo quando se fazendo uso de diferentes sistemas de representação e reconstrução espacial para diferentes situações, como estas implementações diferem entre si em qualidade de representação de aspectos espaciais para a percepção, e quais os aspectos que estão em desenvolvimento nos limites recentes de pesquisa de cada implementação.

Ao mesmo tempo que estes tópicos poderiam ser abordados em trabalhos próprios, o escopo do presente trabalho é o de uma introdução a três diferentes aspectos dentro do campo geral de áudio espacial, sendo eles: o funcionamento psicoacústico da escuta em relação à percepção espacial, os momentos históricos de desenvolvimento de tecnologias que fizeram uso da percepção espacial de forma particular, e alguns aspectos recentes no que diz respeito às principais tecnologias em uso hoje para reprodução espacial no áudio.

Esta abordagem, relativamente limitada se comparada a outros trabalhos particulares a estes aspectos de tecnologias espaciais, se mostrou válida a partir da revisão

bibliográfica inicial para o desenvolvimento desta pesquisa. Notou-se, quando pesquisando a bibliografia de áudio em língua portuguesa, uma ausência de trabalhos gerais sobre o tópico que funcionassem como um meio de introdução a este campo que, como notado anteriormente, se apresenta como um desafio para grande parte dos interessados dado que estas tecnologias estão em relativa distância do corpo de conhecimento comum compartilhado no campo de áudio.

Trabalhos em língua inglesa, como os trabalhos de Rumsey (RUMSEY, 2001) e Geluso e Roginska (GELUSO; ROGINSKA, 2018) mostram-se como interessantes fontes de informação, servindo de guias iniciais para um aprofundamento em tópicos particulares de interesse no campo. Abordando um conjunto de aspectos gerais relacionados a estas tecnologias e fornecendo uma visão abrangente de diferentes formas de se abordar problemas relacionados à gravação e reprodução espacial no áudio, estes trabalhos foram usados como referência à presente pesquisa.

Mesmo nestes trabalhos, no entanto, notou-se que grande parte dos tópicos relacionados a aspectos inerentes aos fundamentos deste campo são apresentados de forma relativamente superficial, tendo sido necessário procurar as publicações originais de autores para se ter um entendimento pleno das intenções e funcionamentos de implementações particulares do áudio espacial. Dado o escopo genérico do presente estudo, sem ter o devido espaço para uma apresentação minuciosa de cada uma das propostas abordadas, procurou-se, sempre quando possível, apontar para as publicações primeiras de cada uma das implementações ou aspectos tratados ao longo do texto.

A bibliografia usada ao longo desta pesquisa procurou fazer uso de fontes originais de textos, artigos e patentes, com o intuito tanto de indicar direções para estudos seguintes que procurem se basear em aplicações referidas neste trabalho, como também com a intenção de evidenciar aspectos que não tenham sido tratados em trabalhos de cunho geral sobre o campo de áudio espacial. Trabalhos não recentes foram citados como reforço de tendências ou evidências de paradigmas de um período em particular. Procurou-se, de resto, referências recentes quanto aos sistemas, tecnologias, e seus aspectos atuais.

## CAPÍTULO 1

### FUNDAMENTOS DA ESCUTA ESPACIAL

Dificilmente estudado em introduções sobre o sistema auditivo, a escuta espacial é uma área de pesquisa que apresenta conceitos bastante particulares. O funcionamento da escuta em relação às funções de atribuição de localização e qualidades espaciais a estímulos externos apresenta diversos comportamentos que não são óbvios, mesmo se comparados com outros modos de funcionamento da audição.

Esta dificuldade de estudo é resultante de uma divisão particular do sistema auditivo em dois segmentos relativamente independentes. Classificados respectivamente como regiões de processamento auditivo periférico e central, a primeira destas tem a função de um processamento inicial de estímulos acústicos, tendo uma importância especial as deformações de sinais resultantes de aspectos físicos característicos de cada indivíduo, que possibilitam a codificação de aspectos espaciais aos sinais antes do estágio de transdução ao ouvido interno (BEGAULT; COOPER; WENZEL, 2018, p. 5-6). A partir deste estágio de transdução, o processamento auditivo central compõe os segmentos neural e cognitivo, cujos funcionamentos resultam na classificação dos estímulos derivados dos sinais acústicos obtidos pelo sistema auditivo periférico em qualidades específicas da percepção humana (BEGAULT; COOPER; WENZEL, 2018, p. 8-9).

O entendimento pleno desta cadeia de processamentos, resultando necessariamente em aspectos neurais particulares e com relativa divergência em casos laboratoriais, se apresenta como um desafio se comparado com estudos de qualidades que se limitam ao sistema auditivo periférico. Estas qualidades, mais facilmente mensuráveis em testes, por terem uma correlação mais direta com um funcionamento fisiológico do ouvido, não apresentam, pelo menos em um estágio superficial de análise, as dificuldades cognitivas que a atribuição interna de qualidade à estímulos mostra ao estudo da escuta espacial.

Soma-se a isto a relativa divergência que diversos resultados experimentais demonstram quando pesquisando sobre um mesmo comportamento da escuta espacial. Apesar de mostrarem tendências gerais de comportamento, a distinção entre qualidades coletivas e individuais, principalmente levando em conta que a escuta espacial é, em grande parte, definida por características particulares, torna-se difícil, necessitando um grande conjunto de experimentos similares para se obter uma conclusão estatística quanto a certos modos de funcionamento da escuta (BLAUERT, 1983, p. 10-12). Mesmo em estudos que procuram avaliar o grau de correlação entre fontes sonoras no espaço e a imagem auditiva formada pelas mesmas no espaço auditivo, nota-se o quão díspares, e até mesmo opostas, estas localizações podem ser (BLAUERT, 1983, p. 4).

É possível notar, a partir destas conclusões, que as atribuições resultantes desta cadeia de análise por parte do sistema auditivo são relativamente inexatas. Por depender de qualidades físicas e fisiológicas que são particulares a cada indivíduo, o estudo da escuta

espacial procura definir tendências de comportamento que sejam relativamente independentes de condicionamento por parte de qualidades individuais.

É necessário notar que, mesmo em estudos consideravelmente controlados em seu espaço acústico e apresentação, pressupõe-se a investigação de um aspecto individual da audição, necessitando, para que este processo idealmente isole um aspecto auditivo, muitas vezes um ambiente artificial de escuta laboratorial. Diversos destes estudos usam de componentes senoidais e outros sinais provenientes de síntese que, apesar de passíveis de derivação a partir de sinais complexos de uma escuta natural, assim como ambientes controlados tal como salas anecóicas, dificilmente são encontrados em uma situação real de audição. Por isto, a extrapolação de seus resultados em comportamentos gerais da escuta necessita repetidos experimentos similares, e esta extrapolação é tida como linha geral de comportamento, raramente uma certeza da função da escuta.

Estas diversas dificuldades, inerentes à investigação de um objeto de estudo como o sistema auditivo, são possíveis razões da complexidade para se abordar este tema. No entanto, mesmo que relativamente incertos, os comportamentos que se pode obter através destas investigações elucidam inúmeros aspectos que não são óbvios simplesmente através de comparação com outros comportamentos da escuta que não sejam espaciais. Inúmeros modos de funcionamento não lineares, dependentes de uma dimensão temporal, ou particulares a aspectos individuais podem ser notados, todos tendo um peso considerável no condicionamento da nossa percepção. Talvez ainda mais importante, especialmente para o desenvolvimento de sistemas imersivos, é o entendimento de aspectos multimodais da percepção espacial através de parâmetros adicionais do sistema auditivo, como a visão e a propriocepção, como fatores essenciais para a formação de imagens auditivas do espaço.

Com o intuito de expor um resumo geral destes aspectos, o presente capítulo procura abordar uma visão geral de qualidades da escuta espacial presentes na bibliografia sobre este campo. O trabalho usado como referência para este capítulo é o trabalho de Jens Blauert (BLAUERT, 1983). Apesar de não ser tão recente como parte da bibliografia presente neste trabalho, o livro de Blauert ainda é citado com uma das principais referências gerais no campo da psicoacústica espacial. Os fundamentos da pesquisa Blauert ainda são válidos, tendo sido, em trabalhos desde então, aprofundados, especialmente através das tecnologias digitais que se tornaram mais acessíveis desde sua primeira edição (BLAUERT, 1983, p. 291). O presente trabalho não procura se aprofundar nos funcionamentos cognitivos do sistema auditivo central, mas somente entender os comportamentos resultantes do processamento deste; no entanto, um trabalho que pode ser abordado para tal é Wenzel et al. (BEGAULT; COOPER; WENZEL, 2018).

Este capítulo não procura comparar os diversos experimentos citados por Blauert e outros para a obtenção dos resultados psicoacústicos apresentados na bibliografia. A apresentação dos diferentes modos de funcionamento do sistema auditivo é apresentada tal qual os resultados abordados por Blauert, de forma a orientar um uso mais consciente dos diversos aspectos da escuta quando tratando de um sistema de áudio espacial ou imersivo. O presente autor acredita que o conhecimento destes resultados, mesmo que não abordados especificamente pelos paradigmas intrínsecos de uma ou outra implementação de sistemas de gravação ou reprodução, são valiosos para se ter uma compreensão plena de quais aspectos se

pode trabalhar, quais as suas importâncias para a formação e condicionamento da percepção espacial, e de que modo eles interagem entre si.

## **1.1 CODIFICAÇÃO ESPACIAL E FUNÇÃO DE TRANSFERÊNCIA**

Quando apresentado em uma situação real de escuta, um sinal é alterado e distorcido ao longo de toda a sua cadeia de propagação, desde a sua emissão até o processo de transdução pelo ouvido interno. Alterações de fase e amplitude ao longo de toda a sua banda de frequência ocorrem por toda a extensão da linha de propagação. Estas alterações são resultantes de diversos aspectos físicos dos meios de propagação, que ocorrem através de fenômenos acústicos particulares à linha de transmissão que define o espaço entre o emissor e o receptor de um sinal.

Um conjunto de alterações similares, resultantes dos mesmos tipos de aspectos que condicionam as qualidades de sinais na linha de transmissão até o ouvido interno, ocorrem quando estes sinais incidem sobre um indivíduo. As qualidades físicas de um objeto, tais como particularidades de seus materiais, a geometria de sua superfície, e suas relativas dimensões em comparação com as extensões de sinais acústicos, são essenciais para se definir como as qualidades de um sinal serão modificadas se aquele interferir no processo de transmissão deste (BLAUERT, 1983, p. 63).

A função através da qual se pode obter as alterações e deformações pelas quais um sinal é alterado é definida como função de transferência (BLAUERT, 1983, p. 78). Aplicável a qualquer meio de propagação de um sinal, uma função de transferência pode ser derivada a partir das qualidades particulares de qualquer meio ou objeto, podendo sintetizar o complexo comportamento físico ou eletroacústico de um elemento da cadeia de um sinal a uma função pela qual se pode derivar a sua resposta a um sinal incidente neste sistema.

Dada a generalidade deste modo de análise, um grande número de funções de transferência poderia ser derivado, como de espaços acústicos, objetos físicos, elementos eletroacústicos, para se obter as particularidades de resposta em diferentes situações de transmissão. No entanto, quando aplicada ao estudo da escuta espacial, o uso de funções de transferência normalmente é feito com o intuito de se derivar o comportamento resultante das qualidades físicas que participam no processo de codificação e alteração de um sinal incidente. Particular para cada pessoa, estas qualidades, desde a geometria e disposição geral de membros até as ressonâncias inerentes à topografia do ouvido externo, contêm os aspectos fundamentais que codificam sinais acústicos à audição.

Essencial para o entendimento deste tipo de codificação é a conclusão de que a contribuição de diferentes aspectos físicos do ouvido externo é moldada pela frequência, distância, e ângulo de incidência de um sinal acústico antes de sua transdução pelo ouvido interno. Segundo Blauert (BLAUERT, 1983, p. 63):

Acoustically the pinna functions as a linear filter whose transfer function depends on the direction and distance of the sound source. By distorting incident sound signals linearly, and differently depending on their direction and distance, the pinna codes spatial attributes of the sound field into temporal and spectral attributes.

Este conjunto complexo de comportamentos acústicos que definem a função da codificação espacial pelo ouvido externo se dispõe de forma irregular ao longo do ouvido externo, de modo que a percepção de diferentes aspectos espaciais usa de diferentes parâmetros deste conjunto. Em termos de frequências, diferentes ressonâncias do pavilhão auricular se encontram em diferentes pontos deste, criando um conjunto de contribuições dependentes do ângulo de incidência e distância da fonte sonora para perfis específicos de amplificação (Fig. 1). Assim, é notável que os parâmetros codificados pelo ouvido externo e interpretados pelo sistema auditivo quando analisando a distância de eventos auditivos não são os mesmos que aqueles analisados na lateralização da escuta horizontal, por exemplo.

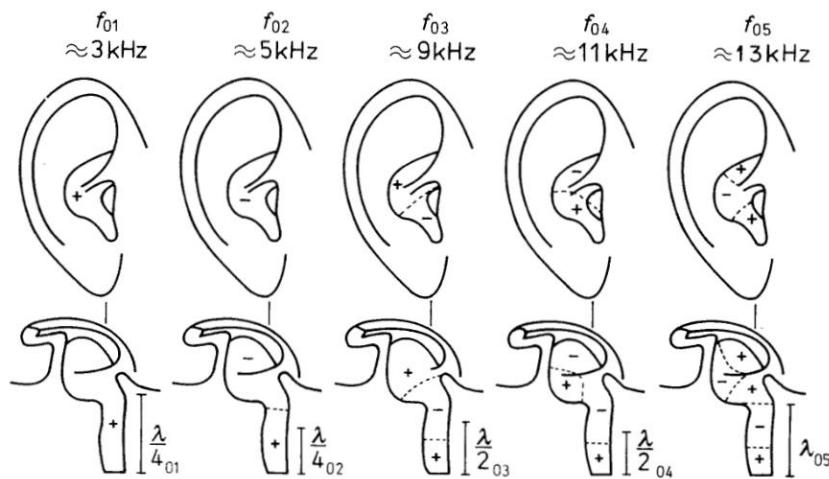


Fig. 1 – Distribuição da pressão sonora para diferentes ressonâncias naturais de um modelo do pavilhão auricular. As linhas tracejadas indicam planos nodais de pressão sonora (FONTE: BLAUERT, 1983, p. 67).

É necessário reafirmar que, apesar das funções de transferência serem capazes de resumir o resultado das particularidades físicas de diversos aspectos essenciais para a codificação de sinais acústicos incidentes durante a escuta, estas não compõem a totalidade dos estímulos que são usados pelo sistema cognitivo para a atribuição de localização à percepção espacial (BEGAULT; COOPER; WENZEL, 2018, p. 22). Apesar de constituir todos os elementos espaciais relevantes à escuta, funções de transferência não contêm o condicionamento de estímulos visuais ou proprioceptivos, fundamentais para a correção de ambiguidades resultantes das próprias qualidades acústicas captadas por sinais pelo ouvido externo em situações particulares de audição. Estas ambiguidades resultam em anomalias no sistema auditivo, gerando percepções múltiplas ou errôneas consequentes de estímulos insuficientes ou contraditórios para a definição clara de distância, direção e dimensão da fonte sonora. O sistema auditivo depende necessariamente do uso de outros modos de percepção para a resolução destas anomalias a uma definição clara de espacialidade da escuta (BLAUERT, 1983, p. 43-44).

## 1.2 CONDICIONAMENTOS COMUNS DA ESCUTA ESPACIAL

São notáveis os aspectos comuns que condicionam o comportamento geral dos diferentes modos como a escuta atribui qualidades espaciais em diferentes situações, exposta a diferentes tipos de condições e sinais. Em um primeiro momento, é possível que se tenha uma visão simplificada das diferentes funções de atribuição de aspectos espaciais da escuta, observando estas funções como sistemas contínuos e constantes, sem divergências drásticas em suas atribuições para além de parâmetros próprios de um sinal incidente como fase e amplitude.

No entanto, os resultados de pesquisas apontam que há uma enorme dependência no comportamento da escuta a diversos parâmetros além de qualidades fundamentais de sinais incidentes. A escuta é fortemente condicionada por estes parâmetros, alterando de forma considerável a percepção de fontes sonoras e as qualidades a elas atribuídas pelo sistema auditivo. Em casos excepcionais, alguns destes parâmetros em particular são as únicas variáveis às quais estão em função o resultado da audição.

A necessidade de medição e comparação destes resultados requer a definição clara de aspectos psicoacústicos da escuta que possam ser obtidos em situações experimentais. Apesar de a percepção espacial conter diversos aspectos que possam ser de utilidade para o entendimento pleno da escuta em relação ao espaço, a presente pesquisa se atentara a dois parâmetros fundamentais ao estudo da escuta espacial: a localização pela escuta em suas três coordenadas de elevação, lateralização, e distância, como também a alteração mínima possível de ser percebida pela escuta em relação a estas três dimensões de coordenadas pelo ouvinte (BLAUERT, 1983, p. 43-44).<sup>1</sup>

Estes dois aspectos, a tridimensionalidade da localização de uma fonte no espaço e os desvios mínimos que esta fonte apresenta em suas coordenadas, são as duas medições que são usadas por Blauert para a análise fundamental do comportamento da escuta espacial. Apesar da interação entre a percepção auditiva e o espaço com o qual este se relaciona serem mais abrangentes que somente a definição de localização de fontes sonoras em relação a um ponto de referência da escuta, este processo se apresenta como o fundamento para o qual outros modos de interação se baseiam, e, portanto, se apresenta como fundamental para o início de um estudo neste campo. As qualidades que estes dois aspectos contêm dentro de seus limites serão exploradas nas subseções seguintes. No entanto, seus funcionamentos são condicionados por mecanismos que são independentes aos sinais individuais, e, portanto, são de interesse comum à escuta espacial como um todo.

Os modos de localização da escuta apresentam condicionamento pelo domínio do tempo e frequência. Atributos relacionados ao tempo existem tanto em escalas de curto e longo prazo, enquanto atributos relacionados à frequência são decorrentes da dependência freqüencial que aspectos de codificação da escuta espacial tem em termos de fenômenos acústicos do ouvido externo como ressonâncias e difrações.

Os tipos de condicionamentos no domínio do tempo se apresentam nas formas de exposição ao estímulo de sinais, na adaptação e aprendizado da escuta, e na persistência de posicionamentos de eventos auditivos. Por exemplo, Blauert nota que, em experimentos

---

<sup>1</sup> O presente autor não conseguiu encontrar uma tradução de uso comum para este conceito na língua portuguesa. Definido por Blauert como *localization blur*, este conceito é definido como a menor mudança possível de uma direção ou posicionamento de um evento auditivo que possa ser percebido como tal pelo ouvinte (BLAUERT, 1983, p. 37). Procurou-se usar consistentemente o termo ‘desvio mínimo’ para se referir a este conceito ao longo do texto.

laboratoriais, maiores durações de exposição ao estímulo de um sinal diminuíram o desvio mínimo necessário para a percepção de alterações de localizações de eventos auditivos pela escuta (BLAUERT, 1983, p. 42), mostrando uma dependência de curto tempo para a resolução de posicionamentos no espaço ocorrendo no campo de centenas de milissegundos. A adaptação da escuta pode ocorrer tanto em curto prazo, apresentando maior facilidade de localização a partir da identificação prévia da fonte sonora em posições de referência (BLAUERT, 1983, p. 45-46), como também em longo prazo, como em casos de adaptação da localização e perda parcial, simétrica ou assimétrica, da escuta (BLAUERT, 1983, p. 48-49). Nestes casos, a localização é adaptada aos parâmetros, agora em desequilíbrio, do ouvido interno, retendo parte das funções de localização de uma escuta sem alterações; Blauert nota este formato de adaptação em testes de localização com indivíduos com perda parcial de audição (BLAUERT, 1983, p. 48-49). A persistência de uma imagem auditiva à escuta mostra-se como um parâmetro inerente à audição, sendo este a taxa máxima de mudança que pode ser identificada pela audição dentro de um período de tempo (BLAUERT, 1983, p. 47). Isto resulta em uma dimensão limite de tempo para a identificação e localização de uma fonte móvel pela escuta, para além da qual a localização é comprometida.

A relação frequencial existente na escuta resulta em condicionamentos resultantes pela frequência de um sinal e pela largura de banda do mesmo. A frequência de um sinal define as regiões de estímulo do ouvido externo durante a audição, podendo apresentar ambiguidades ao sistema auditivo caso não gere clareza a este através dos parâmetros necessários para a localização de fontes. Anomalias auditivas resultam destas situações, comumente notadas em experimentos que usem de sinais senoidais (BLAUERT, 1983, p. 42-43). Em casos extremos, a localização de uma fonte sonora é definida em função somente da frequência, podendo ter posicionamentos específicos para respectivas frequências (BLAUERT, 1983, p. 45), como ilustrado abaixo (Fig. 2). Esta dependência frequencial para codificação de sinais incidentes gera uma consequente dependência da largura de banda de sinais. Sinais de banda larga permitem uma maior exposição de conteúdo frequencial aos elementos que codificam estímulos externos para a escuta, diminuindo a possibilidade de anomalias ou ambiguidades. Esta relação é notada em experimentos com maior conteúdo frequencial nos quais os sinais são mais facilmente localizáveis (BLAUERT, 1983, p. 42-43).

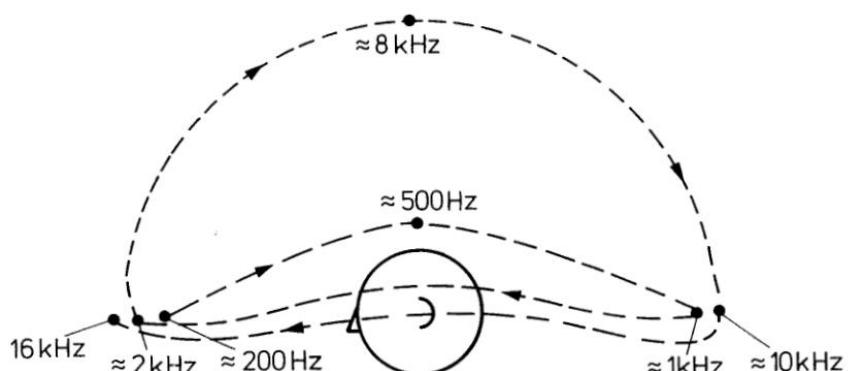


Fig.2 – Localização de eventos sonoros ao longo do plano vertical em função de sinais de banda estreita em diferentes frequências, indicando distância em função da frequência central de um sinal (FONTE: BLAUERT, 1983, p. 45).

Um último aspecto de importância é a familiaridade de sinais incidentes para um ouvinte. Uma maior familiaridade com as qualidades acústicas de uma fonte sonora permite uma localização de maior definição se comparada com fontes não familiares (BLAUERT, 1983, p. 45). Em situações particulares, como a definição da localização de tipos específicos de fala associados a distâncias ou posicionamentos em uma escuta natural, estas associações definem fortemente a atribuição de localização no espaço aos sinais, mesmo que outros aspectos deste contradigam esta atribuição (BLAUERT, 1983, p. 45-46). Um exemplo deste aspecto da escuta é a associação de diferentes tipos de fala com distâncias específicas de eventos sonoros (Fig. 3).

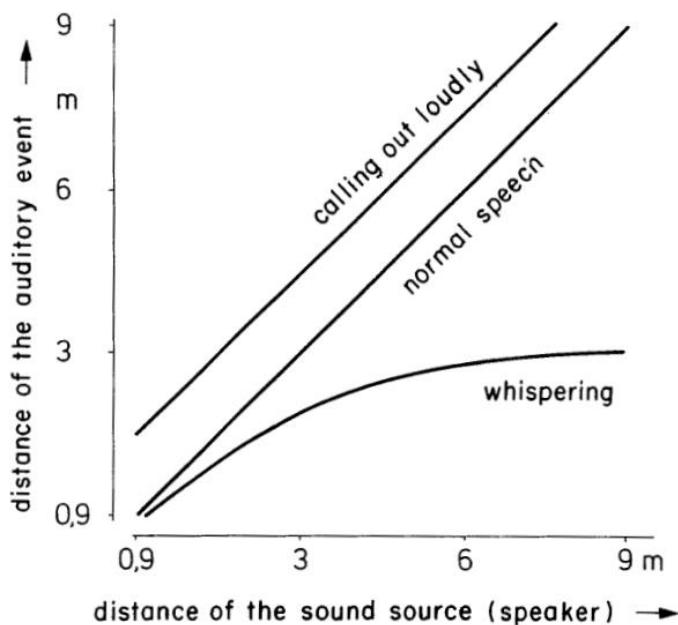


Fig. 3 – Localização entre a distância de uma fonte sonora e a percepção do evento sonoro para diferentes tipos de fala, emitidos à frente do ouvinte (FONTE: BLAUERT, 1983, p. 46).

Deste modo, é notável que diversos aspectos, tanto momentâneos como obtidos por associações próprias de escuta, acabam por moldar constantemente a localização resultante da percepção auditiva. A individualidade de parte destes aspectos é um fator que evidencia as possíveis diferenças que uma mesma fonte sonora pode ter em localização se apresentada a diferentes ouvintes. Como consequência, o estudo da localização na escuta espacial não se torna, então, um estudo da correta localização de uma fonte sonora, dado que as percepções individuais são fortemente moldadas por aspectos particulares (BLAUERT, 1983, p. 4), mas do modo geral de funcionamento que os mecanismos da escuta usam para esta localização. Nas seções seguintes, procura-se apresentar estes mecanismos em cada forma de localização dentro das dimensões da escuta.

### 1.3 ESCUTA NO PLANO HORIZONTAL

A lateralização define, no contexto da escuta espacial, as formas pelas quais eventos auditivos são identificados em relação ao plano horizontal da escuta. As posições formadas pelo sistema auditivo em casos de lateralização são resultantes da comparação de diferenças entre parâmetros de sinais incidentes, tendo ênfase principalmente parâmetros de diferenças interaurais de tempo e amplitude entre os sinais de ambos os lados do sistema auditivo (BLAUERT, 1983, p. 137-138). Ambas as desigualdades destas diferenças interaurais de sinais afetam a lateralização de uma fonte sonora, mas a importância que elas têm à escuta apresenta uma dependência frequencial que define bandas de frequência as quais diferenças de tempo ou amplitude tem maior peso no resultado final da audição.

Estas bandas de frequência dividem duas regiões onde um ou outro modo de lateralização é o principal modo pelo qual fontes são localizadas. Dois gráficos abaixo indicam os comportamentos de alteração de parâmetros de tempo e amplitude resultantes de ângulos de lateralização (Fig. 4-5). As frequências de corte para estas bandas ocorrem entre 800 e 1600 Hz. Esta divisão ocorre por coincidir com frequências cujas distâncias apresentam uma duração de meio-ciclo igual à diferença interaural de tempo de uma fonte em máxima lateralização (BLAUERT, 1983, p. 143). Esta duração máxima de atraso, de aproximadamente 630  $\mu$ s até 1 segundo, aponta o limite de uma resposta linear do posicionamento de fontes através de diferenças de tempo à escuta (BLAUERT, 1983, p. 143-144). Uma região de ambiguidade entre parâmetros interaurais ocorre até 1600 Hz, a partir da qual a escuta começa a depender de diferenças de amplitude para localização (BLAUERT, 1983, p. 149-150).

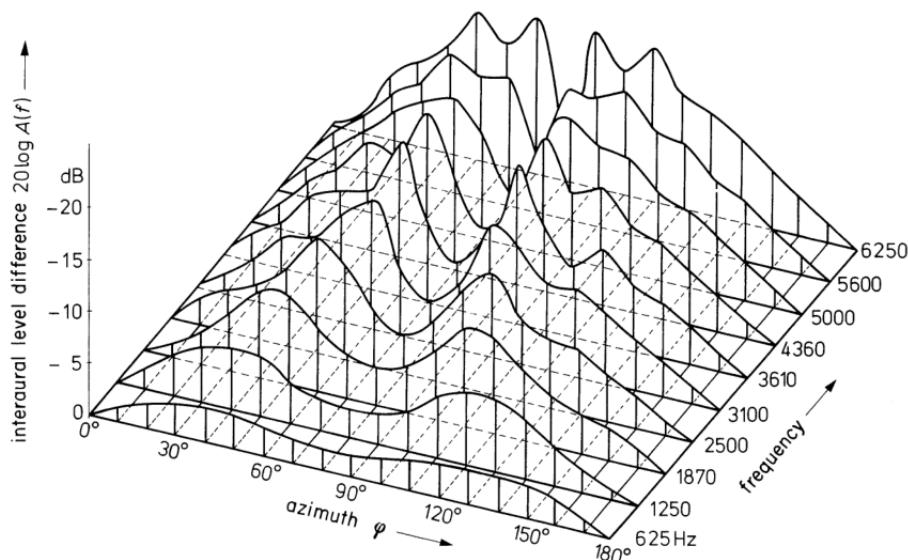


Fig. 4 – Diferenças de amplitude interaural, captadas por microfones dispostos em uma esfera, em função do ângulo de incidência lateral e frequência do sinal (FONTE: BLAUERT, 1983, p. 73).

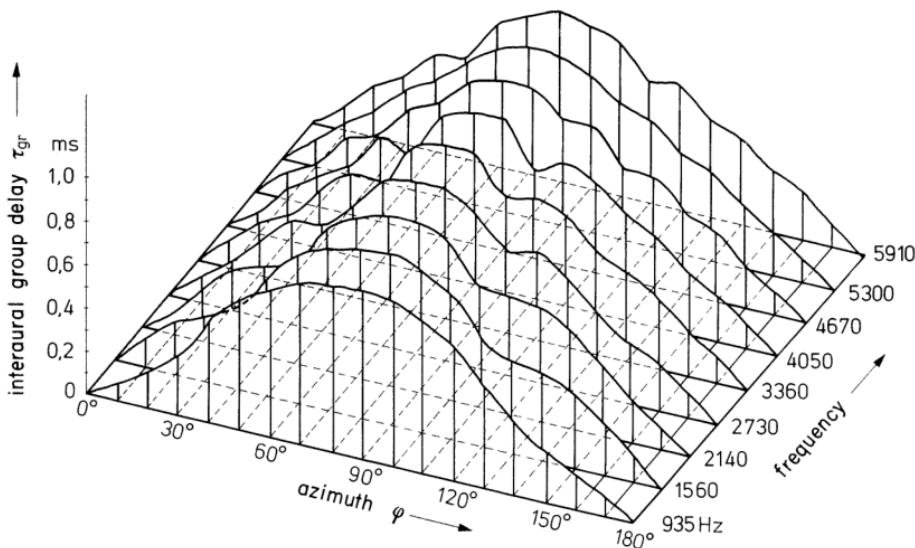


Fig. 5 - Diferenças de tempo de bandas de frequência, captadas por microfones dispostos em uma esfera, em função do ângulo de incidência lateral e frequência do sinal (FONTE: BLAUERT, 1983, p. 74).

Mesmo em casos de comparação de diferenças interaurais de amplitude, aspectos de tempo podem ser usados pelo sistema auditivo através da diferença de envelopes de amplitude. Usando dos mesmos mecanismos de lateralização para componentes freqüenciais abaixo de 800 Hertz, freqüências em todo o espectro podem ser localizadas ao longo do plano horizontal pela comparação destes envelopes. Nestes casos, o sistema auditivo não faz comparações entre componentes de fase, mas desloca a percepção com base nas diferenças de tempo do envelope geral do som (BLAUERT, 1983, p. 151). Em casos de sons com espectros contendo somente componentes acima de 1600 Hz, diferenças de fase entre dois sons que usam o mesmo envelope dinâmico não afetam a posição final da percepção (BLAUERT, 1983, p. 153).

Em casos nos quais as diferenças de fase e envelope dinâmico não podem ser usados pelo sistema auditivo, a lateralização torna-se em função de diferenças interaurais de amplitude. A lateralização máxima de um evento auditivo ocorre com diferenças entre 15 e 20 dB, com desvios mínimos aumentando ao passo que esta lateralização prossegue (BLAUERT, 1983, p. 157-158). Estas diferenças, que dependem da percepção de amplitude em cada sinal individual, são condicionadas por aspectos da audição, como a curva de audibilidade da escuta e respostas do ouvido interno à exposição prolongada a ruídos de grande intensidade (BLAUERT, 1983, p. 162). Estes aspectos da escuta, um inerente à desigualdade da percepção de volume ao longo do espaço da escuta, o outro um reflexo de proteção do tímpano que acaba por atenuar de forma desigual a sensibilidade a diferentes freqüências de sinais externos, condicionam a percepção de espaço baseada em amplitude.

A escuta fora de situações experimentais não faz um uso único de um ou outro parâmetro para a localização de fontes no espaço, mas de todas as diferenças interaurais disponíveis para o sistema auditivo. Estudos procuraram definir a importância relativa entre tempo, amplitude, e envelopes para a formação de imagens na audição, mas os valores obtidos dificilmente são coincidentes. As razões de igualdade entre microssegundos e decibels encontrados por estudos citados por Blauert (BLAUERT, 1983, p. 172) não são os mesmos que

aqueles definidos por Williams (WILLIAMS, 1984, p. 2-4) ou Malham e Myatt (MALHAM; MYATT, 1995, p. 61), por exemplo.

#### **1.4 ESCUTA NO PLANO VERTICAL**

Para a localização de eventos dentro do plano vertical da escuta, o ouvido humano se baseia fundamentalmente em diferenças de fase e amplitude geradas por fenômenos acústicos resultantes entre sinais incidentes e as qualidades individuais do ouvido externo (BLAUERT, 1983, p. 99). As distorções de parâmetros de sinais a serem avaliados pelo sistema auditivo causadas pela interação de ondas acústicas com o ouvido externo, interagindo principalmente com o pavilhão auricular, causam alterações que são próprias do ângulo de incidência, o que o codifica em ângulos de incidência específicos.

Outra decorrência destas alterações é a impressão de posicionamento frontal ou posterior de uma fonte sonora. Experimentos relacionados a esta avaliação de posicionamento indicam a correlação clara entre a codificação por parte do pavilhão auricular e a identificação correta da posição de um evento sonoro (BLAUERT, 1983, p. 99-100).

No entanto, esta codificação não é o único parâmetro de avaliação usado pela escuta. Dependências mencionadas anteriormente são fatores que condicionam a percepção de elevação de uma fonte sonora. Estas dependências são principalmente a largura de banda de um sinal, os componentes freqüenciais do espectro, e a duração de exposição (BLAUERT, 1983, p. 102). Uma menor margem de erro foi notada em experimentos que fizeram uso de sinais com uma duração maior de exposição, qualidade também notada em experimentos com um uso de sinais de banda larga. Não somente a largura de banda do espectro é importante, mas mostrou-se fundamental a presença de componentes freqüenciais acima de 7 kHz (BLAUERT, 1983, p. 102). Em menor grau, a escuta também é auxiliada se esta apresentar maior familiaridade com o sinal usado para avaliação. Este aspecto de familiaridade somente é uma variável à escuta em situações que apresentem ambiguidade ao sistema auditivo. A falta de familiaridade com as características de um sinal gera não somente uma maior margem de erro com a localização, mas uma tendência maior à localização de sinais com um posicionamento posterior, independentemente da localização da fonte sonora (BLAUERT, 1983, p. 103-104).

O fator condicional mais marcante da escuta de fontes em elevação é a sua dependência freqüencial para a localização de fontes sonoras. Em experimentos que investigaram esta dependência, resultados mostram que a localização de componentes freqüenciais específicos é elevada o quanto maior é a freqüência do componente (BLAUERT, 1983, p. 106-107). Mesmo alterações de sinais, volume, movimentação da cabeça, entre outros, não afetaram os resultados em consideração, mostrando uma forte correlação entre bandas freqüenciais específicas e a percepção de elevação destas à escuta. Segundo Blauert (BLAUERT, 1983, p. 107)

Estudos sobre esta tendência de resultados condicionados pela freqüência de um sinal notaram a existência de regiões de elevação relativas a um ponto de escuta cujas divisões tem maior probabilidade de localização de fontes sonoras. Esta probabilidade apresenta, segundo os resultados destes experimentos, a freqüência central de um sinal como a única

variável considerável à probabilidade de localização da fonte reproduutora. Como consequência destas tendências, estudos propõe um modelo de análise do processo de localização da elevação de fontes sonoras através de bandas de direcionalidade dependentes da frequência (BLAUERT, 1983, p. 108).

A ligação entre componentes freqüenciais também ocorre em sua conexão com as ressonâncias particulares do pavilhão auricular. Análises feitas das diferentes regiões de ressonância, essenciais para a codificação de sinais incidentes, notou, nestas regiões, comportamentos diversos quando apresentadas sons reproduzidos por fontes atrás e à frente do ouvinte (BLAUERT, 1983, p. 111). Diferentes regiões de ressonância são mais presentes através da amplificação de bandas de frequência de sinais incidentes do que outras para uma ou outra posição do reproduutor. A presença ou ausência de componentes espectrais nestas bandas de frequência é um fator determinante para a maior ou menor ambiguidade da localização de uma fonte reproduutora à escuta (BLAUERT, 1983, p. 112).

A interação entre estes dois modelos permite a identificação de fontes sonoras no espaço elevado. É notável a relação entre a amplitude de bandas freqüenciais, componentes de alta freqüência e a percepção de altura. Reproduções eletroacústicas, através de processamentos ou particularidades de equipamento, podem comprometer ou induzir a percepção de elevação a direções ou ângulos de incidência em particular. Sinais provenientes de síntese podem, devido a limitações de banda de freqüência ou distribuição de energia ao longo do espectro, ser mais ou menos fáceis de localização dependendo de como estes se disponham em comparação com as particularidades da escuta apresentadas. Em casos de gravação e reprodução binaural, alterações de fase e amplitude pelo equipamento de reprodução, assim como a recodificação de sinais reproduzidos ao ouvido, podem comprometer as relações próprias de fase e amplitude inicialmente gravadas (BLAUERT, 1983, p. 105).

## **1.5 ESCUTA EM DISTÂNCIA**

Dentre os três eixos de localização da escuta, a localização de distâncias de fontes sonoras pela audição é a mais complexa em termos de entendimento e reprodução (BLAUERT, 1983, p. 117). Alguns parâmetros importantes para a localização da distância de fontes já foram definidos em experimentos, mas, ainda assim, a reconstrução fiel ou a síntese deste tipo de percepção espacial ainda é relativamente limitada se comparada com a localização direcional em elevação e horizontalidade.

Há, para a identificação de distâncias, regiões relativamente delimitadas nas quais alguns parâmetros específicos são as principais variáveis para que a escuta. Estas regiões se limitam em distâncias de até três metros, entre três e quinze metros, e para além de quinze metros (BLAUERT, 1983, p. 118).

Dentro de uma distância de três metros entre a fonte reproduutora e o ponto de audição, existem distorções dependentes da freqüência ao longo do espectro de um sinal que alteram a amplitude e a fase de componentes como resultado da curvatura da onda irradiada pela fonte sonora em contato com a geometria da cabeça (BLAUERT, 1983, p. 127). A avaliação destas alterações relativas à curvatura do sinal irradiado parece ser o principal aspecto

de consideração para a escuta definir distâncias nesta região, auxiliada também das diferenças na percepção de volume resultante da distribuição de pressão sonora definida genericamente pelo inverso do quadrado da distância. Em distâncias muito curtas, a partir de vinte e cinco centímetros, as distorções aos componentes do espectro parecem não ter a mesma importância à escuta que em posições mais distantes (BLAUERT, 1983, p. 131).

Para além da distância de três metros, a localização de fontes sonoras depende principalmente da percepção de volume da fonte, apesar de existir a necessidade de algum tipo de referência prévia de amplitude e distância desta fonte para alguma localização em particular (BLAUERT, 1983, p. 119). Ao longo das distâncias nesta região de percepção, diferenças de fase e amplitude dos componentes freqüenciais do espectro do sinal incidente resultantes das diferenças de pressão sonora pelas curvas de igual audibilidade são, junto com a amplitude do sinal, um dos fatores principais usados para localização (BLAUERT, 1983, p. 119-120).

Se uma fonte sonora for posicionada para uma distância maior que quinze metros, o sinal irradiado por esta tem distorções ao longo do espectro resultantes da interação com a umidade ao longo da linha de transmissão, a velocidade do ar, e possíveis turbulências ao longo do caminho (BLAUERT, 1983, p. 126). Estas distorções são dependentes da frequência do componente espectral, e criam alterações no sinal que atenuam principalmente componentes de alta frequência. Esta sonoridade similar a um filtro de corte de alta frequência resultante da atenuação pela umidade do ar entre a fonte e o ponto de escuta torna-se mais importante para a localização de distância do que o volume do sinal percebido pelo sistema auditivo.

Mesmo que estes sejam os principais parâmetros que o sistema auditivo faz uso para a definição da distância em cada respectiva região, outros aspectos condicionais apresentados anteriormente, como familiaridade, tipo de sinal, e largura de banda, ainda são fatores que afetam a definição da localização de uma fonte sonora. Não somente a definição de localização depende destes aspectos condicionais, mas também a diminuição de desvios mínimos em diferentes distâncias é variável a estes aspectos.

Uma anomalia importante relacionada à identificação de distância de uma fonte sonora é a localização de fontes sonoras dentro da cabeça. Segundo experimentos, esta anomalia é resultante do fator de correlação de fase entre sinais no ouvido interno (BLAUERT, 1983, p. 133), ou em situações de escuta nos quais componentes de um sinal não permitam uma localização através dos processos empregados pelo sistema auditivo. Desta última situação, um exemplo comum é o uso de fones de ouvido para escuta, criando uma forma de representação de fontes sonoras cuja ausência de interação com aspectos habituais à audição, resumidos pela função de transferência individual, impede que a escuta atribua qualidades espaciais às fontes reproduzidas. Uma forma de diminuir esta limitação de externalização de fontes sonoras em sistemas de reprodução, segundo Blauert, é o uso de sistemas de equalização que permitam aproximar a resposta de uma sistema de reprodução à equalização da escuta natural, na qual haja interação entre todos os elementos que codificam a informação espacial de sinais incidentes (BLAUERT, 1983, p. 135).

## **1.6 ASPECTOS MULTIMODAIS DA ESCUTA ESPACIAL**

Investigações relacionadas à escuta espacial e à possibilidade de reconstrução da percepção de qualidades espaciais no áudio comumente se limitam a aspectos exclusivos da capacidade de reprodução de sistemas de áudio. Apesar desta abordagem parecer óbvia dentro dos limites do campo em questão, estudos parecem apontar para uma necessidade mais abrangente de estudo e implementação de tecnologias em diferentes modalidades da percepção para que se possa mais fielmente reproduzir a percepção do espaço através do som.

Como dito previamente, parâmetros presentes em sinais de áudio, mesmo quando obtidos através de métodos que permitem a captação e reprodução de elementos espaciais como os provenientes de funções de transferência da escuta, não são capazes de exaurir todos os estímulos usados para que se construa uma ilusão espacial tal qual numa escuta não mediada por elementos tecnológicos. Aspectos auxiliares como a propriocepção, através do sistema vestibular, podem afetar indiretamente a percepção do espaço visual e auditivo, ainda que somente em situações de extremo distúrbio da sensação de equilíbrio (BLAUERT, 1983, p. 198-199).

No entanto, o uso de movimento arbitrário do corpo e do aparelho auditivo durante uma situação de escuta mostra-se fundamental para a formação de imagens auditivas do espaço. A possibilidade de movimentação é um fator essencial na diminuição de ambiguidades provenientes de sinais acústicos codificados pelo ouvido externo. Um dos principais usos do movimento em uma situação de escuta é a centralização da direção de uma fonte sonora no espaço para dentro do eixo frontal da escuta (BLAUERT, 1983, p. 179). O eixo logo à frente do pavilhão auricular apresenta os menores erros de localização e os menores desvios mínimos para todos os três modos de percepção espacial. A movimentação permite diminuir erros de localização ao possibilitar a centralização de fontes na região de maior definição da escuta.

Além disto, ambiguidades geradas através de anomalias da codificação espacial, como inversões de posicionamento à frente e atrás de um ponto de escuta em relação a uma fonte sonora, são possíveis de serem resolvidas através da movimentação, dado que a rotação ao redor de um eixo permite a comparação de codificações únicas para um trajeto em particular relativo à fonte (BLAUERT, 1983, p. 180-181), como ilustrado abaixo (Fig. 6). Estas rotações podem ocorrer ao longo dos três eixos da escuta, permitindo que as alterações no tempo resultantes dos trajetos deem ao sistema auditivo a capacidade de distinguir ambiguidades particulares da codificação espacial (BLAUERT, 1983, p. 182-184). Posicionamentos formados através somente de parâmetros exclusivos da escuta podem ser alterados pela movimentação mesmo quando os sinais acústicos são suficientemente apresentados e uma imagem auditiva já tenha sido formada (BLAUERT, 1983, p. 190). A importância que este aspecto tem à formação da imagem espacial pela escuta já tem sido estudada para aplicações binaurais, e implementações de sistemas de reprodução que possibilitem uma reconstrução usando da movimentação em espaços reconstruídos através de sistemas de áudio têm sido pesquisadas nos últimos anos.

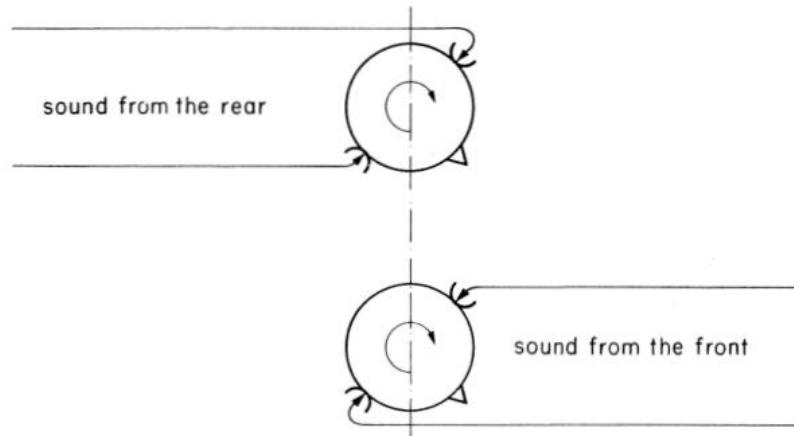


Fig. 6 – Ilustração da percepção oposta de aspectos interaurais de codificação da escuta para sinais à frente e atrás de um ouvinte quando este rotaciona em relação à fonte sonora, permitindo a resolução de ambiguidades de localização entre os planos frontal e traseiro (FONTE: BLAUERT, 1983, p. 180).

De forma semelhante, o uso da visão como complemento aos estímulos externos condiciona a percepção de espaço do sistema auditivo. Através de diversos experimentos, notou-se a conexão entre estímulos visuais com a localização de fontes sonoras, principalmente no que diz respeito à localização de uma fonte em uma direção ou distância, mesmo que parâmetros dos sinais à escuta discordem desta localização (BLAUERT, 1983, p. 193-194). Tendências de expectativa em relação a elementos da visão, como a presença de alto-falantes antes da exposição a sinais acústicos discordantes àquele posicionamento, podem ter maior importância em certas situações de audição do que aspectos derivados pelo sistema auditivo a partir de estímulos acústicos (BLAUERT, 1983, p. 194-195). Este efeito de ventriloquismo é um dos fatores que é essencial à percepção da direcionalidade de eventos auditivos a telas quando um sinal é reproduzido junto a uma imagem (KUSUMOTO; KYTO; OITTINEN, 2015, p. 49).

Parece existir uma tendência, citada por Blauert como complicação espacial (BLAUERT, 1983, p. 195), na qual uma mescla de estímulos provenientes de diversos modos de percepção resultem em uma conclusão sobre posicionamentos espaciais destes estímulos, ainda que eles indiquem localizações contrastantes. Esta tendência parece poder ser direcionada para pesos distintos aos diferentes modos de percepção (BLAUERT, 1983, p. 196), mas a sua importância é notável. Em alguns casos, por exemplo, diferenças em parâmetros acústicos resultantes da interação com o espaço através da movimentação podem ter uma importância até maior que o próprio sinal acústico na formação de imagens espaciais no sistema auditivo quando estes apresentam conflitos de localização (KENDALL, 1995, p. 32). A necessidade da interação entre diversos estímulos, mais do que somente a reprodução fiel de sinais acústicos, aponta para caminhos de desenvolvimento de tecnologias imersivas. Esta, uma amalgama de diversos modos de percepção em uma exposição unificada de estímulos, concordantes ou não, sugerem a possibilidade de um modo próprio de reconstrução espacial à escuta que vai além do uso exclusivo de alto-falantes.

## CAPÍTULO 2:

### HISTÓRICO DE SISTEMAS DE ÁUDIO ESPACIAL

Apesar de todas as implementações de tecnologias de reprodução acústicas serem inherentemente espaciais de alguma maneira, intencionalmente ou não, para a escuta humana, o avanço destas tecnologias somente apresentou preocupação com estas qualidades inerentes de seus sistemas em alguns momentos de sua história. Durante a maior parte do desenvolvimento das tecnologias de áudio, tanto os sistemas, quanto as propostas de reprodução de qualidades espaciais, estiveram, em grande parte, em segundo plano se comparadas com a preocupação apresentada com outros aspectos da reprodução eletroacústica.

A importância dada a aspectos como linearidade de resposta, extensão de resposta de frequência, e reprodução de transientes, em detrimento de outros aspectos relacionados à reprodução de sinais eletroacústicos e, por consequência, de outros aspectos inerentes à escuta humana, explicitam um escopo particular de interesses mais recorrentes do que outros quando se investiga a história do desenvolvimento tecnológico de sistemas de reprodução de áudio.

No entanto, durante períodos específicos, em grande parte por desenvolvimentos externos ao áudio em si, as possibilidades de reprodução de qualidades espaciais de sinais, gravados ou não, receberam atenção momentânea, o que possibilitou a criação de sistemas que foram desenvolvidos ao longo das décadas seguintes. O pleno desenvolvimento destes sistemas, mesmo que algumas vezes já viáveis em projetos de teste anteriores, em várias ocasiões somente veio a ocorrer anos depois, quando as execuções de algumas propostas se tornaram mais viáveis.

Dentre os principais momentos em que ocorreram inovações consideráveis no campo da reprodução de sinais espaciais, pode-se notar avanços desde o início da gravação. Boren remonta aspectos de sistemas de reprodução até o próprio início da gravação acústica, no final do século XIX, com a invenção do fonógrafo. Segundo Boren (BOREN, 2018, p. 40):

The advent of audio recording in the 19<sup>th</sup> century led to the development of zero-dimensional (mono) sound, and later one-dimensional (stereo), and two-dimensional (quad and other surround formats) reproduction techniques. Due to the greater sensitivity of the human auditory system along the horizontal plane, early technology understandably focused on this domain. Our capability to mechanically synthesize full 3D auditory environments is relatively recent, compared to our long history of shaping sound content and performance space.

Estes avanços dimensionais no potencial de representação espacial tiveram protagonismos particulares que se tornaram referências para pesquisas seguintes. Boren lista os principais sistemas iniciais desenvolvidos para diversas abordagens de representação de sinal: o *Theatrophone* como um primeiro marco para transmissão binaural de sinais eletroacústicos, as pesquisas de Harvey Fletcher e Alan Blumlein em diferentes concepções de codificação espacial na década de trinta, os experimentos de áudio multicanal em filmes como *Fantasia* (1940), as diferentes propostas de sistemas quadrifônicos, o desenvolvimento de sistemas

ambissônicos a partir das pesquisas de Michael Gerzon na década de setenta, e as propostas mais recentes de particulares sistemas *surround* (BOREN, 2018, p. 48-55).

De forma simplificada, nota-se tendências no desenvolvimento ao longo destas tecnologias. Pode-se notar implementações binaurais, apesar de rudimentares, tendo seu início com Ader nas suas transmissões telefônicas, implementações estereofônicas, com seus fundamentos sendo definidos pelas pesquisas de Alan Blumlein e Harvey Fletcher, e os princípios fundamentais da síntese de campo de onda<sup>2</sup>, também desenvolvidos por pesquisadores ao lado de Fletcher nos laboratórios Bell durante a década de trinta.

Estas três tendências de implementação de sistemas de reprodução são as principais guias de implementações mais tarde ao longo do século vinte: vê-se desenvolvimentos que procuram refinar aspectos da gravação ou reprodução de sinais em certos sistemas, ou pesquisas que procuram complementar aspectos da audição que haviam sido inicialmente deixados de lado. No entanto, os fundamentos destas três abordagens aos processos de gravação e reprodução espacial mantiveram-se como guias até mesmo às implementações recentes do século XXI, mesmo que aspectos desenvolvidos nos dias de hoje não estivessem contemplados nas primeiras proposições destes sistemas.

Neste capítulo, discorre-se sobre os principais momentos na história das tecnologias de reprodução de áudio durante os quais foi dada uma importância particular em pesquisas e propostas de sistemas que possibilitassem a melhor reprodução de qualidades espaciais. Mesmo com um grande distanciamento temporal entre certos eventos, os princípios gerais de certas implementações de sistemas de reprodução espacial não apresentaram drásticas mudanças ao longo de sua história. Em grande parte, as diferenças entre as propostas de sistemas de reprodução se diferem principalmente entre números de canais passíveis de serem reproduzidos e de princípios inerentes a cada sistema no que diz respeito a quais aspectos de um espaço devem ser dados prioridade em uma situação de reprodução. Novas propostas de sistemas de representação espacial surgem nas últimas décadas do século passado, apesar do interesse e desenvolvimento destas propostas terem sido impulsionados principalmente nas últimas décadas com o ressurgimento de interesse em implementações de aplicações imersivas. Torna-se notável, ao final, que importantes paradigmas, particulares a implementações passadas, ainda se mantêm, mesmo que os contextos que os tornaram presentes tenham sido superados tecnologicamente.

## 2.1 THEATROPHONE

A primeira proposta de um sistema de reprodução de sinais acústicos que empregasse mais que um reproduutor, com a intenção de reprodução das qualidades acústicas de um espaço, é datada de 1881 por Clément Ader. Cinco anos após a patente de Alexander Graham Bell de 1876 (BELL, 1876), que define o marco inicial para transmissões telefônicas, Ader desenvolve um sistema de captação, transmissão e reprodução de dois canais independentes ao longo de todo o sistema. Fundamentado nas pesquisas anteriores de escuta

---

<sup>2</sup> O termo *wave-field synthesis* foi traduzido para o português ao longo do texto como síntese de campo de onda.

espacial feitas por pesquisadores desde pelo menos 1796 (WADE; DEUTSCH, 2008, p. 16), Ader desenvolve o primeiro sistema que põe em prática as noções de diferenciação de amplitude interaural como modo de indução da percepção espacial de múltiplas fontes em reprodução acústica.

Este modelo de funcionamento da escuta espacial humana, fundamentado na diferenciação de parâmetros interaurais, foi a primeira proposta amplamente difundida de como a escuta humana atribui características espaciais a diferentes estímulos externos. Wade e Deutsch citam diversas pesquisas que ocorreram antes da proposição destes modelos para a escuta binaural, principalmente por experimentação procurando a procura de aspectos comuns entre o comportamento da visão binocular e a escuta binaural, que serviram de base para os primeiros sistemas espaciais de reprodução acústica (WADE; DEUTSCH, 2008, p. 17):

Among those who later pursued binocular color combination were Wells, Venturi, Wheatstone, Ernst Heinrich Weber (1795–1878) and August Seebek (1805–1849), and it is notable that all these thinkers were to consider similar aspects of binaural combination.

Estas pesquisas culminaram em um modelo que atribuía a distinção do posicionamento de diferentes fontes sonoras no espaço auditivo às diferenças de fase e amplitude que eram transmitidas independentemente a cada ouvido, criando diferenças interaurais nestes parâmetros dos sinais incidentes. A concepção deste modelo da escuta espacial pode ser notada nos trabalhos iniciais de escuta espacial de Lord Rayleigh, em 1876. Segundo Yost (YOST, 2017):

Rayleigh describes in some detail the ease with which he can locate in left-right and front-back dimensions the location of a person on the circle who speaks. He uses calculations, shown in the paper, to argue for a ratio of the intensities at the two ears as a basis for such an ability: a “binaural ratio” due to reflection and diffraction of sound from and around the head (head-shadow).

Esta razão binaural entre valores de amplitude nos ouvidos direito e esquerda que Rayleigh menciona é tida como o principal parâmetro usado pelo sistema auditivo para posicionamento de fontes sonoras no espaço da escuta nas primeiras propostas de modelos para a escuta humana, e influencia consideravelmente as primeiras décadas de desenvolvimentos de sistemas de gravação e transmissão de áudio.

Os primeiros fonógrafos e gramofones, no entanto, apesar de contemporâneos a estas experimentações psicoacústicas, não vieram a incluir qualidades espaciais além do que era incidental ao próprio processo de gravação que era empregado pelos estúdios nascentes. Apesar de inicialmente promoverem um meio de escuta que transmitia diretamente o som através de tubos ao ouvinte do fonógrafo, estes registros eram exclusivamente monofônicos. Rumsey menciona que, apesar de alguns fonógrafos terem tido duas ou mais cornetas, e terem tido a possibilidade de reprodução de mais de um registro ao mesmo tempo no mesmo reproduutor, estas gravações ainda eram exclusivamente monofônicas, seus únicos aspectos espaciais sendo as sugestões de distância e profundidade resultantes da reverberação captada no momento da gravação (RUMSEY, 2001, p. 10).

Clément Ader é o primeiro a viabilizar um sistema de transmissão de canais de áudio independentes com o intuito de recriar posicionamentos de fontes sonoras. Dentro os primeiros usos da telefonia, as transmissões de apresentações de ópera ao vivo já haviam se

firmado como uma prática paga em cidades da Europa neste período. Ader propõe um sistema de canais independentes como integração do serviço de telefonia existente para óperas e das pesquisas de escuta espacial daquela época (PAUL, 2009, p. 2-3).

A proposta de Ader se baseia na simples disposição de microfones ao longo do fim do palco onde se encena o que seria transmitido, com microfones com distâncias aproximadas de um metro entre cada microfone e seu respectivo par, e transmitir estes canais independentemente, garantido que eles mantivessem independência também no ato de reprodução (ADER, 1882). Este distanciamento entre pares de microfones cria diferenças de tempo e amplitude que codificam a impressão de espacialização das fontes captadas, mesmo que não coincida a aspectos da escuta binaural em situações naturais. Sem experimentos prévios de implementações similares no campo da transmissão estéreo, a tentativa que Ader implementa acaba por comprometer aspectos essenciais para a recriação precisa de uma cena auditiva, se contempladas as necessidades que se conhece hoje pela psicoacústica. Procurando categorizar o *Theatrophone*, Boren afirma (BOREN, 2018, p. 50):

[...] it is probably safest to say that Ader's achievement was the earliest reproduction of binaural sound under the 19<sup>th</sup>-century understanding of the term. Since the transmissions did not make use of a dummy head to obtain level differences, time differences, or spectral cues corresponding to the actual filtering effects of the head, a present-day understanding might instead classify the Theatrophone as a very effective form of 2-channel stereo, distributed over several listening points.

O funcionamento desta proposta, tanto a parte de implementação elétrica e telefônica, como a concepção psicoacústica dos comportamentos da escuta, eram claramente entendidos. Na patente entregue um ano após as primeiras demonstrações por Ader em Paris, há explicações, ainda que simplificadas, da diferenciação de pressão sonora que era resultante do uso de microfones distanciados no espaço como havia sido as primeiras instalações, e do efeito que isto teria no ouvinte do outro lado da linha. Segundo o próprio Ader (ADER, 1882, p. 5):

Hence when an actor or singer is at the right of the stage the sound heard by the auditor's right ear is the louder, and as he moves toward the left of the stage the sound heard by the auditor's left ear becomes the louder, while that heard by his right ear decreases. An effect is thus produced upon the ears akin to that which the stereoscope produces on the eyes, so that the auditor is enabled to follow the actor's movements about the stage [...]

Uma explicação semelhante, mostrando o conhecimento de implementação e comportamento psicoacústico por trás do Theatrophone de Ader, ocorre também em um artigo da revista *Scientific American* que comentou o sistema durante sua segunda exposição, em 1881 (SCIENTIFIC..., 1881, p. 422-423).

Apesar destes conceitos terem sido difundidos poucos anos após a invenção do fonógrafo e da telefonia, tanto em propostas como as de Ader, como na difusão científica das revistas na época, e terem definido claramente a possibilidade e os meios pelos quais seria possível a inclusão de aspectos dimensionais do espaço em difusões telefônicas e registros físicos de gravação, pouco foi feito nas décadas seguintes para esta inclusão. Sunier menciona um segundo sistema, praticamente idêntico ao de Ader, mas nenhum outro desenvolvimento além desta reimplementação da mesma ideia do *Theatrophone* parece ter ocorrido (SUNIER, 1960, p. 27-28). Transmissões de óperas e teatros anos após as exposições do Theatrophone,

como em 1891 pela American Telephone and Telegraph Company, se mantiveram monofônicas, evidenciando o aparente desinteresse, na época, pela implementação de transmissões estereofônicas (TORICK, 1998, p. 27-28). A queda de interesse na estereofonia decorreu, possivelmente, das falhas do Theatrophone durante as suas exposições, dado o nível de ruído da transmissão e consequentemente a suspensão das audições para o público durante as primeiras exposições (BURNS, 2006, p. 128).

## **2.2 ALAN BLUMLEIN**

Esta pausa na integração de possibilidades de captação e reprodução estereofônica se estendeu até o final da década de vinte. Neste período entre as experimentações de Ader e as pesquisas da década de 30, descobertas como transmissões radiofônicas e métodos de gravação elétrica surgiram, mas não se procurou usar destes novos métodos para a reinvenção da estereofonia integrada a estes novos meios.

Os primeiros desenvolvimentos consideráveis que ocorreram neste então surgiram como decorrência do interesse do áudio para cinema, campo crescente desde os primeiros filmes sincronizados a som durante a segunda metade dos anos vinte, principalmente desencadeado pelas novas possibilidades de materiais e qualidade que as gravações proporcionadas por mecanismos elétricos.

Este interesse evidencia-se nas pesquisas de Alan Blumlein, um dos dois principais nomes relacionados ao desenvolvimento da estereofonia. Em 1929, Blumlein foi empregado para o desenvolvimento de um sistema elétrico concorrente às novas pesquisas empregadas por Maxfield e Harrison. Estes, empregados pela Bell Labs, haviam possibilitado o emprego de equipamentos elétricos, principalmente trazidos do uso em rádio e telefonia, e diminuído assim diversos limites que a gravação acústica apresentava previamente (MAXFIELD; HARRISON, 1926, p. 246-248). Nos anos de 1929 e 1930, Blumlein chega a desenvolver avanços à gravação elétrica: microfones, linhas de transmissão, e modos de gravação de disco foram desenvolvidos na EMI durante este tempo (ROBERTS, 2013, p. 38-60).

Após a conclusão das pesquisas em gravação elétrica, Blumlein pesquisou tópicos sobre estereofonia, com uma visão principalmente para a melhora de atribuição de localização a fontes sonoras em filmes, procurando uma realização mais fiel à escuta natural quando atores se locomoviam de um ponto a outro em uma cena. O resultado destas pesquisas preliminares culmina nas suas patentes de 1931, onde Blumlein delineia todos os aspectos necessários para gravação, transmissão, e reprodução de sinais estereofônicos a partir da reprodução em disco (BLUMLEIN, 1931, p. 2-5). Apesar de ter trabalhado nos anos seguintes em execuções concretas que demonstrassem os princípios definidos nestas pesquisas, os fundamentos para o entendimento do comportamento estereofônico de sistemas, assim como os modos como se poderia criar a ilusão de posicionamento espacial em sistemas de dois canais, já se encontravam definidos nas suas patentes deste ano. Gravações de áudio que exploravam vários tipos de movimentação em cena foram feitas em 1933 como demonstrações do funcionamento da proposta de Blumlein, tendo tido um “definitivo efeito binaural” (ROBERTS, 2013, p. 76-78). Blumlein empregou novos testes, tendo criado diversos filmes que serviram como demonstração prática de suas ideias (ROBERTS, 2013, 80-92). No entanto, mesmo após estas

demonstrações, a EMI, companhia que empregou Blumlein durante suas pesquisas, não chegou a usar nenhuma de suas proposições em gravações, filmes ou transmissões comerciais. Segundo Roberts, (ROBERTS, 2013, p. 91-92):

As for binaural sound recordings, suffice to say that it was not until 1958, some 16 years after Alan Blumlein had died and 27 years after it had first been patented, that ‘binaural’ made its next appearance as ‘stereo’ on long playing records.

Burns menciona inadequações do material usado na época na indústria fonográfica como um fator limitante na criação de discos estéreo (BURNS, 2006, p. 140):

In 1935 the commercial prospects for stereo records and domestic stereo gramophones also seemed bleak. The shellac discs of the 1930s were noisy and did not permit the subtle sounds associated with present-day stereophonic LP and CD recordings to be reproduced with sharp clarity. Blumlein's process required a record material which had a much lower surface noise than the shellac—slate mixture then in current use.

As pesquisas de Blumlein foram interrompidas pela Segunda Guerra, tendo sido movido para áreas de pesquisa científica empregada durante a guerra. Mesmo com esta pausa, sem terem sido retomados após o falecimento de Blumlein, todos os resultados de seus experimentos da década de trinta definiram práticas que ainda hoje se mantêm como padrões de gravação e reprodução de áudio estereofônico. Boren cita, por exemplo, padrões para uso de microfones para codificação estéreo, balanço panorâmico de dois sinais ao longo de um sistema multicanal, e uma proposta concreta de codificação estereofônica em um único sulco de mídia física de discos (BOREN, 2018, p. 51). Mesmo décadas depois, quando o formato estéreo estava a ser preparado para comercialização, e com diversas outras propostas para o mesmo problema, os modelos iniciais de Blumlein ainda se mantiveram como padrões da indústria fonográfica (TORICK, 1998, p. 29). A retomada da aplicação de princípios estereofônicos, fortemente baseados na pesquisa de Blumlein, só foram retomados para desenvolvimento comercial na final de década de 50, com o desenvolvimento de sincronização entre fitas magnéticas e menor ruído de superfície em discos de vinil (CLARK; DUTTON; VANDERLYN, 1957, p. 417-419).

## 2.3 PESQUISA ESTEREOFÔNICA NOS LABORATÓRIOS BELL

Ao mesmo tempo que as pesquisas lideradas por Blumlein, que procuravam viabilizar a estereofonia, ocorriam na Inglaterra, pesquisas similares eram feitas nos Laboratórios Bell. Tendo sido o centro de pesquisa que possibilitou o avanço de meios elétricos para gravação sonora, grande parte dos meios necessários para se desenvolver propostas que possibilitassem a independência de canais em transmissão e reprodução já se encontravam presentes para os pesquisadores da Bell.

Ao que parece, nas primeiras décadas do século vinte, sempre esteve em questão uma escuta binaural, mas sem as possibilidades de implementação existentes até o avanço da gravação elétrica. Transmissões multicanal existiram após Ader por outros, apesar de terem sido reproduzidas em somas monofônicas (TORICK, 1998, p. 27-28), e tecnologias que tiravam proveito da possibilidade de localização espacial da escuta humana também foram desenvolvidas durante o período acústico de gravação, como, por exemplo, escutas amplificadas para localização aérea durante a Primeira Guerra (SUNIER, 1960, p. 28). Estudos

aprofundados em aspectos da escuta e da fala, procurando melhorar as qualidades de transmissões de telefonia e rádio, foram feitos na Bell, o que possibilitou um maior entendimento de comportamentos da escuta e a presença de equipamento de pesquisa que viabilizasse a criação de novos meios de gravação. Segundo McGinn (MCGINN, 1983, p. 39):

Prior to 1930 advances in the art of telephony had given rise to extensive instrument development work, which, together with the fundamental knowledge of speech, hearing, and music it made possible, provided [Bell Labs] with a solid basis for its investigations in high-fidelity recording and reproducing.

Esta pesquisa de alta fidelidade desenvolvida na Bell foi liderada por Harvey Fletcher, o diretor de pesquisas acústicas na época. Fletcher tinha como principal interesse a melhora da comunicação telefônica, procurando limitar todos os problemas de transmissão, como a banda funcional de frequência, inteligibilidade, e também o que ele definia como “perspectiva auditiva,” a percepção espacial das fontes sendo comunicadas (MCGINN, 1983, p. 55). Ao longo de toda a década de trinta, Fletcher desenvolveu diversas abordagens e sistemas de reprodução espacial para estas questões, principalmente com seu trabalho paralelo em gravações com o regente Leopold Stokowski.

A mais importante contribuição resultante deste trabalho conjunto de Fletcher e Stokowski, se analisado retrospectivamente, foi o entendimento de reprodução de frentes de onda e a sua conexão com os respectivos sistemas de reprodução eletroacústica. Este modelo de reprodução de frentes de onda e os fundamentos para a aplicação de sistemas multicanal foram amplamente documentados em um conjunto de seis artigos publicados em 1934 por pesquisadores da Bell como resultado de experimentações durante uma transmissão multicanal de música orquestral ocorrida no ano anterior (FLETCHER, 1934, p. 9). Fletcher explica o princípio básico do funcionamento da síntese de campo de onda (FLETCHER, 1934, p. 10):

Assume each microphone to be connected with a perfect transmission line which terminates in a projector occupying a corresponding position on a similar curtain in hall R. By a perfect transmission line is meant one that delivers to the projector electrical energy equal both in form and magnitude to that which it receives from the microphone. If these sound projectors faithfully transform the electrical vibrations into sound vibrations, the audience in hall R should obtain the same effect as those listening to the original music in hall O.

Este formato de reconstrução de um espaço acústico em outro local através do uso de um número suficientemente grande de alto-falantes parece ter sido notado primeiro como decorrência das pesquisas ocorridas por Fletcher. A figura abaixo indica um diagrama do sistema explicado por Snow (Fig. 7). No entanto, por causa das limitações técnicas nas linhas de transmissão, microfones, e alto-falantes da época, não foi possível a construção de um sistema que pudesse evidenciar na prática a aplicação do modelo teórico proposto por Fletcher. A transmissão na qual foram baseados os resultados para os artigos de 1934 fizeram uso de três alto-falantes apenas, possivelmente procurando a sugestão, mais do que a ilusão, da localização precisa de fontes sonoras (FLETCHER, 1934, p. 11).

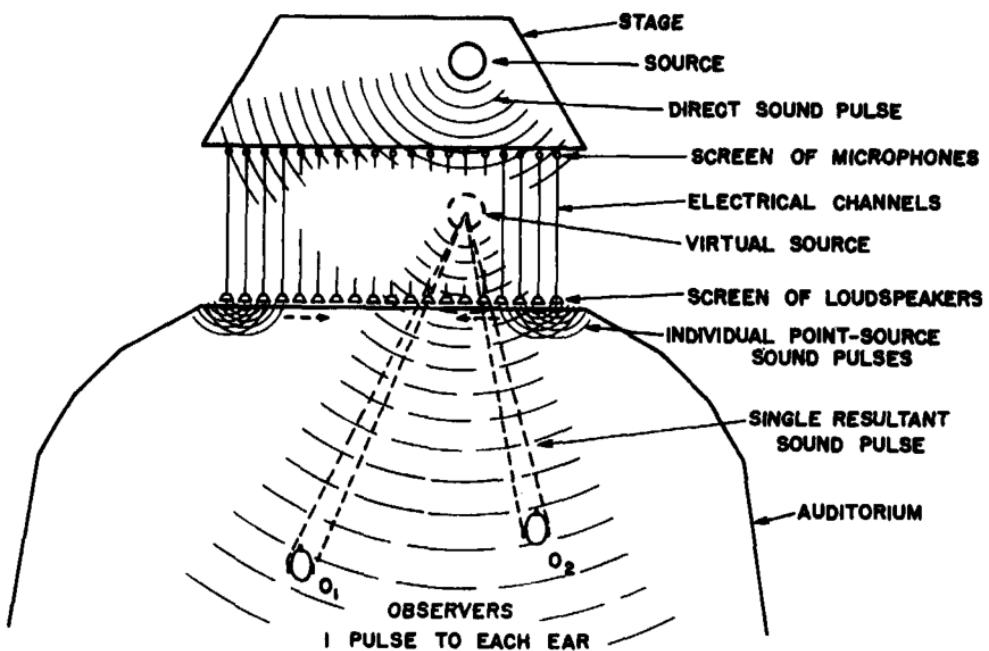


Fig. 7 – Diagrama explicativo do sistema de reconstrução de frente de onda inicialmente proposto por Fletcher  
(FONTE: SNOW, 1953, p. 572).

Procurando uma similaridade entre uma reconstrução fiel do espaço acústico e um sistema prático, um dos artigos da Bell mostra como foram posicionados microfones e alto-falantes nas transmissões de experimentação de 1933. O uso destes procurou formar um espaçamento igual entre as posições tradicionalmente definidas de público e orquestra em uma situação de concerto, os três microfones posicionados ao longo da divisória entre palco e público, e os reprodutores localizados em igual disposição na sala para onde foi transmitido o concerto (BEDELL; KERNEY, 1934, p. 217-218). Há menções que uma matriz de três colunas e fileiras de microfones e reprodutores tenha sido testada antes dos testes finais por Fletcher, procurando incluir elementos verticais de reprodução acústica, mas estes não foram tidos como necessários (MCGINN, 1983, p. 55).

Este comprometimento com um número particularmente limitado de alto-falantes, se comparado com a intenção inicial de recriação fiel do espaço acústico de uma sala de concerto, parece ter ocorrido em parte pelo programa escolhido para reconstrução. Tendo escolhido uma fonte sonora, uma orquestra sinfônica dentro de um espaço reverberante de uma sala de concerto, relativamente difusa espacialmente, principalmente pela distância considerável de grande parte dos ouvintes, parece não se ter notado grande diferença se incluídos mais alto-falantes como tentativa de formar uma imagem espacial mais nítida de diferentes instrumentos no palco. Neste teste, a distância do ouvinte comum em relação aos membros da orquestra ultrapassa a distância crítica da escuta por causa das qualidades de reverberação do espaço acústico de uma sala de concerto, e as contribuições das diversas reflexões para a escuta se tornam maiores do que a presença do som direto das fontes sonoras. Neste tipo de espaço, somente seria necessário, para se criar uma impressão de posicionamento espacial, um sistema minimamente definido espacialmente, sem serem necessários inúmeros microfones e alto-falantes, dado que a precisão de localização espacial contribuída por estes

seria suprimida pelo caráter reverberante das salas onde o programa seria reproduzido. Esta parece ter sido a realização de Fletcher quando ele diz (FLETCHER, 1934, p. 10):

Theoretically, there should be an infinite number of such ideal sets of microphones and sound projectors, and each one should be infinitesimally small. Practically however, when the audience is at a considerable distance from the orchestra, as usually is the case, only a few of these sets are needed to give good auditory perspective; that is, to give depth and a sense of extensiveness to the source of the music.

Os primeiros testes críticos de escuta de espacialização de fontes sonoras ocorreram como um meio de avaliação das diversas sessões de gravação e propostas decorrentes dos testes de Fletcher. Apesar dos primeiros modelos de reconstrução de frentes de onda propostas por Fletcher terem sido uma visão original para a localização de fontes virtuais, e ter sido o ponto de partida para modelos de síntese de campo de onda, estes testes críticos de escuta demonstraram que era desnecessária e impraticável, naquela situação, uma quantidade tão grande de equipamentos como inicialmente se propunha, principalmente com os aspectos técnicos dos equipamentos na época (MCGINN, 1983, p. 55).

Este sistema de três canais idealizado por Fletcher foi o principal formato de reprodução sonora empregado para o cinema durante algumas décadas, dado que ele garantia uma região de cobertura considerável para as salas de cinema da época. Em um artigo mais tarde, Snow explica esta aplicação ao cinema (SNOW, 1953, p. 578):

Three channels appear to be a good economic choice for ordinary stages and auditoriums. Good accuracy of localization can be achieved for favorable observing positions, with reasonable results at other seating locations. The center channel is a great aid for solo and close-up work, as well as removing the “hole-in-the-center” effect mentioned above. For unusually wide stages, additional channels have been found necessary.

Aplicações diretas ao cinema comercial, no entanto, não ocorreram imediatamente após estas pesquisas. Mesmo quando usando somente um subconjunto dos canais disponíveis, somente eram posicionadas fontes monofônicas reproduzidas em um único alto-falante; somente se usavam os canais laterais para se sugerir uma correlação de fontes sonoras em tela e no espaço, sem a intenção de se recriar movimentação ou a ilusão do espaço sugerido em filme. É notável, no entanto, que estas primeiras pesquisas tenham tido grande impacto em propostas mais tardias de sistemas de reprodução de áudio, principalmente para situações comerciais. Estes sistemas propostos foram resultantes de ambas as limitações de equipamentos eletrônicos de época, como também de uma escolha específica tanto de um programa particular, como também de uma localização predefinida de escuta em relação ao contexto de apresentação. Propostas mais tardias de adequação destes sistemas de reprodução para locais menores, quando procurando formatos comerciais para a introdução de sistemas de reprodução estéreo a ambientes caseiros, parecem ter se baseado, no entanto, nestes mesmos testes, reproduzindo, talvez anacronicamente, os limites que existiam nestes experimentos na década de trinta.

## **2.4 FANTASIA (1940) E O SISTEMA FANTASOUND**

Apesar de o cinema ter tido a possibilidade técnica de expansão para sistemas multicanais durante a década de trinta como resultado das experimentações nos Laboratórios Bell e na EMI, houve um grande atraso para estas mudanças acontecerem. Os filmes produzidos após as publicações dos artigos de Fletcher e Blumlein parecem ter se mantido com os mesmos sistemas que já haviam sido desenvolvidos décadas atrás. Isto se baseava em uma reprodução monofônica, gravada também em mono, e entre um, dois, ou três alto-falantes usados para se manter uma área de cobertura adequada, dependendo do tamanho da sala.

O primeiro filme a propor um sistema multicanal foi o filme de animação *Fantasia*, estreada em 1940. Sendo um filme baseado na conexão entre a animação e o repertório orquestral que compunha toda a trilha sonora ao longo do filme, a qualidade sonora tinha grande importância para o produto final. Os sistemas de reprodução que eram então usados não pareciam satisfatórios para as intenções e usos em filmes musicais. Tanto os limites técnicos de qualidade, como as limitações criativas que as disposições e equipamentos de então tinham, pareciam estar cada vez mais perceptíveis, tanto para o público como para profissionais. Segundo um engenheiro da produção de *Fantasia* (GARITY; HAWKINS, 1941, p. 127):

While dialog is intelligible and music is satisfactory, no one can claim that we have even approached perfect simulation of concert hall or live entertainment. It might be emphasized that perfect simulation of live entertainment is not our objective. Motion picture entertainment can evolve far beyond the inherent limitations of live entertainment.

Garity aponta diversos fatores limitantes do que se tinha como padronização de equipamentos de som em salas de cinema na época: a limitação de intervalo dinâmico decorrente das mídias de gravação e reprodução e do nível de ruído inerente nestas; a inerente localização pontual de fontes sonoras quando reproduzidas em um sistema de um único alto-falante, impedindo a distribuição de fontes sonoras no espaço físico e psicoacústico; a fixação da fonte sonora, e consequentemente a ação em cena, no centro da imagem; e a impossibilidade de efeitos dramáticos decorrentes da movimentação e posicionamento de fontes sonoras no espaço (GARITY; HAWKINS, 1941, p. 128).

*Fantasia*, e o Fantasound, sistema de som o qual idealmente teria acompanhado a animação em todas as distribuições, foi produzido procurando superar estas limitações. O processo de gravação, que começou em 1939, teve início com diversas tentativas de gravação que, para que tivessem maior flexibilidade no estágio de edição, utilizaram de uma abordagem multicanal que procurava gravar as diversas seções da orquestra o mais isoladamente possível, principalmente para que fosse viável a alteração de equilíbrio dinâmico entre as partes instrumentais mais tarde na produção (GARITY; JONES, 1941, p. 7).

É possível que as gravações que foram feitas para o isolamento dos diversos segmentos da orquestra neste projeto tenham sido a primeira gravação multicanal no áudio com o intuito de reprodução também em multicanal (HOPE, 1979, p. 30). Um filme anterior a *Fantasia*, cuja trilha foi gravada em 1937 por Stokowski, com um formato de gravação quase idêntico ao empregado mais tarde, usando gravadores ópticos em diversos canais isolando segmentos da orquestra, teve a sua mixagem final distribuída em formato mono. (HOPE, 1979, p. 29).

Nas sessões de gravação para a trilha de *Fantasia*, a disposição e atribuição de cada um dos nove canais seguiam o seguinte formato (HOPE, 1979, p. 29):

In all, 33 microphones were used to cover the orchestra and six of the available nine channels were delegated to cover the sections' such as violin, celli, basses, woodwinds and so on. The seventh channel was used to record a mono mix of all six and the eighth channel recorded a distant pickup of the entire orchestra. The ninth channel was designated as metronome or click track for use by the animators.

Apesar deste número de canais durante o processo de gravação, a redublagem do material gravado durante a mixagem final, destinada a situações sem os preparos do sistema de reprodução total inicialmente planejado, acabava por usar somente três canais, dispostos como esquerda, direita e centro, sem que houvessem canais para efeitos ou sons ao redor da plateia; uma canal extra era usado somente para a gravação de um sinal de controle para amplificação que permitia a automação de ganho dentre os três canais para possibilitar uma redução de ruído durante a reprodução final (GARITY; HAWKINS, 1941, p. 145-146).

Os custos para as salas de cinema se adequarem às diversas necessidades que o sistema Fantasound requeria foram possivelmente o principal fator limitante que impediu que a trilha final fosse ouvida da forma como foi inicialmente idealizada. O equipamento de reprodução necessário para a sincronia dos diversos canais era atípico para instalações da época, e as trilhas impressas não se adequavam ao equipamento já existente nas salas de cinema, sendo inviável a readequação daquelas ao formato menor, dado que isto diminuiria o intervalo dinâmico do sistema, um dos principais ganhos obtidos durante o desenvolvimento do sistema (HOPE, 1979, p. 29). Sobre estes custos, Hope menciona (HOPE, 1979, p. 30):

The cinema installation was very exotic and very expensive. When *Fantasia* was premiered in 1940 there was just one cinema, the Broadway in New York, fully equipped as Disney intended. A total of 90 speakers, 39 behind the stage and 54 spread around the auditorium, were installed along with ganged picture and sound film projectors. The total cost of installation, even in 1940, was \$85,000. The object of spreading speakers round the auditorium was to reproduce sound sections of the track in true surround-sound fashion.

Além disto, outras dificuldades para a adaptação de salas na época também acabaram limitando a abrangência de cinemas que pudesse apresentar *Fantasia* em multicanal. Garity aponta o tempo necessário para instalação de equipamentos, as condições de disponibilidade de recursos e tempo durante o período de inauguração durante a Segunda Guerra, as diferenças de regulamentações técnicas em diferentes locais de apresentação, e fatores de espaço para equipamentos como reprodutores, amplificadores, cabos e alto-falantes (GARITY; JONES, 1942, p. 15).

Apesar das dificuldades que acabaram sendo apresentadas ao longo do período de gravação, produção, e apresentação de *Fantasia*, outras gravações foram feitas usando o sistema como novos testes de possibilidades e potenciais da reprodução multicanal, principalmente procurando viabilizar uma maior distribuição no espaço e uma localização mais precisa de segmentos da orquestra ao longo da cena recriada pelos falantes no espaço (PLUMB, 1942, p. 18). Estas gravações não chegaram a ser usadas em nenhuma produção cinematográfica e, apesar das intenções iniciais dos criadores do sistema Fantasound, nenhum outro filme parece ter sido produzido usando o sistema após *Fantasia*.

Os resultados destes experimentos foram retomados na década de cinquenta, quando houve uma coincidência de fatores que impulsionaram avanços técnicos no cinema. Com o aparecimento de novas propostas de diversos aspectos para a produção e reprodução de

filmes, novos métodos de reprodução acústica foram recuperados. A disponibilidade de gravação em fita magnética, recentemente desenvolvida durante a Segunda Guerra, e sistemas de reprodução como o Cinerama, que permitia a reprodução de sete canais distribuídos ao longo do espaço da sala, e o Cinemascope, acabaram por padronizar formatos para além do estéreo no processo de produção cinematográfica. Após estas mudanças, Fantasia foi readaptado a formatos estéreo e multicanal e redistribuído (HOPE, 1979, p. 30)

## 2.5 SISTEMAS QUADRIFÔNICOS

As recentes inovações no áudio cinematográfico na década de cinquenta foram um considerável salto para o desenvolvimento e para a expansão de sistemas multicanal no áudio. Tendo tido um impulso com outras tecnologias paralelas, como o avanço da televisão comercial e a viabilidade da gravação em fita magnética, o cinema veio a implementar diversos métodos de apresentação multicanal. Com o advento de grandes sistemas sendo desenvolvidos e amplamente divulgados, e com soluções concretas para problemas que o áudio cinematográfico tinha lidado em décadas passadas, o interesse e o envolvimento da indústria cresceram e possibilitaram o avanço de implementações de áudio multicanal duradouras. Os altos custos para adaptação de salas para os novos formatos de distribuição sonora em cinemas não se apresentaram como fator limitante por causa da popularidade notada com os primeiros filmes produzidos até então (DIENTSFREY, 2016, p. 174).

Diferentes sistemas implementaram soluções similares ao problema de como codificar um número de canais em películas para reprodução, algo que ainda se mostrava sem um padrão fixo naquele momento. A necessidade de se produzir diferentes películas em diferentes formatos para diferentes sistemas de reprodução se mostrava como um entrave para grandes produtores.

A codificação de sinais em mídia se mostrou como uma resposta a este entrave. Em filmes distribuídos em formato Cinerama, por exemplo, limites entre quatro a seis canais, dependendo do formato aplicado, eram codificados em filme óptico. Dentro de uma lógica que dava preferência aos canais frontais e limitava o áudio direcionado a canais laterais e posteriores a somente uma das faixas disponíveis, usando esta como canal de efeito ou auxiliar à ação em cena, os canais eram codificados e dependiam de um sistema de decodificação específico para a matriz ou processo usado na codificação da película no processo de gravação. Diversos filmes durante a década de cinquenta até os anos oitenta usaram destes formatos simplificados de distribuição espacial de sinais, com a intenção de tornar canais laterais e traseiros como complementares à ação em tela, onde se localizavam preferencialmente a música e o diálogo (DIENTSFREY, 2016, p. 174-175).

Esta simplificação de processos de gravação e distribuição de sinais em diferentes sistemas procurava tornar mais fácil a distribuição de filmes para diferentes redes, possibilitando uma independência de sistemas de reprodução em relação às fitas magnéticas produzidas. Deste modo, o processo da produção de películas poderia ser unificado em um único método que dependeria de decodificadores particulares em cada sala para recriar as trilhas de áudio corretamente para cada respectiva instalação (DIENTSFREY, 2016, p. 176).

Um avanço similar, seguindo semelhantes diretrizes, ocorreu anos depois com a iniciativa de desenvolvimento do áudio multicanal para instalação caseira. A continuação de pesquisas decorrentes dos trabalhos de Blumlein e Fletcher na década de trinta teve sequência nas décadas seguintes, dando maior entendimento a métodos de implementação de sistemas estereofônicos. Estas pesquisas, assim como aspectos econômicos, sociais e técnicos, possibilitaram a introdução, apesar de lenta, de sistemas estereofônicos em nível caseiro, iniciado durante o final da década de cinquenta. Segundo Valiquet (VALIQUET, 2012, p. 408):

Multichannel film sound continued to gain in popularity through the 1950s with the development of systems like 'Cinerama,' 'WarnerPhonic,' and four-channel 'Cinemascope.' The simultaneous development of competing multi-channel systems for musical media continued until 1957, when the adoption of the two-channel Westrex vinyl cutting method was forced through the intervention of the small American record label, Audio Fidelity.

Mesmo com um curto período de tempo desde a introdução do áudio estereofônico caseiro, deu-se início na década de setenta a uma tentativa de introdução de sistemas quadrifônicos neste mesmo mercado. Desenvolvido a partir de pesquisas voltadas para o funcionamento de canais adicionais laterais em sistemas estéreo, a rápida tentativa de introdução de diferentes formatos individuais possivelmente foi incentivada por uma intenção de se dominar uma tecnologia através de patentes. Mencionando o avanço de sistemas de áudio quadrifônico, referido como *quad*, Davis diz (DAVIS, 2003, p. 561):

[...] quad avoided the compromise in stereo between precise imaging and enveloping ambience, with the front channels providing the imaging and the surround loudspeakers providing the ambience. But part of the underlying motivation was almost certainly commercial one-upmanship, arguably resulting in products and systems being rushed into the marketplace prematurely.

Davis aponta como a codificação de canais em uma mídia limitada em espaço físico foi um dos principais desafios de desenvolvimento de sistemas quadrifônicos. Diferentes sistemas discretos, que carregavam cada canal individualmente para reprodução, não foram tão viáveis em prática como sistemas que implementavam matrizes de decodificação para reprodução, extraíndo os quatro canais de somente dois canais físico em mídia (DAVIS, 2003, p. 561-562).

Dientsfrey aponta que distorções consequentes dos métodos de codificação dos quatro canais eram comuns nos discos quadrifônicos, causando alterações de fase e volume em todos os quatro canais e gerando vazamentos de todos os canais entre si (DIENTSFREY, 2016, p. 184). Não havia a possibilidade, mesmo com diferentes sistemas de decodificação quadrifônica, de se eliminar os vazamentos decorrentes da codificação multicanal nos discos por completo; o resultado era a imprecisão de localização de fontes e falha em criar segmentações claras de posicionamento espacial ao longo dos quatro alto-falantes. Davis menciona como técnicas de automação de ganho foram usadas para aumentar o nível de isolamento entre canais, similar a processos de compressão, mas estas automações não funcionavam de forma segura, podendo criar diversos desequilíbrios no ato de reprodução (DAVIS, 2003, p. 562).

Estas falhas não eram aparentes, apesar de existirem, em sistemas multicanais no cinema, de onde as implementações caseiras se basearam. As aplicações para o cinema de sons que não fossem frontais era conscientemente evitada pela crença de que não se deveria fazer

uso de sons fora do campo de visão da tela de projeção, acreditando que esta localização tornaria difícil a atenção na narrativa (DIENTSFREY, 2016, p. 185). Como consequência, dificuldades técnicas que eram possíveis em sistemas multicanais em outras aplicações não ocorreram durante os diversos filmes que foram feitos em sistemas como o Cinerama antes do desenvolvimento de aplicações quadrifônicas para ambientes caseiros.

A grande dispersão de diversos sistemas concorrentes de áudio quadrifônico, os altos custos de adaptação de ambientes caseiros para implementações particularmente invasivas de espaço, e dificuldades técnicas de entregar uma experiência imersiva como inicialmente era proposto acabaram por dificultar a longevidade das implementações quadrifônicas em situações comerciais (HALLUM, 2017, p. 8).

## 2.6 DESENVOLVIMENTO DO ÁUDIO AMBISSÔNICO

Mesmo antes das propostas de reprodução quadrifônica da década de 70, já haviam ocorrido algumas pesquisas que procuravam maneiras de introduzir o áudio multicanal e a sensação de espaço para além de salas de cinemas. Estas pesquisas iniciais visando a possibilidade de uma expansão da reprodução monofônica ou estereofônica comum se baseava fortemente em uma manipulação específica dos canais já existentes na mídia física, usando de equipamentos de processamento de sinal para alterar e reproduzir sinais que sugeriram localização ou ambientação espacial para mais do que um alto-falante. Um exemplo deste sistema é a proposta chamada ambiofônica, desenvolvida por Keibs em 1960 (RUMSEY, 2001, p. 15).

O trabalho de Keibs se diferencia em abordagem das pesquisas que foram desenvolvidas na década de 70 no campo quadrifônico, dado que Keibs propõe um sistema mais flexível de representação, algo estranho à concepção frontal-lateral de canais principais e auxiliares ao programa reproduzido que foi amplamente aceita pelas patentes quadrifônicas na época, tendo se diferenciado majoritariamente os modos de codificação e matrizes de reprodução entre as diversas patentes, mas não a concepção de reprodução espacial em si.

Keibs não se apresenta como o único indivíduo a sugerir diferentes concepções de espaço no áudio durante as primeiras décadas da segunda metade do século vinte. As propostas de Michael Gerzon, desenvolvidas no começo da década de setenta em paralelo com as pesquisas quadrifônicas para áudio de consumo, se assemelhavam em visualizar a possibilidade de diferentes modos de extrair e trabalhar com a informação espacial em áudio.

As primeiras propostas de codificação ambissônica foram desenvolvidas como consequência da pesquisa pessoal de Gerzon sobre esquemas de codificação espacial relacionados à gravação e reprodução quadrifônica (BOREN, 2018, p. 53). Diferente das propostas prévias de reprodução espacial baseadas em *downmixing* e codificação de canais, a concepção de Gerzon se fundamentava na gravação e armazenamento de componentes espaciais de ordens específicas de detalhamento, necessitando uma matriz de decodificação para que se obtivesse estes componentes para reprodução.

O início das pesquisas de Gerzon se deram com a procura de um modo de se obter uma ambientação similar ao que se tinha em reproduções quadrifônicas, mas se limitando a

somente dois canais de mídia. Seus resultados foram publicados em um artigo em 1970, onde Gerzon explica a proposta (GERZON, 1970a, p. 1104):

The experimental set-up used by the author included four loudspeakers and two stereo amplifiers, [...]. The sounds fed to the four speakers were as follows: the left-hand speaker was fed with the left stereo channel, the right-hand speaker was fed with the right stereo channel, the ‘front’ speaker was fed with the sum of the two stereo channels, and the rear speaker was fed with the difference ‘between’ the two stereo channels.

Esta proposta tornava possível uma recriação de espacialidade em multicanal a partir de uma gravação estereofônica. Propostas similares, que fazem uso de matrizes simples com inversões de fases e somas em níveis específicos, são amplamente usadas para criação de pseudoestereofonia a partir de gravações monofônicas, sendo conhecidas pelo menos desde a década de 50 (SCHROEDER, 1958).

No entanto, a abordagem que Gerzon faz uso se assemelha ainda mais com as matrizes de diferença que foram usadas por Blumlein no início de sua pesquisa estereofônica. Os usos que Blumlein fez da gravação com microfones bidirecionais coincidentes permitia, com o uso de uma matriz de decodificação específica, que fossem extraídos ambos os lados de captação dos microfones, permitindo uma gravação de todo o plano horizontal em sua totalidade ao redor da posição de gravação (GERZON, 1970a, p. 1108).

Esta possibilidade do uso de conjuntos de microfones e matrizes de decodificação para a captação de sinais que seriam componentes de representação de dimensões do espaço, e não canais individuais de gravação pontuais, foi o que possibilitou as ideias iniciais de desenvolvimento para sistemas ambissônicos.

A primeira proposta de uma aplicação ambissônica foi tratada por Gerzon em um artigo de 1970, onde Gerzon já tinha definido claramente as necessidades da implementação de um sistema ambissônico. Neste artigo, Gerzon mostra as comparações entre os sistemas quadrifônicos propostos até então, baseados em reproduções de canais e centrados na captação e reprodução de um palco acústico limitado, mostrando uma possibilidade da captação e derivação dos componentes dos canais destes formatos quadrifônicos a partir de somente três microfones (GERZON, 1970b, p. 338-339). Os diferentes valores de ganho para cada canal de um conjunto de três microfones que possibilitariam a derivação de componentes quadrifônicos já eram apresentados por Gerzon neste artigo, tão quanto outros métodos e valores que possibilitariam a conversão entre formatos quadrifônicos e ambissônicos (GERZON, 1970b, p. 341). Gerzon também propõe um quarto componente deste conjunto de microfones para a captação de um elemento que não teria direcionalidade, podendo ser usado para controlar o foco, ou dispersão, de fontes sonoras de uma cena acústica gravada durante a reprodução (GERZON, 1970b, p. 342).

Como continuação destas propostas, testes são feitos em 1971 e seus resultados são publicados por Gerzon no mesmo ano. Dentre os resultados que foram obtidos, em grande parte sendo primeiras experiências com um sistema experimental, é enfatizada a possibilidade de reprodução de componentes de altura a partir da gravação. O formato de reprodução proposto nesta ocasião espelhava a disposição tetraédrica das cápsulas dos microfones no formato original de captação. Mesmo não tendo tido um número maior de alto-falantes destinados para reprodução destes componentes de altura, é notada a distinção entre a presença da reprodução

destes em comparação com sistemas quadrifônicos contemporâneos. Segundo Gerzon (GERZON, 1971a, p. 397):

There are many who regard the height effect as an altogether unnecessary luxury and, at first sight, our choice of music with virtually no vertical spread seems to suggest they are right. Yet the listening tests showed quite the opposite - the height effect on the reverberation added very considerably to the realism. Indeed, several listeners standing outside the tetrahedron still found the spaciousness of the recording to be superior to that obtained from most conventional four-channel recordings within the square of speakers.

A exploração da possibilidade e da importância de componentes de altura na reprodução de sinais gravados já tinha se mostrado como interesse de Gerzon em publicações anteriores (GERZON, 1970c, p. 381). Outras propostas de sistemas que valorizavam uma reprodução mais abrangente de elementos espaciais no áudio não se mostravam particularmente interessadas em representações de elementos verticais. No entanto, a desconexão do sistema ambissônico com a direcionalidade da gravação de seus canais individuais, em comparação com a direcionalidade limitada da captação tradicional baseada em canais, possibilitava a extração de todos os componentes de altura durante o processamento em pós-produção. É possível que a diferença na percepção de realismo em reproduções com componentes verticais tenha sido possibilitada primeiro pela viabilidade de se captar e reproduzir estes componentes.

Destes experimentos de 1971, Gerzon obtém os fundamentos do que viriam a ser implementações ambissônicas. Durante este período, as matrizes de conversão, algumas possibilidades de disposição de reprodutores, e um uso prático de um conjunto de microfones para a gravação dos diferentes componentes dimensionais foram desenvolvidos e padronizados. Este conjunto de técnicas e teorias, encapsulando um sistema de gravação possível em todas as direções, era chamado de perifônico (GERZON, 1971b, p. 513).

A adição de Gerzon a estes fundamentos foi feita através da inclusão de aspectos psicoacústicos, os quais não figuravam em sistemas multicanal na época. O conceito de áudio multicanal até então era o de criação de um espaço acústico dentro de um volume circundado por reprodutores, tendo a sugestão de direcionalidade resultando das fontes sonoras projetadas para dentro deste volume no espaço. A alteração que Gerzon viu como possível era a de recriar o comportamento acústico dentro de um volume no espaço, tendo a percepção espacial como consequência das projeções. Este paradigma de reprodução é também compartilhado com formatos de campo de onda. Segundo Gerzon (GERZON, 1974, p. 485):

[...] if a surround-sound system is to work optimally, it must be capable of capturing all nuances of reverberant sound and of reproducing these uniformly around the listener. Certain popular commercial matrix systems assign the original sound field to the two available channels in such a discontinuous manner that these criteria cannot be satisfied.

Esta mudança na posição do funcionamento ideal de sistemas multicanal foi uma consequência de pesquisas de Gerzon em psicoacústica, com foco principalmente em gravações binaurais. Estas inclusões foram publicadas em dois artigos em 1975 por Gerzon e Fellgett. Uma definição resumida do sistema, enfatizando a flexibilidade do sistema e seus fundamentos psicoacústicos, é dada por Fellgett no começo de seu artigo (FELLGETT, 1975, p. 20):

Ambisonics is a technology for surround -sound which aims specifically at not making four (or any other number) of loudspeakers audible as separate sources of sound. It is

designed using appropriate engineering methods and psycho-acoustic theory that has shown good predictive value to make best use of available channels of communication (two or more), and of loudspeakers (a limitation often forgotten), to give stable and uncoloured acoustic images in any position, keeping the physical means of reproducing the sound as unobtrusive as possible.

Nestes artigos, dentre particularidades das então recentes aplicações e avanços de sistemas ambissônicos, os autores discorrem sobre as comparações entre suas propostas e os sistemas quadrifônicos em voga na época, assim como os diferentes formatos para armazenamento, produção, e distribuição de gravações ambissônicas em casos comerciais (FELLGETT, 1975, p. 22), e as possibilidades de alteração e processamento dos espaços gravados através de matrizes de rotação, por exemplo (GERZON, 1975, p. 30).

O desenvolvimento final destas primeiras aplicações ambissônicas pode ser visto na conclusão do processo de criação de microfones independentes que viabilizassem uma captação em formato ambissônico sem a necessidade de um conjunto de microfones direcionais em disposição tetraédrica. Este processo foi concluído em 1978, apesar de ter sido iniciado como projeto de desenvolvimento em 1975 (GERZON, 1975, p. 24-25), com a comercialização de modelos de microfones que implementavam as necessidades de captação e disposição de sistemas ambissônicos em sua construção.

Se analisado retrospectivamente, é possível notar que, apesar de desenvolvido em paralelo com outras propostas e sistemas multicanal durante o começo da década de 70, as ideias de Gerzon eram únicas em sua abordagem. Seu foco em um material de gravação flexível e sem comprometimento particular com uma visão ou disposição espacial de materiais, assim como a viabilização, através de matrizes, para a possibilidade de processamentos e reconstruções em qualquer quantidade de alto-falantes mais tarde, mostra como o sistema ambissônico se propunha como mais do que outra entre tantas patentes de propostas de reprodução multicanal. Os formatos de gravação e distribuição ambissônicos, diferente de outros formatos baseados em canais, possibilitam a reconstrução, alteração, e reinserção de perspectivas dentro de uma cena acústica. Esta flexibilidade da informação adquirida através destes métodos torna formatos ambissônicos particularmente valorizados nos dias atuais, onde processamentos digitais, interfaces entre usuário e computador, e um uso refinado de metadados possibilita novas práticas e usos de gravações de espaços acústicos.

## **2.7 PROPOSTAS DE ÁUDIO MULTICANAL SURROUND**

A distribuição de filmes e programas que usavam de tecnologias de decodificação para derivação de canais laterais e traseiros para o mercado caseiro facilitou o acesso às implementações de tecnologias similares ao cinema para além de usos cinematográficos. O formato *Dolby Surround*, cujo nome se tornou sinônimo a estas tecnologias, foi decorrente da distribuição de filmes originalmente em dois canais que faziam uso de matrizes, similares às implementadas em sistemas quadrifônicos, para extrair dois canais, central e lateral, durante a reprodução, tendo compatibilidade com formatos que não usassem destas codificações.

Estes usos foram viabilizados para comercialização para além de aplicações diretas a salas de cinema com a introdução na década de oitenta de mídias de vídeo com qualidade

suficiente para decodificação por algoritmos em equipamento de vídeo (DAVIS, 2003, p. 563). Esta viabilização também ocorreu poucos anos depois em sinais usados para distribuição televisiva e radiofônica.

Um esforço para a convenção de formatos multicanal foi feito durante a década de noventa, procurando uma concordância entre práticas já implementadas em sistemas cinematográficos e resultados experimentais da percepção e localização espacial em diferentes situações de reprodução. Os resultados destes esforços foram publicados por Günther Theile em 1993. A convenção proposta, desenvolvida por um conjunto de grupos relacionadas à transmissão e distribuição de rádio, cinema, e televisão, figurava como um formato que se adequava entre o uso de aplicações com e sem o acompanhamento de imagens (THEILE, 1993, p. 2). Esta proposta seguia em parte as implementações práticas dos formatos *Dolby Stereo* e *Dolby Surround*, principalmente por retomar, por influência destes, o paradigma de disposição de reprodutores para a criação de um espaço acústico dividido entre imagens frontal e lateral, algo que não tinha figurado, pelo menos em parte, em outros sistemas multicanal durante as décadas anteriores. Este paradigma é evidenciado por Theile (THEILE, 1993, p. 2):

[...] the 3/2- or 3/4-loudspeaker arrangement offers enhanced directional stability and clarity of the frontal sound image and improved realism of auditory ambience. However, the addition of side/rear surround loudspeakers does not enlarge the listening angle (by delivering genuine surround localization of phantom sources); rather, it adds an acoustic environment to the frontal stereophonic presentation of directional sound. [...] Thus, in contrast to Quadraphony, the locations of the side/rear surround loudspeakers are largely noncritical with respect to both direction and distance, and they can be placed to fit into an existing living-room environment.

Theile define este sistema como um avanço a um sistema estéreo tradicional, sem a intenção de criar um espaço bem definido e com fontes sonoras de localização precisa para além do espaço acústico frontal. A introdução de mais canais a este formato é tida como uma possibilidade, fazendo uso de processamentos similares aos usados por formatos no cinema em décadas anteriores para diminuir a correlação entre um canal, gerando sinais ideais para a simulação de um espaço reverberante. No entanto, estes processamentos visam tornar os canais adicionais em canais auxiliares a este espaço reverberante, sem a intenção de que eles tornem mais bem definidas as fontes sonoras posicionadas ao redor do ouvinte. Mesmo quando listando as possibilidades de canais laterais e traseiros, Theile menciona somente aplicações de espaço sem localização precisa, como ambiências de salas, espaços ou efeitos (THEILE, 1993, p. 7).

Nota-se, então, a intenção da convenção do uso comercial de sistemas multicanal como uma transferência de formatos, e, por consequência, também de seus paradigmas, do áudio cinematográfico para o espaço caseiro. Estes formatos tinham como intenção a inclusão, desde seus primeiros modelos, de uma compatibilidade reversa com programas anteriores a suas aplicações. Como decorrência, programas estéreos que não usassem elementos de imagem na sua reprodução acabavam por ser reconstruídos dentro dos mesmos paradigmas da reprodução visual, ainda que pesquisas quadrifônicas durante a década de 70 tenham procurado, mesmo que nem sempre, uma exploração de novos modos de apreciação de experiências acústicas enquanto experimentação do espaço.

Desenvolvimentos recentes no campo de formatos *surround*, principalmente na década de 90 com uma maior interação entre as aplicações de áudio e a tecnologia digital, viabilizaram a gravação e distribuição de programas multicanal através de novas mídias comuns

em circulação. Sem a limitação de canais em mídias analógicas, o uso de canais discretos para todos os sinais em uma mídia digital fez com que novos sistemas não necessitassem de processamento de decodificação e extração de sinais para a reprodução (KYRIAKAKIS, 1998, p. 943). Limitações inerentes ao processo de decodificação por matrizes desde os primeiros formatos quadrifônicos, como modulação intercanal e desequilíbrios em localização resultantes de diferenças de fase e amplitude nos sinais, e a necessidade de gerar a percepção de sinais reverberantes a partir do processamento de um único sinal mono para alto-falantes laterais (TOMLINSON, 2008, p. 11), são superadas com a distribuição em canais discretos. No entanto, os paradigmas que definiram os modos como a reconstrução de espaços acústicos ocorreriam ainda se mantiveram, mesmo sem a necessidade prévia de se privilegiar o armazenamento de canais frontais como em mídias analógicas.

Mesmo em aplicações mais recentes, os paradigmas do áudio do cinema continuam sendo um elemento presente no desenvolvimento de novos sistemas multicanal. Propostas como as da NHK apresentam a possibilidade de reprodução de fontes ao longo do espaço, mas estão diretamente ligadas ao uso de imagem como paradigma (HAMASAKI, 2011, p. 14-15). Usos que fogem a estes paradigmas se encontram em sistemas binaurais e ambissônicos, ambos tendo em seus fundamentos uma intenção de maior igualdade em distribuição e representação do espaço.

## CAPÍTULO 3

### TÓPICOS RECENTES EM SISTEMAS DE ÁUDIO ESPACIAL

Avanços recentes nas últimas décadas no campo do áudio, resultantes principalmente da implementação de sistemas anteriores em plataformas digitais, possibilitaram desenvolvimentos que impulsionaram as tecnologias espaciais de áudio. O aumento da capacidade de processamento digital tornou viável a modificação de sinais digitais em tempo real, possibilitando alterações como filtros, convoluções, simulações de espaços acústicos, e o cálculo de funções de transferência para reprodução serem usados em aplicações que procuram recriar diversos aspectos da percepção espacial.

A precisão destes processos se distingue de implementações anteriores que se fundamentavam em circuitos analógicos: implementações em sistemas analógicos muitas vezes não eram capazes de apresentar a definição necessária para aplicações espaciais, acabando por serem aproximações de sistemas mais bem definidos em teoria (BLAUERT, 1983, p. 135-136), ou tendo as limitações de muitas vezes depender de mídias físicas e terem, consequentemente, o resultado de implementações comprometido por erros inerentes aos sistemas relacionados a estas mídias, como no caso de decodificadores de discos quadrifônicos (DAVIS, 2003, p. 562).

O espaço digital não somente permitiu este aumento de precisão em processamentos ligados à reprodução espacial, mas também incentivou a implementação de novos métodos de relação com a informação codificada em formatos espaciais. Novos modos de representação, como sistemas espaciais baseados em objetos digitais, alteram a concepção do uso de formatos reprodutores de tal modo que implementações tenham a flexibilidade de reproduzir disposições de fontes sonoras independentemente do número de alto-falantes (AHRENS; GEIER; SPORS, 2010, p. 219-220). Aplicações como esta, que usam de informações pertinentes à reprodução espacial, permitem que a associação entre sinal e metadados possa ser processada, transmitida, e alterada com facilidade através de implementações que reconheçam estes dados.

Mesmo aspectos fundamentais da escuta espacial podem ser analisados e reconstruídos através de meios eletroacústicos a partir de implementações digitais. A diafonia,<sup>3</sup> cuja importância como fator limitante da reprodução de gravações binaurais em alto-falantes já era conhecida durante muito tempo (BLAUERT, 1983, p. 50-51), pode ser atenuada somente através de processamentos precisos resultantes de meios digitais (CHOUEIRI, 2018, p. 125), o que viabiliza pesquisas que investiguem a reprodução de, por exemplo, gravações binaurais para além de fones de ouvido.

Talvez o maior fator resultante desta interação entre as propostas de sistemas acústicos espaciais e as possibilidades de implementações digitais é a relativa facilidade de integração entre sistemas de áudio e outras formas de correlação como vídeo e posicionamento. Como mencionado anteriormente, os parâmetros acústicos derivados a partir de funções de

---

<sup>3</sup> Esta é a melhor tradução encontrada pelo autor para o termo inglês *crosstalk*.

transferência da escuta e outros aspectos descritores do aparelho auditivo não são o bastante para uma reconstrução fiel dos modos como a escuta se comporta em uma situação não mediada tecnologicamente. Implementações digitais permitem que se integrem aspectos para além de sinais de áudio, possibilitando a inclusão de outras modalidades de relação com o espaço que complementam a escuta na formação da percepção espacial.

Neste capítulo, serão apresentados tópicos ou conceitos que sejam de importância atual para os principais sistemas de áudio espacial em uso. Quando possível, procurou-se enfatizar a presença de tecnologias que complementem aplicações de áudio através de outras modalidades de percepção.

### **3.1 IMPLEMENTAÇÕES DE ÁUDIO BINAURAL**

Os fundamentos da gravação e reprodução binaural no áudio foram entendidos relativamente cedo se comparados com outras formas de representação espacial. Possivelmente pela proximidade com pesquisas diretamente relacionadas ao estudo do funcionamento da audição no espaço, as particularidades de sistemas binaurais, como as diferenças de tempo, amplitude, construção de imagens no espaço reproduzido e as formas pelas quais sinais são transformados ao interagir ao longo da linha de transmissão até o ouvido, já haviam sido pesquisadas e publicadas desde pelo menos os anos trinta (BOER, 1940, p. 107-109).

Mesmo assim, apesar de sua natureza relativamente simples, tanto em gravação como em reprodução, sistemas de áudio binaural tem apresentado aperfeiçoamentos recentes que são de fundamental importância para o avanço de sistemas imersivos. A relevância destes avanços é mais abrangente do que somente a aplicações binaurais em si, dado que outras tecnologias espaciais de áudio, como sistemas ambissônicos, usam métodos de conversão entre formatos para possibilitar a audição de informação espacial através de fones de ouvido, dependendo de transformações binaurais para esta conversão (WIGGINS, 2017, p. 1-2).

Grande parte das limitações atribuídas às atuais formas de emprego de áudio binaural está relacionada com o uso de funções de transferência. Kyriakakis, em um artigo sobre os fatores restritivos de implementações no áudio espacial, menciona três itens que impedem um maior uso de sistemas binaurais para aplicações em geral: a inexistência de funções de transferência individuais para cada ouvinte; erros de reprodução de posicionamentos por gravações apresentadas através de fones de ouvido; e dificuldade de externalização da escuta de fontes, resultando na localização de eventos auditivos dentro da cabeça (KYRIAKAKIS, 1998, p. 945).

A inviabilidade do uso de funções de transferência individualizadas para cada ouvinte sempre foi um fator limitante na abrangência do uso de aplicações binaurais, dado que as medições necessárias para que se possa derivar uma função de transferência somente são possíveis de serem obtidas em centros específicos de estudos acústicos. Métodos alternativos para estas medições, desde modelos geométricos virtuais até modelos estatísticos derivados de populações, já foram testados, mas seus resultados não foram suficientes para correção das anomalias auditivas resultantes de funções não-individualizadas (GARDNER, 2004, p. 39-41). Recentemente, bases de dados que disponibilizam diversas funções de transferência tornaram-

se amplamente acessíveis, possibilitando à cada ouvinte a pesquisa individual à procura de funções que resultem em um menor número de anomalias e erros para a sua escuta individual (WU et al., 2018, p. 194; BOMHARDT; FELS; KLEIN, 2016, p. 2).

No entanto, apesar de haver indicações de que algumas funções de transferências são localizadoras mais eficientes do que outras quando apresentadas os mesmos estímulos espaciais (ARRUDA et al., 1993, p. 113), resultados ainda apontam para a necessidade de funções individualizadas à escuta para que o sistema auditivo possa resolver anomalias resultantes de alterações de componentes espectrais, como as distinções de elevação no plano vertical e a oposição entre posicionamentos frontais e traseiros (ARRUDA et al., 1993, p. 122).

Um método recentemente proposto como solução ao problema de medição de funções de transferência é o uso de processos fotogramétricos para a derivação do comportamento acústico do pavilhão auricular através de imagens. Apesar de não apresentarem ainda a mesma definição de comportamento de medições feitas em laboratórios, os resultados destes métodos exibem uma qualidade suficiente para que se possa ter uma definição melhor em testes de escuta do que com funções genéricas a partir de bases de dados (JOHANSSON et al., 2020). Este tipo de medição já é implementado em aplicações fora do campo de experimentos acústicos, tornando a viabilidade de funções de transferência individualizadas uma realidade possível caso tornem-se acessíveis ao público em geral.

A restrição da localização apresentada por gravações binaurais quando estas são ouvidas através de fones, notada por Kyriakakis como um fator limitante existente no modo como sistemas binaurais são implementados, é um aspecto que procura ser remediado através do desenvolvimento de aplicações que resolvam as interações que ocorrem entre sinais binaurais reproduzidos e o pavilhão auricular. O comprometimento das qualidades do espectro de sinais originalmente captados por gravações usando técnicas binaurais quando reproduzidos através de fones já era citado por Blauert como um problema, dado que a interação do sinal reproduzido novamente com o ouvido afeta os parâmetros do espectro em altas frequências, descaracterizando aspectos espaciais previamente codificados (BLAUERT, 1983, p. 51). No entanto, recentemente, com a introdução de processamentos digitais de sinais, programas que possibilitam a eliminação da diafonia na escuta de gravações binaurais através de alto-falantes tornaram-se possíveis.

Este conceito foi inicialmente introduzido por Bauer como uma forma de adequação da reprodução de sinais gravados para alto-falantes serem reproduzidos em fones de ouvido sem que houvesse uma mudança na percepção de espaço entre os dois modos de reprodução (BAUER, 1961). Esta adequação era obtida através da introdução de sinais que simulavam a diafonia em uma escuta estereofônica. O processo inverso, de se eliminar, através de processamentos, a diafonia em um ponto de escuta, possibilitando a escuta de gravações feitas para audição em fones através de alto-falantes sem que haja uma distorção da percepção de espaço, tem tido implementações robustas recentemente.

Implementações anteriores permitiam o cancelamento da diafonia durante a reprodução, mas tinham diversas limitações resultantes como diminuição de extensão dinâmica, forte coloração do sinal reproduzido, e artefatos espaciais ocasionais decorrentes dos processamentos (CHOUEIRI, 2018, p. 156). Implementações atuais se baseiam na adequação dos sinais reproduzidos através de funções de transferência individuais ao invés de formas anteriores de funções analíticas implementadas por equalização. Isto permite uma melhor

performance, com menor taxa de erros na percepção espacial e coloração do sinal resultante (CHOUEIRI, 2018, p. 160-165). No entanto, esta abordagem ainda apresenta problemas práticos; segundo Rumsey (RUMSEY, 2001, p. 75):

The most important limitation is that the crosstalk-cancelling filters are only valid for a very narrow range of listening positions. Beyond a few tens of centimeters away from the ‘hot spot’ the effect often disappears almost completely. The effect is sometimes perceived as unnatural, and some listeners find it fatiguing to listen to for extend periods.

Uma possível solução para o âmbito de reprodução espacial estreito decorrente destes processamentos é a integração de tecnologias que sejam capazes de seguir o trajeto e posicionamento do ouvinte em relação a reprodutores e ajustar os filtros em uso respectivamente. Este tipo de adaptação por parte de decodificações binaurais já é implementado em aplicações virtuais como simulações e jogos (FARKAS, 2018, p. 35), onde a trajetória de uma visão virtual é simples de se extrair; implementações fazendo uso de câmeras ou acelerômetros sem fio somente foram possíveis recentemente. Nestas aplicações reais, não somente viabiliza-se uma maior zona de cobertura, mas torna-se possível a interação entre a direção de escuta do ouvinte e o espaço virtual sendo representado (ROGINSKA, 2018, p. 98). Deste modo, a reprodução binaural é capaz de incluir aspectos multimodais da percepção essenciais para uma ilusão de fontes sonoras, dado que o movimento do ouvinte em relação a um espaço de escuta é fundamental para desambiguação de diversas anomalias resultantes da audição.

Uma hipótese para a incapacidade de externalização de fontes sonoras durante reproduções binaurais, mesmo quando estas corretamente gravam e reproduzem os estímulos necessários para a localização espacial segundo funções de transferência, é a inexistência de alterações ao sinal quando ocorre movimentação por parte do ouvinte (KENDALL, 1995, p. 38): os únicos sinais naturais que tem este comportamento são a própria voz, que é localizada naturalmente dentro da cabeça, e sinais reproduzidos que se comportam de maneira similar tem uma localização espacial semelhante.

Esta integração entre sistemas de reprodução eletroacústica e sensores de posicionamento são essenciais para a sensação de externalização de fontes sonoras durante a escuta. Sem uma adequação de sinais acústicos em resposta a movimentações pelo ouvinte, a audição tem dificuldade de perceber os estímulos externos como fontes no espaço, resultando em localizações internas aos ouvidos, impossibilitando a externalização (GRIESINGER, 1998, p. 10). Roginska menciona (ROGINSKA, 2018, p. 107):

In a non-interactive listening environment, the listener perceives the audio world from a frozen perspective. In this non-interactive situation, having the audio world move in synchrony with the listener’s movements results in an unnatural listening situation. [...] by accounting for head movements, the perception of inside-the-head images can be minimized.

Assim, com estas adições ao formato tradicional de reprodução binaural, as limitações apresentadas por Kyriakakis são corrigidas. Estas novas tecnologias complementares à simples reprodução binaural tida atualmente como padrão possibilitam a derivação de funções de transferência individuais com relativa facilidade, a reprodução de sinais binaurais em alto-falantes, e a experiência interativa da audição de um espaço sonoro que permite a adequação das perspectivas auditivas em imersão. Neste formato, erros de localização, ambiguidades da

escuta e restrições de externalização são atenuados através da integração de outras modalidades de percepção à escuta e melhores capacidades de decodificação de sinais em espaços virtuais.

O uso de processamentos digitais para correção da diafonia permite também que outras formas de processamentos através de matrizes possam ocorrer antes da reprodução do sinal, como a adequação da reprodução binaural para mais de um ouvinte ou para mais do que o formato transaural padrão de dois alto-falantes apenas (BAUCK; COOPER, 1996, p. 688-696). Através de funções de transferência obtidas em pontos de escuta de salas específicas, processamentos digitais também permitem a auralização de sinais multicanal através de reproduções binaurais, possibilitando a audição de outros formatos específicos de áudio espacial através da simulação de espaços acústicos (RUMSEY, 2001, p. 76-78).

É necessário apontar que, como visto previamente, uma situação ideal de reprodução binaural em fones faz uso de processos de equalização que procuram eliminar tanto a resposta do próprio fone, como também o efeito acústico do volume entre o fone e o ouvido, que pode servir de resistência ao sinal reproduzido e introduzir distorções que impeçam uma percepção espacial correta (KENDALL, 1995, p. 38). No entanto, estas correções, feitas através de filtros de equalização ou convolução, são específicos de cada reproduutor.

### **3.2 IMPLEMENTAÇÕES DE ÁUDIO AMBISSÔNICO**

As propostas iniciais para sistemas ambissônicos apresentadas por Gerzon na década de setenta já continham limitações que foram trabalhadas ao longo das décadas seguintes com o intuito de aprimorar o sistema em diversas formas. Estas limitações não foram descobertas após o sistema ser concebido, já sendo conhecidas como consequência das formas como o sistema é fundamentado, mas foram toleradas em contexto, dado que a intenção então era a de propor um sistema mais flexível de representação espacial em relação às propostas quadrifônicas da mesma época.

Dentre estas restrições, a restrição do número de ordens para um sistema ambissônico apresenta-se como o principal fator limitante de novas aplicações. Sistemas de ordens maiores apresentam uma relativa diminuição em diversos defeitos pelo sistema, tais como *aliasing*, amplificação de níveis de ruído (DANIEL; MOREAU; NICOL, 2003, p. 9), frequência limite de representação espacial, área de cobertura de localização correta (NICOL, 2018, p. 288), e alterações de fase resultante de filtros necessários para se corrigir efeitos acústicos no processo de captação (BERTET; DANIEL; MOREAU, 2006, p. 10-11). Apesar destes defeitos provenientes da aplicação prática das bases teóricas de sistemas ambissônicos já terem sido identificados por Gerzon durante suas primeiras proposições, sistemas ambissônicos limitaram-se durante muitos anos a representações de primeira ordem pela praticidade de construção de equipamentos de gravação e reprodução (NICOL, 2018, p. 278).

Os erros decorrentes de ordens menores em sistemas ambissônicos são variados, mas seus resultados para a audição em uma situação de reprodução são comumente percebidos como incertezas de direcionalidade e definição na localização de fontes sonoras em um espaço reproduzido. Segundo Bertet et al. (BERTET; DANIEL; MOREAU, 2006, p. 2-3):

[...] 1<sup>st</sup> order encoding offers a quite limited angular discrimination for virtual sources that are relatively close to each other. This is probably a first cause of the lack of sound image precision and robustness criticized by detractors of Ambisonics (and of coincident microphone techniques more generally). By introducing “higher order ambisonic” encoding functions (1), i.e. functions which variation is greater for the same angular variation, a better angular discrimination is possible.

Atualmente, microfones ambissônicos de quarta ordem são os microfones de maior ordem disponíveis. Testes entre estes equipamentos e microfones ambissônicos semelhantes de ordens mais baixas confirmam suposições teóricas de melhor definição espacial, localização e área de cobertura correta (BRAUN; FRANK, 2011, p. 5-6). No entanto, o acréscimo linear de ordens em um sistema ambissônico implica em um crescimento exponencial de números de sinais a serem captados, processados, armazenados e reproduzidos. Isto dificulta o uso destes sistemas. Outras formas de captação e reprodução, fundamentadas em concepções de sistemas ambissônicos, têm sido testadas, usando de conjuntos de microfones distribuídos irregularmente no espaço. Apesar de serem capazes, dentro de certos limites, de obterem respostas de espacialização semelhantes a sistemas ambissônicos de quinta ordem (BRUNO; LABORIE; MONTOYA, 2003, p. 17), estes testes ainda não são capazes de superar totalmente as restrições de representação espacial inerentes à captação e reprodução ambissônica. Mesmo em sistemas com trezentos e cinquenta alto-falantes, desenvolvidos e distribuídos especificamente para a reprodução de sinais no espaço, é possível somente a reprodução de sinais de décima segunda ordem no máximo (BARRETT et al., 2016, p. 22).

Esta correlação entre um número exponencialmente grande de captadores ou reprodutores e as ordens de um sistema ambissônico limitam aplicações práticas que procurem uma representação espacial sem apresentar distorções. Mesmo com a possibilidade de uma melhor resolução espacial proveniente de sistemas de ordens maiores, aspectos de amplificação limitam as faixas de frequência nas quais a maior resolução de sistemas de ordens maiores pode ser usada, impedindo baixas frequências de melhor definição no espaço (MEYER, 2018). Em cálculos procurando os limites associados às ordens de sistemas ambissônicos, modelos de reprodução que não apresentassem erros de representação ao longo do espaço reproduzido necessitariam de 38<sup>a</sup> ordem (HOLD, 2018), enquanto outros modelos preveem uma necessidade de 35<sup>a</sup> ordem para a reprodução de frequências acima de 10 kHz sem que haja distorções de parâmetros de localização (WIGGINS, 2017, p. 8).

Atualmente, estas limitações, inerentes aos fundamentos do próprio sistema, impedem que aplicações ambissônicas de ordens maiores viabilizem a captação e reprodução de espaços sonoros sem que haja comprometimentos de parâmetros auditivos como coloração de sinais, localização, área de cobertura, ou banda de frequência.

No entanto, apesar destes comprometimentos e limitações, o uso de formatos ambissônicos atualmente se justifica pela característica própria do sistema em flexibilidade de armazenamento e representação de espaços acústicos. Outros formatos ligados ao áudio espacial seguem, na maior parte dos casos, formas de armazenamento e representação de seus sinais em canais individuais para a reprodução em uma disposição específica de reprodutores; sistemas ambissônicos, por outro lado, tem a possibilidade de armazenamento e reprodução em um número indefinido de reprodutores, tanto em disposição quanto em número de canais (NETTINGSMEIER, 2010, p. 8-9). Principalmente em casos de codificação de espaços através de síntese, a representação de um espaço sonoro não se torna obsoleta com o avanço da

tecnologia da mesma forma que formatos anteriores baseados em canais, dado que estes têm direta dependência de sistemas de reprodução específicos aos respectivos formatos de representação. Esta versatilidade em representar um espaço sonoro dá a aplicações ambissônicas a possibilidade de complementar formas de captação e reprodução do sistema em formatos futuros, vindo a superar, ao longo dos anos, limitações próprias aos fundamentos do sistema, sem que haja uma perda de parte da produção prévia que já tenha ocorrido em versões anteriores. Sem a associação direta com qualquer tipo de forma de reprodução, formatos ambissônicos mostram-se como a melhor forma, atualmente, de representação espacial no áudio, permitindo compatibilidade futura, escalabilidade e universalidade entre suas representações (NICOL, 2018, p. 292-293).

Este paradigma de representação de espaços no áudio, definido como sistemas baseados em campos sonoros (NICOL, 2018, p. 300), apresenta-se como uma possível solução para a ausência de um formato universal de armazenamento e transmissão de informação espacial no áudio. Rumsey menciona esta dificuldade como um dos principais fatores a serem superados por sistemas de áudio imersivo atualmente (RUMSEY, 2014). Mesmo com avanços futuros em diversos aspectos restritivos de representações atuais de espaço, o formato fundamental de codificação do sistema ainda apresentaria uma compatibilidade para extração dos espaços registrados, permitindo a independência destas gravações a aplicações individuais. Convenções para o ordenamento de componentes individuais de diferentes ordens permite a facilidade de transmissão de informação garantindo invariância ao longo do processo (DELEFLIE et al., 2011, p. 4-7). Deste modo, a codificação ambissônica de aplicações de áudio espacial mostra-se particularmente vantajosa como um formato intermediário para produção a reprodução (SCHUTZE, 2018, p. 44).

Uma forma de expansão da versatilidade de formatos ambissônicos se dá através da integração das representações espaciais próprias deste formato às tecnologias de metadados e objetos. A primeira destas fundamenta-se em convenções, como o MPEG-4, de uso de metadados associados à transmissão de áudio, permitindo adaptações e alterações em outros estágios de uso baseados nas informações agregadas aos sinais na forma de metadados (BOSI; GOLDBERG, 2003, p. 412-413). Sistemas de representação baseados em objetos são sistemas fundamentados no uso de metadados para criação de uma representação flexível de objetos de mídia, definindo aspectos de objetos sonoros representados, por sinais que não necessariamente contenham toda a informação espacial exigida para reprodução, mas que possam ser atribuídos qualidades espaciais, como posicionamento, extensão e direvidade, através da interpretação e decodificação por outros sistemas (BREEBART et al., 2008, p. 1-2). A associação destes metadados a objetos de mídia permite uma taxa menor de transferência quando transmitindo sinais, dado que a transmissão de canais codificados através de objetos é limitada em comparação com formatos multicanal, ao mesmo tempo que não compromete aspectos espaciais de reprodução, desde que sistemas de interpretação da informação relevante existam no receptor. Sistemas baseados em objetos também permitem a possibilidade de adaptação do espaço reproduutor, possibilitando processamentos do espaço sonoro, como adaptações relativas à direção do ouvinte ou alterações baseadas em funções de transferência (BREEBART et al., 2018, p. 493-494), o que permite o suporte a um conjunto de relações multimodais necessárias à escuta em uma situação de reprodução.

Esta associação de informações diversas a objetos de mídia tem sido particularmente usada em situações que têm como aspecto crítico o número de canais para

transmissão, como em representações de espaços sonoros. Sistemas ambissônicos de ordens maiores incorrem na dificuldade de transmissão de um número impraticável de canais para grande parte das situações comuns de transmissão (AHRENS; GEIER; SPORS, 2010, p. 220); no entanto, formas de codificação específicas para formatos espaciais permitem que somente objetos de mídia individuais sejam armazenados ou transmitidos, tornando mais viável o uso de sistemas multicanais de grande escala (AHRENS; GEIER; SPORS, 2010, p. 227).

Aplicações específicas para áudio no cinema tiveram um avanço recente com o emprego de paradigmas baseados em objetos de áudio durante os estágios de produção. A integração deste modo de produção no cinema supriu a dificuldade de não haver uma maneira simples de se abordar diversos sistemas com crescentes números de canais disponíveis nas últimas décadas em termos de criação de espaços sonoros (RUMSEY, 2014). Testes psicoacústicos apontam uma semelhança em aspectos de localização e percepção espacial quando comparando a reprodução de formatos *surround* padrão com a decodificação de sistemas baseados em objeto reproduzidos no mesmo número de canais (NEUBAUER; ORAMUS, 2019, p. 7). As implementações para decodificação de objetos de áudio, então, já apresentam uma igualdade em aspectos de reprodução do espaço, ao mesmo tempo que possuem as vantagens já apresentadas anteriormente.

Este conjunto de características e avanços representa uma visão geral dos principais tópicos em desenvolvimento na área de captação, processamento e reprodução de campos sonoros no áudio atualmente. Pode-se ver duas tendências em crescimento que parecem se somar aos paradigmas anteriores no áudio espacial: sistemas que procuram captar ou reproduzir um campo sonoro, e o uso de objetos de mídia para a transmissão, armazenamento, modificação e reprodução de representações espaciais. Ambas são provenientes do avanço de sistemas multicanais prévios, lidando com os problemas resultantes da expansão do número de canais a limites praticáveis atualmente.

As aplicações imersivas no áudio atual parecem se dividir em duas linhas gerais de implementação: a individualização da escuta, voltada para a adaptação de um espaço sonoro a um ponto único de percepção, ocorrendo principalmente através de tecnologias binaurais, e a reprodução de campos sonoros no espaço, através de sistemas multicanais. Mesmo implementações com origens distintas, mas paradigmas comuns, como sistemas ambissônicos e de síntese de campo de onda, se deparam com semelhantes desafios atuais para a superação de suas limitações (DANIEL; MOREAU; NICOL, 2003, p. 11). Uma conexão entre estas duas linhas gerais parece ocorrer na virtualização de espaços sonoros, possibilitando a decodificação de reproduções para diferentes paradigmas de representação, como através de funções de transferência em sistemas binaurais ou na reconstrução de espaços sonoros em formatos ambissônicos; uma generalização ainda maior é permitida através do uso de objetos de mídia, que reforçam a generalidade de diferentes implementações, possibilitando ainda mais um intercâmbio entre diferentes paradigmas de gravação e reprodução.

Deste modo, o principal avanço das novas tecnologias espaciais no áudio em comparação com as formas de gravação e reprodução do século passado está na adaptabilidade proveniente do uso de campos sonoros como formato intermediário, em comparação com formatos anteriores baseados em canais. Esta abordagem de representação permite um processo de produção e reprodução de espaços independente dos sistemas finais. É possível que avanços futuros destaquem estes estágios de representação e reprodução independentemente, e que,

como resultado, novas formas de interação e experimentação de aspectos espaciais resultem da pesquisa própria de formas de reprodução e integração de percepções multimodais do espaço.

## CONCLUSÃO

Este trabalho procurou apresentar três campos de interesse que compõe, para o autor, áreas de estudo de grande importância para aqueles que procuram adentrar no campo de áudio espacial. As áreas da psicoacústica, dos precedentes históricos e das abordagens atuais no desenvolvimento das tecnologias em uso neste campo hoje mostraram-se, ao longo do desenvolvimento desta pesquisa, como maneiras relevantes de se abordar um assunto amplo e complexo como o atual objeto de estudo.

Apesar do escopo deste trabalho não permitir uma investigação mais a fundo em cada uma destas áreas, o autor acredita que as apresentações aqui feitas possibilitam uma indicação inicial de tópicos e aspectos referentes a cada assunto em particular, permitindo pelo menos uma introdução inicial a estas diferentes áreas de interesse ao áudio espacial.

Como mostrado ao longo da exposição feita nesta pesquisa, o áudio imersivo é um campo de estudo de grande extensão, crescendo principalmente nos últimos anos por conta do renovado interesse resultante de aplicações em diversas áreas, possibilitado por novas tecnologias digitais. Seus avanços atuais têm sido consideráveis: em um primeiro momento, estes avanços podem parecer como revoluções tecnológicas, especialmente levando em conta as aparentes inovações de suas abordagens. No entanto, uma investigação mais localizada mostra tendências de desenvolvimento e correlações entre diferentes sistemas dentro de paradigmas gerais compartilhados entre diferentes abordagens dos mesmos problemas comuns.

Como exposto anteriormente, este campo comum entre diferentes abordagens ao áudio espacial nem sempre existiu ao longo da história destas tecnologias. Atualmente, no entanto, o autor acredita que o intercâmbio promovido por novos paradigmas de representação e reprodução do áudio espacial indique um momento propício de estudo e interesse nesta área em particular. Diferentes perspectivas de se abordar problemas semelhantes permitem diferentes abordagens às mesmas dificuldades, e a comunicação agora possível entre diversas implementações existentes torna vantajoso o estudo de tecnologias nesta área.

Apesar de se encontrar fora do presente escopo desta pesquisa, a continuidade do presente estudo poderia ocorrer tanto na área técnica como na área artística. Aplicações práticas na forma de análises individuais de diferentes formatos de gravação, representação e reprodução espacial podem indicar aspectos da escuta que não foram inclusos ao longo deste estudo. De forma semelhante, a comparação, através de gravações de mesmos programas em diversos formatos de representação, poderia evidenciar aspectos práticos que não são explicitados em artigos técnicos. A associação entre os parâmetros técnicos apresentados aqui com processos criativos possibilita a avaliação de intenções artísticas mediadas por meios tecnológicos, viabilizando diferentes formas de criação a partir das mesmas ferramentas.

Assim, acredita-se que esta pesquisa possa servir como uma exposição geral, ainda que breve, de um campo que encontra neste momento uma ocasião oportuna da promoção de desenvolvimento e pesquisas.

## BIBLIOGRAFIA

ADER. Clément Ader. *Telephonic Transmission of Sound from Theaters*. No. 257453, 9 Mai 1882, 13 Jan. 1882.

AHRENS, Jens; MATTHIAS, Geier; SPORS, Sascha. Object-Based Audio Reproduction and the Audio Scene Description Format. *Organised Sound*, v. 15, n. 3, 2010, p. 219-227.

ARRUDA, Marianne et al. Localization Using Nonindividualized Head-Related Transfer Functions. *Journal of the Acoustical Society of America*, v. 94, n. 1, p. 111-123, 1993.

BARRETT, Natasha et al. Holophonic Sound in IRCAM's Concert Hall: Technological and Aesthetic Practices. *Computer Music Journal*, v. 40, n. 4, 2016, p. 14-34.

BAUCK, Jerry; COOPER, Duane. Generalized Transaural Stereo and Applications. *Journal of the Audio Engineering Society*, v. 44, n. 9, 1996, p. 683-705.

BEDELL, E.; KERNEY, Iden. Auditory Perspective: System Adaptation. *Transactions of the American Institute of Electrical Engineers*, v. 53, n. 1, p. 216-219, 1934.

BEGAULT, Durand; GODFROY-COOPER, Martine; WENZEL, Elizabeth. Perception of Spatial Sound. In: ROGINSKA, Agnieszka; GELUSO, Paul. *Immersive Sound: The Art and Science of Binaural and Multi-channel Audio*. 1. ed. New York: Routledge, 2018. p. 5-39.

BELL. Alexander Graham Bell. *Improvement in Telegraphy*. No. 174465, 7 Mar. 1876, 14 Fev. 1876.

BERTET, Stephanie; DANIEL, Jerome; MOREAU, Sebastien. 3D Sound Field Recording with Higher Order Ambisonics – Objective Measurements and Validation of a 4<sup>th</sup> Order Spherical Microphone. In: 120<sup>TH</sup> AUDIO ENGINEERING SOCIETY CONVENTION, 2006, Paris. **Proceedings...** Paris: AES, 2006, p. 24.

BLAUERT, Jens. *Spatial Hearing: The Psychoacoustics of Human Sound Localization*. 2. ed. Cambridge: The MIT Press, 1983.

BLUMLEIN, Alan. Alan Dower Blumlein. *Improvements in and Relating to Sound-Transmission, Sound-Recording and Sound-Reproducing Systems*. No. 394325, 14 Dec. 1931.

BOER, K. Stereophonic Sound Reproduction. *Philips Technical Review*, v. 5, n. 4, p. 107-114, 1940.

BOMHARDT, Ramona; FELS, Janina; KLEIN, Matias. A High-Resolution Head-Related Transfer Function and Three-Dimensional Ear Model Database. In: PROCEEDINGS OF MEETINGS ON ACOUSTICS, 172., 2016, Honolulu. **Proceedings...** Honolulu: Acoustical Society of American, 2017, 11 p.

BOREN, Braxton. History of 3D Sound. In: ROGINSKA, Agnieszka; GELUSO, Paul. *Immersive Sound: The Art and Science of Binaural and Multi-channel Audio*. 1. ed. New York: Routledge, 2018. p. 41-62.

BOSI, Marina; GOLDBERG, Richard. Introduction to Digital Audio Coding and Standards. 1<sup>a</sup> ed. New York: Springer, 2003. 432 p.

BURNS, Russel. *The Life and Times of A. D. Blumlein*. 1. ed. London: The Institution of Engineering and Technology, 2006.

BRAUN, Sebastian; FRANK, Matthias. Localization of 3D Ambisonic Recordings and Ambisonic Virtual Sources. In: INTERNATIONAL CONFERENCE ON SPATIAL AUDIO, 2011, Detmold. **Proceedings...** Detmold: 2011, 6 p.

BREEBART, Jeroen et al. Spatial Audio Object Coding (SAOC) – The Upcoming MPEG Standard on Parametric Object Based Audio Coding. In: 124<sup>TH</sup> AUDIO ENGINEERING SOCIETY CONVENTION, 2008, Amsterdam. **Proceedings...** Amsterdam: AES, 2008, 15 p.

\_\_\_\_\_ et al. Spatial Coding of Complex Object-Based Program Material. *Journal of the Audio Engineering Society*, v. 67, n. 7/8, p.486-497, 2019.

BRUNO, Remy; LABORIE, Arnaud; MONTOYA, Sebastien. A New Comprehensive Approach of Surround Sound Recording. In: 114<sup>TH</sup> AUDIO ENGINEERING SOCIETY CONVENTION, 2003, Amsterdam. **Proceedings...** Amsterdam: AES, 2003.

CHOUEIRI, Edgar. Binaural Audio Through Loudspeakers. In: ROGINSKA, Agnieszka; GELUSO, Paul. *Immersive Sound: The Art and Science of Binaural and Multi-channel Audio*. 1. ed. New York: Routledge, 2018. p. 124-1.

CLARK, H.; DUTTON, G.; VANDERLYN, P. The Stereosonic Recording and Reproducing System: A Two Channel System for Domestic Tape Records. *The Proceedings of the Institution of Electrical Engineers*, v. 104, n. 17, p. 417-432, 1957.

DANIEL, Jerome; MOREAU, Sebastien; NICOL, Rozenn. Further Investigations of High Order Ambisonics and Wavefield Synthesis for Holophonic Sound Imaging. In: 114<sup>TH</sup> AUDIO ENGINEERING SOCIETY CONVENTION, 2003, Amsterdam. **Proceedings...** Amsterdam: AES, 2003, 18 p.

DELEFLIE, Etienne et al. AmbiX – A Suggested Ambisonics Format. In: AMBISONICS SYMPOSUM, 2011, Lexington. **Proceedings...** Lexington: 2011, 11 p.

FARKAS, Tomas. Binaural and Ambisonic Sound as the Future Standard of Digital Games. *Acta Ludologica*, v. 1, n. 2, 2018, p. 34-47.

FELLGETT, Peter. Ambisonics: Part One. General System Description. *Studio Sound*, v. 17, n. 8, p. 20-24, 1975.

FLETCHER, Harvey. Auditory Perspective: Basic Requirements. *Transactions of the American Institute of Electrical Engineers*, v. 53, n. 1, p. 9-11, 1934.

GARDNER, William. Spatial Audio Reproduction: Toward Individualized Binaural Sound. *The Bridge*, v. 34, n. 4, 2004, p. 37-42.

GARITY, William; HAWKINS, John. Fantasound. *Journal of the Society of Motion Picture Engineers*, v. 37, p. 127-146, 1941.

GARITY, William; JONES, Watson. Experiences in Road-Showing Walt Disney's Fantasia. *Journal of the Society of Motion Picture Engineers*, v. 39, p. 6-15, 1942.

GELUSO, Paul; ROGINSKA, Agnieszka. *Immersive Sound: The Art and Science of Binaural and Multi-Channel Audio*. 1. ed. New York: Focal Press, 2018.

GERZON, Michael. Surround Sound from Two-Channel Stereo. *Hi-Fi News*, v. 15, p. 1104-1109, 1970a.

\_\_\_\_\_. The Principles of Quadraphonic Recording Part One. *Studio Sound*, v. 12, n. 8, p. 338-342, 1970b.

\_\_\_\_\_. The Principles of Quadraphonic Recording Part Two. *Studio Sound*, v. 12, n. 9, p. 381-384, 1970c.

\_\_\_\_\_. Experimental Tetrahedral Recording Part One. *Studio Sound*, v. 13, n. 8, p. 396-398, 1971a.

\_\_\_\_\_. Experimental Tetrahedral Recording Part Three. *Studio Sound*, v. 13, n. 10, p. 511-515, 1971b.

\_\_\_\_\_. Psychoacoustics of Surround Sound. *Wireless World*, v. 80, n. 1468, p. 483-486, 1974.

\_\_\_\_\_. Ambisonics: Part Two. Studio Techniques. *Studio Sound*, v. 17, n. 8, p. 24-30, 1975.

GRIESINGER, David. General Overview of Spatial Impression, Envelopment, Localization, and Externalization. In: 15<sup>TH</sup> INTERNATIONAL CONFERENCE OF THE AUDIO ENGINEERING SOCIETY, 1998, Copenhagen. **Proceedings...** Copenhagen: AES, 1998, 15 p.

HOLD, Christoph. Improvements on Higher Order Ambisonics Reproduction. Disponível em <<https://youtu.be/f4oE3i2tjyA>>. Gravação de evento. 2018.

JOHANSSON, Jaan. Accuracy of Photogrammetric Extraction of the Head and Torso Shape for Personal Acoustical HRTF Modelling. In: 94<sup>TH</sup> AUDIO ENGINEERING SOCIETY CONVENTION, 2020, Online. **Proceedings...** Online: AES, 2020.

KENDALL, Gary. A 3-D Sound Primer: Directional Hearing and Stereo Reproduction. *Computer Music Journal*, v. 19, n. 4, p. 23-46, 1995.

KUSUMOTO, Kenta; KYTO, Mikko; OITTINEN, Pirkko. The Ventriloquist Effect in Augmented Reality. In: 2015 IEEE INTERNATIONAL SYMPOSIUM ON MIXED AND AUGMENTED REALITY, 2015, Fukuoka. **Proceedings...** Fukuoka: IEEE, 2015, 5 p.

KYRIAKAKIS, Chris. Fundamental and Technological Limitations of Immersive Audio Systems. *Proceedings of the IEEE*, v. 86, n. 5, p. 941-951, 1998.

LIU, Yu et al. Near-Field Head-Related Transfer-Function Measurement and Database of Human Subjects. *Journal of the Acoustical Society of America*, v. 143, n. 3, p. 194-198, 2018.

MALHAM, David; MYATT, Anthony. 3-D Sound Spatialization Using Ambisonics Techniques. *Computer Music Journal*, v. 19, n. 4, p. 58-70, 1995.

MAXFIELD, J. P.; HARRISON, H. C. Methods of High-Quality Recording and Reproducing of Music and Speech Based on Telephone Research. *Journal of the A.I.E.E*, v. 45, p. 243-253, 1926.

MCGINN, Robert. Stokowski and the Bell Telephone Laboratories: Collaboration in the Development of High-Fidelity Sound Reproduction. *Technology and Culture*, v. 24, n. 1, p. 38-75, 1983.

MEYER, Jens. Higher Order Ambisonic Microphones From Theory to Application: Spatial Audio Summer Seminar 2018. Disponível em <<https://youtu.be/f4oE3i2tjyA>>. Gravação de evento. 2018.

NETTINGSMEIER, Jorn. Higher Order Ambisonics – A Future-Proof 3D Audio Technique. In: 26<sup>th</sup> TONMEISTER TAGUNG INTERNATIONAL CONVENTION, 2010, Leipzig. **Proceedings...** Leipzig: 2010, 12 p.

NEUBAUER, Petr; ORAMUS, Tomas. Comparison Study of Listeners' Perception of 5.1 and Dolby Atmos. In: 147<sup>TH</sup> AUDIO ENGINEERING SOCIETY CONVENTION, 2019, New York. **Proceedings...** New York: AES, 2019, 8 p.

NICOL, Rozenn. Sound Field. In: ROGINSKA, Agnieszka; GELUSO, Paul. *Immersive Sound: The Art and Science of Binaural and Multi-channel Audio*. 1. ed. New York: Routledge, 2018. p. 276-310.

PAUL, Stephan. Binaural Recording Technology: A Historical Review and Possible Future Developments. *Acta Acustica United with Acustica*, v. 95, n. 5, p. 767-788, 2009.

PLUMB, Edward. The Future of Fantasound. *Journal of the Society of Motion Picture Engineers*, v. 39, p. 16-21, 1942.

REEVES, Hazard. This is Cinerama. *Film History*, v. 11, n. 1, p. 85-97, 1999.

ROBERTS, Alexander. *The Inventor of Stereo: The Life and Works of Alan Dower Blumlein*. 1. ed. Oxford: Routledge, 2013.

ROGINSKA, Agnieszka. Binaural Audio Through Headphones. In: ROGINSKA, Agnieszka; GELUSO, Paul. *Immersive Sound: The Art and Science of Binaural and Multi-channel Audio*. 1. ed. New York: Routledge, 2018. p. 88-123.

RUMSEY, Francis. Spatial Audio. 1<sup>a</sup> ed. Oxford: Routledge, 2001. 240 p.

\_\_\_\_\_. Spatial Audio – Reconstructing Reality or Creating Illusion. Disponível em <<https://youtu.be/y82nth2Pnwk>>. Gravação de evento. 2014.

SCHROEDER, Manfred. An Artificial Stereophonic Effect Obtained from Using a Single Signal. *Journal of the Audio Engineering Society*, v. 6, p. 74-79, 1958.

SCHUTZE, Stephan; SCHUTZE, Anna. New Realities in Audio. 1<sup>a</sup> ed. Boca Raton: Taylor and Francis, 2018. 328 p.

SCIENTIFIC AMERICAN. New York: Scientific America, 1881- .

SNOW, William. Basic Principles of Stereophonic Sound. *Journal of the Society of Motion Pictures and Television Engineers*, v. 61, p. 567-589, 1953.

SUNIER, John. *The Story of Stereo: 1881- .* 1. ed. New York: Gernsback Library, 1960.

THEILE, Günther. The New Sound Format “3/2-Stereo”. In: 94<sup>TH</sup> AUDIO ENGINEERING SOCIETY CONVENTION, 1993, Berlin. **Preprint...** Berlin: AES, 1993, 17 p.

TORICK, Emil. Highlights in the History of Multichannel Sound. *Journal of the Audio Engineering Society*, v. 46, n. 1/2, p. 27-31, 1998.

WADE, Nicholas; DEUTSCH, Diana. Binaural Hearing – Before and After the Stethophone. *Acoustics Today*, Melville, v. 4, n. 3, p. 16-27, 2008.

WIGGINS, Bruce. Analysis of Binaural Cue Matching Using Ambisonics to Binaural Decoding Techniques. In: 4<sup>TH</sup> INTERNATIONAL CONFERENCE ON SPATIAL AUDIO, 2017, Graz. **Presentation...** Graz: 2017, 11 p.

WILLIAMS, Michael. Stereophonic Zoom: A Practical Approach to Determining the Characteristics of a Spaced Pair of Microphones. In: 75<sup>TH</sup> AUDIO ENGINEERING SOCIETY CONVENTION, 1984, London. **Preprint...** Londres: AES, 1984, 10 p.

YOST, William. History of Sound Source Localization: 1850-1950. In: 173<sup>RD</sup> MEETING OF ACOUSTICAL SOCIETY OF AMERICA AND 8<sup>TH</sup> FORUM ACUSTICUM, 173, 2008, Boston. **Proceedings...** 2017: Acoustical Society of America, 2017, 15 p.